J Inf Process Syst, Vol.13, No.4, pp.834~851, August 2017
https://doi.org/10.3745/JIPS.02.0034

ISSN 1976-913X (Print)
ISSN 2092-805X (Electronic)

JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

Boosting the Reasoning-Based Approach by Applying Structural Metrics for Ontology Alignment

Abderrahmane Khiat* and Moussa Benaissa*

Abstract

The amount of sources of information available on the web using ontologies as support continues to increase and is often heterogeneous and distributed. Ontology alignment is the solution to ensure semantic interoperability. In this paper, we describe a new ontology alignment approach, which consists of combining structure-based and reasoning-based approaches in order to discover new semantic correspondences between entities of different ontologies. We used the biblio test of the benchmark series and anatomy series of the Ontology Alignment Evaluation Initiative (OAEI) 2012 evaluation campaign to evaluate the performance of our approach. We compared our approach successively with LogMap and YAM++ systems. We also analyzed the contribution of our method compared to structural and semantic methods. The results obtained show that our performance provides good performance. Indeed, these results are better than those of the LogMap system in terms of precision, recall, and F-measure. Our approach has also been proven to be more relevant than YAM++ for certain types of ontologies and significantly improves the structure-based and reasoning-based methods.

Keywords

Description Logics Inference, Intra-Taxonomy Measures, Ontology Alignment, Semantic Interoperability, Semantic Web, Structural Similarity

1. Introduction

The semantic web community relies on ontologies to overcome the crucial problem of semantic heterogeneity. However, these ontologies are heterogeneous. This heterogeneity may occur at the syntactic, terminological, conceptual, and semiotic levels [1].

Ontology alignment, which is defined as the process of identifying semantic correspondences between entities of different ontologies to be aligned [2], represents the solution to the problem of semantic interoperability between different sources of distributed information.

Due to the size and number of ontologies, ontology alignment cannot be done manually beyond a certain complexity. For that reason, automatic techniques, or at least semi-automatic ones, should be developed to reduce the burden of having to manually create and maintain the ontologies alignment [3]. However, automatically identifying the correspondences between ontologies is very difficult due to their conceptual divergence [4].

Manuscript received January 28, 2015; accepted April 20, 2015; onlinefirst December 17, 2015.
Corresponding Author: Abderrahmane Khiat (abderrahmane_khiat@yahoo.com)
* LITIO Laboratory, University of Oran 1 Ahmed BenBella, BP 1524 El M'Naouer Oran, Algeria (abderrahmane_khiat@yahoo.com, moussabenaissa@yahoo.fr)

Numerous ontologies have been developed in recent years in different domains and several methods have been elaborated upon to semantically align their entities.

These methods are generally based on the similarity of computing their names, relationships, and instances. These methods can be terminological, linguistic, structural, semantic, or extensional; and most of the systems often combine these approaches [2,3]. Furthermore, ontology alignment approaches can be split globally into two main categories [2]: reasoning-based approaches and approaches based on the similarity computing.

As such, the objective of this paper is to present a new approach for ontology alignment in order to find new semantic correspondences between the entities of the two ontologies that are to be aligned.

It consists of combining reasoning-based and structure-based approaches. The goal is to apply the intra-taxonomic structural metrics that have been well-established in the literature [2] in the case of aligning two ontologies that do not necessarily share the same taxonomy. Therefore, the idea is to create an inferred shared taxonomy by the two ontologies to align, which serves as support for the application of intra-taxonomic metrics. This creation occurs in two steps. First, an initial reference alignment is exploited to merge the two ontologies to align. Thus, a first shared taxonomy is created. Next, the inference services of description logics [5] are operated to identify the inferred taxonomy. The idea behind this inference is to discover all implicit subsumption relations and thereby enrich the shared taxonomy. As a direct consequence, this can improve the similarities calculated by the intra-taxonomic metrics and thus new semantic correspondences can be discovered.

The rest of the paper is organized as follows: first, background is cited in Section 2. The motivation and illustrative example are presented in Section 3 and the related work is presented in Section 4. Section 5 describes our contribution by giving a detailed account of our approach and Section 6 presents the aspects related to the implementation of our system. We provide the evaluation results in Section 7, in order to show the efficiency of our approach. Section 8 contains our conclusion and establishes the direction for future work.

2. Background

In this section we present the preliminary notions of ontology alignment. We outline the concepts of ontology, similarity, and alignment based on previous work carried out in [2,3].

2.1 Formal Definition of Ontology

Ontology is a six tuple [3]: $O = \langle C, R, I, H^C, H^R, A \rangle$ where:

- C: set of concepts.
- R: set of relations.
- I: set of instances.
- H^C: hierarchy of concepts.
- H^R: hierarchy of relations.
- A: set of axioms.

2.2 Definition of Similarity

DEFINITION 1. A similarity [2] σ: o × o → R is a function from a pair of entities to a real number expressing the similarity between two objects such that:

- ∀ x, y ∈ o, σ (x, y) ≥ 0 (positiveness).
- ∀ x ∈ o, ∀ y, z ∈ o, σ (x, x) ≥ σ (y, z) (maximality).
- ∀ x, y ∈ o, σ (x, y) = σ (y, x) (symmetry).

2.3 Similarity Computing Methods

There are essentially five types of methods for calculating similarities [2]:

1. *Terminological Methods*: these methods are based on string matching and can be applied to the names, labels, and descriptions of the entities to calculate their similarities. As an example of a matcher for this category, we cited the edit distance.
2. *Linguistic Methods*: these methods are based on external resources, such as the dictionaries and thesaurus, in order to calculate the similarities between the names, labels, and descriptions of the entities. As an example of a matcher for this category, we cited the similarity based on WordNet (Wu-Palmer).
3. *Structure-based Methods*: these methods exploit the internal structure (domain, range, properties, cardinality, etc.) and the external structure (hierarchy and the relationship between other entities) of the entities in order to calculate their similarities. As an example of a matcher for this category, we cited the Resnik similarity.
4. *Semantic-based Methods*: these methods are essentially deductive and inferential methods based on the formal semantics of the generic or specific domain. As an example of a matcher for this category, we cited the SAT solvers.
5. *Instance-based Methods*: these methods exploit the instances associated with the concepts (extensions) to calculate the similarities between them. As an example of a matcher for this category, we cited the Jaccard similarity.

2.4 Ontology Alignment

Alignment is a process Fig. 1 that starts from the two representations of O and O' and produces a set of correspondences between pairs of (simple or complex) entities <e, e'> belonging to O and O', respectively [3].

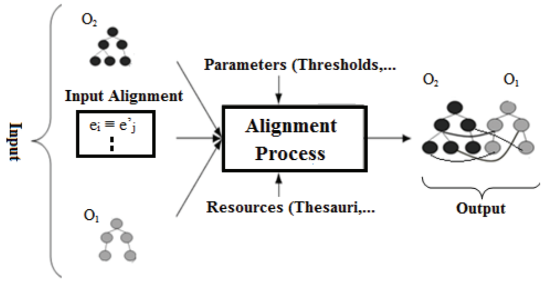

Fig. 1. Ontology alignment process.

There are some other parameters that can be used in the alignment process [2]:

- A: an input alignment, which is to be completed by the process.
- p: the alignment parameters (e.g., weights, thresholds).
- r: the external resource used by the alignment process (e.g., common knowledge and domain-specific thesauri.

2.4.1 Definition of correspondence

Correspondence is a formal expression that states the semantic relationship between two entities belonging to different ontologies [2]. Given two ontologies of O_1 and O_2, aligning one ontology with another means that for each entity (concept C, relation R, or instance I) in ontology O_1, we try to find a corresponding entity that has the same intended meaning in ontology O_2.

In general, a correspondence can be described as a quadruple: <e, e', n, R>:

- e and e' are the entities between which a relation is asserted by the correspondence.
- n is a degree of trust (confidence).
- R is the relation associated to a correspondence, where R identifies the relationship that exists between e and e', which can be a simple set-theoretic relation, a fuzzy relation, or a similarity measure.

2.4.2 Ontology alignment process

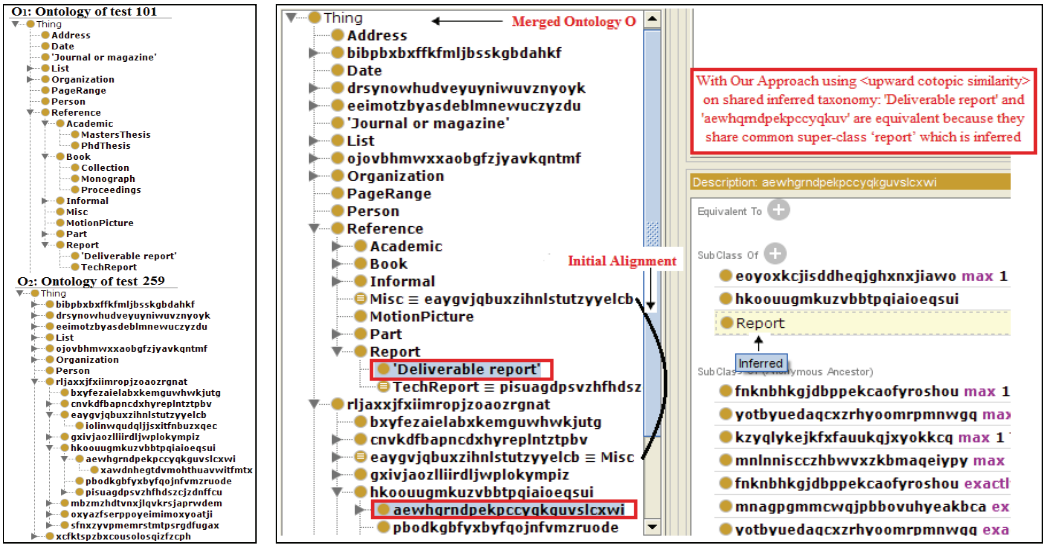

Fig. 2. Illustrative example: our approach combining structure-based and reasoning approaches.

The ontology alignment is a set of correspondences between two or more ontologies (in the case of multi-alignment). The alignment process generally consists of the following steps:

1. Analysis: this step consists of extracting both entities (concepts, relations, instances) of the two ontologies O and O' and their characteristics, which will be used to identify the alignment.

2. Similarity Computing: this step consists of executing the different matchers in order to calculate the similarities between the entities that are to be aligned.
3. Similarity Values Aggregation: this step consists of combining the similarity values calculated by the matchers in the previous step into one value.
4. Selection: this step consists of applying a strategy (i.e., a threshold strategy) in order to filter the alignment defined in the previous step. Other optimization techniques can also be applied at this level to optimize the extraction of the final alignment.
5. Improvement of the Alignment: descriptive logic techniques can be applied at this level to improve the final alignment by diagnosing and repairing any inconsistencies identified in the final alignment.

3. Motivation and Illustrative Example

We present in this section an example that illustrates the principle of our approach Fig. 2, in which the two ontologies to be aligned O_1: 101 and O_2: 259 of the biblio test of benchmark (http://oaei.ontologymatching.org/2012/Benchmarks/index.html#datasets) series of the Ontology Alignment Evaluation Initiative (OAEI) 2012 evaluation campaign.

We began by merging O_1 and O_2 with an initial alignment selected randomly from the reference alignment established by an expert between O_1 and O_2. Where "eaygvjqbuxzihnlstutzyyelcb," the concept of O_2, was subsumed by the concept "Misc" of O_1 and vice versa.

In this section we will show the contribution of the combination of reasoning-based approaches with structure-based approaches.

– First, if we apply only the reasoning-based approaches Fig. 3 using Pellet [6], a DL reasoner, on the merged ontology, we can only deduce that the concept "aewhgrndpekpccyqkguvslcxwi" of O_2 has the super-concept "Report" of O_1, but we will not discover the semantic correspondence (equivalence relation) between the concepts "Deliverable report" of O_1 and "aewhgrndpekpccyqkguvslcxwi" of O_2.

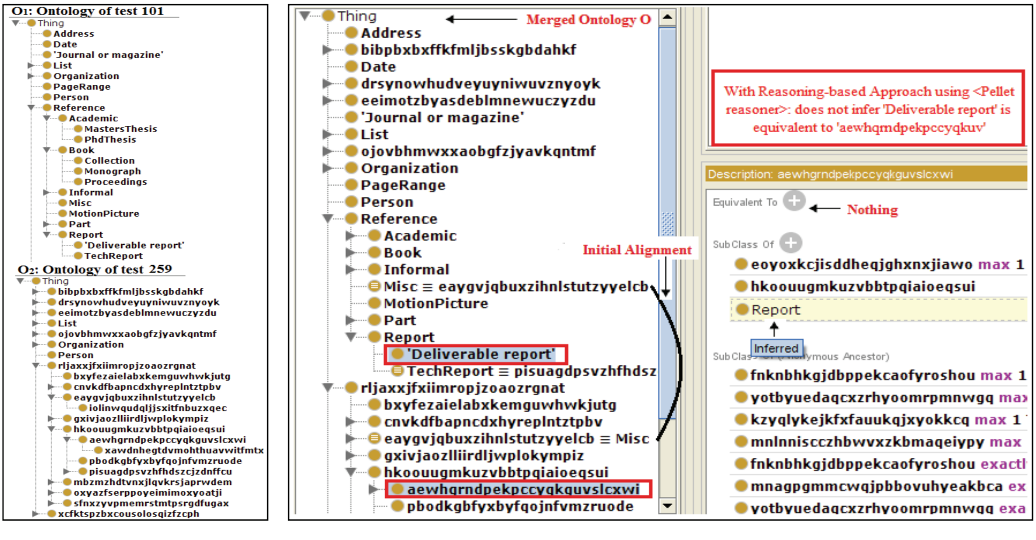

Fig. 3. Illustrative example: insufficiency of reasoning-based approach.

– Second, if we apply only the intra-taxonomic structure-based approaches Fig. 4 using the Upward Cotopic Similarity measure of UPC (Section 4.2.1) on the merged ontology, we will not discover the semantic correspondence (equivalence relation) between the concept “Deliverable report” of O_1 and the concept “aewhgrndpekpccyqkguvslcxwi” of O_2 because the two concepts do not share super-concepts.

– Finally, if we apply our approach Fig. 2, which consists of the invocation of the DL reasoner, Pellet [6], followed by the similarity computation between the two concepts of the two ontologies that are to be aligned by applying the intra-taxonomic metric Upward Cotopic Similarity measure of UPC (Section 4.2.1), we discover that the concept “aewhgrndpekpccyqkguvslcxwi” of O_2 and the concept “Deliverable report” of O_1 have the same super-concept inferred “report.” This means that the discovery of a new semantic correspondence (i.e., the concept “aewhgrndpekpccyqkguvslcxwi” of O_2) is equivalent to the concept “Deliverable report” of O_1.

Matchers that use different terminological and linguistic information cannot discover this semantic correspondence. This is why the majority of systems that have participated in the campaign of OAEI are failing in this test, because the ontologies in this test lack the terminological and linguistic information (http://oaei.ontologymatching.org/2012/results/Benchmarks/Biblio-Benchmarks-r1.html). The process continues until all inferred concepts are covered.

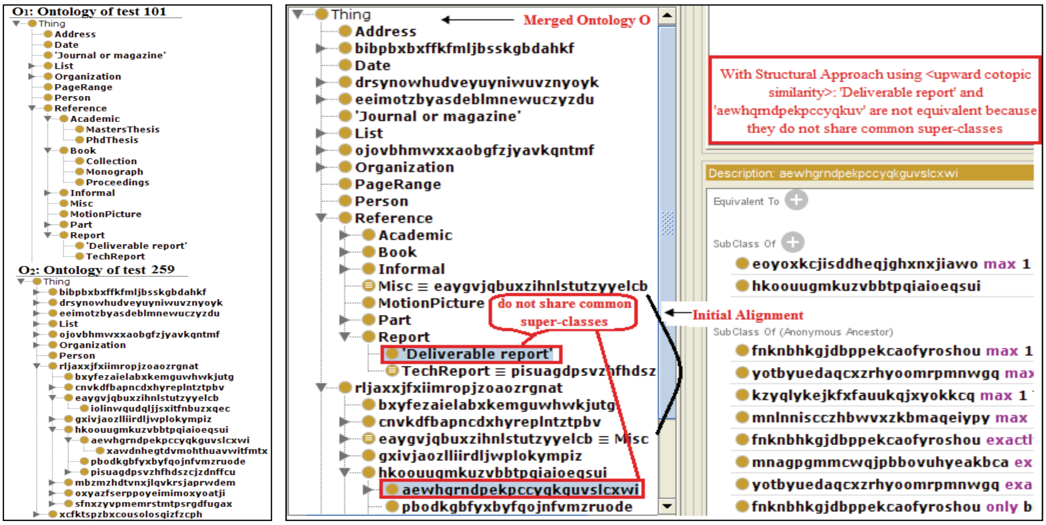

Fig. 4. Illustrative example: insufficiency of structure-based approach.

4. Related Work

In this section we will briefly introduce some relevant works on reasoning-based and structure-based approaches to better position our contribution.

4.1 Reasoning-Based Ontology Alignment

Several reasoning-based ontology-alignment approaches have been developed. We categorized the three approaches as: methods using 1) specific algorithms [7,8], 2) SAT solvers [9,10] and 3) and Description Logics reasoning [11-13]. The principle of the approaches that fall under the third category consists first of identifying an initial reference alignment and then applying the Description Logics reasoning.

There are two aspects that have received a particular attention in semantic-based ontology alignment using Description Logics reasoning, namely the discovery of new semantic correspondences [11] and the improvement of an existing alignment [12,13].

Other systems, such as LogMap [14], incorporate both repair capabilities and discovery of semantic correspondences. In this section, we present some approaches that are relevant to the category of reasoning-based methods.

4.1.1 Discovering new semantic correspondences

The proposed work in [11] consists of first, expressing the initial alignment using the OWL-Full language and then, translating it into OWL-DL + SWRL / OWL2-RL + SWRL—OWL-Full is the most expressive sub-language of OWL (the Web Ontology Language is a family of knowledge representation languages). OWL is an AI-inspired markup language and supports reasoning. Also, it has been adopted by W3C as the standard for representing ontology [15]. OWL-Full supports applications that require more expressiveness, but it can lose computational completeness. OWL-DL supports applications that require maximum expressiveness without losing computational completeness. The Semantic Web Rule Language (SWRL) is an expressive OWL-based rule language. SWRL allows users to write rules that can be expressed in terms of OWL concepts to provide more powerful deductive reasoning capabilities than OWL alone. OWL2-RL profile supports applications that require scalable reasoning without sacrificing too much expressive power.

The objective is to have more expressiveness while preserving decidability.

4.1.2 Improvement of an existing alignment

The proposed work in [12,13] consists of reasoning on an existing alignment to detect potential inconsistencies. More specifically, the authors identify properties (consistency, embedding, containment and minimality) that reflect the quality of alignment and provide algorithms for their verification.

This work is very useful to improve the generated alignment, but it does not discover new semantic correspondences. This approach can be used as the final step for all approaches that discover new semantic correspondences.

4.1.3 Repair and discovery of new semantic correspondences

LogMap [14] is a reasoning-based ontology alignment system. To align two ontologies, it begins by creating an initial alignment using lexical and structural indexing that operates on the ontologies that are to be aligned. Then, it generates the final alignment by alternating iteratively the repair and discovery of new semantic correspondences. It incorporates both the repair capabilities and discovery of semantic correspondences.

Contrary to LogMap, after having identified an initial reference alignment we applied a structure-based intra-taxonomic metric on the shared inferred taxonomy to calculate the similarities. Thus, we were able to identify new semantic correspondences, rather than continue with logical reasoning, as LogMap does.

4.2 Structure-Based Ontology Alignment

Structure-based ontology alignment approaches are a promising family of solutions to find semantic relations among concepts of ontologies. We distinguished two main categories of structure-based ontology alignment at this level: approaches developed specifically for ontology alignment and approaches developed elsewhere that require adaptation in order to be applicable to the case of ontology alignment. We will now present some approaches relevant to these categories.

4.2.1 Structural approaches developed for ontology alignment

Several structural approaches have been developed in order to align two ontologies. As an example of these approaches we have quoted the works presented in [16-18]. It is often required to combine theses approaches with other methods.

Listed below is a description of one of the approaches that is relevant to the category of structural approaches developed for ontology alignment.

Anchor Prompt: This approach uses linguistic matchers in order to identify similar concepts. Then, the algorithm compares the paths connecting these concepts into sub-graphs and the similarity increases between the concepts of the two paths that appear frequently in the same positions on the paths of the same length [16].

4.2.2 Structural approaches used in a single ontology

Several structural measures have been developed using the same taxonomy. These include the structural topological dissimilarity on hierarchies (STDH) method [19], the similarity distance method [20], and the upward cotopic similarity (UCS) method [21], which we will detail below.

Upward Cotopic Similarity:

Let δ: O × O → R is a similarity over a hierarchy H = (h ≤ O), such that:

$$\sigma(c, c') = \frac{UC(c,H) \cap UC(c',H)}{UC(c,H) \cup UC(c',H)} \tag{1}$$

where UC (c, H) = {c' ∈ H; c ≤ c'} is the set of super-classes of c.

These structural measures cannot be used directly for ontology alignment [22] and require adaptation. One way to adapt this metric to the alignment of two ontologies consists of using WordNet [2,23]. However, we used a logical method instead of an adaptation method. Our method consists of starting with an initial alignment that serves as a primer for an inference based on description logics. Our method has the advantage, compared to one based on WordNet, to enable the support of consistency problems and the appearance of the cycles that can occur.

5. Our Ontology Alignment Approach

The approach that we propose in this paper consists of combining reasoning-based and structure-based approaches. More specifically, it exploits the reasoning-based approach in order to create a support that uses the intra-taxonomic structural metrics that have been well established in the literature.

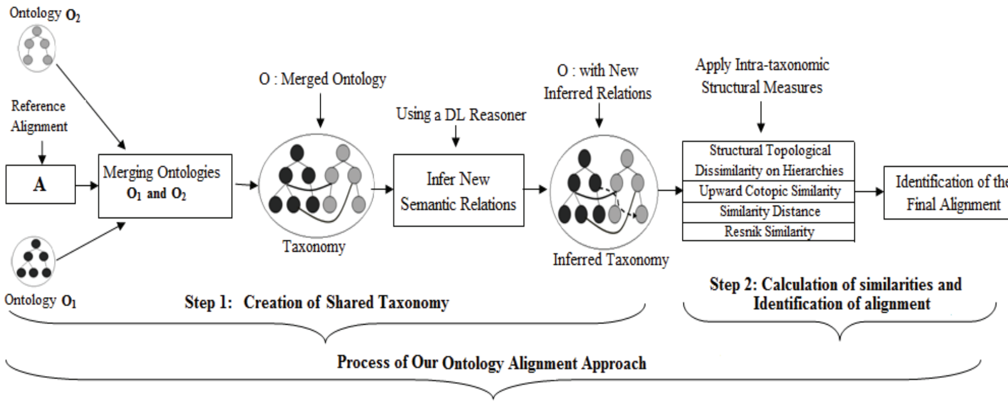

Fig. 5. Process of our ontology alignment approach.

We summarize above the process of our approach shown in Fig. 5 to provide a general idea of the proposed solution. It consists of two successive steps described below.

5.1 Step 1: Creation of Shared Inferred Taxonomy

This step is dedicated to the creation of shared inferred taxonomy between the two ontologies to be aligned. It consists of two phases that are described below.

5.1.1 Phase 1: Merging ontologies to be aligned

This phase consists of merging the ontologies to align O_1 and O_2 using an initial reference alignment A. This alignment can be obtained either by applying existing automatic discovery algorithms, such as terminological, structural methods, etc., or by a domain expert. O is the resulting ontology, as illustrated in Fig. 2. In our case, the reference alignment was obtained from a domain expert.

5.1.2 Phase 2: Inference of new semantic relations

This phase consists of operating the following reasoning services on the ontology O to infer new semantic relations that are subsumption relations:

- Identify all implicit subsumption relations between the concepts of O by operating description logic reasoning (i.e., the creation of inferred shared taxonomy, as shown in Fig. 2).
- Use the properties of relations, such as transitivity, to infer new relations.

5.2 Step 2: Similarity Computing and Alignment Identification

This step is dedicated to identifying the alignment between the entities of the two ontologies. It consists of two successive phases, which are explained below.

5.2.1 Phase 1: Similarity computing between concepts

This phase consists of applying intra-taxonomic structural metrics to calculate the similarities between the concepts of the ontologies that are to be aligned. The structural intra-taxonomic metric that we used in our approach is the UCS (1). We chose this metric because it provides better results than other metrics in the same category as STDH according to the results presented in [22]. We confirmed this result after a comparative study of the metrics UCS and STDH.

5.2.2 Phase 2: Identification of alignment

This phase consists of identifying the semantic correspondences by exploiting similarities calculated in the previous step. Our method was combined with our system, which we created and participated in the OAEI 2014 evaluation campaign. We used the average aggregation strategy (the similarity is obtained by calculating the average of the similarity value calculated by our method and our system). The combination was applied only in the case of making comparisons with other systems: LogMap and YAM++, as illustrated in Figs. 6–9. We selected the threshold S to operate the filtering of the matrix, which results from the combination. This filter allows us to select the semantic correspondences.

The choice of our system, which consists of combining different terminological matchers (e.g., Levenshtein distance [25]) with the average aggregation strategy after local filtering, is motivated by the fact that our system shows good results in terms of the F-measure on different tracks of OAEI 2014, especially the benchmark track of OAEI 2014.

6. Conception and Implementation

In order to test our approach we implemented a Java platform. The merging of the two ontologies that are to be aligned was produced with the PROMPT plug-in [24] of the Protégé platform. The Pellet reasoner was operated to calculate the inferred taxonomy. To parse the two ontologies that are to be aligned, we used the Jena plug-in [26]. The tasks of calculating the similarity and alignment extractions were performed using the Java tool that we developed.

7. Evaluation and Experimental Results

In order to evaluate the performance of our approach we used the biblio test of the benchmark series and anatomy series of the OAEI 2012 evaluation campaign.

The benchmark series consists of a reference ontology that is modified in various ways to analyze the behavior of matchers face these modifications. The biblio test consists of the reference ontology of a bibliographical references domain that has undergone several modifications (a changing of the names of concepts, modification of the hierarchy, etc.). Our choice of the biblio test is justified by the fact that it offers many useful changes in hierarchy to test our approach: suppressed, expanded and flattened. The anatomy series consists of two large ontologies, which respectively describe the anatomical structures of mice and people. Our choice of anatomy test is justified by the fact that it allows us to check if the size of the initial alignment affects the number of semantic correspondences discovered on one side and if our approach works for ontologies of large sizes. The description of the biblio test of the benchmark series and anatomy series in terms of the number of entities is given on OAEI 2012's website of http://oaei.ontologymatching.org/2012. As evaluation criteria we used the standard metrics of precision, recall, and the F-measure adopted in the OAEI 2012 campaign [27].

We considered the following four scenarios in our experimental analysis:

1. In the first scenario, we compared our results with the reasoning-based system of LogMap (http://oaei.ontologymatching.org/2012/results/benchmarks/biblio-benchmarks-r1.html), which has participated in the OAEI 2012 campaign, by using the biblio test of the benchmark series. We chose to compare our system with the LogMap system for the following reasons:
   - LogMap belongs to the same category as ours.
   - LogMap is among the most known systems in its category.
2. In the second scenario, we compared our results with the YAM++ system (http://oaei.ontologymatching.org/2012/results/benchmarks/biblio-benchmarks-r1.html), the best performing system of the OAEI 2012 campaign by using the biblio test (where the H-mean of the F-measure was equal to 0.83 for the biblio test).
3. In the third scenario, we evaluated our approach with the structure-based and reasoning-based methods alone and in combination on the biblio test. The objective of this experiment consisted of evaluating the contribution of our method as compared to structure-based and reasoning-based methods.
4. In the fourth scenario, we studied the correlation between the size of the initial alignment and the number of new identified correspondences using the anatomy series of the OAEI 2012 campaign.

For the first three scenarios above, we merged the two ontologies to be aligned with an initial reference alignment. This fusion was performed using the PROMPT plug-in of the protégé platform.

This initial alignment for the biblio test was a random sample consisting of 11 correct correspondences out of 97, which corresponds to a percentage of 10% to 30% depending on the size of the reference alignment of the test considered, that were drawn from the reference alignment.

In addition, for scenarios one and two, we combined our approach with a simple terminological matcher using our system. For the selection of correspondences shown in the results below, we chose a threshold of S = 0.70, because it provides better results in terms of precision, recall and F-measure for this threshold.

For each situation, we calculated precision, recall, and the F-measure (the number of correspondences of the initial alignment were not taken into account in all cases and scenarios).

7.1 Comparison with the LogMap System

We compared our approach with the LogMap system, which is an ontology alignment system that has participated in the OAEI 2012. The LogMap system uses a reasoning-based method with an H-mean of an F-measure that is equal to 0.56. Figs. 6 and 7 summarize the results obtained in the comparison of our system with the LogMap system in detail and overall on the biblio benchmark test of the OAEI 2012.

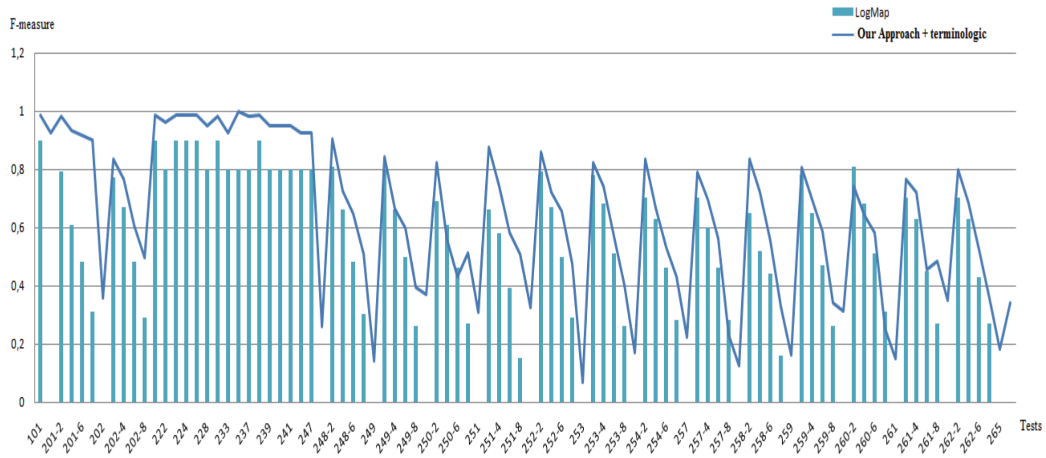

Fig. 6. Comparison of our system with the LogMap system in detail.

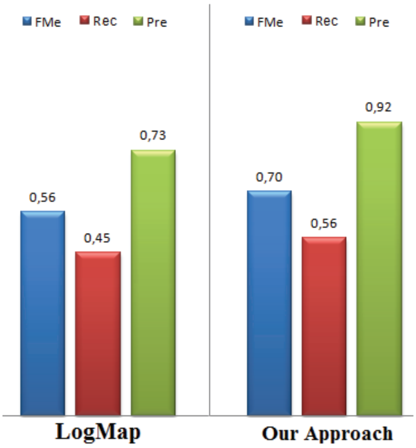

Fig. 7. Comparison of H-mean of our system with the LogMap system.

The analysis of experimental results shows that our results are better than those of the LogMap system in all tests except for the tests 250-4, 250-6, 257-8, 260-2, 260-4, and 260-8, where LogMap's results are slightly better than ours (shown in Fig. 6). The results illustrated in Fig. 7 shows that, overall, our system is better than the LogMap system in precision, recall, and the F-measure on the biblio test of the benchmark track.

7.2 Comparison with the YAM++ System

We compared our approach with the best performing system in the OAEI 2012, namely the YAM++ system with an H-mean of an F-measure that is equal to 0.83 (Figs. 8 and 9).

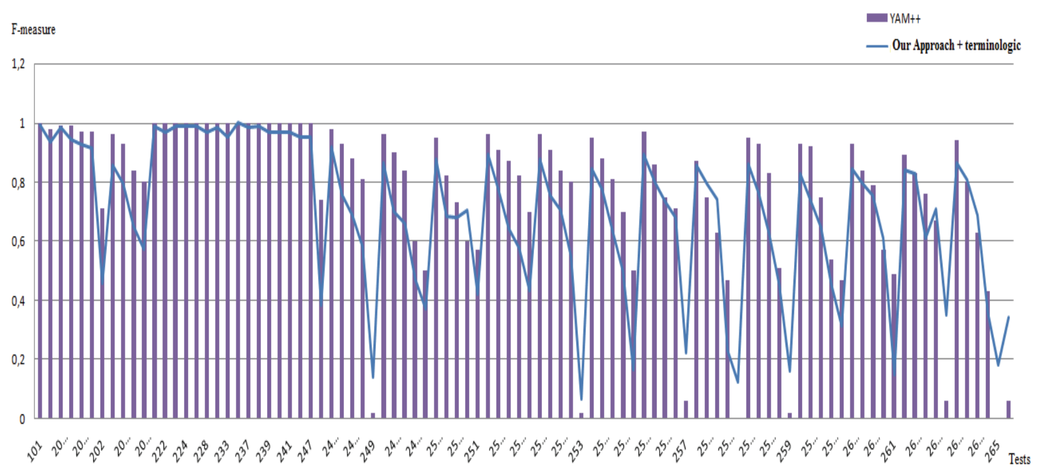

Fig. 8. Comparison of our system with the YAM++ system in detail.

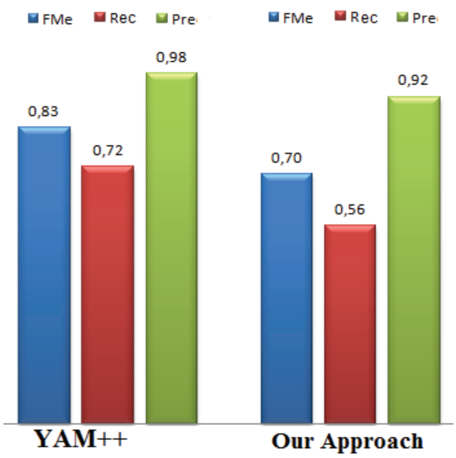

Fig. 9. Comparison of H-mean of our system with the YAM++ system.

An analysis of our experimental results shows that our system has better performance than the YAM++ system only on the tests: 249, 253, 257, 258, 259, 262, 265, and 266.

The explanation for these results is that the ontologies for these tests were constructed by deleting the terminological and linguistic information (names and comments), instances, and properties (i.e., no names: random string, no labels, no properties, and no instances) from the reference ontology. This information is necessary for the YAM++ system to properly function, unlike our approach that behaves correctly in the absence of such information, as illustrated in the concrete example shown Fig. 2. In conclusion, our approach can be recommended for this type of ontology.

7.3 Comparison of Methods

The methodology that we applied to create our experiment is as follows: we considered the biblio test and we separated test cases Fig. 10 depending on the type of hierarchy (flattened, expanded, unmodified, and suppressed). The hierarchies of these ontologies are described on the OAEI 2012's website of http://oaei.ontologymatching.org/2012/benchmarks/index.html#datasets. For each test case we calculated the F-measure (the precision and recall, respectively) both for the structure-based and reasoning-based methods; for the combining of reasoning-based and structure-based methods (our approach) on the biblio test for both cases the global case Fig. 11 (the entire set of the biblio test); and also separately, depending on the type of ontology hierarchy Fig. 10. Then, we calculated the H-mean of the F-measures (the precision and recall, respectively) of the various tests in order to get a representative value of the entire set of the biblio test. We proceeded as follows:

- We applied the structure-based method only on the ontology resulting from the merger.
- We applied the reasoning-based method only on the ontology resulting from the merger.
- We applied our approach, which consists of combining the two methods mentioned above, on the ontology resulting from the merger.

The results are shown in Figs. 10 and 11.

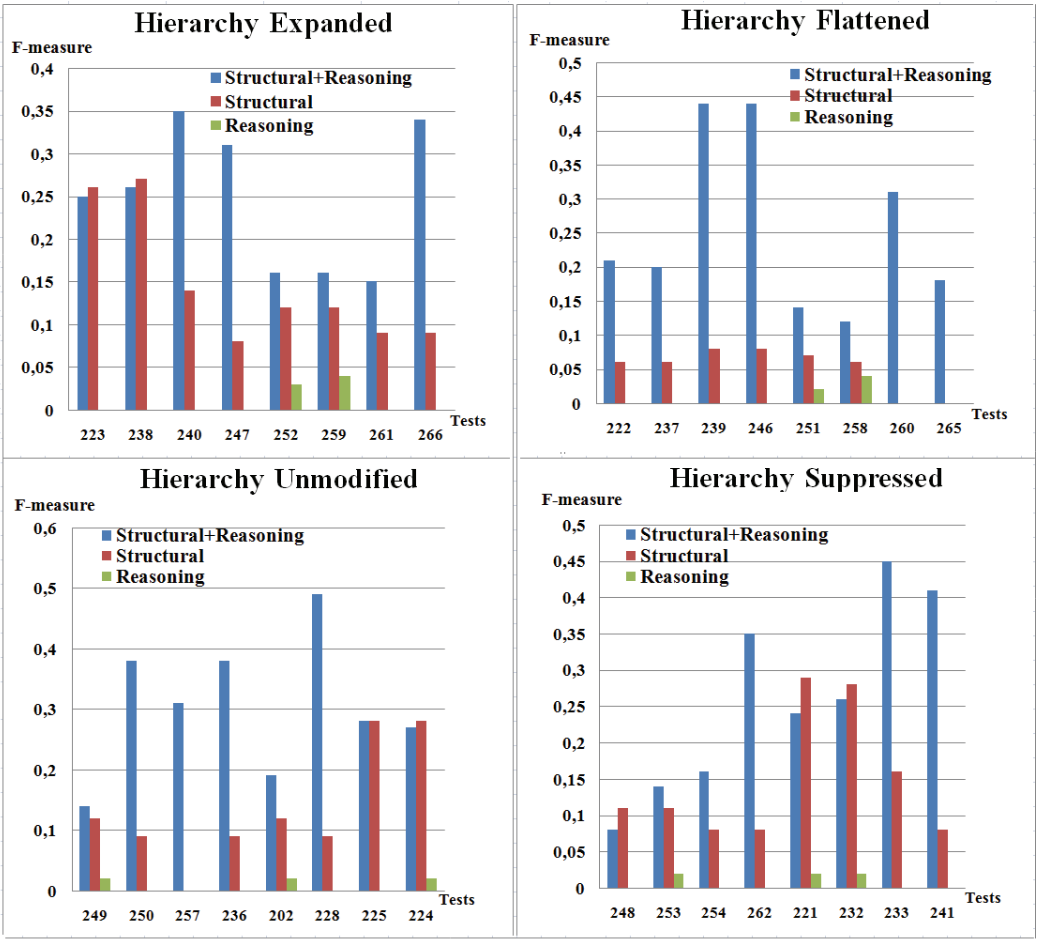

Fig. 10. Comparison of our approach with reasoning-based and structure-based methods on different types of ontologies hierarchies.

An analysis of our experimental results showed that:

- The reasoning-based method was improved (discovery of new correct correspondences) by the structural method being applied downstream of the reasoning-based approach (the growth of the F-measure can reach up to 30%), as shown in Fig. 11.
- The structural method was improved (discovery of new correct correspondences) by the reasoning-based method being applied upstream of the structure-based approach (the growth of the F-measure can reach up to 20%), as shown in Fig. 11.
- This improvements for points 1) and 2) are mainly due to good recall.
- Our approach is reliable for different types of hierarchies of ontologies (It is important to note that the proposed approach performs well on different types of hierarchies).

In conclusion, we can confirm that the combination improves both the structure-based and reasoning-based methods. In other words, we have found these methods to mutually benefit from being combined while overall preserving a good recall.

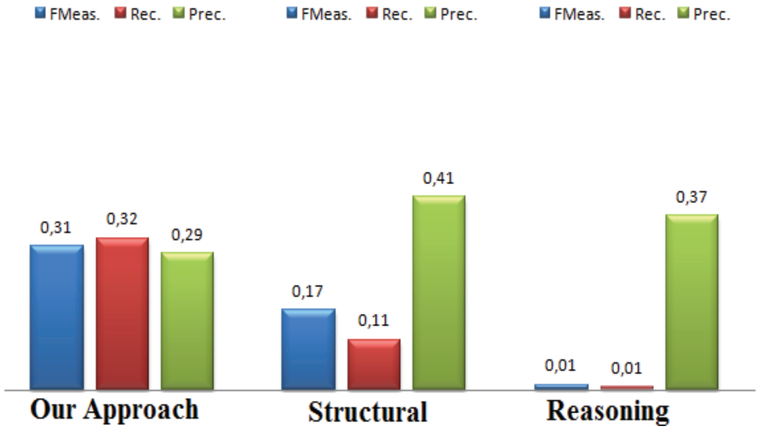

Fig. 11. Comparison of H-mean of our approach with the structure-based and reasoning-based methods.

7.4 Study of the Initial Alignment and the Identified Correspondences

In our experiment, we varied the size of the initial alignment and calculated the number of newly-identified correspondences by our approach without using a terminological matcher in order to study the correlation between the size of the initial alignment and the number of newly-identified correspondences.

Fig. 12 shows that, overall, the quality of the results is globally proportional to the sample size. The peaks are explained by the fact that the initial alignment may influence the semantic correspondences to identify i.e., inference of more relations: richer inferred shared taxonomy.

Ultimately, the results of our experiment show that our approach is reliable and efficient. We found that our approach is relevant in the sense that the reasoning-based approach is improved by the structural metrics of similarity computing for ontology alignment (i.e., the recall of our method is generally good compared to other systems).

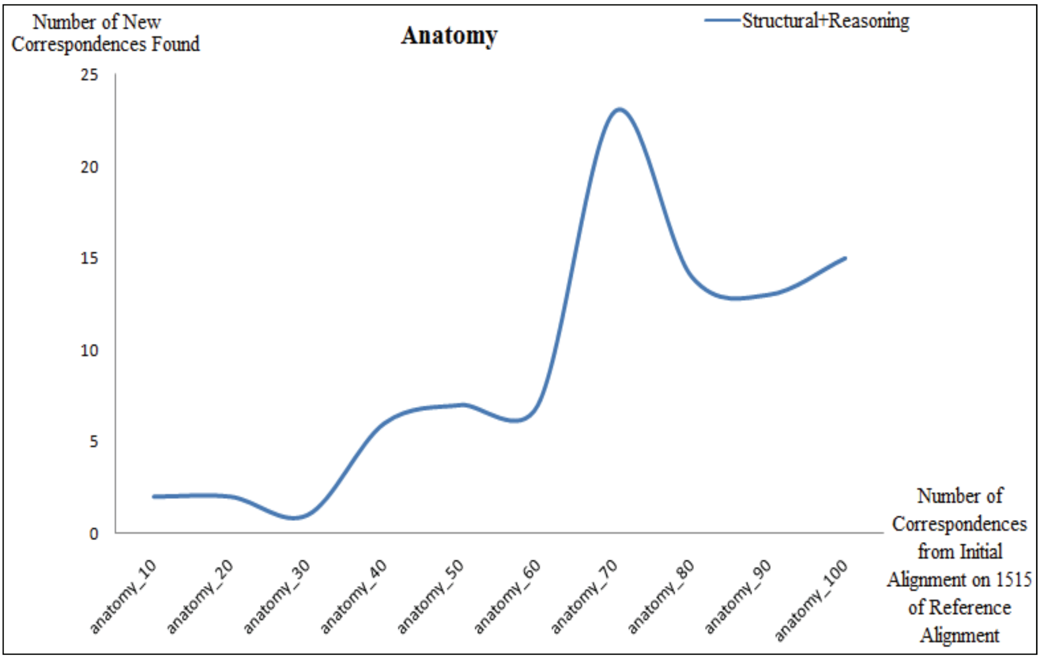

Fig. 12. Our approach on the different sample sizes on anatomy series.

8. Conclusions

In this paper, we have introduced a new ontology alignment approach, which consists of combining reasoning-based and structure-based approaches.

Our approach is based on description logic inference and structural similarity computing. It consists of obtaining the shared inferred taxonomy by first performing logical reasoning on the ontology that results from the merger, which is from an initial reference alignment of the two ontologies that are to be aligned. Then, an intra-taxonomic structural measure is applied.

In addition, we implemented our approach by developing a software tool that allowed for the evaluation of our proposed method.

We have shown the good performance of our approach through the experimental results that we obtained. In fact, these results are better than those of LogMap in terms of precision, recall, and F-measure. The approach has also been revealed to be more pertinent than YAM++ for certain types of ontologies and to significantly improve the structure-based and reasoning-based methods.

Furthermore, one of the advantages of our approach is that it is recommended to align ontologies that do not contain terminological information.

In regards to future perspectives for our work, we envision exploiting other inferred relations, such as disjunction and equivalence, in order to broaden the experiments on one side, along with diversifying the intra-taxonomic structural metrics and judiciously choosing an initial reference alignment. Indeed, the starting point may condition the quality of the final alignment. For instance, an initial alignment that touches the tips of the ontologies that are to be aligned could allow for the discovery of advantageous semantic correspondences.

References

[1] K. Todorov, P. Geibel, and K. U. Kuhnberger, "Mining concept similarities for heterogeneous ontologies," in *Proceedings of the 10th Industrial Conference on Advances in Data Mining: Applications and Theoretical Aspects (ICDM 2010)*, Berlin, Germany, 2010, pp. 86-100.

[2] J. Euzenat, and P. Shvaiko, *Ontology Matching,* 2nd ed. Heidelberg: Springer-Verlag, 2013.

[3] M. Ehrig, *Ontology Alignment: Bridging the Semantic Gap*. New York, NY: Springer, 2007.

[4] P. Bouquet, M. Ehrig, J. Euzenat, E. Franconi, P. Hitzler, M. Krotzsch, et al., "D2.2.1 Specification of a common framework for characterizing alignment," 2005; http://knowledgeweb.semanticweb.org/semanticportal/deliverables/D2.2.1v2.pdf.

[5] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. Patel-Schneider, *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge: Cambridge University Press, 2003.

[6] E. Sirin, B. Parsia, B. C. Grau, A. Kalyanpur, and Y. Katz, "Pellet: a practical OWL-DL reasoned," *Journal of Web Semantics*, vol. 5, no. 2, pp. 51-53, 2007.

[7] A. Maedche, B. Motik, N. Silva, and R. Volz, "Mafra—a mapping framework for distributed ontologies," in *Proceedings of 13th International Conference on Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web (EKAW2002),* Siguenza, Spain, 2002, pp. 235-250.

[8] J. Euzenat, "An API for ontology alignment," in *Proceeding of International Semantic Web Conference (ISWC2004)*, Hiroshima, Japan, 2004, pp. 698-712.

[9] E. Goldberg, and Y. Novikov, "BerkMin: a fast and robust SAT-solver," in *Proceedings of Design, Automation and Test in Europe Conference and Exhibition (DATE2002)*, Paris, France, 2002, pp. 142-149.

[10] J. P. Marques-Silva and K. A. Sakallah, "GRASP: a search algorithm for propositional satisfiability," *IEEE Transactions on Computers*, vol. 48, no. 5, pp. 506-521, 1999.

[11] B. Jafarpour and S. S. R. Abidi, "A semantic web based ontology mapping and instance transformation framework," in *Proceedings of Canadian Semantic Web Symposium (CSWS2013)*, Montreal, Canada, 2013, pp. 5-11.

[12] C. Meilicke, H. Stuckenschmidt, and A. Tamilin, "Improving automatically created mappings using logical reasoning," in *Proceedings of the 1st International Workshop on Ontology Matching (OM2006) co-located with the 5th International Semantic Web Conference (ISWC 2006),* Athens, GA, 2006, pp. 61-72.

[13] H. Stuckenschmidt, L. Serafini, and H. Wache, "Reasoning about ontology mappings," 2005; http://ub-madoc.bib.uni-mannheim.de/1140/1/kr06.pdf.

[14] E. Jimenez-Ruiz and B. C. Grau, "LogMap: logic-based and scalable ontology matching," in *Proceedings of 10th International Semantic Web Conference (ISWC2011)*, Bonn, Germany, 2011, pp. 273-288.

[15] J. H. Yang and I. J. Chung, "A method for automatic generation of OWL-S service ontology," *International Journal of Information Processing Systems*, vol. 2, no. 2, pp. 114-123, 2006.

[16] N. F. Noy and M. A. Musen, "Anchor-PROMPT: using non-local context for semantic alignment," in *Proceedings of the Workshop on Ontologies and Information Sharing at the International Joint Conference on Artificial Intelligence (IJCAI 2001)*, Seattle, WA, 2001, pp. 63-70.

[17] S. Melnik, H. Garcia-Molina, and E. Rahm, "Similarity flooding: a versatile graph matching algorithm and its application to schema matching," in *Proceedings of the 18th International Conference on Data Engineering (ICDE2002)*, San Jose, CA, 2002, pp. 117-128.

[18] J. Euzenat, and P. Valtchev, "An integrative proximity measure for ontology alignment," in *Proceedings of Workshop on Semantic Information Integration co-located with the 2th International Semantic Web Conference (ISWC2003)*, Sanibel Island, FL, 2003, pp. 33-38.

[19] P. Valtchev and J. Euzenat, "Dissimilarity measure for collections of objects and values," in *Proceedings of 2nd International Symposium on Advances in Intelligent Data Analysis, Reasoning about Data (IDA1997)*, London, 1997, pp. 259-272.

[20] J. Zhong, H. Zhu, Y. Li, and Y. Yu, "Conceptual graph matching for semantic search," in *Proceedings of 10th International Conference on Conceptual Structures (ICCS2002)*, Borovets, Bulgaria, 2002, pp. 92-106.

[21] A. Maedche and V. Zacharias, "Clustering ontology-based metadata in the semantic web," in *Proceedings of 6th European Conference on Principles of Data Mining and Knowledge Discovery (PKDD2002)*, Helsinki, Finland, 2002, pp. 348-360.

[22] B. B. Hariri, H. Abolhassani, and A. Khodaei, "A new structural similarity measure for ontology alignment," in *Proceedings of the 2006 International Conference on Semantic Web & Web Services (SWWS2006)*, Las Vegas, NV, 2006, pp. 36-42.

[23] F. Boudries, A. K. Tari, and C. Juiz, "Mapping d'ontologies dans un environnement distribué," in *Proceedings of Colloque sur l'Optimisation et les Systèmes d'Information (COSI2012)*, Tlemcen, Algeria, 2012.

[24] N. F. Noy and A. M. Musen, "The PROMPT suite: interactive tools for ontology merging and mapping," *International Journal of Human-Computer Studies*, vol. 59, no. 6, pp. 983-1024, 2003.

[25] I. V. Levenshtein, "Binary codes capable of correcting deletions, insertions, and reversals," *Soviet Physics Doklady*, vol. 10, no. 8, pp. 707-710, 1966.

[26] D. W. Jeong, H. Y. Shin, D. K. Baik, and Y. S. Jeong, "An efficient web ontology storage considering hierarchical knowledge for Jena-based applications," *Journal of Information Processing Systems*, vol. 5, no. 1, pp. 11-18, 2009.

[27] J. Euzenat, "Semantic precision and recall for ontology alignment evaluation," in *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI2007)*, Hyderabad, India, 2007, pp. 348-353.

Abderrahmane Khiat http://orcid.org/0000-0002-7708-7971

He received the Master's degree in computer science, specialty "Computer Science Theory" from the University of Djilali Liabes, Sidi Bel Abbes, Algeria, in 2011. He is currently working toward the PhD degree at the University of Oran 1 Ahmed Ben Bella, LITIO laboratory, Oran, Algeria. He obtained the price of The Best Paper Awards for Young Scientists Researchers at the International Conference COSI 2014, Béjaia, Algeria. His research covered several areas, including Ontology Matching, Linked Data, Ontology Engineering and Knowledge, Semantic Web, Natural Language Processing and Data Mining, Artificial Intelligence, Human-Machine Interaction and Robotics.

Moussa Benaissa http://orcid.org/0000-0002-6869-469X

He received his engineering degree and his master degree in computer science from the University of Oran 1 Ahmed Ben Bella, Algeria, in 1988 and 1992 respectively. He is currently a lecturer at University of Oran 1 Ahmed Ben Bella. His research interest domains are: Semantic Web, ontology engineering and e-learning.