Dynamic Facet Ordering for Faceted Product Search Engines

Bhagyashree.S, Bindu.S, Meghana.K, Nisha. H. N, Manjunath S*

Department of CSE, GSSSIETW, Mysuru, Karnataka, India

DOI: https://doi.org/10.21467/proceedings.1.11

* Corresponding author email: manjunaths@gsss.edu.in

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. In this work 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

Index Terms-Facets, e-commerce, Reviews, Ratings, static, dynamic, product search, user interfaces.

1 INTRODUCTION

The work presents a framework for dynamic facet ordering in e-commerce Faceted Browsing is widely used in web shops and product comparison sites. The faceted search system proposed in existing focuses on both on structure and textual content. The main contribution of this work is the navigational expectation which is according to the authors, a novel interesting measure achieved through judicious application of p-values. Most commercial applications that use faceted search manual expert-based selection procedure for facets or relatively static facet list. However, selecting and ordering static facets manually require a significant amount of manual effort. To overcome the limitations of static facet ordering, this work proposes an approach for dynamic facet ordering in the e-commerce domain. Based on measures for

© 2018 Copyright held by the author(s). Published by AIJR Publisher in Proceedings of the 3rd National Conference on Image Processing, Computing, Communication, Networking and Data Analytics (NCICCNDA 2018), April 28, 2018.
This is an open access article under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license, which permits any non-commercial use, distribution, adaptation, and reproduction in any medium, as long as the original work is properly cited. ISBN: 978-81-936820-0-5

specificity and dispersion of facet values the fully automated algorithm ranks those properties and facets on top that leads to quick drill-down for any possible target product. This work aims to learn the user interests based on the user interaction with search engine throughout the search session assume that there exists a single product that user wants to find, and that the user will eventually able to find it. In a large-scale simulation and user study, this 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

2 LITERATURE SURVEY

A Literature Survey is a text of a scholarly paper, which includes the current knowledge including substantive findings, as well as theoreticals and methodological contributions to a particular topic. The Literature Survey is therefore a review of a circumscribed area of a primary literature, i.e. it should draw its main sources of evidence from peer-reviewed journal articles. Anna Kutikova, Jan Balata, Zdenek Mikovec Propsed Explorations into ICT Usage and Behavior in Travel Related Activities of Wheelchair Users People with limited mobility such as wheel chair users have to tackle with barriers of built environment, author focus on information and communication technologies such as ICT system. System includes google maps and handling street views [19]. Adrian Hub, Daniel Blank, Andreas Henrich proposed Picadomo, Faceted Image Browsing for Mobile Devices. Mobile devices are equipped with digital cameras and memory cards. Picadamo technique combines content-based image retrieval and faceted search on mobile devices and it is designed for finding images with desired visual properties or other known meta data [20]. Hangjung Zo and K. Ramamurthy proposed Consumer Selection of E-Commerce Websites in a B2C Environment: A Discrete Decision Choice Model. E-Commerce in a business to consumer environment has been growing rapidly, it influences persived value of the products and consumers website choice behaviour. It is used product attributes and value comparision [1].

Qi Liu, Hui Xiong, Chris H. Q. Ding, Jian Chen proposed Enhancing Collaborative Filtering by User Interest Expansion via Personalized Ranking.A personalized ranking strategy is developed for predicting a user's possible interest expansion. Moreover, a diverse recommendation list isgenerated by using user latent interests as an intermediate layer between the user layer and the item layer [12]. Senjuti Basu Roy, Haidong Wang, Ullas Nambiar, Gautam Das and Mukesh Mohania proposed DynaCet: Building Dynamic Faceted Search Systems over Databases Extracting information and insights from large databases is a time-consuming activity. In this demo, we present DynaCet a domain independent system that provides effective minimum-effort based dynamic faceted search solutions over enterprise databases. At every step, Dynacet suggests facets depending on the user response in the previous step. Facets are selected based on their ability to rapidly drill down to the most promising tuples, as well as on the ability of the user to provide desired values for them [7].

Yongzhong Wu, Mianmian Huang, Yuxin Lu proposed Association Rules and Collaborative Filtering on Sparse Data of a Leading Online Retailer Personalized recommender systems are important for online shopping retailers to recommend items to potential customers. However, data sparsity is a key problem leading to poor recommendations. In this paper, we established two recommender models, i.e., the one based on association rules and the one based on collaborative filtering (CF), and tested them on a large set of sparse data obtained from a Chinese leading online shopping retailer [15].

Yang Chen, XiaoYan Sun, DunWei Gong, Yong Zhang, Jong Choi, and Scott Klasky proposed Personalized Search Inspired Fast Interactive Estimation of Distribution Algorithm and Its Application Interactive evolutionary algorithms have been applied to personalized search, in which less user fatigue and efficient search are pursued. Motivated by this, we present a fast-interactive estimation of distribution algorithm by using the domain knowledge of personalized search. We first induce a Bayesian model to describe the distribution of the new user‟s preference on the variables from the social knowledge of personalized search [8]. Saeed Mohajeri, Davood Rafiei, Hamman W. Samuel, Osmar R. Zaiane proposed BubbleNet: An Innovative Exploratory Search and Summarization Interface with Applicability in Health Social Media We analyse the application of various interfaces to facilitate exploratory search and summarization of documents, especially BubbleNet, an innovative interface for summarizing corpus that also allows discovery of new knowledge that the user may not have previously been looking for. BubbleNet is a visual force directed graph that displays an interactive and dynamic network of topics, semantic relationships, and related documents based on a corpus [18].

3 PROPOSED APPROACH

The proposed approach is 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. The solution aims to learn the user interests based on the user interaction with the search engine.

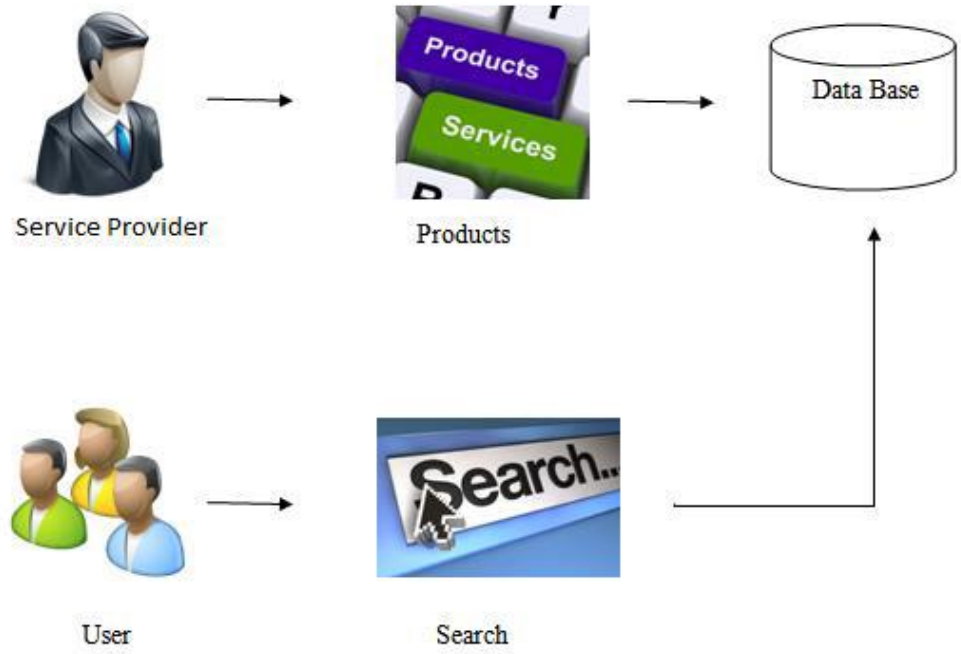

Fig 3.1: System Architechture

4 Methodology

There are three drill-down models that we consider, based on the ones proposed in [14], [17]. These drilldown models rely on five key assumptions, i.e., (1) rationality: the user will end the session once target product is found, (2) practicality: the user will use no more than a fixed number of clicks when looking for the target product, (3) feasibility: the user will perform a roll-up when the target product disappears from the result set, (4) omnisciency: once presented with the facets, the user knows which ones belong to the target product, and (5) linearity: the user scans the properties from top to bottom. Because some of these assumptions are very restrictive, all drill-down models relax one or more of these assumptions. It is, however, useful to identify the theoretical boundaries that may apply to user behavior in order to make a simulation that is more realistic. In the Least Scanning Drill-Down Model, MS, the user u scans the list of facets F starting from the top. When u encounters a facet f ∈ Fdu (a facet associated with the target product), (s)he will select that facet without further scanning.

The Best Facet Drill-Down Model, MB, assumes that when u is searching for du and is scanning F, u identifies the single facet that will reduce the result set size most, while du is still included in the result set. In other words, the user will choose the „best" drill-down option, regardless of the property or facet rank. The Best Facet Drill-Down Model minimizes the number of clicks at the expense of possibly scanning more facets. This is very useful for comparison with the results from the Least Scanning Drill-Down Model. Last, the Combined Drill-Down Model MC provides a more realistic simulation of user behavior by allowing faulty selections (i.e., clicks that will exclude the target product from the result set). This model assumes that the user u scans the list of facets F starting from the top. When u encounters a facet f (s)he will consider selecting f with probability αf when the target product du is associated with this facet, and βf when it is not. For αf and βf we use:

αf =

α |Fp ∩Fdu|

, βf =

β |Fp \Fdu|

(6) where f ∈ Fp and α + β = 1. Once u has a certain facet in consideration, the decision whether to select it will be made stochastically using the Facet Importance Factor γf, defined as follows: γf =(1− rO q (f)−1 |Fdu\q|−1 if f ∈ Fdu (α case) 1 if f 6∈ Fdu (β case) (7) where rO q (f) is a function that returns the rank of f in a list of candidate facets Fdu \q (unselected facets associated with du), and the fraction denominator |Fdu \ q|− 1 is a normalization factor to bring the measure between 0 and 1. When a facet is not selected during a scan, either due to the stochastic effect from αf or βf, or due to its Facet Importance Factor γf, the user will resume scanning the following facet until a selection has been made.

5 EXPERIMENTAL RESULTS

Our approach shows the results for Least Scanning, Best Facet, and Combined Drill-Down models, respectively. We can make several important observations. First, 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-up sand the highest percentage of successful sessions.

Second, we observe that our approach, in most cases, performs best in terms of property and facet scan effort, except for the Combined and Least Scanning Drill-Down Model, respectively. However, although the found differences are statistically significant, it can be argued that they are not relevant, as there were no large effect sizes found. Furthermore, we assume that in practice the property and facet scanning efforts are not the key factors that contribute to the true perceived user effort. We assume that the number of clicks and the responsiveness of the approaches play a much more important role here.

Third and last, in terms of computational time, our approach outperforms the other automatic approaches, often needing orders of magnitude less time to return the sorted facets for a query. For example, the total computation time for the Kim et al. method, on average, is more than 1 second per click. Our approach needs approximately 100 milliseconds per click, which fits the requirements of Web shops and other e-commerce applications, where latencies in terms of seconds are found to be highly undesired [27]. The reason for why the method of Kim et al. is slower stems from the fact that it relies on computing the the conditional entropy for every property pair

6 CONCLUSIONS

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 E -Commerce 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,"SynthesisLectureonInformation 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] Senjuti Basu Roy, Haidong Wang, Ullas Nambiar, Gautam Das and Mukesh Mohania " DynaCet: Building Dynamic Faceted Search Systems over Databases Extracting information and insights from large databases"1084-4627/09 $25.00 © 2009 IEEE DOI 10.1109/ICDE.2009.117

[8] Yang Chen, XiaoYan Sun, DunWei Gong, Yong Zhang, Jong Choi, and Scott Klasky "Personalized Search Inspired Fast Interactive Estimation of Distribution Algorithm and Its Application Interactive evolutionary algorithms" DOI 10.11.09/TEV.2017.2657787,IEEE transaction on evolutionary computation

[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] Yongzhong Wu, Mianmian Huang, Yuxin Lu “ Association Rules and Collaborative Filtering on Sparse Data of a Leading Online Retailer “in Proceedings of the 2017 IEEE IEEM

[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.

[18] Saeed Mohajeri, Davood Rafiei, Hamman W. Samuel, Osmar R. Zaiane “ BubbleNet: An Innovative Exploratory Search and Summarization Interface with Applicability in Health Social Media” 2016 International Conference on Digital Economy (ICDEc)

[19] Anna Kutikova, Jan Balata, Zdenek Mikovec,”Explorations into ICT Usage and Behavior in Travel Related Activities of Wheelchair Users” 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2017) • September 11-14, 2017 • Debrecen, Hungary

[20] Andrian Hub,Daniel Blank,Andreas Henrich “Picadomo,Faceted Image Browsing for Mobile Devices”in CBMI 09 proceedings of the 2009 Seventh international workshop on Content based multimedia indexing,pp 249-254,2009,doi>10.1109/CBMI.2009.34

ISBN: 978-81-936820-0-5

Proceedings DOI: 10.21467/proceedings.1