Fanjuan Shi

A Smart Marketing Personalization Methodology Based On Real-Time Contextualisation In The Era of Big Data
A smart marketing personalization methodology based on real-time contextualization in the era of big data

Devant le jury composé de :

Jean-Jack CEGARRA, Professeur des universités, Université Jean-Moulin Lyon 3, Co-directeur de thèse
Amos DAVID, Professeur des universités, Université de Lorraine, Suffragant
Chirine GHEDIRA, Professeur des universités, Université Jean-Moulin Lyon 3, Co-directrice de thèse
Jean-Luc MARINI, Président et Directeur de la Recherche, Search’XPR SAS, Suffragant
Widad MUSTAFA EL HADI, Professeur des universités, Université de Lille 3, Rapporteur
Catherine PIVOT, Professeur Emérite, Université Jean-Moulin Lyon 3, Co-directrice de thèse
Mohammed SAAD, Professeur des universités, University of the West of England, Suffragant
Jie-Ru ZHOU, Maître de Conférences HDR, Shanghai Jiao Tong University, Rapporteur
Cognition and conduct.

Dynasty MING - WANG Shou-Ren (1472 – 1529)

This work is dedicated to my families.
Acknowledgements

The journey to complete the Ph.D. study is just like mountain-climbing. The target is quite attractive and irresistible but the journey is tedious. By the time I reach the peak, the first thing I want to do, is to express my sincere gratitude to all the people around me, who give me the courage, the power, the ability, and the support to complete this amazing journey.

Firstly, I want to thank my lovely family (my parents-in-law, my parents, and my wife) in Shanghai and in Lyon. Without your encouragement, comprehension, and support, it would be impossible to handle the tremendous challenges of the new environment, new language, and new culture in France, that is very different from that in my mother country. It is your love that keeps me moving forward without fear and hesitation.

Secondly, I want to express my gratitudes to all the reviewers of this thesis, which include Mrs. Widad Mustafa El Hadi, Mrs. Jie-Ru Zhou, Mr. Amos David, Mr. Mohamed Saad, and Mr. Jean-Luc Marini. They set aside their precise time to review my work and provide me with valuable suggestions which enhance the quality of this thesis.

Thirdly, I would like to thank my three co-directors Mr. Jean-Jack CEGARRA, Mrs. Chirine GHEDIRA and Mrs. Catherine PIVOT. Mr. CEGARRA and Mrs. PIVOT have provided me with the valuable academic suggestions and administrative supports at
each critical step of my research. Mrs. GHEDIRA not only helped me to integrate into the local research community but also spent a lot of time in my research to suggest interesting and relevant research ideas, which lead to my five research publications in academic journal and conferences over the past four years.

Then, I would like to thank the two co-founders of the company Search’XPR: Mr. Jean-Luc MARINI and Mr. Olivier FIGON. These two gentlemen have a real passion for challenges and for excellence. During my four years stay with Search’XPR, they have provided me with countless support (financially, operationally, and spiritually) and advice that allowed me to conduct and complete my research. More importantly, they have provided me with the experiment field where I can apply my thoughts and ideas to the real business world. They regard me as their own family member and I will memorize this for the rest of my life.

Next, I would also like to express my gratitude to Mrs. Sandrine BRUNET of the Ecole Doctorale, to Mrs. Catherine Vulcain of the Laboratoire Magellan, and to my colleague Vincent JUSTIN of Search’XPR for their support and advice that facilitate my research and my work in Jean Moulin Lyon University and in Search’XPR.

Finally, I want to thank the “Association Nationale de la Recherche et de la Technologie” (ANRT) for financing this research. The financial support of ANRT constitutes an indispensable part of my research and help improve the quality and market value of the product developed by our team in Search’XPR.
Table of Contents

Acknowledgements ............................................................................................................. 3
Table of Contents ............................................................................................................... 5
Disclaimer .......................................................................................................................... 11
Introduction ....................................................................................................................... 12
  0.1 Evolution of the business to consumer marketing .................................................... 12
  0.2 Challenges of B2C marketing .................................................................................. 17
     0.2.1 A Motivation case ......................................................................................... 17
     0.2.2 Analysis of challenges in B2C marketing ....................................................... 20
     0.2.3 Research goal and objectives ....................................................................... 23
  0.3 Structure of the thesis ............................................................................................. 24
Chapter 1 .......................................................................................................................... 28
  1.1 Introduction ............................................................................................................. 28
  1.2 Shopping journey, consumer decision, and sales channels ................................... 29
     1.2.1 Shopping journey ......................................................................................... 29
     1.2.2 Consumer decision-making ....................................................................... 30
     1.2.3 Sales channels ............................................................................................ 31
     1.2.4 Cross-channel shopping behavior ............................................................... 32
  1.3 Personalization ....................................................................................................... 33
     1.3.1 Consumer profiling ..................................................................................... 34
     1.3.2 Content-based personalization ..................................................................... 34
     1.3.3 Collaborative filtering .................................................................................. 35
     1.3.4 Hybrid approach ......................................................................................... 36
  1.4 Context and contextual factors .............................................................................. 37
1.4.1 Defining context ........................................................................................................... 37
1.4.2 Contextual factors ........................................................................................................ 39
  1.4.2.1 Position .................................................................................................................... 39
  1.4.2.2 Time ........................................................................................................................ 40
  1.4.2.3 State of mind .......................................................................................................... 42
  1.4.2.4 Shopping goal ....................................................................................................... 43
  1.4.2.5 Shopping budget ................................................................................................. 44
  1.4.2.6 User Interface ..................................................................................................... 44
  1.4.2.7 Other contextual factors ..................................................................................... 46
1.4.3 Discovering contextual factors .................................................................................. 46
1.5 Data mining and machine learning .............................................................................. 50
  1.5.1 Data mining ............................................................................................................. 51
  1.5.2 Machine learning .................................................................................................. 53
1.6 Big Data ....................................................................................................................... 54
  1.6.1 What is big data? ..................................................................................................... 55
  1.6.1 What is its impact on B2C marketing strategy? ..................................................... 57
1.7 Summary ....................................................................................................................... 59
Chapter 2 .............................................................................................................................. 60
  2.1 Introduction .................................................................................................................. 60
  2.2 Contextualized B2C marketing personalization: a research overview ..................... 61
  2.3 Cross-channel shopping and consumer intentions ...................................................... 67
  2.4 Cross-channel consumer tracking ............................................................................. 71
  2.5 Personalization Method ............................................................................................... 73
  2.6 Contextualization Approach ..................................................................................... 77
    2.6.1 Contextual pre-filtering ....................................................................................... 80
2.6.2 Contextual modeling ................................................................. 81
2.6.3 Contextual post-filtering .......................................................... 84
2.6.4 Other limitations ................................................................. 85

2.7 Marketing performance improvement ........................................ 87
  2.7.1 Performance analysis ......................................................... 87
  2.7.2 Performance improvement ................................................. 91

2.8 Summary .................................................................................. 92

Chapter 3 ....................................................................................... 94
  3.1 Introduction ........................................................................... 94
  3.2 The cognition-behavior model ................................................ 96
  3.3 The intention-based personalization methodology ...................... 99
  3.4 The contextualization methodology ........................................ 104
  3.5 The Context-aware marketing personalization system ............... 106
    3.5.1 Business logic of the system ............................................ 108
    3.5.2 The infrastructure layer .................................................. 109
    3.5.3 The module layer ............................................................ 111
    3.5.4 The service layer ............................................................. 113
    3.5.5 The application layer ....................................................... 114

3.6 Summary .................................................................................. 114

Chapter 4 ....................................................................................... 116
  4.1 Introduction ........................................................................... 116
  4.2 The encoding protocol .......................................................... 117
    4.2.1 The scope of data collection ........................................... 117
      4.2.1.1 Data Collection on the Online Platforms ................. 119
      4.2.1.2 Data Collection in Physical Stores ....................... 121
4.2.2 The modality of data encoding and collection ................................. 123
  4.2.2.1 UPID and UCID Management ...................................................... 125
4.3 The consolidation protocol .................................................................... 126
  4.3.1 Cross-channel consumer identification ............................................. 127
  4.3.2 The data consolidation protocol ....................................................... 128
4.4 The fragmentation protocol .................................................................... 130
  4.4.1 The milestone event ......................................................................... 130
  4.4.2 The preparation of sub-streams ....................................................... 132
4.5 Summary ................................................................................................. 133
Chapter 5 ....................................................................................................... 135
  5.1 Introduction .............................................................................................. 135
  5.2 The intention management module ...................................................... 136
    5.2.1 The discovery of intention ............................................................... 137
      5.2.1.1 The selection of the intention discovery method......................... 137
      5.2.1.2 The selection and definition of the features ............................... 140
      5.2.1.3 The selection of machine learning algorithms .......................... 142
    5.2.2 The budget estimation protocol ...................................................... 146
  5.3 The preference management module .................................................... 148
    5.3.1 The feature extraction process ....................................................... 149
      5.3.1.1 Item feature extraction ............................................................. 149
      5.3.1.2 Consumer feature extraction .................................................... 150
    5.3.2 The similarity matrix ................................................................. 152
  5.4 The production management module .................................................. 153
    5.4.1 The personalization algorithms ...................................................... 153
    5.4.2 The delivery measures .................................................................... 158
5.4.3 The prioritization algorithm ...................................................................................... 160
5.5 The performance improvement module ........................................................................ 162
  5.5.1 Evaluating the performance of the CAMPS ............................................................ 162
  5.5.2 Improving the CAMPS’s performance ................................................................. 166
5.6 Summary ..................................................................................................................... 169

Chapter 6 .......................................................................................................................... 170
  6.1 Introduction ................................................................................................................ 170
  6.2 The brownfield development of the CAMPS in COMPANY X’s infrastructure 171
    6.2.1 Deploy the event management module .............................................................. 174
    6.2.2 Deploy the preference management module ...................................................... 175
    6.2.3 Deploy the intention management module ......................................................... 178
    6.2.5 Deploy the performance improvement module ................................................ 185
  6.3 Evaluating the performance of the CAMPS .............................................................. 185
    6.3.1 Analyzing consumer data for the preparation phase ........................................ 187
    6.3.2 Analyzing consumer feedbacks for the test stage ............................................. 192
      6.3.2.1 Conversion ability ....................................................................................... 192
      6.3.2.2 User experience ......................................................................................... 196
      6.3.2.3 Value creation ............................................................................................ 199
  6.4 Implications to retailers ............................................................................................. 201
  6.5 Summary ..................................................................................................................... 202

Conclusion .......................................................................................................................... 203
  7.1 Summary of the research project .............................................................................. 203
  7.2 Major contributions .................................................................................................. 205
    7.2.1 Theoretical contributions ............................................................................... 205
    7.2.2 Methodological contributions ........................................................................... 206
Disclaimer

This thesis is prepared under the framework of “Conventions Industrielles de Formation par la Recherche (CIFRE) – CONVENTION No.2014/170” between “L’Association Nationale de la Recherche et de la Technologie (ANRT)” and “Search’XPR SAS (SXPR)”, which provide the financial resources to complete this research task.

The author of this thesis has been authorized by SXPR and its clients (i.e. B2C retailing firms) to access and analyze their documents, information, databases, records, and systems in order to complete the research. Meanwhile, the author is also subject to the confidentiality responsibility regulated by the non-disclosure agreement. To this regard, the name of the researched B2C firm is mentioned as “COMPANY X” and certain sensitive data and trade secrets related to the research have been kept undisclosed.
Introduction

0.1 Evolution of the business to consumer marketing

Mankind has a long history exchanging their goods in order to fulfill their diversified needs. In a marketplace where there are plenty of sellers and buyers, the fundamental and permanent business management question is: “how can firms better market their products to potential customers?” During the process to find an answer, firms and researchers find that creating and delivering marketing messages, namely the information of a product, price, promotion, and place, to potential customers can influence their shopping decisions (Dodds et al., 1991).

Customers choose to receive firms’ marketing information because a shopping decision, no matter an impulsive or a planned one, is not easy to make (Sprotles and Kendall, 1986). In order to evaluate its usefulness, consumers need to obtain information about the (physical and mental) utility of a product or service. If there are multiple options, they also need to compare the price, quality, service, and other criteria with an aim to choose the best one. Hence, consumers, no matter intentionally or unintentionally, are in permanent search for marketing information with an objective to substanti-
ate their shopping decisions. Marketing messages in newspapers, magazines, radio stations and TV channels facilitate their information gathering process and drive them to try new products/services in local stores (Kotler and Levy, 1969).

In view of the impact of marketing information on consumers’ purchase decisions, B2C (Business to Consumer) firms make great efforts to create attractive marketing messages and communicate them to as many potential customers as possible. Before the 90s’, firms made huge investments on TV commercials, radio ads, newspaper ads, billboards ads and direct mails (Borden, 1964). When the Internet age arrives, they construct catalogue websites, develop mobile applications, invest in the search engine optimization, and pay for online banners in order to enhance their visibility to consumers (Kaptein & Parvinen, 2015; Li, 2015).

From an economic perspective, non-customized marketing messages are less expensive to create and more suitable for a mass diffusion. However, they are also less relevant to the specific requirements and needs of each individual (Churchill, 1979). Since a human brain is inherently tuned to ignore irrelevant information (Posner and Petersen, 1990), it is not surprising that traditional “above the line” marketing measures (e.g. TV commercials, radios ads, newspaper ads, billboards, standard online banners) yield a very low conversion rate, because they do not specify audience or customize messages.

The drawbacks of the above standard marketing measures call for the advent of the marketing personalization technology. The concept of personalization originates from the retailing industry where firms adapt their products or services to meet the particular requirements of a specific customer or group of customers (defined by the American Heritage® Dictionary). In the B2C marketing domain, marketing personalization refers to the ability to create tailor-made marketing contents and diffuse them to a specific
consumer or consumer groups based on their personal preference, likes, and interests. Common personalized marketing contents include tailor-made advertisements, personalized recommendations, remarketing contents, and customized marketing emails.

Before the thrive of computer sciences and Internet, marketing personalization, even at a customer cluster level, is costly and time-consuming. As a result, tailor-making marketing offers were only possible in the luxury goods industry, where customers pay enough premiums to justify the additional efforts. Computer programs have significantly reduced the time to customize marketing messages and the Internet has greatly improved the efficiency of marketing information diffusion. As a result, marketing personalization is available for ordinary consumers. From a technical perspective, there are two kinds of marketing personalization approaches.

(i) The **first kind** seeks to create and distribute unique marketing contents for each customer based on his (or her) personal characteristics, needs, and situations (Kay and McCalla, 2012). It allows firms to attract and impress potential customers that are of great value and importance to their business. Individual-level marketing personalization necessitates the creation of consumer profiles for each consumer. The task requires sophisticated modeling techniques, expert knowledge, a tremendous amount of consumer data, as well as high-performance data processing and storage capabilities, which most B2C firms do not have. For these firms, personalizing marketing contents at the individual level is too costly and unrealistic.

(ii) As a compromise, firms switch to the **second kind** of marketing personalization approach, which aims at creating customized marketing contents for specific consumer clusters (Churchill, 2013). A consumer cluster is an agglomeration of consumers who
have common characteristics, needs, or situations in common. By customizing marketing contents at the cluster level, firms can diversify their marketing communication while achieving the economies of scale. With the help of computer programs, the identification of consumer clusters and the customization of marketing contents can be completed semi-automatically or automatically (Mobasher et al., 2000).

Regardless of their differences, both approaches strive to achieve the same objective, that is, to create and diffuse relevant marketing contents to potential customers in a more efficiently and cost-effective way so as to provoke their purchase intention, facilitate their decision-making process, and enhance their level of satisfaction and loyalty. Nowadays, we can find that marketing personalization technology has been applied to almost every step of a shopping journey (Leeflang et al., 2014), which is initiated by a preliminary shopping need (or impulse), substantiated by the process to collect and analyze relevant information of a product or service (e.g., attributes, specifications, conditions, expert recommendations, customer reviews), and concluded by the decision to purchase (or abandon) a specific product. (i) For example, the personalized customer experience created by personalized marketing measures demonstrate to customers that a firm is willing to walk extra miles to care about their personal requirements, needs, and likes. It allows firms to manage their relationship with individual customers and potential customers so as to enhance their engagement (Thongpapanl and Ashraf, 2011). (ii) When firms display preference-based advertisements to provoke the purchase intention of their potential customers, they have more chance than non-personalized ads to persuade a prospect and invite him (or her) to visit their (online and offline) stores (Subramaniam, 2014). (iii) In addition, firms may provide personalized promotion for customers’ favorite products. Such personalized marketing measure can enhance customers’ satisfaction level and loyalty to the brand (Turban et al., 2015).
Meanwhile, the prevalence of the marketing personalization has redefined the way consumers make their purchase decisions. (i) In the past, consumers might turn to a shop assistant or a friend when they need recommendations on which product to choose. Nowadays, they prefer to listen more to a marketing personalization system, which makes suggestions based on the analysis of their historical preferences, tastes, and behavioral patterns. (ii) Many years ago, only the VIP (very important person) customers of some premium brands have the privilege to enjoy personalized offers. Nowadays, ordinary consumers may switch to a new supermarket, bookstore, or sushi restaurant just because it provides tailor-made discount coupons applicable to their favorite products. (iii) For quite a long time, product catalogs, websites, and mobile apps have a standard appearance for every customer. Marketing personalization allows firms to create a unique shopping journey for every customer. The ambiance of exclusiveness created by a personalized shopping scenario (e.g. a user interface or some user experience) can reduce the psychological distance between a consumer and a brand/firm. At a time when customer experience is becoming the key success factor of B2C marketing, personalized marketing contents are the indispensable catalysts that motivate customers to engage with a firm and its brands.

To summarize, marketing personalization enables firms to better market their products to potential customers. Thanks to the advent of new technologies, firms in the mass market can personalize their marketing contents and ordinary consumers are able to enjoy the benefits of marketing personalization.
0.2 Challenges of B2C marketing

Consumers have a permanent appetite for exclusive and tailor-made shopping journeys. Though the personalization approach refreshed consumers’ view of B2C marketing contents, it also raised their expectation for more customized contents in marketing communications (Lee and Lin 2005). Therefore, marketing personalization reaches its development bottlenecks after the initial market success.

The author of this thesis works in a French company (i.e. Search’XPR) that helps B2C firms improve their ability to market products to consumers. In the daily work, the author has encountered a wide range of marketing challenges associated with the current marketing personalization approaches. To illustrate these challenges and their adverse consequences on firm’s business performance, we present the case of COMPANY X, a real client of Search’XPR. The name of the client is kept undisclosed due to the requirement of the confidentiality agreement.

0.2.1 A Motivation case

COMPANY X is a leading multinational sports equipment retailer in the world. It actively engages in the marketing and sales of sports apparel, footwear, equipment, and accessories. The firm has a full spectrum of retail and communication channels to reach potential customers. In addition to selling a large assortment of branded products (e.g. Nike, Adidas, Puma, Wilson...), COMPANY X aggressively promotes and sells its private label products, which focus on the balance between price and quality. The firm embraces innovation and regards it as the core competitive edge of its business. Due to this innovative culture, COMPANY X is keen on applying cutting-edge technologies to all retailing channels in order to reach and impress their potential customers.
COMPANY X is very proud of its marketing personalization system (MPS), which is capable of generating preference-based marketing contents and delivering them to potential customers from all retailing channels (e.g. physical stores, e-commerce website, and mobile commerce application). However, the marketing personalization service is also criticized by many consumers, who had unpleasant experience with it. Here are some typical examples.

(1) Complaint about the irrelevancy of personalized marketing contents. Many customers of COMPANY X claimed that shopping suggestions made by the firm's MPS were not relevant to their needs. For instance, a customer who consulted a pair of roller skates on COMPANY X's e-commerce website quickly abandoned the purchase idea because she thought her 3 years old daughter was too young to participate in roller skating. However, in the following weeks, she kept receiving emails from COMPANY X suggesting her to consider roller skaters of different brands, which confused her and resulted in her decision to unsubscribe to the firm's marketing emails.

(2) Complaint about the quality of personalized marketing contents. COMPANY X made a customer survey on their satisfaction with the personalized marketing contents and discovered that customers were disappointed by the lack of novelty and originality of the firm's MPS. The survey found that many customers visit COMPANY X's website and mobile application to look for innovative shopping ideas. However, since most shopping recommendations were generated based on customers’ choices, they rarely played a role in enlightening customers.

(3) Complaint about the uncoordinated personalized marketing contents. Customers claimed that sometimes COMPANY X's personalized marketing contents were con-
fusing. *For example,* a frequent customer of COMPANY X consulted a high-quality tennis racket on COMPANY X’s website and purchased it from the firm’s local store. He paid the full price and scanned his membership card to accumulate the “points”, a typical CRM trick of all firms. One week later, however, COMPANY X sent him a digital coupon inviting him to purchase the same tennis racket online with 25 euros of discount. This attractive offer did not motivate the customer. On the contrary, he was very disappointed at the late arrival of the coupon.

(4) **Complaint about the quantity and frequency of personalized marketing contents.** Customers complained that COMPANY X presented too many recommendations on its website and mobile application. Besides, the firm’s marketing emails were too frequent. Excessive marketing contents had disrupted customers’ shopping journey, prevented them from making their own decisions, and deprived them of the pleasure of shopping. Based on COMPANY X’s internal analysis, the complaint about excessive marketing contents is the key reason that drives customers to unsubscribe COMPANY X’s marketing emails, block its online advertisements, abandon its website, and uninstall its mobile application.

(5) **Complaint about the breaches of privacy.** A great number of COMPANY X’s customers expressed their aversion to remarketing, which is a new technology allowing COMPANY X to display personalized marketing contents to its customers and potential customers when they visited other websites and mobile applications. In particular, these personalized remarketing contents are often items that have already been consulted by the same customer on COMPANY X’s website or mobile application. Due to
the lack of transparency on the way remarketing technology tracks and monitors Internet users, it is perceived by many consumers as a spying tool that breaches the boundary of privacy (Weinsberg et al., 2015), which must be removed immediately.

Facing these criticisms and complaints, the Digital Marketing Director of COMPANY X is in a dilemma. The MPS employs the state-of-art technologies and infrastructures. It has cost the firm millions of euros to build and to maintain the system. Abandoning it not only means the loss of a huge capital investment, but also a disaster to the firm’s marketing operation. However, the root cause of the criticisms and complaints are complex and systematic. Fixing the MPS before customers become impatient and disappointed is a challenging task. Therefore, COMPANY X invited SearchXPR and the author of this thesis to help. In this project, the author was responsible for problem analysis, conception, methodology, and system design. He was supported by several colleagues who were responsible for programming and implementation.

0.2.2 Analysis of challenges in B2C marketing

The dilemma of COMPANY X is a typical management challenge encountered by many B2C firms, which base their marketing personalization models on the analysis of customers’ past and current behavior (Ghorab et al., 2013). Though firms are able to customize advertisements, recommendations, emails, and remarketing contents for each customer (Zeng et al., 2002; Arya et al., 2014, Jagan and Rajagopalan, 2015), these marketing messages do not necessarily improve the effectiveness and efficiency of B2C marketing efforts, because they are not relevant or useful to customers.

The argument that personalized information is not necessarily useful or relevant to customers may surprise many marketing practitioners, who regard personalization as
the panacea for the B2C marketing issues. However, we can easily present many cases to prove this argument. For example, a consumer Michael may take into account his girlfriend’s preference when he chooses a birthday gift for her. In this case, Michael’s own preference, selection criteria, and past shopping behavior do not have a lot to do with his shopping decision. Similarly, a consumer Lucy may click the web page of a new movie just because of the curiosity. Nevertheless, it does not indicate her like (or dislike) of this movie or movies of this genre. If we do not know why a consumer views some products and how he (or she) think of them, it is arbitrary and dangerous to correlate these products to his (or her) preferences and needs. In COMPANY X’s case, the first and the second type of customer complaints provide a perfect footnote to this argument.

Some inherent imperfections of the personalization technology may also undermine the relevance and usefulness of personalized marketing contents. Here are some examples. (i) It is very hard to draw any inferences for new customers who have not yet provided sufficient information for analysis (Lam et al., 2008). This is known as the “cold start problem”. (ii) If preference data is sparsely populated, firms may not be able to obtain sufficient “similar cases” for a customer in order to make meaningful predictions and inferences (Grčar et al., 2005). This is known as the “data sparsity challenge”. (iii) In order to enhance the accuracy of preference prediction, firms need to fit their prediction model to a customer’s past choices. Accordingly, they become less capable of recommending novel items and innovative shopping ideas that may potentially be enlightening to a customer (McNee et al., 2006). This is known as the “accuracy-diversity dilemma”. (iv) A customer may view an item somewhere and purchase it elsewhere. If firms are not able to synchronize consumer knowledge across their sales and marketing channels, they risk of basing their predictions on incomplete and biased information (Shi, 2016). This is known as the “cross-channel challenge”. In COMPANY X’s
case, the second and the third type of customer complaints reinforce the argument. The encouraging news is that many researchers are working hard to tackle these problems.

However, even technology improvement can significantly enhance the relevance and usefulness of personalized marketing contents, they can still be rejected or ignored by a consumer, who is not able to perceive the relevance and usefulness. Such inefficiency of marketing communication can be caused by the inappropriate design, layout, form, and presentation of the message (Chen and Pu, 2014) or it can be attributed to the inappropriate moment when firms approach their customers (Fischer, 2012). Researchers found that these non-content factors have a psychological impact on consumers’ attitude and receptiveness to personalized marketing messages (Li, 2016). The fourth and the fifth type of customer complaints in COMPANY X’s case provide more evidence to this argument.

Based on the above three arguments, we can make a review of the drawbacks of COMPANY X’s MPS, which may enlighten many B2C firms with the same situation.

(1) The MPS does not analyze customer’s instantaneous intentions and needs. Instead, it arbitrarily assumes that such knowledge can always be inferred from customers’ past and present choices. Hence, its predictions are not always relevant to customers’ needs.

(2) Since the MPS has no knowledge of customers’ instantaneous intention and needs, it must present a lot of shopping recommendations in order to enhance its chance of success. This leads to customer information overload, confusion and choice difficulty (Chen et al., 2009; Bollen et al., 2010).
3. The MPS has no visibility to the cross-channel shopping behavior because it does not consolidate customer data obtained from different sales platforms. Hence, marketing contents generated by the MPS can be inconsistent or uncoordinated, which result in customer confusion and complaint.

4. The MPS does not analyze customers’ cognitive state in real time. Therefore, it risks of disturbing customers by sending marketing messages at an inappropriate moment or through an inappropriate medium.

5. The MPS does not have an effective privacy management mechanism. Accordingly, some of its personalized marketing contents may irritate customers who pay great attention to their privacy.

0.2.3 Research goal and objectives

Based on the above analysis, it is evident that personalization is not the panacea for all the B2C marketing issues. In view of this managerial challenge, the main goal of this thesis is to research and propose new methodology enabling firms to enhance the relevance and usefulness of their personalized marketing contents so that they can better market their products to customers.

Using COMPANY X’s case as an example, we identify four improvement needs for MPS (i.e. research objectives) that can help achieve our research goal. Though more comprehensive and profound analysis is performed in Chapter 2 and 3 to justify them, we make a brief discussion here to rationalize our propositions.

1. MPS should keep track of customers’ instantaneous intention. The intention is an aim that guides action (defined by the American Heritage® Dictionary). By clarifying a
customer’s instantaneous intention, MPS is able to choose the most appropriate personalization strategy that complies with his (or her) subsequent actions. Besides, real-time knowledge of intention enables MPS to adjust the personalization strategy when a customer’s intention changes. Accordingly, MPS may have more chance to succeed in influencing customers’ shopping decisions.

(2) MRS must base their personalization strategy on a holistic view of customers’ cross-channel shopping journey. Comprehensive and complete customer knowledge is the foundation of relevant and useful recommendations. Not only does it help MPS evaluate the maturity of a customer contact opportunity, but also gives MPS the indispensable information to create consistent and relevant marketing messages.

(3) MPS must choose the appropriate medium and moment of contact based on a customer’s preference and cognitive state. The brain of mankind is tuned to focus on information that is relevant to the primary task (Posner and Petersen, 1990). Hence, the best moment and medium of contact change with customers and situations. Delivering messages to customers at their preferred moment through their preferred medium reduces their chances of being rejected without reading, which is one of the major threats to the success of MPS.

(4) MPS need to incorporate a customer privacy management mechanism. Each customer has his (or her) unique requirements and preference for privacy (Shyong et al., 2006). Embracing and respecting such difference can prevent MPS from irritating privacy-sensitive customers.

0.3 Structure of the thesis

This thesis is organized in three parts (related work, methodology, and application) and
completed by an annex providing fundamental knowledge related to the research. Figure 0.1 summarizes the composition of this thesis.

Chapter 1 & 2 constitute the first part of the thesis, where we present the background concepts and the state-of-art research works related to B2C marketing personalization.

- **In Chapter 1**, we present the *background concepts* related to the B2C marketing personalization research. We begin by presenting the consumer shopping journey, shopping channels, and the cross-channel shopping behaviors. Next, we introduce the personalization techniques used by the existing marketing systems. After that, we discuss the definition of context, review various kinds of contextual factors, and their importance to cross-channel marketing. Then, we make an overview of the data mining process and machine learning techniques used to identify contextual factors. We conclude this chapter by discussing the big data analysis issues associated with the machine learning techniques.
In Chapter 2, we review the state-of-art research works related to B2C marketing personalization and identify their improvement needs. We start with a general review of related research papers to highlight the focus and hot topics of the research on context-aware marketing personalization for consumers. Afterward, we make a review of the current model to explain a cross-channel shopping journey, present the existing knowledge on intentions of cross-channel shoppers, and highlight their limitations. Next, we discuss current customer tracking techniques and the main challenges to track the cross-channel shopping behavior. Then, we discuss the current process-driven personalization method and its drawbacks. After that, we review the existing contextualization methods and their limitations. Finally, we present the current performance tracking and improvement methods for B2C marketing personalization systems and highlight their limitations. These discussions lay a foundation for our proposition.

The second part of the thesis is composed of three chapters. In this part, we present our theories, framework, methodologies, and techniques.

In Chapter 3, we put forward a cognition-behavior model to unify various kinds of shopping intentions and journeys. Using this theoretical framework, we propose the intention-based personalization methodology and the cognition-based contextualization methodology, which can enhance the relevance and usefulness of personalized marketing contents.

In Chapter 4, we propose an event streaming methodology to encode, consolidate, and analyze cross-channel shopping behaviors. Thanks to this methodology, firms can determine the real-time intention and cognitive state of their customers. It also allows firms to track the variability of customer intentions so as to
make necessary adjustments to the marketing personalization strategy.

- Based on the methodologies presented in Chapter 3 & 4, we propose a context-aware marketing personalization system (CAMPS) in Chapter 5 capable of identifying consumers’ real-time intention, adjust its personalization strategy and analyze consumer feedbacks to make improvements. The business logic, architecture, and modules of the CAMPS are detailed respectively.

The third part of this thesis, which is composed of two chapters (Chapter 6 & 7), aims at applying our proposed theories, methodologies, and system to the real B2C marketing context.

- In Chapter 6, we apply the CAMPS concept to COMPANY X’s situation and demonstrate that a CAMPS can be created by revamping a legacy marketing personalization system. This is attractive to most firms who expect huge benefits but do not want to take big risks or invest too many resources. Meanwhile, we compare the performance of the CAMPS and COMPANY X’s legacy system in the same experimental conditions to prove that our theory and methodology can improve firms’ B2C marketing capability.

- In Chapter 7, we summarize the main findings, contributions, and provide the implementation suggestions to firms and marketers who want to embrace the challenges of the cross-channel shopping era and prepare their marketing personalization systems for more intense competitions. We also discuss the limitations of our methodology and highlight the directions and suggestions for the future research.
Chapter 1

From propaganda to interaction: the evolution of the B2C marketing

1.1 Introduction

The rapid evolvement of consumer needs and technology is reshaping the way firms generating and presenting personalized B2C marketing contents to their customers. Nowadays, B2C marketing personalization is a highly sophisticated technology that involves knowledge of a wide range of research disciplines (Fan and Poole, 2006). Before talking about how to make new contributions to this domain, it is essential to make a comprehensive overview of the fundamentals in consumer needs, retail channels, personalization, contextualization, machine learning, and big data. Knowledge in these disciplines constitutes the foundation of B2C marketing personalization.

This chapter seeks to lay a foundation for the subsequent discussions by presenting the background concepts, theories, methodologies, and knowledge related to the B2C marketing personalization topic.

We start with a discussion of consumer journey, their decision-making pattern, and the
shopping channels (Section 1.2). Next, we make an overview of the methods to generate personalized marketing contents for consumers (Section 1.3). Then, we discuss the importance of context in the personalization technology as well as the identified contextual factors (Section 1.4). After that, we review the machine learning techniques that enable firms to discover context automatically and unobtrusively (Section 1.5). Finally, we discuss the big data issues that need to be addressed in the data mining process (Section 1.6). The Section 1.7 serves as the summary of the Chapter 1.

1.2 Shopping journey, consumer decision, and sales channels

1.2.1 Shopping journey

A shopping journey is a process or progress to fulfill a shopping need. It is initiated by a preliminary shopping need (or impulse), substantiated by the process to discover and analyze information related to product (e.g. attributes, specifications, conditions, expert recommendations, customer reviews), and concluded by the decision to purchase (or abandon) a specific product.

Until recently, a shopping journey is mainly controlled by firms, who decide where to sell products, how to promote, and which price to sell (Nenonen et al., 2008). In most cases, consumers have to go through a few “touch points” designed and put in place by firms to collect and analyze product information so that they can make a final decision. However, with everything and everyone being connected to the Internet, consumers can obtain product information and place their orders from a wide range of sources beyond the control of a firm (Edelman, 2010). In addition, the powerful search engine and social network multiply consumers’ exposure to new products and shopping ideas (Kucuk and Krishnamurthy, 2007). As a result, consumers becomes more and more
capable of designing their own shopping journey.

B2C marketing personalization aims at providing consumers with useful product or marketing information in order to accelerate a shopping journey or to change its course and direct it towards a specific firm or product.

1.2.2 Consumer decision-making

A shopping journey is featured by a series of consumer decisions, which are conclusions or judgments of a consumer reached after analysis and consideration (Olshavsky and Granbois, 1979). Consumers make decisions on a wide range of topics: the decision to examine a product in detail, the decision to compare two products or offers, the decision to put a product on the wish list, the decision to purchase a product, and the decision on the delivery measure and payment method etc. (Payne et al., 1991).

Decision-making is a process to evaluate a shopping idea, a product, or an offer provided by firms (Sprotles and Kendall, 1986). Various assessment criteria, such as product utility, function, specification, design, price, service and so on, can be applied during this process (Crotts, 1999). Each consumer may have his (or her) own list and priority of assessment criteria (Punj and Stewart, 1983). Besides, the composition of the list and the weight of each criterion may change with the shopping objectives and contextual factors (Menon and Kahn, 1995).

B2C marketing personalization seeks to achieve two objectives. (1) The first objective is to directly influence a consumer so that he (or she) can make a purchase decision favorable to a specific firm or product. (2) If the first objective cannot be achieved immediately, it is also beneficial to influence the decision-making process of a consumer so as to cultivate his (or her) positive impression on a specific firm or product.
1.2.3 Sales channels

In order to gather and assess product information, consumers must visit a firm’s sales channels. A sales channel is a way to bring products to markets so that they can be presented and sold to customers.

Until recently, consumers have two ways to discover and purchase products: in-store or online. In-store shopping refers to the browse, compare, experience, and purchase behavior occurred in a physical store. Online shopping refers to the behavior of searching, evaluating, comparing, and purchasing products in virtual shopping platforms such as websites, mobile Apps, and so on (Vanheems, 2009).

The advantage of a store (e.g. department stores, showrooms, supermarkets and convenient stores) is that customers can have physical touch with a product, which gives them first-handed and intuitive feelings of a product. However, a store has its temporal and spatial constraints, which limit its customer-serving capability (Schoenbachler and Gordon, 2002). On the contrary, an online store (e.g. e-commerce website or mobile commerce application) never closes. In addition, there is almost no limit to the amount of product and customer review information it can present to consumers (Hu et al., 2008). Besides, online stores often offer a competitive price and flexible delivery time, which make their offers attractive to many consumers. However, an online shop is not able to provide customers with the touch and feel experience, which is sometimes indispensable in the shopping journey (Lieber and Syverson, 2012). Today, the ubiquity of the Internet-connected devices has enabled dozens of new transactional touch points, blurring the boundary between online and offline. As a result, cross-channel shopping behaviors becomes prevalent.
1.2.4 Cross-channel shopping behavior

Cross-channel shopping behavior is described as examining a product in one sales channel and purchase it elsewhere (Chiu et al., 2011). The distinct capabilities of different channels are the main motivation of cross-channel shopping behaviors (Chatterjee, 2010). In its narrow sense, cross-channel refers to the alternative use of a firm’s multiple sales channels. In a broad sense, it can also be referred to as the switch from one firm to another during the journey to evaluate and purchase a specific product. The key motivation of cross-channel shopping is to avoid making bad decisions because of the inadequacy of information. By referring to multiple channels, consumers aim at finding lower prices, better services, first-handed product reviews, or simply more comprehensive introductions of a product (Heitz-Spahn, 2013).

In early times, consumers prefer to switch between channels in the same class (intra-class cross-channel behavior). For example, they may visit several furniture stores before buying a closet that fits the theme of their living room. Similarly, they may open several B2C websites to compare their offers for the same product in order to get the best deal. Such cross-channel behaviors are not new to firms, which have already calculated the corresponding risks in their business model (Verhoef et al., 2015).

However, with the prevalence of mobile computing and internet technologies, recent cross-channel behaviors are featured by their inter-class nature, also known as the simultaneous use of online and offline channels (Shi, 2016). For instance, consumers may test a digital camera in a local BestBuy mart and purchase it using Amazon’s mobile application. As a result, retailing business model and its corresponding marketing personalization strategies are under threat of free-riders, who use a specific channel only for free product information and first-handed user experience (Verhoef et al., 2015; Shi,
Recently, due to the rapid development of “marketplace”, consumers are more accustomed to the cross-channel shopping behaviors (Cao and Li, 2015). A marketplace is a retailing platform owned by a major retailer but is open for other retailers to exhibit and sell products. The world’s online retailing giant Amazon and Taobao are both marketplaces. Since different vendors use the same modality to display product and service information, it is quite easy for consumers to compare their offers and conditions and there is almost no switching cost. Accordingly, consumers may use one vendor to fulfill their research needs (because this vendor provides detailed product information) and buy the product from another vendor (because it offers a lower price or better service).

### 1.3 Personalization

Consumers are in permanent search for relevant marketing information. Today, firms use personalization to fulfill their information requirements. Personalization is defined as any action that adapts information or services to the needs of a particular user or a set of users, taking advantage of the knowledge gained from the users’ navigational behavior and individual interests, in combination with the content and the environment (Eirinaki and Vazirgiannis, 2003). Personalization is used to enhance customer satisfaction, improve sales conversion, and facilitate purchase decision (Vesanen, 2007). The objective of a personalization system is to provide users with the information they want or need, without expecting from them to ask for it explicitly (Mulvenna et al., 2000; Montgomery and Smith, 2009).

In this section, we make an overview of the personalization methodologies used by B2C
firms in their marketing personalization system. A detailed description of each methodology is available in the Annex.

1.3.1 Consumer profiling

A consumer profile is a portrait telling us the characteristics and the behavior patterns (i.e. purchase and consumption of products) of one customer or a group of them. The characteristics recorded in a profile range from demographic, economic, geographic, and psychographic features to preferences, shopping patterns, purchase history, and financial capability (Gunter and Furnham, 2014). There are three methods to create a consumer profile.

(1) The first method, known as “labeling”, deals with factual information that does not change frequently.

(2) The second method, known as “rule-based profiling”, seeks to depict conditional facts about one or a group of consumers.

(3) The third method, known as “sequence”, aims at portraying the behavioral patterns that signify the owner(s) of the profile.

Consumer profiles are widely used in personalization tasks to identify “similar neighbors” for the active consumer, whose preferences need to be predicted. The underlying assumption is that consumers may keep to the characteristics and behavior patterns that constitute their profiles. When the assumption is not fulfilled, consumer profiles cannot be used to make predictions.

1.3.2 Content-based personalization

Content-based personalization seeks to refine consumers’ interests in their precedent choices and use acquired knowledge to predict their subsequent needs. This approach
is able to identify items and topics similar to those have been liked by a consumer in the past (Van Meteren and Van Someren, 2000). A typical content-based personalization system is composed of three key components, namely an item analyzer, an interest identifier, and a filtering module (Lops et al., 2011).

1. The item analyzer represents items using structured features extracted in contents.
2. The interest identifier discovers the attributes liked and detested by a consumer.
3. The filtering module finds items relevant to consumers’ interests.

Content-based personalization is indispensable in the B2C marketing personalization tasks. However, it has some limitations. (1) It requires historical consumer feedbacks, which makes it vulnerable to the impact of the cold start problem. (2) It is not capable of presenting novel items (i.e. the accuracy-diversity dilemma). Hence, it is often reinforced by some other methods to form a hybrid approach.

1.3.3 Collaborative filtering

Collaborative filtering (CF) approaches aim at suggesting personalized products to consumers by predicting their rating patterns on items based on their explicit and implicit rating history. There are two types of collaborative filtering methods: memory-based methods and model-based methods (Su and Khoshgoftaar, 2009).

1. Memory-based methods perform personalization tasks based on item-item or user-user relations. The item-based method predicts an active user’s rating of a new item based on his (or her) ratings of similar items in the past (Sarwar et al., 2001). The user-based method discovers a group of users similar to the active user, and use their rating history of the new item to predict the active user’s rating of the new item (Marlin, 2003).
Model-based methods seek to predict an active user’s rating of new items by modeling the components and processes that determine the rating pattern (Sarwat et al., 2002). Common model-based methods include Naïve Bayes (NB), Associated Rule Mining (ARM), Clustering, and latent factor models (Koren, 2008).

CF is another indispensable method widely used by many B2C firms to generate personalized marketing contents. The major challenge of the CF method is the data sparsity problem. To overcome the problem, researchers have proposed methods to collect consumer data from different sources and consolidate them.

### 1.3.4 Hybrid approach

The hybrid approach is referred to the simultaneous use of more than one personalization method. In general, there are two hybrid approaches to integrating multiple personalization methods.

- **Mixing** is referred to as the combination of items predicted by different personalization or non-personalization methods with a purpose to diversify personalization results.

- **Twisting** is referred to as the use of several personalization methods in personalization modeling. Common twisting techniques include:
  
  - *using a two-staged (e.g. content-based + CF method) approach to predict ratings*
  
  - *preparing a dataset using non-personalization features and making predictions using a personalization method*
  
  - *using several collaborative filtering methods in parameter learning and prediction,*
  
  - *re-ranking results predicted by a personalization model using some non-personalization features.*
In practice, most firms use a hybrid approach to obtain more balanced personalization results which take into account the relevance and novelty (Burke, 2002; Adomavicius and Tuzhilin, 2005). To achieve this objective and to enhance the relevance and usefulness of personalization results, it is necessary to incorporate the knowledge of context and the contextual factors (Gorgoglione et al., 2006; Burke, 2007; Palmisano et al., 2008; Rooderkerk et al., 2011).

1.4 Context and contextual factors

In this section, we make a discussion on the definition of context and make an overview of the contextual factors used in the B2C marketing personalization.

1.4.1 Defining context

Many researchers have tried to define the term context. Based on a comprehensive review, Dey et al. discovered two clusters of context definitions and highlighted their limitations (Dey et al., 2001).

(1) The first cluster defined context by itemizing context examples. In these definitions, a context was referred to as the location, identities, time of the day, season, temperature, environment, and so forth (Schilit and Theimer, 1994; Brown et al., 1997; Ryan et al., 1997). Dey et al. claim that definitions based on examples were difficult to apply as they cannot help researchers discover new types of contextual factors (Dey et al., 2001).

(2) The second cluster used a more generalized method. It defined context by its synonyms. Based on such method, context was defined as “the elements of the users’ environment that the computer knows about” (Brown, 1995), “the state of the application’s surroundings” (Ward et al., 1997), “the situation of the user” (Franklin and Flaschbart,
Dey et al. argue that though the second cluster of definitions is more generalized, it provides little clue to analyze the components of context. As a result, definitions of this type cannot be used to identify contextual factors in real-world user cases. Following the analysis, Dey et al. provided their definition of context (Dey and Abowd, 2000b):

“Context is any information that can be used to characterize the situation of entities (i.e., whether a person, place, or object) that are considered relevant to the interaction between a user and an application, including the user and the application themselves.”

Compared to the previous definitions, we believe this definition has the following advantages. (1) It highlights the active role of context in the interaction between a user and an application, which is the theoretical foundation of the context-aware systems. (2) It claims that context can be identified by analyzing the “information that characterizes the situation of entities”, meaning that context must be any factor that can be identified using information or data gathered by virtual or physical sensors. Considering these advantages, we accept the definition proposed by Dey et al. in this research.

Context is essential in enhancing the relevance and usefulness of marketing personalization in the cross-channel shopping context. The knowledge of context enables firms to provide more relevant information to consumers’ current task, to choose the right moment to interact with consumers, and to select the right communication medium to deliver the message. For instance, if a recommendation engine detects that a consumer consistently pays more attention to the performance of products than to their prices in his (or her) shopping journey, proposing higher-end products to him (or her) is apparently a better strategy. When consumers are busy examining a product, it makes no sense to disturbing them. It is agreeable to send product and promotion information to
consumers proactively when they need it.

1.4.2 Contextual factors

Achieving the above functionalities requires firms to identify and make use of contextual factors. In view of the importance of context on the relevance and quality of B2C marketing personalization, we make an overview of the contextual factors frequently used in the B2C marketing personalization domain.

1.4.2.1 Position

A position is a place where a consumer can be located. A consumer’s position can be described by the longitude-dimension-altitude system or the relative coordinates in a coordinate system. In the B2C marketing personalization practice, position has two main functions (Ono et al., 2007; Park et al., 2007; Yang et al., 2008; Shi, 2016).

(1) Position information can enhance the relevance of personalized recommendation by filter or exclude personalization inputs or results that are too far away from the active user (Abbar et al., 2009; Fang et al., 2012; Ström et al., 2014). For example, large airports often have several terminals and each terminal is divided into a few public and restricted areas (i.e. areas that require security clearance). Position information allows personalization systems to filter out unreachable stores and suggest more relevant duty-free shops, cafes, and restaurants to tourists (Jones et al., 2004; Mastoris et al., 2013). In addition, positions can also be used to generate location-based promotions. Researchers have found that location-based promotions not only stimulate immediate purchase impulse but also affect future purchase decisions (Luo et al., 2013).

(2) Position information can also be used to determine the real-time interests of consumers.
When a consumer stops in front of a shelf or a booth in the physical store and starts interacting with items displayed nearby, it is likely that he (or she) is attracted by some specific products. By analyzing the characteristics of the items consulted by this consumer, firms can infer the real-time interest of him (or her). In addition, firms can construct a traveling trajectory for a consumer by recording and analyzing the changes of position over time. By analyzing the trajectory, firms may find some specific shopping patterns or habits demonstrated by the consumer. Automated and unobtrusive perception, record, and analysis of a consumer’s real-time position information enable firms to discover his preferences and habits without disturbing him (Yang et al., 2008). It has a major impact on the in-store shopping experience because shop assistants no longer need to follow their customers in order to find the right moment to provide supports or shopping advice (Yin et al., 2013).

As mobile devices are becoming increasingly indispensable in consumers’ daily life, consumers indoor and outdoor position information are more and more accurate and available (Zheng et al., 2014). To succeed in attracting consumer attention and usage, a marketing personalization system must make creative use of consumers’ position information. However, there are still some key problems to be solved. For instance, it is still a challenge to correlate consumers’ position information to other critical information (e.g. behavior, interest) so as to have a profound understanding of their shopping intention and decision-making pattern. New research efforts need to be made in order to close gaps of this kind.

1.4.2.2 Time

*Time* is a non-spatial continuum in which events occur in apparently irreversible succession from the past through the present to the future. In B2C marketing personalization domain,
it can be used to determine consumers’ behavioral pattern, preference, and urgency of their task (Wang et al., 2012; Braunhofer et al., 2013).

(1) Time as a contextual factor provides rationales to many consumer behaviors (Baltrunas and Amatriain, 2009), which can be explained by the time when they occur (Lee et al., 2010). For example, the download volume of restaurant coupons increases significantly when lunchtime is coming. Similarly, the click-through rate of marketing contents delivered during business hours is often low, because most consumers need to focus themselves on work and have no time for advertisements. Another example is the holiday shopping behavior. Affected by the happy atmosphere, consumers tend to increase their shopping spending and buy unwanted items. By correlating a specific consumer behavior to the time of its occurrence, firms are able to uncover the underlying motivation. Such knowledge may help enhance the relevance and usefulness of marketing personalization (Lamche et al., 2015).

(2) The duration of shopping time as a contextual factor can be used to evaluate consumers’ interest as well as their readiness to personalized marketing contents (Mallat et al., 2009). If consumers spend a long time interacting with an item, it not only indicates that they are interested in this item but also means that they are not in a rush to make a decision. This might be a good moment to present some marketing contents to extend their knowledge and accelerate their shopping decisions. When consumers are in a hurry, it means both a challenge and an opportunity to firms (Lin and Chen, 2013). On one hand, consumers may not have time for recommendations. On the other hand, firms stand a chance to sell their product as long as it facilitates consumers’ decision-making process and fulfills their needs. To this regard, choosing the right moment is essential (Fischer, 2012; Wang and Yang, 2013).
The timestamp as a clue help firms uncover the shopping decision-making patterns (Cho et al., 2002). Using timestamps to track a consumer’s online and in-store behaviors, firms can construct a behavior data stream for him (or her). It allows firms to track specific behaviors and compute their frequency and velocity (Kim et al., 2005). Based on such information, firms can analyze the intensity and magnitude of behaviors, which are indicators of the underlying intentions of a consumer. At the same time, behavior data stream enables firms to uncover patterns and habits behind intricate behaviors (Bucklin and Sismeiro, 2009). Based on such knowledge, firms can predict their subsequent behaviors and propose relevant personalization contents to them in order to foster their shopping journey and enhance their satisfaction.

1.4.2.3 State of mind

A state of mind can be referred to the state of a person’s cognitive process or a temporal mental condition in which the qualities of a state are relatively constant even though the state itself may be dynamic (Baltrunas, 2011; Braunhofer et al., 2013; Kim et al., 2014). Researchers found that it has a significant impact on consumers’ decision making process and behavioral patterns (Yu et al., 2006; Wang et al., 2012; Panniello et al., 2015).

Consumers’ state of mind and their purchase intentions are closely related (Ono et al., 2007; Baltrunas et al., 2012). According to some researchers, a positive state of mind can boost consumers’ curiosity as well as the ability of memorization, imagination and association (Campos et al., 2013), which enable them to adapt to the shopping environment and interact with the products and humans within so as to enjoy the shopping process (Underhill, 2010). As a result, happy shoppers spend more time and money on shopping and they are also more willing to encourage other consumers to visit the same shop (Arnold and Reynolds, 2009). On the contrary, a negative state of mind inhibits
consumers’ purchase intention. The conclusion is validated by experiments (Broekeier et al., 2008).

(2) Consumers’ state of mind has a pervasive impact on their attitude and reaction to a firm’s personalized marketing propositions. Researchers found that a pleasant and relaxed state of mind can reduce consumers’ hostility and resistance to unexpected information (e.g. personalized marketing information) from firms (Schwarz, 2013). Hence, consumers are more willing to interact with personalized information proposed by firms (Lee et al., 2010). On the contrary, stress or negative state of mind provoked by a pressing deadline, a challenging task, or some unpleasant experience reduces consumers’ willingness and ability to process unexpected information. As a result, they either turn off the marketing personalization option or simply ignore the personalized information (Shi et Marini, 2016).

1.4.2.4 Shopping goal

A shopping goal is an objective that consumers want to achieve in the shopping journey. Researchers found that consumers’ attitude and reaction to marketing contents and personalized recommendations are affected by several factors that determine their shopping goals.

(1) The first factor to consider is the motivation of a purchase. That is, whether consumers purchase an item for themselves or for other people (Gorgoglione et al., 2006). In the first case, consumers’ attitude and reaction to the recommended contents are dependent on their own assessing criteria. In the second case, however, they may take into account the opinion of the person that a purchase is intended to. Such difference may lead to distinct reactions to recommendations.
(2) The second factor to consider is the concreteness of the shopping goal (Ariely et al., 2006). When consumers have a concrete shopping goal, they are less likely to listen to marketing personalization systems. However, they might be interested in recommended items which fit with the item they want to buy. When consumers’ shopping goal is vague, they may need advices from a marketing personalization system with which they can make the best decision.

1.4.2.5 Shopping budget

A shopping budget is the total amount of money allocated by a consumer for the purchase of a product. Researchers find that consumers’ opinion on personalized marketing contents is contingent with this monetary contextual factor (Baltrunas et al. 2012).

Consumers’ shopping budget not only represents their buying power but also indicates their preference (Braunhofer et al., 2013). If a consumer prepares a budget that is more than sufficient to buy a high-end product of its kind, it suggests that the consumer pays more attention to the quality and performance than to the price. If a consumer prepares a budget that is merely enough for a low-end product, it indicates that his focus is the basic function. Shopping budget changes with the consumer and with the shopping objective. Firms need to adjust their personalization strategies based on this contextual knowledge so as to make marketing contents relevant and useful (Shi et al., 2015).

1.4.2.6 User Interface

A user interface is a system of interaction or communication between a computer and another entity (e.g. a human, a device, a network). Information created by a marketing personalization system needs to be diffused to consumers through a user interface. From a consumer’s perspective, a user interface is a layout or design of interactive elements of a
computer program, an online service, or an electronic device (Knijnenburg et al., 2012). In online and cross-channel shopping contexts, the user interface is an important contextual factor which can impact consumers’ decision-making and behavioral patterns in the following aspects.

(1) *The user experience of a marketing personalization system starts with its appearance and layout,* which is composed of pictures, interactive objects, and text that varies in color, size, and style (Konstan and Riedl, 2012). Researchers find that improving a user interface with its position, design, layout, typography, color scheme, and information organizing structure can significantly improve consumers’ impression on personalized marketing contents (Cosley et al., 2003; Pu et al., 2011; Hu and Pu, 2011; Chen and Pu, 2014). Poorly designed interfaces cast a shadow on the credibility of a marketing personalization system (Tintarev and Masthoff, 2007; Knijnenburg et al., 2012).

(2) *An interactive user interface can enhance consumers’ receptiveness to personalized marketing contents* (Swearingen and Sinha, 2002). Providing interactive functions on the user interface enables consumers to have a more intuitive understanding of the personalization process so that their confidence in the proposed marketing contents can be enhanced (Spiekermann and Paraschiv, 2002). In the same time, interactivity allows consumers to participate in the personalization process by themselves, which makes personalized contents more relevant to them (O’Donovan et al., 2008; Parra et al., 2014).

(3) A *concise and intuitive user interface can enhance consumers’ receptiveness to personalized marketing contents* (Scheibeheenne et al., 2010). (Chen et al., 2009). Although personalization technologies significantly reduce the cost and time to produce personalized marketing contents, it is in appropriate to display all of them to a consumer at a time,
because excessive information can lead to consumer information overload and choice
difficulty (Bollen et al., 2010).

(4) Besides, the specification and capability of a user’s device may have an influence on his (or 
her) reaction to the personalization contents (Boutemedjet and Ziou, 2008; Su et al., 2010),
since a proper display enhances the attractiveness and credibility of personalization 
contents (Chen and Pu, 2010; Knijnenburg et al., 2012). Firms need to identify the 
most effective user interface for a consumer and adapt personalized marketing con-
tents to the interface in order to provide a better user experience for consumers (Shi 
and Ghedira, 2016).

1.4.2.7 Other contextual factors

In addition, the physical condition that surrounds a user is also important, because 
contextual information such as local temperature, traffic situation, noise level, and con-
gestion can affect his (or her) decision model and behavioral pattern (Park et al., 2007; 
Hidasi and Tikk, 2010; Hussein et al., 2013; Lamche et al., 2015).

1.4.3 Discovering contextual factors

The ability to discover and make use of contextual factors is called context-awareness 
(Abowd et al., 1999). A system is context-aware if it can extract, interpret, and use con-
textual factors to adapt its functionality to the current context of the user (Baldauf et 
al., 2007). The task to discover context can be interpreted as the identification of quali-
tative or quantitative contextual information from users as well as their surrounding 
environment (Dey and Abowd, 2000b).

Aiming at guiding the discovery of different contextual factors, researchers develop a
two-dimension method to classify contextual factors (Adomavicius and Tuzhilin, 2011). The two dimensions are observability and variability.

The observability determines if a contextual factor is explicitly known to firms. According to this definition, there are three kinds of contextual factors.

- **Detectable factors.** The value of these factors can be directly detected and recorded by virtual and physical sensors connected to a firm’s information system. Except for certain format conversion and adaptation tasks, no further computation or data processing work is needed. Examples of such factor include the location of a consumer and the corresponding local time.

- **Computable factors.** The value of these factors cannot be directly obtained by sensors. However, using first-hand real-time data collected by the sensors as well as certain arithmetic or geometrical methods, we can compute the value for these factors. Examples of such factor include the budget of consumer and their distance to a store.

- **Predictable factors.** First-hand data captured by sensors are either insufficient or unable to give us a definitive value for these factors. The alternative is to discover the underlying rules and patterns within the historical data and use the identified knowledge to predict the current value of the factor. This task usually involves machine learning and data mining methods, which are discussed in Section 2.3. Examples of such factor include the intention of consumer and their readiness for the personalized marketing contents.

The variability depicts if a contextual factor changes over time. The value of the static factors
is constant and the value of *dynamic factors* changes with time. According to this definition, contextual factors can be classified by the following framework (Table 1.1):

**Table 1.1. Classifying Contextual Factors by Acquisition Method and Variability**

<table>
<thead>
<tr>
<th></th>
<th>Detectable</th>
<th>Computable</th>
<th>Predictable</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Static</strong></td>
<td>Constant and detectable</td>
<td>Constant and computable</td>
<td>Constant and predictable</td>
</tr>
<tr>
<td><strong>Dynamic</strong></td>
<td>Variable and detectable</td>
<td>Variable and computable</td>
<td>Variant and predictable</td>
</tr>
</tbody>
</table>

Accordingly, there are two methods to handle the variability of contextual factors (Figure 1.1): *frequency-based method* and *event-based method*. The *frequency-based method* allows firms to detect or compute contextual data periodically. The *event-based method* enables firms to initiate a detection or computation task when certain events occur. The value of contextual factors can be computed in real time, over a period of time (from \( T_m \) to \( T_n \)), or based on a cumulative manner which tracks back to the beginning of the shopping journey (back to \( T_0 \)).

*Figure 1.1. Tracking the Variability of Contextual Factors*
A large proportion of contextual information, such as time, weather, temperature, location, traffic, stock availability, and store opening hours can be collected by observation or measurement (Chen, 2005; Fang et al., 2012; Zhu et al., 2012). Some other contextual information, such as a consumer’s instantaneous speed and his (or her) distance to a specific location, can be computed using contextual information directly available (Park et al., 2007; Gallego and Huecas, 2012). Certain contextual information, which is usually contingent upon consumers’ instantaneous states (e.g. state of mind, intention, shopping goal), can neither be obtained through observation nor through computation. At this point, there are two methods to obtain such implicit contextual information.

The first method is to ask consumers to feedback the required context information to the system constantly (Palmisano et al., 2008; Lacic et al., 2015). This method is widely used by many experiments due to its ease of implementation. However, due to three reasons, firms and marketers consider this method less feasible for discovering context information in real-world systems and application (Shi et Ghedira, 2016). (i) Real-world consumers are not always willing to share contextual information with a personalization system. Unlike in the movie rating scenario where users are motivated to share their opinions, sending personal information to a personalization system is considered unnecessary and sometimes annoying in the shopping context (Shyong et al., 2006). (2) It is very hard to ensure the quality of the provided information, nor to mention its consistency across different consumers. For example, consumers may provide incorrect context information because of misunderstanding. (iii) Certain contextual information changes over time. Frequently requesting consumers to feedback ruins their shopping experience. Thus, firms and marketers have to find another method to discover such implicit context information.
The second method is to infer contextual information using contextual and non-contextual data gathered from consumer and from environment (Palmisano et al., 2008). There are two main approaches to inference: unsupervised and supervised learning. The former learns associations between context and user data (or transaction data) without the explicit intervention of consumers (Hastie et al., 2009). Whereas unsupervised learning algorithms can cluster unlabeled consumer data effectively and unobtrusively, they are not able to provide firms and marketers with the names, features and characteristics of the clustered data segments. The latter (supervised learning algorithms) requires consumer intervention at some point to label the context or define some of its characteristics manually (Kotsiantis et al., 2007). These algorithms can classify consumer data and predict their contextual labels simultaneously. However, they require intervention of consumers or experts so as to prepare the training cases, which can be challenging in some application domains where personalization systems and applications are far away from their users.

In view of the advantages and disadvantages of the above two methods, it is necessary to find an improved method that can retain the advantages of both methods and overcome their shortcomings.

1.5 Data mining and machine learning

B2C marketing personalization necessitates the context knowledge of a vast quantity of consumers. To fulfill this requirement, firms need to conduct sophisticated and numerous analysis and computation. At this time, data mining and machine learning methods can be helpful.
1.5.1 Data mining

The process to analyze large quantities of data, discover the underlying insights (e.g. patterns, association rules, anomalies), and transform them into an apprehensible structure for future use is called data mining (Fayyad et al., 1996). In most cases, a data mining task is composed of problem setting, data preparation, exploration, modeling, and evaluation phase (Han et al., 2011).

(1) **Problem setting.** The starting point of a data mining process is to formulate the objective of the analytical task, namely the result to be expected (Maimon and Rokach, 2005). Generally, there are two kinds of objectives. The first is to discover the underlying features, patterns, and characteristics of the data. Such objective can be realized by inferential analysis, which uses knowledge gathered in the sample data to draw inference on the population represented. The second is to predict the unknown information. Such objective can be realized by predictive analysis, which focuses on the relationship between dependent and independent variables. Based on the known information (i.e. current and historical data), the predictive analysis creates heuristics models which can use learned patterns and association rules to predict unknown information.

(2) **Data preparation.** In most cases, raw data are not suitable for analysis. Depending on the requirements of the analytical tool and process, the data preparation process may involve one of the following tasks (Kantardzic 2011). Trimming is a process to remove irrelevant fields from data points, which may affect the speed of the analysis. Data cleansing aims at detecting, correcting, or removing corrupt or inaccurate data points from the dataset so that their impacts on the analysis quality are minimized. Data reformatting seeks to convert raw data into recognizable and computable values. Data transformation is an effort to transform raw data by certain rules (e.g. normalization) so that
it can be accepted by certain analytical models. Data merging is a procedure to reconstitute a complete view of the data by retrieving variables scattered in different datasets.

(3) **Exploration.** Exploratory data analysis is defined as the preliminary investigation of a dataset aiming at discovering main descriptive characteristics of a dataset. It is considered an indispensable step before the modeling phase, since it depicts the feature of each variable, uncovers the relation between variables, suggests the hypothesis/assumptions associated with the problem, and supports the selection of data analysis method(s). Exploratory data analysis can be performed in a quantitative or a graphical manner (Keim, 2002). Both insights are helpful in the next stage.

(4) **Modeling.** This phase is associated with the procedures to develop a data mining model so as to solve the formulated problem (Larose, 2014). Based on the insights discovered in the exploration phase, firms and researchers can select corresponding data analysis method, introduce appropriate variables, adjust model parameters, and make necessary modifications to the model. Afterwards, the data mining model can be used to process the dataset autonomously in order to discover the underlying knowledge and insights.

(5) **Evaluation.** Different data mining models can be compared in two dimensions. The quality measures if a model is able to produce the expected results, which is the features or patterns in the inferential analysis case and the accuracy of prediction in the predictive analysis case. The efficiency measures the amount of time and computation resource required by a model to complete the designated data mining task. This dimension becomes important as the datasets processed by data mining models are usually quite huge. If a model fails to fulfill the requirements of one or two dimensions, it is
necessary to re-initiate the modeling phase, where the model can be modified for possible enhancement (Kantardzic 2011).

Data mining is a process dedicated to discover knowledge in large datasets. Consequently, the techniques to process data must be automatic and self-regulated so that they can fulfill the requirement of quality and efficiency. In the next section, we introduce the machine learning techniques that are developed to cope up with these requirements.

1.5.2 Machine learning

Due to the rapid growth of the quantity and variety of data being processed, firms find it less efficient and effective to discover knowledge in data based on rules learned by human specialists. As datasets become larger in size and more complicated in structure, the exigence for a more rapid and reliable data analyzing solution emerges. Machine learning (ML) methods are the response to this exigence. By automating the creation of data analytic models, ML methods enable computers to learn from data without human intervention or assistance. Using iterative algorithms, ML method allows data analytic models to adapt to the variation of new data.

Compared to the manual analysis methods, the ML methods have two advantages. (1) They can cope with large datasets which cannot be analyzed by human or human-driven analytic tools. (2) They can discover the underlying connections, association rules, or patterns which are too complex for a human to identify. Therefore, ML methods are becoming increasingly popular with firms and researchers. ML tasks are performed by ML algorithms. An ML algorithm is a set of iterative computational procedures which are designated to find connections, rules, patterns, or anomalies in data.
Depending on the nature of the problem to be addressed, ML algorithms can be classified into four kinds (Figure 1.2). The details of each kind of ML algorithm are presented in the Annex.

![Machine Learning Taxonomy]

**Figure 1.2. Machine Learning Taxonomy**

### 1.6 Big Data

With the rapid development of Internet technology, more and more people and objects are connected to each other. Connectivity not only allows firms to collect and analyze consumers’ contextual data that are valuable to the B2C marketing process but also bring major challenges to firms which rely on these data to develop their consumer and market strategy.
1.6.1 What is big data?

**Big data** is described as the voluminous amount of structured, semi-structured, or unstructured data that can be mined for knowledge. Compared to the traditional database management, big data management is a new information management strategy that incorporates various new types of data and data management techniques (Beyer and Laney, 2012). The objective of big data management is to exploit a large amount of data in a cost-effective and innovative manner so as to enable enhanced insight, decision-making, and process automation (Gartner’s IT dictionary, 2015). As more and more firms begin to base their marketing strategy on big data, the corresponding challenges, featured by “7V”, must be addressed.

**Volume.** With the prevalence of web analysis technologies, mobile phones, beacons, sensors, scanners, and other devices deployed in the digital and the real world, firms are now able to capture every click, text message, search query, transaction, movement, location, and more about consumers if they want. Considering the number of consumers involved, the volume of data can be massive (Wu et al., 2014).

**Variety.** Due to distinct data collection devices and sources, data may come in all structures, types, shapes, and formats (Sagiroglu and Sinanc, 2013). Such diversity can be observed in several dimensions. (1) Unstructured data (e.g. texts, images, video clips, soundtracks) are playing an increasingly important role in the knowledge discovery process because they contain more valuable information that can be extracted for analysis. (2) Distinct configurations of data gathering devices contribute to a large assortment of data formats and types, which need to be unified for the subsequent data mining job. In practice, this task is often time-consuming. (3) Merging or organizing data collected by various virtual and physical devices may provide us with a complete view
of the research subject.

**Velocity.** Nowadays, data are generated at an unprecedentedly high speed, which in return requires data processing tasks to be performed more swiftly and efficiently (Kaisler et al., 2013). In practice, many applications (e.g. location based applications, marketing personalization systems) rely on real-time data processing capability to function. In other situations, rapid data analysis and decision-making capability can be the key factor that secures a firm's leading position in the market competition. Given the infrastructure of data management system, the enhancement of data processing speed concerns the optimization of the architecture and the algorithms.

**Variability.** Sometimes, the intrinsic meaning of data is associated with the context (Katal, 2013). Taking semantic analysis as an example, language processing technologies allow firms to gather data from texts, which can reveal valuable information that cannot be obtained in numeric data. However, decoding the exact meaning of textual information can be challenging, as the meaning of a word or phrase is determined by other words, phrases, or passages that surround it. Accordingly, semantic analysis necessitates a different set of skills and techniques than the numeric data analysis.

**Veracity.** In the field of computer science and information technology, the famous “garbage in, garbage out” principle refers to the fact that inaccurate or nonsensical data inputs may lead to erroneous or absurd outputs. The same principle applies to the mining of big data. Before the mining operation, dataset must be properly “cleansed” for two reasons (Demchenko et al., 2013). (i) Data captured by different systems need to be merged and matched so that their value could be exploited. (ii) Errors and anomalies produced during the data gathering and storage process need to be identified and properly dealt with so that their impacts on the analysis result are minimized. Due to
the messy and chaotic nature of certain datasets, such tasks can be quite challenging.

**Visualization.** Big data analysis can yield complicated results that are recorded in numbers and spreadsheets. Whereas the classic analysis reports are content-rich, they are less user-friendly. It is important that the discoveries, insights, and knowledge derived from data are presented in a comprehensible form like charts and infographics so as to make their meanings clear to users (Keim et al., 2013). Lucid visualization of data is always challenging, as it requires a profound knowledge of data as well as the creativity.

**Value.** By making all the efforts to cope with the challenges of volume, velocity, variety, variability, veracity, and visualization, the ultimate objective is to uncover the hidden value in big data. One of the key characteristics that distinguish big data analysis from classic data analysis is that the value of big data lies not only in their literal forms but also in relations that connect data from various sources to each other (LaValle et al., 2011). Certain types of information may not have any explicit value if they are separately analyzed. When properly related, however, they may provide valuable insights to users. One must be able to see through the messy and noisy appearance of big data in order to fully uncover its value.

**1.6.1 What is its impact on B2C marketing strategy?**

In the B2C marketing domain, the key objective of a firm is to better promote their products to customers. Today, more and more firms are aware that big data is set to bring opportunities and challenges to them in their course to achieve the objective.

It is apparent that big data extends and enriches the way firms develop and implement their B2C marketing strategies.

(1) Big data is rich in format and source, which allows firms to have a more profound
knowledge of their customers and potential customers from new angles. As a result, it may reshape the way firms develop their customer strategy. For example, B2C firms rely heavily on surveys to know their customers in the past. Nowadays, such insights can be obtained directly from social networks and many other information platforms.

(2) Big data provides a vast volume of customer profiles and cases, which makes it possible for firms to discover the commonalities of consumer interests, preferences, and behavior patterns. When firms need to develop and elaborate their B2C marketing strategies, these insights can be the foundation that helps firms clarify the needs and intentions of their customers.

Meanwhile, big data is putting forward challenges to B2C firms which are not prepared to embrace the new changes.

(1) The advent of the big data era means that firms need to make real-time decision and take quick actions based on data, which is the expectation from their customers and potential customers. The firms which are the first to make right decisions based on data can win the competition. As a result, a company may only have a few hours (rather than a few weeks) before they must present a new marketing strategy.

(2) The exponential growth of the quantity and complexity of consumer data reduces firms' visibility of the market. If firms do not have the competencies and tools to extract useful marketing insights from a vast amount of unstructured and invalidated raw data, they not only lose the competition to other capable firms in terms of timeliness but also run the risk of making wrong business decisions and strategies.
1.7 Summary

In this chapter, we presented the knowledge of shopping journey, personalization, context, data mining, machine learning, and big data. It lays the foundation for our subsequent discussions on the context-aware marketing personalization approach (CAMPA), which involves the application of the above-mentioned knowledge. In the next chapter, we review the researches relevant to the development and application of the CAMPA based on a corpus of research papers focusing on the intersection of personalization and contextualization in the cross-channel shopping context.
Chapter 2

Panacea or paradox? The challenges of existing B2C marketing personalization methodologies.

2.1 Introduction

Today, more and more firms and researchers believe that the context-aware marketing personalization approach (CAMP A) is the future of B2C marketing personalization (Abar et al., 2009; Wang and Wu, 2011; Baltrunas et al., 2012; Zheng et al., 2014; Lacic et al., 2015; Lamche et al., 2015). As a result, research of this cross-disciplinary subject is getting increasing popularity among researchers and firms. A context-aware marketing personalization approach aims at gathering contextual information surrounding a user (aka a consumer in the shopping context), decoding the underlying impact of context on his (or her) decision model and behavioral pattern, and suggesting products or services relevant to his (or her) interest based on the acquired knowledge (Adomavicius and Tuzhilin, 2012).

In order to obtain a holistic view of the research on the context-aware marketing personalization for cross-channel shoppers in the B2C marketing domain, we made a comprehensive review of the related research papers. By studying the origin and the recent
developments of context-aware personalization approaches, we seek to base our work on the state-of-art development in this domain. The remainder of this chapter is organized as follows.

We start with a general review of related research papers to highlight the focus and hot topics of the research on context-aware marketing personalization for consumers (Section 2.2). Afterward, we make a review of the current model to explain a cross-channel shopping journey, present the existing knowledge on intentions of cross-channel shoppers, and highlight their limitations (Section 2.3). Next, we discuss current customer tracking technics and the main challenges to track the cross-channel shopping behavior (Section 2.4). Then, we discuss the current process-driven personalization approach and its drawbacks (Section 2.5). After that, we review the existing contextualization methods and their limitations (Section 2.6). Finally, we present the current performance tracking and improvement methods for B2C marketing personalization systems and highlight their limitations (Section 2.7). Section 2.8 is the summary of this chapter.

2.2 Contextualized B2C marketing personalization: a research overview

The last decade witnesses the prevalence of context-aware approaches, which are able to perceive the status or change of surrounding context, gather contextual information, discover the underlying implication, and make necessary adaptations (Dey and Abowd, 2000b). The ability to adopt distinct strategies according to the change of context makes context-aware systems suitable for the personalization tasks, which aim at providing tailor-made suggestions to consumers based on their current situation (Bal-
Hence, more and more consumer-oriented industries, such as retailing, tourism, financial service, health-care, media, entertainment, learning, and social media, started to incorporate contextualization approaches to their marketing personalization systems.

With an objective to clarify the development roadmap of the context-aware personalization approaches, we start by resorting to existing literature review papers for clues. We identify a few important papers (Adomavicius et al., 2005; Baldauf et al., 2007; Hong et al., 2009; Adomavicius and Tuzhilin, 2011), which lay the foundation for the development of context-aware marketing personalization approaches. Based on these insights, we make a literature research focusing on the recent advancement of the context-aware personalization approaches. Our quest starts with a few recent papers published on renowned journals and conferences. They are frequently quoted by researchers in this domain. Then, we expand our search scope with regards to several dimensions, which include: (1) the objective and application field of a paper; (2) its impact on the academic community; (3) the contextualization techniques used; (4) the field of application. Accordingly, we present our corpus in Table 2.1, which highlights the research domain, the type of contextual information, and the contextualization method used in each research paper of our corpus. Whereas the corpus cannot cover all the papers related to the context-aware personalization research, the selection criteria ensure that they represent the major and recent advancement of the research works in this domain.
<table>
<thead>
<tr>
<th>Author</th>
<th>Domain</th>
<th>Year</th>
<th>Personalization Method</th>
<th>Context Features</th>
<th>Contextual information</th>
<th>Contextualization Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen</td>
<td>General</td>
<td>2005</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adomavicius et al.</td>
<td>General</td>
<td>2005</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Yu et al.</td>
<td>Media</td>
<td>2006</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Gorgoglione</td>
<td>Shopping</td>
<td>2006</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Wowrell et al.</td>
<td>Point of Interest</td>
<td>2007</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Berkovskiy et al.</td>
<td>Media</td>
<td>2007</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Ono et al.</td>
<td>Media</td>
<td>2007</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Park et al.</td>
<td>Point of Interest</td>
<td>2007</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Yang et al.</td>
<td>Shopping</td>
<td>2008</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Tao-Sutter et al.</td>
<td>Media</td>
<td>2008</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Xu et al.</td>
<td>General</td>
<td>2008</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Boutemedjet and Ziou</td>
<td>Media</td>
<td>2008</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Abhar et al.</td>
<td>Shopping</td>
<td>2009</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Hong et al.</td>
<td>Shopping</td>
<td>2009</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Gallego and Huecas</td>
<td>Shopping</td>
<td>2012</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Baltrunas and Anastajic</td>
<td>Media</td>
<td>2009</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Baltrunas and Rici</td>
<td>General</td>
<td>2009</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Janus et al.</td>
<td>Interest</td>
<td>2002</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Pan et al.</td>
<td>Media</td>
<td>2002</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Casello</td>
<td>Media</td>
<td>2002</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>General</td>
<td>Media</td>
<td>2002</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Ono et al.</td>
<td>Media</td>
<td>2002</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Park et al.</td>
<td>Point of Interest</td>
<td>2007</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Yang et al.</td>
<td>Shopping</td>
<td>2008</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Tao-Sutter et al.</td>
<td>Media</td>
<td>2008</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Xu et al.</td>
<td>General</td>
<td>2008</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Boutemedjet and Ziou</td>
<td>Media</td>
<td>2008</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Abhar et al.</td>
<td>Shopping</td>
<td>2009</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Hong et al.</td>
<td>Shopping</td>
<td>2009</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Gallego and Huecas</td>
<td>Shopping</td>
<td>2012</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Baltrunas and Rici</td>
<td>General</td>
<td>2012</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Author</td>
<td>Domain</td>
<td>Year</td>
<td>Personalization Method</td>
<td>Context Features</td>
<td>Contextual information</td>
<td>Contextualization Method</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------</td>
<td>------</td>
<td>------------------------</td>
<td>------------------</td>
<td>------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Hidasi and Tikk</td>
<td>General</td>
<td>2012</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Zhu et al.</td>
<td>General</td>
<td>2012</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Baltrunas et al.</td>
<td>Point of Interest</td>
<td>2012</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Wang et al.</td>
<td>Media</td>
<td>2012</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Fang et al.</td>
<td>Shopping</td>
<td>2012</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Liu and García-Valdez</td>
<td>Point of Interest</td>
<td>2014</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Zheng et al.</td>
<td>Media</td>
<td>2014</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Neves et al.</td>
<td>Interest</td>
<td>2014</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Lommatzsch</td>
<td>Interest</td>
<td>2014</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Kim et al.</td>
<td>Shopping</td>
<td>2014</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Zheng et al.</td>
<td>General</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Zhang et al.</td>
<td>General</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Shi et al.</td>
<td>Shopping</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Panniello et al.</td>
<td>Shopping</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Chen et al.</td>
<td>Shopping</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Lamche et al.</td>
<td>Shopping</td>
<td>2015</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
The identified papers cover a wide range of research domains related to the context-aware marketing personalization domain. The five major research domains are: consumer interest discovery, point of interest recommendations, shopping suggestions, consumer media recommendations, and generic context-aware marketing personalization approaches (Figure 2.1).

![Figure 2.1. Classification of Research Papers by Application Domain](image)

The research of generic context-aware marketing personalization methods (i.e. methods to discover “user-user”, “item-item”, or “user-time” relations) remains active during the past ten years. Various consumer profile similarity comparison techniques such as vector similarity, singular value decomposition, and tensor factorization were proposed and applied to the research field (Chen, 2005; Adomavicius et al., 2005; Baltrunas and Ricci, 2009; Karatzoglou et al., 2010; Hidasi and Tikk, 2012; Zhang et al., 2015).

Meanwhile, researchers’ interest for specific application fields has shifted during the
past ten years. For instance, the research on consumer media personalization, which seeks to proposing images, music, movies, and video clips related to consumers’ context and interest has significantly reduced since there is no major breakthrough in the discovery of consumers’ contextual factors on the social network platforms (Liu and Aberer, 2013). On the contrary, the research on consumer interest discovery, point of interest (POI) recommendations, and shopping personalization is gaining increasing attention, because the emergent technologies (e.g. location-based services) and infrastructures (e.g. beacons) allow researchers to gather and analyze new contextual factors, which extend the research scope of context-aware personalization approaches (Yang et al., 2008; Abbar et al., 2009; Woerndl et al., 2011; Yin et al., 2013; Ramirez-Garcia and Garcia-Valdez, 2014).

Although the research works cover a wide range of topics on the B2C marketing personalization domain, they have their limitations. We mentioned in the Introduction Chapter that a mono-channel perspective could have an adverse impact on the relevance and usefulness of marketing personalization. However, we are not able to find any research paper that takes into account the cross-channel shopping journey or the cross-channel consumer intentions, which are apparently among the most emergent and important research subject in this domain. In fact, the cross-channel shopping behavior itself is an important contextual factor, which indicates that consumers are not satisfied with (some of) the shopping experience in the current shopping channel. Hence, before exploring the limitations of current personalization and contextualization methods, we would like to discuss about the cross-channel shopping journey, the intentions of cross-channel shoppers, and the way to track cross-channel shopping behaviors in the next section.
2.3 Cross-channel shopping and consumer intentions

As the cross-channel shopping behavior becomes popular with consumers, researchers make efforts to analyze and explain this new phenomenon. They suggest that consumer’s choice of a retail channel to complete purchase transactions is significantly affected by costs and benefits incurred to satisfy shopping goals (Kim et al., 2002). Channel types differ in their abilities in performing various retail service outputs and the benefits and costs they impose on consumers (Bucklin et al., 1996; Verhoef et al., 2015). Consumers adopt cross-channel behavior to take advantage of the specific characteristics and advantages of each channel (Verhoef et al., 2007; Shi, 2016) and satisfy their shopping needs (Konus et al., 2008). The research on the goals that consumers seek to satisfy during the transaction stage of the purchase process suggests that customers’ desire for convenience and their quest for self-affirmation related to decision expertise and thrift can drive their selection of channels when pursuing purchase transactions (Balasubramaniam et al., 2005).

Based on the analysis of the cross-channel shopping motivations, researchers try to establish theoretical frameworks to explain the cross-channel shopping behavior. Some claim that consumers select channels at each stage of the decision process to fulfill utilitarian and hedonic needs at the lowest costs relative to benefits (Noble et al., 2005). Others try to explain the cross-channel shopping behavior from a behavior perspective. They claim that cross-channel shopping behaviors can be explained by three types of consumer intentions (Rose and Levinson, 2004; Jansen et al., 2008). (1) The informational intention is defined as a need to acquire useful data or knowledge. This intention is featured by the efforts to look for detailed and specific information. (2) The navigational intention is defined as the need to search for a hub that can guild users to another place.
This intention is featured by users’ cursory behaviors. (3) The *transactional intention* is defined as a need to complete some specific tasks. This intention is featured by the stepwise user behaviors that are specified by a transaction process.

The way cross-channel shoppers think and behave is quite different from mono-channel shoppers. It is not just the number of channels they use that makes the difference. Mono-channel shoppers perform various kinds of shopping related tasks all in the same channel, while cross-channel shoppers hold that different shopping related tasks, such as searching, browsing, comparing, experiencing, and purchasing are better to be performed in different channels (Shi, 2016).

As cross-channel shopping behaviors become popular, firms and marketers find that more and more “free-riders” who examine products in one channel and purchase them elsewhere (Verhoef et al., 2015). Accordingly, it is increasingly difficult for firms and marketers to tell if consumers use a particular shopping channel to search for shopping ideas (Baeza-Yates et al., 2006), enrich product knowledge (Rose and Levinson, 2004; Verhoef et al., 2007), compare various products (Moe, 2003; Close et al., 2010), obtain using experience (Verhoef et al., 2015), make a purchase, or just for hedonic purposes (Jones et al., 2006).

In short, more and more consumers no longer consider buying a product as their only purpose to visit a shopping channel. If personalization systems still stick to the dogma of "finding items that fit consumers’ current interest," they are not able to accommodate the increasingly diverse and complex needs of consumers.

Although the above research works significantly extend our understanding of the cross-channel shopping behavior and its underlying intentions, they have some limitations that requires improvement.
(1) The diversity of cross-channel shoppers’ intentions brings challenges to current contextualized marketing personalization systems, which base their predictions on consumers’ (past or current) preferences rather than their underlying intentions. As a result, these methods cannot cope with the following complicated scenarios. For example, these systems cannot take initiative to differentiate gift purchases from personal purchases so that sometimes they make some absurd predictions by incorrectly applying consumers’ preference in one context to another context (Gorgoglione et al., 2006; Palmisano et al., 2008). Similarly, they do not have tools that can analyze the degree of clarity of a consumer’s need (i.e. vague or clear) and they tend to perform item-level predictions without knowing whether a consumer needs a particular item or a more general idea (Shi and Ghedira, 2016). As a result, a great proportion of personalization contents were rejected or ignored by consumers. In addition, current contextualized marketing personalization systems are not capable of predicting whether a consumer is looking for similar items, accessory items, or novel items. Aiming at maximizing conversions, these systems methods a mixture of different kinds of predictions to consumers, leading to excessive information load and waste of computational resources.

(2) Cross-channel shopping behaviors also pose a threat to the prediction accuracy of current marketing personalization models, which collect and analyze consumer and item data using a platform-specific method (Fang et al., 2012; Yin et al., 2013; Hussein et al., 2013; Chen et al., 2015; Panniello et al., 2015). Since behavioral and transactional data are scattered in different platforms, analyzing consumer data gathered in one platform may provide incomplete and sometimes biased consumer knowledge to researchers and marketers, leading to confusing personalization results. A solution to this threat is to consider consumers’ cross-channel activities as integrated steps of a shopping journey, combine consumer data from different platforms, and analyze them as a whole.
This is an emerging subject for consumer behavior research and the method to consolidate and analyze cross-channel consumer data is yet to be developed and improved.

(3) These research works cannot provide a theoretical framework based on which we can unify the cross-channel behaviors happened in three different situations, namely the planned purchase, impulsive purchase and repeat purchase.

- **A planned purchase** is a series of actions consumers take to fulfill a shopping need that is proposed by themselves. Consumers know what they want to buy (roughly) and they make an informed decision based on a series of deliberate and thoughtful evaluations, comparisons, and reflections (Cobb and Hoyer, 1986).

- **An impulsive purchase** refers to the purchase of an unexpected or unplanned product that triggers consumers’ feelings or emotions (Cobb and Hoyer, 1986). Impulsive buyers do not follow a regular decision-making model that requires them to evaluate and compare products (Piron, 1991). Instead, they tend to make a quick decision and purchase the item immediately to fulfill their instant self-gratification needs (McGoldrick, 1982).

- **A repeat purchase** refers to the purchase of the same items that have been purchased by a consumer in the past (Hoyer, 1984). Since the consumer is familiar with the items, they rarely spend additional time in evaluation and comparison again (Chiu et al., 2014). As a result, the shopping journey is quick and concise.

These different situations have an impact on consumers’ channel choice (Li et al., 1999) and their reactions to personalized marketing contents (De Kervenoael et al., 2009). However, the current theoretical framework based on financial and behavioral per-
spective cannot unify consumers’ motivations and behavior patterns in different shopping situations. The lack of a systematical theoretical framework means that it is hard to incorporate differences in shopping situations to the contextualized marketing personalization approach.

(4) Current explanations to consumers’ cross-channel shopping behaviors are not systematic and specific. As a result, it is not possible to infer consumers’ information needs on each stage of a shopping journey (Kumar and Vankatesan, 2005). It also means that firms cannot evaluate consumers’ cognitive state as well as the urgency of their information needs. This means that firms are not able to adjust their marketing personalization strategy for consumers who are at different stages of a shopping journey. It is also challenging to adapt personalized marketing messages to consumers with different cross-channel shopping intentions.

In view of these limitations, we propose in Chapter 3 a cognition-behavior model to close the gaps.

2.4 Cross-channel consumer tracking

Another reason that explains the lack of marketing personalization research for the cross-channel shopping context is the difficulty in tracking and analyzing consumer behaviors in a cross-device and cross-platform context (Wijaya et al., 2014).

The methods to track online and offline behaviors have been developed for years. In an online context, firms can deploy cookies on their websites which enable them to keep track of the product viewed by consumers and the quantity of time consumers spend with these products (Mayer and Mitchell, 2012). These cookies also allow firms to capture consumers’ preferences in terms of website design, language, layout, and
content so that firms can display personalized contents to consumers next time they visit the website (Englehardt et al., 2015). Meanwhile, the methods to track and interpreted consumer behaviors are developing rapidly. Using various technologies such as Bluetooth, WiFi, Radio Frequency Identification (RFID), ultra-wide band, and near field communication (NFC), firms can detect consumers’ position in real time and transfer such knowledge into consumer preference by analyzing which product do consumers pay attention to and how do they interact with these products (Liu et al., 2007).

However, the rapid development of online and indoor consumer tracking methods and technologies do not bring to the birth of the cross-channel consumer tracking methods. (1) So far there is no research that proposes a unified method to encode the online and the instore shopping behavior, making it hard for researchers to analyze the motivations and underlying consumer intentions of the cross channel behaviors. (2) The current methods are not able to consolidate and synchronize the online and the instore shopping behaviors of the same consumer, making it hard to follow consumers change of preference and needs. (3) The current methods study the cross-shopping behaviors from a macro perspective. That is, to analyze these behaviors by each session (the time between consumers’ arrival and departure of an online or physical store). However, a shopping journey can be much longer or shorter than a session. The session-based approach may be arbitrary and may provide firms with a biased view of consumers’ intentions and needs.

It is also necessary to mention that the difficulty in cross-channel consumer identification is also caused by legal constraints. In fact, existing biometrics technologies can track consumers across platforms effectively and efficiently (Unar et al., 2014). However, applying biometric technologies in the retailing context is a very controversial
topic as most regulatory authorities and consumers around the world have great concerns about allowing retailing firms to gather, store, and analyze biometric data, which is considered a threat to consumer privacy (Jain and Nandakumar, 2012). As a result, firms and marketers have to choose some less effective or less reliable methods compliant with the laws and regulations.

In view of these limitations, we propose a new method to encode, consolidate, and fragment consumer cross-channel behavior streams in Chapter 4.

2.5 Personalization Method

Even though the challenges of cross-channel consumer tracking can be overcome, the current marketing personalization method is far from being satisfactory. Personalization method is the fundamental of a marketing personalization system. It regulates the behavior of a marketing personalization system in three aspects. (1) It defines the role a marketing personalization system plays in consumer’s shopping journey. (2) It depicts the personalization process of a particular marketing personalization system. (3) It determines the type of consumer data to be used and the way to exploit them in order to obtain personalization results.

A considerable amount of existing personalization systems, including the ones of Amazon, Netflix, and COMPANY X, follow the process-driven personalization method (Figure 2.2, next page), which regards consumers’ behaviors and preferences as input, uses a predefined algorithm or algorithm group to process the data, and provides output.
Figure 2.2. The Process-Driven Personalization Method

According to this method, the main role of marketing personalization systems is to accelerate consumer conversion (Adomavicius et al., 2005; Woerndl et al., 2011). When consumers visit a website, marketing personalization systems use their past or current choices (e.g. purchase, positive rate) as input, identify similar or related items using personalization algorithms (e.g. content-based algorithm, collaborative filtering algorithms), and propose the outputs to them (Figure 2.2a). When consumers leave a website or a physical store without purchasing anything, marketing personalization systems regularly remind them of the updates about the products they viewed in order to bring them back to the website or store and reinitiate a purchase process (Figure 2.2b). In both scenarios, consumers’ transaction data (including but not limited to product searched, viewed, liked, rated, or purchased by consumers) are in the center of a marketing personalization process (Gallego and Huecas, 2012; Campos et al., 2013; Kim et al., 2014; Chen et al., 2015).
In spite of its popularity with famous retailing firms such as Amazon (US), Taobao (China), Fnac (France), Rakuten (Japan), COMPANY X and so on, the linear logic is far from being a perfect personalization method. By studying the personalized contents of the above five major retailers, we identified two major drawbacks of the process-driven personalization method.

- The first drawback is that the process-driven personalization method assumes that consumers need personalization information at all times. Therefore, personalization systems keep sending recommendations, event updates, coupons, and promotion announcements to consumers who expressed their interest explicitly (e.g. by making subscription) or implicitly (e.g. by purchase or consume some specific items). In reality, many expressed preferences (e.g. gift offering purchase, random exploration activities) are incidental and a considerable number of consumers prefer not to be disturbed during the shopping journey. Subsequently, the majority of personalized contents are ignored by consumers.

- The second drawback is that a process-driven method pre-defines the type(s) of information that a consumer looks for. To support our argument, we prepared a list of 50 products of different types, which were then submitted one by one to the above mentioned retailing firms’ website as search keywords. We analyzed the personalized contents provided by these systems and classified them into four clusters based on their relation to the products we submitted as search keywords.
Table 2.2. Composition of Personalized Contents by Firm

<table>
<thead>
<tr>
<th>Firm</th>
<th>similar</th>
<th>related</th>
<th>popular</th>
<th>novel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon</td>
<td>50-70%</td>
<td>15%-35%</td>
<td>10%</td>
<td>5%</td>
</tr>
<tr>
<td>Taobao</td>
<td>50%-60%</td>
<td>20%-30%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Fnac</td>
<td>40%-70%</td>
<td>25%-50%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Rakuten</td>
<td>70%-80%</td>
<td>10%-20%</td>
<td>10%</td>
<td>0%</td>
</tr>
<tr>
<td>COMPANY X</td>
<td>40%-60%</td>
<td>10%-30%</td>
<td>5%</td>
<td>5%</td>
</tr>
</tbody>
</table>

The table 2.2 highlights the composition of personalization results of Amazon, Taobao, Fnac, Rakuten, and COMPANY X. For instance, if a consumer views a scooter in the Amazon website, 50%~70% of the personalized marketing contents are about similar scooters, 15%~35% of them are about accessories (e.g. helmet, knee pads), nearly 10% of them are about sports related items loved by the consumer’s neighbors (i.e. person who have similar preference), and the remaining 5% are about sports related items that are just released. Does this composition reflect consumers’ information need? The answer is “no”.

The type of information consumers may look for depends on their intention. It is rare that consumers are interested in two or more kinds of information at the same time. Therefore, proposing several kinds of personalized contents simultaneously is both ineffective and counterproductive.

To summarize, the process-driven personalization method is problematic because it makes two assumptions that are not compliant with consumers' shopping habit and decision-making process, leading to a high rejection rate of personalized marketing contents. In view of the drawbacks, we present an intention-based personalization methodology based on the analysis of their shopping journey and decision model. The new personalization methodology is used as the methodological framework to build the contextualized marketing personalization system for COMPANY X.
2.6 Contextualization Approach

In this section, we make a review of the current approach to applying contextual factors to the marketing personalization process. The premise to apply contextual information to personalization systems is that the same consumer can adopt different decision strategies and prefer different products or brands depending on the context (Lussier and Olshavsky, 1979; Rooderkerk et al., 2011). The idea to present different products in different ways and at different times is referred to as “contextual marketing” (Luo, 2003), which argues that right marketing contents should be diffused to the right consumer through the right medium and at the right moment (Fischer, 2012). Hence, accurate predictions of consumer preferences should not only be contingent upon the recommendation algorithm with which we predict consumers’ interest but also upon the degree to which we incorporate relevant contextual factors. To this aim, personalized systems and applications need to solve two problems. (i) How to obtain contextual information? (ii) How to exploit contextual information?

From a structural point of view (Figure 2.3), a personalization process consists of three interdependent phases (Vesanen, 2007).
In the input phase, a personalization system or application collects, compiles, and organizes information that is useful for the personalization process (Eirinaki and Vazirgiannis, 2003; Montgomery and Smith, 2009), such as consumer information (e.g. user id, gender, age, feature labels), transaction information (e.g. item id, purchase time, unit price, quantity, payment method, etc.), and contextual information (e.g. location, landing platform/interface, time, state of mind, budget, shopping goals, etc.). Some information is retrieved from the database, and the other is gathered in real time.

In the processing stage, the above information is processed by personalization algorithm(s) to predict consumers’ subsequent behaviors or preferred items (Sarwar et al., 2002; Pazzani and Billsus, 2007; Su and Khoshgoftaar, 2009; Lops et al.,
Depending on the goal of a personalization task, information can be processed in a variety of ways. For instance, statistical algorithms can identify items that are most popular with consumers; collaborative filtering algorithms can identify consumers who are similar to the current consumer in terms of the behaviors or preferences; content-based algorithms can locate items (or consumers) whose features and characteristics are similar to the current item; and association rule-based algorithms can predict consumers’ subsequent behaviors. In practice, personalized systems or applications often use several personalization algorithms in a nested or superimposed manner in order to make the prediction results more accurate and diversified. These systems are called hybrid personalization systems (Burke, 2002; Adomavicius and Tuzhilin, 2005; Burke, 2007).

In the output phase, a personalized system or application presents its predictions to a user interface (for recommended items) or a marketing system (for predicted user behaviors) to complete the personalization process (Adomavicius et al., 2005). Personalization systems or applications with learning capabilities also record and analyze consumers’ feedbacks on the predictions in order to improve the accuracy of their prediction models (Manouselis et al., 2011; Rubens et al., 2011).

Based on these methods, Adomavicius and Tuzhilin makes a framework of contextualization consisting of three approaches (Adomavicius and Tuzhilin, 2011).

(i) The first method is “contextual pre-filtering”, which uses contextual information to choose the most relevant data for the following personalization process. This method allows personalization systems to analyze a sub-dataset rather than the whole dataset. Hence, the speed of the personalization process and the quality of the personalization results can be enhanced.
(2) The second method is “contextual post-filtering”, which uses contextual information to adjust the personalized results. Since the contextualization task is subsequent to the personalization process, firms do not need to modify their established personalization algorithms to integrate the contextual information.

(3) The third method is “contextual modeling”, which incorporates contextual information to a personalization algorithm as an explicit predictor of a user’s preference. Instead of processing two-dimension datasets (user-item), personalization models based on this method deal with multi-dimensional datasets (user-item-context).

According to their framework, contextual information can be exploited by three methods, namely contextual pre-filtering, contextual modeling, and contextual post-filtering (Adomavicius and Tuzhilin, 2011).

2.6.1 Contextual pre-filtering

Contextual pre-filtering aims at choosing the relevant data for a personalization process using contextual information $K^q$ as a filter, allowing personalization systems to retrieve and analyze a sub-dataset $S_{k=k}^q = \{s_1, s_2, ..., s_m\}$ from the whole dataset $S = \{s_1, s_2, ..., s_n\}$ (Adomavicius et al., 2005; Xu et al., 2008). In a recent experiment (Lamche et al., 2015), researchers use several types of contextual information to pre-filter input information for a personalization system, which provides information of relevant nearby shops to consumers. The first contextual filter is a shop’s distance to an active consumer. Shops that are not within the specified distance were removed from the shop dataset. The second contextual filter is a shop’s opening hours. Shops that cannot open during the time specified by an active user are removed. The third filter removes shops that cannot fulfill consumers’ crowdedness criterion. And the last filter disqualifies shops which
do not have the item in stock. Since fewer cases are analyzed by a personalization model, it took less time to complete a personalization process.

Whereas the pre-filtering method is easy to implement and efficient for most of the time, it has some drawbacks and limitations. (1) The pre-filtering method does not work with all types of contextual information. More precisely, the method is useful only when the item data or user data to be processed by a personalization system contain the contextual information used by the filter. This limits its potential to become a general method since some critical contextual information, such as consumers’ state of mind and intention, have no direct relation to the items. (2) The pre-filtering method may over-prune the input dataset so that a personalization model cannot have sufficient cases to analyze, leading to a negative impact on the prediction result. Realizing this possibility, Lamche relaxes the above four contextual conditions in the experiment when the number of remaining cases was less than 300 (Lamche et al., 2015). As a result, the minimum requirement for the number of cases was met at the cost of introducing irrelevant cases, which is exactly the situation contextual pre-filtering method tries to avoid. In view of these drawbacks and limitations, researchers turn to other contextualization methods.

2.6.2 Contextual modeling

*Contextual modeling* is the most commonly used alternative. It regards contextual information as the third dimension (the first is consumer information and the second is item information) and incorporates contextual information directly to the personalization model as an explicit predictor of a user's preference (Adomavicius et al., 2005). Instead of analyzing two-dimension (user-item) datasets, contextualized personalization models analyze multi-dimensional datasets that contain information about consumer, item,
and context (Chen, 2005; Abbar et al., 2009; Zhang et al., 2015).

For instance, in a renowned online shopping research on context-based shopping recommendations, researchers assumed that consumers shopping behaviors were contingent upon their intent of purchase (Palmisano et al., 2008). As a result, they incorporate different intents of purchase (i.e. personal purchase for work; personal purchase for other; gift purchase for friend/partner; gift purchase for parent) to the predictive models they build.

Figure 2.4. Context Prediction Models (a & b) and Contextualized Predictive Models for Consumer Choices and Behaviors (c & d)

- The first model (Figure 2.4a) learns to predict contextual features $K_q$ based on a consumer’s behaviors $X_p$ and $f$ is a predictive function learning via machine learning methods.

- The second model (Figure 2.4b) is a Bayesian Net where a finer contextual feature $K_q$ is dependent on a consumer’s behaviors/features/choices $X_p$ as well as the identified/predicted coarser contextual feature $K_i$ (e.g. a coarser context can be a gift purchase and a finer context can be a gift purchase for parent).

- The third model (Figure 2.4c) learns to predict a consumer’s subsequent behavior
or choice $Y$ taking into account his (or her) behaviors/features/choices $X_p$ as well as the predicted intent of purchase $K^q$. The variables are considered independent.

- The last model (Figure 2.4d) learns to predict a consumer's subsequent behavior or choice $Y$ using the same variables. However, the variable $K^q$ is dependent on variable $X_1, X_2, \ldots, X_p$. They found that the contextualized models (3.3c and 3.3d) performed much better than a non-contextualized model $Y=f(X_1,X_2,\ldots,X_p)$ in predicting consumers' subsequent behaviors.

When researchers use contextualized models to predict consumers' behaviors, as is the case in the above research, non-contextual features $X_1, X_2, \ldots, X_p$ are pre-selected (Palmisano et al., 2008). The same situation can be found in the prediction of consumers’ choices, where personalization algorithms are pre-selected (Adomavicius and Tuzhilin, 2011). At this time, some interesting questions emerge. *Does the selection of different contextual information affect the prediction accuracy? Could a model’s prediction accuracy drop due to the use of non-contextual variables or personalization algorithms that are not suitable to the current context? How to choose the most relevant contextual information, non-contextual variables, and personalization algorithms so as to enhance the prediction accuracy?*

*Some research papers may give us a preliminary answer to the first two questions.* Campos and colleagues compared the performance of different personalization models using different contextual information (Campos et al., 2013). They found that using all available contextual information does not necessarily yield to the best prediction performance, which suggests that using different contextual information in a prediction model can affect the prediction accuracy. However, they did not explain which contextual information is more effective in different situations. Parra and colleagues created a personalization system which allows users to obtain information relevant to their
needs by letting them determine their own recommendation strategy (Parra et al., 2014). The results show that user selected strategies outperformed the pre-defined strategy, which provides an indirect evidence that pre-defined personalization strategy cannot fulfill the needs of all users and a smart personalization strategy should change with the context.

Whereas the conclusion is evident, no research paper has systematically provided an answer to the third question, namely how to choose the prediction model variables (contextual and non-contextual) as well as the algorithm according to the identified context. Therefore, more efforts should be made on the research to discover and exploit of the relationship between elements of a personalization model and contexts so that both the performance and the user interface of a personalization system can be improved thanks to faster, fewer, and more relevant personalization contents.

### 2.6.3 Contextual post-filtering

According to Adomavicius and Tuzhilin’s framework, there is another contextualization method, which is called contextual post-filtering (Adomavicius and Tuzhilin, 2011). This method does not intervene with the personalization process, but it adjusts preliminary personalization results \( R = \{r_1, r_2, ..., r_m\} \) to ensure that the filtered predictions \( R'_{K=K^q} = \{r_1, r_2, ..., r_n\} \) fit with the identified context \( K^q \). To ensure the normal operation of the post-filtering method, preliminary personalization results must contain the contextual information used by the filter (Panniello et al., 2009). As a result, the post-filtering method is suitable for rejecting unfitting cases which can be identified using such explicit contextual information such as item price, stock availability, shop distance, time, crowdedness, weather, and consumers’ expressed moods (Woerndl et al., 2011; Braunofer et al., 2013; Ramirez-Garcia and García-Valdez, 2014; Panniello et al., 2015). It
cannot enhance a personalization model’s prediction accuracy or relevancy if firms want to use some implicit contextual information, which is not available in the prediction results, to filter the results.

2.6.4 Other limitations

In addition to the limitations identified in Section 2.6.1, 2.6.2, and 2.6.3, we find that existing contextualization framework do not pay sufficient attention to the user interface and the moment of contact, which can greatly affect consumers’ impression and reaction to the contextualized marketing personalization contents.

(1) **User interface.** The idea that the user interface is important when displaying personalized marketing contents is not new and much supporting evidence has been presented in the popular press (Cosley et al., 2003; Pu et al., 2011; Hu and Pu, 2011; Chen and Pu, 2014). However, few studies have been devoted to the use of the most appropriate user interface and communication medium to disseminate personalized marketing information in different contexts. Accordingly, a few interesting research questions are yet to be answered. For example, which interface (e.g. pop-up ad, static banner, short message, email) should a personalization system use to display recommendations? Is the answer context specific? What kind of contextual information is the most efficient in explaining the context-specific preference for user interface? Answers to these questions are of great importance and value to firms and marketers. Therefore, more in-depth studies should be carried out.

**Moment of contact.** The moment of contact requires the selecting of the best time for doing or saying something in order to achieve the desired (best) effect. Current marketing personalization method assumes that consumers are always ready for personalized marketing
information (Gorgoglione, 2006; Abbar et al., 2009; Gallego and Huecas, 2012; Kim et al., 2014). As a result, they either diffuse personalized marketing information to consumers right after it is generated or at a particular moment specified by firms and marketers based on some heuristic rules (e.g. $x$ hours after an item is viewed; $y$ weeks since an item has been purchased).

Indeed, consumers may take the initiative to request for personalized information and expect immediate response. However, they may also be busy with their current task at hand and do not want to be disturbed by marketing messages. Therefore, intelligent context-aware personalization methods must proactively predict consumers’ needs and provide suggestions. To this aim, it is necessary to consider whether an unexpected display of personalized marketing information fosters or disturbs consumers’ shopping journey. Whereas this research topic is considered quite interesting and valuable by firms and marketers, the research on this domain is still preliminary.

Our previous experiment indicates that consumers’ instantaneous state of mind can affect their reaction to unexpected display of personalized contents (Shi et Marini, 2016). Compared to stressful online shoppers, relaxed online shoppers are more willing to interact with unexpected recommendations from an e-commerce personalization system. Although the finding indicates that the state of mind plays an important role in determining consumers’ attitude toward personalized marketing contents, there are two challenges that prevent us from predicting the best moment accurately.

1. A new contextualization method needs to be developed so that we can integrate consumers’ preference for interface and communication medium to the marketing personalization system.
A content display strategy needs to be developed based on the knowledge of consumers’ states of mind in real time so that we can predict their readiness to receive personalized information.

In view of these limitations, we present our cognition-based contextualization methodology in Chapter 3, which can make up for the above mentioned research gaps.

2.7 Marketing performance improvement

2.7.1 Performance analysis

When a marketing system is in place, it is necessary to evaluate its performance. Most academic researches rely on offline consumer data to develop personalization models. Accordingly, researchers have developed a number of methods to evaluate the performance of the personalization algorithms. We present the commonly used metrics and discuss their major drawbacks.

(1) Mean Absolute Error (MAE) and normalized MAE (NMAE). These metrics measure the total deviation of the predicted rating of consumer on item P [$r'_{A,P}$] from his (or her) actual rating [$r_{A,P}$] in the dataset $D_s$. Although a smaller MAE or NMAE value indicate that the prediction made by an algorithm is relevant to consumers’ preference, they cannot tell if a system is good at predicting items liked by consumers or excluding items disliked by them (Herlocker et al., 1999). Apparently, the latter is not a helpful function for marketers.

$$MAE = \frac{1}{|D_s|} \sum_{A,P \in D_s} |r'_{A,P} - r_{A,P}| \quad NMAE = \frac{MAE}{r_{max} - r_{min}}$$

(2) Root Mean Square Error (RMSE). RMSE also measures the deviation of the predicted rating from the actual one. It is calculated using the square root of the squared differences between the predicted and actual values. A lower RMSE value indicates a better performance of the algorithm.

$$RMSE = \sqrt{\frac{1}{|D_s|} \sum_{A,P \in D_s} (r'_{A,P} - r_{A,P})^2}$$
rating of consumer on item P \([r'_{A,P}]\) from his (or her) actual rating \([r_{A,P}]\) in the dataset \(D_s\).

However, algorithms that produce greater absolute errors may receive more punishment than systems making smaller absolute errors. This metric also has the same problem as MAE, which is not able to distinguish a preference prediction algorithm from an aversion prediction algorithm (Balabanović and Shoham, 1997).

\[
RMSE = \sqrt{\frac{1}{|P_s|} \sum_{A,P \in D_s} (r'_{A,P} - r_{A,P})^2}
\]

(3) **Revised Kendall’s Tau.** This metric uses a method to measure the prediction accuracy that is independent of rating values. The approximation of the \(\tau\) is indicated as follows, where \(C\) stands for the number of concordant item pairs, \(D\) is the number of discordant item pairs, \(T_A\) is the number of pairs that have tied ratings in consumer’s list, and \(T_P\) is the number of pairs that have tied ratings in the predicted list. \(\tau=1\) indicates that all the ratings are correctly predicted. \(\tau=-1\) means that the prediction is the opposite of the actual list. This method applies equal weight to any interchange of equal distance, no matter where it occurs. Therefore, the list \([2, 1, 3, 4, 5, 6, 7, 8, 9, 10]\) is as good as the list \([1, 2, 3, 4, 5, 6, 7, 8, 10, 9]\). However, if a consumer only considers the first three items on the list, the second list is apparently better than the first list. Unfortunately, such difference cannot be indicated by the revised Kendall’s Tau (Herlocker et al., 2004).

\[
\tau \approx \frac{C - D}{\sqrt{C + D + T_A\sqrt{C + D + T_P}}}
\]

(4) **Precision (positive prediction rate, PPR) and Recall (true positive rate, TPR).** These metrics measure the prediction accuracy of a personalization algorithm from different angles. They both start with a taxonomy which classifies a product list \(L\) into four categories. \(l_{TP}\) stands for a set of items liked by a consumer and predicted as “liked” by the
personalization algorithm (aka true positive). $l_{FP}$ represents items disliked by a consumer but predicted as “liked” (aka false positive). $l_{FN}$ is a set of items liked by a consumer but predicted as “disliked” (aka false negative). $l_{TN}$ stands for items disliked by a consumer and predicted as “disliked” (aka true negative). Based on this taxonomy, the **Precision (PPR)** is the proportion of correctly predicted items in all items that are predicted as “liked”. The **Recall (TPR)** is defined as the proportion of consumer liked items that are correctly predicted by the algorithm (Sarwar et al., 2000).

$$\text{Precision (PPR)} = \frac{l_{TP}}{l_{TP} + l_{FP}} \quad \text{Recall (TPR)} = \frac{l_{TP}}{l_{TP} + l_{FN}}$$

Precision and Recall as performance metrics have some major drawbacks. (i) Both metrics rely on consumers’ rating information. In the real retailing context, only a small portion of consumers evaluate items after they purchase them. (ii) Since different consumers have their own interpretation of the same rating score, the rating information is not consistent across consumers. As a result, the aggregated precision may be meaningless. (iii) It is not viable to ask all consumers to provide a list of their disliked items, which happen to be sold by a retailer (i.e. $l_{FN}$). Therefore, the recall is not computable in practice, making it suitable for evaluating offline consumer datasets and models only.

(5) **The area under the receiver operating characteristic curve (AUC of ROC).** The ROC curve illustrates a plot of **recall (TPR)** versus **fallout** [FPR = $l_{FP} / (l_{FP} + l_{TN})$], where points on the curve correspond to each value of the prediction threshold. Increasing the threshold to the maximum may eliminate the positive predictions. As a result, the point is moving towards the origin (0,0) because TPR=FPR=0 when the threshold is maximized. On the contrary, decreasing the threshold to the minimum may eliminate the negative predictions. Accordingly, the point is moving to (1,1) since TPR=FPR=1 when the
threshold is minimized. To evaluate different personalization algorithms, one can compare their AUC of ROC. If an algorithm P has a greater AUC value than an algorithm Q, P is considered better (Hanley and McNeil, 1982). However, the method also has several drawbacks. (i) The difference of ordering among items cannot be reflected by the metric; (ii) It is almost impossible to get the false negative data in the real world retailing scenario; (iii) Like the Kendall’s Tau, AUC cannot reflect the difference caused by the equally distant swap of two items (Lobo et al., 2008).

(6) Coverage metric. This metric provides a new perspective for evaluating personalization algorithms. The coverage indicates the proportion of item types that can be penetrated by a personalization algorithm, where $N(L)_d$ is the amount of item types or categories in the proposed items and $N$ stands for the total amount of item types (Herlocker et al., 2004).

$$\text{coverage} = \frac{N(L)_d}{N}$$

Low coverage suggests that an algorithm does not provide enough choices to consumers. For retailers, it also means that the majority of its products cannot be presented by the algorithm. The coverage metric must be used with other metrics (apparently those accuracy metrics) to evaluate the performance of an algorithm. Besides, it does not consider the ordering of items in the personalized contents, which is an important factor to consumer experience and satisfaction.

In addition to the above drawbacks, these metrics also have several common disadvantages.

(i) Whereas they can indicate from a theoretical perspective the difference between
consumers’ actual preference and the predicted preference, these metrics cannot simulate the experience and opinion of the real-world consumers.

(2) The capability of the personalization algorithm is only one of the key components that determines the performance of a personalization system. Other components might include the use of appropriate algorithm in different circumstances, the appropriate communication medium, and the selection of appropriate moment of delivery. Apparently, these important factors cannot be correctly measured by the above performance metrics.

2.7.2 Performance improvement

Current researches on marketing personalization use a static method to improve the performance of their methods (Zhang et al. 2015). They rely on offline assessments to evaluate and compare the performance of different methods (Hidasi and Tikk, 2012; Hussein et al., 2013), make modifications to the method or parameters (Bonnefoff et al., 2012; Liu and Aberer, 2013; Neves et al. 2014), repeat the test until they find the best method and its parameters (Lommatzsch, 2014; Panniello et al. 2014; Chen et al., 2015). When the best method is selected, no further improvement can be made. We believe that this approach has two drawbacks.

(1) Current improvement approach involves a lot of manual process. Since today’s personalization models are more and more sophisticated, it becomes increasingly difficult for human experts to exhaust all the improvement possibilities, not to mention to discover underlying rules or patterns that are not intuitive to human brains.

(2) Consumers preference and habits are changing every day. Today’s best method can
quickly become less effective or less efficient tomorrow. Current one-off approach cannot cope with the dynamic evolution of consumer requirements, behavior patterns, and trends. Meanwhile, it is apparently financially and operationally unfeasible to assign a team to only monitor and make improvements to the current marketing personalization approach.

In view of the drawbacks and limitations of these performance, we propose an autonomous performance analysis and improvement method for our context-aware marketing personalization system in Chapter 5.

2.8 Summary

This chapter serves as a complete review of the research on improving marketing personalization systems with contextual information. In this chapter, we introduced the cross-channel shopping behavior and its implications on marketing personalization. We also presented the current personalization and contextualization methodologies and their limitations. Through extensive and in-depth discussion of a large number of research literatures, we have found multiple levels of research opportunities.

At the theory level, it is necessary to develop a model to unify different motivations of the cross-channel shopping journey and pave the way for the identification of consumer intention, cognitive state, and their information needs.

At the methodology level, it is necessary to replace the existing process-driven personalization methodology with a new methodology so as to overcome the limitations and drawbacks of the former. Meanwhile, it is also essential to propose a new personalization methodology that is adaptive to the cross-channel shopping context.
At the technical level, it is indispensable to develop new technics to track consumers across shopping channels in real time. Meanwhile, it is necessary to track and analyze consumers’ cognitive state and preference for communication medium in real time so that firms can approach their customers in the right moment through the right way.
Chapter 3

Right content, right moment, and right medium for the right customer: a context-aware marketing personalization approach

3.1 Introduction

The real-world business situation is complex. To achieve the marketing and promotion objectives of different business units, product lines, and brands, firms like COMPANY X develop and maintain a sophisticated marketing personalization systems to maintain brand exposure (by regularly pushing new products and promotion campaigns related to consumers’ interest), increase consumer loyalty (by regularly announcing customer events, new products and news), and enhance shopping experience (by helping consumers find what they want in real time).

Different marketing personalization systems of the same firm sometimes provide excessive, redundant, and contradictory information to consumers. For instance, consumers may be encouraged to purchase peripherals related to the product they just returned. They may also be informed of a discount for a product they just purchased at
full price. When they need fresh and creative ideas, they may find that personalization systems are not able to propose better ideas than theirs. When they prefer to focus on a group of similar alternatives to make a choice, they may be disturbed by some so-called novel suggestions of personalization systems. As a result, firms may expect consumer confusion and personalization failures (Li, 2016).

Before making any improvement to their marketing personalization methods, firms and marketers should ask themselves three fundamental questions:

“Do we know the underlying intentions and information needs of our customers?”

“How do we process a marketing personalization task to fulfill these needs?”

“How do we integrate contextual factors to the personalization process?”

In this chapter, we try to provide answers to these questions. In view of the lack of a theoretical model to unify different shopping scenarios (i.e. planned, impulse, and repeat shopping scenarios) and motivations in a shopping journey, we propose a cognition-behavior model to pave the way for the explanation of the cross-channel shopping behavior in different shopping scenarios and for the identification of consumers’ underlying intentions and their information needs (Section 3.2). Based on this model, we put forward an intention-based personalization methodology to make up for the drawbacks of the current process-driven personalization methodology (Section 3.3). Meanwhile, we also propose a novel cognition-based contextualization methodology that takes into account a wider range of important contextual factors into the marketing personalization process (Section 3.4). Based on these new model and methods, we put forward our context-aware marketing personalization system (CAMPS) and its architecture (Section 3.5). Section 6 serves as the summary of the chapter.
3.2 The cognition-behavior model

In this section, we propose a cognition-behavior model to explain the cross-channel shopping behavior in different shopping scenarios (i.e. planned, impulse, and repeat shopping scenarios) in a shopping journey and identify consumers’ underlying intentions and their information needs.

*A shopping journey* is a process or progress to fulfill a shopping need. It is initiated by a preliminary shopping need (or impulse), substantiated by the process to discover and analyze information related to product (e.g. attributes, specifications, conditions, expert recommendations, customer reviews), and concluded by the decision to purchase (or abandon) a specific product. It is a process where consumers make decisions based on various kinds of information collected, filtered, analyzed, and synthesized.

A shopping journey begins with a shopping motivation that emerges in a consumer’s mind deliberately or spontaneously. A *Repeat purchase* motivation drives consumers to purchase a product they bought in the past. They shopping journey can be short if they are familiar with the product and the transactional conditions. A *Planned purchase* motivation drives consumers to enrich their product knowledge by gathering and analyzing a large amount of first-hand and second-hand information so that they can select and prioritize their selection criteria. Sometimes, an *impulsive purchase* motivation may occur to consumers. Under such circumstance, they either accept the product with no hesitation, or accept the “idea” and start looking for better options.

In Section 2.3, we made a review of the research works on the cross-channel shopping behavior and highlighted that current consumer decision-making models cannot provide a framework to explain the shopping journeys in different shopping motivations, nor can they help to indicate the cognitive state and the information needs of cross-
channel shoppers. In view of this research gap, we propose a cognition-behavior model for the cross-channel shopping journey (Figure 3.1, next page).

![Cognitive-behavior model for cross-channel shopping journey]

**Figure 3.1. A Cognitive-Behavior Model for the Cross-Channel Shopping Journey**

According to our cognition-behavior model, a real-world cross-channel shopping journey is composed of six cognitive stages, namely “intend, search, learn, try, compare, and purchase”. The shopping journey is non-linear and consumers may choose to return to a precedent cognitive state at any time.

- **At the intending stage**, consumers (no matter with an impulse purchase, planned purchase, or a repeat purchase motivation) begin their shopping journey by visiting a retailer’s sales channel. At this stage, consumers require information and inspirations that can help define their shopping ideas.

- **At the searching stage**, consumers try to skim through different items or categories related to their shopping idea. In the repeat purchase scenario, consumers just need to retrieve the item that they bought in the past. In the planned or impulse purchase
scenario, consumers may need a great number of products and categories information to refine their shopping ideas.

- At the learning stage, consumers focus on substantiating their shopping idea. In a planned shopping scenario, consumers are interested in any features, criteria, or second-hand information that are helpful when they prepare a shortlist of products for further evaluation. Such information can be anything about the brand and the product. In a repeat shopping scenario, consumers may simply skip this stage. In an impulsive purchase scenario, they seek for emotional factors that can fulfill their utilitarian or hedonic needs. In a cross-channel shopping context, if consumers find that they cannot obtain sufficient useful information from the current channel, they switch to another no matter it belongs to the same retailer or not.

- At the trying stage, consumers focus on the touching, feeling, and evaluating experience that allows them to obtain first-hand or secondary user experience of the products. Repeat purchase customers may skip this stage since they already know the product they selected. However, this stage is important for impulse purchase and planned purchase customers who have little knowledge of the product. In an offline context, consumers may try the sample. In an online context where there is no physical sample, they may refer to other online shoppers’ opinions. In a cross-channel shopping context, they can simultaneously do both to have a full-rounded opinion on the product.

- At the comparing stage, they use the learned knowledge and experience to evaluate the options and make their purchase decision. Again, repeat purchase customers may skip this stage because they already have a choice. However, the comparing stage is critical in a cross-channel shopping journey because it is in this stage that
planned and impulsive purchase consumers may change their mind and choose another retailer because of the differences in price, promotion, stock availability, service condition, or vendor (e.g. credibility, service quality).

- At the purchasing stage, consumers look for supplementary information in order to ensure that they make a smart purchase decision and the transaction can be completed as they planned. At this stage, consumers pay great attention to the total price, the payment term, the items in the order, the service packages (e.g. warranty, freebies), and the useful tips. They may be interested in other products or accessories that are complementary to their purchase decision.

The key feature of our cognition-behavior model is that it depicts a dynamic, interactive and non-linear shopping journey to simulate the reality of a cross-channel shopping journey. According to our model, consumers who are at different cognitive states are interested in different kinds of information. Such information in return drives them to a new cognitive state. As a result, we see consumers move back and forth between different cognitive stages. Due to the fact that each sales channel has its unique advantage in satisfying consumers’ informational and transactional needs, the cross-channel shopping behavior becomes logical and necessary. Our cognition-behavior model allows us to unify the three shopping motivations, explain the cross-channel shopping behaviors, and identify the intentions and information needs of consumers at different cognitive states.

3.3 The intention-based personalization methodology

Based on the diversified motivations and information needs presented by the cognition-behavior model, we propose a taxonomy of shopping intentions. We define a shopping intention as an aim that guides the subsequent shopping actions and information needs.
The taxonomy of shopping intentions can allow us to follow consumers’ information needs and their cognitive states in the cross-channel shopping context.

We develop the cross-channel shopping intention taxonomy based on a two-dimension framework (Figure 3.2). The first dimension is the similarity of product, which reflects whether the items viewed by a consumer are similar or different in terms of their major features. The second dimension is the cognitive state of a consumer, which reflects whether a consumer’s attention is concentrated or deconcentrated. Concentrated consumers are immersed in the interaction with an item or a category. Deconcentrated consumers switch between items or categories frequently. The two-by-two matrix in Figure 3.2 indicates that there are four kinds of cross-channel shopping intentions.

![Figure 3.2. The Taxonomy of Cross-Channel Shopping Intentions](image)

(1) Being an informational intention, the **learning intention (LIT)** aims at gaining insight into an item that is unfamiliar to a consumer. The LIT consumers are likely to spend
time examining (or trying) the interested item in detail and collecting information about its features. These behaviors enable the LIT consumers to enrich their domain knowledge and provide them with the first-hand experience that are useful for the subsequent purchase decision-making. The LIT consumers tend to exhibit a focused interaction pattern because they are not familiar with the product they are interacting with. They prefer to spend most of their time on a few items and rarely resort to categories because items provide more detailed information than categories.

(2) Being another informational intention, the evaluation intention (VIT) seeks to evaluate the benefits and drawbacks of various similar items in order to select the best one. The VIT consumers are familiar with the items as well as the standards used to evaluate items. They switch between different options frequently to analyze the features and specifications of the candidate items. These behaviors enable the VIT consumers to identify items that can fulfill the evaluation standards. Since they need to interact with multiple items simultaneously, the VIT consumers tend to exhibit a deconcentrated behavioral pattern. Though they also spend much more time on items than on categories, the VIT consumers distinguish themselves from the LIT consumers by frequent switching between items.

(3) Being a navigational intention, the mining intention (NIT) aims at elaborating a relatively vague shopping idea so that a more focused search can be performed. The NIT consumers narrow down the search perimeter by exploring different kinds of items (a.k.a. referential items) that have some characteristics in common. The commonality can be a feature that is collectively possessed by all the referential items (e.g. color, brand, price). It can also be a shopping idea that requires further elaboration and refinement (e.g. birthday present). Although most referential items may not be purchased
by the NIT consumers, they contribute to the formation of the final purchase decision. The NIT consumers exhibit a deconcentrated behavioral pattern because they do not need to pay too much attention to the detailed information of the referential items. Instead, the NIT consumers use more often categories than the LIT and VIT consumers because categories can guide them to a wide range of referential items.

(4) The purchasing intention (PIT) is a type of transactional intention, which seeks to acquire the selected items by completing the required transactional process. The PIT consumers exhibit a concentrated behavioral pattern because they need to ensure that key transactional information is properly provided and validated.

Since consumers may reach a sales channel because of their diversified shopping intentions, sticking to an invariant personalization methodology, as is required by the current process-driven personalization methodology presented in Section 2.5, cannot fulfill the different information needs of customers. To close this research gap, we propose a novel intention-based personalization methodology in Figure 3.3.

![Figure 3.3. Intention-Based Personalization Methodology](image_url)
(1) According to our personalization methodology, a personalization system accompanies consumers during a shopping journey and predict their instantaneous intention in real time by monitoring their activities in each cognitive step.

(2) When the system predicts that a consumer requires personalized information support, it looks for the most relevant type of information requirement(s) in the consumer behavior database where consumers’ historical preferences for the same situation are recorded.

(3) According to the requirement(s), the system uses the most performant personalization algorithm(s), the corresponding transactional and behavior data, as well as the contextual information to generate personalized marketing contents.

(4) At the same time, the system chooses the appropriate medium and the moment to present the personalized contents according to consumers’ historical preference in the same situation.

(5) The system records and analyzes consumers’ feedback to the personalized contents. If consumers accept the personalized contents (e.g. interact with the contents), the corresponding personalization effort is labeled “effective”. Otherwise, it is labeled “ineffective”. These labeled data are studied by the system to constantly improve the prediction accuracy of the personalization system.

Compared to the process-driven personalization methodology (Figure 2.2) currently used by major retailing firms, our intention-driven personalization methodology (Figure 3.3) has three advantages.

(1) Our methodology uses a flexible personalization method, which allows personalization models to selects personalization building blocks (i.e. algorithm, consumer data,
delivery moment and delivery measure) automatically according to the evolvement of consumers’ instantaneous intention linked to their real-time cognitive step. The flexible method can prevent a personalization system from proposing irrelevant contents to consumers, which is beneficial to consumers (i.e. their information load reduced and user experience enhanced) and firms (i.e. their user interface simplified and computational cost declined).

(2) Our personalization methodology incorporates new critical contextual factors (i.e. right moment, right medium, shopping budget) into the personalization model. It provides firms and marketers with new ways to improve the relevance and usefulness of their personalized marketing messages and it enables marketing systems to adapt to the cross-channel shopping context.

(3) The resilient learning tactic of our personalization methodology enables a personalization system to learn from its previous successful and failed efforts so that the system can make progress in the future by choosing the personalization strategy and building blocks that are most likely to success in a specific cognitive state.

3.4 The contextualization methodology

In Section 2.6, we made a review of the three existing contextualization methodologies (i.e. contextual pre-filtering, contextual modeling, and contextual post-filtering) and their limitations. With an aim to extend the current methodologies, we propose a new contextualization methodology in this section. Our methodology adopts a cognitive perspective, which believes that the way to generate, deliver, and present information is as important as the content of information. It decomposes a contextualization task into three phases, namely input phase, processing phase, and output phase.
Figure 3.4. Cognition-Based Contextualization Methodology

Our cognition-based contextualization methodology is presented in Figure 3.4. In addition to the three existing contextualization methods, we propose three other methods respectively at the “processing phase” (where consumer and contextual data are processed) and the “output phase” (where personalized marketing messages are delivered and presented).

- **Model elements contextualization.** This method is about the selection of the appropriate personalization building blocks (i.e. consumer data, personalization algorithm, contextual factor) associated with the current context, namely the shopping intention of a consumer. To this aim, this method enables a marketing personalization system to activate the “relevant” personalization building blocks, which are determined based on the analysis of the historical performances of each building block in different contexts.

- **Contextualization of the moment of contact.** This is the method to determine the appropriate moment of contact. To this aim, the historical performances of each
“moment of contact” under different online shopping intentions and cognitive states need to be tracked and analyzed. Such knowledge allows a marketing personalization system to predict the right moment to deliver personalized marketing messages.

- **Contextualization of the display interface.** This is the method to determine the appropriate user interface (i.e. communication medium) to present the personalized marketing contents. To this aim, the historical performances of each “communication medium” used under different online shopping intentions and cognitive states need to be tracked and analyzed. Such knowledge enables a marketing personalization system is able to predict the right medium to deliver personalized marketing message.

### 3.5 The Context-aware marketing personalization system

Based on the personalization methodology and the contextualization methodology presented in Section 3.3 and 3.4, we present our “Context-Aware Marketing Personalization System” (CAMPS), which is able to identify consumers’ real-time intention and present their preferred marketing contents to them at the appropriate moment and through the appropriate communication medium (Figure 3.5, next page).
Figure 3.5. Context-Aware Marketing Personalization System (CAMPS) Architecture
3.5.1 Business logic of the system

The functionality of the CAMPS is fueled by the evolvement of consumers’ intentions. When consumers shop in a store, the CAMPS starts to keep track of their activities and convert their behaviors into data. These data are analyzed by the CAMPS in real time to generate insights (e.g. consumer intention, cognitive step, budget, information requirement etc.) necessary for the personalization task. The CAMPS takes actions to choose the appropriate personalization strategy and building blocks (e.g. personalization algorithm, consumer data, contextual information) according to the identified intention. It generates personalized contents and decides when and how to present them to consumers. The work of the CAMPS does not end with the delivery of personalized contents. It analyzes consumers’ feedbacks to find the best personalization strategy in different contexts and use the acquired knowledge to make improvements in future personalization tasks.

Contextual information plays a critical role in the CAMPS. (1) It determines the elements, namely the personalization algorithm and the consumer data of the personalization model. (2) It contextualizes personalization results by filtering irrelevant contents. (3) It determines the best moment and medium to deliver the personalized marketing content.

The architecture of the CAMPS is composed of four layers. On the infrastructure layer, servers and devices provide the hardware to support various computation, storage, transmission, and interaction tasks performed by CAMPS. The module layer is the place where databases, algorithms, and protocols work together to provide intermediary data and values indispensable for the personalization task. On the service layer, these intermediary data are aggregated, integrated, or consolidated so that they are ready to be
exploited by consumers, marketers, firms, and other applications and services which request for personalized contents. On the application layer, the CAMPS presents these deliverables (e.g. personalized contents, CAMPS performance reports) to corresponding users (e.g. consumers, marketers, firms) in an appropriate manner.

3.5.2 The infrastructure layer

The infrastructure layer is responsible for providing data gathering measures, connectivity, storage capacity, computational capability, and human-computer interface for the proper operation of the CAMPS.

*Personal computers (PC), laptops, tablets, smartphones, and connected objects* (e.g. beacons, anchors, smart screens) are the consumer touch points of the CAMPS. Consumers use these devices to interact with items, contents, and agents (i.e. computer programs capable of completing certain pre-designated tasks) of deployed in the online and physical stores. They are also the platforms where consumers can submit service requests and receive personalized marketing contents (e.g. real-time recommendations, emails, pushed messages, remarketing contents).

*The website server and the App server* are the carrier of virtual stores where consumers can view and purchase products. These servers use various technologies to collect consumer data. For example, they deploy cookies on consumers’ devices to record behavior data. They also use tags embedded in websites and Apps to track consumer activities. Nowadays, behavior data are becoming one of the main sources of information for firms and marketers who want to better understand consumers. Recently, firms start to deploy *Internet of Things (IOT) servers* and connected objects (e.g. beacon, anchor, smart
screens) in their physical stores to keep track of consumer behaviors in the real shopping environment. An IOT server can record consumers’ position in real time. Based on these real-time positional data, personalization systems can find out consumers’ location, speed, moving direction, acceleration, as well as the frequently visited zones, shelves, or stands in a physical store. More and more firms are using such information to improve their marketing personalization systems and enhance consumers’ cross-channel shopping experience.

The transaction server is responsible for maintaining item (e.g. product, content) data and consumer data. On one hand, it is the custodian of the database of all the items and contents a firm may propose to consumers. Information in this database includes but not limited to article name, item identification number, model, specification, description, price, attribute tags, consumer ratings, stock availability, popularity, and so on. On the other hand, the transaction server manages the consumer preference database. It records consumers’ past choices (e.g. purchase history, shopping cart history, browsing history, review history, wish list) and infers their instantaneous and long-term preferences. Consumer and item database are the cornerstones of any personalization systems. The transaction server regularly updates both databases in order to make them up-to-date.

The event server encodes events collected from different platforms, consolidates these cross-channel shopping activities, synthesizes behavior streams, and then separates them into sub-streams that can be analyzed the intention prediction module. During the process, the event server is also responsible for cleansing, aggregating, caching, storing, and sharing data among various modules, services, and applications of the CAMPS.
The *analytics server* provides the computational capability for algorithms and protocols to compute the value of various variables and metrics, which are indispensable for the personalization task. At the same time, the analytics server is also responsible for analyzing consumers’ feedbacks to the personalized contents in order to produce performance reports and improve the context-aware personalization strategy for the CAMPS.

The *personalization server* is the place where personalized marketing contents are created. It chooses the appropriate moment and medium and instructs website server, App server, email campaign server, and servers of ad publishers and exchanges to deliver personalized marketing contents to consumers according to its instructions (content, moment, medium).

### 3.5.3 The module layer

The module layer hosts the algorithms, protocols and database that are responsible for the preparation of consumer behavior streams, the prediction of consumer intention, the selection of personalization strategy, and the creation of personalized marketing contents.

The *event management module* is composed of the encoding protocol, the consolidation protocol, and the fragmentation protocol. The *encoding protocol* establishes a set of rules to convert information about consumers’ online and offline activities to standardized behavior data. The *consolidation protocol* defines the way to synthesize data and create behavior streams. The *fragmentation protocol* sets the standard to divide a behavior stream into several sub-streams, which is required by both the intention prediction task and the personalization task.

The *preference management module* is composed of the feature extraction process, the
consumer-item matrix, and the similarity matrix. The feature extraction process specifies the kinds of features to be extracted in consumer and item data and determines their importance. The similarity matrix enables the CAMPS to find similar items or consumers based on the similarity coefficient of each item (or consumer) pairs. The preference management module is the cornerstone of the personalization system. It is frequently accessed by personalization algorithms to create personalized marketing contents.

The intention management module is composed of the clustering algorithm, the classification algorithm, and the budget estimation protocol. The clustering algorithm discovers various kinds of shopping intentions by studying the features of hundreds of thousands of sub-streams. The classification algorithm predicts consumers’ intention by analyzing the ways consumers do shopping and the types of information they try to obtain. The budget estimation protocol focuses on determining consumers’ shopping budget for particular items or contents based on their historical browsing and shopping habit.

The production management module is composed of the context-aware personalization strategy, the personalization algorithms, and the prioritization algorithm. Unlike traditional personalization systems which only use pre-defined personalization building blocks (i.e. personalization algorithms and delivery measures), the CAMPS chooses appropriate building blocks for a personalization task by itself based on identified consumer intention and the corresponding context-aware strategy. The personalization algorithms (e.g. community-based, popularity-based, content-based, and collaborative filtering algorithms) are responsible for finding or generating appropriate personalized contents. The prioritization algorithm is responsible for prioritizing conflicting personalized contents or delivery measures so as to maximize the expected performance of the CAMPS.
The performance improvement module is composed of the feedback tracking protocol and the learning algorithm. The feedback tracking protocol establishes the methods, rules, and standards to track consumers’ feedbacks to personalized contents in different devices and platforms. The learning algorithm analyzes the performance of the CAMPS, produces performance reports, and updates the context-aware personalization strategy if it discovers better compositions of personalization algorithms and delivery measures.

3.5.4 The service layer

On the service layer, the CAMPS coordinates individual tasks, synchronizes the working process, and consolidates the data products of different modules so as to ensure that a personalization task can be performed accordingly to the defined business logic.

The data collection and pre-processing service prepares consumer data for the CAMPS. Based on these data, the preference extraction service, the event streaming service, and the intention prediction service identify insights and intelligence that are essential for a personalization task. The A/B offloading service determines whether a consumer is able to receive personalized contents or not. It is an important function when firms and marketers want to evaluate the performance of the CAMPS. If a consumer is eligible for personalized contents, the algorithm selection, moment selection, medium selection, and prioritization service work together to present corresponding contents to him (or her). At the same time, the feedback tracking and analysis services work closely to evaluate if the contents are appreciated by the consumer. The feedback analysis service prepares the performance report of the CAMPS. If it finds better personalization strategies for a particular context, it updates the context-aware personalization strategy database.
3.5.5 The application layer

The application layer is the place where CAMPS users (e.g. consumers, marketers, system developers) can specify their needs and receive corresponding information. Consumers can receive personalized marketing contents that are adapted to their instantaneous intention. These contents may be presented as recommendations or remarketing ads in real time or as emails and App reminders that arrive at an appropriate moment. At the same time, marketers and system developers can monitor the performance of the CAMPS in real time. If needed, they can also examine the instantaneous intention and preference of particular consumers.

In this section, we presented the architecture of the CAMPS, a context-aware marketing personalization system. We introduced the business logic of the system and detailed its infrastructures, modules, services, and applications. The CAMPS has three new features compared to the existing personalization systems. (i) It chooses by itself the most appropriate building blocks to complete a personalization task rather than following invariant rules made by human programmers. (ii) It is a context adaptive system driven by consumers’ instantaneous intention. The idea to provide information at the right moment through the appropriate medium makes the CAMPS more flexible and relevant than the existing systems. (iii) The CAMPS is able to learn from its previous efforts and make improvement in the future. Hence, it is less likely to make the same mistakes.

3.6 Summary

This chapter aims at establishing a context-aware marketing personalization approach (CAMPA). We begin by presenting a theoretical model to describe and explain the non-linear cross-channel shopping behaviors with different shopping motivations. The
model also allows firms to track and analyze consumers’ real-time cognitive states and information needs. Based on this model, we proposed our intention-based personalization methodology and the cognition-based contextualization methodology, which lay the foundation of the subsequently proposed context-aware marketing personalization system (CAMPS). The framework, architecture and processes of the CAMPS are detailed in this Chapter.

Our propositions focus on consumers’ real-time intentions and cognitive states, which give firms and marketers a new and insightful perspective to track and analyze their customers and potential customers. With the real-time knowledge of consumer intention and cognitive state, firms are able to predict the information needs of their customers in a more accurate and relevant manner. As a result, it is less likely that they use less relevant resources to generate irrelevant marketing contents, which waste company resources and confuse customers. Such real-time knowledge also enables firms to use more user-friendly media to reach their customers through their preferred moment of content. Accordingly, customers’ level of engagement and satisfaction can be enhanced.
Chapter 4

Cross-channel consumer tracking: towards a holistic view of consumer journeys

4.1 Introduction

As we discussed in Chapter 1, more and more consumers are becoming cross-channel shoppers who use two or more retailing channels or platforms simultaneously. The new way of shopping brings a question to the designers of marketing personalization approaches. That is, “how should marketing personalization systems treat data and information that come from different retailing channels?”

If they maintain a mono-channel perspective as is required by the process-driven personalization methodology, these systems may not need to be revamped. However, consumers may receive weird personalization contents because the marketing personalization system only has partial knowledge of consumers’ shopping intention, cognitive state, information needs, and behaviors. To enhance the relevance of personalized marketing contents to consumers’ requirements for personalized information, we need to regard consumers’ cross-channel shopping behaviors as an integral shopping journey and consolidate data coming from different channels.
To respond to this challenge, we propose in this chapter a cross-channel event management method, which necessitates three steps. (1) A protocol must be established to determine the types of consumer data to be collected and the method to standardize the data collection process across different channels. These topics are discussed in Section 4.2. (2) A second protocol needs to be created so that various types of consumer data coming from different retailing channels and platforms can be consolidated and converted into a behavior stream according to the standard rules. These topics are covered in Section 4.3. (3) Due to the variance and evolvement of consumer intention, it is necessary to divide a behavior stream into several sub-streams so that they can be analyzed and compared. In Section 4.4, this topic is discussed. Section 4.5 serves as the summary of the chapter.

4.2 The encoding protocol

In this section, we discuss the first step of the cross-channel shopping event management. That is, to collect and encode consumer data from different shopping channels and platforms. The discussion is composed of two topics. (1) Which kinds of data should be gathered? (2) How to encode these data?

4.2.1 The scope of data collection

There are three kinds of retailing platforms that cross-channel shoppers are most likely to use, which are e-commerce websites, mobile Apps, and physical stores. The first two platforms are also called “online platforms” in that consumers have to connect to these virtual stores from distance. Physical stores used to be ignored by marketing personalization systems several years ago because there was no reliable measure to track consumers’ in-store behavior. However, thanks to the prevalence of indoor positioning
technology, more and more kinds of consumer data are made available to marketing personalization systems.

Consumer data and information generated on the e-commerce websites, m-commerce Apps, and physical stores have common characteristics because these online platforms are created to imitate and enhance the physical stores. Therefore, we can map different products and product segments displayed in a physical store to a directed acyclic graph (DAG) that signifies the structure of an e-commerce website or an m-commerce App (Figure 4.1).

![Figure 4.1. Mapping Virtual Categories and Items to Store Shelves and Products](image)

According to the above mechanism, shelves and product segments in a physical store are represented by the “category pages” of different hierarchical levels, which can group similar products and categories together. A higher-level category page contains several lower-level category pages. The category pages of the lowest level are the containers of products or contents that have certain characteristics in common. These products and contents are called “items”. For example, C1 represents all the beauty products of the
store. C2a signifies the products for hair care. C3a is the container of all the conditioner products. And Ia is the product information page (i.e. item page) that presents the detail information of a particular conditioner to consumers.

Regardless of which retail platform a consumer may use, the CAMPS is interested in two types of information: (1) the characteristics of items and categories a consumer interacts with. (2) the types of consumer behaviors associated with these items and their corresponding timestamp. On the online platforms, information is collected by cookies or tags, which are embedded in e-commerce websites and Apps. In the offline stores, the CAMPS must rely on anchors, access points, beacons, and smart devices deployed in a store to gather consumer data and transmit them to the IOT server for cleansing and pre-processing. We define the types of data to be collected by the CAMPS as follows.

4.2.1.1 Data Collection on the Online Platforms

According to the cognitive steps defined by our intention-based personalization methodology in Chapter 3, consumers’ online shopping intentions can be reflected by their efforts to search, click, browse, like, choose, purchase, or reject items on e-commerce websites and m-commerce Apps. Accordingly, the following kinds of consumer behaviors need to be recorded by the CAMPS.

*Reaching and leaving a retailing platform.* These events indicate the arrival and departure of a consumer. The time in between these events is called a “session”. When consumers arrive in a retailing platform, the CAMPS record the event, the timestamp, the information of the platform, and his (or her) unique user identification (UUID). It also creates a session identification (SSID) for him (or her).
Search. The search behavior indicates a precise or vague consumer requirement for information. Frequent search efforts in a short period of time may suggest that consumers cannot find what they want or that they cannot articulate their requirements. The CAMPS must record every search event, the platform where the event occurs, the timestamp as well as the associated UUID and SSID. Search keywords are not recorded, since they are not indispensable for the intention prediction task.

Click (website) and tap (App) of an item or category. The click/tap behavior indicates that consumers are interested in an item or a category. Even though they may not purchase the clicked item or category, the click/tap behavior itself suggests that the clicked/tapped item or category is somehow related to consumers’ instantaneous needs. The CAMPS records every click/tap event linked to an item or category, the timestamp, the platform information, the item or category’s unique identification (UPID or UCID), the price (for an item) as well as the associated UUID and SSID.

Display. The display event is defined as the system actions to provide contents, which can be either consumer requested contents or personalized marketing propositions, to consumers. The CAMPS records every display event, the timestamp, the platform, the price of the item, the UPID and UCID as well as the associated UUID and SSID.

Detach. The detach event is defined as any behaviors that result in consumers’ departure from the current item page. Detach events are derived based on the pre-defined behaviors, which include viewing another item, closing a web page or App, returning to a category, or adding an item to the shopping cart. The CAMPS records the detach event, its timestamp, the platform information, the UPID and UCID as well as the UUID and SSID associated with the detach event.

Adding item to the shopping cart. This behavior indicates that the added item is highly
relevant to a consumer’s instantaneous need. The CAMPS must record every add-to-cart event, the timestamp, the platform information, the price of the item, the corresponding UPID as well as the associated UUID and SSID.

Removing item from the shopping cart. This behavior indicates that the removed item is no longer relevant to a consumer’s instantaneous needs. The CAMPS records the remove event, the timestamp, the platform information, the price of the item, the corresponding UPID as well as the associated UUID and SSID.

Purchase. The click/tap of the “pay” button suggests that a consumer has made the decision. The CAMPS must record the purchase event, the timestamp, the platform, the corresponding UPID, the retailing price of the item, the quantity, the size of the shopping cart (i.e. the monetary value and the quantity of each item), the transaction identification (TRID) as well as the associated UUID and SSID.

Scroll / slide. This behavior indicates that consumers are skimming a web or App page and they pay no attention to the detailed information. The CAMPS record every scroll/slide event, the timestamp, the platform, the UUID, and the SSID.

4.2.1.2 Data Collection in Physical Stores

The CAMPS relies on various Bluetooth, Wi-Fi, UWB, or RFID technology based connected objects (e.g. anchors, access points, beacons, smart devices) to collect consumer data in physical stores (Figure 4.2). The CAMPS measures a consumer’s instantaneous positions based on his (or her) distance to nearby anchors or access points. Using positional information, it infers a consumer’s velocity and nearby shelves/products.

$$v_j = \sqrt{\frac{(x_{je} - x_{js})^2 + (y_{je} - y_{js})^2}{(T_{e_j} - T_{s_j})}}$$ (Equation 4.1)
The computation is conducted once every X seconds. Depending on the capability of the hardware, X ranges from 1 to 10 seconds. The CAMPS records the starting timestamp (Tsj) and ending timestamp (Tej) of each computation period. It also records the coordinates of a consumer’s starting position (xjs, yjs) and ending position (xje, yje). These data are used to compute the average velocity of a consumer (vj) during the period [Equation 4.1]. According to the velocity, consumer’s status can be described as high-speed walking, low-speed walking, and stop-over. In addition, the CAMPS records the UCID of the nearest shelf to a consumer for each starting and ending position based on the [Equation 4.2]. Such information can be used to infer a consumer’s instantaneous preference and needs.

\[
D_{p1,p2} = |x_{p1} - x_{p2}| + |y_{p1} - y_{p2}| \quad \text{(Equation 4.2)}
\]
When consumers enter or leave a physical store, the connection and disconnection event can be recorded by CAMPS according to the same way used by the online. When consumers pay for the products they have chosen, the purchase data are uploaded to the CAMPS as well.

### 4.2.2 The modality of data encoding and collection

Based on the defined data collection protocol, consumer data from three platforms can be gathered using a standard way. We use Table 4.1 (next page) to illustrate all kinds of consumer data the CAMPS may collect. We use codes to distinguish between different events and retailing platforms. Other consumer data are recorded according to their original formats.

In the example, a consumer (XS30MW6092) began his shopping journey on a retailer’s e-commerce website. He found an item (105.210.301.401) and added it to his shopping cart. Then, he found a better item (105.210.301.402) using the retailer’s m-commerce App and updated his shopping cart. Then he paid for the new item and quitted the retailer’s website and App. On the other day, he went to one of the retailer’s store. He went directly to a shelf, located the product he wanted, purchased it, and left the store.
<table>
<thead>
<tr>
<th>eventName</th>
<th>timestampS</th>
<th>timestampE</th>
<th>platform</th>
<th>UUID</th>
<th>SSID</th>
<th>UPID</th>
<th>UCID1</th>
<th>UCID2</th>
<th>price</th>
<th>quantity</th>
<th>cartValue</th>
<th>cartItem</th>
<th>cartQty</th>
<th>TRID</th>
</tr>
</thead>
<tbody>
<tr>
<td>LD (arrive)</td>
<td>20161225D18:30:45T</td>
<td>20161225D18:30:45T</td>
<td>W (web)</td>
<td>XS30MW6092</td>
<td>U3PV690G286REC09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH (search)</td>
<td>20161225D18:31:08T</td>
<td>20161225D18:31:08T</td>
<td>W (web)</td>
<td>XS30MW6092</td>
<td>U3PV690G286REC09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT (click/tap)</td>
<td>20161225D18:31:52T</td>
<td>20161225D18:31:52T</td>
<td>W (web)</td>
<td>XS30MW6092</td>
<td>U3PV690G286REC09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OL (scroll)</td>
<td>20161225D18:32:17T</td>
<td>20161225D18:32:17T</td>
<td>W (web)</td>
<td>XS30MW6092</td>
<td>U3PV690G286REC09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD (arrive)</td>
<td>20161225D18:40:40T</td>
<td>20161225D18:40:40T</td>
<td>A (app)</td>
<td>XS30MW6092</td>
<td>U3PV2533G8HNO4Z</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH (search)</td>
<td>20161225D18:46:33T</td>
<td>20161225D18:46:33T</td>
<td>A (app)</td>
<td>XS30MW6092</td>
<td>U3PV2533G8HNO4Z</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD (add)</td>
<td>20161225D18:50:01T</td>
<td>20161225D18:50:01T</td>
<td>A (app)</td>
<td>XS30MW6092</td>
<td>U3PV2533G8HNO4Z</td>
<td>105.210.301.402</td>
<td>105.210.301.400</td>
<td>120.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QT (quit)</td>
<td>20161225D18:50:21T</td>
<td>20161225D18:50:21T</td>
<td>W (web)</td>
<td>XS30MW6092</td>
<td>U3PV690G286REC09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QT (quit)</td>
<td>20161225D18:50:29T</td>
<td>20161225D18:50:29T</td>
<td>A (app)</td>
<td>XS30MW6092</td>
<td>U3PV2533G8HNO4Z</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD (arrive)</td>
<td>20161226D11:13:55T</td>
<td>20161226D11:13:55T</td>
<td>S (store)</td>
<td>XS30MW6092</td>
<td>U64G1NSCV82KB7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LW (low speed walk)</td>
<td>20161226D11:14:00T</td>
<td>20161226D11:14:05T</td>
<td>S (store)</td>
<td>XS30MW6092</td>
<td>U64G1NSCV82KB7</td>
<td>105.210.301.500</td>
<td>105.210.301.500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP (stop-over)</td>
<td>20161226D11:14:05T</td>
<td>20161226D11:14:10T</td>
<td>S (store)</td>
<td>XS30MW6092</td>
<td>U64G1NSCV82KB7</td>
<td>105.210.301.600</td>
<td>105.210.301.600</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QT (quit)</td>
<td>20161226D11:53:33T</td>
<td>20161226D11:53:33T</td>
<td>S (store)</td>
<td>XS30MW6092</td>
<td>U64G1NSCV82KB7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.2.2.1 UPID and UCID Management

Due to the complexity of the retailing business model, we think it is necessary to extend the discussion on the management of UPID and UCID. As more and more cross-channel retailers (e.g. Amazon, Taobao, Fnac, and Rakuten) choose to open their retailing platforms to other vendors (i.e. marketplace model), it is quite likely that different vendors are selling the same items. If each vendor manages the UPID and UCID based on its own modality, the same item/category is represented by different UPIDs/UCIDs. This is a disaster to these cross-channel retailers because their marketing personalization system cannot extract consumers’ preference correctly.

To avoid such catastrophic consequence, we propose to unify the management modality of UPID/UCID. Vendor must use the taxonomy modality determined by the retailing platform. However, they can add a suffix (based on the rules provided by the platform) to the standard UPID/UUID to identify themselves. As a result, vendors keep their independence and retailing platforms can consolidate item and category data for analytical purposes.

The standard UPID and UCID are not generated randomly. We propose a hierarchical modality to encode product segments (i.e. categories) and products (i.e. items). The catalogues of retailing firms can be represented by a DAG, where higher-level nodes represent major product categories and lower-level nodes represent product segments in each category.

Taking the DAG in Figure 4.3 (next page) as an example, the retailer’s product catalogue is composed of three category tiers and an item tier. Each node (i.e. oval) represents a category and each leaf (i.e. rectangle) is an item. The UCID of the highlighted T-1 node is 103.200.300.400 because it is the third T-1 node (103) that contains various T-2 (200)
and T-3 (300) nodes. The UCID of the highlighted T-2 node is 103.202.300.400 because it belongs to the third T-1 node (103) and it is the second node (202) of the branch. The UCID of the highlighted T-3 node is 103.202.301.400 because it belongs to the third T-1 node (103) and the second T-2 node (202) and it is the first node of the branch (301). And the UPID of the highlighted item is 103.202.301.401 because it is the first leaf (401) of the branch.

![Diagram of hierarchical encoding modality]

Figure 4.3. Hierarchical Encoding Modality

Our hierarchical encoding modality provides each category and item with a unique identification. It also allows marketing personalization systems to quickly determine how similar two or more items and categories are without making tedious semantic analysis. For example, item_a (103.202.301.401) and item_b (103.202.301.402) are brothers because they have the same parent (301). On the contrary, item_a (103.202.301.401) and item_c (103.202.311.402) are cousins because they only have the same grandparent (202).

4.3 The consolidation protocol

Once channel-specific consumer data are gathered, they need to be consolidated according to a protocol so that the CAMPS can have a complete view of consumer preference and activities in the cross-channel shopping journey. The data consolidation
concerns two subjects. (1) How to identify a consumer who uses different retailing channels and platforms simultaneously? (2) How to consolidate the collected consumer data?

4.3.1 Cross-channel consumer identification

We propose to use a unique user identification system (UUID) to identify consumers who use different retailing platforms. (1) To begin, the CAMPS obtain the consumer identities of a retailer across all its channels and platforms. It compares these identities and merges the information of two identities if they have at least three exact matches of key value (i.e. username, email address, mobile phone number, delivery address, fixed phone number, device identification). (2) The comparison and merge process is performed irately until no consumer identity can be merged. At this moment, the CAMPS allocate a UUID to each of the remaining consumer identity. (3) Every time when a consumer lands on the retailing platforms, the system starts to analyze his (or her) identity information. If the consumer has logged in (i.e. an identified consumer), the system tries to collect additional information (if any) that is not known to the system so as to enrich the identity. (4) Otherwise, the system allocates a temporary tracking identification to the consumer for the time being and keeps collecting his (or her) identity information that can help recognize the consumer. If the gathered key information matches an existing identity, the consumer is identified even if he (or she) is not login. (5) If none of the existing identity is matched by the end of a session, the system assigns a UUID to the new identity and update the identity database.

Since each UUID is associated with a consumer’s identity and device information in all retailing channels and platforms, cross-channel shoppers can be identified. The UUID system cannot achieve the level of accuracy of a biometric system. However, the mistake is tolerable due to two reasons. (i) It is almost impossible that two different consumer
identities have three identic key information fields because it is not allowed by the new user registration rule implemented by the retailer. Therefore, the only mistake the UUID system may make is to recognize an existing user as a new one due to insufficient consumer identity information. However, this mistake does not have a huge impact on consumer’s shopping experience and it can be corrected automatically when more identity information is available to the system.

4.3.2 The data consolidation protocol

The CAMPS uses UUID and event timestamp as key to consolidate consumer shopping data gathered from different channels.

(1) When a consumer initiates a session in any of the retailer’s retailing platforms, the CAMPS searches his (or her) UUID in the active UUID container, which caches all the UUIDs that are interacting with the retailer. If there is a match, this new session is consolidated to the existing behavior stream. Otherwise, the CAMPS creates a new behavior stream and update the active UUID container.

(2) Events (i.e. consumer behaviors) of two sessions are consolidated according to the chronological order.

(3) The task to consolidate data and create behavior stream is terminated only if the last session within the stream is terminated. A session ends if a consumer is no longer active (i.e. no events) during a period of time.

We present two exemplary behavior streams created based on the above data consolidation protocol. The stream in Figure 4.4(a) presents that a consumer removed an item that he added through the website channel and purchased it using the
App channel. This is quite common when a retailer provides a channel-specific discount to encourage consumers to use a particular channel. The stream in Figure 4.4(b) depicts a free-rider who went to a store to try a product and then buy it using the App channel where the price is lower.

![Diagram](image)

**Figure 4.4. Consolidating Cross-Channel Data based on UUID and Event Timestamp**

The consolidated behavior streams can give firms and marketers a more holistic view of consumer behaviors. From a mono-channel perspective, the firm’s website and store channel need to follow-up the unfulfilled sales lead while the two impulsive purchases on the App channel is hard to explain. From a cross-channel perspective, the firm’s customer relationship management (CRM) system does not need to make any more follow-up efforts because the purchase behaviors have already been performed on the App channel as the extension of previous shopping journeys.
4.4 The fragmentation protocol

Using the encoding protocol and the consolidation protocol, the CAMPS is able to generate a complete cross-channel behavior stream when a consumer uses more than one retailing channels or platforms simultaneously. However, it is inappropriate for the CAMPS to predict consumer intention and preference based on an entire behavior stream, due to three reasons.

(1) Intention and preference may change over time as consumers continue their shopping journey. If a behavior stream cannot be properly split into several sub-streams, the CAMPS is not able to determine the variance of intention and preference.

(2) Consumers’ behavioral features and patterns are dependent on their intention and preference. Some key behavioral features and patterns that are locally significant may not be prominent in the entire behavior stream. If a stream is not fragmented, these features and patterns may be too feeble to be detected.

(3) The CAMPS is designed to assist consumers and foster their shopping experience throughout the shopping journey. If behavior streams are not fragmented, the CAMPS cannot provide personalized marketing contents to consumers during their shopping journey. In view of the necessity to divide a behavior stream to several sub-streams, we propose the following fragmentation protocol.

4.4.1 The milestone event

Dividing a behavior stream into several sub-streams at a particular time interval is the simplest fragmentation method. However, an arbitrary fragmentation rule of this kind
can lead to misjudgment because there could be more than one intention and preference within a time interval. To avoid such mistake, we propose a cognition-based fragmentation protocol which divides a behavior stream based on the “milestone events”.

The **milestone events** are the consumer behaviors that can indicate the change of cognitive step, intention, preference, and need of consumers. According to the definition, we propose four kinds of behavior as milestone events.

1. **Adding item to the shopping cart (AD).** This event indicates that the added item is considered highly relevant to a consumer’s instantaneous need. Before this event, a consumer makes efforts to search for possible options and compare them. After this event, a consumer may proceed with the payment, search for another type of item, or quite the website.

2. **Removing item from the shopping cart (RM).** This event shows that the removed items are no longer considered relevant by consumers, who are likely to search for new items that may fulfill their needs.

3. **Return to top tier categories (UN).** This event can be inferred by comparing the UCID of the current click event and the UCID of the previous click. Items and segments (i.e. lower tier categories) within two top tier categories have significant differences. Each retailing firm may define its own top tier categories according to the structure of its product catalog. In the physical shops, consumers are either next to one shelf (i.e. product segment) or another. The CAMPS cannot recognize any hierarchical browsing paths. However, it can map the obtained UCID to the firm’s product catalog and determine whether the two segments are within the same top tier category or not.
(4) Purchasing (PU). The purchase event suggests an end of a shopping journey. Before this event, consumers provide necessary information to complete the transaction. As soon as the payment is made, they may leave the retailing platform or start to plan for the next shopping list. The CAMPS need to be informed of the purchase event so that it can adjust its personalization strategy accordingly.

4.4.2 The preparation of sub-streams

The process to prepare sub-streams is composed of the following steps.

(1) When a new behavior stream is created, the CAMPS analyzes the events within. As soon as it recognizes a milestone event (e.g. AD: add to cart, RM: remove from cart, UN: return to top tier categories, PU: purchase), it marks the first sub-stream whose starting time is the timestamp of the first event and the ending time is the timestamp of the first milestone event.

(2) It computes the metrics that are required to predict consumers’ cognitive step, intention, preference, and budget based on events recorded during the first sub-stream.

(3) At the same time, CAMPS keeps analyzing new events that are generated by consumers. When it recognizes a new milestone event, it marks a new sub-stream, which starts right after the first milestone event and ends at the second milestone event.

(4) The CAMPS computes the corresponding metrics using events recorded during this sub-stream.

(5) The CAMPS repeats step (3) and (4) until a consumer is no longer active (i.e. no more events) on any platform of a retailer.
The exemplary behavior stream in Figure 4.5 illustrates the sub-streams that are separated based on the defined milestone events. Sub-streams are of different lengths. It suggests that consumers may generate more events in certain circumstances than in others. This feature may help us to distinguish some intention states from others. Sub-streams may also contain events collected from different retailing platforms. By sorting cross-channel shopping behaviors in chronological order and analyzing the causal relationship between the antecedent behavior and its subsequent behaviors, the CAMPS can decode the complex consumer decision-making process.

![Behavioral Stream Diagram]

Figure 4.5. Dividing a Behavior Stream into Sub-streams based on Milestone Events

4.5 Summary

The normal operation of the CAMPS relies on the behavior streams and sub-streams that contain cross-channel and intention-specific shopping information. In this chapter, we presented our method to manage cross-channel shopping behaviors. We began the discussion by establishing an encoding protocol which specifies the kinds of consumer data to be collected and the way to encode them. The protocol ensures that events collected from different platforms of a retailer are compatible and comparable.

Then, we proposed the protocol to consolidate the cross-channel shopping behaviors. The discussion focuses on two issues: (1) How to identify consumers in a cross-channel shopping environment? (2) How to generate behavior streams by merging consumer
data collected from different retailing channels and platforms?

Finally, we discussed the necessity to divide a behavior stream into several sub-streams and proposed the fragmentation protocol which divides a behavior stream based on certain types of milestone events. The divided sub-streams contain cross-channel shopping information that is associated with consumers’ instantaneous intentions. The CAMPS can use information in each sub-stream to compute the value of features and metrics required by the intention management module, preference management module, production management module, and performance management module. The definition and function of different features and metrics are discussed in Chapter 5.
Chapter 5

From data to marketing success: interpretation of consumer intention and personalization of marketing communication

5.1 Introduction

In the precedent chapter, we proposed the protocols and methodological steps to prepare behavior streams and sub-streams that are indispensable for the proper operation of the modules of the CAMPS. Sub-streams contain a great amount of information about consumers’ intentions, preferences, habits, and needs. This chapter focuses on the extraction and exploitation of the information.

Unlike traditional linear marketing personalization systems that only analyze consumer preferences, the CAMPS takes into account the context information that is critical to consumers’ decision-making and behavioral pattern. Accordingly, the CAMPS has four core modules instead of one. In addition to the preference management module, the CAMPS also relies on the intention management module, the production module, and the performance improvement module to maintain its proper operation.
In Section 5.2, we present the intention management module, which serves as the starting step of the personalization process. The intention management module determines consumers’ instantaneous shopping intention based on their behaviors and interested contents so that the CAMPS may choose the appropriate personalization strategy. This section proposes the algorithms to recognize the shopping intentions defined in Section 3.3. It also details the protocol to estimate consumers’ budget, which is part of their shopping intention.

Since the CAMPS is a context-aware system, it needs to have various personalization algorithms to cope with different situations. In Section 5.3, we introduce the preference management module, which is composed of the personalization algorithms used by the CAMPS as well as the methods to extract and exploit consumer preference data.

Being a self-learning system, the CAMPS “knows” the best personalization strategy for each shopping intention and cognitive state because it infers these insights by analyzing millions of behavior streams. In Section 5.4, we detail the performance improvement module that makes the CAMPS smart through a learning process.

In Section 5.5, we introduce the production module which is responsible for the generation and delivery of the personalized marketing contents.

Section 5.6 serves as the summary of the chapter.

### 5.2 The intention management module

The intention is an aim that guides action. According to the cognition-behavior model (Figure 3.1) and the intention-based personalization methodology (Figure 3.3) we proposed in Chapter 3, the intention is the starting point of a shopping journey and the
source of all the subsequent behaviors. In the mono-channel shopping context, consumers’ intention is to purchase items because they have to complete all the cognitive steps (i.e. search, learn, test, compare, purchase) in the same retailing channel (Shi, 2016). The cross-channel shopping context complicates the problem in that a consumer may use different retailing channels or platforms to achieve objectives that are defined by different cognitive steps (Verhoef et al., 2015). As a result, marketing personalization systems must prepare to serve consumers who are at different cognitive steps and have diversified intentions.

5.2.1 The discovery of intention

Based on the four intention types we defined in Section 3.3, we present the method used by the CAMPS to recognize consumers’ intention in the cross-channel shopping context. The discussion is composed of three topics: (1) the selection of the intention discovery method; (2) the metrics used to recognize intention; (3) the algorithms used to recognize intention.

5.2.1.1 The selection of the intention discovery method

Intention discovery poses a major challenge to the research on consumer behavior in the cross-channel shopping context. Whereas several methods have been proposed, their drawbacks are unneglectable.

(1) The first method is to survey consumers for information, which works well on such websites like MovieLens where users are motivated to provide feedbacks and information so as to get more accurate and personalized services. However, as we have already discussed in Section 1.4, the method has three major drawbacks which make it unsuitable for the retailing context. Therefore, it is not used by the CAMPS.
The second method is to analyze users’ query keywords in order to infer their intention. This method compares users’ query keywords to a database containing baseline keywords and corresponding intention labels predefined by domain experts. When a user keyword matches a baseline keyword, the corresponding intention label is assigned to the user (Liao et al., 2011). Although this method can identify instantaneous intentions automatically and unobtrusively, it loses its predictive power when consumers choose not to use a search engine in the shopping journey. In addition, the accuracy of the intention discovery is contingent with the definition of baseline keywords and their intention labels. For a system that is developed to serve millions of consumers, this method is not reliable.

The third method is to predict consumers’ intention based on the pattern recognition method. Although empirical results suggest that pattern recognition is efficient in predicting purchase intention, it is unable to identify other shopping intentions (Montgomery et al., 2004). We believe that pattern recognition is a promising tool but it should not be used alone to predict consumers’ intentions.

To avoid the above mentioned inconveniences and drawbacks, many real-world marketing personalization systems assume that consumers’ instantaneous intentions are contingent with the item they explored recently (Shi et al., 2014). However, such method is not always effective, especially when consumers’ shopping objective needs refinement and clarification (Shi et al., 2015).

In view of the drawbacks of the above methods, we believe that a robust intention discovery method should fulfill three requirements. (1) It must minimize the disturbance to consumers, meaning that the intention discovery process should be automatic and unobtrusive. (2) It needs to be accurate and complete, meaning that it should be able to
identify different kinds of intentions correctly. (3) It must determine the variance of consumer intention, meaning that the method should be able to keep track of consumers’ cognitive status as well as the contents they are interacting with.

Based on these standards, we propose a progressive intention discovery method based on machine-learning technology for the intention management module of the CAMPS (Figure 5.1), which is composed of four steps.

![Figure 5.1. The Intention Management Module](image)

(i) The CAMPS learns to identify the features (to be defined in the next paragraph) of each intention status based on the offline sub-streams and the clustering algorithm (aka unsupervised learning algorithm). It analyzes each intention cluster and assigns to it an appropriate intention label capable of depicting its characteristics. Labeled data points in each cluster are called seeds.

(ii) The CAMPS uses supervised learning algorithms to analyze the seeds and generate a work-in-progress (WIP) predictive model. The WIP predictive model is iteratively modified by the CAMPS until the model's prediction accuracy reaches an acceptable level.
The CAMPS uses the WIP model to predict the intention of sub-streams in the real retailing environment. As soon as a prediction is made, the CAMPS invites consumers to feedback their instantaneous intention. The enormous numbers of consumers and sub-streams can provide enough labeled data (i.e. new seeds) for the CAMPS.

These new seeds, which are the feedbacks of real-world consumers, are used to further enhance the accuracy of the intention prediction model of the CAMPS.

5.2.1.2 The selection and definition of the features

Based on our two-dimensional intention taxonomy defined in Section 5.2.1.1, the CAMPS can use two features to differentiate four types of intentions. To further improve the prediction accuracy, we propose to use the following three features.

`isPurchase`. This is a Boolean variable which determines if there is a PU event (i.e. purchase) in the sub-stream. If `isPurchase`=1, the sub-stream is labelled “PIT” (i.e. purchase intention) immediately. If `isPurchase`=0, the CAMPS proceeds with the subsequent learning process.

The average time per item interaction (ATIN). This is a numeric variable that reflects the average duration time of each item interaction within the perimeter of a sub-stream. An item interaction is defined as a period of time when a consumer interacts with and only with the particular item. The metric can be computed using the [Equation 5.1].

\[
ATIN_j = \frac{\sum_{x}^{x} (T_{ex,j} - T_{sx,j})}{x} \quad (Equation \ 5.1)
\]

Let us assume that there are `x` item interaction events in the sub-stream `j`. According to our encoding protocol defined in Section 4.2.1, each event has a starting timestamp (`T_{sx,j}`) and an ending timestamp (`T_{ex,j}`). Therefore, the ATIN is equal to the total of the item
interaction time divided by the number of item interactions within the sub-stream. A high ATIN value indicates that a consumer is deeply immersed in the interaction with the item. A low ATIN value suggests that a consumer is more interested in gathering non-item information or comparing different items.

The similarity of interacted items and categories (SMIT). This is a numeric variable that reflects the similarity of a group of items and categories that are explored by a consumer within the perimeter of a sub-stream. The metric can be computed using the following equations (Figure 5.2). According to our encoding protocol defined in Section 4.2.2, each item and category in the virtual and physical shops can be mapped to a DAG that represents the catalog of a retailer.

\[
d_{R_P} = \begin{cases} 0 & (N = 1) \\ \frac{1}{N-1} & (N > 1) \end{cases} \quad \text{Equation 5.2}
\]

\[
d_{R_1} = \begin{cases} 0 & (N = 1) \\ \sum_{i=1}^{N} \frac{1}{N-1} & (N > 1) \end{cases} \quad \text{Equation 5.3}
\]

\[
d_{R_S} = d_{R_1} + d_{R_1} - 2 * d_{R_1} \quad \text{Equation 5.4}
\]

\[
SMIT = \begin{cases} 0 & (X = 1) \\ \frac{1}{X-1} \sum_{n \in \text{NCP}, n \neq X} d_{MN} & (X > 1) \end{cases} \quad \text{Equation 5.5}
\]

Let a catalog has \(N\) tiers. The root node is the only node that is located in Tier 1. The items nodes are located in the \(Nth\) tier. The remaining tiers are reserved for category nodes. We further specify that the distance from any (item or category) node to its direct parent node is \(1/(N-1)\) [Equation 5.2], where \(N\) is the tier number that the node is located.

![Figure 5.2 Method to Compute the SMIT](image-url)
Accordingly, the distance from any node to the root node can be computed using [Equation 5.3]. And the distance from any node R to another node S can be computed using [Equation 5.4], which define the distance as the sum of the two nodes’ respective distance to their common parent node P. For instance, the distance between node R and S is much less than the distance between K and L. Because R and S are “brothers” and K and L are “cousins”. Finally, if a sub-stream contains X nodes (i.e. different UPIDs and UCIDs), the SMIT of these nodes can be computed using [Equation 5.5]. A high SMIT value indicates that these items and categories are different. A low SMIT value suggests they are similar. Specially, the SMIT value is 0 if there is only one item or category in the sub-stream.

Due to technology and capability limitations, current indoor geolocation devices are not able to correctly and stably capture item-level interaction information (i.e. time spent on each item) in the physical shops. In view of this uncertainty, we do not incorporate data from physical shops when computing the ATIN for the moment, even though our method allows for those data to be included. However, the computation of the SMIT is not impacted since the required information (i.e. nearest UPID or UCID when consumers make a stop-over) is easy to obtain.

5.2.1.3 The selection of machine learning algorithms

According to our intention discovery method presented in Figure 6.2, the CAMPS uses an unsupervised machine learning algorithm to identify the intention clusters and then uses a supervised machine learning algorithm to develop the intention prediction model. We start with the selection of the unsupervised learning algorithm for the clustering task.

- Selecting a clustering algorithm
There are two clustering methods (see Annex 2.3), namely the hierarchical method and the partitional method.

(1) The **hierarchical clustering method** uses previously identified clusters to find subsequent clusters. There are two clustering technics. The *divisive (top-down) technic* starts with a whole dataset as one cluster and proceeds to divide it into subsequent smaller clusters. The objective is to maximize the distance between separated clusters. On the contrary, the *agglomerative (bottom-up) technic* begins with each data point as a cluster and merges them into subsequent larger clusters. The objective is to minimize the distance between merged clusters. Hierarchical clustering methods become less accurate when the sizes of clusters are quite different. They are sensitive to outliers. When an enormous amount of data need to be analyzed, the hierarchical clustering process can be time-consuming. Therefore, it is not considered as a commercially viable solution.

(2) Two common **partitional clustering** technics are the *k-means and the k-medoids clustering algorithm*. They follow the same clustering process, which starts by randomly selecting $k$ data points as cluster centroids, assigns the remaining data points to the closest centroid, selects a new centroid for each cluster, re-allocates the remaining data points, and alternates these process until cluster centroids do not change too much (i.e. convergence). Their difference is the definition of the cluster centroid. In the k-means algorithm, the centroid is the mean of all the data points in the cluster, which can be any value in a continuous space. Hence, it can handle numeric variables. In the k-medoids algorithm, the centroid is the median of all the data points, which can only be one of the data points in the cluster. Hence, it is used to analyze category variables. The k-means and the k-medoids method have certain drawbacks. They are less effective when dealing with overlapping clusters. They use a hard clustering approach which assigns
each data point to one and only one cluster. They are sensitive to outliers.

In view of the above-mentioned drawbacks, we choose the Expectation-Maximization (EM) algorithm for Gaussian Mixture Models (GMM) (Bilmes, 1998). A GMM is the linear combination of $K$ Gaussian distributions. Each Gaussian distribution is one of the GMM’s components. The probability density function of a GMM is:

$$p(x) = \sum_{k=1}^{K} p(k)p(x|k) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \Sigma_k)$$

where $p(x|k)$ are the components ($1 \leq k \leq K$); $\mu_k$ is the mean of the $k$th component; $\Sigma_k$ is its variance; $\pi_k$ is its mixing coefficient, $\pi_k = \frac{1}{N} \sum_{n=1}^{N} z_{nk}$; $\mathcal{N}(x|\mu_k, \Sigma_k)$ is the Gaussian distribution, $0 \leq \pi_k \leq 1$; $N$ is the size of the sample dataset.

The process to build GMM for a sample dataset can be translated into a parameter estimation problem, which seeks to use the known data points to estimate the parameter $\mu_k$, $\Sigma_k$, and $\pi_k$. We want to find the maximum likelihood estimate for these parameters. That is, a set of $\mu_k$, $\Sigma_k$, and $\pi_k$ that maximize the possibility for the GMM to generate the given sample dataset. The log likelihood function of the GMM is defined as follows:

$$\ln p(X) = \sum_{i=1}^{N} \ln \left( \sum_{k=1}^{K} \pi_k \mathcal{N}(x_i|\mu_k, \Sigma_k) \right)$$

To solve this equation, it is necessary to introduce a $K$-dimensional latent Boolean variable $z \in \{0, 1\}$ for each data point $x_i$. Let $\sum_{k=1}^{K} z_k = 1$. Therefore, if the GMM assigns a data point $x_i$ (i.e. a behavior sub-stream) to the $j$th component, only $z_j$ is equal to 1 and the value of other $z$’s are 0. If $z$ is known, the dataset becomes $\{X, Z\}$ (i.e. the complete dataset). Therefore, the log likelihood function becomes:

$$\ln p(X, Z|\mu, \Sigma, \pi) = \sum_{i=1}^{N} \sum_{k=1}^{K} z_{nk} \{\ln \pi_k + \ln \mathcal{N}(x_i|\mu_k, \Sigma_k)\}$$
[The expectation step]:

However, the latent Boolean variable $z$ is the intention label of a behavior sub-stream, which is not known to us. Therefore, we replace $z$ with its expectation so that the log likelihood function becomes:

$$E_z \{ \ln p(X, Z | \mu, \Sigma, \pi) \} = \sum_{i=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \{ \ln \pi_k + \ln \mathcal{N}(x_i | \mu_k, \Sigma_k) \}$$

where $\gamma(z_{nk})$ is the expectation of $z$:

$$\gamma(z_{nk}) = \frac{\pi_k \mathcal{N}(X | \mu_k, \Sigma_k)}{\sum_{j=1}^{K} \pi_j \mathcal{N}(X | \mu_j, \Sigma_j)}$$

[The maximization step]:

We differentiate with respect to $\mu$

$$\mu_k = \frac{\sum_{n=1}^{N} \gamma(z_{nk}) x_n}{\sum_{n=1}^{N} \gamma(z_{nk})}$$

We then differentiate with respect to $\Sigma$

$$\Sigma_k = \frac{\sum_{n=1}^{N} \gamma(z_{nk}) (x_n - \mu_k) (x_n - \mu_k)^T}{\sum_{n=1}^{N} \gamma(z_{nk})}$$

And $\pi_k = \frac{1}{N} \sum_{n=1}^{N} \gamma(z_{nk})$

We alternate the expectation step and the maximization step until model parameters do not change too much (i.e. convergence). Accordingly, we obtain the components’ (i.e. intention clusters) mean, variance, and the mixing coefficient, which can help us to determine the intention label of each data point (sub-stream). The EM algorithm for GMM is a soft clustering method that indicates the strength of association between a data point and a cluster. Although its convergence process may take longer time than
the k-means algorithm, it is considered a better solution, which can handle category
data, outliers, and different cluster sizes.

- **Selecting a classification algorithm**

In Annex 2.1, we introduced several supervised machine learning algorithms capable
of classifying unlabeled data points based on the labeled data points. We propose to
use these algorithms to create different intention prediction models and choose the
best-performing for the CAMPS. Since the intention prediction model is used to pro-
cess consumer data in real time, the selection criteria should consider not only include
the prediction, but also the processing time. We propose to consider different algo-

\[
S_i = \beta \times \frac{N_i}{N_a} + (1 - \beta) \times e^{-\frac{T_i}{T_m}}
\]

(Equation 5.6)

where \(N_i\) is the number of cases correctly predicted by model \(i\), \(N_a\) denotes the total
number of cases predicted by model \(i\), \(T_i\) denotes the processing time of model \(i\), and
\(T_m\) denotes the longest processing time of all models. And \(\beta\) denotes the importance of
prediction accuracy against processing time.

### 5.2.2 The budget estimation protocol

A shopping budget is the total sum of money consumers allocate to a shopping task or
a specific item. It is an important constraint that affects consumers' decision making.
Nevertheless, unless consumers explicitly communicate their budget, such infor-

\textit{Such effort}
not only enables the CAMPS to obtain more information about consumers’ price sensitivity but also help clarify their intentions. For example, if the prices of the explored items are consistently similar, it may indicate that a consumer’s shopping intention is becoming specific. Otherwise, it may suggest that a consumer is still refining and elaborating his (or her) shopping idea. In view of the importance of consumer budget, we propose the following algorithm capable of predicting this information when it is not specified by consumers.

The algorithm for consumer budget estimation

---

- $budget_{m,k}$: budget of consumer $m$ in the $k$th sub-stream
- $L_{m,k}$ & $H_{m,k}$: lower and upper limit of the budget range specified by consumer $m$
- $n_{m,k}$: number of UPIDs in the $k$th sub-stream of consumer $m$
- $SMIT_{m,k}$: similarity of items explored in the $k$th sub-stream by consumer $m$
- $\text{threshold} = \lambda$: similarity threshold ($\lambda$ is defined by the marketer)
- $\theta_{\text{min}Tn}, \theta_{\text{max}Tn}$: minimum and maximum price of the $n$ best-selling UPID in the UCID
- $p_{x,m,k}$: price of the $x$th item explored by consumer $m$ in the $k$th sub-stream

Let $\text{threshold} = \lambda$ and other variables = 0.  % Initialization

if consumer $m$ specifies the budget

$budget_{m,k} = [L_{m,k}, H_{m,k}]$.  % use consumer provided budget

else if $n_{m,k} = 0$ % no item was explored during the sub-stream

$budget_{m,k} = budget_{m,k-1}$

decrease

else if $n_{m,k} = 1$ % only one item was explored

$budget_{m,k} = [\theta_{\text{min}Tn}, \theta_{\text{max}Tn}]$  % refer to the top selling items (UCID) in the same segment
else if \( n_{m,k} = 1 \) and \( \text{SMIT} < \lambda \)
% consumer explored several similar items

\[
\text{budget}_{m,k} = \left[ \min(p_{1,m,k}, p_{2,m,k}, \ldots, p_{x,m,k}), \max(p_{1,m,k}, p_{2,m,k}, \ldots, p_{x,m,k}) \right]
\]
% use the implicit price preference

else budget\(_{m,k} = \left[ \theta_{\text{min}T_n}, \theta_{\text{max}T_n} \right]
% refer to the top five best selling items in the segment

(UCID) that is recently explored

end

end

end

end

end

end

========================================================================

In order to implement the budget estimation algorithm, the CAMPS must refer to a retailer’s sales management system which provides the price of the top selling items of every segment (i.e. the lower level tiers that contain items). This index should be regularly updated based on the recent sales performance (e.g. last eight weeks) specified by the marketers so as to ensure that its recency. The estimated budget range is flexible, meaning that it changes with the shift of consumers’ intention and preference dynamically. Meanwhile, the budget range is a soft constraint, meaning that it is the first constraint to be relaxed if qualified personalization contents are not sufficient.

5.3 The preference management module

The performance management module is responsible for managing the consumer-item, item similarity, and consumer similarity relationships. The consumer-item rela-
tionship is also known as the preference. The item similarity and the consumer similarity relationship determine whether a pair or a group of items (or consumers) are similar to each other in terms of their features. In their process to generate personalized marketing contents, the preference and similarity data stored in the performance management module are frequently accessed by the personalization algorithms.

5.3.1 The feature extraction process

In order to compare consumers and items, the CAMPS must define the features to be used for the comparison. In the following paragraphs, we discuss respectively the feature extraction processes for consumer and item.

5.3.1.1 Item feature extraction.

Most retailers encode their items before putting them on the shelves. The encoding process aims at assign structured and unstructured data to the new items so that they can be located, retrieved, aggregated, and analyzed. The structured data is composed of the UPID, the taxonomy information, the price, and the associated keywords (tags). The unstructured data, which must be processed before being used, include the item description, the picture, and other non-textual data. The CAMPS extracts two kinds of features from these data.

1. The item’s taxonomy information is to be converted to the categorical distance between the item to the root node of the retailer’s catalog based on the method presented in Figure 5.2.

2. The TF*IDF score [Equation 5.7] of each keyword is to be computed using the method described by Pazzani and Billsus (Pazzani and Billsus, 2007) and stored in the
item characteristic vector that depicts the features of an item.

\[
w(t, i, I) = \frac{f_{t,i}}{\max\{f_{t', i}, t' \in I\}} \times \log \frac{N_i}{N_{t,i} + 1}
\]  

(Equation 5.7)

where \( f_{t,i} \) is the frequency of keyword \( t \) in item \( i \); \( N_i \) is the number of items; and \( N_{t,i} \) denotes the number of items containing keyword \( t \).

The feature extraction process for items

==================================================================

input: item data

output: item features

step 1: get the UPID of the item

step 2: compute the distance between the item to the root node = \( \Sigma (1/n) \), \( n = \) the tier where the item is located

step 3: get the keywords of the item

step 4: compute the TF*IDF score for each keyword respectively

step 5: store the TF*IDF scores into a vector

==================================================================

5.3.1.2 Consumer feature extraction.

The preference management module collects two kinds of consumer data.

(i) The consumer preference data. In the personalization research, most models rely on users’ rating data to infer their preference and profile. However, compared to the large amount of purchase transactions, the amount of consumer review and rating is negligible in the cross-channel retailing context. Therefore, we propose to use transactional data, which are composed of items explored and purchased by a consumer. Compared
to the rating data, transactional items are more numerous in quantity and they indicate
the implicit choice of a consumer (Koren, 2008). We use two coefficients to adjust the
preference strength of different items. We specify that purchase items are more im-
portant than explored items and recent transactional data are more important.

(2) The consumer profile data. In addition to the preference data, consumers can also
be clustered by their profile features. Whereas each retailer has its own modality to
manage consumer profiles, the recency, frequency, monetary, and demographic data
are commonly used to create profile labels. The profile data can be used to find neigh-
bors for a consumer when his (or her) preference information is vague not available.

The feature extraction process for consumers

input: consumer data

output: consumer features

step 1: get the transactional data (UPID)

step 2: get the keywords of each item and compute their TF*IDF scores respectively

step 3: aggregate the TF*IDF score of the same keyword based on the recency of the interaction and the strength of
the interaction

step 4: store the TF*IDF scores into a vector

step 5: get the customer relationship management (CRM) data

step 6: create group labels for each consumer based on recency, frequency, monetary, and demographic information

step 7: store the labels to the consumer-label Boolean matrix

=================================================================================================
5.3.2 The similarity matrix

The CAMPS use the following similarity matrices to find a community of similar items (or consumers).

(i) The item distance matrix measures the taxonomical similarity of pairwise items based on their distance defined by the Equation 5.4. A smaller distance value indicates that the two items belong to similar categories while a greater distance value suggests that they belong to two categories very different in nature. The CAMPS use this matrix when consumers’ preference and intention is specific.

(ii) The item-feature matrix depicts the soft features (i.e. secondary, non-taxonomical labels) of each item. Accordingly, we can also measure the item similarity relationship based on the Pearson correlation coefficient [Equation 5.8] of their respective characteristic vectors. \(|\rho|=1\) means the two items are perfectly correlated and \(|\rho|=0\) means that there is no linear correlation at all.

\[
\rho = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}} \quad \text{(Equation 5.8)}
\]

(iii) The consumer similarity matrix measures the characteristic similarity of two consumers based on the cosine similarity coefficient [Equation 5.9] of their respective vector in the consumer-label Boolean matrix. \(|\cos \theta|=1\) means that two consumers are identical in terms of the labels and \(|\cos \theta|=0\) means that they are not similar at all.

\[
\cos (\theta) = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2} \sqrt{\sum y_i^2}} \quad \text{(Equation 5.9)}
\]
5.4 The production management module

The production management module is the place where information becomes actions. The module is responsible for the creation and delivery of personalized marketing contents. The former function necessitates the personalization algorithms as well as their associated consumer data. The latter function relies on the delivery measures (i.e. the moment and the medium to present personalized marketing contents) and the prioritization algorithms.

5.4.1 The personalization algorithms

The CAMPS uses a modularized approach to manage its personalization algorithms and switch between them based on the personalization strategy. Each algorithm is connected to one or a few consumer databases containing preference, behavioral, and demographic data. When an algorithm is activated, it is allowed to access specific consumer database(s) and computation resources so as to generate personalized marketing contents. The modularized approach allows firms and marketers to mix their existing personalization algorithms with newly acquired ones as long as they fulfill diversified consumer needs using different methods. Even if many algorithms are embedded, the CAMPS only use the high-performing algorithms (to be defined in Section 5.5.2) for each intention cluster. In our research project, COMPANY X chooses to preserve its existing four personalization algorithms for cost-saving purposes. These algorithms are briefly introduced as follows.

- “Potayto-Potahto” algorithm

“Potayto” and “potahto” are the different ways to pronounce the word “potato” in British and American English. Just like its name, the “Potayto-Potahto” algorithm (Pa-Pa)
seeks to identify the items or contents that are nearly identical to the active item (i.e. the item that is currently being explored by a consumer). The content-based personalization process is composed of four steps.

1. The algorithm constructs a profile for each item explored in the sub-stream. An item profile is a feature vector depicting the weight of each feature $w_f^{x} = (w_{f_1}^{x}, w_{f_2}^{x}, \ldots, w_{f_m}^{x})$ computed based on the text frequency (TF) method, where $w_{f_m}^{x}$ is the weight of the $m$th feature of item $x$; $tf(f_m, x)$ is the frequency of feature $f_m$ in item $x$; $\sum_{i=1}^{m} tf(f_i, x)$ is the frequency of all features in item $x$. Since item features do not change frequently, the item profile can be prepared offline and updated from time to time.

$$w_{f_m}^{x} = \frac{tf(f_m, x)}{\sum_{i=1}^{m} tf(f_i, x)} \quad \text{(Equation 5.10)}$$

2. Then, the algorithm computes the pairwise similarity score of all items using Pearson correlation coefficient [Equation 6.8]. Such task can also be completed offline.

3. Based on the similarity scores and a threshold value, the algorithm prepares a list of similar items as candidate.

4. The taxonomical distance of each candidate is reviewed and items which do not belong to the same product segments (i.e. UCID) as the last active item are removed. The remaining items are submitted to the budget range filter before they can be delivered to consumers.

- “Potato-Tomato” algorithm.

“Potato” and “Tomato” are two types of vegetables. However, these two words have similar phonological features and exactly the same vowels. Just like its name, the content-based “Potato-Tomato” algorithm (Pa-Ta) aims at looking for items that can fulfill two
characteristics: (1) the candidate and the active item should have some common “soft” features (i.e. labels that depicts the secondary or less important features). (2) they should belong to the same product segment. The personalization process is composed of four steps.

(1) The algorithm constructs a “consumer profile” (i.e. a preference vector \( \overrightarrow{w_u} \) that indicates a consumer’s likes and dislikes) based on the soft features of items that were explored by him (or her) in the sub-stream, where \( I_L (or I_{NL}) \) is a set of items liked (or disliked) by a consumer that is known to the system; \( \overrightarrow{w_j} \) (or \( \overrightarrow{w_k} \)) is the vector representing the soft feature of the \( jth \) (or \( kth \)) item in \( I_L \) (or \( I_{NL} \)); \( \beta \) and \( \gamma \) are the coefficients indicating the weight of liked and disliked items in the preference vector \( \overrightarrow{w_u} \).

\[
\overrightarrow{w_u} = \beta \cdot \frac{1}{|I_L|} \sum_{j \in I_L} \overrightarrow{w_j} - \gamma \cdot \frac{1}{|I_{NL}|} \sum_{k \in I_{NL}} \overrightarrow{w_k} \text{ (Equation 5.11)}
\]

(2) The algorithm identifies a group of items based on the item distance matrix (presented in Section 6.3.2) that measures the taxonomical distance between the active item and the other items. Items whose taxonomical distance is greater than a threshold \( \omega \) are rejected. The threshold \( \omega \) needs to be specified with reference to the structure of the product catalog in order to prevent items of irrelevant segments from entering the next step while preserving the diversity of items and information contained in the active segment.

(3) The algorithm computes the Pearson correlation coefficient for the preference vector \( \overrightarrow{w_u} \) and each candidate item in terms of their soft features. The items are ranked according to their similarity to the preference vector.

(4) The similar items (i.e. items that are closest to the preference vector) are verified by the budget range filter before they can be delivered to consumers.
“Potato-Photo” algorithm

From a semantic perspective, “potato” and “photo” are very different. However, they have certain characteristics in common that maintain some weak, indirect, or latent connection between the two words. For instance, they both begin with letter p and end with letter o. Just like its name, the “Potato-Photo” algorithm (Pa-Ph) uses an item-based collaborative filtering method to find items and contents which have indirect and latent connections with consumers’ needs. The personalization process is composed of five steps.

(1) The algorithm prepares a consumer-item matrix evaluating consumers’ preference for items based on their implicit feedbacks (i.e. items they explored and purchased) in the past N sessions. The preference value of a consumer u for an item i is defined as $r_{u,i}^N$, where t is the inversed session number (i.e. $t_{last} = 1$ and $t_{first} = N$); $p_t^i$ ($q_t^i$) is the number of times item i was purchased (viewed) during the session t; $\theta_1$ and $\theta_2$ are the weight of the purchase and the exploration behavior; $(1 + \frac{1}{e})^{\frac{1}{t}}$ is the discount factor that simulates the time-based forgetting effect; $s_t^i$ is the item qualifier which determines if the item has been explored long enough to be included in the computation. $r_{u,i}^N = 0$ means that a consumer had never purchased or explored the item. The use of historical and current data can help alleviate the data sparsity problem.

$$r_{u,i}^N = \sum_{t=1}^{\min(N,t_{last})} \frac{(p_t^i \theta_1 + q_t^i \theta_2)}{(1 + \frac{1}{e})^{\frac{1}{t}}}$$

$\begin{cases} 0 & \text{active page view time } < t_{pv} \\ 1 & \text{active page view time } \geq t_{pv} \end{cases}$

(Equation 5.12)

(2) The algorithm identifies n consumers who have positive preference value for both items i and item j. After that, it creates two vectors $\vec{v}_i^* = (r_{u_1,i}^N, r_{u_2,i}^N, ..., r_{u_n,i}^N)$ and $\vec{v}_j^* = (r_{u_1,j}^N, r_{u_2,j}^N, ..., r_{u_n,j}^N)$ which contain the preference value for the two item respectively. Next, it computes the Pearson correlation coefficient (i.e. similarity) of the two items.
using their preference value vectors [Equation 6.13]. The same process is performed for all the item pairs in the item matrix.

\[
\rho_{i,j} = \frac{\sum_{u=1}^{n}(r^N_{u,i} - \bar{r}_u)(r^N_{u,j} - \bar{r}_u)}{\sqrt{\sum_{u=1}^{n}(r^N_{u,i} - \bar{r}_u)^2} \sqrt{\sum_{u=1}^{n}(r^N_{u,j} - \bar{r}_u)^2}} \quad \text{(Equation 5.13)}
\]

(3) To predict a consumer \(u\)'s preference for a new item \(s\) that he (or she) never purchased or explored before, the algorithm uses the collaborative filtering based item similarity matrix prepared in the previous step to identify a group of similar neighbor items \(D = (i_1, i_2, ..., i_w)\) of item \(s\) based on a Pearson correlation coefficient threshold. Then, it uses consumer \(u\)'s preference value \(r^N_{u,i}\) for these items and their weight (i.e. the similarity coefficient) to compute the weighted average preference value, which is the predicted preference value for item \(s\).

\[
r^N_{u,s} = \frac{\sum_{w \in D_{similar}} \rho_{s,w} \cdot r^N_{u,w}}{\sum_{w \in D_{similar}} |\rho_{s,w}|} \quad \text{(Equation 5.14)}
\]

(4) The algorithm predicts the preference value for all the new items for consumer \(u\) and rank them in regards to their predicted preference values. The items which have a high preference value (i.e. \(r^N_{u,s} > \delta\)) are selected as candidates.

(5) The algorithm then measures the taxonomical distance of these items to the active item explored by the consumer so as to remove all the items that are in the same product category as the last active item. The remaining items are verified by the budget range filter before they can be delivered to consumers.

- “Chewing Gum” algorithm

In the supermarkets, chewing gums are usually placed next to the cashier. They are popular items with low price, which are capable of provoking impulsive purchase. Just
like its name, the “Chewing Gum” algorithm (Ch-Gm) proposes low price items that are frequently purchased by other consumers who have the same preference as the active consumer. The personalization process is composed of four steps.

(i) The “Chewing Gum” algorithm analyzes the item(s) in the active consumer’s shopping cart and identify other shopping carts which also contain the same item(s).

(ii) It ranks the items in these shopping carts with reference to their appearance frequency and select the top-N most popular items.

(iii) It identifies the cheapest item in this top-N list and propose it to a consumer.

5.4.2 The delivery measures

After personalized contents are created, the CAMPS needs to choose a delivery measure to present them to consumers. A delivery measure is defined as a mechanism that specifies the medium and the moment to deliver personalized marketing contents. The CAMPS is equipped with three delivery measures.

- **Recommendations**

*Recommendations* are personalized marketing contents presented to consumers when they visit a retailer’s physical store, e-commerce website, or m-commerce App. A recommendation has three components. The content component provides textual and image information about the recommended item. The interactive component enables consumers to explore the recommended item in detail. The communication component explains how the recommendation is produced and allows consumers to inform the CAMPS if it makes a mistake. Once a recommendation is generated, it is examined by the prioritization algorithm before being presented to a consumer immediately. In
the majority of cases, there is no time lag between the production and delivery of recommendations. Since personalized contents are displayed on a retailer’s own e-commerce website and m-commerce App, the delivery cost of a recommendation is contingent with the operating cost of the CAMPS.

- **Remarketing ads**

*Remarketing ads* are marketing contents that a retailer (i.e. the advertiser) displays on other websites or mobile Apps (i.e. the publisher). Advertisers must pay for every click on the displayed remarketing ads. The price of each click is determined in the auction organized by ads agencies or an ads exchanges, where higher price guarantees better display positions and platforms. The content to be displayed on the remarketing ad banner is determined by the advertisers. Instead of displaying items that were antecedently explored by a consumer or presenting non-personalized marketing contents, a retailer may deliver to consumers the personalized marketing contents created by the CAMPS, after they are examined by the prioritization algorithm. In this case, when consumers explore the publisher’s website or mobile App, they see remarketing ads related to their previous shopping journeys with the advertiser. The display cost of each remarketing ad is contingent with the cost of clicked remarketing ads as well as the operating cost of the CAMPS.

- **Emails**

*Emails* are one of the most common means of communication between a retailer and consumers. Compared to recommendations and remarketing ads, emails enable firms and marketers to provide more information and functions in their marketing messages. The CAMPS can use emails to deliver personalized marketing contents to consumers after their cross-channel shopping journey. There is a lag between the production and
the delivery of marketing contents. Personalized marketing emails relevant to recent unfulfilled consumer needs help to bring cross-channel shoppers back to a retailer’s sales network and accelerate their decision-making process. The display cost of each email is contingent with the operating cost of the email system and the CAMPS.

Consumers’ preferences for personalization algorithms and delivery measures are dependent on their instantaneous intention. To select the appropriate algorithm and measure for each intention cluster, the CAMPS refers to the historical performance (i.e. purchase rate) of different personalization strategies (i.e. personalization algorithm and delivery measure) in different intention contexts. High-performing strategies may have more chances to perform the personalization task in their competent intention contexts. The process is detailed in Section 5.5.

5.4.3 The prioritization algorithm

To prevent consumer confusion and information overload, personalized marketing contents must be examined by the prioritization algorithm before being delivered. The prioritization algorithm seeks to identify inappropriate marketing contents and determine the order of content delivery. The examination process is explained as follows.

At the end of each sub-stream, the prioritization algorithm examines the nature (i.e. recommendation, remarketing ad, email) of the delivery measure and take corresponding actions to handle different situations.

Recommendations are immediately pushed to a consumer’s active device (i.e. a desktop, a laptop, or a mobile device a consumer is currently using to explore the retailing channel) after the stock availability check with an objective to avoid proposing unavailable items to consumers. Recommendations on the mobile devices are displayed as pop-up
message boxes at the bottom of the screen while recommendations on the laptops and
desktops are displayed on the right-lower corner of the e-commerce website.

A remarketing ad is examined by the prioritization algorithm to determine if the asso-
ciated UPID belongs to the segments whose items were purchased by the same con-
sumer during the past $X_{rm}$ days. If the answer is positive, the remarketing ad is canceled.
Otherwise, the remarketing ad is encrypted as an “instruction package” and submitted
to the data management platform (DMP) of an online ads agency or exchange. Each
instruction is composed of a universal consumer identification that the DMP uses to
track consumers across different devices and web domains, the identification of the re-
tailer, and the UPID of the personalized content. If the retailer has already provided an
instruction for the consumer in the DMP, the current instruction is replaced by the new
one to ensure that personalized remarketing ads are always relevant to a consumer’s
latest unfulfilled shopping needs.

Personalized marketing emails are temporarily stored by the retailer’s email system,
which can store up to 8 emails for each consumer. Each email is given a deliverable date,
which indicates the earliest time the email can be delivered. The prioritization algo-
rum examines all the deliverable emails on a daily basis. First, it determines if a mar-
keting email has been sent to a consumer in the past $X_{ms1}$ hours. If yes, it is not take any
further actions. Otherwise, it deletes emailing containing items that are taxonomically
close to the items purchased by a consumer during the past $X_{ms2}$ days (i.e. taxonomical
distance less than the threshold). Next, it ranks the remaining emails according to their
similarity with the last item explored but not purchased by the consumer. The email
containing the most similar item is sent immediately and the remaining emails are kept
by the email system for the same process the next day. The objective of the examination
and prioritization process is to avoid spamming consumers.

Prioritization is the last step before personalized marketing contents of the CAMPS are delivered to consumers. However, it is not the end of the personalization task. The CAMPS needs to analyze consumers’ feedback information so that it can improve its personalization strategies. The performance improvement process is discussed in the next Section.

5.5 The performance improvement module

The performance improvement module is the place where information becomes insights. The module is responsible for two tasks. (1) It evaluates the performance of the CAMPS based on a series of metrics to be defined in the Section 5.5.1. (2) With the purpose to discover effective personalization model for each intention status, it tracks and analyzes consumers’ reactions to marketing contents proposed by different personalization models (i.e. composition of different algorithms, consumer data, moments of delivery, and communication media) in each intention context. If certain personalization models consistently outperform others in a particular intention context, they are used more frequently by the CAMPS in similar personalization tasks so as to improve the consumer conversion rate.

5.5.1 Evaluating the performance of the CAMPS

In Section 2.7, we made a review of the existing performance evaluation metrics and methods and their limitations. In view of these limitations, we propose the following metrics to assess the performance of the CAMPS (Table 5.1, next page). These metrics take into account various components of a marketing personalization system and their
impact on consumers’ decision.

Instead of measuring just one aspect of the personalization system (e.g. accuracy, coverage) at a time, they provide an integrated view of the system’s capability in improving conversion and enhancing consumer satisfaction. Meanwhile, the metrics is also capable of informing firms and marketers of the best-performing personalization strategy (i.e. composition of algorithms, media, and moments of delivery), due to the rich information collected by the CAMPS based on the encoding protocol.

In our project, the purchase rate \( r_{pur} \) is used as the CAMPS’s key performance indicator (KPI) because COMPANY X regards revenue generation as the priority of the marketing personalization system. However, firms and marketers may also consider to use other metrics if their objective is to foster consumer impression (click-through rate, \( r_{ct} \)) or to enhance consumer shopping experience (duration of interaction, \( T_{int} \)).

**Table 5.1. KPI of the CAMPS**

<table>
<thead>
<tr>
<th>category</th>
<th>KPI</th>
<th>definition</th>
<th>calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversion</td>
<td>purchase rate ( r_{pur} )%</td>
<td>proportion of personalized marketing contents purchased by consumers. ( f_{pur} ): the amount of purchased personalized contents.</td>
<td>( r_{pur} = \frac{f_{pur}}{f_{ds}} )</td>
</tr>
<tr>
<td>satisfaction</td>
<td>click/tap delay</td>
<td>The amount of time between the display and the click/tap of the proposed item. ( T_{e_{ct}} ): timestamp of the click/tap event; ( T_{e_{ds}} ): timestamp of the display event.</td>
<td>( T_{ci} = T_{e_{ct}} - T_{e_{ds}} )</td>
</tr>
</tbody>
</table>
The amount of time a consumer spent with the item or service proposed by the CAMPS in a session.

\[ T_{\text{int}} = \sum_{i=1}^{n} (T_{\text{EX}} - T_{\text{C}i}) \]

Due to the complexity of the cross-channel shopping context (i.e. multiple display media and shopping platforms), some rules and protocols need to be defined for the computation of the KPI. These rules and protocols are retailer-specific. In our case, we discuss with the marketing executives of COMPANY X to determine these rules and protocols together.

- **Click-through delay**

COMPANY X uses on-site recommendations, remarketing ads banners, and marketing emails to deliver personalized contents. The rules to validate a successful click-through are different for these media.

1. On-site recommendations displayed on website or App must be clicked/tapped within 30 minutes (or the same session, whichever comes the first) to be validated as a successful click-through. SSID, UUID, UPID, and recommendation identification (RMID) are used to track the consumer behaviors.

2. Remarketing ads banners must be clicked in order to be validated as a successful click-through. There is no need to impose a time constraint because remarketing ads banners are not recoverable once skipped. The information contained in the ads banner’s uniform resource locator (URL) can be used to ensure the traceability of consumers’ behaviors.
(3) Marketing emails must be clicked (i.e. to redirect recipients to COMPANY X's e-commerce website or App) within two days upon reception so as to be validated as a successful click-through.

(4) The CAMPS records a consumer's instantaneous intention (when the content is delivered and when the content is clicked) and the platform where the content is clicked. Meanwhile, the system also records the active algorithm, the consumer data accessed by the algorithm, the medium to deliver the content, and the platform where the personalized content is received in order to make further analysis.

*Purchase rate*

The rules and protocols to validate a purchase of personalized contents is different from the click-through.

(1) If a consumer purchases an item (i.e. UPID), the system trace back to determine if the same item has been proposed to him (or her) within the last 14 days as a personalized marketing content. The UUID, UPID, and the associated timestamps are used to ensure the traceability of the events.

(2) The system validates a personalized content's contribution to sales only if the personalized content is clicked or tapped (i.e. interaction) before being purchased.

(3) If the consumer was exposed to several personalization interactions associated with the same UPID during the period, the purchase is attributed to the last interaction\(^1\).

---

\(^1\) Depending on the requirement of firms and marketers, one may also consider to attribute the purchase to the first interaction or allocate the contribution proportionally to different interactions. We choose the last interaction because we think that it makes the decisive contribution to the final purchase decision.
The system records a consumer’s instantaneous intention (when the content is delivered) and the platform where the content is purchased in order to make further analysis.

Based on the rich information associated with each personalized marketing content delivered to consumers (i.e. instantaneous consumer intention, personalization algorithm, display moment, communication medium, and interaction platform), the CAMPS is able to analyze the performance of different components in a personalization model and determine the best personalization strategy that is proven to be effective for consumers who have a particular shopping intention. For instance, the CAMPS may compute the intention-specific click-through rate for different personalization algorithms. Such data can help the CAMPS to identify the most effective algorithm(s) that are proven to be effective in yielding high click-through rate. Similarly, the CAMPS may further determine the click-through rate by delivery moment and by communication medium so as to identify the appropriate moment and medium to deliver the personalized marketing contents to consumers.

5.5.2 Improving the CAMPS’s performance

The CAMPS’s production management module uses four personalization algorithms (presented in Section 5.4.1) and three delivery measures (presented in Section 5.4.2) to generate and deliver personalized marketing contents to consumers based on their instantaneous shopping intention. The performance improvement process is presented as follows.

Step 1. Choose the objective to be achieved by the CAMPS from two options: (1) maximize consumer purchase rate of the personalized contents; (2) maximize the net profit of
the personalized contents in an intention cluster. This objective shall be chosen by the marketers.

**Step 2.** Provide personalized marketing contents to each intention cluster (i.e. consumers of the same intention) based on the undifferentiated personalization strategy (i.e. using all possible strategies [i.e. 4 algorithms × 3 delivery measures] alternately and equally) for a period of time and record the purchase information of them. The period can range from several weeks to months in order for the CAMPS to gather enough consumer data for analysis, depending on the traffic of the retailing platforms.

**Step 3.1.** Compute the strategy-specific purchase rate $r^k_{pur}$ for each intention cluster. If the conversion objective is selected, record the personalization strategies whose purchase rate is greater than a threshold $\phi$ defined by the marketers for each intention cluster.

**Step 3.2.** In our case, COMPANY X selects the net profit objective. Therefore, it is necessary to find the personalization strategies that can yield greatest expected net profit for each intention cluster. First, compute the unit net profit ($\pi_k$) of each personalization strategy $k$ using the following equation, where $P$ is the expected unit gross profit for a personalized content, which can be provided by the retailer; $C_k$ is the cost to deliver the content, which is contingent with the $k$th personalization strategy; $r^k_{pur}$ is the purchase rate of the $k$th personalization strategy.

\[ \pi_k = P - C_k \]

---

2 For instance, if the CAMPS uses a retailer’s own recommender systems or emailing systems to deliver personalized marketing contents, the per delivery cost is the amortized operating cost (e.g. equipment, personnel, utilities) of these systems. If the CAMPS uses remarketing channels, it will be charged only when the content is clicked by a consumer (i.e. pay-per-click). To simplify the computation of unit cost, the remarketing expenses can also be amortized to all the remarketing ads displays, no matter they are clicked by consumers or not.
Then, compute the expected net profit of the $k$th personalization algorithm for a particular intention cluster. Since firms expect positive net profit, $r_{pur}^k$ must be greater than $c_k/p$ (i.e. the threshold). Accordingly, record all the personalization strategies whose purchase rate is greater than the threshold.

$$
\Pi_k = \pi_k \cdot r_{pur}^k = \left(p \cdot r_{pur}^k - c_k\right) \cdot r_{pur}^k
$$

**Step 4.** Prepare a personalization strategy performance database which records all the personalization strategies, their purchase rates, their net profit values, and their probabilities of being chosen, which is proportional to the expected net profit value it might generate.

$$
prob_{\pi_k} = \frac{\pi_k}{\sum_{k=1}^{m} \pi_k}
$$

**Step 5.** Replace the undifferentiated personalization strategy with the new context-aware personalization strategy. Generating and delivering personalized marketing contents using the context-aware personalization strategy for a period of time and collect the purchase and click information about these contents. This period may range from one week to several weeks depending on the traffic of the retailing platform.

**Step 6.** Regularly repeat Step 3 to 5 to revamp the context-aware personalization strategy so that the CAMPS can use more frequently the most appropriate personalization strategy to provide personalized marketing contents. As a result, the purchase rate or the expected net profit objective can be constantly improved.
5.6 Summary

The generation and delivery of personalized marketing contents rely on the cooperation of different functional modules. In this chapter, we present the key modules, processes, and algorithms of the CAMPS in the order of intention recognition, preference prediction, content generation, content delivery, and performance improvement. The CAMPS predicts consumers’ intention and use the knowledge to optimize the personalization task by choosing the most appropriate personalization algorithm and delivery measure relevant to consumers’ instantaneous intention and needs.

Whereas the above generic personalization process flow (i.e. intention → algorithm → preference → delivery → feedback) paves the way for the normal operation of the CAMPS, it also allows firms and marketers to make retailer-specific modifications and customizations so as to optimize the consumer experience. For instance, marketers can use different behavioral features, personalization algorithms, and delivery measures in order to fulfill the specification and requirements of a retailer. In the next chapter, we demonstrate how the CAMPS can be applied to the cross-channel retailing context.
Chapter 6

Applying the context-aware marketing personalization approach to the retailing world

6.1 Introduction

In the previous chapters, we presented the theoretical framework, approach, methodologies, framework, and architecture of the CAMPS. Retailers may follow the steps defined in Chapter 3, 4, and 5 to develop the CAMPS from scratch if they do not have any marketing personalization system in place. The development process is called the greenfield project. In this chapter, we discuss the application of the CAMPS in the real retailing environment, where most retailers have their marketing personalization system in place (i.e. brownfield development). In this case, marketers and developers may create a CAMPS by revamping legacy systems and adding necessary modules.

Companies operating multiple retailing channels, like COMPANY X, may prefer to choose the brownfield development option due to the following reasons. (1) From a managerial perspective, keeping the existing personalization system would interfere less with the other computer systems (e.g. customer relationship management system and enterprise resource planning system) that are responsible for the normal operation of
a multi-channel retailer. Retailing firms which consider system stability as the key would prefer the less risky brownfield option. (2) From an economic perspective, recycling the existing personalization system means that less manpower and resources would be invested in the development and the personalization system would be ready sooner. If both options can attain the same benefits, the brownfield option would be more beneficial for a retailer’s business performance.

In Section 6.2, we illustrate the CAMPS we developed for COMPANY X based on their legacy marketing personalization system. We are also interested in quantifying the benefits the CAMPS can create for a retailer. In Section 6.3, we analyze the consumer feedback data gathered by the CAMPS and review its performance. The direct and indirect contribution of the CAMPS is illustrated and discussed to prove that the CAMPS is an effective marketing personalization system that can deliver satisfying results and make continuous improvements. Section 6.4 serves as the summary of the chapter.

6.2 The brownfield development of the CAMPS in COMPANY X’s infrastructure

The executives of COMPANY X, who are not satisfied with the performance of their legacy marketing personalization system in the cross-channel retailing context, decided to evaluate our CAMPS in their digital marketing system. Due to the managerial and economic reasons, COMPANY X’s team specified that the new context-aware marketing personalization system must use the existing information delivery measures and preserve the four personalization algorithms. However, we are allowed to revamp the backbone architecture of the legacy system (Figure 6.1, next page) and add new modules
so as to cope with the cross-channel shopping challenges and enhance the performance of the digital marketing system.

Figure 6.1 The Legacy Marketing Personalization System of COMPANY X

The modularized structure of the CAMPS allows us to map the contextualization modules presented in Figure 3.5 to COMPANY X’s legacy system. Figure 6.2 presents the conceptual architecture of the context-aware marketing personalization system we deployed in COMPANY X’s digital marketing infrastructure.
Except for the preserved modules, there are significant differences between the legacy system and the CAMPS.

(i) The CAMPS uses an event management module to process clickstream data for the personalization engine so that the latter can focus on the personalization task (i.e. predicting consumers’ ratings for new contents and identify the most appropriate content). Meanwhile, the event management module provides information for the intention management module, which chooses the most appropriate personalization algorithm and delivery measure for the identified intention. As a result, the workload of the personalization engine is significantly relieved because the data processing task is handled by the event management module and the engine no longer needs to use all the algorithms for one personalization task.

(ii) The CAMPS use a series of modules (i.e. budget filter and prioritization algorithm) to enhance the relevancy and consistency of personalized marketing contents presented to consumers. These modules help reduce the volume of information a consumer may receive from COMPANY X and avoid the chaotic situations mentioned in Section 0.2.1.

(iii) The CAMPS is equipped with a learning algorithm that can inform the system of the most efficient personalization strategy depending on consumers’ intention. Unlike the legacy system which follows the pre-defined and invariable personalization strategies, the CAMPS is capable of constantly creating and optimizing its own strategies based on consumers’ historical behaviors in similar situations so as to maximize the consumer conversion.
In the following paragraphs, we present the process we followed to deploy different modules of the CAMPS and measure the performance of the system.

6.2.1 Deploy the event management module

We deployed an event management server to host the encoding, consolidation, and fragmentation protocol, which are responsible for transforming consumer data collected by COMPANY X’s website server, App server, IOT server, and transaction server into behavior sub-streams (Figure 3.5). The event management server computes the key features of each sub-stream and submit them to the analytics server. Meanwhile, it extracts the UPID and UCID contained in each sub-stream and submit them to the personalization server.

The challenge of deploying the event management module is to reorganize COMPANY X’s product catalog. In theory, COMPANY X’s product catalog is composed of six tiers, where the first tier differentiates different type of sports activities (e.g. hiking, running, fishing, swimming, football, basketball etc.), the second tier specifies the type of the sports equipment (e.g. clothing, shoes, backpack, protection kits etc.), the third tier is organized according to the user of the sports equipment (e.g. men, women, children, professionals), the fourth tier denotes the brand or manufacturer of the equipment, the fifth tier specifies the product segment information, and all the products are located in the sixth tier. Hence, the taxonomical distance from an item to the root node is 2.28 (i.e. \( d = 1 + 1/2 + 1/3 + 1/4 + 1/5 \)).

However, the actual product catalog has two major issues that affect the computation result of the inter-item taxonomical distance. (1) We found that nearly 30% of the items do not have complete classification information, leading to the fact that their distance
to the root node is less than 2.28. (2) We also found that about 40% of the items belong to several categories of different tiers at the same time, which makes it hard to determine which categorical path to follow for these items.

In fact, these two issues are quite common in the retailing organizations because product classification rules are changing constantly and sometimes retailers have to create some categories and delete others to fulfill the requirement of the seasonal promotions and other marketing campaigns.

We fixed the above issues by two measures. (1) We preserve the longest categorical path of an item and remove the other ones by assuming that the longest path contains the most detailed classification information of the item. Thus, each item has only one path to the root node. (2) We create “virtual” categories at different tiers to accommodate items which are supposed to be sorted into the corresponding categories so that each item has complete classification information. The process to reorganize COMPANY X’s product catalog requires manual effort and expert knowledge. However, it is necessary to go through this tedious and time-consuming process so that we can continue with the subsequent task to assign UPID (UCID) to products (categories) and create preference matrix.

6.2.2 Deploy the preference management module

We deployed the preference management module on the transaction server to keep a record of consumers’ implicit preference based on the purchase and exploration information gathered by the event server. Each time when the preference management module receives a new batch of consumer data from the event server, it adds new UUID and UPID to the consumer-item matrix and modifying the elements within. Meanwhile,
the preference management module computes the taxonomical distances between items on a daily basis based on COMPANY X's frequently changing product catalog. The computation task is performed in the midnight so that the new distance matrix can be used for the next day's personalization tasks. Besides, the similarity coefficients for all consumer pairs are also updated on a daily basis according to consumers' shopping behaviors (i.e. purchase and exploration) so as to enable the CAMPS to identify neighbors who have similar taste as the active consumer. Finally, the module is also responsible for making offline computation of the item similarity coefficients based on the TF*IDF method.

The preparation of the taxonomical distance matrix consists three steps. (1) The module compares the new product catalog to the previous version and identifies all the new items and modified items. (2) It computes the pairwise distance of these items. (3) It computes the distance from these items to the remaining unchanged items. According to our method, there is no need to compute the distance between unchanged items, which accounts for the majority of the product catalog. Firms like COMPANY X have more than 50,000 kinds of products in their catalog and there are about 200 to 500 new and modified products every day. According to our method, they can save about 95% of the computation time.

The preparation of the consumer similarity matrix is composed of three steps. (1) The module makes a snapshot of the consumer-item matrix on the cut-off time (23:30 in our case). (2) Based on the snapshot matrix, it computes each consumer's preference value for each items using [Equation 5.12]. We let $\theta_1 = 3\theta_2$ to indicate that purchases are
stronger preference indicators than explorations. We also let $t_{pv} = 30$ seconds (i.e. minimal interaction time) so that products must be explored by a consumer long enough to be included in the preference value. (3) Once all the preference values are computed, the module computes the similarity coefficients for each consumer pairs using [Equation 5.8]. Since COMPANY X has a great volume of consumers and items in the matrix, the computation may take two to three hours. Therefore, the computation task is performed during the midnight (from 00:00 to 03:00). New coefficients are populated to the consumer similarity matrix, which are used for the next day's personalization task.

The preparation of the item similarity matrix is composed of three steps. (1) The module identifies the new items by comparing the new and the old catalog and creates a feature vector for each new item based on [Equation 6.7] as well as a group of pre-defined features. In our case, we use 50 “soft” keywords (i.e. adjectives describing the user experience and the non-taxonomical attributes of the product, such as comfortable, performance, fashionable) provided by the sports equipment expert of COMPANY X as features. (2) It computes the pairwise similarity between new items. (3) It computes the similarity between a new item and an existing item. All the similarity coefficients are populated to the item similarity matrix for the next day's personalization task. Product features do not change frequently and there are less than 50 new products every day in COMPANY X's case. Hence, the computation task can be performed offline after the update of the product catalog and it takes no more than 15 minutes to update the item similarity matrix.

---

3 By analyzing the relationship between a product purchase and the associated product exploration behaviors during the past three months, we discovered that on average products have been explored three times before they were purchased by a consumer (p=0.01).
6.2.3 Deploy the intention management module

The intention management module is responsible for analyzing sub-stream features prepared by the event management module and predicting the underlying intention of a sub-stream based on the intention prediction model. The process to develop the intention prediction model is composed of four steps (Figure 5.1).

(i) We start with a sample dataset containing the behavior features (i.e. isPurchase, ATIN, and SMIT) of 14,369,350 sub-streams gathered from 3,376,217 consumers. We standardize the features of these sub-streams and use the EM for GMM algorithm to classify them into four clusters, which is consistent with our intuitive assumption in Section 6.2.1. Although the selection of initial parameter values, namely the mean and the standard deviation of each Gaussian distribution (i.e. cluster) can be purely random and have very few impact on the final classification result, we try to use more reliable and accountable criteria for the task. To do so, we base our value selection on COMPANY X’s historical consumer conversion analysis, which indicates that nearly 2% of the consumer sessions are purchase oriented (i.e. \( p(\text{PIT}) = 2\% \)) and about 30% of the consumers do not have specific shopping objectives when they visit COMPANY X’s retailing channels (i.e. \( p(\text{NIT}) = 30\% \)). Since we do not know the composition of LIT and VIT consumers, we simply let them be the same \( [i.e. \, p(\text{LIT}) = p(\text{VIT}) = (100\%-2\%-30\%)/2 = 34\%]. \) The standard deviation of the PIT cluster is set at 0.5% and the standard deviation of the remaining clusters is set at 5%. In our project, the EM for GMM algorithm is required to make 1,000 iterations and return the clustering results.

4 Z-score is used in the data normalization step. \( Z = \frac{x - \mu}{\sigma} \), where \( x \) is the original value of an observation, \( \mu \) is the mean of the sample, \( \sigma \) is the standard deviation of the sample, and \( z \) is the standardized score of \( x \).
The result (Table 6.1, next page) suggests that all the PIT sub-streams were correctly classified into a single cluster (i.e. Cluster 4) and the feature of the four clusters (i.e. the cluster centroids) seems consistent with our intuitive assumptions. The centroid of Cluster 1 has a high ATIN value (represented by z-score) and a low SMIT value, suggesting that consumers in this cluster prefer to make in-depth interaction with similar items. This fits the learning intention defined by us in Section 6.2.1. Meanwhile, the centroid of Cluster 2 has a low ATIN value and a low SMIT value, suggesting that consumers are comparing similar items (i.e. the evaluating intention). In addition, the centroid of Cluster 3 has a high SMIT value and a low ATIN value, suggesting that consumers in this cluster prefer to explore different items briefly. Such behavior is consistent with the profile of the mining intention. Based on the classification result, we assign the corresponding intention label to each sub-stream member in a cluster. The labeled data are called the seeds.

Table 6.1 Clustering Result

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Cluster Size</th>
<th>Weight (%)</th>
<th>Z-isPurchase</th>
<th>Z-ATIN</th>
<th>Z-SMIT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>5,169,426</td>
<td>35.975%</td>
<td>-0.068</td>
<td>1.364</td>
<td>-0.805</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>5,610,641</td>
<td>39.046%</td>
<td>-0.068</td>
<td>-0.573</td>
<td>-0.355</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>3,524,041</td>
<td>24.525%</td>
<td>-0.068</td>
<td>0.167</td>
<td>1.549</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>65,243</td>
<td>0.454%</td>
<td>14.807</td>
<td>-1.106</td>
<td>-0.107</td>
</tr>
<tr>
<td>Total</td>
<td>14,369,350</td>
<td><strong>100.000%</strong></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

(2) We use the 10-fold cross validation (10-CV) method to train different classification algorithms and evaluate the classification accuracy and computation time of these models based on the 14,369,350 seeds prepared in the previous step. In COMPANY X's case, intention classification models need to handle a great amount of prediction requests simultaneously. Hence, prediction results must be produced rapidly so that the
CAMPS can perform subsequent personalization task. Meanwhile, intention classification models should be able to process non-normalized data because the standardization process is not suitable for the online situation where the minimum and maximum of a feature is unknown. In addition, the models must be able to accommodate binary feature (i.e. isPurchase) and numeric features (i.e. ATIN, SMIT) at the same time. Based on these criteria, we evaluate the k nearest neighbor (KNN), support vector machines (SVM), decision tree (DT), Naive Bayes (NB), multi-class linear discriminant analysis (mLDA), and artificial neural network (ANN) classifiers. For SVM, we test the linear kernel function and the radial basis function (RBF). Different parameters are tested to find the best model. For DT, we limit the growth of the trees and use different profit/cost parameters to adjust the impact of misclassification. For ANN, different thresholds and weights are tested to find the best learning model. For kNN, different k values are tested to find a balance between prediction accuracy and computation time.

We normalize the features for SVM and mLDA classifiers and we also discretized the features for the Naive Bayes classifier to fulfill the assumption of the algorithm (Xing et al., 2010). We create about 200 models and evaluate their performance using [Equation 5.6], which takes into consideration a model's prediction accuracy and computation time.

We present the performance of the best model created by each classification algorithm in Figure 6.3 (next page), which suggests that the classification accuracy (i.e. $\beta=1.0$) of different models are less dispersed than their computation time (i.e. $\beta=0.0$). The result is consistent with the previous research results (Caruana and Niculescu-Mizil, 2006). Although the most accurate classifier (SVM-based) can classify 91.2% of the seeds correctly, the training process takes 3,266 seconds, making it unsuitable for the online deployment. The fastest classifier takes only 421 seconds to train (NB based). However, its
prediction accuracy is not satisfactory (78.5%), mainly due to the independent sub-stream features. The result also indicates the $\beta$ dependent performance score of each classifier. It suggests that the DT-based intention classifier obtains the highest performance score when $\beta \in [0.19, 0.89]$, a range wide enough to fulfill marketers’ diversified requirement for accuracy and speed. The classifier can predict 87.1% of the seeds correctly and the training process takes only 539 seconds. Due to its sound prediction accuracy and fast training speed, the DT based classification model is chosen for the online training.

![Plot of the Performance Score of the Top-Ranking Intention Classifiers](image)

**Figure 6.3 Plot of the Performance Score of the Top-Ranking Intention Classifiers**

(3) The DT-based classifier is trained by seeds whose intention labels are intuitively assigned by the clustering algorithm. To validate and enhance its predictive power, the
model needs to be tested in the real shopping environment. We use the DT-based classifier to predict the instantaneous intention information for COMPANY X’s customers and present the prediction results to them. Customers are invited to assess the prediction result and inform us their real intention if the classifier makes a mistake. The classification model learns from these new instances and adjusts its parameters to enhance its prediction accuracy. By giving new and misclassified cases a higher weight in the learning process, the model can abandon the less indicative seeds while reserving spaces for new instances that are more significant to the model’s prediction accuracy.

Table 7.2 Confusion Matrix Before and After the Real-World Training

<table>
<thead>
<tr>
<th>Confusion matrix (offline training)</th>
<th>Confusion matrix (online training)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIT prediction</td>
<td>VIT prediction</td>
</tr>
<tr>
<td>LIT-clustering label</td>
<td>4,681,965</td>
</tr>
<tr>
<td>VIT-clustering label</td>
<td>368,931</td>
</tr>
<tr>
<td>MIT-clustering label</td>
<td>21,242</td>
</tr>
<tr>
<td>PIT-clustering label</td>
<td>0</td>
</tr>
</tbody>
</table>

WIP model accuracy & predictive model accuracy

<table>
<thead>
<tr>
<th>LIT prediction</th>
<th>VIT prediction</th>
<th>MIT prediction</th>
<th>PIT prediction</th>
<th>Weight</th>
<th>LIT prediction</th>
<th>VIT prediction</th>
<th>MIT prediction</th>
<th>PIT prediction</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIT - clustering label</td>
<td>90.16%</td>
<td>7.85%</td>
<td>1.96%</td>
<td>0.00%</td>
<td>35.98%</td>
<td>LIT - consumer feedback</td>
<td>93.20%</td>
<td>6.77%</td>
<td>0.03%</td>
</tr>
<tr>
<td>VIT - clustering label</td>
<td>6.58%</td>
<td>90.39%</td>
<td>3.03%</td>
<td>0.00%</td>
<td>39.05%</td>
<td>VIT - consumer feedback</td>
<td>3.85%</td>
<td>92.88%</td>
<td>3.32%</td>
</tr>
<tr>
<td>MIT - clustering label</td>
<td>0.60%</td>
<td>23.24%</td>
<td>77.16%</td>
<td>0.00%</td>
<td>24.52%</td>
<td>MIT - consumer feedback</td>
<td>0.05%</td>
<td>12.02%</td>
<td>87.93%</td>
</tr>
<tr>
<td>PIT - clustering label</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>100.00%</td>
<td>0.45%</td>
<td>PIT - consumer feedback</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Model accuracy | 87.1% | 81.0% |
Sample size | 14,369,350 | 14,369,350 (50% new instances) |

The confusion matrices in Table 7.2 shows the prediction accuracy of the DT-based classification model in two scenarios. The left matrix indicates that all PIT and the majority (90.2%) of LIT and VIT were successfully recognized by the model. However, 22.2% of the MIT were misclassified as VIT. Overall, the model can correctly classify 87.1% of the offline intuitive seeds. After introducing real-world consumer instances, the model’s prediction accuracy is increased across all intention clusters. The right matrix indicates that 91.8% of the cases (i.e. 50% of the indicative seeds and 50% of real-world
consumer instances) are correctly classified by the model. The improvement is mainly contributed by the enhanced ability to recognize MIT (77.2% $\rightarrow$ 87.9%).

(iv) Based on the above results, we believe that the DT-based model is competent for the intention prediction task. Accordingly, it is deployed in the intention management module using massive online analysis (Bifet et al., 2010) to analyze sub-stream features at real-time. The model constantly discards the less indicative instances and adds new instances that can maintain/enhance prediction accuracy in its training dataset so that it can adapt to the evolution of consumer behaviors.

Meanwhile, we deploy the budget estimation protocol, which analyses the price and similarity (i.e. SMIT) information of the products explored by a consumer in the sub-stream and determines a soft price range\(^5\) to filter the proposed marketing contents. The key parameter of the budget estimation protocol is the item similarity threshold $\lambda$. COMPANY X has a six-tier product catalog so that the taxonomical distance between two items ranges from 0.4 to 4.6 (refer to Figure 5.2). Based on preliminary tests, we define $\lambda = 0.9$ so that “similar items” for the budget estimator is defined as the items who have the same grandparent node.

6.2.4 Deploy the production management module

The production management module receives sub-stream intention labels produced by the intention management module and chooses the corresponding personalization algorithm and delivery measure to produce and present personalized marketing contents based on the context-aware personalization strategy. Before a delivery, marketing

\(^5\) If none of the candidate fits into the budget range, the CAMPS chooses one among the Top-5 best-ranking candidates whose price is closest to the upper or lower limit of the budget range so that at least one personalized marketing content is available.
contents are validated by the prioritization algorithm. The parameters of the personalization algorithms and delivery measures are set by COMPANY X’s develop team based on previous experiments. We present these parameters in Table 6.2 (next page).

**Table 6.2 Key Parameters of the Production Management Module**

<table>
<thead>
<tr>
<th>personalization algorithm</th>
<th>Parameter</th>
<th>Setting</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Potato-Potayo</strong></td>
<td>( p \geq 0.7 )</td>
<td>% threshold used in step (3)</td>
<td></td>
</tr>
<tr>
<td><strong>Potato-Tomato</strong></td>
<td>( \alpha = \beta = 1 )</td>
<td>% coefficient used in step (1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \omega = 1.7 )</td>
<td>% threshold used in step (2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p \geq 0.7 )</td>
<td>% threshold used in step (4)</td>
<td></td>
</tr>
<tr>
<td><strong>Potato-Photo</strong></td>
<td>( t_{pv} = 30 \text{ sec} )</td>
<td>% threshold used in step (1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \theta_1 = 3 ) ( \theta_2 = 1 )</td>
<td>% coefficient used in step (1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p \geq 0.7 )</td>
<td>% threshold used in step (3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \delta = 1.2 )</td>
<td>% threshold used in step (4)</td>
<td></td>
</tr>
<tr>
<td><strong>Chewing Gum</strong></td>
<td>( N = 5 )</td>
<td>% parameter used in step (2)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>delivery measure</th>
<th>Setting</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recommendation</strong></td>
<td>none</td>
<td></td>
</tr>
<tr>
<td><strong>Remarketing</strong></td>
<td>( X_{rm} = 90 \text{ days} )</td>
<td></td>
</tr>
<tr>
<td><strong>Email</strong></td>
<td>( X_{ms1} = 90 \text{ days} ) ( X_{ms2} = 4 \text{ days} )</td>
<td></td>
</tr>
</tbody>
</table>

We also deploy a mechanism that can enable the CAMPS to activate and deactivate corresponding personalization algorithms and delivery measures separately according to the context-aware personalization strategy. Depending on the status of the mechanism, there are four possible scenarios. (1) The mechanism is switched off and the CAMPS choose algorithm and delivery measure randomly. (2) The mechanism is switched on only for selecting the personalization algorithm. (3) The mechanism is switched on only for selecting the delivery measure. (4) The mechanism is completed switched on. This mechanism can be helpful when we need to evaluate the CAMPS’s performance in Section 6.3.
6.2.5 Deploy the performance improvement module

The performance improvement module is responsible for discovering the effective personalization strategies for each intention cluster. The COMPANY X chooses to pursue the net profit objective and shares with us its internal profit and cost for different content delivery measures. Using these data, we manage to compute the threshold purchase rate \( r_{purr}^k \geq \frac{c_k}{p} \) for each personalization strategy. It can be used to qualify (disqualify) effective (ineffective) personalization strategies.

6.3 Evaluating the performance of the CAMPS

Once the CAMPS is deployed, we are interested in evaluating its performance and comparing it with COMPANY X’s legacy personalization system to determine if the context-aware and adaptation strategy helps enhancing consumer conversion and engagement.

To achieve this goal, we design an online experiment. When consumers visit the retailing network of COMPANY X, they are randomly and equally assigned to one of the five experiment groups with reference to the SSID. The experiment is composed of two stages (Table 6.2).

<table>
<thead>
<tr>
<th>Group</th>
<th>Group Size</th>
<th>Stage</th>
<th>Intention Prediction</th>
<th>Personalization Algorithm</th>
<th>Delivery Measure Selection</th>
<th>Learning Algorithm</th>
<th>Budget Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group A</strong></td>
<td>20% of the traffic</td>
<td>Preparation (6m)</td>
<td>on</td>
<td>off</td>
<td>off</td>
<td>off</td>
<td>off</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Test (6m)</td>
<td>on</td>
<td>off</td>
<td>off</td>
<td>off</td>
<td>off</td>
</tr>
<tr>
<td><strong>Group B</strong></td>
<td>20% of the traffic</td>
<td>Preparation (6m)</td>
<td>on</td>
<td>off</td>
<td>off</td>
<td>off</td>
<td>off</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Test (6m)</td>
<td>on</td>
<td>off</td>
<td>off</td>
<td>off</td>
<td>on</td>
</tr>
<tr>
<td><strong>Group C</strong></td>
<td>20% of the traffic</td>
<td>Preparation (6m)</td>
<td>on</td>
<td>off</td>
<td>off</td>
<td>off</td>
<td>off</td>
</tr>
</tbody>
</table>
(1) The preparation stage takes six months. It allows the CAMPS to collect adequate consumer feedback data which can reveal the preference of different intention clusters for personalization strategies. During this stage, all CAMPS modules are switched off except for the intention management module that is responsible for predicting consumers’ instantaneous intentions. Hence, all the experiment groups have the exactly same configurations and parameters. Each personalization strategy has an equal chance (i.e. one-twelfth) to be selected by the CAMPS. Based on consumer data collected in the preparation stage, the CAMPS is able to prepare the personalization strategy performance database.

(2) The test stage takes six months and the five experiment groups are configured differently (Table 6.2). Group A keeps all the modules deactivated (except for the intention prediction model that is always activated, same for the following groups) to simulate the situation of the legacy system. Group B activates the budget estimator to determine if the price filter helps enhance the quality and performance of personalized marketing. Group C only activates the learning algorithm and the personalization algorithm selection mechanism in order to determine the effect of smart algorithm selection. Group D only activates the learning algorithm and the delivery measure selection mechanism in order to determine the effect of smart delivery measure selection. Group E activates all the modules and mechanisms so as to determine the capability of the CAMPS.

<table>
<thead>
<tr>
<th>Group</th>
<th>Traffic Percentage</th>
<th>Preparation (6m)</th>
<th>Test (6m)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group D</strong></td>
<td>20% of the traffic</td>
<td>on off off off on off</td>
<td>on off off off off</td>
</tr>
<tr>
<td><strong>Group E</strong></td>
<td>20% of the traffic</td>
<td>on off off off on off</td>
<td>on on on on on</td>
</tr>
</tbody>
</table>
6.3.1 Analyzing consumer data for the preparation phase

The preparation phase focuses on gathering each intention cluster’s preference for personalization strategies. The CAMPS records the intention label of each sub-stream and it delivers and tracks the personalized marketing contents based on the undifferentiated personalization strategy (i.e. each algorithm and delivery measure has an equal chance to be used). The summary statistics about the delivered and purchased personalized marketing contents are presented in Table 6.3.

Table 6.3(a) Purchase Rate by Personalization Algorithm

<table>
<thead>
<tr>
<th>Intention</th>
<th>Personalization Algorithm</th>
<th>Conversion by Intention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(i) Pa-Pa</td>
<td>(ii) Pa-Ta</td>
</tr>
<tr>
<td>LIT (N=4,249,392)</td>
<td>0.45%</td>
<td>0.32%</td>
</tr>
<tr>
<td>NIT (N=5,258,964)</td>
<td>0.01%</td>
<td>0.09%</td>
</tr>
<tr>
<td>PIT (N=71,772)</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>VIT (N=3,761,376)</td>
<td>0.34%</td>
<td>0.39%</td>
</tr>
</tbody>
</table>

Overall, the purchase conversion rate of personalized marketing contents is 0.17% (Table 6.3a: “conversion by intention” column). The purchase conversion rates of LIT, PIT, and VIT consumers are relatively higher (0.20%~0.23%) than that of the NIT consumers (0.10%), suggesting that NIT consumers are generally less interested in purchasing personalized marketing contents. The result agrees with our intuitive profile for NIT consumers (Section 3.3), who seek to clarify their shopping objectives before making a purchase.

By looking at the algorithm-specific purchase conversion rates, it is clear that the Pa-Pa and Pa-Ta algorithm outperform the other algorithms (Pa-Ph, Ch-Gm), even though the difference of performance is not huge (Table 6.3a: “conversion by algorithm”). However, by investigating deeply into intention specific purchase conversion rates of different personalization algorithms (Table 6.3a), the general conclusion does not apply anymore.
LIT consumers prefer the Pa-Pa algorithm, whose conversion rate (0.45%) is 114% higher than the overall conversion rate of the group (0.21%). Pa-Ph is the favorite algorithm of NIT consumers, whose performance (0.28%) is two times better than the second ranking algorithm (Pa-Ta, 0.09%). PIT consumers prefer the Ch-Gm algorithm (0.43%), even though the algorithm is almost not effective in all other intention clusters. VIT consumers prefer the Pa-Ta algorithm, whose conversion rate (0.39%) is 70% higher than the group’s average conversion rate (0.23%). These results confirm our assumption that consumers’ preference for personalization algorithms is contingent with their intention and there is no algorithm that can outperform others in all the intention clusters.

Table 6.3(b) Purchase Rate by Communication Medium

<table>
<thead>
<tr>
<th>Intention</th>
<th>Delivery Measure</th>
<th>Conversion by Intention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(i) Recom</td>
<td>(ii) Remkt</td>
</tr>
<tr>
<td>LIT (N=4,249,392)</td>
<td>0.28%</td>
<td>0.23%</td>
</tr>
<tr>
<td>NIT (N=5,258,964)</td>
<td>0.09%</td>
<td>0.12%</td>
</tr>
<tr>
<td>PIT (N=71,772)</td>
<td>0.28%</td>
<td>0.16%</td>
</tr>
<tr>
<td>VIT (N=3,761,376)</td>
<td>0.22%</td>
<td>0.21%</td>
</tr>
<tr>
<td>Conversion by Algorithm</td>
<td>0.19%</td>
<td>0.18%</td>
</tr>
</tbody>
</table>

At first glance, the delivery measure does not seem to be a decisive factor affecting the purchase conversion rate in different intention clusters. Even though the purchase conversion rate of the email is slightly lower (0.15%), the conversion rate gap to the other delivery measures is minor (Table 6.3b: “conversion by delivery measure”). From an intention cluster perspective, instant recommendations are the favorite delivery measure of LIT and PIT consumers, VIT consumers are more receptive to emails, and NIT consumers prefer buying items presented by personalized remarketing ads. However, the impact of distinct delivery measures is less important than the impact of different personalization algorithms, suggesting that consumers pay more attention to quality than format when evaluating personalized marketing contents.
We further analyze the impact of personalization strategy (i.e. the selection of the personalization algorithm and the delivery measure as a whole) on the purchase conversion rate. The results indicate that certain personalization strategies (highlighted in Table 6.3c) are far more effective than others in converting consumers to customers. The conclusion is based on the fact that each strategy has an equal chance to be used during the preparation stage. Hence, the high-performing personalization strategies do not achieve a high purchase conversion rate by chance. The huge difference in purchase conversion rate is the foundation for the CAMPS to choose high-performing strategies for each intention cluster.

Table 6.3(c) Purchase Rate by Personalization Strategy

<table>
<thead>
<tr>
<th>Intention</th>
<th>PS-01</th>
<th>PS-02</th>
<th>PS-03</th>
<th>PS-04</th>
<th>PS-05</th>
<th>PS-06</th>
<th>PS-07</th>
<th>PS-08</th>
<th>PS-09</th>
<th>PS-10</th>
<th>PS-11</th>
<th>PS-12</th>
<th>Conversion by Intention</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIT (N=4,249,392)</td>
<td>0.75%</td>
<td>0.46%</td>
<td>0.15%</td>
<td>0.30%</td>
<td>0.35%</td>
<td>0.31%</td>
<td>0.04%</td>
<td>0.09%</td>
<td>0.07%</td>
<td>0.02%</td>
<td>0.01%</td>
<td>0.01%</td>
<td>0.21%</td>
</tr>
<tr>
<td>NIT (N=5,258,964)</td>
<td>0.03%</td>
<td>0.00%</td>
<td>0.01%</td>
<td>0.09%</td>
<td>0.15%</td>
<td>0.04%</td>
<td>0.23%</td>
<td>0.34%</td>
<td>0.28%</td>
<td>0.03%</td>
<td>0.01%</td>
<td>0.00%</td>
<td>0.10%</td>
</tr>
<tr>
<td>PIT (N=71,772)</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.43%</td>
<td>0.27%</td>
<td>0.35%</td>
<td>0.70%</td>
<td>0.38%</td>
<td>0.22%</td>
<td>0.20%</td>
</tr>
<tr>
<td>VIT (N=3,761,376)</td>
<td>0.32%</td>
<td>0.29%</td>
<td>0.41%</td>
<td>0.26%</td>
<td>0.26%</td>
<td>0.22%</td>
<td>0.13%</td>
<td>0.08%</td>
<td>0.02%</td>
<td>0.08%</td>
<td>0.08%</td>
<td>0.14%</td>
<td>0.23%</td>
</tr>
<tr>
<td>Conversion by Algorithm</td>
<td>0.34%</td>
<td>0.23%</td>
<td>0.17%</td>
<td>0.26%</td>
<td>0.26%</td>
<td>0.22%</td>
<td>0.13%</td>
<td>0.21%</td>
<td>0.16%</td>
<td>0.02%</td>
<td>0.03%</td>
<td>0.04%</td>
<td>0.17%</td>
</tr>
</tbody>
</table>

In order to validate the above intuitive analysis, we need to determine if the performance of personalization algorithms and delivery measures are intention-dependent. To do so, we make a univariate analysis of the variance, whose null hypothesis (H_0) is that the means of the intention-specific purchase conversion rates are not contingent with the personalization algorithms, delivery moments, or the delivery measures. We
aggregate the purchased personalized marketing contents by 240 groups (i.e. 5 experiment groups * 4 intention clusters * 4 personalization algorithms * 3 delivery measures) and compute the purchase rate of each of these 240 groups for the univariate analysis, whose results are presented in Table 6.4 (next page).

Table 6.4 Result of the Univariate Analysis of Variance

<table>
<thead>
<tr>
<th>Between-Subjects Factors</th>
<th>Value Label</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>intention cluster</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>1.00 LIT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00 NIT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.00 PIT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.00 VIT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>algorithm</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>1.00 Pa-Pa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00 Pa-Ta</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.00 Pa-Ph</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.00 Ch-Gm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>delivery measure</td>
<td></td>
<td>80</td>
</tr>
<tr>
<td>1.00 Recom</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00 Remkt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.00 Email</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>.002</td>
<td>47</td>
<td>3.336E-05</td>
<td>15.525</td>
<td>.000</td>
</tr>
<tr>
<td>Intercept</td>
<td>.002</td>
<td>1</td>
<td>.002</td>
<td>876.232</td>
<td>.000</td>
</tr>
<tr>
<td>intention</td>
<td>.001</td>
<td>3</td>
<td>.000</td>
<td>89.733</td>
<td>.000</td>
</tr>
<tr>
<td>algorithm</td>
<td>2.469E-05</td>
<td>3</td>
<td>8.229E-06</td>
<td>3.830</td>
<td>.011</td>
</tr>
<tr>
<td>delivery</td>
<td>1.210E-05</td>
<td>2</td>
<td>6.052E-06</td>
<td>2.817</td>
<td>.062</td>
</tr>
<tr>
<td>intention * algorithm</td>
<td>.001</td>
<td>9</td>
<td>7.572E-05</td>
<td>35.240</td>
<td>.000</td>
</tr>
<tr>
<td>intention * delivery</td>
<td>4.84E-05</td>
<td>6</td>
<td>8.076E-06</td>
<td>3.759</td>
<td>.001</td>
</tr>
<tr>
<td>algorithm * delivery</td>
<td>2.581E-05</td>
<td>6</td>
<td>4.302E-06</td>
<td>2.002</td>
<td>.067</td>
</tr>
<tr>
<td>intention * algorithm * delivery</td>
<td>.000</td>
<td>18</td>
<td>1.094E-05</td>
<td>5.091</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>.000</td>
<td>192</td>
<td>2.149E-06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.004</td>
<td>240</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>.002</td>
<td>239</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .792 (Adjusted R Squared = .741)

The univariate model explains the total variance of the purchase conversion rate by the main effect (i.e. the impact of different experiment conditions) and their interactions. The models adjusted r-squared is 0.741, suggesting that the model is able to well explain the differences of means based on the selected experiment conditions. The model confirms that the purchase conversion rate is dependent on the intention (p = 0.000 < 0.01, reject H0) and the variance of purchase conversion rate is caused by three factors, namely the interaction between intention and algorithm (p = 0.000 < 0.01, reject H0), the interaction between intention and delivery measure (p = 0.001 < 0.01, reject H0), and the interaction of all the three conditions (p = 0.000 < 0.01, reject H0). The conclusion
of the univariate analysis validates our intuitive analysis. Our context-adaptive personalization strategy is based on this conclusion.

The CAMPS computes the initial frequency matrix for the context-adaptive personalization strategies (Table 6.5 in the next page) with reference to the improvement process presented in Section 5.5.2. Being the backbone of the context-aware personalization strategy, the matrix determines the algorithm and the delivery measure the CAMPS is to use for a sub-stream whose intention label is predicted.

Table 6.5 Initial Frequency Matrix of Personalization Strategy

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LIT</td>
<td>NIT</td>
<td>PIT</td>
<td>VIT</td>
<td>LIT</td>
</tr>
<tr>
<td>PS-01</td>
<td>(i) Pa-Pa+(i) Recom</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>19.6%</td>
</tr>
<tr>
<td>PS-02</td>
<td>(i) Pa-Pa+(i) Remkt</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>19.6%</td>
</tr>
<tr>
<td>PS-03</td>
<td>(i) Pa-Pa+(i) Email</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>19.6%</td>
</tr>
<tr>
<td>PS-04</td>
<td>(i) Pa-Pa+(i) Recom</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>13.1%</td>
</tr>
<tr>
<td>PS-05</td>
<td>(i) Pa-Pa+(i) Remkt</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>13.1%</td>
</tr>
<tr>
<td>PS-06</td>
<td>(i) Pa-Pa+(i) Email</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>13.1%</td>
</tr>
<tr>
<td>PS-07</td>
<td>(i) Pa-Pa+(i) Recom</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>13.1%</td>
</tr>
<tr>
<td>PS-08</td>
<td>(i) Pa-Pa+(i) Remkt</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>13.1%</td>
</tr>
<tr>
<td>PS-09</td>
<td>(i) Pa-Pa+(i) Email</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>13.1%</td>
</tr>
<tr>
<td>PS-10</td>
<td>(i) Pa-Pa+(i) Recom</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>13.1%</td>
</tr>
<tr>
<td>PS-11</td>
<td>(i) Pa-Pa+(i) Remkt</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>13.1%</td>
</tr>
<tr>
<td>PS-12</td>
<td>(i) Pa-Pa+(i) Email</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>8.3%</td>
<td>13.1%</td>
</tr>
</tbody>
</table>

By examining the matrix in detail, we can identify several features.

(1) The personalization strategies in Group A and Group B are identical. This is because we specify that the personalization strategy is randomly chosen in these groups to simulate the legacy marketing personalization system of COMPANY X. The difference between Group A and Group B is that the later activates the budget filter.

(2) In Group C, personalization strategies with the same algorithm have the identical frequency, which indicates that their delivery measure is chosen randomly.

(3) Similarly, strategies with the same delivery measure have the same frequency in Group D, indicating that their personalization algorithm is chosen randomly.
The frequencies in Group E are quite different because both the algorithm and the delivery measure selection mechanisms are activated.

In addition, the blanks in the matrix indicate that these personalization strategies are deactivated for the corresponding intention cluster since their purchase conversion rate is lower than the threshold conversion rate. During the test stage, the frequency matrix is to be updated regularly using new consumer feedback data so that the CAMPS can adapt to the evolvement of consumers’ preference and behavior pattern.

### 6.3.2 Analyzing consumer feedbacks for the test stage

The most important KPIs for a marketing personalization system is its conversion ability and the user experience it may bring to consumers. Following these two directions, we present the performance of the CAMPS as follows.

#### 6.3.2.1 Conversion ability

We analyze the purchase conversion rate data collected during the test stage to verify if the context-adaptive personalization strategy is able to outperform the undifferentiated strategy. We are particularly interested in three questions.

1. **Does the context-adaptive strategy consistently outperform the undifferentiated strategy in all intention clusters?**
2. **Is the context-adaptive strategy able to maintain its advantage over the undifferentiated strategy over time?**
3. **Is the learning mechanism enable the CAMPS to make progressive improvements over time?**
To answer these questions, we compute the purchase conversion rate of personalized marketing contents proposed in different intention states for each group of consumers and plot them in Figure 6.4.

![Figure 6.4 Purchase Conversion Rate by Intention Cluster](image)

To answer the first question, we refer to the curve of Group A as the baseline, which does not activate any context-adaptive modules. By comparison, we can find that the price filter of the budget estimation module (i.e. Group B) does not enhance the purchase conversion rate by itself. There is a minor improvement of purchase conversion rate in Group D, indicating that the high performing delivery measures can provoke the intention to purchase suggested products. A much greater improvement can be observed in Group C, where the high-performing personalization algorithms are selected for each intention state. And the greatest improvement comes from Group E, where all
the context-adaptive measures (i.e. price filter, algorithm selection, delivery measure selection) are activated.

To further validate the answer to the first question, we randomly split all the personalized marketing contents that have the same intention label and belong to the same experiment group into 100 clusters and compute their purchase rates. Accordingly, we have 500 clusters (i.e. 5 experiment groups * 100 clusters/group). We analyze the variance (ANOVA) of these clusters to determine if the “experiment group” (i.e. personalization strategy) have an impact on the purchase conversion rate. The ANOVA results in Table 6.6 confirm that the purchase conversion rate of Group B (budget filter) and Group D (delivery measure) are not different from Group A (the baseline) while that of Group C (algorithm selection) and Group E (all activated) are very different from the baseline⁶. It indicates that activating all the context-adaptive measures, namely the personalization strategy of the CAMPS, can attain the highest purchase conversion rate.

Table 6.6 ANOVA Test for the Purchase Conversion Rate by Intention State

<table>
<thead>
<tr>
<th>Test of Homogeneity of Variances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levene Statistic</td>
</tr>
<tr>
<td>df1</td>
</tr>
<tr>
<td>943</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Subset for alpha = 0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.18354%</td>
</tr>
<tr>
<td>Group A</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Group B</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Group D</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Group C</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Group E</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Sig.</td>
<td></td>
<td>.048</td>
</tr>
</tbody>
</table>

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 100,000.

⁶ The difference is significant at 0.01 level.
The answer to the second question is obvious. Over a period of six months, the performance of Group E is consistently better than Group A across all intention states, indicating that the CAMPS’s advantages over COMPANY X’s legacy system is persistent, even though they use the same personalization algorithms and delivery measures.

The third question is not easy to answer due to the insufficiency of data. We discover that the curve of Group E (i.e. activate all the context-adaptive measures) follows an upward trend. However, there are a few months when the purchase conversion rate of Group E drops slightly. It is possible that these declines are caused by the overall fluctuation of consumer purchase intention. However, we are not sure to what extent certain seasonality effects (e.g. Christmas promotion, summer holiday) can affect consumers’ purchase intention. In addition, we only have six months’ data, it is not sufficient data to perform a seasonal decomposition test so as to validate our assumption.

![Figure 6.5 The Net Contribution of the CAMPS and the Growth Rate](image)

**Figure 6.5 The Net Contribution of the CAMPS and the Growth Rate**
Nevertheless, we manage to use an additive model, which assumes that the performance of Group E is a linear combination of the general purchase intention represented by the purchase rate of the baseline group (Group A) and the CAMPS’s influence, to evaluate the net contribution of the CAMPS. The net contribution of the CAMPS and its growth rate is presented in Figure 6.5 (previous page), which shows that during a six-month period, the CAMPS makes a positive net contribution to the purchase conversion rate and the month on month growth rate of the contribution is also positive. The result suggests that the learning algorithm is effective during the test period and the CAMPS is learning from its past success. We need to gather more data to have a more solid conclusion.

6.3.2.2 User experience

We are also interested in the opinions of the users of the marketing personalization systems. Since it is not feasible to interview hundreds of thousands of consumers who interact with the personalized marketing contents, we choose to evaluate the systems using two metrics (Table 5.1).

The click-through delay measures how much time it takes before a displayed personalized marketing content is clicked or tapped by a consumer. A short delay indicates that the content arrives at the appropriate time. The duration of interaction measures the time consumers spend on personalized contents. A long duration indicates that the personalized content is more attractive.

We use the independent samples T-test to compare the means of these two metrics for the legacy system (Group A) and the CAMPS (Group E). This statistical test is commonly used to determine if the statistical averages (i.e. means) of two metrics are significantly different.
The descriptive statistics (Table 6.7) indicates that the average click through delay of Group A (the legacy system) is 13.2 (N = 35,728) seconds and that of Group E (the CAMPS) is 10.2 seconds (N = 87,269). The Levene’s test for equality of variances indicates that the null hypothesis (H₀: the variances of two groups are equal) should be rejected because the homogeneity of variances is not significant at 0.01 level (p = 0.000 < 0.01). Therefore, we refer to the highlighted row where the equal variance is not assumed. The independent samples test suggests that the null hypothesis (H₀: the distributions of the two sample groups are equal) should be rejected and the means of the two groups are significantly different (p = 0.000 < 0.01). We perform the nonparametric tests for the two groups and the results in the hypothesis test summary of Table 6.7 confirm the previous conclusion.

Table 6.7 Result of the Independent Samples T-Test

<table>
<thead>
<tr>
<th>Descriptive Analysis - Click Through Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>group</td>
</tr>
<tr>
<td>click through delay</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Independent Samples Test - Click Through Delay</td>
</tr>
<tr>
<td>Levene’s Test for Equality of Variances</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>equal var assumed</td>
</tr>
<tr>
<td>equal var not assumed</td>
</tr>
<tr>
<td>t-test for Equality of Means</td>
</tr>
<tr>
<td>t</td>
</tr>
<tr>
<td>lower</td>
</tr>
<tr>
<td>click through delay</td>
</tr>
<tr>
<td>equal var not assumed</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Descriptive Analysis - Duration of Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>group</td>
</tr>
<tr>
<td>duration of interaction</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Independent Samples Test - Duration of Interaction</td>
</tr>
<tr>
<td>Levene’s Test for Equality of Variances</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>equal var assumed</td>
</tr>
<tr>
<td>equal var not assumed</td>
</tr>
<tr>
<td>t-test for Equality of Means</td>
</tr>
<tr>
<td>t</td>
</tr>
<tr>
<td>lower</td>
</tr>
<tr>
<td>duration of interaction</td>
</tr>
<tr>
<td>equal var not assumed</td>
</tr>
</tbody>
</table>

197 / 273
When consumers consider that personalization marketing contents are relevant to their instantaneous needs and arrive at an appropriate time, they quickly click or tap it to explore the details. Apparently, marketing contents in Group E are considered more relevant and appropriate because the click through delay is shorter. Since the two groups use the same user interfaces (i.e. COMPANY X’s website, remarketing ads banners, and COMPANY X’s email templates), we should attribute the performance difference to the CAMPS’s context-adaptive strategies.

### Duration of interaction

Meanwhile, the descriptive statistics (Table 7.9) indicates that the average duration of interaction of Group A (the legacy system) is 120.9 seconds (N = 35,728) and that of Group E (the CAMPS) is 142.1 seconds (N = 87,269). The Levene’s test for equality of variances indicates that the null hypothesis (H₀: the variances of two groups are equal) should be accepted because the homogeneity of variances is significant at 0.01 level (p = 0.697 > 0.01). Therefore, we refer to the highlighted row where the equal variance is assumed. The independent samples test suggests that the null hypothesis (H₀: the distributions of the two sample groups are equal) should be rejected (p = 0.000 < 0.01) and the means of the two groups are significantly different. Due to the homogeneity of variances, there

<table>
<thead>
<tr>
<th>Hypothesis Test Summary</th>
<th>Null Hypothesis</th>
<th>Test</th>
<th>Sig.</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The distribution of click through delay is the same across categories of group.</td>
<td>Independent-Samples Wald-Wolfowitz Rank Test</td>
<td>.001¹</td>
<td>Reject the null hypothesis.</td>
<td></td>
</tr>
<tr>
<td>2. The distribution of click through delay is the same across categories of group.</td>
<td>Independent-Samples Mann-Whitney U Test</td>
<td>.000</td>
<td>Reject the null hypothesis.</td>
<td></td>
</tr>
<tr>
<td>3. The distribution of click through delay is the same across categories of group.</td>
<td>Independent-Samples Kolmogorov-Smirnov Test</td>
<td>.000</td>
<td>Reject the null hypothesis.</td>
<td></td>
</tr>
</tbody>
</table>

¹ Computed using the maximum number of runs when breaking inter-group ties among the records.

Asymptotic significances are displayed. The significance level is 0.01.
is no need to perform the nonparametric tests. We can conclude that the average duration of interaction in Group E is significantly greater that in Group A.

When consumers find that personalized marketing contents are helpful, they are likely to spend a long time exploring them. The results suggest that marketing contents presented by the CAMPS are more appreciated by consumers because consumers in Group E spend more time (i.e. an additional 20%) on personalized marketing contents. Since the users of the CAMPS and the legacy system are randomly allocated, we can attribute the improvement to the CAMPS’s context-adaptive strategies.

6.3.2.3 Value creation

So far, we only analyze the direct impact of personalized marketing contents on consumers’ shopping behavior. That is, consumers purchase the same item as is proposed by marketing personalization systems. Apparently, personalized marketing contents may also inspire consumers to purchase products that are similar to the proposed contents. This is the indirect contribution of marketing personalization systems.

In COMPANY X’s case, we define the indirect contribution as the purchase of items whose taxonomical distance to the proposed item is less than 0.4. That is, the purchased product and the product proposed by a marketing personalization system has the same parent node. A consumer may purchase a similar item just because he (or she) prefers the color or the vendor of the similar item. Marketing personalization systems
may not be sensitive to these detailed features due to technical and managerial constraints\(^7\). We are interested in evaluating the indirect contribution of marketing personalization systems because it may provide directions for the further enhancement of marketing personalization systems.

**Figure 6.6 Direct and Indirect Sales Contribution Ratio**

Based on the sales contribution validation policy we specified in Section 5.5.1, we aggregate the sales revenue directly and indirectly contributed by the marketing personalization system of Group A (the legacy system) and Group E (the CAMPS) respectively and plot them in Figure 6.6, where the indirect sales revenue contribution is presented by its ratio to the direct revenue contribution of the same month. By analyzing the figure, we can draw two conclusions.

---

\(^7\) Many marketplaces have specific policies and criteria to rank vendors, which may not be the same as that of consumers. It results in the fact that marketing personalization systems cannot always propose vendors that is highly evaluated by consumers.
(1) The CAMPS (Group E) contributes more direct revenue than the legacy system (Group A). This is due to the impact of the context-adaptive strategies that help improve the purchase conversion rate of the personalized marketing contents.

(2) The CAMPS also attains a much higher indirect to direct ratio than the legacy system, suggesting that the CAMPS is more capable of inspiring consumers and giving them shopping ideas, though it may not always be precise in predicting the exact items that consumers may purchase eventually.

6.4 Implications to retailers

Recently, marketing personalization systems are mainly regarded as sales tools. Retailers pay great attention to the direct sales conversion ability of them while neglecting their role as an influencer. The indirect sales contribution of the CAMPS suggests that the influencer or inspirer role is not auxiliary at all. By choosing personalization algorithms and delivery measures properly, personalization systems can be very influential when consumers are formulating their purchase plans.

From a technical perspective, the implication is that we can enhance the CAMPS in two ways. We can create a few more taxonomical tiers so that more detailed features of products can be included. As a result, the proportion of the direct sales contribution will increase. We can also modify the performance evaluation mechanism by introducing the indirect contribution factor. Hence, the proportion of the indirect sales contribution will rise. Whereas this subject remains to be the future improvement of the CAMPS, our intuitive judgment is that the second option may be more promising because of three reasons. (i) The first option requires a tremendous amount of workload to renovate the product catalog and similarity matrices. Even though we can accept
these efforts as an initial investment, we are not able to exhaust the detailed features of products. (2) Adding additional taxonomical tiers will increase the computation cost, which will have an adverse impact on the performance of the CAMPS. (3) Modifying the performance measurement mechanism only affects the improvement management module, which means less initial development work. Besides, analyzing indirect sales does not increase too much computation cost because anyway the CAMPS must examine all the purchases one by one to validate the sales contributions.

6.5 Summary

In this chapter, we presented the process to deploy the CAMPS in a brownfield context. By adding the necessary modules, we are able to convert a non-contextualized marketing personalization system into a CAMPS. This easy-to-deploy feature makes the modularized CAMPS not only suitable for retailers who want to construct a brand-new marketing personalization system, but also for firms who want to maintain the backbone of their existing systems and minimize the renovation cost. Based on a real-world experiment, we validated that consumers’ purchase intention is contingent with their intention. It is possible to enhance their intention to purchase by choosing the appropriate personalization algorithm, delivery measure, and budget range for the marketing personalization system based on the context-aware cognitive methods. The experiment also validates that the CAMPS is able to make progressive improvement based on its precious success. For retailers who want to impress their consumers and boost their sales performance, it is worthwhile to integrate the CAMPS to their retailing infrastructure.
Conclusion

7.1 Summary of the research project

The rapid development of technology reshapes the retailing business model as well as the way consumers make their purchase decision. Nowadays, consumers are no longer satisfied with obtaining abundant and diversified product information, which are rich in format and available almost on any platforms at any time, from retailers. More and more consumers expect novel and personalized offers in real time (Konstan and Riedl, 2012). The requirement for rapidity brings a huge challenge to modern retailers who compete aggressively to survive.

Current marketing personalization systems cannot cope with this emerging requirement because of two reasons. (1) Facing an unprecedented amount of users and items, existing personalization systems that use a heuristic approach to mobilize all personalization modules at the same time reach their performance bottleneck in terms of prediction accuracy and computation time (Bobadilla et al., 2013). (2) Excessive consumer contacts ruin the reputation of marketing personalization systems as a trusted advisor or assistant (Nguyen et al., 2013).

In this thesis, we propose a context-aware marketing personalization system, known as the CAMPS. The system is capable of unobtrusively decoding the cognitive intentions
hidden in consumers’ shopping behavior data that are gathered from different online and offline shopping platforms and using the appropriate personalization algorithm, delivery measure, and budget range to present personalized marketing contents that are relevant to consumers’ needs. Compared to the existing systems, the CAMPS has three major advantages.

(1) It consumes fewer computation resources because it uses only the personalization modules that are proven to be high-performing in similar cases. As a result, the time needed to generate a personalized content and the information load of the content is significantly reduced. Such advantage is of high value to retailers and marketers who need to present their marketing contents on small screens and for a great number of consumers.

(2) The CAMPS gathers data from online and offline retailing channels, which enables it to have a much more complete and holistic view of consumers’ cognitive states and needs. Such advantage is valuable for firms who operates multiple retailing channels simultaneously.

(3) The CAMPS has a progressive and automatic learning mechanism, which can identify the most effective personalization strategy and adapt the CAMPS to the evolvement of consumer preferences and behavioral patterns autonomously. As a result, the knowledge discovery and application process becomes automatic, real-time, and intelligent.

The real-world experiment validates that the performance of the personalization algorithms and delivery measures are intention dependent. The results of the experiment also confirm that the CAMPS, which uses an intention-based personalization strategy, outperforms its non-contextual counterpart in terms of the sales conversion ability, user experience, and indirect influence on consumer decisions.
7.2 Major contributions

This thesis achieves its research goal by fulfilling the four improvement needs that are presented in Section 0.2.3. In view of the challenges in B2C marketing personalization, we propose a context-aware marketing personalization approach (CAMPA) consisting of a cognition-behavior model (theoretical framework), a few new methods (i.e. intention-based personalization and contextualization methods) as well as some innovative technics, which allow firms to create a context-aware marketing personalization system (CAMPS). Meanwhile, the improvements we make to the current B2C marketing personalization practice have a few major impact on customer experience.

7.2.1 Theoretical contributions

In view of the complexity of cross-channel shopping behaviors and the lack of a comprehensive theory to explain it, we propose a cognition-behavior model, which contributes to the advancement of B2C marketing theories in three aspects.

(1) **The cognition-behavior model decomposes a complex shopping journey into progressive cognitive steps** so that researchers can explain different customer purchase motivations (i.e. impulsive purchase, planned purchase, and repeat purchase) based on the same theoretical framework. It also facilitates the discovery of the underlying information needs behind the various purchase motivations, which necessitate different B2C marketing and communication strategies.

(2) **The cognition-behavior model can uncover the underlying shopping intentions behind cross-channel shopping behaviors.** Such knowledge can help firms understand the main role of each retail channel in customers' shopping journey. When firms need to enhance their retail channels, they know where resources should be invested.
The cognition-behavior model lays a theoretical foundation for the discovery of distinct customer information needs at different cognitive steps. Based on such knowledge, firms can adjust the weight of past preference, current interest, and contextual factors in their personalization model so as to generate more relevant and useful marketing contents for customers.

Meanwhile, we redefine the contextualization process by proposing a more generic and comprehensive theoretical framework. Our theoretical framework extends the possibility of incorporating contextual factors to B2C marketing models so as to make them more relevant to the needs of customers.

### 7.2.2 Methodological contributions

In view of the drawbacks of the process-driven personalization methodology used by many B2C firms, we develop an intention-based personalization methodology to create and deliver marketing contents. Our methodology focuses on customers’ instantaneous intention and uses contextual factors to enhance the relevance and usefulness of marketing contents delivered to customers. It contributes to the advancement of B2C marketing methodologies in four aspects.

(1) Our intention-based personalization methodology adopts an adaptive method which enables an MPS to adjust its personalization strategy and patterns according to the evolvement of customers’ cognitive step and information needs. Compared to the current process-driven methodology which cannot adjust its personalization method, our cognition-driven methodology can better fulfill the diversified and changing customer needs.
(2) Our intention-based personalization methodology enables an MPS to keep track of customers’ cognitive load and use the proven appropriate medium to deliver marketing messages to customers at the right moment of contact. Compared to the process-driven methodology that pays little attention to customers’ cognitive load, our methodology can enhance customers’ experience with an MPS and increase the probability of recommendation success.

(3) Our intention-based personalization methodology incorporates a resilient learning mechanism allowing an MPS to analyze customer feedbacks, identify best practices, and apply them to upcoming similar cases in an automatic and recurring manner. Current process-driven methodology pays little attention to this domain. Our work advances the development of learning mechanism in the B2C marketing personalization sector.

(4) In addition, we improve the existing contextualization methodology and develop a cognition-based contextualization methodology allowing firms to incorporate contextual factors in a more diversified manner. Such improvement helps improve customers’ experience with firms’ marketing initiatives and enhances the productivity and efficiency of the generation and delivery of marketing contents.

7.2.3 Technical contributions

Our work contributes to the development of B2C marketing technics in the following aspects.

(1) We develop a cross-channel customer tracking mechanism to track, record, consolidate and analyze customer data generated in different retailing platforms (e.g. mobile application, e-commerce website, and physical stores). This technic provides firms
with a full visibility of customers’ cross-channel shopping behavior in the shopping journey, which may redefine the way firms develop and execute their B2C marketing strategy.

(2) We develop a **model capable of predicting customers’ instantaneous intention** based on their real-time cognitive state and behaviors. The model enables firms to track the change of customer motivations, interests, and information needs during the shopping journey. As a result, firms can have more profound and real-time knowledge of their customers, which is an indispensable factor for successful B2C marketing.

(3) We **define, quantify, and incorporate important contextual factors** for the MPS. These contextual factors not only have a huge impact on the relevance and usefulness of personalized marketing contents but also indispensable for other B2C marketing initiatives, such as customer relationship management (CRM), upselling, and referrals.

(4) We propose a **brownfield renovation technic to contextualize legacy MPS** and make them cognition and context-aware. This solution is attractive for most B2C firms who already have an MPS and do not want to suffer from additional investment and disturbance to business because of the introduction of a new MPS.

### 7.2.4 Implications for consumers

Although our CAMPA mainly focuses on helping firms better market their products, it has some positive impacts on consumers as well.

(1) Due to various efforts to improve the quality of the marketing personalization process, consumers have access to more relevant, credible, consistent, and useful marketing contents, which facilitate their shopping decisions. Such improvement not only increases consumers’ willingness to examine and purchase recommended products but
also enhance their confidence in a firm’s marketing contents.

(2) Adapting the creation and delivery of marketing contents to consumers’ cognitive state is a competitive edge in the B2C marketing domain. It means that consumers may receive the marketing contents they want at their preferred moment through a preferred medium. Meanwhile, they are less likely to be disturbed by excessive and irrelevant marketing contents. The better experience created by our CAMPA approach has a positive impact on customers’ satisfaction level and their loyalty to a firm.

7.3 Implications for retailers and marketers

In this section, we discuss the implications of our CAMPA approach and the CAMPS system on B2C marketing and personalization. The discussion focuses on three subjects: (1) How the CAMPA can change the way firms do B2C marketing? (2) How the CAMPA can change the way firms acquire, activate, and retain customers? (3) How can the CAMPA inspire the marketing and advertising industry?

7.2.1 Towards an agile B2C marketing approach

Based on the cross-channel behavior encoding, consolidation, and fragmentation, the CAMPA approach gives firms a full visibility to the shopping journey of their customers in real time. For retailers who struggle to obtain consumer insights from data, this is an unprecedented competitive edge. Compared to the existing data-driven marketing approaches, the CAMPA approach not only enables firms to know what their customers are doing right now (current approach can also do that) but also allows them to find out why customers do so and what the underlying information and transactional requirements are. Such knowledge enables firms to obtain consumer information that would
have taken them weeks or months to get in the past.

The enhanced rapidness of information gathering brings both opportunities and challenges to firms. On one hand, firms now have more profound consumer insights in real time. On the other hand, they are also under great pressure to react to these consumer insights using innovative interactive marketing and communication measures because their customers who generate these insights expect them to do so. If firms cannot turn consumer insights into immediate actions, they are going to be pushed out by competitors who excel in this new competence.

What does this mean to firms? It means that marketing plans can no longer be carried out in its old-school ways. Whereas strategies, brand images, annual media plans, and budgets may still be the backbone of today’s marketing practice, the time when marketers are allowed to take weeks or even months to plan, implement and evaluate a marketing campaign has passed (Figure 7.1). Due to the fact that firms and their consumers are exposed to new market trends, consumer habits and social events every passing minute, firms have to adopt an **agile marketing approach**, which is featured by its rapid reaction, short lead time, and constant improvement mechanism so as to cope up with the evolution of the market dynamics and competitions.

Figure 7.1 From Campaign-based to Agile Marketing Approach
In the short term (one to three years), firms may assign dedicated marketing teams to track, analyze, interpret, and react to the dynamics in consumer preferences, needs, and habits. However, with the exponential growth of new media and new consumer touch points, the main responsibilities of the marketing teams are set to be changed. In the long term (five years), marketers must shift their focus to the development of new systems, applications, and algorithms that can automatically track, analyze, interpret, and react to the consumer dynamics.

7.2.2 A new perspective for customer lifecycle management

Due to the prevalence of the cross-channel shopping behavior, firms suffer a lot from free-riding, namely the action to use a sales channel only for obtaining product information and user experience. Low-cost retailers can use a low price strategy to “steal” the price-sensitive customers from high-cost retailers who invest a lot in building fancy online and physical showrooms.

Recently, BestBuy’s decision to exit China (BestBuy Newsletter) provide a perfect footnote to this argument. The well-located mega stores of this U.S. based consumer electronics and home appliance retailing giant were always full of visitors. However, most of the visitors regarded BestBuy as a free playground to test these high-tech products. After that, they made their purchase order on other e-commerce platforms, where the price can be 10-15% lower.

Cross-channel shopping behaviors, especially the free-riding behaviors have an adverse impact on firms’ financial performance and pose challenges to the customer acquisition, activation, and retention. By providing firms with the visibility to consumers’
intention and cognitive state in real time, our CAMPA approach and the CAMPS system shed the light on the development of a new customer lifecycle management methodology. The knowledge of the maturity of a shopping decision and of the cognitive state of a shopping journey allows firms to develop more attractive and competitive marketing contents that can be more relevant and useful to consumers’ instantaneous shopping intention. As a result, free-riding behaviors can be reduced and the sales conversion can be enhanced.

(1) Today, firms invest a lot to acquire new sales leads and measure their sales conversion rate without knowing their cognitive state (Figure 3.1: intent, search, learn, test, compare, purchase) and intention (Figure 3.2: PIT, NIT, VIT, LIT). However, the traditional channel-specific sales funnel approach does not apply to a cross-channel shopping context, where consumers can quit and rejoin a channel at any time. The CAMPA approach allows firms to track and analyze consumer behaviors correctly based on their cognitive state and shopping intention. As a result, it keeps the channel-specific conversion metrics from being distorted.

(2) Nowadays, consumers have an unprecedented ability to obtain information and to look for alternative offers. Firms which actively acquires and activates consumers without knowing their cognitive state and intention may run the risk of creating sales opportunities for their competitors who know consumers better. This is an even worse result than doing nothing. The CAMPA approach provides firms with the real-time knowledge of their customers’ cognitive state and intention. Such knowledge enables firms to develop attractive and competitive marketing measures that can keep customers from purchasing elsewhere.
In the era of cross-channel shopping, customer retention becomes a very challenging task. Due to the prevalence of new sales channels and consumer touch points, today’s consumers have the freedom to purchase anywhere they like. The astonishing overseas online shopping expenditure made by the Chinese online shoppers can provide a good footnote to this argument (Mintel, 2016). As a result, today’s firms can forget about the loyalty of their customers and start to think about how they should be loyal to their customers. In another word, when a consumer has a purchase need, can a firm’s brands and sales channels be memorized and considered by him (or her) immediately? To this aim, firms need to keep track of their customers’ cognitive state and intention.

7.2.3 A new perspective of online marketing and advertising

The CAMPA also enables firms to use a systematic approach to evaluate and enhance consumers’ experience with online marketing and advertising contents that are created and delivered using different methods. The main objective of CAMPA is not to help firms reach their customers and potential customers but to impress these people using relevant contents and communication strategies. Such new methodology can drive the development of innovative remarketing technology, which is the hybrid version of online marketing and advertising.

Current remarketing agencies have a wide network of user touch points that a consumer may encounter. They (should) also have a profound knowledge of consumers’ online shopping journey and preference thanks to their cross-domain consumer tracking ability. However, do to the fact that current remarketing ads often send contents that consumers are already familiar with, they are considered annoying by many consumers (Couch, 2013). By integrating the CAMPA approach, remarketing agencies can optimize their current display strategy. To this aim, they should not only take into account
the preference and the demographic features of the audience, which they already do, but also pay attention to consumers’ preferred medium of contact and moment of contact under different cognitive state and shopping intention. By adjusting the display strategy in real time, remarketing agencies may revamp their market reputation and enhance their conversion rate.

7.2.4 Implement suggestions to firms

Meanwhile, we also have two suggestions to firms which are interested in applying the CAMPA approach and the CAMPS system.

(i) The processes to discover consumer intention cluster and to predict instantaneous consumer intention highlight the necessity to combine heuristic approaches with machine learning methods. In COMPANY X’s case, we have no information about the taxonomy of consumer intention or the proportion of each intention category. By inferring from consumer analysis data that are available to us and making intuitive assumptions on consumer behaviors, we manage to obtain some “reasonable” initial parameters and features for the intention prediction task. Once the classification model is set, the CAMPS is on its own to rectify the parameters, enhance prediction accuracy of consumer intention with reference to the consumer feedbacks, and identify the most appropriate personalization strategies for each intention cluster autonomously. Some knowledge acquired by the CAMPS may be hard to comprehend according to our intuitions and conventional wisdom (e.g. the best delivery measure for mining consumers is to display remarketing ads to them after their current shopping session). However, this is the true value of machine learning technologies, which can discover latent rules and patterns that are not obvious to business analysts and researchers (Jordan and Mitchell, 2015). If retailers and marketers seek to boost their business performance,
they should be prepared to integrate the new technologies.

(ii) The intention-dependent performance of personalization strategies means that there is no panacea in the domain of marketing personalization. The conclusion overthrows a popular notion that personalization algorithms play a decisive role in determining the relevancy and value of personalized marketing contents to consumers. Retailers and marketers need to personalize their personalization strategies based on diversified consumer contexts so as to create and deliver relevant and high-quality marketing contents to consumers. Due to the rapid evolvement of consumer decision-making patterns and behaviors, the personalization strategy does not remain unchanged. Retailers and marketers must revamp their personalization strategies constantly so as to remain the competitiveness.

7.2.5 Consumer privacy protection

A personalization task is dependent on the tracking, analysis, and exploitation of consumers’ personal data. As a result, if a firm wants to provide more relevant and personalized marketing contents to customers, it has to get more personal data from customers (i.e. the famous “personalization-privacy paradox”). Our CAMPA approach and the CAMPS system is also based on the interpretation of consumer data. However, we think that the conflict between personalization and privacy is not irreconcilable.

Privacy first. We always believe that the privacy has a higher privacy over the personalization, which means that consumers must be given the choice to say no. In our case, the CAMPS system obtain consumers’ consensus before tracking and analyzing any consumer data.

Transparency. Most consumers are not specialist in any discipline related to cross-
channel marketing personalization. As a result, it is the responsibility of the firm to inform customers of the implications of their choice. In our case, the CAMPS provide an easy-to-understand notification explaining to consumers which kinds of data are collected, where are they to be used, how personalized contents are generated, and why they see certain types of marketing contents.

**Data safety.** It is the responsibility of a firm to keep the consumer data in the right hands. Anonymization is not sufficient, because current technologies allow identity thefts to triangulate the identify of a consumer even they do not have his (or her) complete or critical personal information. In our case, the CAMPS uses a software as a service (SaaS) approach to process anonymized data streams stored on COMPANY X’s clouds so that it does not store any consumer data, which minimizes the risk of data manipulation and unauthorized use.

### 7.4 Limitations of the research

Although our research project proposes novel approaches to exploit consumer behavior data, resolves a few technical difficulties and attains the expected results, there remain several areas that requires further improvements.

(1) Due to the constraint of our research agreement with COMPANY X, we are not able to integrate and test a great number of novel personalization algorithms on COMPANY X’s platform. Even though COMPANY X’s personalization algorithms represent the mainstream methods (i.e. content-based, collaborative filtering, and statistical methods) that are popular in the marketing personalization domain, our generalized conclusion will be more convincing if it is based on an experiment that tests a greater number of novel personalization algorithms. This research objective will be realized in our future
research cooperation with COMPANY X.

(2) The experiment result indicates that the budget estimation protocol does not contribute to the enhancement of the purchase conversion rate. The conclusion seems inconsistent with our intuitive judgment. We believe that the ineffectiveness is caused by the personalization algorithms, which implicitly restrict the price when they choose similar items. Since estimating consumers’ budget is not a very hot research topic, we will keep track of the latest development of this research topic and work with the industry experts to propose better solutions.

(3) Due to the limited precision and stability of the indoor positioning devices available in the market, we are only able to obtain the category information of in-store browsing behaviors. As a result, the consumer-item matrix is less precise than it should be. If we can obtain reliable item-level in-store interaction information, the CAMPS’s preference management module will be more accurate in predicting consumers’ tastes and likes. We will keep evaluating different indoor positioning solutions in the market and discuss with the hardware manufacturers to accelerate the advent of the more precise technical solution.

7.5 Next generation marketing personalization system

With marketing personalization systems being the main engine to drive the growth of the retailing business, more and more research efforts are to be made in order to bring the capability of marketing personalization systems to a higher level. Currently, we are developing a research plan for the next generation context-aware marketing personalization system (CAMPS+), which will have the following new features.

With the ability to gather and analyze visual information, the CAMPS+ will be able to...
have a more in-depth knowledge about the in-store shopping activities. We will incorporate the pattern recognition technologies to infer consumers’ preferred products and even the colors. Such capability helps to enhance the accuracy of the consumer preference management module, which will in return improve the quality of the personalized marketing contents.

Supported by wearable devices, the CAMPS+ will be able to collect the biometrics of consumers and use these data to predict their instantaneous emotional state. We will try to discover the correlation between emotional state and the willingness to receive prompted messages. If such correlation exists, the CAMPS+ can be more precise in predicting the appropriate moment to present personalized offers.

The business focus of the CAMPS is to maximize the direct purchase conversion rate or the net profit of personalized marketing contents. Based on the tremendous indirect sales contribution presented in the experiment result, it may be worthwhile to revamp the performance improvement algorithm so that it can provide multiple business objectives for retailers and marketers to choose from.
List of Figures

Page 25  Figure 0.1 Structure of the Thesis
Page 48  Figure 1.1. Tracking the Variability of Contextual Factors
Page 54  Figure 1.2. Machine Learning Taxonomy
Page 65  Figure 2.1. Classification of Research Papers by Application Domain
Page 74  Figure 2.2. The Process-driven Personalization Method
Page 78  Figure 2.3. The Personalization Process Framework
Page 82  Figure 2.4. Context Prediction Models (a & b) and Contextualized Predictive Models for Consumer Choices and Behaviors (c & d)
Page 97  Figure 3.1. A Cognitive-Behavior Model for the Cross-Channel Shopping Journey
Page 100 Figure 3.2. The Taxonomy of Cross-Channel Shopping Intentions
Page 102 Figure 3.3. Intention-Based Personalization Methodology
Page 105 Figure 3.4. Cognition-Based Contextualization Methodology
Page 107 Figure 3.5. Context-Aware Marketing Personalization System (CAMPS) Architecture
Page 118 Figure 4.1. Mapping Virtual Categories and Items to Store Shelves and Products
Page 122 Figure 4.2. Floor Plan of a Physical Store with Connectivity (COMPANY X)
Page 126 Figure 4.3. Hierarchical Encoding Modality
List of Tables

Page 48  Table 1.1. Classifying Contextual Factors by Acquisition Method and Variability
Page 63  Table 2.1. Selected Papers on the Context-Aware Marketing Personalization Approaches
Page 76  Table 2.2. Composition of Personalized Contents by Firm
Page 124 Table 4.1. Examples of Consumer Data Collected by the CAMPS
Page 163 Table 5.1. KPI of the CAMPS
Page 179 Table 6.1 Clustering Result
Page 184 Table 6.2 Configurations for Each Experiment Group
Page 187 Table 6.3(a) Purchase Rate by Personalization Algorithm
Page 188 Table 6.3(b) Purchase Rate by Communication Medium
Page 189 Table 6.3(c) Purchase Rate by Personalization Strategy
Page 190 Table 6.4 Result of the Univariate Analysis of Variance
Page 191 Table 6.5 Initial Frequency Matrix of Personalization Strategy
Page 194 Table 6.6 ANOVA Test for the Purchase Conversion Rate by Intention State
Page 197 Table 6.7 Result of the Independent Samples T-Test
Page 227 Table A.1. Measuring Item-Consumer Pertinence
Page 229 Table A.2. Example of A User-Item Rating Matrix

221 / 273
### List of Acronyms

**-A-**
- AD: add to cart
- ANN: artificial neural network
- ANOVA: analysis of variance
- ANRT: Association Nationale de la Recherche et de la Technologie
- ARM: association rule mining
- ATIN: average time per item interaction
- AUC of ROC: area under the receiver operating characteristic curve

**-B-**
- BN: Bayesian network
- B2C: Business to Consumer

**-C-**
- CAMPA: Context-Aware Marketing Personalization Approach
- CAMPS: Context-Aware Marketing Personalization System
- CB: content-based
- CEO: chief executive officer
- CF: collaborative filtering
- CIFRE: Conventions Industrielles de Formation par la Recherche
- CRM: customer relationship management
- CV: cross validation

**-D-**
- DAG: directed acyclic graph
- DT: Decision tree

**-E-**
- EM: expectation maximization (EM)

**-G-**
- GMM: Gaussian Mixture Models

**-I-**
- IDF: inverse document frequency
- IOT: Internet of things
- IT: information technology

**-K-**
- KPI: key performance indicators
- kNN: k-nearest neighbor

**-L-**
- LDA: linear discriminant analysis
- LIT: learning intention

**-M-**
- MAE: mean absolute error
- ML: machine learning
- MLE: maximum likelihood estimation
- MPS: marketing personalization system

**-N-**
- NB: Naïve Bayes
- NFC: near field communication
- NIT: mining intention
- NMAE: normalized mean absolute error

**-P-**
- PC: personal computer
- PIT: purchase intention
- POI: point of interest
- PPR: precision
- PU: purchase event
RBF: radial basis function
RFID: Radio Frequency Identification
RM: remove from cart
RMID: recommendation identification
RMSE: root mean square error

SaaS: software as a service
SMIT: similarity of items and categories
SSID: session identification
SVD: singular value decomposition
SVM: support vector machines

TF: text frequency
TPR: recall

TVSM: transductive support vector machine
TRID: transaction identification

UCID: unique category identification
UUDI: unique user identification
UN: return to top tier categories
UPID: unique item identification
URL: uniform resource locator
UWB: ultra-wideband

VIP: very important person
VIT: evaluation intention

WIP: work-in-progress
Annex

A1 Personalization

A1.1 Consumer profiling

A consumer profile is a portrait telling us who the customers are and how they behave when purchasing and consuming products. Profiles can be used to depict the characteristics of one consumer or a group of them. The characteristics recorded in a profile range from demographic, geographic, psychographic, and economic features to preferences, shopping patterns, purchase history, and financial capability (Gunter and Furnham, 2014).

There are three methods to create a consumer profile. The first method, known as “labeling”, deals with factual information. The method labels consumers according to their demographic (age, marital status), social (education background), economic (revenue, wealth), geological (home address), preferential, and other states. Each consumer profile distinguishes itself from others by its unique labels. The second method, known as “rule-based profiling”, seeks to depict conditional facts about one or a group of consumers. The rules can be quite diversified in content, but consistent in essence. For example, “Adam buys a bunch of roses every Saturday when he goes for grocery shopping” specifies Adam’s habit (buy flower) during his weekend grocery shopping (condition). Similarly, the rule “when shopping online, Noah always choose the best rated vendor” Noah’s vendor preference given a condition (online shopping). The third method, known as “sequence”, aims at portraying the behavioral patterns that signify the owner(s) of the profile. For example, certain consumers’ shopping pattern can be represented as
“household essential ➔ beauty ➔ beverage ➔ food”. Once created, consumer profiles need constant maintenance to keep up with the evolvement of demographic, psychographic, economic, and behavioral features. Homogenous profiles are to be merged, and new profiles are to be created.

Consumer profiles are widely used in personalization tasks. In COMPANY X’s scenario, selected profiles are regarded as the target audience of the work-in-progress promotion programs or campaigns. Marketing practitioners craft programs with reference to the preferences, habits, and motifs of selected profiles so as to ensure their attractiveness and feasibility. Meanwhile, consumer profiles are the cornerstone of COMPANY X’s email marketing engine, which seeks to reactivate hibernate consumers using personalized messages. Besides, COMPANY X uses profiles as a reference to recommend products relevant to consumers’ purchasing habits. In addition to these applications, some firms also use profiles to anticipate consumers’ attitude towards customized offers and marketing contents.

Regardless of its wide application in personalization systems, consumer profiling is not a panacea for personalization problems. The underlying assumption of the consumer profiling approach is that consumers may keep to the characteristics, rules, and patterns that constitute their profiles. That is why consumer profiles are derived from the historical behaviors and transactions. However, the decision criteria and behavioral patterns of real-world consumers are always affected by the variant contexts. Accordingly, consumer profiling approach can misjudge the preference and needs of a consumer if it is not aware of the surrounding context.
A1.2 Content-based personalization

Content-based personalization approach seeks to refine consumers’ interests in their precedent choices and use the acquired knowledge to predict their subsequent needs. This approach is able to identify items and topics similar to those have been liked by a consumer in the past (Van Meteren and Van Someren, 2000). A typical content-based personalization system is composed of three key components, namely an item analyzer, an interest identifier, and a filtering module (Lops et al., 2011).

An item (e.g. product, topic, document) is usually presented by semi-structured or unstructured textual contents, which depict its features. The task of an item analyzer is to represent items using structured features extracted in contents. Firstly, textual information needs to be processed by some text mining techniques (Das and Martins, 2007) so that the key information can be extracted. Then, the key information must be processed to obtain features, namely the clusters of similar key information. At this step, features can still be unstructured data. Therefore, the second step is to convert extracted features from its original form to the anticipated one (e.g. a vector), which can be used for further analysis and computation. Thus, an item is represented by a set of features. Different items can be compared to determine their similarities. In practice, TF*IDF (text frequency times inverse document frequency) is considered a simple but effective text mining technique by many firms, including COMPANY X. The approach assumes that keywords (or features) with strong indicative ability should be frequently observed in the host item while rarely observed in other items (Pazzani and Billsus, 2007). According to the assumption, the TF*IDF score of a keyword is computed by the following equation, where $f_{t,i}$ denotes the frequency of keyword $t$ in item $i$, $N_i$ denotes the number of items, and $N_{t,i}$ denotes the number of items containing keyword $t$. Thus,
an item can be represented by an array of TF*IDF scores signifying the importance of
the corresponding keywords.

\[ w(t, i, l) = \frac{f_{t,i}}{\max\{f_{t',i} : t' \in i\}} \ast \log \frac{N_t}{N_{t,i} + 1} \]

In order to match relevant items to consumers, the interest of consumer should be iden-
tified as well. The responsibility of an interest identifier is to discover the features (e.g.
keywords) liked and detested by a consumer. Based on the structured representations
of item prepared by the item analyzer, the task to discover consumer interest is com-
posed of three steps. Firstly, the interest identifier needs to gather all the items that have
received consumer's feedback in the past. Next, the identified items are to be combined
with reference to the value of the features. Liked items are assigned positive values and
detested items are assigned negative values. Some weighting techniques can be used to
reflect the magnitude of consumer attitude (e.g. strong preference, consistent aversion).
Finally, the interests of consumers can be represented by a vector demonstrating their
unique attitude towards the features. During this step, the consumer interest vector
may need to be normalized so that it can be used by the filtering module.

<table>
<thead>
<tr>
<th>Features</th>
<th>acceleration</th>
<th>noise</th>
<th>design</th>
<th>space</th>
<th>eco-friendly</th>
<th>equipment</th>
</tr>
</thead>
<tbody>
<tr>
<td>( v_s )</td>
<td>0.7</td>
<td>0.1</td>
<td>0.6</td>
<td>0.4</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>( v_h )</td>
<td>0.2</td>
<td>0.8</td>
<td>0.7</td>
<td>0.6</td>
<td>0</td>
<td>0.7</td>
</tr>
<tr>
<td>( c_i )</td>
<td>0.8</td>
<td>-0.8</td>
<td>0.6</td>
<td>0.7</td>
<td>0.6</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Once item features and consumer interests are obtained, it is time for a filtering module
to find items relevant to consumers’ interests. Since item and consumer are both repre-
sented by a $1 \times n$ vector ($n$ is the number of features), it is possible to measure the relevance of an item to the given interest using some similarity metrics. For example, Table A.1 represents the feature of two cars ($v_s$ and $v_h$) as well as the interest of a consumer ($c_i$). It is obvious that $v_h$ is less relevant to the interest of $c_i$, who clearly demonstrates his (or her) aversion to the noisy and eco-unfriendly cars. Such observation can be validated by the Pearson correlation metric described by the following equation, which indicates a high positive $v_s$-$c_i$ correlation ($corr = 0.81$) and a negative $v_h$-$c_i$ correlation ($corr = -0.40$).

In practice, a filtering module needs to rank thousands of items according to their similarity scores and choose the most “similar” ones as the candidate for the personalization task.

$$\text{Corr}(x, y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2} \sqrt{\sum(y_i - \bar{y})^2}}$$

Content-based algorithm is an indispensable personalization tool for firms and consumers, but it also has some limitations. (i) The discovery of consumer interest requires certain amount of historical consumer feedbacks to items. When encountering new consumers who have not made enough choices yet, content-based algorithms are less effective. (ii) Consumers are not always looking for similar items. However, content-based algorithms are not good at providing “different” or “novel” items. In order to diversify the results of personalization, some other algorithms need be used.

A1.3 Collaborative filtering

Collaborative filtering approach aiming at suggesting personalized products to consumers by predicting their rating patterns on items based on their explicit and implicit
rating history. Generally speaking, there are two types of collaborative filtering methods: memory-based methods and model-based methods (Su and Khoshgoftaar, 2009).

**Memory-based methods** perform personalization tasks based on item-item or user-user relations. For instance, *item-based method* predicts an active user’s rating of a new item based on his (or her) ratings of similar items in the past (Sarwar et al., 2001). Similarly, *user-based method* discovers a group of users similar to the active user, and use their rating history of the new item to predict the active user’s rating of the new item (Marlin, 2003).

<table>
<thead>
<tr>
<th>user/item</th>
<th>$i_1$</th>
<th>$i_2$</th>
<th>$i_3$</th>
<th>$i_4$</th>
<th>$i_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>active user</td>
<td>?</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$u_1$</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>$u_2$</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$u_3$</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>$u_4$</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Table A.2. Example of A User-Item Rating Matrix

Table A.2 provides a simple example of the way *user-based method* rate new items. Our objective is to predict active user’s rating of item $i_1$. Using the Pearson correlation metric, it is obvious that the active user’s rating pattern is more similar with $u_1$ ($\text{corr}_1 = 0.95$) and $u_2$ ($\text{corr}_2 = 0.62$) than $u_3$ ($\text{corr}_3 = 0.19$) and $u_4$ ($\text{corr}_4 = -0.15$). Tough $u_1$ and $u_2$ can both be considered as “similar” users, the magnitude of similarity is different. Apparently, it is necessary to integrate a weighting method to the prediction so as to reflect the impact of user similarity. Hence, the predicted rating can be obtained using one of the following equations. The first equation predicts the active user’s rating of item $i$ ($\hat{r}_{ui}$) with reference to the weighted average rating of item $i$ made by his (or her) similar neighbors
\( r_{a,i} = 4.60 \), where \( r_{u,i} \) denotes a similar neighbor’s rating on item \( I \) and \( w_{a,u} \) denotes the similarity between the neighbor users and the active user. In the second equation, the predicted rating \( (r'_{a,i} = 4.55) \) is determined by the average rating style of the active user \( (r_e) \) on items and his (or her) preference to item \( i \), which is approximated by the preference of the neighbor users to item \( i \).

\[ r'_{a,i} = \frac{r_u \cdot w_{a,u}}{\sum |w_{a,u}|} \quad \text{or} \quad r'_{a,i} = r_a + \frac{\sum (r_{u,i} - r_u) \cdot w_{a,u}}{\sum |w_{a,u}|} \]

The rating can also be predicted using item-based method. In this case, the starting point becomes listing all the items rated by the active user in the past and identifying a neighborhood of items similar to new item \( i \). During this step, one may refer to factors or features characterizing items so as to compute the similarity metrics (e.g. cosine similarity coefficient, Pearson correlation coefficient, Jaccard similarity coefficient). Once the neighbor items are identified, the rating of new item \( i \) can be predicted according to the rating histories and the magnitude of similarity between new item and neighbor items, as is showed in the above equations.

In reality, a memory-based collaborative filtering system must predict the rating of a wide range of products so as to identify a few candidates of high rating value for personalized suggestions. In certain scenarios (e.g. movie rating) where users are frequently shared, memory-based methods are able to collect enough rating data for the collaborative filtering task. However, in some scenarios where the rating matrix is sparse (e.g. online shopping), one must consider some other methods. A possible solution is to implement default rating values (e.g. expert rating, inferred rating) to enlarge the rating database (Chee et al., 2001). Alternatively, one may consider to use machine learning techniques to fill in the missing ratings before performing the collaborative
Model-based methods seek to predict the active user’s rating of new items by modeling the components and processes that determine the rating of a new item. Unlike the memory-based methods, model-based methods can incorporate implicit user feedbacks (e.g. clicks, transactions) as a compensation to the inadequate rating information. In practice, four kinds of model-based methods are widely used by firms.

**Simple Bayes Network** seeks to predict an active user’s rating of new item \(i\) based on its features. Assuming the features are independent given the rating, the probability of a certain rating can be predicted given all features, and the rating with the highest probability is defined as the rating of the new item (Miyahara and Pazzani, 2000). The following equation signifies the computation process, where \(f_j\) denotes the \(j\)th feature of the new item. Simple Bayesian model is efficient in classifying Boolean (e.g. like and dislike) and complete data. Real-world rating prediction tasks are usually multi-class problems with sparse data, where the simple Bayesian model is less accurate (Su and Khoshgoftaar, 2006). To address the challenge, such techniques as extended logistic regression (ELR) are developed to learn the parameters and predict the missing feature values.

\[
\text{rating} = \arg\max \ p(\text{rating}_k) \prod_j^n P(f_j | \text{rating}_k) \quad k \in \text{rating Set}
\]

**Association rule mining** (ARM) method discovers coexistent relationships in items, which makes it useful in predicting items relevant to consumers’ personal needs (Mobasher et al., 2000). In the online shopping context, an ARM-based model discovers the item association rules by mining users’ navigation and transaction history (Mobasher et al.,
For example, if the preference pattern “purchase an expensive mobile phone → purchase a phone protection case” is frequently observed with young female consumers, we may infer that a lady is likely to accept the offer of a phone protection case, given that she belongs to the segment and she just purchased an iPhone6S 64GB. Similarly, if the association rule “Saturday evening, single male consumer → order Pizza” is frequently and consistently observed among the customers of a firm, it may expect high acceptance to the personalized Pizza coupons, which are recommended to similar prospect consumers on Saturday nights.

*Clustering* methods group similar data points and separate different ones with reference to their inherent features. In general, there are three clustering methods, namely partitional method, hierarchical method, and density method, which are introduced in Section 2.4. Considering each item as a data point and its technical specifications as features, we can use this technique to identify items similar to those viewed by users in the past. The similarity of the items can be measured by the Minkowski distance and its variants (i.e. Euclidean, Manhattan, Chebyshev distance), where \( x_{1i} \) denotes the \( i \)th feature of item \( X_1 \), and \( x_{2i} \) denotes the \( i \)th feature of item \( X_2 \).

\[
D_{Minkowski}(X_1, X_2) = \left( \sum_{i=1}^{n} |x_{1i} - x_{2i}|^q \right)^{\frac{1}{q}} \quad \text{when } q \neq 0
\]

\[
D_{Euclidean}(X_1, X_2) = \left( \sum_{i=1}^{n} |x_{1i} - x_{2i}|^2 \right)^{\frac{1}{2}} \quad \text{when } q = 2
\]

\[
D_{Manhattan}(X_1, X_2) = \sum_{i=1}^{n} |x_{1i} - x_{2i}| \quad \text{when } q = 1
\]

\[
D_{Chebyshev}(X_1, X_2) = \lim_{p \to \infty} \left( \sum_{i=1}^{n} |x_{1i} - x_{2i}|^q \right)^{\frac{1}{q}} \quad \text{when } q = \infty
\]
Clustering methods are usually regarded as an intermediary step in the collaborative filtering exercises, which serves to prepare an appropriate item dataset (usually a subset of all the item database) to be processed by other collaborative filtering techniques (Sarwar et al., 2002).

**Latent factor methods** such as singular value decomposition (SVD) can decompose user preference and item features and transform them into the same reduced latent factor space according to descriptive contents and user feedbacks recorded in the past. Accordingly, one may identify similar users, items, or user-item pairs useful for the subsequent prediction of the active user’s rating of a new item. In addition to the explicit user feedbacks (e.g. ratings), latent factor model may also take into account the implicit feedbacks such as web actions and shopping cart maneuvers.

For instance, a rating prediction model can be represented as follows, where the 1st term denotes the average rating of all items by all the users, the 2nd term denotes the deviation of rating style of the active user, the 3rd term denotes the deviation of rating of the new item, the 4th term denotes the explicit and implicit interaction of the active user’s profile with the feature profile of the new item, the 5th term denotes the residual effect of the active user’s rating on items similar to the new item, and the 6th term denotes the offset added to the baseline prediction considering the implicit user feedbacks (Koren, 2008).

\[
\hat{r}_{ui} = \mu + b_u + b_i + q^T \left( p_u + \frac{1}{2} \sum_{j \in N(u)} y_j + \left| R^k(i;u) \right|^{\frac{1}{2}} \sum_{j \in R^k(i;u)} (r_{uj} - b_u) w_{ij} \right. \\
\left. + \left| N^k(i;u) \right|^{\frac{1}{2}} \sum_{j \in N^k(i;u)} c_{ij} \right)
\]
A1.4 Hybrid approach

The hybrid approach is referred to the simultaneous use of more than one personalization method. In general, there are two hybrid approaches to integrating multiple personalization methods.

**Mixing** is referred to as the combination of items predicted by different personalization or non-personalization methods with a purpose to diversify personalization results. Before presenting the personalization result, a ranking system can be deployed to adjust the importance of different types of results and eliminate duplicated items. The mixing approach is easy to implement. However, it risks of providing too many irrelevant contents to users and making them confused when the origin of the mixed results is not properly explained.

**Twisting** is referred to as the use of several personalization methods in personalization modeling. Common twisting techniques include:

- using a two-staged (e.g. content-based + CF method) approach to predict ratings
- preparing a dataset using non-personalization features and making predictions using a personalization method
- using several collaborative filtering methods in parameter learning and prediction,
- re-ranking results predicted by a personalization model using non-personalization features.

In practice, most firms use a hybrid approach to obtain a more balanced personalization result which takes into account the relevance and novelty (Burke, 2002; Adomavicius and Tuzhilin, 2005). However, such objective cannot be achieved without the help of contextual factors (Burke, 2007).
A2 Machine learning algorithms

A2.1 Supervised machine learning

In some cases, a dataset contains examples whose inputs (or features) and the corresponding outputs (or labels) are known to the machine learning algorithms. The objective of learning is to find the association rule that connects the inputs and the outputs (Kotsiantis et al., 2007). A typical learning process serving such objective is composed of the following steps (Figure A.1).

![Supervised Machine Learning Process](image-url)

**Figure A.1. Supervised Machine Learning Process**
(1) The machine learning algorithm creates a model (i.e. a function or a set of functions) to predict outputs based on the provided inputs. (2) It compares its predicted output with the provided (correct) output to find error. Afterwards, the algorithm modifies its model to enhance the prediction accuracy. (3) The revised model is used to predict new outputs based on the previous inputs again. (4) The algorithm iterates the above steps until the prediction error is within an acceptable range specified by the user of the algorithm.

Since this learning technique makes effort to approximate the provided output based on the inputs, it is called “supervised machine learning”. Supervised learning algorithms are widely used in situations where connections between inputs and outputs are stable and consistent over time. In such situations, supervised machine learning algorithms are able to use rules acquired in the exemplary data to predict the outputs of new the inputs whose outputs are unknown (Kotsiantis et al., 2007).

Depending on the method to discover connections between inputs and outputs, supervised learning algorithms can be divided in to logic-based algorithms (e.g. decision tree model), perceptron-based algorithms (e.g. artificial neural networks), statistical learning algorithms (e.g. linear discriminant analysis, naive Bayes, Bayesian Networks), instance-based algorithms (e.g. k-Nearest Neighbor), and support vector machines (SVM). Whereas the objective of this thesis is not to enhance supervised machine learning algorithms, some of them are used to discover contextual information. Therefore, we provide a brief introduction of the algorithms discussed in the thesis.

**Decision tree (DC)** models sort data points based on their feature values (i.e. inputs). The underlying assumption is that data points with different outputs possess different val-
ues in at least one of their features (Safavian and Landgrebe, 1990). Features are successively used according to their ability to separate data. Thus, one may acquire $N+1$ classes if the dataset has $N$ features. Training data are separated into several subsets. A decision model studies all the subsets to modify its classification rule until all the data points are sorted as desired. Decision tree models generally perform well when dealing with categorical or discrete features.

An artificial neural network (ANN) is a computational model inspired by the natural neuron networks. It is used to estimate functions which are dependent on a large number of inputs (Zhang et al., 1998). Each ANN possesses many artificial neurons interconnected in a way that allows them to receive inputs, compute them, and deliver outputs. Basically, an artificial neuron receives input, adjust its value using the weight (i.e. a factor determining the strength of the input), and compute the intermediary output using an activation function. The intermediary output is passed to the next neuron(s) until an output is produced. The computed output is compared to the desired output to find errors. An ANN model is iteratively modified until the error is acceptable.

A Bayesian Network (BN) can be represented by a directed acyclic graph (DAG) consisting of a set of nodes (i.e. variables) and edges (i.e. the conditional dependencies between variables). The building of a BN model is composed of two steps: determining the structure of DAG (i.e. the dependency between nodes) and learning the parameters (i.e. the probability distribution of the node upon its parent nodes) for each node (Friedman et al., 1997). When the structure is known and the training data is complete, one can use maximum likelihood estimation (MLE) method to solve the parameters. When the structure is known and the training data is incomplete, Expectation Maximization (EM) algorithm can be considered. When the structure of a network is unknown, one may
introduce a scoring function to assess the fitness of the structure with respect to the desired training data and look for “better” structures based on the score function. BN algorithms usually perform better in the learning cases where there are less variables and variables are discretized in advance.

**k-Nearest Neighbor** (kNN) is a type of instance-based learning which assumes that data points having similar properties generally exists in close proximity to each other (Cover and Hart, 1967). Based on this assumption, the label of an unclassified data point can be determined by identifying the most common label of its k nearest classified neighbors. Similarly, the output of an unseen data point can be predicted by (weighted) averaging the output of its k nearest neighbors. Taking classification as example, the process is composed of the following steps. First, the algorithm computes the relative distances from an unclassified data point to all the classified data points. Second, it selects k nearest neighbors based on their distances to the unclassified data point. In this step, many distance metrics such as Minkowski distance and its variant \( p=1: \text{Manhattan distance}; p=2: \text{Euclidean distance}; p=\infty: \text{Chebyshev distance} \) can be used, and a guiding principle is that a distance metric should be able to minimize the distance between similar data points and maximize the distance between different data points. Third, the algorithm identifies the most frequent label (among the k nearest neighbors) and assigns it to the unclassified data. The quality of the kNN models is dependent on the selection of features, the value of k, and the selection of distance metric. Hence, the method involves a huge amount of tuning work so as to provide sound results.

**Support vector machines** (SVM) are a group of algorithms aiming at separating training data in a multi-dimension space into two groups by finding an optimum hyperplane that can maximize the distance from the hyperplane to the training data of both sides.
Taking Figure A.2a as an example, once the optimum hyperplane (red continued line) is identified, data points that lie on its margins (blue dashed line) are known as support vectors. Meanwhile, other data points are ignored. SVM algorithms have several advantages. (i) They are able to accept misclassified data (i.e., the outliers) from the training dataset by introducing slack variables to the constraints. (ii) They can classify the linear-inseparable data by mapping them to a higher-dimensional space and finding a hyperplane there (Figure A.2b).

![Figure A.2. Support Vector Machine](image)

**A2.2 Semi-supervised machine learning.**

Sometimes acquiring outputs for data points can be either an expensive and difficult task. It is quite common to receive datasets with incomplete output information and in practice the incomplete data points usually account for the majority of the dataset. In such cases, the objective of learning is to predict the missing outputs for the data points. Taking the classification task as example, the objective of learning is to predict the label of the unlabeled data. To the aim, one may have two options.
The first option is to base a prediction model on labeled data. Accordingly, the knowledge discovery task follows the supervised learning process. However, this option is likely to produce an inaccurate prediction model when the labeled data (i.e. the features-label pairs) cannot represent the population of the whole dataset. Taking Figure A.3 as an example, the red dashed line in the left panel illustrates a classifier we might adopt when examining two data points labeled “A” and “B” respectively. After adding unlabeled data points (see the right panel), two data clusters can be recognized and the original classifier ends up attributing certain data points to the wrong cluster. Obviously, the blue dashed line represents a better classifier which is capable of separating two clusters correctly.

In view of such issue, one may consider the second option, which is to use large amount of unlabeled data, together with the labeled data, to build prediction models (Zhu et Goldberg, 2009). Since the technique uses unlabeled data to enhance the accuracy of the prediction model, it is called “semi-supervised machine learning”.

Figure A.3. Classifier Before and After Using Unlabeled Data
In order to use the semi-supervised machine learning technique, one must assume that the data to be analyzed comply with certain distribution structure. Depending on the type of problem to be addressed, the assumption can be expressed in three ways.

- **Smoothness Assumption**: If two data points belong to the same high-density region and their inputs are close, then so should be their corresponding outputs.

- **Cluster Assumption**: data points in the same cluster are likely to have the same class label.

- **Manifold Assumption**: the high-dimensional data lie on a low-dimensional manifold.

Based on these assumptions, one may consider using self-training, transductive support vector machines, graph-based approaches, or generative models to solve semi-supervised learning problems (Zhu, 2005). We present two types of algorithms that are discussed in this thesis.

**Self-training algorithms** use their predictions to train themselves. (1) The algorithm builds a prediction model based on the training set consisting of labeled data. (2) The model classifies the unlabeled data. (3) The most confident predictions are added to the training set. (4) The model is retrained. To minimize the impact of misclassified data, one may filter the less confident predictions by setting a threshold.

**Transductive support vector machine (TVSM) is an extension of the standard SVM.** TVSM aims at identifying a model to classify the unlabeled data in a way that a linear boundary between two data classes has the maximized margin for both the labeled and the unlabeled data. Intuitively, the boundary usually passes through a low-density region where fewer labeled and unlabeled data points can be found.
A2.3 Unsupervised machine learning.

There are also situations where datasets contain nothing but unlabeled data points. Accordingly, it becomes important to explore the dataset and find some underlying structure or pattern within (Hastie et al., 2009). To achieve this objective, one may consider to segment data into various clusters, which contains data points of similar features. Due to the fact that the learning objective is to discover new knowledge rather than approximating the provided answers, the learning process is described as unsupervised learning. Consequently, algorithms developed to perform such learning tasks are called unsupervised machine learning algorithms.

Unsupervised machine learning algorithms have a wide range of applications in practice. (1) It is able to recognize the main patterns hidden in a dataset, which may suggest the direction of the subsequent research. (2) It allows users to identify the major clusters in a large dataset. Thus, the nature of a cluster can be learned by studying the associated common features. (3) It allows us to discover the features that can effectively separate clusters. Such knowledge is useful for the data classification task. Depending on the way to separate data, unsupervised machine learning algorithms can be classified into several types. We present the algorithms discussed in this thesis.

**Hierarchical algorithms** use previously identified clusters to find successive clusters. The divisive approach starts with a whole dataset as one cluster and proceeds to divide it into smaller clusters. It aims at maximizing the distance between separated clusters. The agglomerative approach begins with each data point as a cluster and merges them into larger clusters. It seeks to minimize the distance between grouped clusters.
Partitional algorithms seek to aggregate data points according to the number of clusters specified by users. The k-means clustering algorithm is one of the commonly used partitional algorithm. The algorithm begins data clustering by randomly choosing k data points (k is specified by the user of the algorithm) as the centroids. Next, it assigns the remaining data points to their nearest centroid, based on the chosen relative distance metric (e.g. Minkowski distance and its special cases). Then, the algorithm re-computes the k centroids using the new cluster membership information. Afterwards, data points are re-assigned to the (new) cluster centroids. The process repeats until the clustering membership does not change any more.
References


Corfu, Greece (pp. 2-7).


in recommender systems. ACM Transactions on Computer-Human Interaction (TOCHI), 17(1), 5.


Couch, E. (2013) DOES REMARKETING WORK, OR IS IT JUST ANNOYING? Available at: http://www.ppcher.com/does-remarketing-work-or-is-it-just-annoying/


Definition of “Intention”: https://ahdictionary.com/word/search.html?q=intention

Definition of “Personalization”: https://ahdictionary.com/word/search.html?q=personalization


Gartner’s IT glossary: http://goo.gl/OTj8y5


Introduction to Psychographics (Vol. 5). Routledge.


Heidelberg.


User Privacy. IEEE Computer, 45(11), 87-92.


Adapted Interaction, 22(4-5), 441-504.


Liao, Z., Jiang, D., Chen, E., Pei, J., Cao, H., & Li, H. 2011. Mining concept sequences from large-scale search logs for context-aware query suggestion. ACM Transactions on Intelligent Systems and Technology (TIST), 3(1), 17.


Merriam-Webster dictionary: www.merriam-webster.com/dictionary/personalization


Nguyen, T. T., Kluver, D., Wang, T. Y., Hui, P. M., Ekstrand, M. D., Willemsen, M. C., &


Rodden, T., Cheverst, K., Davies, K., & Dix, A. (1998, May). Exploiting context in HCI design for mobile systems. In Workshop on human computer interaction with mobile
devices (pp. 21-22).


the workshop on context-aware movie recommendation (pp. 34-40). ACM.


data using belief nets algorithms. In 2006 18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'06) (pp. 497-504). IEEE.


for context-aware recommendations of mobile applications. In Data Engineering Workshop, 2007 IEEE 23rd International Conference on (pp. 871-878). IEEE.


Zheng, Y., Burke, R., & Mobasher, B. (2014, March). Splitting approaches for context-
Symposium on Applied Computing (pp. 274-279). ACM.

aware recommendation engine. In 2015 IEEE International Conference on Data Mining
Workshop (ICDMW) (pp. 1668-1671). IEEE.

recommendation. In International Conference on Web Information Systems Engineer-
ing (pp. 431-447). Springer International Publishing.

Zhu, H., Chen, E., Yu, K., Cao, H., Xiong, H., & Tian, J. (2012, December). Mining per-
sonal context-aware preferences for mobile users. In 2012 IEEE 12th International Con-
ference on Data Mining (pp. 1212-1217). IEEE.


lectures on artificial intelligence and machine learning, 3(1), 1-130.