

Association for Information Systems AIS Electronic Library (AISeL)

MCIS 2016 Proceedings

Mediterranean Conference on Information Systems
(MCIS)

2016

Garment Recommendations for Online and Offline Consumers

Maria Skiada

ELTRUN: The e-Business Research Center, Athens University of Economics and Business, Greece, mskiada@aueb.gr

George Lekakos

Athens University of Economics and Business (AUEB), glekakos@aueb.gr

Sofia Gkika

ELTRUN: The e-Business Research Center, Athens University of Economics and Business, Greece, gkikas@aueb.gr

Cleopatra Bardaki

ELTRUN, Department of Management Science and Technology, Athens University of Economics & Business, Greece, cleobar@aueb.gr

Follow this and additional works at: <http://aisel.aisnet.org/mcis2016>

Recommended Citation

Skiada, Maria; Lekakos, George; Gkika, Sofia; and Bardaki, Cleopatra, "Garment Recommendations for Online and Offline Consumers" (2016). *MCIS 2016 Proceedings*. 62.

<http://aisel.aisnet.org/mcis2016/62>

This material is brought to you by the Mediterranean Conference on Information Systems (MCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in MCIS 2016 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

GARMENT RECOMMENDATIONS FOR ONLINE AND OFFLINE CONSUMERS

Completed Research

Skiada, Marianna, ELTRUN:The e-Business Center, Department of Management Science and Technology Athens University of Economics and Business, Athens, Greece,
mskiada@aueb.gr

Lekakos, Georgios, ELTRUN:The e-Business Center, Department of Management Science and Technology Athens University of Economics and Business, Athens, Greece,
glekakos@aueb.gr

Gkika, Sofia, ELTRUN:The e-Business Center, Department of Management Science and Technology Athens University of Economics and Business, Athens, Greece,
gkikas@aueb.gr

Bardaki, Cleopatra, ELTRUN:The e-Business Center, Department of Management Science and Technology Athens University of Economics and Business, Athens, Greece,
cleobar@aueb.gr

Abstract

Recommender Systems have obvious influence in environments where data size exceeds the capabilities of any user to fully explore the available choices in the store (physical or on-line). Many algorithms and techniques have been used to help recommending useful and interesting items to users. If the user is unidentified, the process is even harder as there are no historical or other data to use as input. Association rules is a popular technique used for many purposes in Recommender Systems such as for building more robust systems, improving quality of recommendations; and even addressing fundamental limitations of recommender systems and, generally, large datasets, e.g. sparsity and cold start. At the same time, efforts have been made to fully understand if and how differently customers are behaving in an online and in a physical environment. This work tries to combine the two efforts. We use association rules to provide recommendations to customers, as well as understand who the customer is, what her needs are and what is her mentality when entering a physical store or the corresponding e-shop. To fulfill our goal, we used descriptive statistics along with Association Rules analysis of the POS transactional data on basket data level and historical data level.

Keywords: Recommender Systems, Association Rules, Basket Analysis, physical vs on-line environment

1 Introduction

The increase of data from business operations and consumer interactions with physical and online stores paved the ground for the emergence of Recommender Systems changing the marketing approach regarding electronic commerce and access to information and personalized services. Recommender Systems (RSs) are software tools and techniques suggesting to users interesting for them items (Resnick and Varian, 1997). These systems have obvious influence in environments where data size exceeds the capabilities of any user to fully explore them; and have become an integral part of many online stores, such as Amazon, Asos and Netflix.

Recommender Systems incorporate in use many algorithms/techniques to address the problem of suggesting interesting items to users. These algorithms are separated into two major groups: a) memory-based algorithms, such as Collaborative and Content Based Filtering and b) model-based techniques, such as clustering, association rules and Bayesian models. Memory-based algorithms approach the recommendation problem by using the entire database, meaning that they use all the available historical data to make predictions (Breese et al., 1998). In contrast, model-based techniques involve building models based on the available dataset. Model-based approaches develop a "model" to make recommendations, without having to use the complete dataset every time a recommendation is requested. This approach is superior in terms of speed and scalability, compared to memory-based algorithms. However, it suffers from inflexibility because building a model is often a time- and resource-consuming process, while less accurate predictions are produced because only a part of the dataset is used to build the model.

However, because of scalability and sparsity problems (in such systems, even the most active users will not have bought/evaluated more than 1% of products), model-based algorithms are used very often. The extraction of association rules is one popular model-based technique implemented in this field that attempts to discover patterns of products that are purchased together. The generated rules can be used for many and different purposes including marketing, inventory management, etc. In the field of recommender systems, it has also been used in several ways, e.g. for shielding against attacks (Sandvig, 2007), improving scalability for memory-based algorithms (Sarwar et al. 2000), improving quality of recommendations combined with Collaborative Filtering (Sarwar et al. 2000) and user's activity (Davidson et al., 2010); and helping with cold start problem where the user has not rated/purchased enough items to build a complete profile (Shawet al., 2010) etc.

Much research has been conducted concerning the online customer behaviour in different contexts, in order to identify what motivates customers to shop online (Donthuand Garcia, 1999; Dubelaar et al. 2003), distinguish customer personas (Bellenger and Korgaonkar, 1980) and discover differences in shopping process between the two genres (Dittmar, 2004; Huang and Yang, 2010) using questionnaires of interviews.

The primary contribution of this paper is to show how the association rules method, supported by descriptive statistics, can process basket and historical data from real POS and online transactions in order to mine patterns of buying behaviour in the different environments and between the two genders. The generated association rules can become the cornerstone of a recommender system that proposes new garments to the customers based on the generated knowledge of their shopping preferences. The experiments have been executed with transactional (sales) data from a European brand in clothing industry. Specifically, we obtained and used data from two different stores (the e-shop and one physical store) to generate the association rules.

The remaining of the paper unfolds as follows. In Section 2, a review of previous research for online customer behaviour and differences between the genres in purchase process is presented. Section 3 describes the available datasets and discusses the utilized approach. Then, the results of each environment are presented in Section 4. Finally, we summarize our contributions with suggestions for future research in Section 5.

2 Related Work

Nowadays, consumers can choose to purchase goods from a variety of different channels including traditional stores, e-shops, TV-shopping, catalogues etc. One of the most common product category that is sold online and offline are clothes. From 2004 to 2009 retail sales increased globally \$144 billion (Euromonitor, 2010). Most of the times retailers have both online and offline presence (Mintel, 2012). Nevertheless, shopping garments online is risky (Kim and Forsyth, 2009) since the customer cannot touch or try on the items before purchasing them; and the items are complex since they are described by many parameters such as size, textile, color, body fit etc.

Studies have shown that online shopping behavior is affected by demographics, channel knowledge, perceived channel utility and shopping orientations (Li et al., 1999; Hargittai et al., 2008; Cho and Workman, 2011; Ruane and Wallace, 2013; Scarpiet al., 2014). Shopping orientation refers to a customer's general attitude about shopping (Brown et al., 2003); and typically customers are distinguished between economic shoppers that act as “problem solvers”; and recreational shoppers that enjoy shopping (Bellenger and Korgaonkar, 1980). Problem solvers shop in an effective and quick way in order to fulfil their needs with the less possible effort, while recreational shoppers seek for “fun, fantasy, arousal, sensory stimulation, and enjoyment” (Hirschman and Holbrook, 1982)

Donthu and Garcia (1999) pointed that there is distinction between “Internet shoppers” and “no shoppers” in the United States. They define users who actually make online purchases as Internet “shoppers” and all the other Internet users as “no shoppers.” Internet “shoppers” are found to be older and with higher income (Donthu and Garcia, 1999; Barnes et al., 2007). Online shopping has been related to privacy concerns, credit card thefts and problems concerning the order that the customers has done (wrong items, problem with return process) etc. These concerns increase customer's risk and influence her willingness to buy online (Dubelaar et al. 2003). Donthu and Garcia (1999) claimed that online shoppers have lower levels of risk aversion, find also seeking more convenient and are more innovative and impulsive than no shoppers.

Thus, it is essential to discover and define which customers choose to shop online; factors such as demographics, education, personality (Burke, 2002;) and Need For Touch (NTF) (Cho and Workman, 2011) seem to be influencing. On the one hand, younger people, who are interested in technology, always want to discover new items, search for more product characteristics and want to compare their alternatives before they purchase (Wood, 2002). On the other hand, older customers believe that the benefits from online shopping are less than the effort they need to make in order to learn how to perform online shopping and therefore they avoid it (Ratchford et al., 2001), leading to low access and usage (Karahasanović, 2009). Oppenheim and Ward (2006) suggested that the primary reason people shop over the internet has now switched to convenience, while before it was almost about finding the lowest price.

Concerning gender, men use various types of technology during the shopping process and studies have shown that men are more likely than women to purchase products/ services from the Internet (Korgaonkar and Wolin, 1999), but women that do choose to shop online, shop more frequently than them (Burke, 2002; Li et al., 1999). Statistics have shown that clothing is one of the very few categories where women buy more than men online (Statistics Denmark, 2007); and that they are more involved in fashion (O'Cass, 2004). Female customers focus on the enjoyable process of shopping and, therefore have hedonic orientation (Scarpiet al., 2014) and are “want based” (Cho and Workman, 2011). Whereas men focus on shopping goods with the less effort (Dittmar et al. 2004), are more utilitarian and “need based”. So women are more “shopping for fun”-oriented compared to men who are more “quick shoppers”, which in turn may have an effect on their respective perceived benefit or barriers when purchasing clothing online, suggesting that men have more positive attitudes toward online buying than women do (Hansen and Jensen, 2009). On the other hand younger woman (Generation Y) that are equally skilled to men (Hargittai et al., 2008) use online stores for research and evaluation

before visiting physical stores (Ruane and Wallace, 2013). Customers who shop for fun and entertainment are more loyal to physical stores (brick and mortar stores) while price is important to both online and offline channels (Scarpinet al., 2014)

3 Association Rules with Basket and Wardrobe data analysis

Association rules technique aims to extract rules such as "if customer purchases A, she will also buy B". So, it can be seen as a method of discovering interesting relations within datasets. The algorithm searches the dataset to find items that appear together. The algorithm then groups these items into sets of items and generates rules from them. These rules are used to predict the presence of an item (or items) in the database, based on the presence of other specific items. Association Rules is a useful data mining technique for analyzing and predicting customer behaviour. They play an important part in basket data analysis, product clustering, catalogue design and store layout.

In a real case scenario, we want to recommend items to a customer for whom we know nothing before she uses her loyalty card. Without usage of the loyalty card, we do not have customer's id and consequently no access to customer's preferences, previous purchases or demographical data. The only thing that can be used as input is what she tries on in the fitting room, holds in hands when she asks for recommendation or has visited in the e-shop during the current session. Hence, based on those selections and on earlier generated rules from the entire dataset, we can recommend other items. Additionally, these rules can be used for making more efficient promotional actions, where the included products are not chosen randomly but based on them.

The analysis can be done in terms of two dimensions: Basket analysis and historical data analysis. The generated rules can regard specific items (lowest unit of analysis) or product category. Item analysis will not lead to significant results when applied to a sparse dataset which includes too many available items and customers. In contrast, analysis per product category can produce long lasting and significant rules.

Basket Analysis aims to answer the question what the customer buys during the shopping journey. To discover this, basket identification is necessary in order to identify which product categories were purchased during each single shopping journey. At this point, customer identification is not really needed because we focus on the basket and not on the customer.

Respectively, historical data analysis focuses on the customer and her purchases throughout her 14 month journey, trying to discover what she will buy, maybe not on the same basket but in a wider time space. In this case, all the purchased categories build her personal wardrobe from the specific store. Garments are usually long lasting products that are used many times. So what was she bought before can influence what she will be buy in the future.

The generated association rules from the implementation of Apriori algorithm (Agrawal and Srikant, 1994; Agrawal et al., 1993) can be used for predictions. These rules predict which items are likely to appear together (e.g. be bought together) given the presence of the items on the left side of each rule. It also provides information about the probability, the support and the importance of each rule in order to decide if it is useful. For example, if X and Y represent two garments in a basket, support is the number of cases in the dataset that contain the combination of items, X and Y. While probability (confidence) represents the fraction of cases in the dataset that contain X and that also contain Y.

The importance (lift) of an itemset is calculated as the probability of the itemset divided by the compound probability of the individual items in the itemset. For example, if an itemset contains {A, B}, all the cases that contain this combination A and B are counted, and divided by the total number of cases, and then normalizes the probability. The importance of a rule is calculated by the log likelihood of the right-hand side of the rule, given the left-hand side of the rule. For example, in the rule If {A} Then {B}, the ratio of cases is calculated with A and B over cases with B but without A, and then normalizes that ratio by using a logarithmic scale. If importance is > 1 , this lets us know the degree to

which those two occurrences are dependent on one another, and makes those rules potentially useful for predicting the consequent in future data sets. Importance (lift) is the strongest and most significant metric in Association Rules' evaluation.

The algorithm's purpose is double, to discover patterns of purchase and understand customers' behaviour in online and offline context for the two genres. These rules can be useful for the following cases: recommend products to unidentified users (eshop, physical store etc), build more effective promotional activities by knowing which products are related, and understand how customers behave in each channel. Finally, long tail products that are difficult to be sold can be sold along with others using the proper association rules.

4 Experiment

4.1 Data description

We want to compare the extracted association rules of the e-shop and the physical store. The ultimate aim is to analyze how these rules express the datasets and provide insights on customers' behavior that can enable a recommender system that recommends items based on these rules.

A European retailer of men and women has provided data, within the context of a EU project. Specifically, we have POS datasets from two stores: the e-shop and a physical store. The available data concern the time period between 1/7/2014 and 30/9/2015, namely 14 months. The e-shop's dataset consists of 535029 rows, which belong to 98567 distinct baskets. From these baskets, 70% belong to identified customers as the remaining 30% cannot be bounded to a specific customer because no loyalty card was used. Respectively, the physical store's dataset consists of 506895 rows, which belong to 89916 distinct baskets, from which only 3,5% are not associated with an identified customer. In our experiments, we used only transactional data that are associated to customers. Table 1 contains the demographical profile of the customers. Tables 2 and 3 present descriptive statistics for purchases of the two stores.

	Physical Store	Eshop
Distinct Customers	46671	52733
Average Birth Year	1953	1961
Women	73,54%	78,24%
Men	25,56%	16,54%
Not Available	0,90%	5,22%

Table 1 Demographical profile Customers

According to Table 1, the e-shop has more distinct customers than the physical store, who are on average 8 years younger and are mainly women.

	Number of Visits	Distinct Purchased Categories	Number of Purchased Items	Items Per Basket
All	3,2	7,0	10,9	3,34
Men	2,9	6,3	10,7	3,64
Women	3,3	7,3	11,0	3,23

Table 2 Descriptive Statistics of Physical store

	Number of Visits	Distinct Purchased Categories	Number of Purchased Items	Items Per Basket
All	1,7	3,8	6,9	3,78
Men	1,6	2,9	7,2	4,4
Women	1,7	3,8	6,5	3,63

Table 3 Descriptive Statistics of Eshop

Customers visit more often the physical store, where they purchase more items from more product categories -almost double in number - but they buy fewer items per baskets (basket volume) compared to the e-shop. In the e-shop, customers make fewer transactions, select items from less product categories, but the volume of their basket is bigger. Comparing men and women, we notice that women visit more times the physical store than men, but their online visits are almost equal. Women's baskets include wider variety of garment categories in both environments, but their online volume of basket is lower than men's. However, men make fewer visits on the physical store but they buy more items each time and have a strong online presence.

4.2 Results

We generated the association n rules (based on Apriori algorithm) for both available datasets, only for the identified customers. The rules were generated per store (e-shop and physical) based on basket data and historical / wardrobe data. We set the same thresholds for both datasets Importance: (lift) over 1 and probability over 70%. The following paragraphs present indicative rules for each case.

Eshop

We extracted association rules for the e-shop, first on basket level and, then, on historical data level:

Probability	Importance	Rule
81 %	2,88	Brand Women Accessories, Brand Women Blouse->Brand Women Jacket Leather
78 %	2,84	Brand Women Jacket Leather, Brand Women Blouse->Brand Women Accessories
70 %	2,54	Men Suits blazer(from suits)-> Men Suits pants(from suits)
70 %	2,45	Men traditional Pants, Men traditional suits-> Men traditional accessories
75 %	2,36	Men Suits classic vest(from suits), Men Suits classic pants (from suits)-> Men Suits classic blazer (from suits)
100 %	1,63	Women traditional Skirt, Women underwear -> Women traditional Jacket
68%	1,35	Large Size Cardigan , Large Size tops->Large Size Shirts

Table 4 Indicative eshop's basket Rules

Apriori algorithm generated 60 rules for basket data (Table 4). The most important rules are related to brand women garments, indicating that customers trust brands for a variety of product categories. Many and significant rules follow in the list related to men's wardrobe and few that combine men's clothing with women's. In addition, we found association between large size clothes. In the bottom of the rules list, we found associations concerning garments of women's wardrobe.

Probability	Importance	Rule
73 %	3,10	Men Suits classic blazer (Large Sizes) (from a suit) -> Men Suits classic pants (Large Sizes) (from a suit)
82 %	2,95	Brand Women Jacket Leather, Brand Women Blouse -> Brand Women Accessories
72 %	2,89	Brand Women Accessories, Brand Women Blouse -> Brand Women Jacket Leather
92 %	2,54	Men Suits classic vest(from suits) , Men Suits classic pants (from suits) -> Men Suits classic blazer (from suits)
79 %	2,50	Men Suits classic blazer (from suits)-> Men Suits classic pants (from suits)
72 %	2,67	Men Suits blazer(from suits) -> Men Suits pants(from suits)
100 %	1,83	Women traditional Skirt, Women traditional Dress -> Women traditional Jacket
86 %	1,51	Large Sizes Vest, Men underwear -> Large Sizes Shirt
71 %	1,50	Plus Size Brand Women Dress, Plus Size Brand Women Jeans ->Plus Size Brand Women Pants
81 %	1,35	Specials Other, Brand Leisure ->Brand House
85 %	1,33	Socks Men Socks, Men Accessories-> Men underwear
75 %	1,32	Brand Leisure, Men Knitwear pullover ->Brand House
79 %	1,31	Brand (men clothing) T-shirt, Brand Half arm Men Polo Shirt -> Men T-Shirt / Sweater shirt Men Polo Shirt
71 %	1,31	Home textiles House , Brand Leisure ->Brand House

Table 5 Indicative eshop's Historical data Rules

On the long run (historical data level), more association rules -235 in number- are generated since a bigger “historical data basket” is build (Table 5). Most significant are the rules related to brand clothes for women’s and men’s garments. Historical data analysis revealed that brands have loyal customers among men too, who repetitively buy from these bands. Traditional clothes generate few, but important rules; and plus size clothes seem to persist over time. Rules including both women’s and men’s clothes are limited; and small items for the household are appearing for the first time. Last, but numerous, in the rules list ordered by importance we found rules associating ladies’ garments.

Physical Store

We generated the association rules (based on the Apriori algorithm) for the physical store to produce important itemsets; first on basket level and, after, on historical data level. In both cases, more rules are generated because the number of distinct purchased garment categories and the number of purchased items are higher than those of the e-shop’s although physical store had fewer customers during the available 14 months of sales.

Probability	Importance	Rule
100 %	3,28	Large SizesMen modular pants, Ties -> Large SizesMen modular jacket

100 %	3,03	Men Suits classic pants(Large Sizes) (from a suit), Men shirts -> Men Suits classic blazer(Large Sizes) (from a suit)
100 %	1,88	Men traditional clothes Pants, Women traditional clothes Accessories-> Men traditional clothes suits
87 %	3,13	Large Sizes Men modular jacket, Men shirts -> Large Sizes Men modular pants
83 %	2,44	Men Suits Men pants(from suits), Men shirts -> Men Suits Men blazer(from suits)
83 %	2,31	Men Suits pants(from suits), Socks Men Socks-> Men Suits blazer(from suits)
81 %	1,96	Women outfits Skirt short festive, Women outfits Blouse festive-> Women outfits Blazer Festive

Table 6 Indicative physical store's Basket Rules

Physical store's rules -40 in total- are mainly about man's wardrobe (Table 6). Many of the rules show baskets including both men's and women's garments. Men's products are more in number and sometimes basic woman items are found in between denoting the presence of woman. Rules that concern entirely women's items are limited. Plus size garments have intense presence and household products appear in between both gender baskets. Finally, we noticed that no significant rule is related to brand clothes for either sex.

Probability	Importance	Rule
73 %	3,10	Men Suits classic blazer(Large Sizes) (from a suit)-> Men Suits classic pants(Large Sizes) (from a suit)
82 %	2,95	Brand Women Jacket Leather, Brand Women Blouse->Brand Women Accessories
72 %	2,89	Brand Women Accessories, Brand Women Blouse->Brand Women Jacket Leather
72 %	2,67	Men Suits Men blazer(from suits)-> Men Suits Men pants(from suits)
100 %	2,64	Men Suits classic pants (from a suit) , Women shirt -> Men classic blazer (from a suit)
100 %	1,83	Women traditional clothes Skirt, Women traditional clothes Dress-> Women traditional clothes Jacket
77 %	1,71	Brand Men Pants short, Jeanswear Shirt Jeanwear-> Jeanswear Short Pants jeanwear
71 %	1,61	Brand Women Jacket, Women outfits Knit Tops-> Women cardigan
86 %	1,51	Large Sizes Vest, Men underwear-> Large Sizes Shirt
71 %	1,50	Plus Size Brand Women Dress, Plus Size Brand Women Jeans->Plus Size Brand Women Pants
83 %	1,35	Brand Seasonal items, Women outfits Women Dress festive->Brand House
81 %	1,35	Specials Other, Brand Leisure->Brand House
85 %	1,33	Socks Men Socks, Men Accessories Men accessories-> men underwear

Table 7 Indicative physical store's Historical data Rules

Respectively, historical data analysis extracted 584 rules (Table 7), in which men's clothing has very strong presence. These rules concern only men's apparel, but have high importance or combine garments of both sexes. Brands and plus size garments also have strong presence. Last in significance are associations concerning traditional clothes and garments related to women's wardrobe.

5 Discussion

Many apparel retailers have presence in both online and offline environments and aspire to implement recommender systems in their stores, in a way to help customers deal with the wide variety of products and improve their shopping experience. But when little or no input is available for the customer shopping preferences, the recommendation process is becoming very demanding. To address this challenge, the garments that the customers choose to try on in the fitting room or the garments that they visit online can be useful input for the model-based technique of Association Rules. Each rule consists of two parts the left and the right. In presence of the left, the right part occurs within a specified importance. So, the selected items by the customer can be used as the left part and based on pre-generated rules we can recommend accordingly the right.

During our research, we applied association rules on real POS data from a physical store and the e-shop of a European apparel retailer. We applied this technique for each dataset on basket data (where basket identification uniquely describes each case); and on historical data (where all transactional data of each Customer is used as a unique case) to generate rules that associate products.

This paper contributes with the comparative analysis of the consumers' shopping behaviour using POS data from a physical and online store of an apparel retailer, while taking into account the factor of gender. The generated rules/ knowledge, expressing the customers' shopping behaviour, can feed a recommender system that proposes new garments to the customers based on these rules. The experiments have been executed with transactional (sales) data from a European brand in clothing industry. Specifically, we obtained and used data from two different stores (the e-shop and one physical store) to generate the association rules between garments. This work doesn't use questionnaires or interviews to capture the customers' shopping behaviour. On the contrary, the results derive from a data mining technique (Association Rules) and descriptive statistics on a real 14month POS dataset.

Customers of the e-shop are younger than the average traditional store customer, which is it can be attributed to the fact that younger people are more familiar with technology. Physical store has more customers who visit it more often, but buy smaller baskets compared to the e-shop. For both environments, woman's participation is larger but, on average, she has stronger presence in the physical store which she visits more than man and more than the respective eshop. Men are more open to online experience and they prefer to buy massively, while women enjoy the shopping experience and prefer buying fewer items per basket and visit more frequently the store. We assume that this is because she does "shopping for fun" (Hansen and Jensen, 2009), while man who is a "quick shopper" (Hansen and Jensen, 2009) that has stronger online presence.

For the e-shop it can be noticed that women brands are significant for both basket and historical rules, suggesting that customers express their trust and likeness by buying repeatedly product from different categories. Moreover, numerous rules exist related to men's garments. Few of them include a woman garment, indicating the presence of a woman. The large number of rules include only men clothing denoting that men shop to build their wardrobe with supplementary garments e.g. blazer -> suit -> pants and their needs are more clear leading to specific choices. So, these rules are applicable for recommendations, when a man or a married woman is identified in any way. Historical rules outnumber basket rules since they are constructed by bigger cases. They repeat the majority of basket rules and they discover customers' tendency to buy household items in different baskets. In the long term, men also express their preference on some brands. So men are less brand-oriented compared to women and household items may not be the main need of customers (they probably express a pattern of spontaneous purchases). Rules for women garments are less significant, denoting that women buy larger garment variety and they don't have the willingness to make online purchases (Hansen and Jensen, 2009) or a structured plan on what to buy from the eshop for themselves (i.e. they are occasional online buyers).

Physical store's rules are mainly about men shopping. Most of the time, the rules include men's products and, often, a woman item is found in between implying her presence. So, when a man visits the physical store, he is probably going to buy (big basket) and he will probably also buy something for a woman, or when a woman purchases stuff for a man she will buy something for herself too. Rules concerning women items are very limited, as they have smaller basket volume and along with their items they also buy things for their house or other male family member. Brands do not appear on basket rules, but they appear in history rules, having the same pattern with traditional clothing and household items. Plus size garments are significant in basket and historical rules, which makes sense since body type doesn't change easily.

The clearest differences between the online and offline environments are that brands and traditional clothing appear more frequently and with greater significance online on both basket and historical level. Brands seem trustful solution to overcome online limitations and traditional clothing are not so important to visit the store only for purchasing them. Moreover, in the online environment fewer rules include men's and women's clothing at the same time, suggesting that customers tend to buy items for themselves as opposed to many mixed baskets and historical data of the physical store. Women seem more brand-oriented than men and, in total, they buy more items as they purchase things not only for personal usage but also for others (men, children, household).

Research results are based on the explanation of the association rules technique, descriptive statistics and literature. To go deeper and earn a greater understanding of them, questionnaires, interviews and focus groups should be conducted to discuss results and behaviors giving customers the chance to explain and compare results and conclusion for online vs offline behavior and differences of the two genres regarding shopping process.

Acknowledgements

The authors acknowledge Sensor-Enabled Real-World Awareness for Management Information Systems (SERAMIS) EU project partners (<http://seramis-project.eu>) for providing the necessary data for this study.

References

- Agrawal, R., Imielinski, T., and Swami, A. (1993). Mining Associations between Sets of Items in Massive Databases. In *Proc. of ACM SIGMOD* (pp. 207-216).
- Agrawal, R., and Srikant, R. (1994, September). Fast algorithms for mining association rules. In *Proc. 20th int. conf. very large data bases, VLDB* (Vol. 1215, pp. 487-499).
- Barnes, S. J., Bauer, H. H., Neumann, M. M., & Huber, F. (2007). Segmenting cyberspace: a customer typology for the internet. *European journal of marketing*, 41(1/2), 71-93.
- Bakewell, C., and Mitchell, V. W. (2006). Male versus female consumer decision making styles. *Journal of Business Research*, 59(12), 1297-1300.
- Bellenger, D. N., and Korgaonkar, P. K. (1980). Profiling the recreational shopper. *Journal of retailing*, 56(3), 77-92.
- Breese, J. S., Heckerman, D., and Kadie, C. (1998, July). Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence* (pp. 43-52). Morgan Kaufmann Publishers Inc.
- Broekhuizen, T. L., and Jager, W. (2004). *A conceptual model of channel choice: Measuring online and offline shopping value perceptions*. University of Groningen.
- Brown, M., Pope, N., and Voges, K. (2003). Buying or browsing? An exploration of shopping orientations and online purchase intention. *European Journal of Marketing*, 37(11/12), 1666-1684.

- Cho, S., & Workman, J. (2011). Gender, fashion innovativeness and opinion leadership, and need for touch: Effects on multi-channel choice and touch/non-touch preference in clothing shopping. *Journal of Fashion Marketing and Management: An International Journal*, 15(3), 363-382.
- Davidson, J., Liebold, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., ...and Sampath, D. (2010, September). The YouTube video recommendation system. In *Proceedings of the fourth ACM conference on Recommender systems* (pp. 293-296). ACM.
- Dittmar, H., Long, K., and Meek, R. (2004). Buying on the Internet: Gender differences in on-line and conventional buying motivations. *Sex roles*, 50(5-6), 423-444.
- Donthu, N., and Garcia, A. (1999). The internet shopper. *Journal of advertising research*, 39(3), 52-52.
- Dubelaar, C., Leong, M., and Alpert, F. (2003). Impact of interactivity on the stickiness of online gift stores. *Journal of Asia Pacific Marketing*, 2(2), 22-41.
- Euromonitor, 2010. Euromonitor International From Trade Sources/National Statistics
- Hansen, T., and Møller Jensen, J. (2009). Shopping orientation and online clothing purchases: the role of gender and purchase situation. *European Journal of Marketing*, 43(9/10), 1154-1170.
- Hargittai, E., & Walejko, G. (2008). The Participation Divide: Content creation and sharing in the digital age 1. *Information, Community and Society*, 11(2), 239-256.
- Holbrook, M. B., and Hirschman, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *Journal of consumer research*, 132-140.
- Huang, J. H., & Yang, Y. C. (2010). Gender differences in adolescents' online shopping motivations. *African Journal of Business Management*, 4(6), 849.
- Kim, J., & Forsythe, S. (2009). Adoption of sensory enabling technology for online apparel shopping. *European Journal of Marketing*, 43(9/10), 1101-1120
- Karahasanović, A., Brandtzæg, P. B., Heim, J., Lüders, M., Vermeir, L., Pierson, J., ...& Jans, G. (2009). Co-creation and user-generated content—elderly people's user requirements. *Computers in Human Behavior*, 25(3), 655-678.
- Korgaonkar, P. K., and Wolin, L. D. (1999). A multivariate analysis of web usage. *Journal of advertising research*, 39, 53-68.
- Oppenheim, C., and Ward, L. (2006, May). Evaluation of web sites for B2C e-commerce. In *Aslib Proceedings* (Vol. 58, No. 3, pp. 237-260). Emerald Group Publishing Limited.
- O'Cass, A. (2004). Fashion clothing consumption: antecedents and consequences of fashion clothing involvement. *European Journal of Marketing*, 38(7), 869-882.
- Resnick, P., and Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, 40(3), 56-58.
- Ratchford, B. T., and Debabrata Talukdar, M. S. L. (2001). A model of consumer choice of the Internet as an information source. *International Journal of Electronic Commerce*, 5(3), 7-21.
- Resnick, P., and Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, 40(3), 56-58.
- Ruane, L., & Wallace, E. (2013). Generation Y females online: insights from brand narratives. *Qualitative Market Research: An International Journal*, 16(3), 315-335.
- Sandvig, J. J., Mobasher, B., and Burke, R. (2007, October). Robustness of collaborative recommendation based on association rule mining. In *Proceedings of the 2007 ACM conference on Recommender systems* (pp. 105-112). ACM.
- Sarwar, B., Karypis, G., Konstan, J., and Riedl, J. (2000, October). Analysis of recommendation algorithms for e-commerce. In *Proceedings of the 2nd ACM conference on Electronic commerce* (pp. 158-167). ACM.
- Shaw, G., Xu, Y., and Geva, S. (2010). Using association rules to solve the cold-start problem in recommender systems. In *Advances in Knowledge Discovery and Data Mining* (pp. 340-347). Springer Berlin Heidelberg.
- Scarpi, D., Pizzi, G., and Visentin, M. (2014). Shopping for fun or shopping to buy: Is it different online and offline?. *Journal of Retailing and Consumer Services*, 21(3), 258-267.

Wood, S. L. (2002). Future fantasies: a social change perspective of retailing in the 21st century. *Journal of retailing*, 78(1), 77-83.