User Involvement in Ontology Matching Using an Online Active Learning Approach

Booma Sowkarthiga Balasubramani, Aynaz Taheri, and Isabel F. Cruz

ADVIS Lab
Department of Computer Science
University of Illinois at Chicago
{bbalas3,ataher2,ifcruz}@uic.edu

Abstract. We propose a semi-automatic ontology matching system using a hybrid active learning and online learning approach. Following the former paradigm, those mappings whose validation is estimated to lead to greater quality gain are selected for user validation, a process that occurs in each iteration, following the online learning paradigm. Experimental results demonstrate the effectiveness of our approach.

1 Introduction

The result of performing ontology matching is a set of mappings between concepts in the *source ontology* and concepts in the *target ontology*. This set is called an *alignment*. The *reference alignment* or *gold standard* is (an approximation of) the set of correct and complete mappings built by domain experts. We consider a semi-automatic ontology matching approach, whereby the mappings are first determined using automatic ontology matching methods, which we call *matchers*, followed by user validation.

We use six of the matchers of the AgreementMaker ontology matching system [3], including the Linear Weighted Combination (LWC) matcher, which performs a weighted combination of the results of the other five matchers, using weights that are automatically determined using a quality metric [4].

We train a classifier and modify the weights of the LWC matcher using an iterative approach, following the on-line learning paradigm. At each iteration, user validation is sought for those candidate mappings that can potentially contribute the most to the quality of the final alignment, following the active learning paradigm. The process continues until there is no significant improvement in F-Measure. We describe this process in Section 2. Experimental results are obtained using the ontology sets from the Ontology Alignment Evaluation Initiative (OAEI) and comparison is made with the results of other systems in Section 3. We discuss related work in Section 4, and conclude with Section 5.

2 Proposed System

After the source and target ontologies are loaded into AgreementMaker, the following steps are executed in sequence:

Automatic matching algorithms execution The following matchers are executed individually and their results are stored in the corresponding similarity matrices: the Advanced Similarity Matcher (ASM) [5], the Parametric String-based

Matcher (PSM) [4], the Lexical Similarity Matcher (LSM) [5], the Vector-based Multi-word Matcher (VMM) [4], and the Base Similarity Matcher (BSM) [5].

Linear weighted combination The Linear Weight Combination (LWC) matcher [6] linearly combines the similarity matrices of the other five automatic matchers using weights determined by the local confidence quality metric, which estimates the quality of the scores produced by each matcher. The new score for each mapping is stored in the LWC matrix. It is up to the selection phase to output only those mappings that are in the final alignment, taking into account the desired cardinality of the mappings (e.g., one-to-one) [4].

Candidate mapping selection Candidate mappings to be presented to the users for validation are based on the combination of the following three criteria: (1) Disagreement-based Top-k Mapping [6], which measures the level of similarity among the five scores, one for each of the matchers considered. If the matchers mostly agree on the scores, then the disagreement is low, but it is high when the matchers disagree on the scores; (2) Cross Count Quality (CCQ), which counts, for a score, the number of non-zero scores in the row and column of that score in the LWC matrix [2]. The count is normalized by the maximum sum of the scores per column and row in the whole matrix; (3) Similarity Score Definiteness (SSD), which is a quality metric that ranks mappings in increasing order of their score [2]. It evaluates how close the score associated with a mapping is to the maximum and minimum possible scores (1 and 0).

User validation The result of this step is a label that has value 1 if the mapping is correct and 0 if the mapping is incorrect. For each iteration, users validate a set of candidate mappings. The validation of each mapping is called an *interaction* by others [7]. There can be any number of interactions per iteration, that is, users can be presented with any number of mappings to validate at a time.

Classification We use a logistic regression classifier, which considers the parametric distribution $P(Y|X)$ where Y is the discrete-valued user label (1 or 0) and the feature vector $X = \langle X_1, \ldots, X_n \rangle$ is the signature vector [6] with n scores computed for a mapping by n individual matchers, and estimates the parameter that is the vector of weights $W = \langle w_1, \ldots, w_n \rangle$ of the LWC matcher. The logistic regression model is based on the following probabilities:

$$P(Y = 1|X) = \frac{1}{1 + e^{w_0 + \sum_{i=1}^{n} w_i X_i}}, P(Y = 0|X) = \frac{e^{w_0 + \sum_{i=1}^{n} w_i X_i}}{1 + e^{w_0 + \sum_{i=1}^{n} w_i X_i}}$$

W is updated during the iterative process by taking the partial derivative of the log likelihood function with respect to each component, w_i. The recursive rule for the update is as follows, where α is the learning rate that determines how fast or slow the weights will converge to their optimal values [10]:

$$W \leftarrow W + \alpha \sum_{i=1}^{m} X^i (Y^i - g(W^T X^i))$$

3 Experimental Evaluation

We use the 2014 OAEI Conference Track ontology sets and their reference alignments to simulate the user validation. The baseline is the F-Measure obtained automatically by the AgreementMaker matchers. Table 1 depicts the average F-Measure after 20 iterations using the three candidate selection criteria individually or in combination with one another. The top performer is the Disagreement-based Top-k Mapping Selection criteria.

	1	2	3	4	5	6	7
Candidate Mapping Selection Strategy	48.08	52.45	60.43	51.42	48.91	52.47	53.18
Baseline (Before User Feedback)	51.8	51.8	51.8	51.8	51.8	51.8	51.8

Strategies: 1. CCQ 2. SSD 3. Disagreement 4. CCQ + SSD 5. CCQ + Disagreement 6. SSD + Disagreement 7. CCQ + SSD + Disagreement

Table 1: Average F-Measure for 20 iterations (123 interactions/iteration).

Matcher	F-Measure with User Feedback	F-Measure w/o User Feedback	F-Measure gain	Relative Number of Interactions
AML	0.801	0.730	0.071	0.497
LogMap	0.729	0.680	0.049	0.391
HerTUDA	0.582	0.600	-0.018	0.996
WeSeE	0.473	0.610	-0.137	0.447
Our Approach	0.604	0.518	0.086	0.470

Table 2: Comparison with the 2014 OAEI Interactive Track results.

Our approach has an average F-Measure gain of 8.6% and an average F-Measure of 60.4%. This is a considerable improvement as we started from an average F-Measure of 51.8%, which was obtained using the automatic matchers along with LWC. Table 2 compares our results with those obtained by other systems that participated in the 2014 OAEI Interactive Track. It performs better than HerTUDA and WeSeE (with F-Measure values of 58.2% and 47.3%, respectively). The F-Measure gain of AML [9] is 7.1% and of LogMap is 4.6%, therefore our approach has the highest F-Measure gain. The table also shows the relative number of interactions, which is the average number of interactions per pair of ontologies divided by the size of the reference alignment for that pair. Our approach shows better improvement in F-Measure with fewer number of interactions when compared to AML that has the highest F-Measure.

Figure 1 shows the effect of the total number of interactions on the F-Measure in our approach. Here, the total number of interactions represent the sum of the number of interactions in each of the 21 reference alignments in the Conference Track dataset (one for each pair of ontologies) up to 123 interactions. The Disagreement-based Top-k Mapping Selection performs better than the other candidate selection strategies. SSD and the combination of SSD+CCQ+ Disagreement have the next highest average F-Measure.

4 Comparison with Related Work

We divide previous work into two categories depending on whether feedback from single or multiple users is considered.

Single user A previous approach that uses AgreementMaker performs updates in the LWC matrix based on user feedback [6], but does not use a classifier to adjust

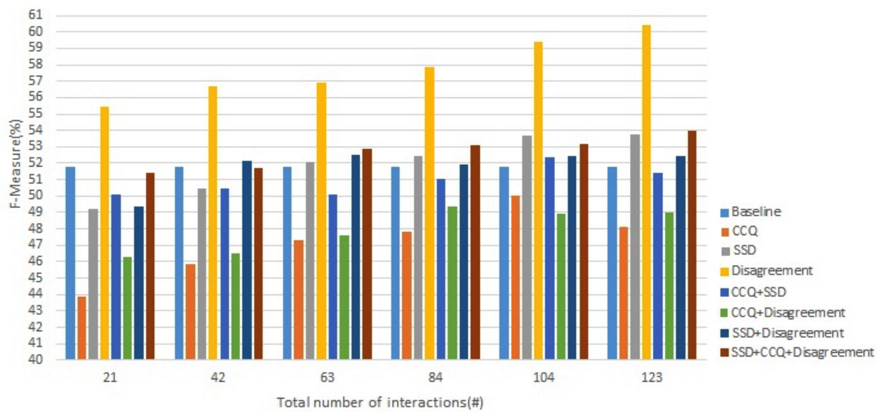

Fig. 1: F-Measure gain as a function of the number of interactions.

the LWC weights. Another method uses logistic regression to learn an optimal combination of both lexical and structural similarity metrics [8]. Compared to our approach, it uses different similarity metrics, candidate selection strategies, and techniques to customize weights for different matching strategies. Another system aggregates similarity measures with the help of self-organizing maps and incorporates user feedback for refining self-organizing map outcomes [11]. There is an active learning approach where the user validation is propagated according to the ontology structure [13]. Another approach makes use of the parameterization of matchers [12]. It uses example mappings to automatically determine a suitable parameter setting for each matcher, based on those examples. However, in our approach, the LWC uses five of the already existing matchers with the same configuration as in AgreementMaker.

Multiple users We discuss two approaches. The first one uses a pay-as-you-go approach and propagates the (possibly faulty) user validation input to similar mappings [2]. In the second approach, a multi-user feedback method that attempts to maximize the benefits that can be drawn from user feedback, by managing it as a first class citizen [1]. None of these approaches uses a classifier.

5 Conclusions and Future Work

In this paper, we have proposed an effective semi-automatic ontology matching approach that combines active learning with online learning. Our experimental evaluation demonstrate that a considerable improvement in F-Measure can be achieved over the base case. Clearly, a combination of user feedback with learning is fertile ground for future research, where the scalability of the methods to large and very large ontologies and the use of a variety of classifiers and of candidate selection strategies would be some of the topics to investigate.

Acknowledgments

This research was partially supported by NSF Awards IIS-1143926, IIS-1213013, and CCF-1331800.

References

1. Belhajjame, K., Paton, N.W., Fernandes, A.A.A., Hedeler, C., Embury, S.M.: User Feedback as a First Class Citizen in Information Integration Systems. In: CIDR Conference on Innovative Data Systems Research. pp. 175–183 (2011)
2. Cruz, I.F., Loprete, F., Palmonari, M., Stroe, C., Taheri, A.: Pay-As-You-Go Multi-User Feedback Model for Ontology Matching. In: International Conference on Knowledge Engineering and Knowledge Management (EKAW), pp. 80–96. Springer (2014)
3. Cruz, I.F., Palandri Antonelli, F., Stroe, C.: AgreementMaker: Efficient Matching for Large Real-World Schemas and Ontologies. PVLDB 2(2), 1586–1589 (2009)
4. Cruz, I.F., Palandri Antonelli, F., Stroe, C.: Efficient Selection of Mappings and Automatic Quality-driven Combination of Matching Methods. In: ISWC International Workshop on Ontology Matching (OM). CEUR Workshop Proceedings, vol. 551, pp. 49–60 (2009)
5. Cruz, I.F., Stroe, C., Caci, M., Caimi, F., Palmonari, M., Palandri Antonelli, F., Keles, U.C.: Using AgreementMaker to Align Ontologies for OAEI 2010. In: ISWC International Workshop on Ontology Matching (OM). CEUR Workshop Proceedings, vol. 689, pp. 118–125 (2010)
6. Cruz, I.F., Stroe, C., Palmonari, M.: Interactive User Feedback in Ontology Matching Using Signature Vectors. In: IEEE International Conference on Data Engineering (ICDE). pp. 1321–1324 (2012)
7. Dragisic, Z., Eckert, K., Euzenat, J., Faria, D., Ferrara, A., Granada, R., Ivanova, V., Jiménez-Ruiz, E., Kempf, A.O., Lambrix, P., Montanelli, S., Paulheim, H., Ritze, D., Shvaiko, P., Solimando, A., dos Santos, C.T., Zamazal, O., Grau, B.C.: Results of the Ontology Alignment Evaluation Initiative 2014. In: ISWC International Workshop on Ontology Matching (OM). pp. 61–104. CEUR Workshop Proceedings (2014)
8. Duan, S., Fokoue, A., Srinivas, K.: One Size Does Not Fit All: Customizing Ontology Alignment Using User Feedback. In: International Semantic Web Conference (ISWC). Lecture Notes in Computer Science, vol. 6496, pp. 177–192. Springer (2010)
9. Faria, D., Pesquita, C., Santos, E., Palmonari, M., Cruz, I.F., Couto, F.M.: The AgreementMakerLight Ontology Matching System. In: International Conference on Ontologies, DataBases, and Applications of Semantics (ODBASE). pp. 527–541. Springer (2013)
10. Halloran, J.: Classification: Naive Bayes vs Logistic Regression. Tech. rep., University of Hawaii at Manoa EE 645 (2009)
11. Jirkovskỳ, V., Ichise, R.: Mapsom: User Involvement in Ontology Matching. In: Joint International Semantic Technology Conference (JIST), pp. 348–363. Springer (2014)
12. Ritze, D., Paulheim, H.: Towards an Automatic Parameterization of Ontology Matching Tools Based on Example Mappings. In: ISWC International Workshop on Ontology Matching (OM). pp. 37–48 (2011)
13. Shi, F., Li, J., Tang, J., Xie, G., Li, H.: Actively Learning Ontology Matching via User Interaction. In: International Semantic Web Conference (ISWC). Lecture Notes in Computer Science, vol. 5823, pp. 585–600. Springer (2009)