Chapter 1
Top–k Relevant Term Suggestion Approach for Relational Keyword Search

Xiangfu Meng
Liaoning Technical University, China

Xiaoyan Zhang
Liaoning Technical University, China

Chongchun Bi
Liaoning Technical University, China

ABSTRACT

This chapter proposes a novel approach, which can provide a list of keywords that both semantically related to the application domain and the given keywords by analyzing the correlations between query keywords and database terms. The database term is first modeled as <attribute, keyword> and suppose each query keyword can map into a database term. Then, a coupling relationship measuring method is proposed to measure both term intra- and inter-couplings, which can reflect the explicit and implicit relationships between terms in the database. Based on the coupling relationships between terms, for a given keyword query, an order of all terms in database is created for each query keyword and then the threshold algorithm (TA) is leveraged to expeditiously generate top-k ranked semantically related terms. The experiments demonstrate that our term coupling relationship measuring method can efficiently capture the semantic correlations between query keywords and terms in database.

1. INTRODUCTION

Keyword query is becoming a very popular way to obtain the information from the relational database along with its wide spread use on the Web. In real applications, however, most of common Web database users usually have insufficient knowledge about the database content and schema, and they are also lack of keywords related to the searching domain. Thus, it is not easy for them to find appropriate keywords

to express their query intentions. To explore the database, the user may issue a query with a few general keywords at first, and then gradually refines the query through observing the query results. In such an iteration, the user needs to check each result to identify whether it is related to his interest or not, which is a time-consuming and tedious work.

Consider a DBLP database consisting of 3 relations connected through primary-foreign-key relationships shown in Figure 1.

Suppose a master student who is a XML beginner just knows a few keywords about XML research field and wants to find chapters about the XML search techniques from DBLP website. Based on the DBLP database, he/she would issue a query Q containing keywords "XML, search". On receiving the query Q, the traditional keyword search approach will return a set of minimal total joint networks (MTJNTs), each of which

1. Is obtained from a single relation or by joining several relations, and
2. Contains all the query keywords.

Since there are too many chapters containing keywords "XML" and "search" in DBLP dataset, there are too many MTJNTs in the query results. In such a case, the user would like the system suggest a list of keywords that are semantically related to Q in order to reduce the searching scope. From Figure 1, it is clearly that the author "Jeffrey" and keywords "XPath", "XQuery", and "twig pattern" are very relevant to Q. That means these terms can refine Q to formulate a more selective query. As an example, the user would execute a query Q'=[Jeffrey, XML, search] to retrieve only the chapters of author Jeffrey on XML searching and the query results are "$a1 \bowtie w1 \bowtie p1$" and "$a1 \bowtie w2 \bowtie p4$". Additionally,

Figure 1. An example of DBLP database
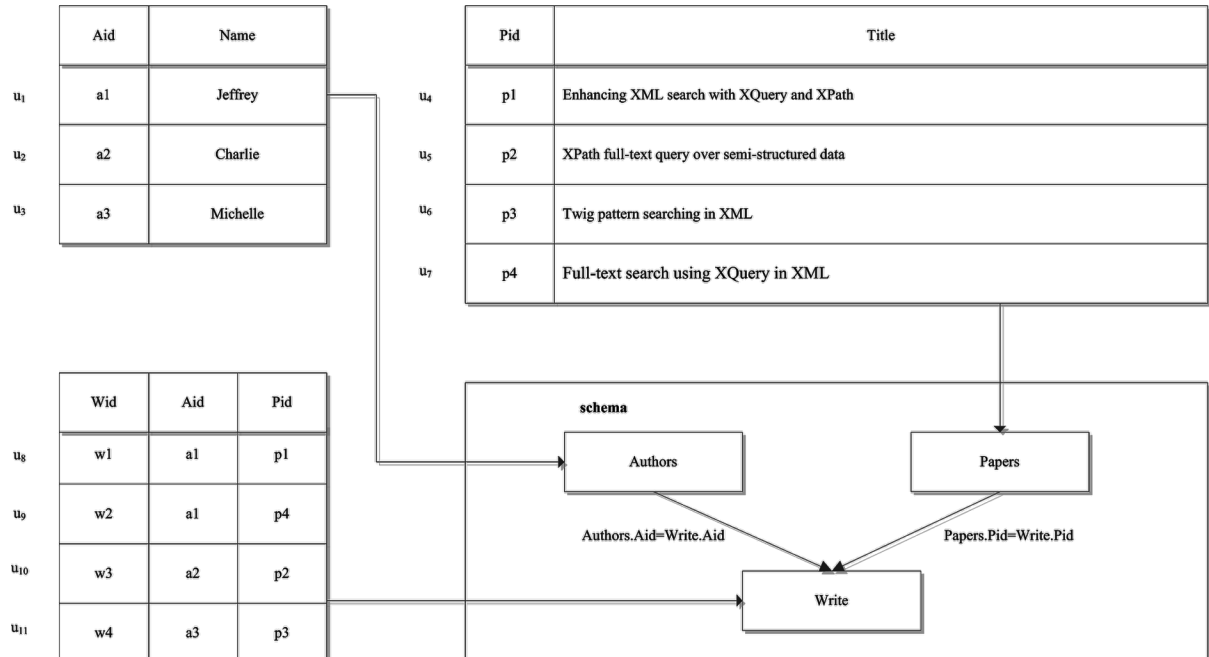
Related Content

User-Oriented IR Research Approaches
Iris Xie (2008). *Interactive Information Retrieval in Digital Environments (pp. 1-28).*
www.irma-international.org/chapter/user-oriented-research-approaches/24522

Personalization Approaches for Ranking: A Review and Research Experiments
Madhuri A. Poteyand Pradeep K. Sinha (2017). *International Journal of Information Retrieval Research (pp. 1-16).*
www.irma-international.org/article/personalization-approaches-for-ranking/165376

Current Status, Requirements, and Challenges of Blockchain Application in Land Registry
Mohammed Shuaib, Shadab Alam, Rafeeq Ahmed, S. Qamar, Mohammed Shahnawaz Nasirand Mohammad Shabbir Alam (2022). *International Journal of Information Retrieval Research (pp. 1-20).*
www.irma-international.org/article/current-status-requirements-and-challenges-of-blockchain-application-in-land-registry/299934

Current Status, Requirements, and Challenges of Blockchain Application in Land Registry
Mohammed Shuaib, Shadab Alam, Rafeeq Ahmed, S. Qamar, Mohammed Shahnawaz Nasirand Mohammad Shabbir Alam (2022). *International Journal of Information Retrieval Research (pp. 1-20).*
www.irma-international.org/article/current-status-requirements-and-challenges-of-blockchain-application-in-land-registry/299934

Question Answering
Ivan Habernal, Miloslav Konopíkand Ondrej Rohlík (2012). *Next Generation Search Engines: Advanced Models for Information Retrieval (pp. 304-343).*
www.irma-international.org/chapter/question-answering/64431