The Value of Personalised Recommender Systems to E-Business: A Case Study

M. Benjamin Dias
Unilever Corporate Research
Sharnbrook, Bedfordshire, UK
Benjamin.Dias@Unilever.com

Dominique Locher
LeShop.ch
Chemin du Dévent, 1024
Ecublens, Switzerland
Dominique.Locher@leshop.ch

Ming Li
Unilever Corporate Research
Sharnbrook, Bedfordshire, UK
Ming.x.Li@Unilever.com

Wael El-Deredy
School of Psychological Sciences
University of Manchester, UK
Wael.El-deredy@manchester.ac.uk

Paulo J. G. Lisboa
School of Computing and Mathematical Sciences
Liverpool John Moores University, UK
P.J.Lisboa@livjm.ac.uk

ABSTRACT

Recommender systems have recently grown in popularity both in e-commerce and in research. However, there is little, if any, direct evidence in the literature of the value of recommender systems to e-Businesses, especially relating to consumer packaged goods (CPG) sold in a supermarket setting. We have been working in collaboration with LeShop (www.LeShop.ch), to gather real evidence of the added business value of a personalised recommender system. In this paper, we present our initial evaluation of the performance of our model-based personalised recommender systems over the 21-month period from May 2006 to January 2008, with particular focus on the added-value to the business. Our analysis covers shopper penetration, as well as the direct and indirect extra revenue generated by our recommender systems. One of the key lessons we have learnt during this case study is that the effect of a recommender system extends far beyond the direct extra revenue generated from the purchase of recommended items. The importance of maintaining updated model files was also found to be key to maintaining the performance of such model-based systems.

Categories and Subject Descriptors

H.3.5 [**Information Systems**]: Online Information Services—*Commercial services*; J.7 [**Computer Applications**]: Computers in Other Systems—*Consumer products*

General Terms

Economics, Experimentation, Performance, Verification

RecSys'08, October 23–25, 2008, Lausanne, Switzerland.

Keywords

Case study, Business value creation, E-commerce, Performance evaluation, Personalisation, Recommender systems

1. INTRODUCTION

As a result of the increasing and ubiquitous use of the Internet, recommender systems are growing in popularity in e-commerce (e.g. www.Amazon.com, www.eBay.com, etc.). Concurrently, and most probably as a result, the area of personalised recommender systems has recently become an important research focus as well [1, 2, 3, 6, 4, 8, 5]. However, most of the research in this area has thus far been focused on the development of new and improved recommender algorithms. Many companies who use personalisation technologies (e.g. Amazon, e-Bay, etc.) often conduct their own unpublished evaluations in order to determine the value of these technologies to their business. However, there have been few published studies that provide direct evidence of the value of recommender systems to e-Businesses, apart from a very few exceptions, such as [7] based on the online bookstore Mitos (www.mitos.co.il). Furthermore, we have found no such published studies relating to consumer packaged goods (CPG) sold in a supermarket setting.

CPG items not only cover a wide and diverse range of categories, but are also purchased repeatedly, and hence are distinctly different to book sales (e.g. on Amazon.com, Mitos) and movie recommendations (e.g MovieLens). We have been very fortunate to collaborate with LeShop (www.LeShop.ch), who are pioneers in their field. Together, we have conducted a case study, in a real-world setting, with the aim of evaluating the business value of deploying a personalised recommender system for CPG items.

2. BACKGROUND

LeShop (www.LeShop.ch) have been the No.1 e-grocer in Switzerland since 1998 and, a couple of years ago, became the first online grocery pure player in Europe to achieve operational break-even. We have been working in collaboration with LeShop to gather crucial evidence of the real added value to the Business, of deploying personalised recommender systems on their e-commerce site, in terms of the increase in sales revenue. The aim of our collaboration was to increase sales volume and share-of-wallet, by providing LeShop's customers with relevant and personalised recommendations in order to introduce them to items which they had not purchased before. The tools we utilized were in-store and checkout recommender systems, which were integrated into the typical shopping session as illustrated by the flow-chart given in Figure 1.

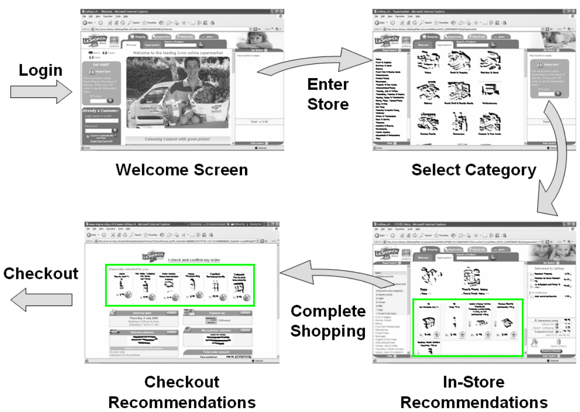

Figure 1: A flow-chart diagram of a typical shopping session on the LeShop.ch e-commerce site. The two recommender interfaces marked by the green boxes are shown in more detail in Figures 2 & 3.

2.1 Checkout Recommender

We started with the deployment of a checkout recommender system, which provided six recommendations, presented to the shoppers at checkout, as illustrated in Figure 2, where the business sensitive information (descriptions, images and prices of the products) has been blurred out.

Figure 2: The user-interface of the checkout recommender system, where the three personalised recommendations are highlighted by the red box.

The first three recommendations, generated by the LeShop system, were "forgotten items" (i.e. items the shopper usually buys, but are not currently in the basket). The three recommendations on the right were the three items (not currently in the basket) with the highest probability of purchase, given the contents of the basket. The probability of item I_k, given the basket B containing items $I_1, \cdots, I_s$ is:

$$P(I_k|B) \propto P(I_k) \cdot P(B|I_k) = P(I_k) \cdot \prod_{i=1}^{s} P(I_i|I_k)\ , \qquad (1)$$

where $P(I_k)$ (the prior probability of purchasing item I_k) and $P(I_i|I_k)$ (the probability of item I_i being bought given that item I_k was bought) are both computed using the frequency of past purchases in the training data. (For more details on the models, please see [8, 5]).

While the "forgotten items" model was regularly updated, we were unable to keep the personalised model updated regularly due to various operational issues. As shown later in Section 3, this inadvertently provided us with evidence that the recency of the training data used to build the model files has an impact on model performance.

2.2 In-Store Recommender

At the same time as the launch of the checkout recommender, LeShop also deployed an in-store recommender system, which initially provided up to eight within-category, non-personalised recommendations, selected by category managers. In our most recent deployment, in October 2007, we personalised two of these in-store recommendations, via our model-based system [8, 5]. The in-store recommendations are presented to the shoppers at the bottom of each category home page, as illustrated in Figure 3, where once again the business sensitive information has been blurred out.

Figure 3: The user-interface of the in-store recommender system, where the two personalised recommendations are highlighted by the red box.

Here, the model-based personalised recommendations are the first two recommendations on the top-left. While the personalised recommendations could be any item available in the store, the programmed recommendations were restricted to within-category items. For example, on the "Dairy" category home page, the personalised recommendations could be a fruit and a light bulb, whereas the programmed recommendations would all be dairy items such as milk, butter and cheese. Furthermore, the programmed in-store recommendations were regularly updated by the category managers, while the model files that generated the personalised recommendations were updated only once, since its launch. The results of the model update can be seen in Section 3.

3. RESULTS ANALYSIS

In this section we present the results of an initial performance evaluation of our live model-based personalised recommender systems between May 2006 and January 2008, with particular focus on their added-value to the business.

3.1 Model Updates

Following the launch of the checkout recommender on the 5^{th} of May 2006, the model files were updated using the latest data on the 20^{th} of October 2006, 7^{th} of March 2007, 22^{nd}

of May 2007, 2^{nd} of October 2007 and the 20^{th} of November 2007. The In-Store recommender model files were updated only once, on the 28^{th} of November 2007, since its launch on the 2^{nd} of October 2007. We use red dotted vertical lines in Figures 4, 5 & 6 to indicate the time points at which the model files were updated.

3.2 Shopper Penetration

We begin with an analysis of how the use of each Recommender System spread over time. This is the most fundamental measure of the value of a recommender system, because no matter how good the recommendations are, if the users (in the case the shoppers) do not use the system it will not be able to generate any value for the Business. Figure 4 shows the penetration of each of our recommender systems over time. The measure of penetration (i.e. the units on the y-axis) used, is the proportion of shoppers who accepted at least one of our recommendations, represented as a percentage of all shoppers who bought at least one item from LeShop, both of which were computed up to and including the month in consideration.

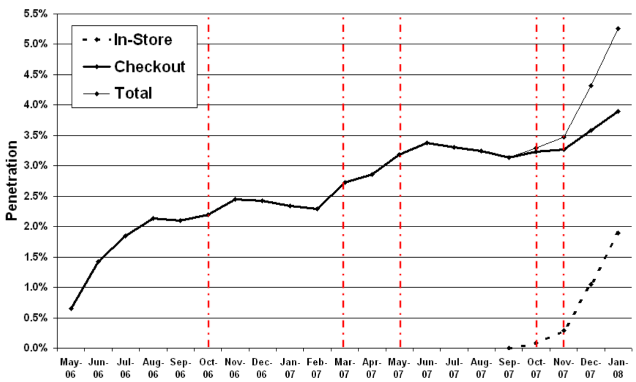

Figure 4: Penetration of our recommender systems over time. The red dotted vertical lines indicate the time points at which the model files were updated.

Here we see that updating the checkout recommender model files consistently resulted in an increase in the number of new shoppers using the recommender system. The increase in penetration resulting from a model update was on average 0.26%. Furthermore, the usage of the recommender systems continues to grow. This could be becuase the shoppers are increasingly becoming more comfortable with the new technology. Unlike that of the checkout recommender, the penetration of the in-store recommender is increasing quite rapidly since its launch. We believe this is because the exposure of the in-store recommender is, on average, nine times that of the checkout recommender. It is also interesting to see that there are a number of users who use only one of the recommenders, since the total penetration is greater than the penetration of either of the recommenders. This may be attributed to the fact that different shoppers have different needs fulfilled by different recommenders.

3.3 Direct Extra Revenue

Next we analyzed the direct extra revenue generated by the recommender systems. The direct extra revenue is simply the total amount of money shoppers spent on purchasing items recommended by our systems. A plot of the direct extra revenue computed on a monthly basis and presented as a percentage of LeShop's total monthly turnover is given in Figure 5.

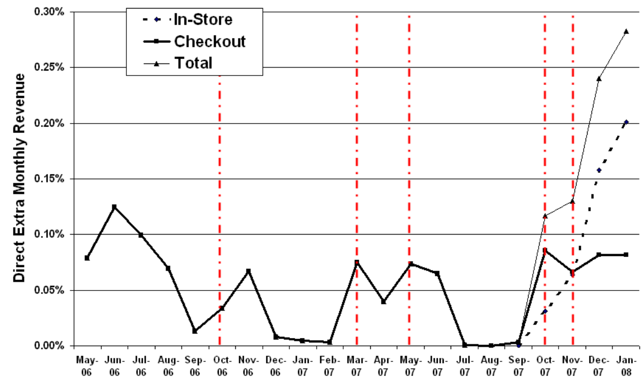

Figure 5: The monthly direct extra revenue generated by our recommender systems, presented as a percentage of the total monthly turnover. The red dotted vertical lines indicate the time points at which the model files were updated.

An important lesson to take away from this analysis is that it is imperative to keep the model files updated, in order to maintain a steady flow of direct extra revenue. We see the performance of the model-based systems fall-off very rapidly until they are updated again using the latest data. We also see that, given its higher exposure (on average ~ 9 times higher), the in-store recommender is able to generate significantly more direct extra revenue than the checkout recommender. It is also important to note that although our highest achieved percentage value of 0.30% seems like a small amount, when you consider the magnitude of the monthly turnover of a successful e-commerce business like LeShop, the actual value generated is substantial.

3.4 Indirect Extra Revenue

The indirect extra revenue generated by our recommender systems comprises two components. The first component is the total amount of money the shoppers spend on repeat purchases of items first introduced to them by our recommender systems. The second component is the total amount of money the shoppers spend on purchases of non-recommended items from categories first introduced to them (via other recommended items from the same category) by our recommender systems. In both cases a particular purchase contributes to the indirect extra revenue only if the purchased item was ***not recommended to the shopper during the current shopping session***. We refer to this as the "indirect" extra revenue, as this extra revenue was not generated directly via the purchase of recommended items, but instead because of the influence of the recommenders. As shown in Figure 6, the indirect extra revenue generated by our recommenders is substantial.

Here we note that within a few months, the indirect extra revenue overtakes the direct extra revenue and then remains consistently above it. Furthermore, even when the model files are out-of-date and, as a result the direct extra revenue falls-off rapidly, the indirect revenue remains relatively stable. We also see that the indirect extra revenue substantially

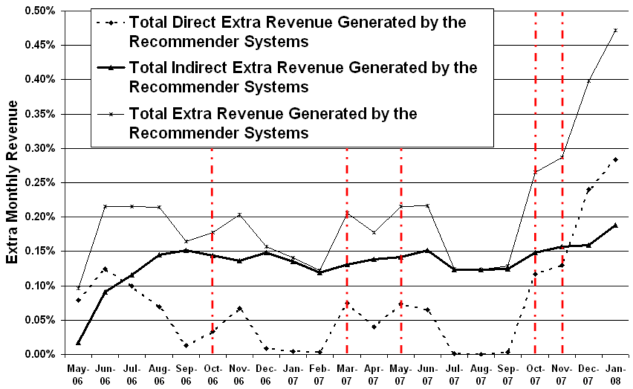

Figure 6: The direct, indirect and total extra revenue generated by our recommender systems. The red dotted lines indicate the time points at which the model files were updated.

(at least by 66% and on average by 336%[1]) increased the total extra value generated by our recommender systems.

If we take a closer look at the categories new to LeShop's shoppers, which were introduced to them by our recommender systems, the top ten categories in terms of extra revenue generation turned out to be Delicatessen (26.02 %), Dairy (19.67 %), Fruits & Vegetables (17.12 %), Butcher (9.04 %), Snacks (8.17 %), Hygiene (4.66 %), Household (4.01 %), Wine (2.26 %), Condiments (2.10 %) and Canned Food (2.04 %) [2]. As we can see the top four categories on this list contain fresh produce items, indicating that recommender systems may be a useful means of overcoming shopper resistance to buying fresh produce on-line.

4. CONCLUSIONS

Recommender systems have recently grown in popularity both in e-commerce and as a research theme. However, most of the literature in this area has been focused on algorithm development, and there is little, if any published direct evidence of the value of recommender systems to e-Businesses, especially relating to consumer packaged goods (CPG) sold in a supermarket setting. Over the past few years, we have been working in collaboration with LeShop to gather such evidence. We deployed in-store and checkout recommender systems on LeShop's e-commerce site, to introduce their customers to items, which they had not purchased before. In this paper, we presented our initial evaluation of the live performance of these recommenders between May 2006 and January 2008, with particular focus on their added-value to the business. Our analysis looked at penetration (i.e. how the use of each recommender system spread over time), and the revenue generated directly from purchases of the recommended items. An important lesson learnt from this analysis is the importance of keeping the model files updated, in order to maintain a steady flow of direct extra revenue, with the performance falling-off very rapidly between model updates.

Another key lessons we have learnt during this case study is that the effect of a recommender system extends far beyond the direct extra revenue generated from the purchase of recommended items. Our analysis shows how our recommenders generated a substantial amount of additional revenue to the business by introducing shoppers to new categories from which they then continued to purchase. A closer analysis indicates that recommender systems may be a useful means of overcoming shopper resistance to buying fresh produce on-line. We are carrying out further analysis on the data we have gathered from this case study. However, even the limited results of our initial evaluation presented here, clearly shows the added-value of personalised recommender systems to e-Businesses.

5. ACKNOWLEDGMENTS

We wish to thank LeShop (www.LeShop.ch - The No.1 e-grocer of Switzerland since 1998) for providing us with the data and the opportunity for this work. We especially thank Valery Blanc and Julien de Perrot, at LeShop, for their invaluable support and collaboration. We are also very grateful to Prof. John Riedl, at the University of Minnesota, for his insightful guidance and feedback. The authors also wish to thank the anonymous reviewers for their helpful and constructive comments.

6. REFERENCES

[1] J. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, 1998.

[2] T. Brijs, G. Swinnen, K. Vanhoof, and G. Wets. Using association rules for product assortment decisions: A case study. In *Knowledge Discovery and Data Mining*, pages 254–260, 1999.

[3] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, 22(1):5–53, 2004.

[4] G. Lekakos and G. M. Giaglis. Improving the prediction accuracy of recommendation algorthms: Approaches anchored on human factors. *Journal of Interacting with Computers*, 18:410–431, 2006.

[5] M. Li, M. B. Dias, W. El-Deredy, and P. J. G. Lisboa. A probabilistic model for item-based recommender systems. In *Proceedings of the ACM Conference on Recommender Systems*, pages 129–132, 2007.

[6] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In *10th International World Wide Web Conference*, pages 285 – 295, May 2001.

[7] G. Shani, D. Heckerman, and R. I. Brafman. An MDP-based recommender system. *Journal of Machine Learning Research*, 6(2):1265–1296, 2006.

[8] C. M. Sordo-Garcia, M. B. Dias, M. Li, W. El-Deredy, and P. J. G. Lisboa. Evaluating retail recommender systems via retrospective data: Lessons learnt from a live-intervention study. In *Proceedings of the International Conference on Data Mining*, pages 197–203, 2007.

[1]Considering only the months during which the direct extra revenue was above 0.01%

[2]Here the percentage values are the contribution of each category to the extra revenue.