Computers in Human Behavior 30 (2014) 745–752

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Establishing agent trust for contradictory evidence by means of fuzzy voting model: An ontology mapping case study

Maria Vargas-Vera [a,*], Miklos Nagy [b]

[a] Facultad de Ingenieria y Ciencias, Centro de Investigaciones en Informatica y Telecomunicaciones, Universidad Adolfo Ibanez, Vinia del Mar, Chile
[b] The Open University (OU), Milton Keynes, UK

ARTICLE INFO

Article history:
Available online 5 November 2013

Keywords:
Ontology mapping
Semantic Web
Multi-agent systems
Uncertain reasoning

ABSTRACT

This paper introduces a novel trust assessment formalism for contradicting evidence in the context of multi-agent ontology mapping. Evidence combination using the Dempster rule tend to ignore contradictory evidence and the contemporary approaches for managing these conflicts introduce additional computation complexity i.e. increased response time of the system. On the Semantic Web, ontology mapping systems that need to interact with end users in real time cannot afford prolonged computation. In this work, we have made a step towards the formalisation of eliminating contradicting evidence, to utilise the original Dempster's combination rule without introducing additional complexity. Our proposed solution incorporates the fuzzy voting model to the Dempster–Shafer theory. Finally, we present a case study where we show how our approach improves the ontology mapping problem.

1. Introduction

High quality semantic meta data is a crucial part of the envisioned Semantic Web. However, the possible number of applications (Memon & Khoja, 2009) that can be developed on the Semantic Web heavily relies on how data can be integrated from distributed and heterogeneous data sources. There are several challenges, which have been identified in the context of ontology mapping by Shvaiko and Euzenat (2008) and Euzenat and Shvaiko (2007). These challenges are considered as roadblocks for developing real applications and they are explained in Section 3. Conflicting information that is inherent to interpreting Semantic Web data is one of these challenges. This conflict can be a result of insufficient or contradicting information of different terms that are similar or even the same. For example, consider two ontologies, which describe scientific publications. Both of these ontologies describe the concepts "paper". In ontology 1 the paper is represented as "Scientific Paper" in the context of "Conference participant" and in Ontology 2 as "Chapter" in the context of "Book". When mapping algorithms extend these contexts using any kind of background knowledge e.g. WordNet sister terms, one can derive that both describes a printed work of someone. The problem is though that this knowledge cannot directly be deduced from the ontologies, because "Scientific Paper" refers to participant, while "Chapter" refers to portion of the book, that has been published. Naturally human experts could easily resolve this contradiction through discussing their point of views and decide if the mapping can be made or not. However, this is not the case for ontology mapping applications that operate without human intervention. In the case of applications, sufficient conflict resolution processes need to be in place, to improve the quality of the mappings by eliminating the contradictions. In this paper, we propose a conflict elimination approach using a fuzzy voting model. Based on our initial approach for eliminating conflicts (Nagy, Vargas-Vera, & Motta, 2008), we propose different fuzzy variables, membership functions and a customised voting algorithm in order to provide more reliable results. The fuzzy voting model allows to detect and eliminate contradictory evidence, instead of discarding the whole scenario or combining them with contradictions. These contradictions can occur on any entities in the ontology e.g. classes, objects, data properties and instances. The main contribution of this paper is that it proposes a conflict elimination method, based on trust and fuzzy voting, before any conflicting beliefs are combined.

The paper is organised as follows. In Section 2 we present the related work. Section 3 describes why conflicts occur in the ontology mapping context and our proposed solution is explained in Section 4. In order to validate our approach we have carried out experiments, which is presented in Section 5. Finally in Section 6 we describe the future research directions and the conclusions of our work.

2. Related work

Managing contradictions in the context of ontology mapping in general has lead to different approaches. We review the most relevant approaches that were also identified as state-of-the-art (Shvaiko & Euzenat, 2013), before we introduce our approach.

* Corresponding author.
E-mail addresses: maria.vargas-vera@uai.cl (M. Vargas-Vera), m.nagy@open.ac.uk (M. Nagy).

0747-5632/$ - see front matter
http://dx.doi.org/10.1016/j.chb.2013.10.032

Since our approach eliminates the contradictions before the judgement of the mapping is established, the relevant work on this area can be found in other ontology mapping approaches. Therefore, in our scenario the test bed that we have used is the OAEI-2008 datasets. The biggest challenge was how to compare our approach with other solutions. It is clear that the only qualitative comparison on the mapping system level can only be made through the Ontology Alignment Initiative, which is an international effort to compare ontology mappings systems. Different approaches to eliminate contradictions for ontology mapping have been proposed by the ontology mapping community. These approaches can be classified into two distinct categories.

First group include solutions that consider uncertainty and fuzziness as an inherent nature of the ontology mapping and tries to describe it accordingly. Ferrara et al. (2008) model the whole ontology mapping problem as fuzzy where conflicts can occur therefore, their approach models the whole mapping process as an uncertain reasoning task, where the mapping results need to be validated at the end of the reasoning process. The reasoning is supported by fuzzy Description Logic approaches. As a consequence, their mapping validation algorithm interprets the output mapping pairs as fuzzy and tries to eliminate the inconsistencies from them.

Tang et al. (2006) formalise the ontology mapping problem as making decisions on mappings using Bayesian reasoning. Their system *RiMOM* Tang et al. (2006) has participated in the OAEI competition as well. Their solution do consider two kinds of conflicts in the ontologies, namely the structure and naming conflicts. However, they use thesaurus and statistical techniques to eliminate them before combining the results. *RiMOM* approach produces ontology mapping using well-defined processing steps like *ontology pre-processing*, *strategy selection*, *strategy execution* and *alignment combination*. *RiMOM* has been very successful during OAEI competitions; however, its strategies have to be defined in advance together with their rules, which are selected during execution time. As a result, it is questionable how the system can be adapted to the Semantic Web environment, where domains can change dynamically. Furthermore, the assumption that ontologies with *similar features* are similar in reality might not be valid in all cases. Another weak point is that large ontologies cannot easily be loaded into the internal model and the approach does not consider optimisation for the mapping process. Nevertheless, the main idea is remarkable since it builds up its own structure and, hence, tries to interpret the ontology before processing it.

The second group, however, differ conceptually because they mainly utilise data mining and logic reasoning techniques in pre and post processing stages of the mapping.

For example, Liu, Wang, and Wang (2006), split the ontology mapping process into four different phases. Their approach first exploits the available labels in the ontologies then it compares the instances. After it recalls mappings from the previous mapping tasks and compares it with the structure of the ontologies. Their approach also tries to eliminate contradictions, using the previous experience and data mining techniques on the relations that are defined on the ontologies.

Similar solution has been proposed by the Jean-Mary, Shironoshita, and Kabuka (2009) and Jean-Mary and Kabuka (2008). *Automated Semantic Mapping of Ontologies with Validation (ASMOV)* automates the ontology alignment process using a weighted average of measurements of similarity along four different features of ontologies, and performs semantic validation of resulting alignments. This system acknowledges that conflicting mappings are produced during the mapping process but they use an iterative post processing logic validation in order to filter out the conflicting mappings.

Anchor-Flood (Seddiqui & Aono, 2009), is an ontology mapping tool conceived in the context of the *International Patent Classification (IPC)*. The mapping approach itself was designed to work as part of a *patent mining* system that assigns patent abstracts to existing IPC ontologies and it also uses a multi-phase approach to create the mapping results. These phases are *pre-processing*, *anchoring*, *neighbouring block collection* and *similarity measures*. Anchor-Flood also uses an internal representation form to which the ontologies are transformed before processing. The system is also reliant on the availability of individuals, which might not be always present in real life scenarios. There are also a number of weaknesses that are related to the fact that the approach is highly dependent on the correctness of the initial anchoring. Inconsistencies might not be eliminated and missed links might not be discovered it they do not fall into the context of already linked entities.

TaxoMap (Hamdi, Niraula, & Reynaud, 2009) is an approach that is based on the assumption that large scale ontologies contain very extensive textual descriptions and well defined class structures but do not contain a large number of properties or individuals. The similarity assessment uses various *Natural Language Processing* techniques and frameworks like *TreeTagger* Schmid (1994) and *structural heuristic-based* similarity algorithms like *Semantic Cotopy* (Ehrig, Koschmider, & Oberweis, 2007). In order to filter out inconsistent mappings, it uses a so-called *refinement module*. End users have the possibility to define constraints and solutions using a logic-based language called *Mapping Refinement Pattern Language (MRPL)*. For example, this language allows the end users to express domain specific constraints that can remove a mapping pair on condition that the classes involved in the mapping do not have an equivalence relation in the source or target ontology. One weakness of the system is that it requires the fine-tuning of nine different threshold values, which is a challenge given the possible combinations and the possible impacts on the result set.

Lily (Wang & Xu, 2009) mapping approach carries out the mapping in different phases. These phases are *pre-processing*, *match computing* and *post processing*. In the last phase, the system extracts the final mapping set based on the similarity assessments, and then verifies that inconsistent mappings are indicated to the user, who can remove them manually. It is important to point out that the mapping approach recognises the fact that the interpretation of the ontologies involves dealing with uncertainty. However, the objective is only to reduce the amount of uncertainty instead of dealing or reasoning with it. As a result, the mapping process only reduces the negative effect of the matching uncertainty. Lily can also deal with large-scale ontology matching tasks thanks to its scalable ontology matching strategy.

3. Conflicting and trust related to Semantic Web data

3.1. Sources of conflicts

As we briefly mentioned earlier (in the paper), in the context of ontology mapping different challenges had been recognised by Shvaiko and Euzenat (2008). These challenges are viewed as roadblocks for implementing ontology mapping applications that can be applied with high confidence in different contexts i.e. real world domains. We have chosen two, which we believe are mostly related to problem of contradictions. In these two cases, the ontology mapping systems have to establish a certain degree of understanding of the meaning of the data that is present in the different ontologies.

Firstly, the uncertainty related to different representations stems from the fact that W3C has proposed different languages that can be used on the Semantic Web e.g. RDF(S),[1] OWL[2] and SKOS.[3] The problem is that ontology engineers can choose any

[1] http://www.w3.org/RDF/.

[2] http://www.w3.org/TR/owl-features/.

[3] http://www.w3.org/TR/skos-reference/.