An Adaptive Electronic Menu System for Restaurants

Paulo Henrique Azevedo Filho
Aberklar GbR
Helene-Mayer-Ring 7
80809, Munich, Germany
paulo.azevedo@aberklar.com

Wolfgang Wörndl
Technische Universität München
Boltzmannstraße 3
85748, Garching bei München, Germany
woerndl@informatik.tu-muenchen.de

ABSTRACT

This work shows the early stages of the development of a collaborative-filtering-inspired adaptive system to streamline the ordering process at restaurants that use electronic menu systems.

Among other results, the proposed system achieved a reduction of the session duration, while increasing feedback given by restaurant guests.

Categories and Subject Descriptors

H.4.m [**Information Systems Applications**]: Miscellaneous

General Terms

Human Factors, Economics

Keywords

Recommender Systems, Adaptive, Collaborative-Filtering, Electronic Menu, Target Variable

1. INTRODUCTION

There is an increasing trend in the number of tablet-based systems for ordering food, in order to streamline the service, increasing efficiency and profit. This is a great opportunity to make a change: instead of mere replacements for the paper menu and replacement for some of the roles of the waiters, such systems can do much more, such as becoming feedback-gathering devices, or helping guests choose their dishes better and increasing their satisfaction by offering meaningful suggestions.

Those meaningful suggestions may stem from collaborative filtering. Well-established algorithms for collaborative filtering exist, as well as several hybridization techniques blending this with other approaches. However, since users of electronic menus are mostly anonymous (no log-in necessary), and the offer of dishes and drinks varies greatly from restaurant to restaurant, the choice for a recommendation technique has to be further tweaked to address these issues.

The present work is an attempt to embrace this opportunity, building an adaptive recommender system for electronic restaurant menus, to aid guests in their ordering process, based on MenuMate.

MenuMate, developed by the German startup Aberklar[1], is an electronic menu that offers users a picture-centric approach for use at restaurants. Through MenuMate users can place orders, request the bill and provide feedback about their experience. This is the system that was extended with the adaptive system described in this paper. Figure 1 gives an idea of the general structure of the dish overview screen, and the arrows suggest the repositioning of dishes of the proposed method, that will be described soon.

The dish overview screen shows a title bar for each category, followed by thumbnails of each dish in that category. From that screen, the user can go to the dish details screen, with a full screen picture and detailed description, from where the dish can be ordered. After orders have been placed and the tab asked for, the user is invited to provide feedback about a few different variables, such as food, drinks, service and electronic menu system, using a star-based rating.

The rest of this document is organized as follows: Section 2 will put this work in perspective in terms of the electronic menu used for its implementation, as well as of the related work. Section 3 will describe in general terms the proposed algorithm and how it fits in. Section 4 describes some of the performed experiments and displays their results. Finally, section 5 offers a brief summary as well as points some possibilities of further work to be explored.

2. BACKGROUND AND RELATED WORK

To achieve the aforementioned goal, once the current state of MenuMate is known, it is necessary to know what the current state of the art for this specific niche is.

Wasinger et al. have proposed an electronic menu with an embedded recommender system, called Menu Mentor [9]. The authors come from a perspective of highly personalized explicit recommendations, processed on the user's phone, and that must be scrutable, that is, allow the users to know why a given recommendation (be it positive or negative) was

[1]www.aberklar.com/en

Figure 1: Screen of MenuMate with the overview of the dishes, as well as a suggestion of the dynamic reordering of the menu items.

given, and give them the chance to override it. It does, however, assume explicit user profiles, acquired through usage of the system on the user's smartphone. In order to minimize user setup, restaurants should be able to offer the system and the hardware.

There are psychological studies about how to organize a restaurant menu, with techniques such as placing items with high prices first to "smoothen" the effect of the lower-priced items following, even if they are not actually cheap [6]. There are also studies evaluating guests' gaze and how it correlates to which dishes get ordered, placing dishes restaurateurs want to have ordered in those positions, such as shown at [1]. There are as well other possibilities that go in a similar direction, such as [8] and [10]. Going for this sort of psychological study, however, would require extensive trials after every change to the menu, as well as the usage of the same menu in a different place where cultural background changes. This type of approach is also sensitive to the direction of reading of the mother tongue of the restaurant guest, such as right to left in case of Arabic speakers.

The goal of this work was to develop an electronic menu system that adapts itself according to a certain target variable, such as session duration, but while avoiding the need both for the setup of a user profile and extensive trials after each change. This is what will be presented in the next section.

3. ALGORITHM AND ARCHITECTURE

The proposed idea is rather simple. It consists in reordering the menu items to optimize a quantifiable target variable, that can be either maximized or minimized. For this work, four target variables were studied: tab value, session duration, revenue rate (cents per minute) and feedback rating. Once this definition is set, as different menu sequences are used the value of each target variable is recorded for each session, as well as the position of each dish in the menu. The optimal menu sequence is computed by calculating the correlation coefficient between each dish's position and the performance, and sorting ascending or descending, depending on whether the action is to maximize or minimize, respectively.

Four correlation coefficients were tested: Spearman's [4], Pearson's [5], Goodman and Kruskal's [2] and Kendall's [3]. In fact, the experiments tried all of them in different sessions and made a comparison between them. More coefficients could be used, the only requirement is that they must yield values between -1 and 1, which could imply a normalization step for coefficients that do not yield results in this range.

In order to vary the position of dishes, there is what we called pre-optimization randomization, which will shuffle the dishes before sorting them, giving the chance to dishes that have the same coefficient (within a small delta) to change places, changing not only absolute but also relative order. A bias could, otherwise, arise from the fact that stable sorting algorithms were used, thus keeping items with a similar correlation coefficient always in the same relative order, preventing them from switching places and moving far away from each other. Also, dishes for which there is no previous information always start with coefficient 0.

The basic principle is as follows: after each session, the tablets send the raw data gathered to the system that generates the sequence (there is one such system per restaurant). This system, based on the chosen coefficient and target variable, calculates the value for that variable and the coefficients, generates the menu sequence, and at the beginning of each session the tablets poll this system for the latest sequence to be employed. This system, called Menu Optimizer, is configurable in respect to the target variable, action (maximize or minimize) and correlation coefficient, together with other parameters relevant for the A/B testing performed, that will be explained in the next section.

4. PRELIMINARY USER STUDIES

Since the method proposed does not explicitly recommend a single item or try to predict ratings, some methods usually employed to evaluate recommender systems cannot be employed, such as cross-validation, recall and precision and accuracy.

There are, however, other methods that can be employed directly: the test was a double-blind A/B test, in which neither the user nor the waiter knew which the target variable at the time was, nor the correlation coefficient used. It was also online (in the sense that real users were using the system, generally spending real money through it [7], as was the case in our tests), and measured a few variables.

Since there is no way to directly measure accuracy for this system, there is employment of efficiency. In traditional recommender systems, the goal is to try to predict what the user would like to find and show it to them, shortening the search. Assuming that users find more easily the information they are looking for, they will more promptly take decisions based on that information, namely order the dish they intend to. Assuming this to be true, it would derive that a reduction of session duration would imply increased accuracy in a way. To assess whether this is true, it should be coupled with an increase of the feedback ratings, which would show that the items found satisfied the user. This metric, that as mentioned before, is called efficiency, and was used to assess the system.

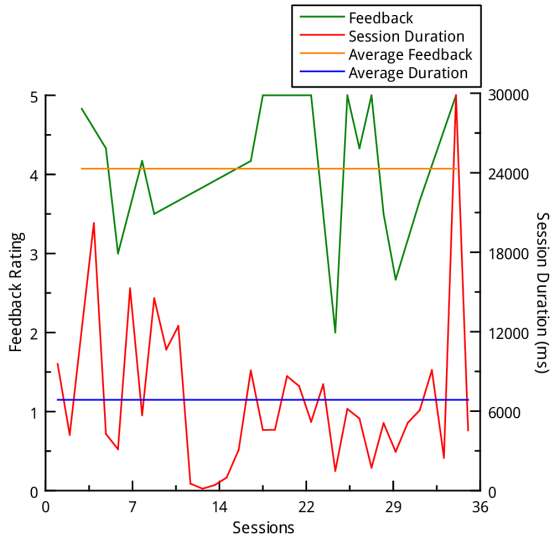

Figure 2: The evolution of feedback ratings and session duration at one of the restaurants, that suggests good efficiency.

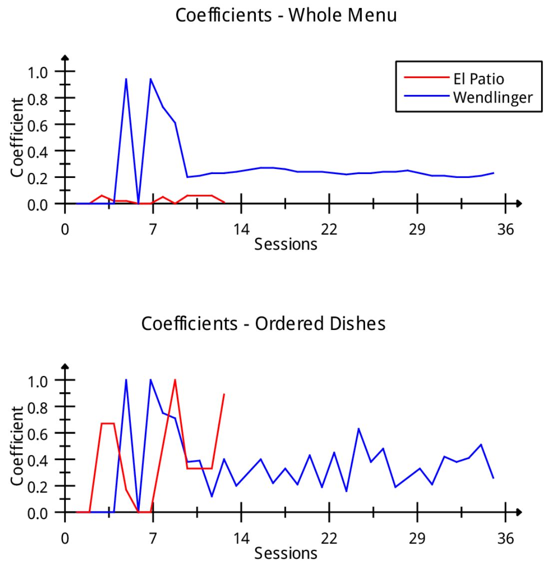

Figure 3: How the values of the correlation coefficients evolved over time at the two restaurants.

Another adaptation made was with regards to serendipity. Instead of measuring it, there was the measurement of catalog coverage: the measurement of how much users tend to order varied dishes, compared to users of plain paper menu. For that, observations were made of tables whose guests did not use MenuMate, and which dishes they ordered, and those were compared to the orders of MenuMate users.

There were two classes of tests performed, the stress tests and the user tests. The stress tests, which will not be displayed in detail here, were used to assess the scalability of the system, which was developed to be run on a Raspberry Pi computer, with an embedded 900MHz ARM processor. It suffices to say that the system could serve between 100 and 160 tablets with unnoticeable performance losses, and that the limits reached were due to the test rig employed, rather than the system itself. Another interesting result on this front is that, on the employed hardware, menu optimization time is increased on average by 3ms for each session stored, which allows prediction of the optimization time based on the size of the history of observations.

The preliminary user studies were performed at two different restaurants in Munich, for about a week in each, time during which the system would gather session data and automatically adapt itself, switching the target variable at regular intervals. The number of observations was rather small due to the reduced number of days allowed for observations. There were 13 sessions observed in one restaurant, called El Patio, and 35 in the other, called Wendlinger. The restaurants had, respectively, 201 and 290 menu items, split in 24 and 34 categories. The original menu sequence was devised by their respective owners.

Figure 2 shows the evolution over time of the efficiency at Wendlinger (the number of sessions to which users gave feedback at El Patio was too low to assume it was meaningful). It suggests that indeed over time, independent of those being the target variables, feedback tends to rise and session duration to reduce, indicating good efficiency.

Table 1 shows how the choice of a target variable influences the performance of that variable, as well as of other variables. In this table only sessions optimized with Spearman's correlation coefficient are used, because it was the only coefficient with a high enough number of observations from which to derive conclusions (21 sessions in total). The "none" line, represents results for a non-optimized menu. The system did well in optimizing for the target variable, with the best results for both feedback and session duration, a close second for tab value, while not delivering good results for revenue rate, because the higher tab value was not enough to compensate for the roughly halved session duration time achieved when the target variable was the duration. It is also worth noticing that indeed, revenue rate is a direct consequence of both tab value and session duration, but the average revenue rate is not necessarily the same as the revenue rate of the averaged tab values and durations. For clarity, the best results for each variable are displayed in boldface.

Figure 3 shows another facet of the inner workings of the system: although the number of observed sessions was low, it is very fast to converge the internal coefficients, suggesting some stability after approximately 10 sessions. Another interesting effect is that the absolute values of the coefficients tend to be higher for dishes that ended up being ordered, which means the system can in fact predict which dishes will be ordered, by checking the dishes with highest absolute value of the coefficients.

Due to operational constraints, catalog coverage was only measured at El Patio. There, 12 of the 13 sessions were with menu optimization enabled, that will be considered for this measurement. 50 sessions with paper menus were also

Target Variable	Feedback	Value (€)	Duration	Rev. Rate
None	2,0	25,17	12.865,6	20,9
Feedback	**5,0**	13,25	15.779,9	33,8
Tab Value	4,4	38,00	7.499,5	59,1
Session Duration	4,3	36,85	**2.406,9**	**154,9**
Revenue Rate	**5,0**	**40,10**	4.400,6	56,6

Table 1: Cross-references between target variables and the results for all interest variables.

observed. In total, 97 different dishes and drinks were ordered, in different quantities. From those 97, 19 were ordered both with and without MenuMate, 63 only by users of the paper menu and 15 only by users of MenuMate with menu optimization.

It is not easy to extrapolate how those proportions would be in case an equal number of observations was available, and discarding paper menu-based sessions could arbitrarily lead to any results. Assuming, however, that for both systems at each session there is an equal probability of adding not previously ordered items, and that this probability is inherent to either MenuMate or the paper menu, a proportion rule may be followed.

In the case of the paper menu, $19 + 63 = 82$ different items were ordered in 50 sessions, which yields an average of 1,64 new items per session. With MenuMate in use, there were $19 + 15 = 34$ different items ordered in 12 sessions, resulting in 2,83 new items per session. That may indicate that thus, the menu optimizations led to guests ordering a bigger variety of items. Alternatively, it could be that as more sessions would be measured, these sessions would progressively get less ”innovative”, in terms of the ordered dishes, which could revert this balance. An extended evaluation, with a similar number of sessions in both conditions would help settle this matter.

5. CONCLUSIONS AND FUTURE WORK

The proposed menu optimizer harnesses principles of recommender systems to improve restaurant menus according to an arbitrary interest variable. In fact, it can be used to facilitate user interaction for any system to which access is anonymous and the number of items not overwhelming. Another use that comes to mind is the choice of which presentations to attend at a conference, or main sights to visit in a city.

One of the main contributions is the strong suggestion that this purely statistical method may improve the menu even if the underlying mechanism that drives the change is not understood by the system (i.e. the psychological implications).

The developed system comprises the algorithm for menu optimization, coupled with a robust and scalable implementation of it. It was followed by qualitative and quantitative tests, with real paying users, of which only very few results were presented this time due to page number constraints, but that nevertheless suggest efficacy of the proposed method.

There are, however, some promising improvements to the method. Among which, the highlights are:

- Feature extraction, to allow dishes of different restaurants to be matched and correlation information exchange, possibly improving results;
- Extra variable isolation, that would allow control over influential external variables such as weather, time of the day or season;
- If individual user profiling is done, further personalization, such as filtering out dishes based on allergies or taste preferences, could be done;
- Stochastic exploration of dishes, which could allow for recommendation of the next course based on previously ordered dishes in a session.

6. REFERENCES

[1] J.-G. Choi, B.-W. Lee, and J.-w. Mok. An experiment on psychological gaze motion: a re-examination of item selection behavior of restaurant customers. *Journal of Global Business and Technology*, 6(1):68, 2010.

[2] L. A. Goodman and W. H. Kruskal. Measures of association for cross classifications iii: Approximate sampling theory. *Journal of the American Statistical Association*, 58(302):310–364, 1963.

[3] M. G. Kendall. A new measure of rank correlation. *Biometrika*, pages 81–93, 1938.

[4] J. L. Myers, A. Well, and R. F. Lorch. *Research design and statistical analysis*. Routledge, 2010.

[5] K. Pearson. Notes on regression and inheritance in the case of two parents. In *Proceedings of the Royal Society of London*, volume 58, pages 240–242, 1895.

[6] W. Poundstone. *Priceless: The myth of fair value (and how to take advantage of it)*. Macmillan, 2010.

[7] G. Shani and A. Gunawardana. Evaluating recommendation systems. In *Recommender systems handbook*, pages 257–297. Springer, 2011.

[8] B. Wansink, J. Painter, and K. Van Ittersum. Descriptive menu labels' effect on sales. *The Cornell Hotel and Restaurant Administration Quarterly*, 42(6):68–72, 2001.

[9] R. Wasinger, J. Wallbank, L. Pizzato, J. Kay, B. Kummerfeld, M. Böhmer, and A. Krüger. Scrutable user models and personalised item recommendation in mobile lifestyle applications. In *User Modeling, Adaptation, and Personalization*, pages 77–88. Springer, 2013.

[10] S. S. Yang, M. M. Sessarego, et al. $ or dollars: Effects of menu-price formats on restaurant checks. *Cornell Hospitality Reports*, pages 6–11, 2009.