

Chapter 8

Rough Set Based Aggregation for Effective Evaluation of Web Search Systems

Rashid Ali

Aligarh Muslim University, India

M. M. Sufyan Beg

Jamia Millia Islamia, India

ABSTRACT

Rank aggregation is the process of generating a single aggregated ranking for a given set of rankings. In industrial environment, there are many applications where rank aggregation must be applied. Rough set based rank aggregation is a user feedback based technique which mines ranking rules for rank aggregation using rough set theory. In this chapter, the authors discuss rough set based rank aggregation technique in light of Web search evaluation. Since there are many search engines available, which can be used by industrial houses to advertise their products, Web search evaluation is essential to decide which search engines to rely on. Here, the authors discuss the limitations of rough set based rank aggregation and present an improved version of the same, which is more suitable for aggregation of different techniques for Web search evaluation. In the improved version, the authors incorporate the confidence of the rules in predicting a class for a given set of data. They validate the mined ranking rules by comparing the predicted user feedback based ranking with the actual user feedback based ranking. They show their experimental results pertaining to the evaluation of seven public search engines using improved version of rough set based aggregation for a set of 37 queries.

INTRODUCTION

In modern society, we can see effect of information technology everywhere. In industrial environment, information technology is being used in one or another way at every step from manufacturing

of goods to marketing and maintenance of the product. Through, e-commerce, big companies are trying to reach the customers in every part of the e-connected world. They are also paying the search engines providers to have their products advertisements placed at a higher position in the search results returned in response to the queries

DOI: 10.4018/978-1-4666-0294-6.ch008

of the target customers. In such circumstances, it is worthy to evaluate search results. Rank Aggregation is the problem of combining given set of rankings from different voters into a single ranking list, which represents consensus. This finds applications in various fields. For example, in sports, it may be used to get an overall ranking of teams from the rankings by different judges. In academics, it can be used to obtain ranking of universities on the whole from the rankings done on the basis of different performance measuring parameters. In industrial environment, it may be used to grade a set of products completely from the grading by a number of parameters like cost, weight, volume etc. It can also be used to select suppliers of the raw material on the basis of different criteria like quality system outcome, on time delivery etc. When applied to the web, this finds applications in meta-searching, spam fighting and evaluating search systems, searching for multiple terms using word association techniques and combining different ranking functions. In this paper, we discuss Rank aggregation in context of the evaluation of Web search systems.

This paper is organized as follows. In section 2, we first discuss the importance of Web searching in industrial environment. Then, we list some important definitions and discuss related work in the area. In section 3, we discuss the basics of rough set theory and mention the advantage of using rough set theory over fuzzy set theory in data analysis. In section 4, we first mention the difficulty in using the rough set based rank aggregation for the evaluation of Web search systems. Then, we discuss the details of the improved version of Rough set based rank aggregation technique in section 5. We present our experimental results in section 6. Finally, we conclude in section 7.

BACKGROUND AND RELATED WORK

Here, we first discuss *industrial web search*.

Industrial Web Search

Web searching is very important in an industrial environment. A large number of online users are using Web search for business information. It is an emerging trend for growing number of people to search for the best product or the products that suit best to their pocket. That means Web searching influences highly the consumers' buying decisions. In turn, this means to industrial marketers or sellers to develop and manage their business presence online for the Web search for the significant business growth and development. On their websites, they must provide information about events, offers and promotions, allow for customer reviews, and include photos to help users familiarize themselves with their products and services. Since, a large number of users, use search engines for finding products, it is essential for the industrial marketers to approach search engines to make their products placed high in the search results. Their presence in the search results ensures that they are reaching their targeted customer groups. This motivates many industrial companies to pay search engines for advertisements. Now, there are a large number of search engines available. Therefore, evaluation of web search engines will help the industrial marketers to select best search engines for the advertisements and they will invest their money for advertisements sensibly.

Important Definitions

Now, we list some of the important definitions.

Definition 1. Rank Aggregation Problem: *Given a set of n candidates say $C=(C_1, C_2, C_3, \dots, C_n)$, a set of m voters say $V=(V_1, V_2, V_3, \dots, V_m)$, and a ranked list l_i on C for each voter i . Then, $l_i(j) < l_i(k)$ indicates that the voter i prefers the candidate j to k . The **rank aggregation problem** is to combine the m ranked lists $l_1, l_2, l_3, \dots, l_m$ into a single list of candidates, say l that represents the collective choice of*

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/rough-set-based-aggregation-effective/64721

Related Content

Two-Decision-Maker Conflict Resolution with Fuzzy Preferences

Mubarak S. Al-Mutairi (2014). *International Journal of Applied Industrial Engineering* (pp. 40-59).
www.irma-international.org/article/two-decision-maker-conflict-resolution-with-fuzzy-preferences/138308

Redesign of the Workplace for Toolmakers Towards Industry 4.0

Ivana Radi, Bojan Rupnik, Simona Šinko, Tomaž Kramberger and Brigita Gajšek (2021). *Research Anthology on Cross-Industry Challenges of Industry 4.0* (pp. 1333-1352).
www.irma-international.org/chapter/redesign-of-the-workplace-for-toolmakers-towards-industry-40/276878

Education in the Era of Industry 4.0: Qualifications, Challenges, and Opportunities

Dharmendra Trikam Lal Patel (2021). *Research Anthology on Cross-Industry Challenges of Industry 4.0* (pp. 1647-1665).
www.irma-international.org/chapter/education-in-the-era-of-industry-40/276894

User Acceptance of eGovernment Services: Analysis of Users' Satisfaction Level Based on Technology Acceptance Model

Serdar Yarikas, Ibrahim Arpacı and Gülgün Afacan (2013). *Industrial Dynamics, Innovation Policy, and Economic Growth through Technological Advancements* (pp. 348-362).
www.irma-international.org/chapter/user-acceptance-egovernment-services/68368

Application of Three Meta-Heuristic Algorithms for Maximizing the Net Present Value of a Resource-Constrained Project Scheduling Problem with Respect to Delay Penalties

Masoud Rabbani, Azadeh Arjmand, Mohammad Mahdi Saffar and Moeen Sammak Jalali (2016). *International Journal of Applied Industrial Engineering* (pp. 1-15).
www.irma-international.org/article/application-of-three-meta-heuristic-algorithms-for-maximizing-the-net-present-value-of-a-resource-constrained-project-scheduling-problem-with-respect-to-delay-penalties/159082