

Result Analysis of Xml Context Diversified Search

Simon Farande
Computer Science and Engineering
AMGOI
Wathar, India.

D.S.Bhosle
Computer Science and Engineering
AMGOI
Wathar, India.

Abstract: user searches a keyword with short and vague keywords and large amount of irrelevant results are shown which are irrelevant. This paper focuses on diversified XML keyword search based on its different contexts in the XML data. If the user types short and vague query it is searched in XML data, appropriate suggestions are given based on context. Then relevant results are obtained based on suggestions.

Keywords—xml; context; diversified; search; insert.

I. INTRODUCTION

Information retrieval based on structured and semi-structured data has recently gain popularity. When the given keyword query only contains a small number of vague keywords, it would become a very challenging problem to derive the user's search intention due to the high ambiguity of this type of keyword queries. Although sometimes user involvement is helpful to identify search intentions of keyword queries, a user's interactive process may be time consuming when the size of relevant result set is large. To address this, Jianxin Li, Chengfei Liu *Member*, and Jeffrey Xu Yu [1] proposed a method of providing diverse keyword query suggestions to users based on the context of the given keywords in the data to be searched. By doing this, users may choose their preferred queries or modify their original queries based on the returned diverse query suggestions. This paper focuses on Result analysis of [2] and [3].

II. SURVEY

Various types of search engine provide different types of results, but one thing common about them is they don't provide diverse query suggestions. Large amount of irrelevant results are shown to the user.

III. CONTEXT DIVERSIFICATION OF RESULTS

When the user searches the keyword from the system, the results are displayed w.r.t. query and the suggestions are given as per the mutual scores of the given keyword search intentions. If the search intention is successful then it is stored for future use, thus automatically updating new query suggestions. The results are pruned if the user decides to go with diversified results, showing only relevant data.

IV. METHODOLOGY

XML context diversified search as mentioned in [2] and [3] works great, but for this paper it is tested on DBLP (digital bibliography library project) dataset [4], for each XML dataset used, select some terms based on the following criteria: (a) a selected term should often appear in user-typed keyword queries;(b) a selected term should highlight different semantics when it co-occurs with feature terms in different contexts. DBLP dataset is used to measure the effectiveness of diversification, here user searches the keyword from DBLP and if the result is successful then it is stored for future use, the keywords from user query are again searched for relevant suggestions for eg. If the user searches "parallel integer", the search is again made for keywords "parallel and "integer", so on searching these two keywords the suggestions would get as *parallel processing, integer sorting, parallel computation* etc. These are called diversified results, user can modify the queries based on diversified results or can directly work on suggestions.

V. RESULT ANALYSIS

As mentioned in the methodology the results we get are based on mutual information of keywords, and the diversified suggestions are based on each keyword for eg. If the user searches "parallel integer" then the diversified query suggestions are shown in Table 1

Keyword	Count
Parallel integer	8
Integer	786
Parallel processing	104
Parallel computation	102
Parallel recognition	4
Integer sorting	6

Table I. Diversified query suggestions

The diversified results shown in Table I. can be shown in column chart as shown in Fig. 1.

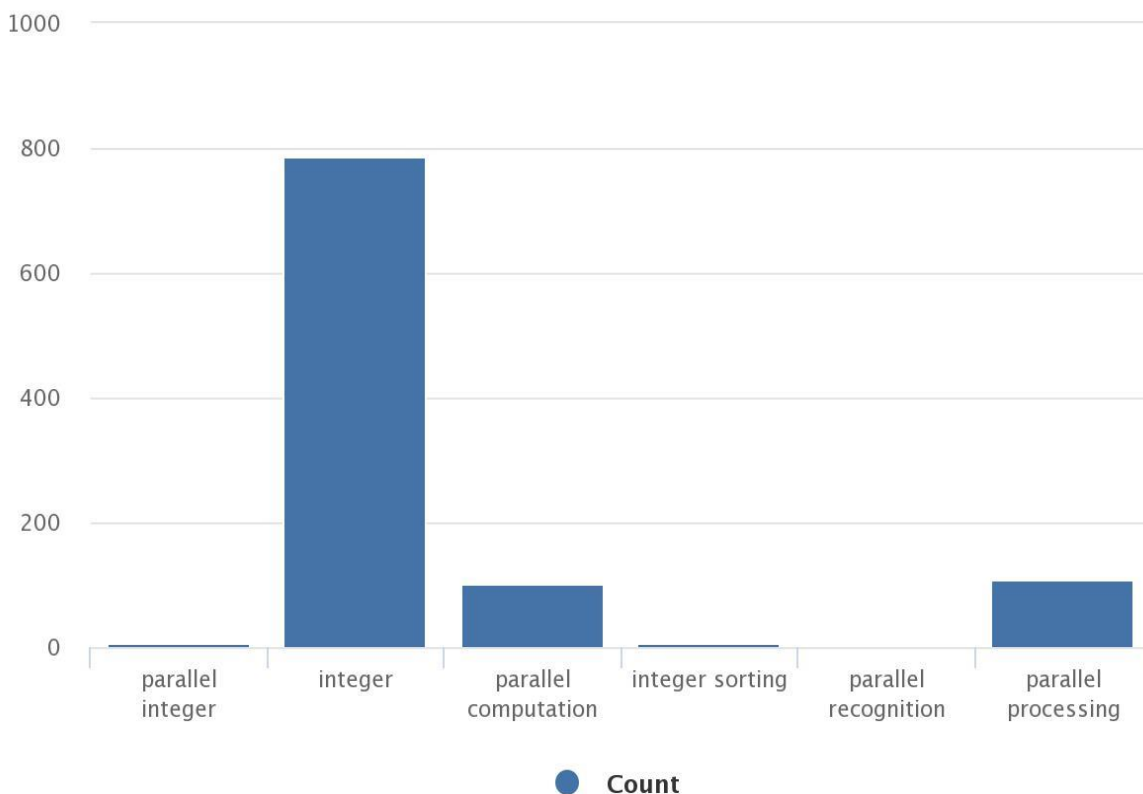


Fig.1 “Parallel integer” Diversified result column chart

The Fig 1.shows the diversified search intentions and the count of results that are obtained.

REFERENCES

[1] Jianxin Li, Chengfei Liu and Jeffrey Xu Yu “Context-based Diversification for Keyword Queries over XML Data” IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING , VOL. 26, NO. 5, MAY 2014

[2]Simon Farande, Dr.D.S.Bhosale, "Survey Paper on XML context diversified search" in International Journal of Advance Research in Science and Engineering Vol. No.5,Special Issue No.01, March 2016,IJARSE,ISSN 2319-8354,page 1-4.

[3] Simon Farande,Dr.D.S.Bhosale "Design and Implementation of XML context diversified search" *International Research Journal of Engineering and Technology (IRJET)*, Volume: 03 Issue: 06 | June-2016 [e-ISSN: 2395 -0056, p-ISSN: 2395-0072]

[4] “http://dblp.uni-trier.de/xml/.”