Anchor-Profiles for Ontology Mapping with Partial Alignments

Frederik C. Schadd Nico Roos
Department of Knowledge Engineering,
Maastricht University, Maastricht, The Netherlands

Abstract.
Ontology mapping is a crucial step for the facilitation of information exchange between knowledge sources. In the industry this process is often performed semi-automatically, with a domain expert supervising the process. Such an expert can supply a partial alignment, known as anchors, which can be exploited with more elaborate mapping techniques in order to identify the remaining correspondences. To do this we propose a novel approach, referred to as anchor-profiles. For each concept its degree of similarity to each anchor is gathered into a profile for comparison. We evaluated our approach on the Ontology Alignment Evaluation Initiative (OAEI) benchmark dataset using partial alignments that are randomly generated from the reference alignments. The evaluation reveals an overall high performance when compared with mapping systems that participated in the OAEI2012 campaign, where larger partial alignments lead to a higher f-measure.

Keywords. semantic web, interoperability, ontology mapping, alignment reuse, anchor profiles

1. Introduction

The ability to access and process information is an ever increasing issue due to the rise of the internet and with it the ability to access knowledge sources across the world. While in the past information domains would be modelled using database schemas, recent developments resulted in the proliferation of semantic web technologies for this purpose. Many companies depend on these data solutions for their business operations and services. However, since every user of such technologies has different requirements and views of a given domain, it is likely that he or she would model this domain differently when compared to another user. This leads to the interoperability problem, where the presence of different domain specifications, known as ontologies, prohibits the exchange of information between the two different knowledge sources. As an example, this leads to problems when two businesses want to cooperate and access each other's knowledge bases, or when one company acquires another and wishes to integrate the new data into its existing knowledge system.

In order to circumvent this issue, the heterogeneous ontologies that are employed by the two different knowledge systems need to be mapped, such that for each data concept a corresponding concept in the other ontology is identified. Traditionally, this task has been done manually by domain experts. However, this solution is rarely appropriate due to the required manual labour being too much when mapping large ontologies. This lead to the development and evaluation of different ontology mapping approaches and systems [6, 17, 18], which aim to aid the domain expert with the mapping task or even perform the mapping of ontologies completely autonomously. The ability of automatically mapping ontologies becomes increasingly important with the ongoing development of the semantic web [3], which would allow autonomous agents to roam the semantic web and access heterogeneous data sources to provide different services to the user.

Following established work [5, 8], we define ontology mapping as a process that receives two ontologies O_1 and O_2 as input and produces an alignment A. Furthermore, there are a series of optional inputs to this process, like a previously computed alignment A', a set of parameters p and a list of external resources r. We define a correspondence between entities of two ontologies O_1 and O_2 as a 5-tuple $< id, e_1, e_2, r, c >$ such that:

- id is a unique identifier allowing the referral to specific correspondences.
- e_1 is a reference to an entity originating from the first ontology. Commonly a Uniform Resource Identifier (URI) is used as referral to a specific entity.
- e_2 is a reference to an entity originating from the second ontology.
- r denotes the semantic relation between e_1 and e_2. Several types of relations can be modelled, which are subsumption ($\sqsubseteq$), generalization ($\sqsupseteq$), undirected disjointness ($\perp$), overlapping ($\sqcap$) and equivalance ($\equiv$).
- c is a confidence value in the interval $[0, 1]$, which is used to express the certainty that the particular relation holds.

Given the definition of a correspondence, an alignment A between two ontologies O_1 and O_2 is defined as a set of correspondences where each correspondence contains a reference to one entity of O_1 and one entity of O_2. In order to evaluate A, it is compared to a reference alignment R, which contains correspondences as specified by a domain expert and thus are assumed to be correct. In this research, we will explore a mapping technique which exploits the optional input alignment A'. Here, we make the distinction between two cases, being (1) a complete alignment and (2) a partial alignment (PA). In a complete alignment, all concepts which a domain expert would map are present in at least one correspondence, where the main focus of the applied techniques lie on refining the already existing alignment. However, in a partial alignment this is not the case, resulting in the problem that one has to discover the remaining correspondences. The proposed technique is designed to deal with the second case, where a partial alignment needs to be completed. Such a partial alignment can originate from a domain expert who, due to time constraints, cannot produce a complete mapping. For this research, we will assume that all correspondences in PA, also referred to as anchors, are correct, such that $PA \subset R$ and included in A, such that $PA \subset A$.

2. Related Work

Several works exist that have tackled approaches which reuse previously generated alignments. This type of approach has initially been suggested by Rahm et al. [13]. Here, the focus lies on finding auxiliary ontologies which are already mapped to the target ontology. This has the intention that, by selecting the auxiliary ontology according to a specific criteria, the remaining mapping problem between the source and auxiliary ontology might be easier to solve than the original problem. Subsequent works have expanded this idea to deriving mappings when both input ontologies have an existing alignment to an auxiliary ontology.

COMA++ employs several strategies with regard to exploiting pre-existing alignments [2]. Most prominently, the system can explore alignment paths of variable lengths between multiple ontologies, which are obtained from a corpus, in order to derive its mappings. It is also possible to explore ontologies from the semantic web for this purpose [14]. The resulting mapping derivations of multiple alignment paths can be combined to form a more reliable mapping.

While the previously mentioned approached utilized complete mappings involving auxiliary ontologies, there has been some research into approaches that exploit partial alignments that exist between the source and target ontologies. These alignments can either be user generated, by for instance using the PROMPT tool [12], or automatically generated from a different system.

The most prominent approach is the Anchor-PROMPT [11] algorithm. Here, possible paths between anchors are iteratively explored in parallel in both ontologies while the encountered concept combinations are registered. The intuition is that concept pairs which have been encountered regularly during the exploration phase are more likely to correspond with each other.

The Anchor-Flood algorithm also features a type of iterative exploration by exploiting anchors [16]. This approach selects a main anchor and iteratively expands the explored neighbourhood of this anchor. At each iteration, a matching step is invoked which compares the concepts in this neighbourhood and updates the alignment if new correspondences are found.

3. Anchor Profiles

A profile similarity gathers context information of ontology concepts and compares these context collections by parsing them into a vector space and comparing the resulting vectors. This context information can consist of data from the concept description and the descriptions of related concepts [10]. The intuition behind this approach is that concepts can be considered similar if they have similar context information. More generally, a profile can be considered as a vector generated from data which describes a concept, hence two concepts are similar if their profiles can be considered similar.

When mapping two ontologies for which a partial alignment is provided by a domain expert, new opportunities arise when selecting similarity measures for a mapping system. Instead of using description information as the basis for a profile, we suggest utilizing the correspondences of the given partial alignment, also referred to as anchors, as basis for a new kind of profile similarity. Here, since the anchors are assumed to be correct, the main intuition is that two concepts can be considered similar if they exhibit a comparable degree of similarity towards a given anchor. More formally, given two ontologies O_1 and O_2, and given an anchor $A_x[C^1, C^2]$ containing a correspondence between the concepts C^1 and C^2 originating from O_1 and O_2 respectively, and given a concept similarity $sim'(E, F) \in [0, 1]$ which expresses the similarity between two concepts, we define an anchor similarity $sim_A(C, A_x)$ between an arbitrary concept C and A_x as:

$$sim_A(C, A_x) = \begin{cases} sim'(C, C^2) & \text{if } C \in O_1 \\ sim'(C, C^1) & \text{if } C \in O_2 \end{cases} \tag{1}$$

Note that C is compared to the concept in the anchor which originates from the other ontology. If one were to compare C to the anchor concept from the same ontology, sim' would be reduced to a structural similarity, similar to a taxonomy distance, making the distinction between classes that are related equivalently close to a given anchor prohibitively difficult. From equation 1 follows that two concepts C and D can be considered similar if $sim_A(C, A_x)$ and $sim_A(D, A_x)$ are similar. Given that a partial alignment most likely contains multiple correspondences, this intuition needs to be expanded for a series of anchors. This brings us back to the generalized idea of a profile, such that we can use the anchor similarities sim_A between a concept C and all anchors as the basis of a profile, referred to as anchor-profile. Figure 1 visualizes the concept of an anchor-profile similarity.

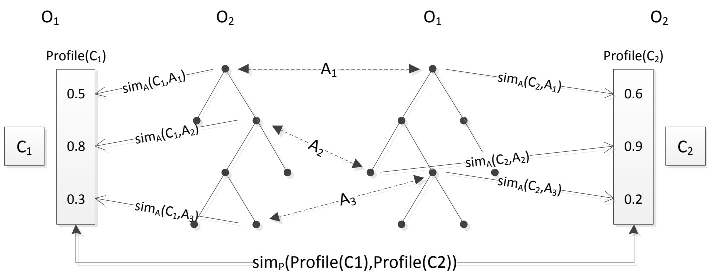

Figure 1. Visualization of an anchor profile similarity.

The example in Figure 1 shows two ontologies, O_1 and O_2, and three anchors A_1, A_2 and A_3. Two concepts C_1 and C_2, originating from O_1 and O_2 respectively, are compared using their respective anchor-profiles $Profile(C_1)$ and $Profile(C_2)$. The profile vectors are compared using the similarity sim_P. While there exist various similarity measures for vectors, for this research the well-known cosine-similarity [19] has been applied as sim_P.

Since the main intuition of this approach is that corresponding concepts should exhibit a comparable degree of similarity towards the given anchors, it is necessary to choose sim' such that this metric is robust under a wide variety of

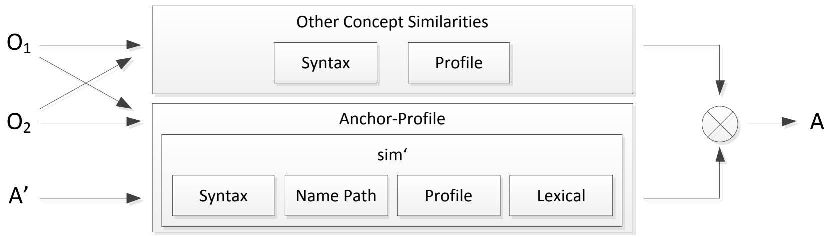

Figure 2. Overview of the tested mapping system.

circumstances. Since every single metric has potential weaknesses [17], it is preferable to aggregate different metrics in order to overcome these. To realise this, sim' utilizes the aggregate of all similarities from the *MaasMatch* system [15]. Figure 2 displays the configuration of the evaluated mapping system. Here, two distinct similarity matrices are computed, being the similarities of the anchor-profiles and an aggregate of other metrics. This second matrix is necessary for the eventuality where the system has to differentiate correspondences that all contain anchor-profiles which closely resemble null vectors, which occurs when a concept displays no similarity to any of the given anchors. This can occur when a given ontology has a considerable concept diversity and the given anchors do not adequately cover the concept taxonomy. The aggregate of these two matrices is then used to extract the output alignment A.

4. Experiments

Evaluation Measures The most established way of evaluating an ontology mapping approach is to produce an alignment between two ontologies O_1 and O_2 for which a correct alignment already exists. This allows for the comparison between the two alignments with the intuition that the correctness of the computed alignment is in correlation with its similarity to the reference. Provided such a golden standard, one can compute measures such as precision, recall and f-measure in order to express to what extent a computed alignment resembles the reference [7]. Given a generated alignment A and a reference alignment R, the precision $P(A,R)$ and recall $R(A,R)$ of alignment A can be computed as follows:

$$P(A,R) = \frac{R \cap A}{A} \qquad (2) \qquad\qquad R(A,R) = \frac{R \cap A}{R} \qquad (3)$$

The F-Measure can be used to express the overall quality of an alignment. It is calculated by computing the harmonic mean of the precision and recall measures. Given the precision $P(A,R)$ and recall $R(A,R)$ of an alignment A, the f-measure $F(A,R)$ of alignment A can be computed as follows:

$$F(A,R) = \frac{2 * P(A,R) * R(A,R)}{P(A,R) + R(A,R)} \qquad (4)$$

Evaluating with Partial Alignments While the measures of precision, recall and f-measure are widely used as evaluation criteria for ontology mapping approaches [6], they are inadequate when evaluating a mapping procedure which utilized a partial alignment as input. This issue stems from the assumption that correspondences in PA are assumed to be correct, and hence included in the output alignment A. Computing the standard measures using A would then result in biased values, concealing the true quality of the new computed correspondences. It follows that in order to measure the quality of the computed correspondences, it becomes necessary to exclude the correspondences of PA from this computation.

To achieve this, the standard measures of precision, recall and f-measure can be modified in order to remove this bias. Given an alignment A, a reference R and a partial alignment PA that was provided as input for a given mapping task, the adapted measures of precision $P^*(A,R,PA)$ and recall $R^*(A,R,PA)$ can be computed as follows:

$$P^*(A,R,PA) = \frac{\mid A \cap R \cap \overline{PA} \mid}{\mid A \cap \overline{PA} \mid} \quad (5) \qquad R^*(A,R,PA) = \frac{\mid A \cap R \cap \overline{PA} \mid}{\mid R \cap \overline{PA} \mid} \quad (6)$$

Given these adapted measures of precision and recall, the adapted f-measure $F^*(A,R,PA)$ can be computed as follows:

$$F^*(A,R,PA) = \frac{2 * P^*(A,R,PA) * R^*(A,R,PA)}{P^*(A,R,PA) + R^*(A,R,PA)} \tag{7}$$

Evaluating Datasets In order to evaluate the performance of a mapping approach which exploits partial alignments, it is necessary to have access to a dataset which not only contains appropriate mapping tasks and their reference alignments, but also partial alignments that can be used as input. However, within the boundaries of the OAEI competition, which allows a comparison with other frameworks, there does not exist a dataset which also supplies partial alignments as additional input. When a dataset does not contain partial alignments, it is possible to generate these by drawing correspondences from the reference alignment at random. However, in order to account for the random variation introduced by the generated partial alignments, it becomes necessary to repeatedly evaluate the dataset using many generated partial alignments for each mapping task. The values of precision, recall and f-measure can then be aggregated using the arithmetic mean.

Next to establishing the mean performance of a system, it is also interesting to see how stable its performance is. Traditionally, this is expressed via the standard deviation. However, given that in this domain the measurements origin from different tasks of differing complexity, this introduces a problem. Given the presence of tasks of varying complexity that can occur in a dataset, it is to be expected that the mean performances of the repeated evaluations differ for each task. Thus, in order to combine the standard deviations of the different tasks, a statistical measure is needed that takes this into account. To do this we propose using the pooled standard deviation of the different measures [4, 9].

Given k samples, the different sample sizes $n_1, n_2, \ldots, n_k$ and sample variances $s_1^2, s_2^2, \ldots, s_k^2$, the pooled standard deviation of the collection of samples can be calculated as follows:

$$s' = \sqrt{\frac{(n_1 - 1) \times s_1^2 + (n_2 - 1) \times s_2^2 + \cdots + (n_k - 1) \times s_k^2}{n_1 + n_2 + \cdots + n_k - k}} \tag{8}$$

In this domain, the repeated evaluation of a single track using randomly generated partial alignments can be viewed as a sample, such that the pooled standard deviation expresses how much the results deviate across all tracks. For the remainder of this paper, we will refer to the pooled standard deviation of P^*, R^* and F^* as s'_{P^*}, s'_{R^*} and s'_{F^*} respectively.

4.1. Evaluation

To evaluate an anchor-profile approach, an ontology mapping system incorporating the proposed similarity has been evaluated on the *benchmark-biblio* dataset originating from the 2012 Ontology Alignment Evaluation Initiative [1]. This synthetic dataset consists of tasks where each task tests a certain limiting aspect of the mapping process, for instance by distorting or removing certain features of an ontology like concept names, comments or properties. Since this dataset does not contain partial alignments that can be used as input, they were randomly generated from the reference alignments. In order to evaluate what impact the size of the partial alignment can have on the mapping process, we evaluated our approach over a spectrum of partial alignment recall values $[0.1, 0.2, \ldots, 0.9]$. Thus, as an example, a partial alignment recall of 0.2 indicates that *PA* was randomly generated from the reference R such that *PA* has a recall of 0.2. In order to mitigate the variance introduced through the random generation of *PA*, each recall level has been evaluated 100 times where each evaluation contained a new set of randomly generated partial alignments. For each evaluation, the adapted measures of precision, recall and f-measure, P^*, R^* and F^* respectively, were computed and aggregated. Table 1 displays the aggregated results of the evaluation.

PA Recall	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
P^*	0.760	0.769	0.779	0.786	0.801	0.817	0.835	0.855	0.866
R^*	0.632	0.641	0.649	0.656	0.663	0.674	0.685	0.702	0.745
F^*	0.668	0.678	0.686	0.693	0.701	0.713	0.726	0.743	0.780
s'_{P^*}	0.094	0.099	0.107	0.112	0.125	0.139	0.155	0.180	0.219
s'_{R^*}	0.049	0.068	0.083	0.092	0.102	0.117	0.133	0.158	0.215
s'_{F^*}	0.038	0.053	0.066	0.074	0.083	0.098	0.115	0.142	0.199

Table 1. Results of the evaluations on the benchmark-biblio dataset using different recall requirements for the randomly generated partial alignments. For each recall requirement, 100 evaluations were performed and aggregated.

From Table 1, several interesting results and trends can be seen. First, we can see that overall for all *PA* recall levels the system resulted in an adapted precision in the interval $[0.76, 0.87]$, adapted recall in the interval $[0.63, 0.75]$ and adapted f-measure in the interval $[0.66, 0.78]$. Thus, for every PA recall level the approach resulted in high precision and moderately high recall measure. Furthermore, we can observe that as the recall of PA increases, the adapted precision, recall and f-measure of A increase as well. This increase is fairly consistent over all PA recall levels, indicating that a larger amount of anchors improves the representative strength of the computed anchor profiles.

Inspecting s'_{P^*}, s'_{R^*} and s'_{F^*} reveals each measure shows a similar trend. For each measure, an increase of the recall level of PA also yields an increase of the pooled standard deviation, with the resulting alignments at PA recall level of 0.1 being fairly stable, while a moderate variance can be observed at a PA recall level of 0.9. This trend is to be expected since any variation in A will have a larger impact on P^*, R^* and F^* if PA has a significant size.

Having established the overall performance of the proposed approach for different size levels of PA, it would be interesting to inspect the performance over the different tasks of the dataset. Table 2 shows the mean f-measure over several task groups when using different PA size levels, of which three were selected for the sake of brevity. The task groups reflect different kinds of alterations in the target ontology, such as altered names and comments (201-202, 248-253, 254-266) and altered or suppressed structures or properties (221-228, 232-247, 248-253, 254-266), where groups *232-247 and 254-266* test combinations of structural changes. The results indicate that the approach is robust against structural deformities, while being susceptible to deformity in the concept names and descriptions. This can be attributed to the underlying similarities used in sim', meaning that a wider spectrum of aggregate similarities could remedy this weakness.

PA recall	101	201-202	221-228	232-247	248-253	254-266
0.1	1.0	0.71	1.0	1.0	0.59	0.57
0.5	1.0	0.75	1.0	1.0	0.63	0.61
0.9	1.0	0.86	1.0	1.0	0.74	0.69

Table 2. Average f-measure of the anchor-profile approach using three different PA size thresholds across the different task groups of the benchmark-biblio dataset.

Next to establishing the overall performance on the benchmark dataset, it is also important to provide some context to that performance. To do this, we will compare the performance of the Anchor-Profile approach with the top 8 frameworks out of 18 frameworks that participated in the OAEI 2012 competition [1] in Table 3. Unfortunately, none of the OAEI evaluations contained a task which also provided partial alignments, however a comparison with state-of-the-art systems which tackled the same task without a partial alignment can still be a useful performance indication. For this comparison, both the smallest and largest evaluated PA size levels were used.

The results of Table 3 indicate that the quality of correspondences produced by our approach is in line with the top ontology mapping frameworks in the field. In fact, when including PA in the evaluation metrics, the anchor-profile approach outperforms these frameworks given a large enough recall level[1] of PA. Using

[1]The results of the experiments indicate that a recall level of 0.5 would suffice.

System	Precision	Recall	F-Measure
MapSSS	0.99	0.77	0.87
YAM++	0.98	0.72	0.83
Anchor-Profile (0.9)	0.866*(0.998)	0.745*(0.967)	0.78*(0.982)
AROMA	0.98	0.64	0.77
WeSeE	0.99	0.53	0.69
AUTOMSv2	0.97	0.54	0.69
Hertuda	0.9	0.54	0.68
Anchor-Profile (0.1)	0.760*(0.88)	0.632*(0.623)	0.668*(0.691)
HotMatch	0.96	0.5	0.66
Optima	0.89	0.49	0.63

Table 3. Comparison of the Anchor-Profile approach, using two different PA thresholds, with the 8 best performing frameworks from the OAEI 2012 competition. An asterisk indicates the value has been adapted with respect to PA, while the values inside the brackets indicate the respective measure over the entire alignment.

partial alignments with a recall of 0.1 resulted in an f-measure similar to the HotMatch framework, ranking at 8th place in this comparison. A *PA* recall level of 0.9 resulted in a sufficiently high f-measure to rank 3rd among the top ranking systems. With regards to precision and recall, our system differentiates itself from other frameworks by having a comparatively lower precision and higher recall. This indicates that our approach is capable of identifying correspondences which other system cannot, while further measures must be implemented to differentiate between correspondences that have similar anchor profiles.

5. Conclusion

In this paper, we proposed the creation of concept profile based on their similarities to given anchor correspondences. We evaluated this approach on the OAEI benchmark-biblio dataset using various sizes of input partial alignments, which are used as anchors. The performed experiments revealed an high performance, where the aggregated f-measure is positively correlated with the size of the input partial alignment. The investigation into the different tracks revealed that, while the approach is robust against distortions of the ontology properties, hierarchy and instances, it is susceptible to disruptions of the concept names and descriptions. However, this tendency can be explained by the used metrics which determine the similarity between the concepts and individual anchor. A comparison with other mapping systems revealed that the quality of the computed correspondences is on-par with state-of-the-art systems when using only small input partial alignments. When using larger partial alignments, the proposed approach outperforms most of the compared systems.

This research assumed that all the correspondences in the input partial alignment are correct, due to the assumption that these would be generated by a domain expert. However, this might not always be the case since even experts can make mistakes. Further research should investigate to what extent this approach can work if there are some anchors which represent incorrect correspondences.

References

[1] Aguirre, J., Grau, B., Eckert, K., Euzenat, J., Ferrara, A., van Hague, R., Hollink, L., Jimenez-Ruiz, E., Meilicke, C., Nikolov, A., Ritze, D., Shvaiko, P., Svab-Zamazal, O., Trojahn, C., Zapilko, B.: Results of the ontology alignment evaluation initiative 2012. In: Proc. of the 7th ISWC workshop on ontology matching. pp. 73–115 (2012)

[2] Aumueller, D., Do, H.H., Massmann, S., Rahm, E.: Schema and ontology matching with coma++. In: Proceedings of the 2005 ACM SIGMOD international conference on Management of data. pp. 906–908. ACM (2005)

[3] Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Scientific American 284(5), 34–43 (2001)

[4] Dodge, Y.: Pooled variance. In: The Concise Encyclopedia of Statistics, pp. 427–428. Springer New York (2008)

[5] Euzenat, J.: An api for ontology alignment. Proceedings of the International Semantic Web Conference (ISWC) pp. 698–712 (2004)

[6] Euzenat, J., Meilicke, C., Stuckenschmidt, H., Shvaiko, P., Trojahn, C.: Ontology alignment evaluation initiative: six years of experience. Journal on data semantics XV pp. 158–192 (2011)

[7] Giunchiglia, F., Yatskevich, M., Avesani, P., Shvaiko, P.: A large dataset for the evaluation of ontology matching. The Knowledge Engineering Review Journal 24, 137–157 (June 2009)

[8] Giunchiglia, F., Shvaiko, P.: Semantic matching. The Knowledge Engineering Review 18(3), 265–280 (2003)

[9] Killeen, P.R.: An alternative to null-hypothesis significance tests. Psychological science 16(5), 345–353 (2005)

[10] Mao, M., Peng, Y., Spring, M.: A profile propagation and information retrieval based ontology mapping approach. In: Proceedings of the Third International Conference on Semantics, Knowledge and Grid. pp. 164–169. IEEE (2007)

[11] Noy, N.F., Musen, M.A.: Anchor-prompt: Using non-local context for semantic matching. In: Proceedings of the ICJAI workshop on ontologies and information sharing. pp. 63–70 (2001)

[12] Noy, N.F., Musen, M.A.: Algorithm and tool for automated ontology merging and alignment. In: Proceedings of AAAI-00. (2000)

[13] Rahm, E., Bernstein, P.A.: A survey of approaches to automatic schema matching. the VLDB Journal 10(4), 334–350 (2001)

[14] Sabou, M., dAquin, M., Motta, E.: Exploring the semantic web as background knowledge for ontology matching. In: Journal on Data Semantics XI, Lecture Notes in Computer Science, vol. 5383, pp. 156–190. Springer Berlin Heidelberg (2008)

[15] Schadd, F., Roos, N.: Maasmatch results for oaei 2012. In: Proceedings of The Seventh ISWC International Workshop on Ontology Matching. pp. 160–167 (2012)

[16] Seddiqui, M.H., Aono, M.: An efficient and scalable algorithm for segmented alignment of ontologies of arbitrary size. Web Semantics: Science, Services and Agents on the World Wide Web 7(4), 344–356 (2009)

[17] Shvaiko, P., Euzenat, J.: A survey of schema-based matching approaches. In: Journal on Data Semantics IV, vol. 3730, pp. 146–171. Springer (2005)

[18] Stuckenschmidt, H., Klein, M.: Structure-based partitioning of large concept hierarchies. The Semantic Web–ISWC 2004 pp. 289–303 (2004)

[19] Tan, P.N., Steinbach, M., Kumar, V.: Introduction to Data Mining. Addison Wesley, 1 edn. (May 2005)