

## Article

# Matching Ontologies through Multi-Objective Evolutionary Algorithm with Relevance Matrix

Hai Zhu <sup>1</sup>, Xingsi Xue <sup>2,\*</sup> and Hongfeng Wang <sup>1</sup>

<sup>1</sup> School of Network Engineering, Zhoukou Normal University, Zhoukou 466001, China; zhuhai@xjtu.edu.cn (H.Z.); cnhfwang@zknk.edu.cn (H.W.)

<sup>2</sup> Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fujian University of Technology, Fuzhou 350118, China

\* Correspondence: xxs@fjut.edu.cn

**Abstract:** The ultimate goal of semantic web (SW) is to implement mutual collaborations among ontology-based intelligent systems. To this end, it is necessary to integrate those domain-independent and cross-domain ontologies by finding the correspondences between their entities, which is the so-called ontology matching. To improve the quality of ontology alignment, in this work, the ontology matching problem is first defined as a sparse multi-objective optimization problem (SMOOP), and then, a multi-objective evolutionary algorithm with a relevance matrix (MOEA-RM) is proposed to address it. In particular, a relevance matrix (RM) is presented to adaptively measure the relevance of each individual's genes to the objectives, which is applied in MOEA's initialization, crossover and mutation to ensure the population's sparsity and to speed up the the algorithm's convergence. The experiment verifies the performance of MOEA-RM by comparing it with the state-of-the-art ontology matching techniques, and the experimental results show that MOEA-RM is able to effectively address the ontology matching problem with different heterogeneity characteristics.

**Keywords:** ontology matching; sparse multi-objective optimization problem; multi-objective evolutionary algorithm; relevance matrix

**MSC:** 68T30, 68W50



**Citation:** Zhu, H.; Xue, X.; Wang H. Matching Ontologies through Multi-Objective Evolutionary Algorithm with Relevance Matrix. *Mathematics* **2022**, *10*, 2077. <https://doi.org/10.3390/math10122077>

Academic Editors: Yu Xue, Chunlin He and Ferrante Neri

Received: 25 April 2022

Accepted: 13 June 2022

Published: 15 June 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Ontology provides a standard and shared representation on domain knowledge, which is regarded as the solution for the data heterogeneity problem [1]. With the development of semantic web (SW) [2,3], more and more ontology-based intelligent systems for E-learning [4], personalized search and browsing [5], and collaborative molecular biology [6] have been developed, which require mutual collaborations to enhance their intelligent behaviors. To this end, it is necessary to find the entity mappings between their ontologies, which is the so-called ontology matching [7]. Essentially, matching two ontologies aims to find a mapping matrix (MM), which describes an alignment through setting rows and columns as two ontologies' entities, respectively, and elements as 1 denotes that two corresponding entities are mapped, otherwise not. Table 1 shows an example of MM, where  $O_1$  and  $O_2$  are two ontologies, and  $O_1.Article$  and  $O_2.Paper$  respectively denote the concepts "Article" and "Paper" in  $O_1$  and  $O_2$ . Since the entities' scale could be very large, most of the MM elements are zero (i.e., it is a sparse matrix), and to determine an optimal MM is essentially a sparse optimization problem [8].

**Table 1.** An example of mapping matrix.

	<i>O<sub>2</sub>.Article</i>	<i>O<sub>2</sub>.Person</i>	<i>O<sub>2</sub>.Evaluation</i>	<i>O<sub>2</sub>.Neutral</i>	<i>O<sub>2</sub>.Positive</i>	<i>O<sub>2</sub>.Negative</i>
<i>O<sub>1</sub>.Author</i>	0	1	0	0	0	0
<i>O<sub>1</sub>.Paper</i>	1	0	0	0	0	0
<i>O<sub>1</sub>.Accept</i>	0	0	0	0	1	0
<i>O<sub>1</sub>.Reject</i>	0	0	0	0	0	1
<i>O<sub>1</sub>.Topic</i>	0	0	0	0	0	0
<i>O<sub>1</sub>.Organization</i>	0	0	0	0	0	0

Recently, the evolutionary algorithm (EA) has become a popular method of addressing ontology matching problem [9], but the existing single-objective EA makes use of f-measure [10] to evaluate the MM quality, which yields the bias improvement on the solutions, i.e., the solution might sacrifice one of the objectives to improve the other one. To ensure the unanimous improvements on these two metrics, in this work, the ontology matching problem is defined as a sparse multi-objective optimization problem (SMOOP), and a multi-objective evolutionary algorithm with relevance matrix (MOEA-RM) is proposed to address it. The traditional MOEA performance usually degenerates when addressing SMOOP since the searching space grows exponentially with the increasing number of decision variables [11,12]. To face this challenge, various strategies have been proposed in the past decades, such as decision variable grouping [13], decision variable analysis [14] and special initialization and evolutionary operators [15–17]. However, it is still difficult to maintain the sparsity of the population, and the algorithm always converges slowly. To overcome these drawbacks, MOEA-RM first introduces a relevance matrix (RM) to adaptively measure each gene or correspondence’s relevance to the objective, which is then used to initialize the population to ensure the population’s sparsity, and improve the algorithm’s converging speed as well as the sparsity of generated individuals when executing the crossover and mutation operators. In particular, this work’s contributions are as follows:

- The multi-objective ontology matching problem is formally defined;
- A MOEA-RM is presented to address the ontology matching problem, which uses RM-based initialization, crossover and mutation to adaptively maintain population’s diversity and improve the algorithm’s converging speed;
- The proposed MOEA-RM is employed on 39 different ontology matching tasks, and the experimental results show its effectiveness.

The rest of this paper is organized as follows: Section 2 overviews the existing MOEAs for addressing the ontology matching problem; Section 3 provides the preliminary background knowledge and defines the problem investigated; Section 4 presents the RM-based MOEA, and Section 5 shows the experimental results; finally, Section 6 concludes this work and points out the future work.

## 2. Related Work

Comparing with the popular artificial intelligence techniques, such as neural network [18–21], data mining [22,23], etc., the evolutionary improvement on alignment’s quality is able to better refine their quality. In recent years, various MOEA-based ontology matching techniques have been proposed. To trade off the alignment’s completeness and correctness, Xue et al. propose a NSGA-II [24] and MOEA/D [25] to tune the matching systems’ parameters. These MOEA based matching techniques are able to provide different Pareto solutions for decision makers. Acampora et al. also use NSGA-II to optimize the alignment’s quality [26], and they further make comparisons among different MOEA-based matching techniques’ searching performance results [27]. To face the challenge of a large-scale ontology matching problem, Xue et al. present a general framework of MOEA-based

large-scale matching technique, which first divides two ontologies into several similar segments, and then use MOEA to match them separately. After that, to ensure the diversity of population, they use an adaptive strategy to guide the algorithm's search direction [28]. They also try to reduce the MOEA's computational complexity by using the compact encoding mechanism, which is able to address the ternary compound ontology matching problem [29], where an entity correspondence might consist of more than two entities. To enhance the converging speed, Lv et al. [30] get an expert involved in MOEA's evolving process, and make use of his knowledge to improve the alignment's quality. The interactive MOEA is also used to match the sensor ontology on the internet of things [31].

The existing MOEA-based ontology matching techniques model the ontology matching process as a continuous optimization problem. However, they need to construct several similarity matrices to maintain the candidate entity mappings' similarity values, which requires huge computational complexity. To overcome this drawback, in this work, the ontology matching is defined as a 0–1 integer optimization problem, and considering its characteristics of sparsity, a RM is proposed to adaptively maintain the population's sparsity and guide the algorithm's searching direction.

### 3. Ontology Matching Problem

An ontology consists of the classes that define the domain concepts, the datatype properties that describe the class's feature, and the object properties that present the relationships between classes [32,33]. Different ontologies might define an entity in different ways, yielding the heterogeneity problem. To address this issue, it is necessary to find the heterogeneous entity correspondences in automatic or semi-automatic ways. The found correspondences between the entities are called ontology alignment, where each correspondence mainly consists of two entities, the relationships (typically equivalence  $\equiv$ ) and their similarity value. The similarity value of two entities is an important metric that measures whether they are similar or not, which is typically calculated through the similarity measure [34].

To measure the quality of an alignment, the classic metrics are recall, precision and f-measure [10]. However, these metrics require using the reference alignment, which is not always available in practical matching tasks. To this end, this work uses the approximate metrics, i.e., MatchCoverage and Frequency [35], which respectively estimate an alignment's recall and precision. To be specific, given two ontologies  $O_1$  and  $O_2$ , their alignment  $A$ 's MatchCoverage and Frequency are respectively defined as follows:

$$\text{MatchCoverage}(A) = \frac{|MatchedEntity_{O_1}| + |MatchedEntity_{O_2}|}{|O_1| + |O_2|} \quad (1)$$

$$\text{Frequency}(A) = \frac{2 \times \sum sim_i}{|MatchedEntity_{O_1}| + |MatchedEntity_{O_2}|} \quad (2)$$

where  $|MatchedEntity_{O_1}|$ ,  $|MatchedEntity_{O_2}|$  are respectively the cardinality of the matched entity sets in  $O_1$  and  $O_2$ ,  $|A|$  is the number of correspondences in  $A$ , and  $sim_i$  is the  $i$ -th correspondence's similarity value.

Given an alignment  $A$ , the ontology matching problem is defined in Equation (3):

$$\begin{cases} \max & f(M) = (\text{MatchCoverage}(M), \text{Frequency}(M)) \\ \text{s.t.} & M_{|C_1| \times |C_2|} \\ & m_{ij} \in \{0, 1\}, i = 1, 2, \dots, |C_1|, j = 1, 2, \dots, |C_2| \end{cases} \quad (3)$$

where  $|C_1|$  and  $|C_2|$  are the number of concepts in ontologies  $O_1$  and  $O_2$ , the decision variable is the MM  $M_{|C_1| \times |C_2|}$ , and the two objectives are to maximize the MM corresponding alignment's MatchCoverage and Frequency.

#### 4. Multi-Objective Evolutionary Algorithm with Relevance Matrix

Before executing the matching process, the entities, such as class, datatype property and object property, are extracted from two ontologies. Their names are pre-processed through tokenization and stemming, which are then used to construct the entity similarity matrix for evaluating the individual's fitness value. In particular, the similarity matrix's rows and columns are respectively two ontologies' entities, and the elements are the corresponding entities' similarity value. After that, the ontology alignment is optimized through MOEA-RM, whose framework is presented in Algorithm 1.

---

**Algorithm 1** The framework of multi-objective evolutionary algorithm with relevance matrix

---

```

Initialization( $P$ );
Non-dominated_Sort( $P$ );
while Terminating condition is not met do
    Update( $RM$ );
    Generate  $P'$  via  $RM$  based operators;
    Non-dominated_Sort( $P \cup P'$ );
    Select_Next_Generation( $P \cup P'$ );
    Non-dominated_Sort( $P$ );
end while

```

---

The MOEA-RM framework is similar to that of NSGA-II, while the novelties lie in the RM maintenance in each generation and RM-based evolutionary operators. This work uses MM to encode an individual, which is described in Table 1. A RM is a statistical matrix with exactly the same number of rows and columns as MM, which reflects the current generation's Pareto front (PF) gene distribution. In each generation, we sum all the PF solution's corresponding MMs to obtain RM, which is defined as follows:

$$RM_{i,j} = \sum_k MM_{i,j}^k \quad (4)$$

where  $MM_{i,j}^k$  is the  $k$ -th MM's  $i$ -th row and  $j$ -th column element's value. The higher value of a RM's element means more gene bits in the current population's PF solutions have chosen corresponding entity pairs, which can be utilized to guide the algorithm in either speeding up the converging speed or enhancing the population's diversity. In the following, we respectively describe RM-based initialization and RM-based evolutionary operators.

##### 4.1. Initialization

Given two ontologies  $O_1$  and  $O_2$ , the population  $P$  with size  $N$ , a RM  $RM_{|O_1| \times |O_2|}$ , an individual  $ind$  is initialized according to Algorithm 2.

Here, we first randomly initialize the population with uniform probability, and then all MMs are summed up to obtain a RM for initialization. After that, for each individual's element, we randomly pick up one other element with the same value in this MM to compare the corresponding RM values. If the former's RM value is bigger, i.e., more individuals have chosen this gene bit as 1, we change the MM value from 1 to 0; otherwise, the value is changed from 0 to 1.

**Algorithm 2** Initialization

---

```

for  $k = 0; k < N; k++$  do
  for  $i = 0; i < |O_1|; i++$  do
    for  $j = 0; j < |O_2|; j++$  do
       $MM_{i,j}^k = \text{rand}\{0,1\};$ 
    end for
  end for
  end for
  update RM with  $P$ ;
  for  $k = 0; k < N; k++$  do
    for  $i = 0; i < |O_1|; i++$  do
      for  $j = 0; j < |O_2|; j++$  do
        if  $MM_{i,j}^k == 1$  then
          Randomly select  $MM_{i',j'}^k$  with value 1;
          if  $RM_{i,j} > RM_{i',j'}$  then
             $RM_{i,j} = 0;$ 
          end if
        end if
        if  $MM_{i,j}^k == 0$  then
          Randomly select  $MM_{i',j'}^k$  with value 0;
          if  $RM_{i,j} < RM_{i',j'}$  then
             $RM_{i,j} = 1;$ 
          end if
        end if
      end for
    end for
  end for
end for

```

---

**4.2. Relevance Matrix Based Evolutionary Operator**

We use RM to execute the crossover and mutation to adaptively trade off the algorithm's convergence and divergence, whose pseudo-codes are shown in Algorithm 3.

During the crossover, the offspring individual  $z$  is set the same as one of its parent solutions, assuming  $p$ . Then, for each different gene in  $p$  and  $q$ , we need to first decide whether the current evolutionary strategy is focused on convergence or diversity by the parameter  $\theta$ . Assuming it is the former, and we compare the RM corresponding element  $RM_{i,j}$  with randomly selected element  $RM_{i',j'}$ . If  $RM_{i,j}$  is larger, we set  $z_{i,j}$  as 1, otherwise as 0. When it is the latter, if  $RM_{i,j}$  is larger, we set  $z_{i,j}$  as 0, otherwise as 1. With respect to the mutation, we first judge whether this gene bit should execute the mutation according to the mutation rate  $p_m$ . If it is so, we need to further judge whether the current evolutionary strategy should prefer convergence or diversity by the parameter  $\theta$ . The rest of the operations are similar to those in the crossover.

In particular, the parameter  $\theta \in [0, 1]$  controls the algorithm's preference on convergence; we update it in each generation and adaptively trade off the algorithm's exploration and exploitation. In the early stage,  $\theta$  should be small to ensure the population's diversity, while in the late stage, it should be large to speed up the algorithm's convergence. On this basis, given current generation  $Gen$  and the maximum generation  $MaxGen$ ,  $\theta$  is updated as follows:

$$\theta_{Gen} = \frac{Gen}{MaxGen} \quad (5)$$

**Algorithm 3** Relevance matrix-based crossover and mutation

---

\*\*\*\*\* Crossover \*\*\*\*\*

```

[p, q] = randomly select two parents from the population;
z = p; //Initialize the offspring individual z;
for i = 0; i < p.|O1|; i++ do
  for j = 0; j < p.|O2|; j++ do
    if pi,j ≠ qi,j then
      if rand(0,1) < θ then
        //Enhance the convergence
        if RMi,j > RMi',j' then
          zi,j = 1;
        else
          zi,j = 0;
        end if
      else
        //Enhance the Diversity
        if RMi,j > RMi',j' then
          zi,j = 0;
        else
          zi,j = 1;
        end if
      end if
    end if
  end for
end for

```

\*\*\*\*\* Mutation \*\*\*\*\*

```

for k = 0; k < N; k++ do
  for i = 0; i < |O1|; i++ do
    for j = 0; j < |O2|; j++ do
      if rand(0,1) < pm then
        if rand(0,1) < θ then
          //Enhance the convergence
          if RMi,j > RMi',j' then
            MMi,jp = 1;
          else
            MMi,jp = 0;
          end if
        else
          //Enhance the diversity
          if RMi,j > RMi',j' then
            MMi,jp = 0;
          else
            MMi,jp = 1;
          end if
        end if
      end if
    end for
  end for
end for

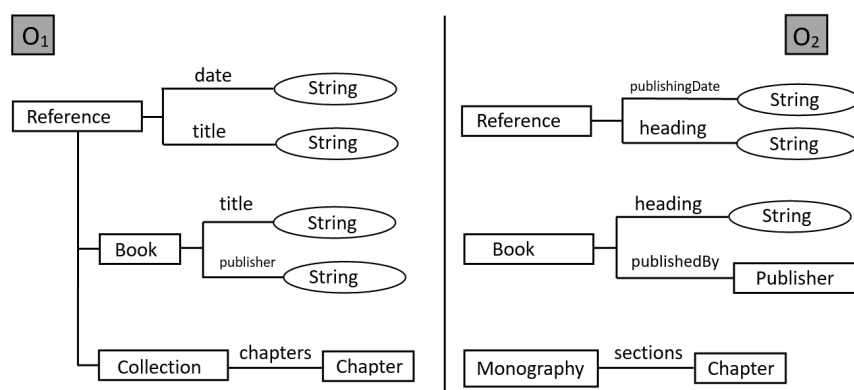
```

---

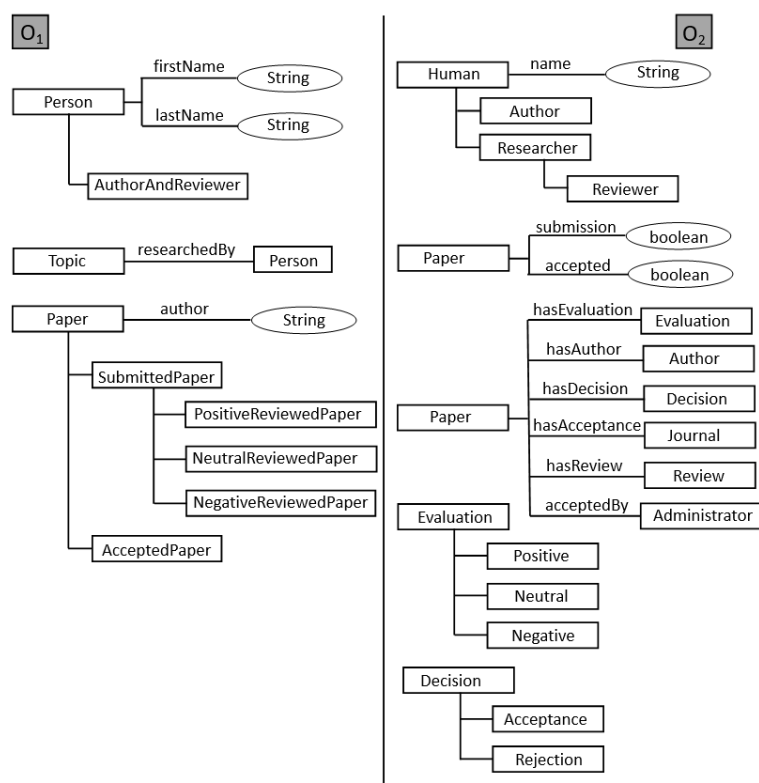
**5. Experiment****5.1. Experimental Setup**

We use the Ontology Alignment Evaluation Initiative (OAEI) benchmark track (<http://oaei.ontologymatching.org/2008/benchmark> (accessed on 14 June 2022)) and conference track (<http://oaei.ontologymatching.org/2020/conference> (accessed on 14 June 2022)) to

test the MOEA-RM performance on schema-level matching tasks. The benchmark track requires matching two bibliographic ontologies, and the target ontology's entity names could be random strings or synonyms; the hierarchy could be expanded or flattened; the properties could be suppressed; and the classes could be refined by several sub-classes or flattened. The conference track requires matching 16 different ontologies on the conference organization, which have been used in some actual conference series and the corresponding conference web sites. Figures 1 and 2 show the examples of testing cases from the OAEI benchmark track and conference track, respectively.



**Figure 1.** A segment of testing case from OAEI benchmark track.



**Figure 2.** A segment of testing case from OAEI conference track.

We compare MOEA-RM with EA [36], NSGA-II [26] and OAEI participants [37]. We use the profile-based similarity measure [38] to distinguish heterogeneous entities, and MOEA-RM's parameter setting is as follows: the population size  $N = 80$ , the maximum generation  $MaxGen = 2000$  and the mutation rate  $p_m = 0.02$ , and the EA and NSGA-II configurations are referred to in their literature. The EA, NSGA-II and MOEA-RM results are the average of 30 independent runs, and the OAEI participants' results are from the



OAEI official web site (<http://oaei.ontologymatching.org> (accessed on 14 June 2022)). In this work, we select the solution with the best f-measure from PF as the algorithm's output, and Tables 2 and 3 respectively compare MOEA-RM with EA, NSGA-II and OAEI participants on the OAEI benchmark and conference tracks in terms of recall, precision and f-measure.

**Table 2.** Comparison in terms of recall and precision. The symbols  $r$  and  $p$  respectively denote recall and precision.

OAEI's Benchmark Track							
Testing Case	AMLC [39]	LogMap [40]	LogMapLt [40]	XMap [41]	EA	NSGA-II	MOEA-RM
	$r(p)$	$r(p)$	$r(p)$	$r(p)$	$r(p)$	$r(p)$	$r(p)$
101	1.00 (1.00)	0.88 (0.96)	0.78 (0.64)	0.93 (1.00)	0.78 (0.84)	1.00 (1.00)	1.00 (1.00)
202	0.80 (0.92)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.72 (0.87)	0.80 (0.91)	0.95 (0.95)
221	0.49 (0.53)	0.87 (0.98)	0.76 (0.69)	0.95 (1.00)	0.87 (0.87)	0.97 (0.92)	1.00 (1.00)
222	0.71 (0.32)	0.00 (0.00)	0.76 (0.69)	0.80 (0.75)	0.78 (0.85)	0.97 (0.92)	1.00 (1.00)
223	0.40 (0.62)	0.90 (0.98)	0.76 (0.69)	0.98 (0.96)	0.87 (0.87)	0.86 (0.95)	1.00 (1.00)
224	0.58 (0.45)	0.90 (0.98)	0.82 (0.98)	0.98 (0.96)	0.94 (0.85)	0.86 (0.95)	1.00 (1.00)
225	0.51 (0.52)	0.92 (0.97)	0.76 (0.69)	0.98 (0.96)	0.78 (0.85)	0.82 (0.87)	1.00 (1.00)
228	1.00 (1.00)	0.92 (0.97)	0.58 (0.40)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
232	0.51 (0.52)	0.87 (0.98)	0.88 (0.93)	0.98 (0.96)	0.81 (0.95)	0.86 (0.95)	1.00 (1.00)
233	1.00 (1.00)	0.92 (0.97)	0.58 (0.40)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
236	1.00 (1.00)	0.92 (0.97)	0.72 (0.87)	1.00 (1.00)	0.82 (0.88)	0.92 (0.92)	1.00 (1.00)
237	0.42 (0.58)	0.00 (0.00)	0.88 (0.93)	0.80 (0.75)	0.87 (0.80)	0.82 (0.87)	0.93 (0.98)
238	0.51 (0.52)	0.96 (0.93)	0.88 (0.93)	0.98 (0.96)	0.82 (0.92)	0.86 (0.95)	1.00 (1.00)
239	1.00 (1.00)	0.91 (0.93)	0.58 (0.40)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
240	1.00 (1.00)	0.91 (0.93)	0.58 (0.40)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
241	1.00 (1.00)	0.91 (0.93)	0.72 (0.87)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
246	1.00 (1.00)	0.88 (0.96)	0.72 (0.87)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
247	1.00 (1.00)	0.88 (0.96)	0.72 (0.87)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
average	0.77 (0.77)	0.75 (0.80)	0.69 (0.68)	0.91 (0.90)	0.89 (0.91)	0.93 (0.95)	0.99 (0.99)
OAEI's Conference Track							
Testing Case	AMLC	LogMap	LogMapLt	XMap	EA	NSGA-II	MOEA-RM
	$r(p)$	$r(p)$	$r(p)$	$r(p)$	$r(p)$	$r(p)$	$r(p)$
cmt-conference	0.53 (0.67)	0.53 (0.73)	0.33 (0.56)	0.00 (0.00)	0.68 (0.76)	0.68 (0.76)	0.83 (0.86)
cmt-confOf	0.56 (0.90)	0.31 (0.83)	0.38 (0.67)	0.44 (0.88)	0.65 (0.68)	0.65 (0.68)	0.75 (0.81)
cmt-edas	0.77 (0.91)	0.62 (0.89)	0.62 (0.73)	0.69 (0.75)	0.65 (0.72)	0.68 (0.76)	0.80 (0.86)
cmt-ekaw	0.55 (0.75)	0.55 (0.75)	0.45 (0.56)	0.64 (0.70)	0.65 (0.68)	0.65 (0.68)	0.82 (0.89)
cmt-iasted	1.00 (0.80)	0.84 (0.80)	0.90 (0.89)	0.93 (0.80)	0.75 (0.89)	0.83 (0.87)	0.91 (0.94)
cmt-sigkdd	0.92 (0.92)	0.88 (0.95)	0.67 (0.89)	0.83 (0.91)	0.75 (0.75)	0.87 (0.90)	0.95 (0.95)
conference-confOf	0.87 (0.87)	0.73 (0.85)	0.60 (0.90)	0.80 (0.71)	0.78 (0.67)	0.80 (0.88)	0.88 (0.87)
conference-edas	0.65 (0.73)	0.65 (0.85)	0.53 (0.75)	0.65 (0.79)	0.78 (0.67)	0.68 (0.78)	0.86 (0.85)
conference-ekaw	0.72 (0.78)	0.48 (0.60)	0.32 (0.62)	0.60 (0.58)	0.74 (0.66)	0.70 (0.78)	0.84 (0.85)
conference-iasted	0.36 (0.83)	0.50 (0.88)	0.29 (0.80)	0.36 (0.62)	0.68 (0.52)	0.75 (0.60)	0.75 (0.74)
conference-sigkdd	0.73 (0.85)	0.73 (0.85)	0.53 (0.80)	0.60 (0.58)	0.75 (0.75)	0.75 (0.75)	0.92 (0.85)
confOf-edas	0.58 (0.92)	0.53 (0.77)	0.58 (0.58)	0.53 (0.91)	0.65 (0.72)	0.65 (0.72)	0.71 (0.76)
confOf-ekaw	0.80 (0.94)	0.70 (0.93)	0.50 (0.77)	0.80 (0.76)	0.88 (0.75)	0.88 (0.75)	0.85 (0.90)
confOf-iasted	0.44 (0.80)	0.54 (0.89)	0.54 (0.90)	0.67 (0.43)	0.62 (0.51)	0.69 (0.51)	0.72 (0.78)
confOf-sigkdd	0.88 (0.95)	0.81 (0.90)	0.68 (0.88)	0.57 (0.80)	0.88 (0.73)	0.89 (0.78)	0.93 (0.95)
edas-ekaw	0.48 (0.79)	0.52 (0.75)	0.43 (0.59)	0.52 (0.75)	0.65 (0.68)	0.65 (0.68)	0.84 (0.75)
edas-iasted	0.47 (0.82)	0.37 (0.88)	0.37 (0.88)	0.42 (0.57)	0.63 (0.57)	0.62 (0.82)	0.70 (0.85)
edas-sigkdd	0.75 (0.84)	0.47 (0.88)	0.47 (0.88)	0.62 (0.81)	0.75 (0.75)	0.68 (0.76)	0.78 (0.82)
ekaw-iasted	0.70 (0.84)	0.70 (0.78)	0.60 (0.60)	0.70 (0.58)	0.68 (0.76)	0.82 (0.74)	0.80 (0.80)
ekaw-sigkdd	0.73 (0.80)	0.70 (0.78)	0.70 (0.78)	0.64 (0.78)	0.78 (0.67)	0.70 (0.81)	0.75 (0.82)
iasted-sigkdd	0.87 (0.81)	0.88 (0.82)	0.73 (0.73)	0.87 (0.68)	0.76 (0.75)	0.80 (0.85)	0.85 (0.86)
average	0.68 (0.83)	0.62 (0.82)	0.53 (0.75)	0.61 (0.68)	0.70 (0.69)	0.73 (0.55)	0.82 (0.84)



**Table 3.** Comparison in terms of f-measure.

OAEI's Benchmark Track							
Testing Case	AMLC	LogMap	LogMapLt	XMap	EA	NSGA-II	MOEA-RM
101	1.00	0.95	0.71	0.97	0.81	1.00	1.00
202	0.86	0.00	0.00	0.00	0.79	0.85	0.95
221	0.51	0.94	0.72	0.97	0.87	0.95	1.00
222	0.50	0.00	0.72	0.78	0.82	0.95	1.00
223	0.51	0.94	0.72	0.97	0.87	0.90	1.00
224	0.51	0.94	0.90	0.97	0.90	0.90	1.00
225	0.51	0.95	0.72	0.97	0.82	0.85	1.00
228	1.00	0.92	0.48	1.00	1.00	1.00	1.00
232	0.51	0.94	0.90	0.97	0.88	0.90	1.00
233	1.00	0.92	0.48	1.00	1.00	1.00	1.00
236	1.00	0.92	0.80	1.00	0.85	0.92	1.00
237	0.50	0.00	0.91	0.78	0.84	0.85	0.95
238	0.51	0.95	0.90	0.97	0.87	0.90	1.00
239	1.00	0.92	0.48	1.00	1.00	1.00	1.00
240	1.00	0.92	0.48	1.00	1.00	1.00	1.00
241	1.00	0.92	0.80	1.00	1.00	1.00	1.00
246	1.00	0.92	0.80	1.00	1.00	1.00	1.00
247	1.00	0.92	0.80	1.00	1.00	1.00	1.00
average	0.77	0.78	0.68	0.91	0.91	0.94	0.99
OAEI's Conference Track							
Testing Case	AMLC	LogMap	LogMapLt	XMap	EA	NSGA-II	MOEA-RM
cmt-conference	0.59	0.62	0.42	0.00	0.72	0.72	0.84
cmt-confOf	0.69	0.45	0.48	0.58	0.66	0.66	0.78
cmt-edas	0.83	0.73	0.67	0.72	0.68	0.72	0.83
cmt-ekaw	0.63	0.63	0.50	0.67	0.68	0.72	0.85
cmt-iasted	0.89	0.89	0.89	0.89	0.82	0.85	0.92
cmt-sigkdd	0.92	0.91	0.76	0.87	0.75	0.89	0.95
conference-confOf	0.87	0.79	0.72	0.75	0.78	0.83	0.87
conference-edas	0.69	0.73	0.62	0.71	0.73	0.73	0.82
conference-ekaw	0.75	0.53	0.42	0.59	0.70	0.74	0.86
conference-iasted	0.50	0.64	0.42	0.45	0.59	0.68	0.74
conference-sigkdd	0.79	0.79	0.64	0.69	0.75	0.75	0.83
confOf-edas	0.71	0.62	0.58	0.67	0.68	0.68	0.79
confOf-ekaw	0.86	0.80	0.61	0.78	0.82	0.82	0.88
confOf-iasted	0.57	0.62	0.62	0.52	0.58	0.60	0.75
confOf-sigkdd	0.92	0.83	0.73	0.67	0.80	0.85	0.94
edas-ekaw	0.59	0.62	0.50	0.62	0.66	0.62	0.70
edas-iasted	0.60	0.52	0.52	0.48	0.60	0.66	0.77
edas-sigkdd	0.80	0.61	0.61	0.64	0.75	0.72	0.80
ekaw-iasted	0.78	0.74	0.60	0.64	0.72	0.78	0.80
ekaw-sigkdd	0.76	0.74	0.74	0.70	0.73	0.76	0.78
iasted-sigkdd	0.84	0.85	0.73	0.76	0.76	0.82	0.85
average	0.74	0.70	0.61	0.64	0.71	0.74	0.83

### 5.2. Experimental Results

As shown in Tables 2 and 3, MOEA-based matching technique's results are generally better than single-objective EA since MOEA can overcome the negative impact brought by the bias improvement on solutions. With the introduction of RM, MOEA-RM is able to effectively search for more potential feasible regions, which is of help to overcome the NSGA-II premature convergence. Thus, MOEA-RM outperforms NSGA-II on all testing cases. Since MOEA-RM takes into consideration more similarity measures to distinguish the heterogeneous entities, its precision value is generally high, and with the help of its powerful searching ability, the MOEA-RM recall values are also high on all testing cases.

Comparing with the state-of-the-art ontology matching systems (AMLC aggregates various matchers' alignments through a pre-defined framework, LogMap uses a reasoning-based matching strategy to improve the alignment's quality, and XMap utilizes the structural approach to determine the final alignment), the MOEA-RM average recall, precision and f-measure values are the highest, which shows that the mechanism of iteratively refining the alignment is able to determine high-quality ontology alignments. To conclude, the MOEA-RM-based ontology matching technique is able to effectively match ontologies with different heterogeneous characteristics.

## 6. Conclusions and Future Work

Matching ontologies is critical to SW development, and to determine the high-quality ontology alignment, this work models the ontology matching problem as a SMOOP, and proposes a MOEA-RM to adaptively measure each gene or correspondence's relevance to the objectives to effectively address it. To maintain the population's diversity and overcome the algorithm's premature convergence, MOEA-RM uses RM to adaptively measure each gene or correspondence's relevance to the objective, which is then used to initialize the population to ensure the population's sparsity, and guide the crossover and mutation operators. The experiment uses the OAEI benchmark and conference tracks to test the MOEA-RM performance, and the experimental results show that our approach is able to effectively match the ontologies with various heterogeneous characteristics.

In the future, we are interested in using MOEA-RM to determine m:n correspondences, which is more challenging in terms of complicate semantics and high computational cost. Additionally, the similarity measure should be improved to distinguish the complex style of correspondences, and the corresponding semantic reasoning techniques could be of help to achieve high precision value. Last but not least, the evolutionary operators could also be improved to enhance the algorithm's efficiency.

**Author Contributions:** Conceptualization, H.Z. and X.X.; methodology, X.X. and H.W.; software, X.X.; validation, H.Z. and H.W.; formal analysis, H.Z.; investigation, H.Z.; resources, X.X.; data curation, X.X.; writing—original draft preparation, X.X.; writing—review and editing, H.Z.; funding acquisition, H.Z. and X.X. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the National Natural Science Foundation of China (Nos. 62172457 and 62172095), the Natural Science Foundation of Fujian Province (Nos. 2020J01875 and 2019J01771) and the Scientific Research Foundation of Fujian University of Technology (No. GY-Z17162).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Guarino, N.; Oberle, D.; Staab, S. What is an ontology? In *Handbook on Ontologies*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 1–17.
2. Berners-Lee, T.; Hendler, J.; Lassila, O. The semantic web. *Sci. Am.* **2001**, *284*, 34–43. [[CrossRef](#)]
3. Cardoso, J.; Sheth, A. The Semantic Web and its applications. In *Semantic Web Services, Processes and Applications*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 3–33.
4. Shishehchi, S.; Banihashem, S.Y.; Zin, N.A.M. A proposed semantic recommendation system for e-learning: A rule and ontology based e-learning recommendation system. In *Proceedings of the 2010 International Symposium on Information Technology*, Kuala Lumpur, Malaysia, 15–17 June 2010; Volume 1, pp. 1–5.
5. Gauch, S.; Chaffee, J.; Pretschner, A. Ontology-based personalized search and browsing. *Web Intell. Agent Syst. Int. J.* **2003**, *1*, 219–234.
6. Altman, R.B.; Buda, M.; Chai, X.J.; Carillo, M.W.; Chen, R.O.; Abernethy, N.F. RiboWeb: An ontology-based system for collaborative molecular biology. *IEEE Intell. Syst. Their Appl.* **1999**, *14*, 68–76. [[CrossRef](#)]

7. Shvaiko, P.; Euzenat, J. Ontology matching: State of the art and future challenges. *IEEE Trans. Knowl. Data Eng.* **2011**, *25*, 158–176. [\[CrossRef\]](#)
8. Xue, X.; Liu, W. Integrating Heterogeneous Ontologies in Asian Languages through Compact Genetic Algorithm with Annealing Re-sample Inheritance Mechanism. *Trans. Asian Low-Resour. Lang. Inf. Process.* **2022**, *2022*, 1–17. [\[CrossRef\]](#)
9. Naya, J.M.V.; Romero, M.M.; Loureiro, J.P.; Munteanu, C.R.; Sierra, A.P. Improving ontology alignment through genetic algorithms. In *Soft Computing Methods for Practical Environment Solutions: Techniques and Studies*; IGI Global: Hershey, PA, USA, 2010; pp. 240–259.
10. Rijsberge, C.J.V. *Information Retrieval*; University of Glasgow: Glasgow, UK, 1975.
11. Tan, Z.; Wang, H.; Liu, S. Multi-stage dimension reduction for expensive sparse multi-objective optimization problems. *Neurocomputing* **2021**, *440*, 159–174. [\[CrossRef\]](#)
12. Xue, Y.; Wang, Y.; Liang, J.; Slowik, A. A self-adaptive mutation neural architecture search algorithm based on blocks. *IEEE Comput. Intell. Mag.* **2021**, *16*, 67–78. [\[CrossRef\]](#)
13. Ma, X.; Liu, F.; Qi, Y.; Wang, X.; Li, L.; Jiao, L.; Yin, M.; Gong, M. A multiobjective evolutionary algorithm based on decision variable analyses for multiobjective optimization problems with large-scale variables. *IEEE Trans. Evol. Comput.* **2015**, *20*, 275–298. [\[CrossRef\]](#)
14. Zille, H.; Ishibuchi, H.; Mostaghim, S.; Nojima, Y. A framework for large-scale multiobjective optimization based on problem transformation. *IEEE Trans. Evol. Comput.* **2017**, *22*, 260–275. [\[CrossRef\]](#)
15. Tian, Y.; Zhang, X.; Wang, C.; Jin, Y. An evolutionary algorithm for large-scale sparse multiobjective optimization problems. *IEEE Trans. Evol. Comput.* **2019**, *24*, 380–393. [\[CrossRef\]](#)
16. Zhang, Y.; Tian, Y.; Zhang, X. Improved SparseEA for sparse large-scale multi-objective optimization problems. *Complex Intell. Syst.* **2021**, 1–16. [\[CrossRef\]](#)
17. Xue, Y.; Zhu, H.; Liang, J.; Slowik, A. Adaptive crossover operator based multi-objective binary genetic algorithm for feature selection in classification. *Knowl.-Based Syst.* **2021**, *227*, 107218. [\[CrossRef\]](#)
18. Bento, A.; Zouaq, A.; Gagnon, M. Ontology matching using convolutional neural networks. In Proceedings of the 12th Language Resources and Evaluation Conference, Marseille, France, 11–16 May 2020; pp. 5648–5653.
19. Khoudja, M.A.; Fareh, M.; Bouarfa, H. Ontology matching using neural networks: Survey and analysis. In Proceedings of the 2018 International Conference on Applied Smart Systems (ICASS), Medea, Algeria, 24–25 November 2018; pp. 1–6.
20. Zhang, Y.; Wang, X.; Lai, S.; He, S.; Liu, K.; Zhao, J.; Lv, X. Ontology matching with word embeddings. In *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data*; Springer: Cham, Switzerland, 2014; pp. 34–45.
21. Xue, X.; Huang, Q. Generative adversarial learning for optimizing ontology alignment. In *Expert Systems*; Wiley: Hoboken, NJ, USA, 2022; pp. 1–12.
22. Belhadi, H.; Akli-Astouati, K.; Djenouri, Y.; Lin, J.C.W. Data mining-based approach for ontology matching problem. *Appl. Intell.* **2020**, *50*, 1204–1221. [\[CrossRef\]](#)
23. David, J. Association rule ontology matching approach. *Int. J. Semant. Web Inf. Syst.* **2007**, *3*, 27–49. [\[CrossRef\]](#)
24. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [\[CrossRef\]](#)
25. Zhang, Q.; Li, H. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evol. Comput.* **2007**, *11*, 712–731. [\[CrossRef\]](#)
26. Acampora, G.; Kaymak, U.; Loia, V.; Vitiello, A. Applying NSGA-II for solving the ontology alignment problem. In Proceedings of the 2013 IEEE International Conference on Systems, Man, and Cybernetics, Manchester, UK, 13–16 October 2013; pp. 1098–1103.
27. Acampora, G.; Ishibuchi, H.; Vitiello, A. A comparison of multi-objective evolutionary algorithms for the ontology meta-matching problem. In Proceedings of the 2014 IEEE Congress on Evolutionary Computation (CEC), Beijing, China, 6–11 July 2014; pp. 413–420.
28. Xue, X.; Tsai, P.W.; Zhuang, Y. Matching Biomedical Ontologies through Adaptive Multi-Modal Multi-Objective Evolutionary Algorithm. *Biology* **2021**, *10*, 1287. [\[CrossRef\]](#)
29. Xue, X.; Lu, J.; Chen, J. Ternary Compound Matching of Biomedical Ontologies with Compact Multi-Objective Evolutionary Algorithm Based on Adaptive Objective Space Decomposition. In Proceedings of the 2020 16th International Conference on Computational Intelligence and Security (CIS), Nanning, China, 27–30 November 2020; pp. 121–125.
30. Lv, Q.; Jiang, C.; Li, H. An interactive multi-objective ontology matching technique. In Proceedings of the International Conference on Advanced Machine Learning Technologies and Applications, Cairo, Egypt, 20–22 March 2021; pp. 955–964.
31. Xue, X.; Chen, J. A preference-based multi-objective evolutionary algorithm for semiautomatic sensor ontology matching. *Int. J. Swarm Intell. Res.* **2018**, *9*, 1–14. [\[CrossRef\]](#)
32. Acampora, G.; Loia, V.; Vitiello, A. Enhancing ontology alignment through a memetic aggregation of similarity measures. *Inf. Sci.* **2013**, *250*, 1–20. [\[CrossRef\]](#)
33. Xue, X.; Wang, Y. Optimizing ontology alignments through a Memetic Algorithm using both MatchFmeasure and Unanimous Improvement Ratio. *Artif. Intell.* **2015**, *223*, 65–81. [\[CrossRef\]](#)
34. Jean-Mary, Y.R.; Shironoshita, E.P.; Kabuka, M.R. Ontology matching with semantic verification. *J. Web Semant.* **2009**, *7*, 235–251. [\[CrossRef\]](#) [\[PubMed\]](#)

35. Xue, X.; Wang, Y. Using memetic algorithm for instance coreference resolution. *IEEE Trans. Knowl. Data Eng.* **2015**, *28*, 580–591. [[CrossRef](#)]
36. Wang, J.; Ding, Z.; Jiang, C. Gaom: Genetic algorithm based ontology matching. In Proceedings of the 2006 IEEE Asia-Pacific Conference on Services Computing (APSCC'06), Guangzhou, China, 12–15 December 2006; pp. 617–620.
37. Caraciolo, C.; Euzenat, J.; Hollink, L.; Ichise, R.; Isaac, A.; Malaisé, V.; Meilicke, C.; Pane, J.; Shvaiko, P.; Stuckenschmidt, H.; et al. Results of the ontology alignment evaluation initiative 2008. In Proceedings of the 3rd ISWC Workshop on Ontology Matching (OM), Karlsruhe, Germany, 26 October 2008; pp. 73–119.
38. Xue, X.; Chen, J. Matching Biomedical Ontologies through Compact Differential Evolution Algorithm with Compact Adaption Schemes on Control Parameters. *Neurocomputing* **2021**, *458*, 526–534. [[CrossRef](#)]
39. Lima, B.; Faria, D.; Couto, F.M.; Cruz, I.F.; Pesquita, C. OAEI 2020 Results for AML and AMLC. In Proceedings of the Fifteenth International Workshop on Ontology Matching, Athens, Greece, 2 November 2020; pp. 154–160.
40. Jiménez-Ruiz, E.; Grau, B.C.; Cross, V. LogMap family participation in the OAEI 2017. *Ceur Workshop Proc.* **2017**, *2032*, 1–5.
41. Djeddi, W.E.; Khadir, M.T.; Yahia, S.B. XMap: Results for OAEI 2015. In Proceedings of the Tenth International Workshop on Ontology Matching, Bethlehem, PA, USA, 12 October 2015; pp. 216–221.