International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 06 | June -2020 ISSN: 2582-3930

# Dynamic Facet and Ordering for Faceted Product Search Engines

**V. Satish Kumar** Department of Master of Computer Application, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur

**Mr. B. Ajay Kumar** Assistant Professor, M. Tech Department of Information Technology, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur

## 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, favorably compared to a facet list created by domain experts, a greedy approach as baseline, and a state-of-the-art entropy-based solution.

## Introduction

This study examined how searchers interacted with a web-based, faceted library catalog when conducting exploratory searches. It applied eye tracking, stimulated recall interviews, and direct observation to investigate important aspects of gaze behavior in a faceted search interface: what components of the interface searchers looked at, for how long, and in what order. It yielded empirical data that will be useful for both practitioners (e.g., for improving search interface designs), and researchers (e.g., to inform models of search behavior). Results of the study show that participants spent about 50 seconds per task looking at (fixating on) the results, about 25 seconds looking at the facets, and only about 6 seconds looking at the query itself. These findings suggest that facets played an important role in the exploratory search process (7). 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 I Expand. The goal is to build an item-oriented model-based collaborative-filtering framework. The I Expand 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, I Expand 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 I Expand on three benchmark data sets, and experimental results show that I Expand can lead to better ranking performance than state-of-the-art methods with a significant margin (12).

Faceted search is becoming a popular method to allow users to interactively search and navigate complex information spaces. A faceted search system presents users with key value metadata that is used for query refinement. While popular in e-commerce and digital libraries, not much research has been conducted on which metadata to present to a user in order to improve the search experience. Nor are there repeatable benchmarks for evaluating a faceted search engine. This paper proposes the use of collaborative filtering and personalization to customize the search interface to each user’s behavior. This paper also proposes a utility-based framework to evaluate the faceted interface. In order to demonstrate these ideas and better understand personalized faceted search, several faceted search algorithms are proposed and evaluated using the novel evaluation methodology (14).

Multifaceted search is a popular interaction paradigm for discovery and mining applications that allows users to digest, analyze and navigate through multidimensional data. A crucial aspect of faceted search applications is selecting the list of facet values to display to the user following each query. We call this the facet value selection problem. When refining a query by drilling down into a facet value, documents that are associated with that facet value are promoted in the rankings. We formulate facet value selection as an optimization problem aiming to maximize the rank promotion of certain documents. As the optimization problem is NP-Hard, we propose an approximation algorithm for selecting an approximately optimal set of facet values 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 value selection algorithms appearing in the literature. The results show that our algorithm significantly outperforms those baseline schemes (17).


In the era of information revolution, the amount of digital contents is growing explosively with the advent of personal smart devices. The consumption of the digital contents makes users depend heavily on search engines to search what they want. Search requires tedious review of search results from users currently, and so alleviates it; predefined and fixed categories are provided to refine results. Since fixed categories never reflect the difference of queries and search results, they often contain insensible information. This paper proposes a method for the dynamic generation of refining categories under the ontology-based semantic search systems. It specifically suggests a measure for dynamic selection of categories and an algorithm to arrange them in an appropriate order. Finally, it proves the validity of the proposed approach by using some evaluative measures (20).

## System Analysis

## Existing System:

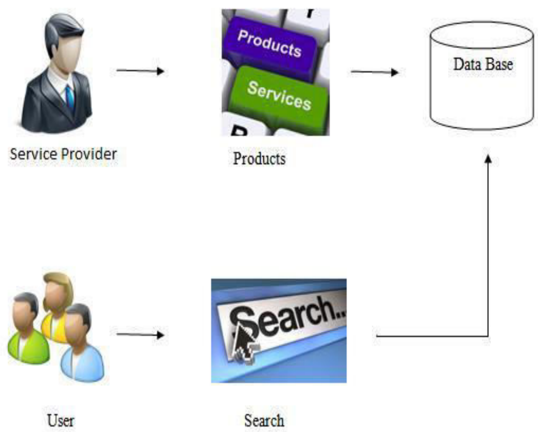
The faceted search system proposed in existing focuses on both textual and structured content. Given a keyword query, the proposed system aims to find the interesting attributes, which is based on how surprising the aggregated value is, given the expectation. The main contribution of this work is the navigational expectation, which is, according to the authors, a novel interestingness measure achieved through judicious application of p-values. These solutions often assume that there is a ranking of the results, based on a preceding keyword-based query or external data, which is often not the case for e-commerce.

## Disadvantage:

Large number of facets are available. Displaying all facets may be a solution when a small number of facets is involved, but it can overwhelm the user for larger sets of facets 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.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 06 | June -2020 ISSN: 2582-3930

## System Design




## System Architecture

## 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. 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 favor facets that match many products over the ones that match only a few products, taking into account the importance of facets.

My solution aims to learn the user interests based on the user interaction with the search engine.

## Advantage:

In our study, we use the common disjunctive semantics for values and conjunctive semantics for properties and take into account the possibility of drill-ups. This means that result set sizes are expected to both increase and decrease during the search session, either by deselecting a facet or choosing an addition facet in a property

In terms of the number of clicks, our approach seems to outperform the other methods, except in the case of the Best Facet Drill-Down Model, where each approach performs equally well. Furthermore, for the Combined Drill-Down Model, our approach results in the lowest number of roll-ups and the highest percentage of successful sessions.

## Implementation:

### Search Sessions

A query in a search session is defined as a collection of previously selected facets. We have decided to apply disjunctive semantics to a selection of facets within a property. For facets across different properties, we use a conjunctive semantics. For example, selecting the facets Brand: Samsung, Brand: Apple, and Color: Black results in (Brand: Samsung OR Brand: Apple) AND Color: Black. Several ecommerce stores on the Web (e.g., Amazon.com and BestBuy.com) use the same principle, which, from a user experience point-of-view, is very intuitive. Our approach assumes that users can undertake two types of actions: drill-down and roll-up. A drilldown is defined as an action of selecting one or more facets, leading to a reduction of the result set size. A roll-up action increases the result set size, which is likely to happen when the user notices that the selected facets are too strict

### Computing Property Scores

In this module, we now discuss the details of computing property scores, shown as one of the first two processes. The outcome of the property scores is used to first sort the properties, after which the facet scores, are used to sort the values within each property. We zoom into the main steps of computing the property score. As shown by the diagram, the score for each property is computed separately and can thus be done in parallel. We designed the proposed algorithm in such a way that more specific facets and properties are ranked higher. To support the algorithm in identifying more specific facets, we introduce the disjoint facet count. This metric is used to compute the score for qualitative properties.

### Product Count Weighting

For numeric properties, we have chosen to use the knowledge about the distribution of the numeric values for computing property scores. It is fairly straightforward to imagine that it may be useful to drill-down using a numeric property when the values for the result set are widely dispersed. When the facets are nearly uniformly distributed over the complete range of values, a drill-down using a user-defined range would lead to a large reduction of the result set. With the Gini impurity and the Gini coefficient, we now have metrics to score both qualitative and numeric properties. This score is independent from the number of products on which it is based. This could possibly lead to problems, as properties that occur within few products will obtain a relatively high score. To compensate for this, we introduce the product count weighting. The product count weighting is used to normalize the Gini indices, resulting in the final property score.

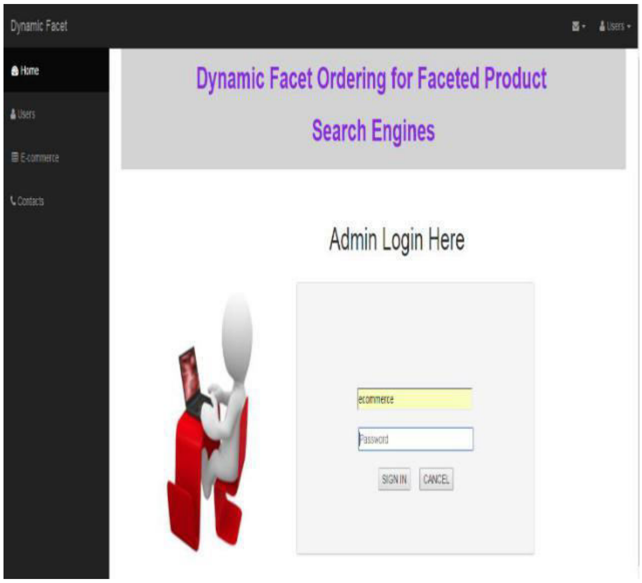
### Computing Facet Scores

In this module, we have explained how we compute scores for properties. We now discuss the details of computing facet scores, shown as one of the first two processes. However, our approach also sorts the values within each property in order to reduce the value scanning effort. This is in contrast to for instance the approach in exiting, which considers property ranking but disregards facets ranking. For numeric properties, value ordering is neglected, as these are often represented with a slider widget in user interfaces. The slider widgets, of which an example, give an indication of the minimum and maximum values for a property, and allow the user to freely define a range of facets within these boundaries. For qualitative properties our approach employs the facet count, ranking facets descending on count, per property. As the target product is unknown to the system, this will increase the

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 06 | June -2020 ISSN: 2582-3930

chance that a facet matching the target product is placed on top.

## Screen Shorts




## System Study:

## Feasibility Study

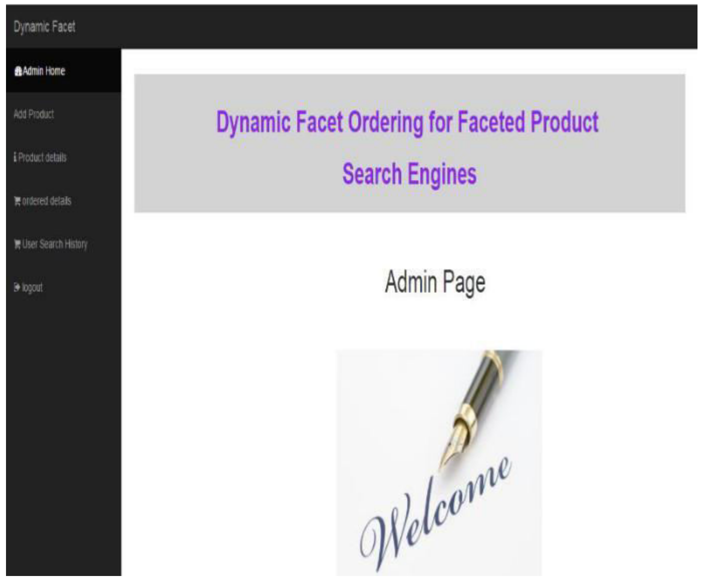
The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

ECONOMICAL FEASIBILITY
TECHNICAL FEASIBILITY
SOCIAL FEASIBILITY

## Economical Feasibility

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and




development of the system is limited. The expenditures must be justified. Thus, the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

## Technical Feasibility

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

## Social Feasibility

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

## Conclusion

In this work, we 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 to sort properties based on their facets 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.

In future we would like to replicate our study on a different domain than cell phones, thereby addressing one of the limitations of the current evaluation. Also, we would like to investigate the use of other metrics, such as facet and product popularity, for determining the order and optimal set of facets.

## References

[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.

[3] D. Tunkelang, "Faceted Search," Synthesis Lectures on Information 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:ALiteratureReview," 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.