Listwise v.s. Pagewise: Towards Better Ranking Strategies for Heterogeneous Search Results

Junqi Zhang*
Department of Computer Science and Technology, Institute for Artificial Intelligence, Beijing National Research Center for Information Science and Technology, Tsinghua University
Beijing 100084, China
zhangjq17@mails.tsinghua.edu.cn

ABSTRACT

As heterogeneous verticals account for more and more in search engines, users' preference of search results is largely affected by their presentations. Apart from texts, multimedia information such as images and videos has been widely adopted as it makes the search engine result pages (SERPs) more informative and attractive. It is more proper to regard the SERP as an information union, not separate search results because they interact with each other. Considering these changes in search engines, we plan to better exploit the contents of search results displayed on SERPs through deep neural networks and formulate the pagewise optimization of SERPs as a reinforcement learning problem.

KEYWORDS

Whole-page optimization; Multimedia; Ranking

ACM Reference Format:
Junqi Zhang. 2019. Listwise v.s. Pagewise: Towards Better Ranking Strategies, for Heterogeneous Search Results. In *The Twelfth ACM International Conference on Web Search and Data Mining (WSDM '19), February 11–15, 2019, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3289600.3291596

1 MOTIVATION AND RESEARCH QUESTIONS

Great changes have taken place in Web search recently. A decade ago, search engines returned a list of organic results (one blue hyperlink with short snippet contents). However, to make search engine result pages (SERPs) more informative and attractive, modern search engines incorporate more and more heterogeneous verticals apart from organic results, such as images, videos, news, applications and local maps. Besides, rather than arranging search results as a single list, SERPs are separated into several parts: some contain main search results and the others show additional information such as knowledge cards, advertisements and related searches. These changes recast users' browsing habits to a large extent and lead to three main challenges.

First, homogenous ten blue links have changed to heterogeneous verticals. Results belonging to the same source like news or images are aggregated as a vertical. Different types of search results are presented according to pre-designed structures (as shown in Figure 1). The distinct presentations make it more efficient for users to seek information, and also alter users' preference for search results.

Second, multimedia information is incorporated in the snippet of search results. Apart from texts, images occupy a large proportion of SERPs. For example, images for the query, thumbnails of news and products, and preview frames of videos convey equally or even more important information than texts. Multimedia contents make SERPs more informative and attractive, which makes a vital influence on users' preference. However, this influence is largely neglected by previous work and needs better investigation and mining.

Third, the objective of search engines changes from listwise ranking to pagewise optimization. Many approaches have been proposed to rank search result lists in descending order of relevance scores. However, as the listwise ranking changes to whole-page optimization, more factors need to be considered to assign a proper position for a search result. For example, we need to consider the sizes of the search results and whether they contain redundant information if we want to make the best use of the SERP's space. Because a search result is not isolated but can interact with other results, its potential usefulness for users depends on its context. To take all these factors into consideration, we will design a new pagewise evaluation metric, to properly model the interactions between search results, and a pagewise ranking algorithm, to optimize the pagewise metric.

To address these challenges, we propose three research questions respectively.

RQ1: How can we represent heterogeneous search results in a unified way?

To optimize the whole-page presentation according to the contents of SERPs, an algorithm needs to encode the heterogeneous search results as a unified representation. Search results consist of information items such as texts and images according to some pre-designed structures. Many approaches in computer vision (CV) and natural language processing (NLP) fields can be adopted to represent the semantic information of visual and textual modalities. Besides, the structures of search results are defined by HTML. Thus, it is possible to aggregate the multimedia information items through the parse tree of HTML source codes.

RQ2: How can we exploit multimedia information?

For textual modality, we can exploit the textual semantics of titles and snippets through NLP methodologies. However, not only the general semantics but also the matching between text contents and user's information need that matters in search. For example, if the user searches for the weather, temperature information is essential.

*2nd year Ph.D. student

WSDM '19, February 11–15, 2019, Melbourne, VIC, Australia

ACM ISBN 978-1-4503-5940-5/19/02.
https://doi.org/10.1145/3289600.3291596

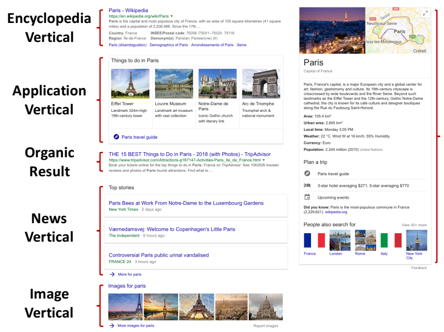

Figure 1: The SERP of query "Paris".

For visual modality, many deep neural approaches have been proposed in computer vision fields to detect and recognize the contents in images. However, Web images are noisy and diverse, which include icon images of companies, cartoon images, images containing texts and so on. How to exploit the open-domain Web images is still challenging.

RQ3: How to optimize whole-page considering interactions between search results?

The optimization of whole SERP is not simply ranking search results according to relevance scores. The position, presentation and context of search results are all vital for users' preference. Search results affect each other in satisfying users' information needs and the information they provide is not isolated but perceived as whole by users. It is necessary to optimize the SERP as an information group, not separate search results.

To optimize and evaluate the organization of the whole SERP, we need to design a pagewise evaluation metric. Whether the most relevant search results are placed at the most attractive positions and the presentation of search results are all crucial for the metric.

2 RELATED WORK

2.1 Whole-page Optimization

The whole-page presentation optimization problem is first introduced by Wang et al.. They optimize item position, image size, text font and other styles to render heterogeneous results onto SERPs. However, they utilize statistical summary features to encode heterogeneous results, which are not capable to capture the complicated presentation and multimedia information. The interactions between search results can also be better exploited. The whole-page optimization is defined as the complex ranking settings by Oosterhuis and de Rijke. They rank both the documents and positions to meet users' preference, but mainly concentrate on listwise optimization. To model the dependencies between documents, Wei et al. adopt Reinforcement Learning to rank as a sequential decision making process.

3 METHODOLOGY

The procedure of whole-page optimization is shown in Figure 2. The first stage is to extract features of search results through deep neural networks. For query q, the features of all search results displayed on the SERP (including the knowledge cards, advertisements, related searches et al.) are denoted as $\mathbf{f} = \{f_i, i = 1, 2, \ldots, N\}$. The objective of whole-page optimization is to find the optimal presented positions of the search results $\mathbf{p} = \{p_i, i = 1, 2, \ldots, N\}$. Given $\mathbf{f}$ and $\mathbf{p}$, the SERP of q is determined.

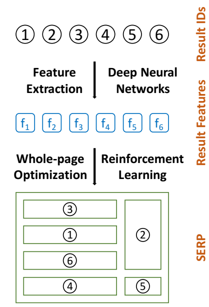

Figure 2: Whole-page optimization.

3.1 Deep Neural Networks

To address the *RQ*1, we adopt Recursive Neural Networks to encode the heterogeneous verticals, while utilizing Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to exploit the multimedia information for *RQ*2.

- Exploit Visual Information by Convolutional Neural Networks. CNNs have show superior performance in extracting visual semantics of images.
- Exploit Textual Information by Recurrent Neural Networks. RNNs have been widely used in natural language processing tasks, which perform well in eliciting high level structure and semantic information of natural language.
- Exploit Structure Information by Recursive Neural Networks. We plan to utilize Recursive Neural Networks to represent the complicated structure of search results. All the information items can be recursively aggregated by the parse tree of the search result.

3.2 Reinforcement Learning

For *RQ*3, we formulate the whole-page optimization as a reinforcement learning problem [2]. There are two possible ways:

(1) The optimization is regarded as a one-step learning procedure as in [2]. The whole-page presentation $\mathbf{p}$ is determined by the policy function with only one action.

(2) The optimization is regarded as a multi-step learning procedure as in [1, 3]. With each action, the policy function decides the position for only one search result. The whole SERP is filled by search results step-by-step.

To make the policy function rank search results not only according to the relevance, but also the interactions between them, we need to design an effective reward. Users's clicks can be regarded as the implicit feedback of interactions between search result. Thus, we design a reward function considering both the relevance scores of search results and the clicks of users, which is defined as $reward = \delta(\mathbf{relevances}, \mathbf{clicks} \mid \mathbf{f}, \mathbf{p})$. The optimization objective can be achieved by the maximization of the expected cumulative reward.

REFERENCES

[1] Harrie Oosterhuis and Maarten de Rijke. 2018. Ranking for Relevance and Display Preferences in Complex Presentation Layouts. *arXiv preprint arXiv:1805.02404* (2018).
[2] Yue Wang, Dawei Yin, Luo Jie, Pengyuan Wang, Makoto Yamada, Yi Chang, and Qiaozhu Mei. 2018. Optimizing Whole-Page Presentation for Web Search. *ACM Transactions on the Web (TWEB)* 12, 3 (2018), 19.
[3] Zeng Wei, Jun Xu, Yanyan Lan, Jiafeng Guo, and Xueqi Cheng. 2017. Reinforcement Learning to Rank with Markov Decision Process. In *ACM SIGIR 2017*.