Information Sciences 303 (2015) 101–119

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Mapping evolution of dynamic web ontologies

A.M. Khattak [a], Z. Pervez [b], W.A. Khan [c], A.M. Khan [d], K. Latif [e], S.Y. Lee [c,*]

[a] College of Technological Innovations, Zayed University, United Arab Emirates
[b] School of Computing, University of the West of Scotland, United Kingdom
[c] Department of Computer Engineering, Kyung Hee University, Republic of Korea
[d] Department of Computer Science, Innopolis University, Russia
[e] School of Electrical Engineering and Computer Science, NUST, Pakistan

ARTICLE INFO

Article history:
Received 31 October 2013
Received in revised form 11 December 2014
Accepted 27 December 2014
Available online 3 January 2015

Keywords:
Ontology change
Change management
Change history
Mapping reconciliation

ABSTRACT

Information on the web and web services that are revised by stakeholders is growing incredibly. The presentation of this information has shifted from a representational model of web information with loosely clustered terminology to semi-formal terminology and even to formal ontology. Mediation (i.e., mapping) is required for systems and services to share information. Mappings are established between ontologies in order to resolve terminological and conceptual incompatibilities. Due to new discoveries in the field of information sharing, the body of knowledge has become more structured and refined. The domain ontologies that represent bodies of knowledge need to be able to accommodate new information. This allows for the ontology to evolve from one consistent state to another. Changes in resources cause existing mappings between ontologies to be unreliable and stale. This highlights the need for mapping evolution (regeneration) as it would eliminate the discrepancies from the existing mappings. In order to re-establish the mappings between dynamic ontologies, the existing systems require a complete mapping process to be restructured, and this process is time consuming. This paper proposes a mapping reconciliation approach between the updated ontologies that has been found to take less time to process compared to the time of existing systems when only the changed resources are considered and also eliminates the staleness of the existing mappings. The proposed approach employs the change history of ontology in order to store the ontology change information, which helps to drastically reduce the reconciliation time of the mappings between dynamic ontologies. A comprehensive evaluation of the performance of the proposed system on standard data sets has been conducted. The experimental results of the proposed system in comparison with six existing mapping systems are provided in this paper using 13 different data sets, which support our claims.

1. Introduction

The increasing amount of information available on the web places a heavy computational load on the systems that are designed to access, interpret, manipulate, maintain, merge, integrate, infer, and mine this information [21]. The fundamental requirement of information exchange among applications, systems, system agents, and web services is the development of a

* Corresponding author. Tel.: +82 31 201 2514; fax: +82 31 202 2520.
E-mail address: sylee@oslab.khu.ac.kr (S.Y. Lee).

http://dx.doi.org/10.1016/j.ins.2014.12.040
0020-0255/

consistent and comprehensive model for knowledge representation, which is essential for the sharing of knowledge pertaining to research outcomes, sharing information among independent organizations [6], and the exchange of information among healthcare systems [31] and among heterogeneous systems and services [3]. In order to make the sharing of information possible, there is a need to model the information more appropriately while preserving its semantics.

Ontology provides a formal structure (model) with semantics with regard to how an expert perceives the domain of interest. Ontology is defined as *a formal, explicit specification of a shared conceptualization*. Ontology is the main source of semantic web information and its services, which helps to clearly define the meaning of resources and achieve a better understanding of the work that is shared between a human and computer systems [35,40]. Service-Oriented Architecture (SOA) and Semantic Web Services Technology are becoming more mature and are now widely used [10]. The meaningful information and the machine interpretable information that is contained in ontology helps to create semantic web services that are automated with regard to service discovery, selection, and interoperability [15].

Current web information can be viewed as the evolution of traditional web information, which ranges from a collection of web pages to the integration of those pages with services that these sites can use to interoperate with one another. Interoperability is both collaborative and multifaceted and is needed to overcome the problems of incompatibilities among organizations, structures, data, architecture, services, and business rules [51]. However, since the data, architecture, and services are usually provided by autonomous parties, often high interface, structural, and semantic heterogeneities exist with regard to information storage and exchange [8,14,18,20,21,26,32,35,37,43,44,49,56,63]. In order to overcome this issue, we utilize the value of data and schema mapping [8,11,18,32,37,44,47,49,52]; in other words, the mapping among schema or ontology elements is the definition of semantic relatedness. Use of ontology in systems dealing with information extraction from a large and complex structured source of information and web services can yield valuable results [4,8,11,18,20,32,47,59]. The increased use of ontology in Information Systems and Knowledge Sharing Systems also increases the significance of ontology maintenance [21,37]. However, the large and complex structure and the decentralized nature of the web compel communities to create their own ontologies to represent information [14,21,59]. Thus, mediation among distributed and autonomous sources is required for exchange of information [8,18,21,32,44,49,63].

The number of information sources is increasing significantly, and this increases the importance of having a sophisticated mechanism to extract information and to manage the heterogeneity among these information sources. Mediation (mapping) is used to align two or more ontologies (information sources) for the purpose of information sharing [5,8,32,35,44,46,49,57,63]. These mappings are generated by mapping systems with two main concerns: accuracy and efficiency (the time required to produce the mappings). Existing mapping systems, such as Falcon [32], FOAM [18], Lily [63], AgreementMaker [12,13], Prompt [49], H-Match [8], and MAFRA [44], are currently considered the best matching and mapping systems. These systems consume a lot of time when mapping large knowledge databases such as Google Classification,[1] Wiki Classification,[2] ACM Classification Hierarchy,[3] and MSC Classification Hierarchy.[4] Data-sources are provided by autonomous and independent providers, which means that these data-sources evolve independently from one another and with flexible structures [27]. This results in a change to the existing mapping methodologies, which makes these mappings unreliable with regard to the sharing of information. This is why there is a need for a system that supports mapping for evolving ontologies. Existing systems complete the mapping process by completely re-creating the mappings among the evolved ontologies, which is a very time consuming process.

Re-creation of mappings is required for mapped ontologies that are dynamic and subject to change. Existing systems take more time to re-create mappings as compared to the process of creating the initial mappings as these systems start the mapping process from scratch; however, the changes in the mapped schemas and regenerated mediation are not significant [27]. Consequently, a less time consuming scheme that can be used in the reconciliation of ontology mappings (mapping evolution) in dynamic and evolving ontologies is proposed in this research paper in order to support information exchange and reliable service interoperability. The hypothesis of the proposed approach is to only consider the changed resources in the mapping regeneration process that will not only reduce the time required for mapping regeneration but will also support updated and reliable mappings for information sharing and eliminate stale mappings while preserving the same level of accuracy. To achieve this, our approach uses the Change History Log (CHL) [38] (i.e., local, centralized, and distributed) to map reconciliation in less time than existing systems. The proposed technique drastically reduces the time required for the re-creation of mappings between dynamic ontologies. The CHL is used to store the changes occurring in dynamic ontologies, which are later used for mapping reconciliation. The use of the CHL in ontology matching/mapping helps in the reconciliation of mappings in dynamic and evolving web ontologies by overcoming the staleness problem associated with these mappings and reducing the time required to reconcile these mappings. During the reconciliation of ontology mapping, only the outdated mappings are updated, which saves both time and resources. We have tested the Falcon, Lily, FOAM, Prompt, AgreementMaker, and H-Match algorithms on 13 different data sets that are available online and then extended these algorithms with the proposed scheme by incorporating the use of CHL. Our proposed extensions have been tested on the same data sets and have shown a drastic reduction in the amount of time required for the reconciliation of these mappings. Detailed experimental results that support our claims are provided in this paper.

[1] http://www.google.com/Top/Reference/Libraries/Library_and_Information_Science/Technical_Services/Cataloguing/Classification/.
[2] http://en.wikipedia.org/wiki/Taxonomic_classification.
[3] http://www.acm.org/about/class/1998/.
[4] http://www.math.niu.edu/~rusin/known-math/index/index.html.