Automatic Ordering of Facets Based on Machine Learning

[1]Diksha M, [2]Raghavendra T.S
[1]MTech Student, [2]Associate Professor
Dept. Of CSE, Cambridge Institute of Technology, Bangalore

***Abstract*-Faceted browsing is widely used in Web shops and product comparison sites. In these cases, a fixed ordered list of facets is often employed. This approach suffers from two main issues. First, one needs to invest a significant amount of time to devise an effective list. Second, with a fixed list of facets it can happen that a facet becomes useless if all products that match the query are associated to that particular facet. In this work, we present a framework for dynamic facet ordering in e-commerce. Based on measures for specificity and dispersion of facet values, the fully automated algorithm ranks those properties and facets on top that lead to a quick drill-down for any possible target product. In contrast to existing solutions, the framework addresses e-commerce specific aspects, such as the possibility of multiple clicks, the grouping of facets by their corresponding properties, and the abundance of numeric facets. In a large-scale simulation and user study, our approach was, in general, favourably compared to a facet list created by domain experts, a greedy approach as baseline, and a state-of-the-art entropy-based solution.**

***Keywords*— facet ordering, product search, user interfaces.**

I. INTRODUCTION

Studies from the past have shown that other factors than the price play a role when a consumer decides to choose where to buy a product online . Therefore, online retailers pay special attention to the usability and efficiency of their Web shop user interfaces. Nowadays, many Web shops make use of the so-called faceted navigation user interface , which is in literature also sometimes referred to as 'faceted search'. Facets are used by some users as a search tool, while others use it as a navigation and/or browsing tool . One of the reasons why faceted search is popular among Web shops is that users find it intuitive. The term 'facet' has a rather ambiguous interpretation, as there are different types of facets. In this work, I refer to facets as the combination of a property and its value.

Currently, most commercial applications that use faceted search have a manual, 'expert-based' selection procedure for facets , or a relatively static facet list . However, selecting and ordering facets manually requires a significant amount of manual effort. Furthermore, faceted search allows for interactive query refinement, in which the importance of specific facets and properties may change during the search session. Therefore, it is likely that a predefined list of facets might not be optimal in terms of the number of clicks needed to find the desired product. In order to deal with this problem, I propose an approach for dynamic facet ordering in the e-commerce domain. The focus of our approach is to handle domains with sufficient amount of complexity in terms of product attributes and values. Consumer electronics (in this work 'mobile phones') is one good example of such a domain. As part of our solution, we devise an algorithm that ranks properties by their importance and also sorts the values within each property. For property ordering, we identify specific properties whose facets match many products (i.e., with a high impurity). The proposed approach is based on a facet impurity measure, regarding qualitative facets in a similar way as classes, and on a measure of dispersion for numeric facets. The property values are ordered descending on the number of corresponding products. Furthermore, a weighting scheme is introduced in order to favour facets that match many products over the ones that match only a few products, taking into account the importance of facets. Similar to existing recommender system approaches , our solution aims to learn the user interests based on the user interaction with the search engine.

From the perspective of user interface design, we distinguish between two main facet types: qualitative facets (e.g., WiFi:true) and numeric facets (e.g., Lowest price (e):64.00). We further distinguish between two types of qualitative facets: nominal facets and Boolean facets. Nominal facets are, for example, those for the property Display Type, and can have any nominal value. Boolean facets are for instance Multitouch, and have only three options from an interface perspective: true, false, or no preference.

II. LITERATURE SURVEY

A. Consumer Selection of E-Commerce Websites in a B2C Environment: A Discrete Decision Choice Model

Despite the suggestions of friction-free information availability, considerable price dispersions for the same product are not uncommon across online retailers in the business-to-consumer (B2C) segment. Online customers do not necessarily always buy from the site with the lowest price, suggesting that other forces are at work. This paper presents and empirically examines a model that proposes that Web site value in terms of (perceived) Web site quality as well as awareness of the site and consumer differences (on price sensitivity) are key variables in explaining online consumer behavior in their choice of Web site despite the existence of price dispersions. Two hundred ninety-three students participated in a series of controlled laboratory experiments making use of two different types of products in terms of complexity and expensiveness (Canon digital camera Powershot S400

and digital versatile disc full-screen edition of Star Wars: Episodes I and II) that required them to interact with three different real-world Web sites offering each of these two branded products and make their decision on which of the three Web sites they will chose to buy the product from. The prices varied across the Web sites, as did the quality of the sites on various dimensions and site awareness of the participants. Conditional logit models of discrete choice for each of the two product types indicate differential influences of Web site quality dimensions and price sensitivity. A number of interesting implications emerge, and pointers to further extensions of the research theme are discussed.

B. Enhancing Collaborative Filtering by User Interest Expansion via Personalized Ranking.
Recommender systems suggest a few items from many possible choices to the users by understanding their past behaviors. In these systems, the user behaviors are influenced by the hidden interests of the users. Learning to leverage the information about user interests is often critical for making better recommendations. However, existing collaborative-filtering-based recommender systems are usually focused on exploiting the information about the user's interaction with the systems; the information about latent user interests is largely underexplored. To that end, inspired by the topic models, in this paper, we propose a novel collaborative-filtering-based recommender system by user interest expansion via personalized ranking, named iExpand. The goal is to build an item-oriented model-based collaborative-filtering framework. The iExpand method introduces a three-layer, user-interests-item, representation scheme, which leads to more accurate ranking recommendation results with less computation cost and helps the understanding of the interactions among users, items, and user interests. Moreover, iExpand strategically deals with many issues that exist in traditional collaborative-filtering approaches, such as the overspecialization problem and the cold-start problem. Finally, we evaluate iExpand on three benchmark data sets, and experimental results show that iExpand can lead to better ranking performance than state-of-the-art methods with a significant margin.

C. Approximately Optimal Facet Value Selection.
Multifaceted search is a popular interaction paradigm for discovery and mining applications that allows users to analyze and navigate through multidimensional data. A crucial aspect of faceted search applications is selecting the list of facets to display to the user following each query. We call this the facet selection problem. When refining a query by drilling down into a facet, documents that are associated with that facet are promoted in the rankings. We formulate facet selection as an optimization problem aiming to maximize the rank promotion of certain documents. As the optimization problem is NPHard, we propose an approximation algorithm for selecting an approximately optimal set of facets per query. We conducted experiments over hundreds of queries and search results of a large commercial search engine, comparing two flavors of our algorithm to facet selection algorithms appearing in the literature. The results show that our algorithm significantly outperforms those baseline schemes.

D. Dynamic Faceted Navigation in Decision Making using Semantic Web Technology.
Categorization in the decision making classifies decision makers' experiences about the world and provides a guide to reach a goal. This implies that dynamically providing categories reflecting the given decision context gives a great enhancement in decision quality. This study discusses the dynamic category selection under the Semantic Web environment, focusing on an implementation of a decision support system, the dynamic facet navigation system working with an ontology. Predefined fixed categories are provided to refine search results.

E. Usability Studies of Faceted Browsing: A Literature Review.
Faceted browsing is a common feature of new library catalog interfaces. But to what extent does it improve user performance in searching within today's library catalog systems? This article reviews the literature for user studies involving faceted browsing and user studies of "next-generation" library catalogs that incorporate faceted browsing. Both the results and the methods of these studies are analyzed by asking, What do we currently know about faceted browsing? How can we design better studies of faceted browsing in library catalogs? The article proposes methodological considerations for practicing librarians and provides examples of goals, tasks, and measurements for user studies of faceted browsing in library catalogs.

III. PROPOSED SYSTEM

we propose an approach for dynamic facet ordering in the e-commerce domain. The focus of our approach is to handle domains with sufficient amount of complexity in terms of product attributes and values. we devise an algorithm that ranks properties by their importance and also sorts the values within each property. The proposed approach is based on a facet impurity measure, regarding qualitative facets in a similar way as classes, and on a measure of dispersion for numeric facets. The property values are ordered descending on the number of corresponding products. Furthermore, a weighting scheme is introduced in order to favor facets that match many products over the ones that match only a few products, taking into account the importance of facets.

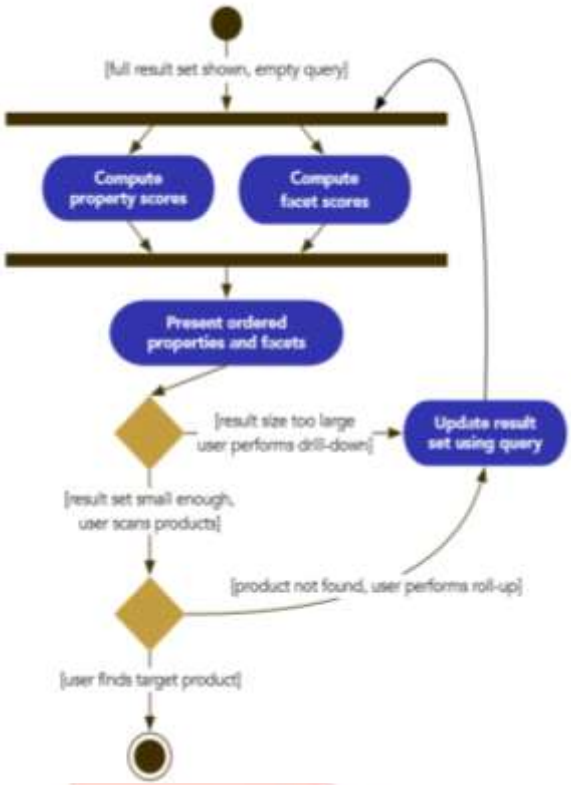

Fig 1: System Architecture

Advantages of Proposed System

1.we employ a weighting scheme based on the number of matching products to adequately handle missing values and take into account the property product coverage.

2.While analyzing the user effort, especially in terms of the number of clicks, we can conclude that our approach gives a better performance than the benchmark methods and in some cases even beats the manually curated 'Expert-Based' approach.

3.In addition, the relatively low computational time makes it suitable for use in real-world Web shops, making our findings also relevant to industry

IV. CONCLUSION

In this work, proposed an approach that automatically orders facets such that the user finds its desired product with the least amount of effort. The main idea of our solution is and then, additionally, also sort the facets themselves. We use different types of metrics to score qualitative and numerical properties. For property ordering we want to rank properties descending on their impurity, promoting more selective facets that will lead to a quick drill-down of the results. Furthermore, we employ a weighting scheme based on the number of matching products to adequately handle missing values and take into account the property product coverage. We evaluate our solution using an extensive set of simulation experiments, comparing it to three other approaches. While analyzing the user effort, especially in terms of the number of clicks, we can conclude that our approach gives a better performance than the benchmark methods and in some cases even beats the manually curated 'Expert-Based' approach. In addition, the relatively low computational time makes it suitable for use in real-world Web shops, making our findings also relevant to industry. These results are also confirmed by a user-based evaluation study that we additionally performed.

V. REFERENCE

[1] H. Zo and K. Ramamurthy, "Consumer Selection of ECommerce Websites in a B2C Environment: A Discrete Decision Choice Model," IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, vol. 39, no. 4, pp. 819–839, 2009.

[2] M. Hearst, "Design Recommendations for Hierarchical Faceted Search Interfaces," in 29th Annual International Conference on Research & Development on Information Retrieval (ACM SIGIR 2006). ACM, 2006, pp. 1–5.

[3]D.Tunkelang,"FacetedSearch,"SynthesisLecturesonInformation Concepts, Retrieval, and Services, vol. 1, no. 1, pp. 1–80, 2009.

[4] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst, "Faceted Metadata for Image Search and Browsing," in Proceedings of the SIGCHI Conference on Human factors in Computing Systems. ACM, 2003, pp. 401–408.

[5] J. C. Fagan, "Usability Studies of Faceted Browsing: A Literature Review," Information Technology and Libraries, vol. 29, no. 2, p. 58, 2010.

[6] M. Hearst, A. Elliott, J. English, R. Sinha, K. Swearingen, and K.P. Yee, "Finding the Flow in Web Site Search," Communications of the ACM, vol. 45, no. 9, pp. 42–49, 2002. [7] B. Kules, R. Capra, M. Banta, and T. Sierra, "What Do Exploratory Searchers Look at in a Faceted Search Interface?" in 9th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2009). ACM, 2009, pp. 313–322.

[8] Amazon.com, "Large US-based online retailer," http://www. amazon.com, 2014.
[9] V. Sinha and D. R. Karger, "Magnet: Supporting Navigation in Semi-structured Data Environments," in 24th ACM SIGMOD International Conference on Management of Data (SIGMOD 2005). ACM, 2005, pp. 97–106. [10] Kieskeurig.nl, "Major Dutch price comparison engine with detailed product descriptions," http://www.kieskeurig.nl, 2014.
[11] Tweakers.net, "Dutch IT-community with a dedicated price comparison department," http://www.tweakers.net, 2014.
[12] Q. Liu, E. Chen, H. Xiong, C. H. Ding, and J. Chen, "Enhancing Collaborative Filtering by User Interest Expansion via Personalized Ranking," IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 42, no. 1, pp. 218–233, 2012.
[13] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, "An Algorithmic Framework for Performing Collaborative Filtering," in 22nd Annual International Conference on Research and Development in Information Retrieval (ACM SIGIR 1999). ACM, 1999, pp. 230–237.
[14] J. Koren, Y. Zhang, and X. Liu, "Personalized Interactive Faceted Search," in 17th International Conference on World Wide Web (WWW 2008). ACM, 2008, pp. 477–486.
[15] G. M. Sacco and Y. Tzitzikas, Dynamic Taxonomies and Faceted Search. Springer, 2009, vol. 25.
[16] D. Dash, J. Rao, N. Megiddo, A. Ailamaki, and G. Lohman, "Dynamic Faceted Search for Discovery-Driven Analysis," in Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM 2008). ACM, 2008, pp. 3–12.
[17] S. Liberman and R. Lempel, "Approximately Optimal Facet Value Selection," Science of Computer Programming, vol. 94, pp. 18–31, 2014.