International Journal of Web Portals
Volume 9 • Issue 2 • July-December 2017

Interrogation Based on Semantic Annotations:
Context-Based Construction of Formal Queries from Keywords

Latreche Abdelkrim, EEDIS Laboratory, Djillali Liabes University, Sidi Bel Abbes, Algeria
Lehireche Ahmed, EEDIS Laboratory, Djillali Liabes University, Sidi Bel Abbes, Algeria
Kadda Benyahia, Laboratory Technology of Communication, University of Tahar Moulay. Saida, Algeria

ABSTRACT

Traditional information search approaches do not explicitly capture the meaning of a keyword query, but provide a good way for the user to express his or her information needs based on the keywords. In principle, semantic search aims to produce better results than traditional keyword search, but its progression has retarded because of to the complexity of the query languages. In this article, the authors present an approach to adapt keyword queries to querying the semantic web based on semantic annotations: the approach automatically construct structured formal queries from keywords. The authors propose a new process where they introduce a novel context-based query autocompletion feature to help the users to construct their keywords query by suggesting queries given prefixes. They also address the problem of context-based generating formal queries by exploiting user's query history, where previous queries can be used as contextual information for generating a new query. With the first tests, the authors' approach achieved encouraging results.

KEYWORDS

Auto-Completion, Keyword Search, Query Context, RDF Graphs, Semantic Annotation, Semantic Search

1. INTRODUCTION

Traditional information search approaches do not explicitly capture the meaning of a keyword query, but provide a good way for the user to express his or her information needs based on the keywords. Semantic search through the exploitation of semantic annotations has become a highly explored way and aims to produce better results than traditional keyword research. Currently, there is billions of RDF[1] triplets stored in multiple Web data sources from different domains connected by millions of RDF links. This evolution allows users to express more complex information needs. However, the progress of semantic research has been retarded because of to the complexity of query languages e.g. SPARQL[2]). However, for users to use these query languages, they must master the complex representation of formal logic and be familiar with the structure and content of the underlying ontology. This creates an obstacle between semantic search and end users. Therefore, it is important to allow users to perform semantic searches simply by entering keyword queries. To adapt keyword query to semantic search, the following obstacles can be overcome: (1) Vocabulary lag: Traditional web users usually have no knowledge about the content and structure of the KB (Annotations and underlying

DOI: 10.4018/IJWP.2017070103

ontologies), (2) Absence of relations: relations between KB resources must be explicitly formulated in semantic queries, which often are not stated in user's keyword queries.

To overcome the obstacles mentioned above, various solutions on the translation of keyword queries into semantic queries have been proposed (Royo et al., 2005; Bernstein et al., 2005; Stojanovic et al., 2003; Bobed & Mena, 2016). Graphical based search (Athanasis et al., 2004; Lei et al., 2006; Kandogan et al., 2006; Li et al., 2007; Tran et al., 2007) also contributes in this way to create graph queries by exploiting and browsing ontology. Other relevant works in some context are (Tran et al., 2011; Wang et al., 2008; Zhou et al., 2007). In these systems, the queries are generated from the all the paths derived from an RDF graph. In the same context, other more relevant systems (Tran et al., 2009; Fu et al., 2011; Teng & Zhu, 2015; Le et al., 2014; Latreche et al., 2009) have been realized, where authors present, keyword search approaches on the RDF data graph. In these approaches, the general process of constructing formal requests consists of the following phases: (1) Keywords to semantic entities (elements defined in KB) mapping; (2) Constructing query graphs connecting previously detected elements by exploring the KB; (3) ranking the constructed queries;

Each approach has focused on one or more steps of this process.

In the systems discussed above, in particular, systems for keyword search on RDF data graph, it is observed that: (1) the end users express their need for information intuitively, without having prior knowledge of vocabulary, knowledge representation formalism or the query language used; (2) users find it difficult to correctly formulate their keyword query which satisfies their information needs. Then, it is important to incorporate techniques that help users to formulate correctly their keyword queries; (3) Some queries are more ambiguous than others; which implies that for a given keyword query a multitude of formal queries are generated, it is therefore necessary to propose effective techniques for solving this type of problem; (4) the calculation of all possible combinations of keyword interpretations, greatly increases the time complexity of the exploration algorithms; (5) most of these approaches produce most popular query interpretations, but may not be desired by the user.

All these difficulties contribute to reducing the performance of these systems. Then, faced with these difficulties, we propose our novel approach, which can automatically generate context-aware formal queries from keyword queries on the graph-structured data, and particularly RDF data model, by exploiting information from a user's query history, in order to allow users accustomed to formulating keyword queries to make semantic searches transparently on the basis of semantic annotations.

We are making the following contributions to meet these challenges:

1. Introduce our novel process, which can context-based generation of formal queries from keyword queries on the graph-structured data. This task requires the following steps: (1) context-based query auto-completion; (2) Top-k query graphs construction; (3) query graphs mapping; (4) weighting model (update scores).
2. Introduce and formalize novel context-based query autocompletion (QAC) functionality. QAC to helps user to construct his keywords query (sets of terms) by suggesting queries given prefixes. Given a prefix, possible query completions are ranked according to a predefined criterion, and then some of them are returned to the user.
3. Exploit user's previous queries by using them as contextual information to generate a new query.
4. Design and implement an efficient graph exploration algorithm that enhances existing algorithms.
5. Present a complete evaluation and demonstrate that our approach produces better results than those of the state-of-the-art.

The rest of this paper is organized as follows: In Section 2 defines the problem of formal queries construction. We then discuss specific aspects of data and query processing in more detail in section 3. In section 4 we detail context-based query auto-completions, formal queries construction in section 5, and weighting model in Section 6. Implementation and experimental results are presented in Section 7, a discussion of related work in Section 8. We conclude and future work in the last section.