

A System for Recommending Items Based on Viewing-Time-Weighted Preferences for Attributes

Jeffrey Parsons¹, Paul Ralph², Katherine Gallagher¹

¹Faculty of Business Administration
Memorial University of Newfoundland
St. John's, NL, Canada
jeffreyp@mun.ca, kgallagh@mun.ca

²Sauder School of Business
University of British Columbia
Vancouver, BC, Canada
paulralph@gmail.com

Research Question

This paper outlines the motivation, design, and preliminary evaluation of a recommender system that infers user preferences from product viewing times. Unlike existing systems, we use viewing time as an indicator of preference for items, based on findings that people look at objects they like, or find interesting, for a longer time than objects they do not like, or do not find interesting (Berlyne and Lawrence, 1964; Faw & Nunnally, 1967; Oostendorp and Berlyne, 1978; Konstan et al. (1997). In an earlier study, we found a positive relationship between time spent viewing an item in an online catalog and revealed preference for that item as indicated by subsequent selection of an item for purchase (Parsons et al., 2002).

Research Approach

DESIRE Recommender System

DESIRE is an item-to-user recommender system that combines a viewing time- and attribute-based preference inference algorithm with an attribute-based recommendation engine. Consider a user interacting with an online catalog. Each time the user views an item page (any page containing an item description), an implicit rating is calculated (process 1 in Figure 1).

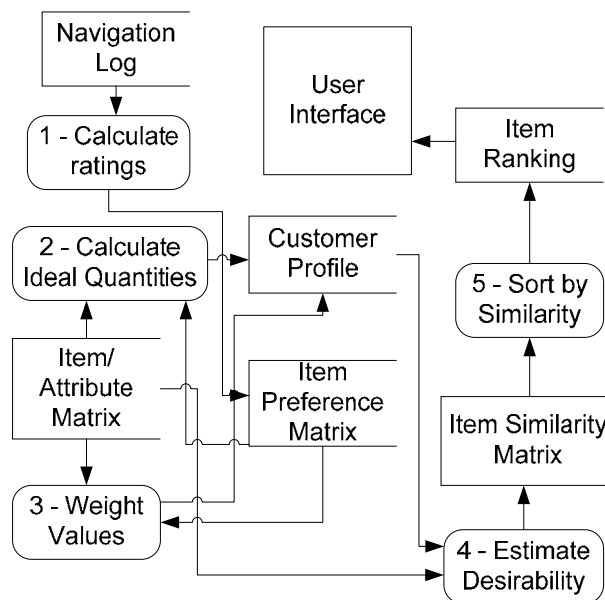


Figure 1: Implementation of DESIRE

The recommendation engine component of DESIRE decomposes each item into a collection of properties. For each numeric property, this user's projected ideal quantity is calculated (process 2 – details omitted). For each value of each textual property, a preference weight is calculated (process 3 – details omitted). Then, for each item in the item set, the desirability is calculated by comparing the item's property values to these ideal quantities and value weights (process 4). Finally, the item set is sorted by desirability (process 5) and a recommendation set is returned.

DESIRE does not treat all properties equally. The desirability of an item, *i*, is calculated as a weighted average of *i*'s property desirabilities, where the weights correspond to the relative importance attributed to each property. The relative importance of attributes could be determined by asking users to rank or rate the importance of each attribute for their decision-making, or calculated from industry-based surveys of the importance of various attributes to consumers of specific products. Weights may vary based on the type of users (i.e., corporate vs. private), type of item (i.e., clothing vs. electronics) the domain (i.e., information search vs. e-commerce), or other factors.

Empirical Test of DESIRE

We conducted a study to assess the quality of DESIRE's recommendations. We compared DESIRE's ratings to explicit user ratings on a set of items, employing Mean Absolute Error (MAE) (Herlocker et al., 2004). We expected DESIRE to produce better recommendations than a system that makes naïve or random recommendations. We test DESIRE against two such recommendation strategies: (1) always assign the mean rating to items; (2) assign ratings to items using a uniform distribution.

Procedure

We conducted a laboratory study, consisting of a simulated shopping exercise in which participants viewed a series of items from several product categories. Participants were then asked to rate a series of items they had not previously seen. These ratings form the basis for our analysis.

Data Collection

Each item in the catalog was described in terms of values for a set of quantitative and categorical attributes that were used as input to the DESIRE algorithm. In addition, the system recorded the time spent viewing each catalog item. These viewing times were used by DESIRE to determine the preferences for attribute values. After users browsed the catalog, they were asked to rate a series of items not contained in the catalog. DESIRE calculated the expected ratings of these items. These rating pairs formed the basis for the analysis of the quality of DESIRE's ratings.

Main Findings & Ongoing Work (Synopsis)

To test Hypothesis 1, we examined the quality of DESIRE's ratings by looking at the size of errors using MAE (Herlocker et al., 2004). Table 1 reports the MAE for DESIRE recommendations.

To put these numbers in context, we examine how well DESIRE's performance compares to a system that recommends items based on randomly assigned ratings. Such a system would assign ratings to items randomly, so that each score on a 9-point scale would be expected to be assigned to one-ninth of the items. The performance of a random recommender will depend on the distribution of user ratings of items in a catalog. For our study, if the distribution of user ratings is uniform, the MAE would be 2.84 on a 9-point scale. In comparison, if there is no distribution (all ratings at one point), the MAE would be 2.22 if that rating was 5, and higher otherwise. This latter figure is also the MAE if the user ratings are uniformly distributed, and the recommender always guesses '5.' In this context, DESIRE ratings were substantially better than random for bicycles, cameras, and notebooks, lending support to Hypothesis 1. In contrast, ratings were no better than random for boots (as expected), and portable music players. This at least raises the possibility that preferences for boots and portable music players might be based on more than their measurable properties.

Table 1: MAE of DESIRE Ratings by Product Category

| CATEGORY | MAE | MAE as % |
|------------------------|------|----------|
| bicycles | 1.14 | 12.7 |
| Boots | 2.35 | 26.1 |
| digital cameras | 0.83 | 9.2 |
| notebook computers | 1.57 | 17.4 |
| portable music players | 2.60 | 28.9 |

We have conducted additional tests of the quality of DESIRE's recommendations, including a general comparison to published systems on MAE performance. These are omitted here due to space limitations, but will be presented at the conference. Also, we are using the Netflix database (<http://www.netflixprize.com>) of movie ratings to test the effectiveness of the attribute-based preference inference algorithm.

The ideas embodied in DESIRE are intended to supplement, rather than replace, collaborative and content-based recommenders. Since the results of any two recommenders can be combined in such a way that the combination can do no worse than the better of the two recommenders, DESIRE can be combined with existing recommenders to improve accuracy. Evaluating the resulting combinations is an important area for future research.

Also, further research is needed to examine the impact of other factors that influence viewing time (Figure 2) and control for these in DESIRE's algorithm. By isolating the impact of preference from other factors that influence viewing time, we to produce further improvements in DESIRE's recommendation quality.

References

- Berlyne, D. & Lawrence, G. 1964. Effects of complexity and incongruity variables on GSR, investigatory behavior and verbally expressed preference. *The Journal of General Psychology*, 71: 21-45
- Faw, T. & Nunnally, J. 1967. The Effects on Eye Movements of Complexity, Novelty, and Affective Tone. *Perception & Psychophysics*. 2 (7): 263-267.
- Herlocker, J., Konstan, J., Terveen, L & Riedl, J. (2004) Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, 22 (1): 5-53.
- Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L., Riedl, J. 1997. GroupLens: Applying collaborative filtering to Usenet news. *Communications of the ACM*, 40 3: 77-87.
- Oostendorp, A., & Berlyne, D. E. 1978. Dimensions in the perception of architecture II: measures of exploratory behavior. *Scandinavian Journal of Psychology*, 19 1: 83 – 89.
- Parsons, J., Gallagher, K., and Ralph, P., Inferring Preferences from Viewing Time: Implications for the Design of Electronic Catalogs. *INFORMS Annual Meeting*, San Jose, CA, November, 2002