International Journal of Strategic Information Technology and Applications
Volume 10 • Issue 2 • April-June 2019

Large-Scale Ontology Alignment- An Extraction Based Method to Support Information System Interoperability

Mourad Zerhouni, EEDIS Lab., University Djilali Liabes, Sidi Bel Abbès, Algeria

Sidi Mohamed Benslimane, LabRI Laboratory, Ecole Superieure en Informatique, Sidi Bel Abbes, Algeria

https://orcid.org/0000-0002-7008-7434

ABSTRACT

Ontology alignment is an important way of establishing interoperability between Semantic Web applications that use different but related ontologies. Ontology alignment is the process of identifying semantically equivalent entities from multiple ontologies. This is not always obvious because technical constraints such as data volume and execution time are determining factors in the choice of an alignment algorithm. Nowadays, partitioning and modularization are two main strategies for breaking down large ontologies into blocks or ontology modules respectively to align ontologies. This article proposes ONTEM as an effective alignment method for large-scale ontology based on the ontology entities extraction. This article conducts a comprehensive evaluation using the datasets of the OAEI 2018 campaign. The obtained results are promising, and they revealed that ONTEM is one of the most effective systems.

KEYWORDS

Alignment, Extraction, Large Ontologies, Matching, Modularization, Partitionnement, Semantic Interoperability

INTRODUCTION

Over the last several years, diverse ontologies have been created to enable interoperability across multiple and diverse information systems, especially with the advent of the Semantic Web. Ontologies are at the kernel of Semantic Web by allowing the explicitness of the semantic purpose for structuring different fields of interest. They play a key role in annotating web pages or services by modeling the concepts, attributes and relationships used to annotate resource content.

Nowadays the use of ontologies in many contexts such as science, e-commerce, e-health, and industry has been increased (Stuckenschmidt, 2009). With the rapid development of ontology applications, domain ontologies became very large in scale, which generates several difficulties in ontology constructing, matching, reusing, maintenance, reasoning, etc.

In many application contexts, several ontologies covering the same or related fields are developed independently of each other by different communities, which raises the issue of being able to exchange, integrate and transform data. At this stage, the problem of interoperability arises, allowing heterogeneous systems to communicate and cooperate, and to this end, semantic links must be established between entities belonging to two different ontologies, and the transition to the web is a real challenge that requires researchers to make efforts to optimize content management, which can be constantly enriched and developed. To this end, it is necessary to improve the quality of the

DOI: 10.4018/IJSITA.2019040104

organization, structuring, research, identification, access, use, integration, and automated processing of this content. In the past several years, a panoply of systems has been developed to automatically generate, match, and integrate ontologies in a process called ontology alignment. In each of these cases, it is important to know the relationships between the terms in the different ontologies. It has been realized that this is a major issue and much research has recently been done on ontology alignment, i.e. finding mappings between terms in different ontologies.

Due to their sizes and monolithic nature, large ontologies lead to a new challenge to the state-of-the-art ontology alignment technology. Ontology alignment is an important way of establishing interoperability between applications that use different but related ontologies (Hu et al.,2008).

In the past several years, an enormous effort has been invested in aligning and harmonizing ontologies including rule-based and statistical matchers. Existing matchers rely on entity features such as names, synonyms, as well as relationships to other entities (Shvaiko and Euzenat, 2013; Otero-Cerdeira et al., 2015). However, manually aligning ontologies is a labor-intensive process and often semi-automatic approaches are sought.

All alignment techniques are required to scale up to handle large ontologies (Jimenez-Ruiz et al., 2012; Diallo, 2014). However, this is not always obvious because the creation of multiple ontologies, sometimes for the same domain, leads to heterogeneity between the knowledge expressed within each of them that must be resolved: it is the problem of interoperability.

When ontologies cover overlapping topics, the overlap can be represented using ontology alignments. These alignments need to be continuously adapted to changing ontologies. Especially for large ontologies this is a costly task often consisting of manual work (Jurisch and Igler, 2018).

The research field of ontology alignment is very active with its yearly workshop as well as a yearly event, the Ontology Alignment Evaluation Initiative (OAEI, e.g. (Euzenat, 2011)), that focuses on evaluating systems that automatically generate mapping suggestions. OAEI systems are typically able to cope with small and medium size ontologies, but fail to complete large tasks in a given time frame and/or with the available resources (e.g., memory). Large ontologies still pose serious challenges to state-of-the-art ontology alignment systems (Jimenez-Ruiz et al., 2018).

Our work aims to meet the challenge of scaling up the alignment method. In particular, we propose a novel algorithm to extract common concepts and labels to both ontologies for alignment purposes. Our algorithm has been tested on the OAEI campaign. Satisfactory results have been achieved compared with the reference alignments contained in the LargeBio Track 2018 section.

This paper is organized as follows: Section 2 presents the different alignment strategies. Section 3 scrutinizes the main related work. In section 4, we describe our extraction-based method for large ontologies alignment. Section 5 is an experimental study that illustrates the results and performance of our method. Section 6 presents a discussion of the obtained results. Finally, Section7 concludes the paper and points out future work.

ONTOLOGY ALIGNMENT STRATEGIES

Alignment consists of determining the set of correspondences between two ontologies by using or implementing solutions to different heterogeneity problems. Nowadays there are a lot of techniques and tools for addressing the ontology alignment problem. However, the complex nature of this problem means that the existing solutions are unsatisfactory.

Several alignment techniques, based on different criteria, are currently proposed in the literature. (Ardjani et al., 2015) provides a synthesis of alignment techniques. The choice of one technique or a composition of several of them is not an easy task. Many alignment methods dedicated to ontologies have emerged in the last decade. However, these methods are designed to align small ontologies. Partitioning (Pereira et al. 2017) and modularization (Santos et al., 2015) are the two main strategies for breaking down large ontologies into blocks or ontology modules, respectively. These methods can only work if the number of concepts at the input of the alignment tool is limited.