Dynamic Facet Ordering for Faceted Product Search Engines (Extended Abstract)

Damir Vandic#1, Steven S. Aanen*2, Flavius Frasincar#3, and Uzay Kaymak†4

Econometric Institute, Erasmus University Rotterdam
PO Box 1738, NL-3000 DR
Rotterdam, the Netherlands
* *Grible.co*
Coolsingel 104, NL-3011 AG
Rotterdam, the Netherlands
† *Information Systems IE&IS, Eindhoven University of Technology*
P.O. Box 513, NL-5600 MB
Eindhoven, the Netherlands
[1] vandic@ese.eur.nl [2] aanen@grible.co [3] frasincar@ese.eur.nl [4] u.kaymak@tue.nl

***Abstract*—Currently many webshops rely on a fixed list of product facets to help users find the products of interest. Such a solution suffers from two problems: (1) it is difficult to devise a fixed list of facets that would satisfy all user interests, and (2) the top facets could become obsolete when all product results have these facets. To address these two problems we propose a novel algorithm for dynamic ordering of the product facets based on the query results. This algorithm relies on measures such as specificity and dispersion for qualitative and quantitative facets, respectively, to rank the properties associated with these facets so that users are able to find the products of interest with a minimum number of drill-down steps. Using a large-scale simulation study and a user-based evaluation, we show that our algorithm outperforms the expert-based fixed facets approach, a greedy baseline, and a state-of-the-art entropy-based solution. This paper is an extended abstract of our previous work [1].**

***Index Terms*—Facet ordering, product search, user interfaces**

I. Introduction

Many of the current Web shops that use faceted search have a manual, 'expert-based' selection procedure for displaying product facets. However, manually ordering these facets requires a significant amount of work. In addition, such approaches fail to consider that after each query the resulting set of products changes, and thus the previously fixed ordering of facets might not be that relevant anymore. Therefore, it is highly unlikely that the fixed set of facets will be optimal for the duration of the user search session.

In order to address the previously identified issues we propose an approach for dynamic facet ordering for Web shops. Our solution ranks the properties associated to facets (a property has many underlying facets, e.g., the property 'Featured Brands' has facets 'Samsung', 'Motorola', 'Nokia', etc.) by means of specificity measures for qualitative properties (e.g., 'Color' property) and dispersion measures for quantitative properties (e.g., 'Price') so that the most discriminative properties are shown on top of the facet list.

As part of our solution, we devise an algorithm that ranks properties by their importance and also sorts the facets within each property. For property ordering, we identify specific properties whose facets match many products (i.e., with a high impurity) or have a high degree of dispersion. The specificity of these aspects is weighted so that facets that match many products should be ranked higher compared to facets that match only a few products. The facets are ordered, per property, descendingly based on the number of products that have these facets.

II. Related Work

Many of the recent studies [2], [3] do not differentiate between qualitative and quantitative properties, thus failing to consider the specificity of ratio scale data. Furthermore, some of the existing solutions are not able to cope with the e-commerce domain, where queries are Boolean and thus there is no ranking of products with respect to a query [2]. Also e-commerce queries have a conjunctive semantics for properties and disjunctive semantics for facets, something that none of the previous works considered.

In the proposed solution we consider Boolean queries that use a conjunction of properties, and disjunctions of facets, as typically encountered in e-commerce systems. We also propose the use of a dispersion measure to properly deal with quantitative properties, along with a property weighting based on the popularity of facets in products, to cater for property popularity and missing values. In addition, differently from the existing solutions, we perform both a large-scale simulation study and a user-based evaluation, demonstrating the superiority of our approach over a state-of-the-art solution [3] for e-commerce and two baselines.

III. Facet Optimization Algorithm

Figure 1 shows the main flow in a user search session in our approach. The user may not know the name of the product it is looking for, but knows the characteristics of the desired product. At the beginning, the user is presented with the complete set of products available in the Web shop. Our solution computes the property and the facet scores in parallel, after which the user is presented an ordered list of properties and, per property, an ordered list of facets. If the result set is too large, the user performs a drill-down (by selecting additional properties and their corresponding facets) and thus the query is updated triggering a new search iteration. If the result set is small enough, the user scans the returned products. If the desired product is not found, the user performs a roll-up (deselecting properties or selecting additional facets) and thus the query is updated again, triggering a new iteration. On the other hand, if the product is found, the search session ends.

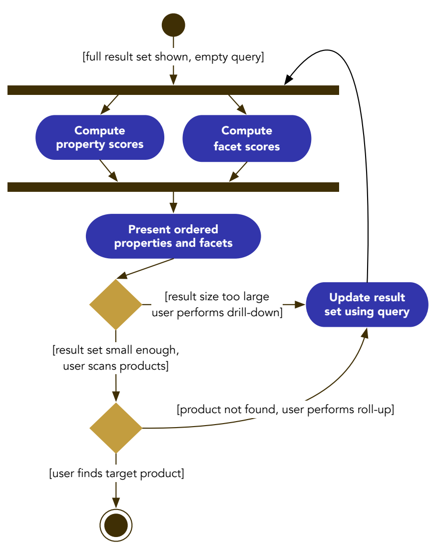

Fig. 1. Activity diagram describing the main flow of a search session.

Figure 2 shows the steps taken to compute the property score (one of the steps in Figure 1). If the property is a qualitative property, our approach computes the disjoint facet count for this property. Using the disjoint facet count a temporary property score (as a measure of specificity) is computed based on the Gini impurity measure. If the property is a quantitative one, in order to exploit the ratio scale of this data, we compute the Gini coefficient, which represents the temporary property score (as a measure of dispersion). The temporary scores are weighted based on the property popularity among products to favor properties that have facets spread among many products and to better cope with missing facets in the dataset. At the end, the properties are sorted in the descending order of their scores. The facets associated to a property are sorted in the descending order of facet scores computed as the number of products associated to a facet (one of the steps in Figure 1).

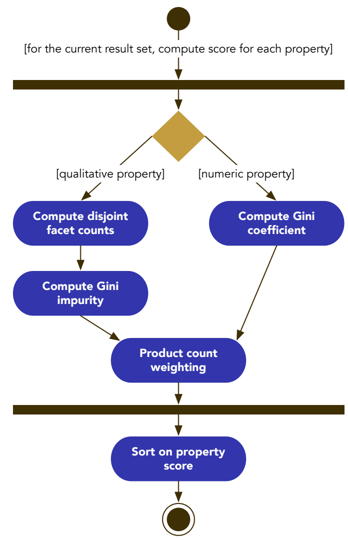

Fig. 2. Activity diagram describing the property score computation.

IV. Evaluation

For our experiments, we have gathered data from Tweakers Pricewatch [4]. The complete catalog contains 794 mobile phones, 53 properties, and 1,816 facets, from which 348 are qualitative and 1,468 are quantitative. We performed a large-scale simulation study based on 3 drill-down models (Least Scanning, Best Facet, and Combined Model), 4 ordering schemes (Expert-based, Greedy Count, Kim et al., and our approach), 794 target products, and 50 repetitions for the Combined Model. Our approach performs better in terms of number of clicks, has the lowest number of roll-ups, and the highest percentage of successful sessions.

In addition to the simulation study, we also performed an experiment with 27 real users that were given 10 tasks to find products that match given product descriptions. Each user performed half of the experiments with our system and half of the experiments with the expert-based fixed facets system. Again the experiments show that the users need less clicks to find the desired product using our system.

V. Conclusion

In this work we have presented an approach to automatically and dynamically order properties and their facets in e-commerce to help users quickly find the desired products. As future work we plan to extend our solution by accounting for the property and facet popularity with respect to users.

References

[1] D. Vandic, S. S. Aanen, F. Frasincar, and U. Kaymak, "Dynamic facet ordering for faceted product search engines," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 5, pp. 1004–1016, 2017.

[2] S. Liberman and R. Lempel, "Approximately optimal facet value selection," *Science of Computer Programming*, vol. 94, pp. 18–31, 2014.

[3] H.-J. Kim, Y. Zhu, W. Kim, and T. Sun, "Dynamic faceted navigation in decision making using semantic web technology," *Decision Support Systems*, vol. 61, pp. 59–68, 2014.

[4] Tweakers.net, "Dutch IT-community with a dedicated price comparison department," http://www.tweakers.net, 2014.