Annals of Emerging Technologies in Computing (AETiC)
Vol. 9, No. 3, 2025

Research Article

CNPMap: A Novel Approach for Ontology Alignment Capturing Beyond-Neighbourhoods Semantic Similarities

Abderrahmane Messous[*] and Fatiha Barigou

Computer Science Laboratory of Oran, University of Oran 1, Algeria
messous.abderrahmane@edu.univ-oran1.dz; barigou.fatiha@univ-oran1.dz
*Correspondence: messous.abderrahmane@edu.univ-oran1.dz

Received: 13th September 2024; Accepted: 29th June 2025; Published: 1st July 2025

Abstract: Ensuring semantic interoperability between heterogeneous systems remains a challenging task due to the structural complexity and diversity of ontological representations. Traditional ontology alignment methods often focus on local features, overlooking important semantic relationships beyond direct neighbourhoods. Here, we introduce CNPMap, a novel alignment approach that addresses this limitation by capturing non-local semantic similarities using a critical node-based partitioning strategy. CNPMap operates in three stages. First, it generates an initial alignment using a hybrid linguistic similarity measure. Then, a graph-based partitioning method exploits the Critical Node Detection Problem (CNDP) to divide ontologies into semantically coherent components. Finally, a context-aware similarity enhancement phase refines the alignments using a sigmoid function that modulates similarities based on both partition-level and entity-level relationships. We evaluated CNPMap on the OAEI 2023 Conference track. The approach improved the F-measure on several ontology pairs by 3% to 6% compared to baseline lexical matchers. For instance, the F-measure increased from 0.69 to 0.74 on the cmt–conference pair and from 0.76 to 0.82 on the cmt–sigkdd pair. CNPMap also achieved a precision of 0.75, outperforming most participating systems. However, its recall was slightly lower due to the conservative threshold used during the initial alignment phase. Our study reveals that integrating partition-based context into similarity computation significantly improves alignment quality, especially for complex ontologies. Future enhancements will focus on improving recall through adaptive thresholds and learning-based parameter tuning.

Keywords: ***CNPMap; Critical Nodes Detection Problem; Graph matching; Ontology matching; Ontology partitioning; Similarity enhancement***

---

1. Introduction

Ontologies serve as formal and explicit representations of knowledge within specific domains, capturing concepts and the relationships among them [1]. They are essential for data representation, integration, and interoperability, facilitating seamless information exchange across heterogeneous systems and applications. However, as ontologies continue to grow in complexity and volume, achieving accurate semantic interoperability through ontology alignment has become an increasingly challenging task.

Ontology matching, defined as the process of identifying correspondences between entities in different ontologies, is a core function in the Semantic Web and knowledge integration frameworks. It enables the reuse and combination of distributed knowledge sources and is critical in domains such as e-commerce, healthcare, and biomedicine [2-3]. Numerous systems have been developed to perform ontology alignment, often relying on combinations of linguistic and structural techniques [4].

Linguistic methods focus on comparing the textual representations of entities (e.g., labels, comments, or descriptions). While they are effective in identifying lexical similarities, these methods

Abderrahmane Messous and Fatiha Barigou, "CNPMap: A Novel Approach for Ontology Alignment Capturing Beyond-Neighbourhoods Semantic Similarities", *Annals of Emerging Technologies in Computing (AETiC)*, Print ISSN: 2516-0281, Online ISSN: 2516-029X, pp. 69-84, Vol. 9, No. 3, 1st July 2025, Published by International Association for Educators and Researchers (IAER), DOI: 10.33166/AETiC.2025.03.005, Available: http://aetic.theiaer.org/archive/v9/v9n3/p5.html.

often fail to uncover deeper semantic correspondences. Structural approaches, by contrast, analyse the graph-based representation of ontologies to identify relationships between entities, such as hierarchical, sibling, or neighbourhoods' relations. Hybrid techniques that combine both approaches have proven more effective overall. However, most existing systems limit structural similarity computations to an entity's direct neighbourhoods, which can miss valuable non-local information. On the other hand, approaches that incorporate the entire ontology structure tend to suffer from increased computational overhead and may introduce noise from irrelevant or weakly related entities.

In this work, we propose CNPMap, a novel ontology alignment approach designed to overcome these limitations by capturing semantic similarities beyond direct neighbourhoods. CNPMap is composed of three main phases: (i) an initial alignment phase based on a hybrid linguistic similarity measure; (ii) an ontology partitioning strategy that leverages the Critical Node Detection Problem (CNDP) [5-7] to decompose the ontology into semantically coherent subgraphs; and (iii) a similarity refinement phase that combines entity-level and partition-level similarities through a parameterized sigmoid function.

A key innovation of CNPMap lies in its context-aware partitioning algorithm, which directly reuses the generated partitions to compute final similarities, eliminating the need for inter-partition correspondence computation. This not only simplifies the alignment process but also improves its semantic precision. Furthermore, the algorithm includes a strategy for handling isolated nodes, which are not initially integrated into any partition, thereby ensuring a more complete and consistent alignment.

We evaluate the effectiveness of our approach using the OAEI 2023 Conference track, a benchmark featuring real-world ontologies with rich structural and terminological diversity. The results show that CNPMap significantly improves alignment quality across several ontology pairs, achieving competitive F-measure scores while maintaining high precision.

The main contributions of this work are summarized as follows:

- A hybrid linguistic similarity measure combining syntactic and semantic techniques for initial matching;
- A graph partitioning algorithm based on CNDP tailored to produce balanced, semantically coherent ontology partitions;
- An effective treatment of isolated nodes during partitioning to enhance alignment coverage;
- A context-aware similarity refinement mechanism that integrates both entity and partition-level similarities using a sigmoid-based modulation function.

2. Background and related work

2.1 Background

2.1.1. Ontology and ontology matching

An ontology is a formal and explicit specification of a shared conceptualization of a domain [1]. It defines the entities (concepts), their properties, and the relationships between them. Ontologies are often represented as graphs, where nodes denote concepts or classes, and edges represent relationships or properties between them. Formally, an ontology graph is defined as a labelled directed graph $G = (V, E)$, where V is the set of entities and E is the set of directed edges denoting relationships.

Ontology matching, also referred to as ontology alignment or ontology mapping, is the task of determining correspondences between entities in different ontologies [4]. It enables interoperability across systems by bridging heterogeneous conceptual models. The output is typically a set of mappings $M = (e_1, e_2, conf)$, where e_1 and e_2 are entities from the source and target ontologies respectively, and $conf \in [0,1]$ denotes the confidence score of the correspondence.

Various similarity measures, including string-based, structural, and semantic approaches, are employed to compute these scores. However, most techniques are either restricted to local structural comparisons (e.g., direct neighbourhood) or incur significant computational cost when attempting global similarity propagation.

2.1.2. Critical Node Detection Problem (CNDP)

Critical nodes in a graph are those whose removal significantly decreases the graph's overall connectivity [6]. Detecting these nodes is the objective of the Critical Node Detection Problem (CNDP), which is typically formulated as an optimization problem [5]. Various connectivity metrics can be used to identify critical nodes, such as maximizing the number of connected components, minimizing pairwise connectivity, or minimizing the size of the largest connected component [7]. Identifying such nodes facilitates a deeper understanding of the graph's structure, properties, and function.

In our approach, we adopt the minimum pairwise connectivity criterion, as it directly reflects the effect of node removal on connectivity. As a result, the partitions obtained correspond to semantically coherent subgraphs, which aligns closely with the goals of ontology alignment.

We rely on the heuristic described in [5]. in which the problem is formalized as follows:

INPUT : An undirected Graph $G = (V, E)$ $and\ an\ Integer\ K;$

OUTPUT : $A = Argmin \sum_{i,j\in(V\backslash A)} U_{ij}(G(V\backslash A)) : |A| \leq K$

$$U_{ij} = \begin{cases} 1, if\ i\ and\ j\ are\ in\ the\ same\ component\ of\ G(V/A), \\ 0, Otherwise \end{cases}$$

The term $\sum_{i,j\in G(V\backslash A)} U_{ij}$ measures the overall pairwise connectivity of the residual graph. An equivalent formulation of the objective function is given by:

$$f(s) = \sum_{Ci\in G(v\backslash s)} \frac{\delta i(\delta i - 1)}{2} \tag{1}$$

where δ_i denotes the size of the C_{ith} connected component.

Figure 1 shows an example of a graph to which the heuristic was applied. The output is the set of nodes {S4, S8}, whose removal decomposes the graph into three connected components:

$$G \setminus \{S4, S8\} = \{\{S2, S1, S3\}, \{S5, S6, S7\}, \{S9, S10, S11, S12\}\}$$

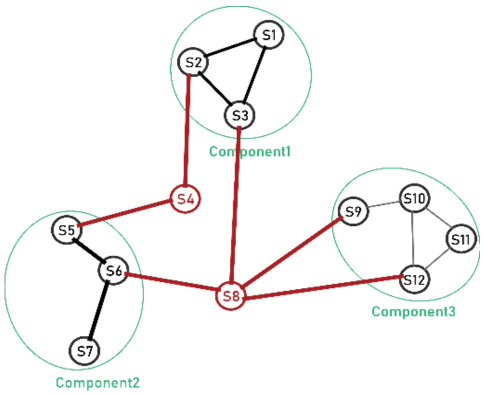

Figure 1. Example of Graph Partitioning through Critical Node Detection

2.2. Related work

Several ontology alignment approaches go beyond simple lexical comparison by incorporating structural and contextual information to improve similarity estimation. These methods can be broadly grouped into three categories: **profile-based and context-based similarity**, **propagation-based similarity**, and **vector space models**.

2.2.1. Profile-Based and Context-Based Similarity

These approaches enrich the representation of entities by leveraging their local and extended context—such as ancestors, descendants, and annotations—resulting in more semantically informed similarity computations. For instance, **CFA** [8] formulates profile-based similarity in biomedical ontologies and applies a firefly optimization algorithm to improve matching quality. **CroMatcher** [9] extracts semantic context from annotated descendant classes, building richer profiles for each entity. **ServOMap** [10] adopts an information retrieval perspective by generating two virtual documents per entity (direct and extended context) to evaluate similarity through textual analysis and Lucene scoring. **YAM++** [11] combines graph indexing with supervised learning to capture both local and contextual features for candidate pair generation. **PoMap** [12] and **MC-ESM** [13] further enhance context-aware matching by modeling entity neighborhoods using structural and lexical signals.

Recent advancements also explore deep representation learning. For example, [14] introduces a dual-attention mechanism that captures multifaceted context by leveraging both textual semantics and structural signals within ontologies. Meanwhile, [15] proposes hybrid embeddings that integrate ontological structure and textual semantics for robust cross-ontology similarity. Additionally, anchor-based partitioning guided by relevance-driven genetic programming [16] provides scalable context exploitation mechanisms suitable for large ontologies.

2.2.2. Propagation-Based Similarity

Propagation-based methods refine similarity values by diffusing them across the ontology graph, allowing indirect relationships to influence the alignment process. Similarity Flooding [17] propagates similarity iteratively based on topological relations, reinforcing correspondences through mutual structural support. Anchor-PROMPT [18] improves efficiency by restricting propagation to subgraphs anchored on high-confidence matches. Anchor-Flood [19] extends this by combining lexical-based anchor generation with graph-based propagation and structural filtering.

These methods are particularly effective at increasing recall by identifying semantically distant yet related matches. However, they require careful tuning of propagation parameters to prevent noise and over-alignment, particularly in large or sparsely connected graphs.

2.2.3. Vector Space Approaches

In vector-based ontology alignment, entities are projected into high-dimensional vector spaces using lexical, structural, or learned features. Similarity is then computed via distance metrics or machine learning algorithms. Traditional vector models [20-22] encode concept features into structured vectors, facilitating rapid alignment computation. In [22], discrete optimization and compact evolutionary algorithms are employed to enhance the alignment process within the vector space.

Although these approaches offer strong performance in specific domains, they often demand extensive feature engineering or training data, and may underperform when ontologies involve deep hierarchies, semantic inconsistencies, or sparsely labelled entities.

2.3. Limitations of Existing Methods and Positioning of CNPMap

Many existing ontology alignment systems suffer from inherent limitations related to how they handle structural context. A common drawback is their local focus, relying exclusively on direct neighbours to compute similarity. This restricted view often overlooks semantically relevant information that exists beyond the immediate neighbourhoods of entities. In contrast, methods that take the entire ontology structure into account may capture broader relationships but typically face significant challenges in terms of scalability and noise amplification, as irrelevant or weak signals are included in the computation.

Another key limitation is that most systems treat ontology partitioning as a preprocessing step, used solely to reduce the search space. They seldom integrate the partition structure into the final similarity computation, thereby missing the opportunity to model contextual coherence more effectively.

To address these shortcomings, CNPMap introduces a novel ontology alignment framework that combines critical node-based partitioning with context-aware similarity refinement. Unlike traditional methods that decouple partitioning from matching, CNPMap integrates partition structure directly into the final alignment process, enhancing both semantic precision and computational efficiency.

This integration enables CNPMap to:

- Capture semantic similarities beyond direct neighbourhoods,
- Leverage the semantic coherence of partitions to provide contextual boundaries,
- Preserve scalability by avoiding unnecessary cross-partition computations.

By incorporating heuristics from the Critical Node Detection Problem (CNDP) and combining them with a hybrid linguistic matcher and a sigmoid-based similarity modulation function, CNPMap offers a novel and effective contribution to the field of ontology alignment. Its design directly tackles the limitations of existing systems while providing a scalable, context-sensitive alternative.

3. CNPMap Methodology

CNPMap is a three-phase ontology alignment framework designed to capture semantic similarities beyond direct structural neighbourhoods. It integrates linguistic similarity computation, critical node-based ontology partitioning, and context-aware refinement to produce more informed and coherent alignments.

The approach consists of the following phases:

- **Phase 1 – InitialMatching:** An initial set of correspondences is computed using a word-level linguistic similarity function. This yields a high-precision set of anchor pairs that serve as guidance for the structural decomposition of the ontologies.
- **Phase 2 – Ontology Partitioning via Critical Node Detection:** Each ontology is partitioned by solving the Critical Node Detection Problem (CNDP), which identifies and removes nodes that play a central role in graph connectivity. The resulting partitions form structurally coherent subgraphs that define localized semantic contexts.
- **Phase 3 – Context-Aware Similarity Refinement:** Within each aligned pair of partitions, entity-level similarity scores are refined using a sigmoid function that integrates both lexical similarity and partition-level context. This process enables the detection of semantically related entities that may not be directly connected in the graph.

By leveraging both linguistic and structural signals, CNPMap aims to enhance the semantic relevance of alignments, particularly in cases where relationships extend beyond immediate neighbourhoods. Figure 2 provides an overview of the approach, and the following sections describe each phase in detail.

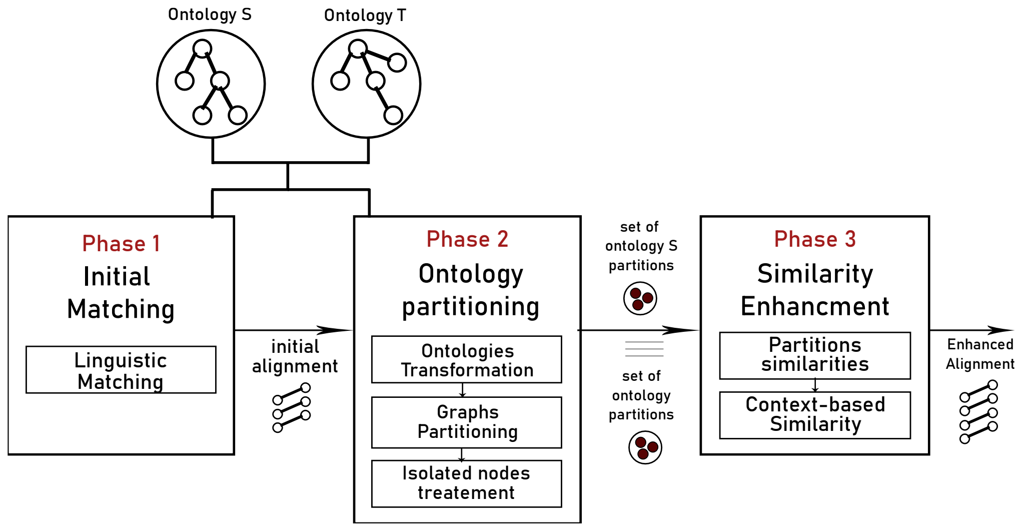

Figure 2. Three-Phase Architecture of CNPMap

3.1. Initial Matching

The method begins by computing an initial alignment using a linguistic similarity measure with a predefined threshold $\boldsymbol{\alpha}_{\boldsymbol{Ini}}$. This measure analyses the words forming the annotations of two entities. To prepare these annotations, we apply a preprocessing pipeline consisting of: (1) removing numbers, punctuation, and stop words; (2) segmenting labels into individual words and converting them to lowercase; and (3) applying stemming. This process produces two sets of words, denoted w_i^s and w_i^t, corresponding to the source and target entities respectively.

The linguistic similarity $Sim_{ling}(d_i^s, d_j^t)$ is then computed between the two word sets w_i^s and w_i^t as follows:

$$Sim_{ling}(d_i^s, d_j^t) = \frac{1}{Min(|w_i^s|, |w_j^t|)} \sum_{V_k \in W_i^s} \left(\max_{V_l \in W_j^t} (Sim_w(v_k, v_l)) \right) \quad (2)$$

Here, $Sim_w(v_k, v_l)$ evaluates the similarity between two individual words using two methods:

- Syntax-based similarity, via the Levenshtein distance [23], which counts the minimum number of edit operations (insertions, deletions, substitutions) to transform one string into another.
- Semantic similarity, using JWNL Distance [24], based on WordNet [25], which combines information content and edge-counting to assess conceptual closeness.

The final word-level similarity is defined as:

$$Sim_w(v_k, v_l) = max(\text{Levenshtein}(v_k, v_l), \text{JWNLDistances}(v_k, v_l)) \quad (3)$$

Only correspondences with a similarity score exceeding $\boldsymbol{\alpha}_{\mathbf{Ini}}$ are retained in the initial alignment set.

3.2. Ontology partitioning

The objective of this phase is to decompose each ontology into multiple partitions based on semantic information. The underlying principle is to group semantically related entities within the same partition.

3.2.1. Ontology transformation

In this study, we propose an ontology matching approach that leverages graph algorithms from the literature. Specifically, ontologies are represented as undirected graphs, where concepts are transformed into nodes and relationships into edges. We focus exclusively on hierarchical relationships, as they represent a fundamental aspect of ontological structure. To facilitate this transformation, we employ the Jena Framework [10], a robust tool for parsing and manipulating ontologies. By leveraging Jena's capabilities, we convert ontologies into graph structures suitable for analysis and comparison using graph-based techniques.

3.2.2. Graph partitioning

In the field of ontology matching, ontology partitioning is widely used in large ontology alignment systems [10, 26-29], and it is a crucial step to reduce the complexity of the matching process. The main objective of graph partitioning is to divide the ontology graph into smaller and more manageable components, while preserving the semantic relationships between the nodes. In the literature, various algorithms have been proposed for ontology partitioning, such as AHSCAN [30], PATO [31], Karger [29, 32], and CP (Clique Percolation) [29, 33]. However, most of these algorithms rely on local search strategies, which may not always lead to optimal partitioning.

The proposed approach adopts a global search strategy for graph partitioning, aiming to produce structurally and semantically similar partitions across heterogeneous ontologies. The algorithm systematically explores possible partitioning configurations and applies heuristic techniques to determine an optimal graph division that preserves semantic coherence. A key objective of this method is to minimize variance among partitions, ensuring a balanced and meaningful distribution of entities. By aligning partitions of comparable structure, the approach improves the likelihood of identifying correct correspondences, thereby enhancing the accuracy and reliability of the ontology matching process.

To achieve similar partitions, our approach relies on a clear yet effective principle: the more two partitions share similarities, the higher their level of similarity. However, it is essential that each partition from the source ontology shares exclusive similarities with only one partition from the target ontology, and vice versa. This constraint ensures that the resulting mappings are both unique and semantically relevant. To implement this, we developed a variant of the CC-CNP algorithm [34], enhanced with a new constraint in the form of a predefined list of nodes that must not appear within the same partition. This addition reinforces semantic cohesion within partitions while preserving their distinctiveness, ultimately enabling more precise and meaningful alignments.

To operationalize this strategy, we propose a constrained variant of the CC-CNP algorithm, detailed in Algorithm 1.

Algorithm 1. CNPMap Algorithm

Procedure ConstrainedCriticalNode (G, β, LIST)
1. MIS ← MaximalIndepSet(G)
2. NoAdd ← 0
3. **while** $(NoAdd = |V| - |MIS|)$ **do**
4. $i \leftarrow argmin\left\{ \sum_{h \in M_j} \frac{\delta_h(\delta_h - 1)}{2} : j \in V \setminus MIS \right\}$
5. **If** $\left(|h| \leq \beta \ \ \forall h \in H_i \subseteq G\ (MIS \cup \{i\}) \right)$ **then**
6. **if** $(p \in h$ ***NAND*** $k \in h \,/\, \forall h \in H_i \subseteq G(MIS \cup \{i\}) : p, k \in LIST)$ **then**
7. $MIS \leftarrow MIS \cup \{i\}$
8. **else**
9. $NoAdd \leftarrow NoAdd + 1$
10. **end if**
11. **end if**
12. **end while**
13. **IsolatedNodesTreatment()**
14. **Return** G(MIS) / ∗set of graph components ∗ /

End procedure ConstrainedCriticalNode

Algorithm 1 presents the pseudo-code of our constrained CC-CNP variant applied to graph partitioning. The algorithm takes three main inputs: an undirected ontology graph $G = (V, E)$, an integer β specifying the maximum allowed size of each partition, and a predefined list of nodes, referred to as $LIST$, which must not be assigned to the same partition. This list typically includes anchor entities identified during the initial matching phase. By incorporating this additional constraint, the algorithm aims to produce partitions that are both semantically cohesive and structurally distinct, thereby enabling more accurate and context-aware ontology alignment.

3.2.2.1. Computation of the Maximum Independent Set (Line 1)

The Maximum Independent Set (MIS) is defined as a set of nodes in a graph that are mutually non-adjacent. The construction of the MIS begins with an empty set. Initially, a single node is added. The algorithm then iteratively examines the remaining nodes, adding any node that is not adjacent to those already included in the MIS. This process continues until no further nodes can be added. The resulting set is considered maximal because no additional nodes can be included without violating the independence condition.

3.2.2.2. Calculation of partitions (lines 3 to 12)

After identifying the MIS, we introduce a variable called $NoAdd$ that tracks the number of nodes whose inclusion could potentially violate the predefined maximum size limit of partitions (β). This variable serves a dual purpose: it ensures that the partitioning process adheres to the predefined constraints and acts as a termination condition for the algorithm.

The algorithm then evaluates each node exhaustively within the set ($V \setminus MISV$) to determine its suitability for reintroduction into the graph while maintaining the prescribed objective function (1). The algorithm enters a loop spanning lines 2 to 5, employing a greedy approach to identify a node i from the set V that is not currently part of the MIS. The goal of this selection is to minimize the objective function for the graph $G(MIS \cup \{i\})$, where i can be any node in the set V.

The selection process for adding node i to the MIS is governed by two crucial conditions, as outlined in lines 5 and 6 of the algorithm. The first condition requires that the size of each partition h in the graph $G(MIS \cup \{i\})$ does not exceed the predefined limit β. This ensures that the partitioning process maintains a balanced distribution of nodes among the partitions.

The second condition stipulates those nodes listed in $LIST$ must not share common partitions. This constraint is essential to guarantee that each partition in the source ontology corresponds to at most one partition in the target ontology, thereby maximizing the relevance of the matching process.

Deviations from these conditions trigger the incrementation of $NoAdd$, and this iterative process continues until $NoAdd$ equals $|\ V\ | - |\ MIS|$.

3.2.2.3. Treatment of isolated nodes

Isolated nodes are those which are not part of a partition. Consequently, the nodes concerned include the partitions of dimension 1 as well as the critical nodes. The isolated node is denoted as N_{iso}. The objective of this phase lies in determining the partition which will accommodate the isolated node among those connected to it. We reapply the objective function to the graph $G(MIS \cup \{N_{iso}\})$ during the process of selecting the appropriate partition. Unlike the previous phase (phase 2), for the calculation of "pairwise connectivity", only the links to a single partition are added to the graph $G(MIS \cup \{N_{iso}\})$. This is done to avoid the merging of several partitions, thus making it possible to only increase the "pairwise connectivity" of the selected partition. Therefore, the objective function minimization aims to incorporate the isolated node into the partition that requires the fewest additional nodes.

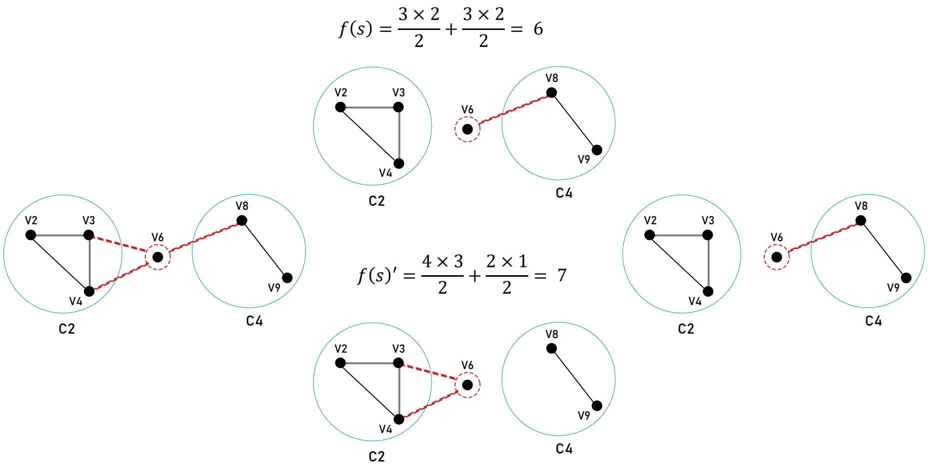

Figure 3. Illustration of the treatment of an isolated node during the partitioning process

Figure 3 illustrates a critical node, V6, which is connected to two partitions, $C2$ and $C4$. To determine the appropriate partition for this isolated node, the objective function is computed twice: $f(s)$ and $f(s)'$. In $f(s)$, only the relations with $C4$ are retained, whereas in $f(s)'$, only the relations with $C2$ are considered. Since $f(s)$ yields a lower value than $f(s)'$, node $V6$ is assigned to partition $C4$.

3.3. Similarity Enhancement

After generating an initial set of correspondences and partitioning the ontologies, it becomes necessary to refine similarity scores by integrating structural and contextual relationships among entities. This step helps uncover semantically meaningful correspondences that may not be apparent through lexical similarity alone.

3.3.1. Partitions similarity

Once the source and target ontologies are partitioned, it is necessary to assess the similarity between corresponding partitions $C_i^s \in O_s$ and $C_j^c \in O_t$. To this end, we introduce a novel similarity measure, $SimCC$, which evaluates the similarity between partitions based on the entities they contain. Specifically, $SimCC$ computes the similarity score by aggregating the similarities between the constituent nodes of each partition pair. The formula used for this computation is provided in Equation (4).

$$SimCC(C_i^s, C_j^c) = \frac{1}{Min(|C_i^s|, |C_j^c|)} \sum_{V_k \in C_i^s} \left(\max_{V_l \in C_j^c} \left(Sim_{ln}(V_k, V_l) \right) \right) \tag{4}$$

Where:

- C_i^s and C_j^c are the source and target partitions, respectively.
- $|C_i^s|$ and $|C_j^c|$ represent the number of entities in each partition.
- $Sim_{ln}(V_k, V_l)$ is the linguistic similarity score between entities $V_k \in C_i^s$ and $V_l \in C_j^c$.

Linguistic similarities are calculated by comparing the textual descriptions of classes. For this purpose, we apply the string comparison method described in [35], which evaluates similarity by combining commonality, difference, and refinement components:

$$Sim_{ln}(V_i, V_j) = comm(d_i, d_j) - diff(d_i, d_j) + winkler(d_i, d_j) \quad (5)$$

Where:

- d_i (respectively d_j) denotes the description of the entity V_i (respectively V_j).
- $comm(d_i, d_j)$ measures the commonality between the two descriptions.
- $diff(d_i, d_j)$ measures the difference between them.
- $winkler(d_i, d_j)$ applies a refinement based on the Winkler method, as introduced in [35].

$SimCC$ compares two sets of entities: C_i^s from the source ontology and C_j^c from the target ontology. It computes the similarity between these sets by aggregating the maximum pairwise similarity scores Sim_{ln} between their respective entities. The final similarity score is normalized by the minimum cardinality of the two sets, i.e., $min(|C_i^s|, |C_j^c|)$.

3.3.2. Context-based similarity

In the final phase of our ontology alignment process, we refine the similarity scores between entities belonging to similar partitions. This step aims to enhance alignment quality by identifying new correspondences that were not evident in the initial matching. We treat partitions as semantic contexts for the entities they contain and adjust the entity-level similarity computation accordingly, leveraging the similarity between their respective partitions.

To compute the final similarity score, denoted as $Sim_{context}$, we introduce a parameterized variant of the sigmoid function (Equation 6). Widely used in neural networks, the sigmoid function is valued for its smooth, bounded, and differentiable nature, making it ideal for modulating confidence scores derived from heterogeneous similarity signals. In our approach, the sigmoid curve is modulated by the inter-partition similarity $SimCC$, which governs how strongly partition-level similarity influences the final similarity score between entities:

$$Sim_{context} = \frac{1}{1 + e^{-(f \times SimCC) \times Sim_{ln}}} \quad (6)$$

Where:

- Sim_{ln} is the initial linguistic similarity between two entities (as defined in Equation 5),
- $SimCC$ is the similarity between their respective partitions (Equation 4),
- f is a scaling factor used to normalize the impact of $SimCC$.

By adjusting the slope of the sigmoid function through $f \times SimCC$, we dynamically control the function's sensitivity to similarity variations. A shallow curve (low $SimCC$) filters out weak matches, whereas a steep curve (high $SimCC$) amplifies even moderate similarities, encouraging alignment. This mechanism ensures that contextual information significantly influences the final similarity score, improving both the precision and coherence of the alignment.

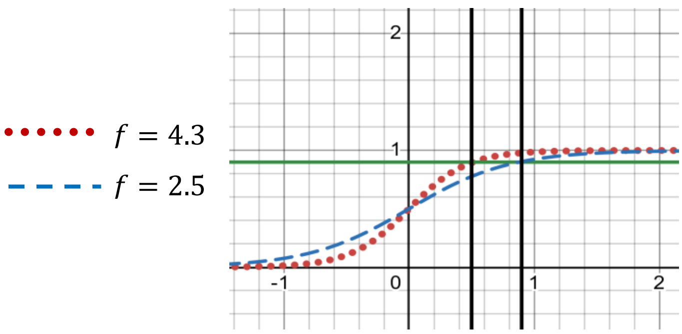

Figure 4. Impact of the scaling factor f on the sigmoid curve.

As shown in Figure 3, setting $f = 4.3$ allows entity similarities ≥ 0.5 to be considered valid alignments, whereas $f = 2.5$ restricts this to similarities ≥ 0.9. In our design, the similarity between entities serves as the input to the function, and the curve's steepness is regulated by the similarity between their respective partition. This parametric control enhances both the precision and interpretability of the similarity scores.

The parameter f was selected through empirical tuning based on performance across benchmark datasets. Nevertheless, a more systematic approach to parameter optimization, potentially leveraging learning-based methods, represents a promising avenue for future research.

4. Experimental study and results discussion

In the experiments, we utilize the benchmark track provided by the Ontology Alignment Evaluation Initiative (OAEI 2023)[1] to assess the performance of our approach. Each test case in the benchmark track consists of two ontologies, one serving as the source ontology and the other as the target ontology, along with a reference alignment used to evaluate the effectiveness of the ontology matcher. To enable comparisons with various matching techniques, we evaluate the alignment's quality in this study using recall, precision, and the F-measure [36].

Given a reference alignment A_{ref}, the definitions of recall (R), precision (P), and F-measure (F) for an alignment A are as follows:

$$Recall(A) = \frac{|A \cap A_{ref}|}{|A_{ref}|} \tag{7}$$

$$Precision(A) = \frac{|A \cap A_{ref}|}{|A|} \tag{8}$$

$$F - mesure(A) = \frac{2 \times Recall(A) \times Precision(A)}{Pecall(A) + Precision(A)} \tag{9}$$

4.1. Experiment configuration

In this work, we use the following parameters:

- $\boldsymbol{\alpha_{Ini}}$ = 0.95: Threshold for the initial similarity between entities in the source and target ontologies.
- $\boldsymbol{\beta}$ = 6: Maximum allowed size for each partition.
- $\boldsymbol{f}$ = 5: Curve factor controlling the slope of the sigmoid function.

The choice of these parameter values follows several guiding principles:

- A high threshold ($\boldsymbol{\alpha_{Ini}}$) was selected to ensure high-quality initial alignments. This is crucial, as the subsequent partitioning phase directly depends on the accuracy of this initial alignment. Reliable anchors serve as the foundation for effective and semantically coherent partitions.
- The value of $\boldsymbol{\beta}$ directly impacts computational complexity. As the number of nodes in a partition increases, so does the processing time. To identify a suitable value, we conducted preliminary experiments on the OAEI 2023 track by systematically varying $\boldsymbol{\beta}$ and observing its effect on runtime and alignment quality.
- The parameter $\boldsymbol{f}$ governs the steepness of the sigmoid curve used in the final similarity computation. A larger value of $\boldsymbol{f}$ leads to a steeper curve, allowing moderate similarity values to reach the alignment threshold. This parameter was set empirically based on its observed influence on F-measure during pre-experiments.

We conducted a pre-experimental study on the OAEI 2023 benchmark to observe the impact of these parameters on alignment quality. Our findings indicate that the chosen configuration yields the best performance across the tested ontology pairs.

The experiments were run on a machine with the following specifications:

- Processor: Intel Core i7-10750H
- Clock Speed: 2.60 GHz × 6 cores
- RAM: 16 GB

[1] https://oaei.ontologymatching.org/2023/conference/

The detailed experimental results are presented in the next section.

4.2. Evaluation of our approach in terms of Recall, Precision and F-measure

To evaluate our ontology alignment approach, we selected the OAEI 2023 Conference track. This choice is motivated by the fact that this track includes real-world ontologies characterized by diverse terminology and complex structures, making it particularly suitable for assessing context-aware and structure-based alignment methods. Additionally, as a widely adopted benchmark in the ontology matching community, it allows for rigorous and reproducible comparisons with state-of-the-art systems. Its moderate size also makes it ideal for detailed evaluation, without the computational burden often associated with larger datasets such as Anatomy or LargeBio.

In our experiments, we compared the initial alignments, based purely on lexical similarity, with the refined alignments generated by our complete approach. The results, presented in Table 1, show that our method leads to tangible improvements in alignment quality for several ontology pairs. For example, the alignment between the cmt and conference ontologies showed a noticeable increase in F-measure from 0.69 to 0.74, while the alignment between cmt and sigkdd improved from 0.76 to 0.82. A more moderate gain was observed between edas and ekaw, where the F-measure rose from 0.60 to 0.63. However, the alignment between cmt and ekaw experienced a slight degradation, with a drop of 3% in the F-measure.

Although the performance remained stable for the majority of ontology pairs, these results illustrate that our method can yield significant gains when contextual information plays a decisive role. The observed improvements confirm the relevance of the Conference track as a benchmark and underscore the value of our context-driven approach for enhancing alignment precision.

4.3. Evaluation of results with different partition size

According to our hypothesis, unlike traditional systems, structural similarity should not be computed solely based on entities in the immediate neighbourhoods. Instead, we propose extending the structural context to include all entities within the same partition. We argue that this broader context enhances the precision and reliability of ontology alignment. To validate this assumption, we conducted additional experiments focusing specifically on alignments that improved during the similarity enhancement phase. In each experiment, we varied the maximum allowed partition size β from 1 to 10.

Table 1. Evaluation of our approach in terms of Recall, Precision and F-measure

		Initial Alignment			Enhanced Alignment			
Source ontology	**Target Ontology**	**R**	**P**	**F**	**R**	**P**	**F**	**Enhancement**
cmt	conference	0.6	0.81	0.69	0.67	0.83	0.74	5%
cmt	confOf	0.31	0.83	0.45	0.31	0.83	0.45	0%
cmt	edas	0.62	0.89	0.73	0.62	0.89	0.73	0%
cmt	ekaw	0.55	0.75	0.63	0.55	0.67	0.6	-3%
cmt	iasted	1	0.67	0.8	1	0.67	0.8	0%
cmt	sigkdd	0.67	0.89	0.76	0.75	0.9	0.82	6%
conference	confOf	0.67	0.83	0.74	0.67	0.83	0.74	0%
conference	edas	0.53	0.75	0.62	0.53	0.75	0.62	0%
conference	ekaw	0.44	0.58	0.5	0.44	0.58	0.5	0%
conference	iasted	0.43	0.55	0.48	0.43	0.55	0.48	0%
conference	sigkdd	0.67	0.83	0.74	0.67	0.83	0.74	0%
confOf	edas	0.47	0.75	0.58	0.47	0.75	0.58	0%
confOf	ekaw	0.7	0.88	0.78	0.7	0.88	0.78	0%
confOf	iasted	0.78	0.64	0.7	0.78	0.64	0.7	0%
confOf	sigkdd	0.57	0.8	0.67	0.57	0.8	0.67	0%
edas	ekaw	0.52	0.71	0.6	0.57	0.72	0.63	3%
edas	iasted	0.47	0.64	0.55	0.47	0.64	0.55	0%
edas	sigkdd	0.47	1	0.64	0.47	1	0.64	0%
ekaw	iasted	0.7	0.58	0.64	0.7	0.58	0.64	0%
ekaw	sigkdd	0.73	0.89	0.8	0.73	0.89	0.8	0%
iasted	sigkdd	0.73	0.73	0.73	0.73	0.73	0.73	0%

Table 2 presents the variation in F-measure as a function of β for different ontology pairs. For the first alignment (cmt–conference), the improvement occurred when β = 3, with the F-measure rising from 0.69 to 0.74. However, in this case, it is unclear whether the improvement stems from our method, since the similarity computation might still rely solely on the direct neighbourhoods. In contrast, for the cmt–sigkdd and edas–ekaw alignments, improvements were only observed starting from β = 6. The F-measure increased from 0.76 to 0.82 for cmt–sigkdd and from 0.60 to 0.63 for edas–ekaw, indicating that entities beyond the immediate neighbourhoods contributed to the discovery of new correspondences.

These results support our hypothesis: expanding the structural context to include all entities within a partition enables the identification of additional meaningful alignments that would otherwise be missed. Nonetheless, the optimal value of β appears to be ontology-dependent, suggesting that further experimentation is needed to adapt partition granularity to the characteristics of each ontology pair.

Table 2. Evaluation of Alignment Performance with Varying Partition Size (β)

	The maximum partition size (β)									
Ontologies	**β=1**	**β=2**	**β=3**	**β=4**	**β=5**	**β=6**	**β=7**	**β=8**	**β=9**	**β=10**
cmt-conference	69%	69%	74%	74%	74%	74%	74%	74%	74%	74%
cmt-sigkdd	76%	76%	76%	76%	76%	82%	82%	82%	82%	82%
edas-ekaw	60%	60%	60%	60%	60%	63%	63%	63%	63%	63%

4.4. Comparison with OAEI 2023 Systems

To assess the performance of our approach, we compared CNPMap against several systems that participated in the OAEI 2023 Conference track. As shown in Table 3, CNPMap achieved an F-measure of 0.66, indicating a strong balance between precision (0.75) and recall (0.61).

Although this result demonstrates the competitiveness of our method, there remains potential for improvement, particularly in terms of recall. Increasing recall would enable the system to capture a broader set of correct correspondences, thereby improving the completeness of the alignments without compromising the current level of precision.

Table 3. Comparison of Proposed Approach with OAEI 2023 Participants' Approaches

Matchers	R	P	F
GraphMatcher[37]	0.77	0.71	0.74
CNPMap	0.61	0.75	0.66
SORBETMtch[38]	0.61	0.73	0.66
LogMap[39]	0.56	0.76	0.64
Matcha[40]	0.62	0.62	0.62
OLaLa[41]	0.61	0.59	0.60
ALIN[42]	0.44	0.82	0.57
LogMapLt[39]	0.47	0.68	0.56
LSMatch[43]	0.41	0.83	0.55
AMD[44]	0.41	0.82	0.55
TOMATO[45]	0.47	0.57	0.52
PropMatch[46]	0.08	0.86	0.15

In terms of precision, CNPMap outperforms many of the compared systems, highlighting its effectiveness in generating accurate mappings. Notably, CNPMap ranks second in terms of overall F-measure, underscoring its robustness and reliability for ontology alignment, especially when contextual and structural features are critical.

In conclusion, CNPMap demonstrates strong alignment capabilities, particularly through its precision-driven matching process. With future enhancements focused on recall, the approach has the potential to further increase its alignment coverage while maintaining its high-quality results. This makes CNPMap a promising solution for scalable and context-aware ontology alignment.

5. Conclusion

In this study, we introduced CNPMap, an innovative ontology alignment approach that leverages the Critical Node Detection Problem (CNDP) to partition ontologies and compute similarities beyond direct neighbourhoods. By integrating both structural and contextual information, CNPMap addresses a key limitation of traditional methods that often rely exclusively on local features.

Our approach was evaluated on the OAEI 2023 Conference track, a benchmark composed of real-world ontologies with rich structures and diverse terminologies. The experimental results demonstrated notable improvements in alignment quality. For example, the F-measure increased from 0.69 to 0.74 for the cmt-conference alignment, from 0.76 to 0.82 for cmt-sigkdd, and from 0.60 to 0.63 for edas-ekaw. These improvements validate the effectiveness of our context-based similarity refinement strategy, particularly in cases where local structural information alone is insufficient.

We also analyzed the role of the partition size parameter (β) and observed that values greater than 5 are necessary to capture non-local similarities. This supports our hypothesis that semantically coherent partitions facilitate the discovery of additional valid correspondences, even among non-adjacent entities in the ontology graph.

In comparison with systems participating in OAEI 2023, CNPMap achieved a precision of 0.75 and an F-measure of 0.66, ranking second overall. These results underscore the robustness and reliability of our method in maintaining a strong balance between precision and recall.

Nevertheless, recall remains a challenge, primarily due to the strict configuration used during the initial alignment phase (α_Ini = 0.95). While this threshold ensures high-precision anchor correspondences and improves partitioning reliability, it also filters out many potentially valid but lower-confidence matches. Additionally, the strict contextual segmentation induced by partitioning can limit the discovery of cross-partition alignments.

To address these limitations, we propose several future research directions:

- Implement adaptive thresholding mechanisms to retain a broader range of candidate correspondences.
- Explore multi-context propagation techniques that permit controlled overlap between partitions, enhancing recall.
- Explore machine learning techniques to automatically calibrate the sigmoid parameter f enabling dynamic adaptation of similarity modulation across diverse datasets and domains.
- Develop scalable implementations using parallel computing or graph summarization techniques to handle large ontologies efficiently.
- These enhancements aim to improve the scalability, recall, and alignment coverage of CNPMap while preserving its high precision.

In summary, CNPMap represents a promising and effective framework for context-aware ontology alignment. With further refinement, it can become a high-performance solution for aligning complex, large-scale ontologies across diverse domains and applications.

CRediT Author Contribution Statement

Abderrahmane Messous: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Visualization, and Writing – Original Draft; Fatiha Barigou: Project administration, Supervision, and Writing – Review & Editing.

References

[1] Thomas R. Gruber, "Toward Principles for the Design of Ontologies Used for Knowledge Sharing", *International Journal of Human-Computer Studies*, Print ISSN:1071-5819, Online ISSN: 1095-9300, Vol. 43, No. 5–6, November 1995, pp. 907–928, Published by Elsevier B. V., DOI: 10.1006/ijhc.1995.1081, Available: https://www.sciencedirect.com/science/article/pii/S1071581985710816.

[2] Pavel Shvaiko and Jérôme Euzenat, "Ontology Matching: State of the Art and Future Challenges", *IEEE Transactions on Knowledge and Data Engineering*, Print ISSN:1041-4347, Online ISSN: 1558-2191, Vol. 25, No. 1, 13 December 2013, pp. 158–176, Published by IEEE, DOI: 10.1109/TKDE.2011.253, Available: https://ieeexplore.ieee.org/document/6104044.

[3] Lorena Otero-Cerdeira, Francisco J. Rodríguez-Martínez and Alma Gómez-Rodríguez, "Ontology Matching: A Literature Review", *Expert Systems with Applications*, ISSN: 0957-4174, Vol. 42, No. 2, 1 February 2015, pp. 949–971, Published by Elsevier, DOI: 10.1016/j.eswa.2014.08.032, Available: https://www.sciencedirect.com/science/article/abs/pii/S0957417414005144.

[4] Jérôme Euzenat and Pavel Shvaiko and Ontology Matching, *Ontology Matching*, 2nd ed. Heidelberg, Germany: Springer, Available: https://link.springer.com/book/10.1007/978-3-642-38721-0.

[5] Ashwin Arulselvan, Clayton W. Commander, Lily Elefteriadou and Panos M. Pardalos, "Detecting Critical Nodes in Sparse Graphs", *Computers & Operations Research*, ISSN:0305-0548, Vol. 36, No. 7, 2009, pp. 2193–2200, July 2009, Published by Elsevier B. V., DOI: 10.1016/j.cor.2008.08.016, Available: https://www.sciencedirect.com/science/article/abs/pii/S0305054808001494.

[6] Arulselvan Arulselvan, Clayton W. Commander, Panos M. Pardalos and Oleg Shylo, "Managing Network Risk via Critical Node Identification", *Risk Management in Telecommunication Networks*, Springer, 2007, Published by Springer, Available: https://pureportal.strath.ac.uk/en/publications/managing-network-risk-via-critical-node-identification.

[7] Mouloud Lalou, Mohammed Amine Tahraoui and Hamamache Kheddouci, "The Critical Node Detection Problem in Networks: A Survey", *Computer Science Review*, Print ISSN: 1574-0137, Online ISSN: 1879-0755, Vol. 28, 1 June 2018, pp. 92–117, Published by Elsevier, DOI: 10.1016/j.cosrev.2018.02.002, Available: https://www.sciencedirect.com/science/article/abs/pii/S1574013716302416.

[8] Xiaoyong Xue, "A Compact Firefly Algorithm for Matching Biomedical Ontologies", *Knowledge and Information Systems*, ISSN: 0219-1377, Vol. 62, No. 8, 1 August 2020, pp. 3171–3187, Published by Springer, DOI: 10.1007/s10115-020-01443-6, Available: https://link.springer.com/article/10.1007/s10115-020-01443-6.

[9] Mladen Gulić, Blaz Vrdoljak and Mladen Banek, "CroMatcher: An Ontology Matching System Based on Automated Weighted Aggregation and Iterative Final Alignment", *Journal of Web Semantics*, ISSN:1570-8268, Vol. 41, 1 December 2016, pp. 50–71, Published by Elsevier, DOI: 10.1016/j.websem.2016.09.001, Available: https://www.sciencedirect.com/science/article/abs/pii/S1570826816300361.

[10] Guéla Diallo, "An Effective Method of Large Scale Ontology Matching", *Journal of Biomedical Semantics*, ISSN: 2041-1480, Vol. 5, No. 1, 1 December 2014, p. 44, Published by BioMed Central, DOI: 10.1186/2041-1480-5-44, Available: https://jbiomedsem.biomedcentral.com/articles/10.1186/2041-1480-5-44.

[11] Duy Hoa Ngo, Zohra Bellahsene and Rémi Coletta, "YAM++: A Combination of Graph Matching and Machine Learning Approach to Ontology Alignment Task", *Journal of Web Semantics*, Print ISSN: 1570-8268, Online ISSN: 1570-8276, Vol. 16, 1 December 2012, pp. 30–44, Published by Elsevier, DOI: 10.1016/j.websem.2012.07.002, Available: https://www.sciencedirect.com/science/article/abs/pii/S1570826812000526.

[12] Amir Laadhar, Faiza Ghozzi, Imen Megdiche, Franck Ravat, Olivier Teste *et al.*, "POMap: An Effective Pairwise Ontology Matching System", in *Proceedings of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (KEOD 2017)*, 1-3 November 2017 , Madeira, Portugal, ISBN: 978-989-758-272-1, pp. 161–168, Published by SciTePress, DOI: 10.5220/0006492201610168, Available: https://www.scitepress.org/Papers/2017/64922.

[13] Xiaoyong Xue and Chuan Jiang, "Matching Sensor Ontologies with Multi-Context Similarity Measure and Parallel Compact Differential Evolution Algorithm", *IEEE Sensors Journal*, Print ISSN: 1530-437X, Online ISSN: 1558-1748, Vol. 21, No. 21, 1 November 2021, pp. 24570–24578, Published by IEEE, DOI: 10.1109/JSEN.2021.3115471, Available: https://ieeexplore.ieee.org/document/9547321.

[14] Vivek Iyer, Arvind Agarwal and Harshit Kumar, "VeeAlign: multifaceted context representation using dual attention for ontology alignment", *in Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Republic, 7-11 November 2021, Punta Cana, Dominican Republic, ISSN:1063-6382, pp. 10780–10792, Published by Association for Computational Linguistics, DOI: 10.18653/v1/2021.emnlp-main.842, Available: https://aclanthology.org/2021.emnlp-main.842.

[15] Zhigang Hao, Wolfgang Mayer, Jingbo Xia, Guoliang Li, Li Qin *et al.*, "Ontology Alignment with Semantic and Structural Embeddings", *Journal of Web Semantics*, Print ISSN: 1570-8268, Online ISSN: 1570-8276, Vol. 78, 1 July 2023, Article No. 100798, Published by Elsevier, DOI: 10.1016/j.websem.2023.100798, Available: https://www.sciencedirect.com/science/article/abs/pii/S1570826823000276.

[16] Donglei Sun, Qing Lv, Pei-Wei Tsai, Xingsi Xue and Kai Zhang, "Anchor-Based Ontology Partitioning and Genetic Programming with Relevance Reasoning for Large-Scale Biomedical Ontology Matching", *Expert Systems with Applications*, Print ISSN: 0957-4174, Online ISSN: 1873-6793, Vol. 270, 25 April 2025, Article No. 126445, Published by Elsevier, DOI: 10.1016/j.eswa.2025.126445, Available: https://www.sciencedirect.com/science/article/abs/pii/S0957417425000673.

[17] Sergey Melnik, Hector Garcia-Molina and Erhard Rahm, "Similarity Flooding: A Versatile Graph Matching Algorithm and Its Application to Schema Matching", in *Proceedings of the 18th International Conference on Data Engineering (ICDE 2002)*, 26 February - 01 March 2002, San Jose, CA, USA, ISSN:1063-6382, pp. 117–128, Published by IEEE, DOI: 10.1109/ICDE.2002.994702, Available: https://ieeexplore.ieee.org/document/994702.

[18] Natalya F. Noy and Mark A. Musen, "Anchor-PROMPT: Using Non-Local Context for Semantic Matching", in *Proceedings of the Workshop on Ontologies and Information Sharing at the 17th International Joint Conference on*

Artificial Intelligence (IJCAI 2001), 4 August 2001, Seattle, USA, ISBN: 978-1-55860-812-2, pp. 63–70, Published by AAAI Press, Available: http://dit.unitn.it/~accord/RelatedWork/Matching/noy.pdf.

[19] Md. Hanif Seddiqui and Masaki Aono, "An Efficient and Scalable Algorithm for Segmented Alignment of Ontologies of Arbitrary Size", *Journal of Web Semantics*, Print ISSN: 1570-8268, Online ISSN: 1570-8276, Vol. 7, No. 4, 1 December 2009, pp. 344–356, Published by Elsevier, DOI: 10.1016/j.websem.2009.09.001, Available: https://www.sciencedirect.com/science/article/abs/pii/S1570826809000432.

[20] Rubén Tous and Jaime Delgado, "A Vector Space Model for Semantic Similarity Calculation and OWL Ontology Alignment", in *Proceedings of the 17th International Conference on Database and Expert Systems Applications (DEXA 2006)*, Lecture Notes in Computer Science, 4-8 September 2006, Kraków, Poland, Print ISBN:978-3-540-37871-6, Online ISBN:978-3-540-37872-3, Vol. 4080, pp. 307–316, Published by Springer, DOI: 10.1007/11827405_30, Available: https://link.springer.com/chapter/10.1007/11827405_30.

[21] Zahra Eidoon, Nasser Yazdani and Farhad Oroumchian, "Ontology Matching Using Vector Space", in *Proceedings of the 30th European Conference on Information Retrieval (ECIR 2008)*, Lecture Notes in Computer Science, 30 March - 3 April 2008, Glasgow, United Kingdom, Print ISBN:978-3-540-78645-0, Online ISBN:978-3-540-78646-7, Vol. 4956, pp. 472–481, Published by Springer, DOI: 10.1007/978-3-540-78646-7_45, Available: https://link.springer.com/chapter/10.1007/978-3-540-78646-7_45.

[22] Shu-Chuan Chu, Xingsi Xue, Jeng-Shyang Pan and Xiaojing Wu, "Optimizing Ontology Alignment in Vector Space", *Journal of Internet Technology*, Print ISSN: 1607-9264, Vol. 21, No. 1, 1 January 2020, pp. 15–22, Published by National Dong Hwa University, DOI: 10.3966/160792642020012101002, Available: https://jit.ndhu.edu.tw/article/view/2218.

[23] Vladimir I. Levenshtein, "Binary Codes Capable of Correcting Deletions, Insertions, and Reversals", *Soviet Physics Doklady*, Vol. 10, No. 8, February 1966, pp. 707–710, Published by the Academy of Sciences of the USSR, Available: https://nymity.ch/sybilhunting/pdf/Levenshtein1966a.pdf.

[24] Zhen Zhao, Jiajun Yan, Lijun Fang and Peng Wang, "Measuring Semantic Similarity Based on WordNet", in *Proceedings of the 2009 Sixth Web Information Systems and Applications Conference (WISA 2009)*, 18-20 September 2009, Xuzhou, China, ISBN:978-0-7695-3874-7, pp. 89–92, Published by IEEE, DOI: 10.1109/WISA.2009.14, Available: https://ieeexplore.ieee.org/document/5368123.

[25] George A. Miller, "WordNet: A Lexical Database for English", *Communications of the ACM*, 1 November 1995, New York, United States, Print ISSN: 0001-0782, Vol. 38, No. 11, pp. 39–41, Published by ACM, DOI: 10.1145/219717.219748, Available: https://dl.acm.org/doi/10.1145/219717.219748.

[26] Xingsi Xue and Jeng-Shyang Pan, "A Segment-Based Approach for Large-Scale Ontology Matching", *Knowledge and Information Systems*, Print ISSN: 0219-1377, Online ISSN: 0219-3116, Vol. 52, No. 2, 1 August 2017, pp. 467–484, Published by Springer, DOI: 10.1007/s10115-016-1018-9, Available: https://link.springer.com/article/10.1007/s10115-016-1018-9.

[27] Archana Patel, Sarika Jain "A partition based framework for large scale ontology matching", *Recent Patents on Engineering*, ISSN: 1872-2121, Vol. 14, No. 3, 2020, pp. 488-501, Published by Bentham Science Publishers, DOI: 10.2174/1872212113666190211141415, Available: https://www.eurekaselect.com/article/96545.

[28] Xingsi Xue and Jie Zhang, "Matching Large-Scale Biomedical Ontologies with Central Concept Based Partitioning Algorithm and Adaptive Compact Evolutionary Algorithm", *Applied Soft Computing*, Print ISSN: 1568-4946, Online ISSN: 1872-9681, Vol. 106, 1 October 2021, DOI: 10.1016/j.asoc.2021.107343, Available: https://www.sciencedirect.com/science/article/abs/pii/S1568494621002660.

[29] Fatmana Şentürk and Vecdi Aytac, "A Graph-Based Ontology Matching Framework", *New Generation Computing*, Print ISSN: 0288-3635, Online ISSN: 1882-7055, Vol. 42, No. 1, 1 January 2024, pp. 33–51, Published by Springer, DOI: 10.1007/s00354-022-00200-3, Available: https://link.springer.com/article/10.1007/s00354-022-00200-3.

[30] Nurcan Yuruk, Mutlu Mete, Xiaowei Xu and T. A. Schweiger, "AHSCAN: Agglomerative Hierarchical Structural Clustering Algorithm for Networks", in *Proceedings of the 2009 International Conference on Advances in Social Network Analysis and Mining (ASONAM 2009)*, 20-22 July 2009, Athens, Greece, ISBN: 978-0-7695-3689-7, pp. 72–77, Published by IEEE, DOI: 10.1109/ASONAM.2009.74, Available: https://ieeexplore.ieee.org/document/5231935.

[31] Heiner Stuckenschmidt and Anne Schlicht, "Structure-Based Partitioning of Large Ontologies", in *Modular Ontologies: Concepts, Theories and Techniques for Knowledge Modularization*, Lecture Notes in Computer Science, Print ISBN: 978-3-642-01906-7, Online ISBN: 978-3-642-01907-4, Vol. 5445, 1 June 2009, pp. 187–210, Published by Springer, DOI: 10.1007/978-3-642-01907-4_9, Available: https://link.springer.com/chapter/10.1007/978-3-642-01907-4_9.

[32] David R. Karger, "Global Min-Cuts in RNC, and Other Ramifications of a Simple Min-Cut Algorithm", in *Proceedings of the 4th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 1993)*, ISBN:0898713137, 25-

27 January 1993, Texas, Austin, USA, pp. 21–30, Published by ACM, DOI: 10.5555/313559.313605, Available: https://dl.acm.org/doi/10.5555/313559.313605.

[33] Afnizanfaizal Abdullah, Safaai Deris, Siti Zaiton Mohd Hashim and Hamimah Mohd Jamil, "Graph Partitioning Method for Functional Module Detections of Protein Interaction Network", in *Proceedings of the 2009 International Conference on Computer Technology and Development (ICCTD 2009)*, ISBN:978-0-7695-3892-1, Vol. 1, 13-15 November 2009, Kota Kinabalu, Malaysia, pp. 230–234, Published by IEEE, DOI: 10.1109/ICCTD.2009.168, Available: https://ieeexplore.ieee.org/document/5359747.

[34] Mohammed Lalou, Mohammed Amine Tahraoui and Hamamache Kheddouci, "Component-Cardinality-Constrained Critical Node Problem in Graphs", *Discrete Applied Mathematics*, Print ISSN: 0166-218X, Online ISSN: 1872-6771, Vol. 210, 1 January 2016, pp. 150–163, Published by Elsevier, DOI: 10.1016/j.dam.2015.01.043, Available: https://www.sciencedirect.com/science/article/abs/pii/S0166218X15000797.

[35] William E. Winkler, "The State of Record Linkage and Current Research Problems", Statistical Research Division, U.S. Bureau of the Census, 1 October 1999, Available: https://www.census.gov/content/dam/Census/library/working-papers/1999/adrm/rr99-04.pdf.

[36] C. J. Van Rijsbergen, "Foundation of Evaluation", *Journal of Documentation*, Print ISSN: 0022-0418, Vol. 30, No. 4, 1 December 1974, pp. 365–373, Published by Emerald, DOI: 10.1108/eb026584, Available: https://www.emerald.com/insight/content/doi/10.1108/eb026584/full/html.

[37] Sefika Efeoglu, "GraphMatcher: A Graph Representation Learning Approach for Ontology Matching", in *Proceedings of the 17th International Workshop on Ontology Matching (OM 2022)*, 1 October 2022, Hangzhou, China, pp. 174–180, Published by CEUR-WS.org, Available: https://ceur-ws.org/Vol-3324/oaei22_paper7.pdf.

[38] Francis Gosselin and Amal Zouaq, "SORBET: A Siamese Network for Ontology Embeddings Using a Distance-Based Regression Loss and BERT", in *Proceedings of the 22nd International Semantic Web Conference (ISWC 2023)*, Lecture Notes in Computer Science, 6-10 November 2023, Athens, Greece, Print ISBN:978-3-031-47239-8, Online ISBN:978-3-031-47240-4, Vol. 14265, pp. 561–578, Published by Springer, DOI: 10.1007/978-3-031-47240-4_30, Available: https://link.springer.com/chapter/10.1007/978-3-031-47240-4_30.

[39] Jiménez-Ruiz Ernesto, "LogMap Family Participation in the OAEI 2023", in *Proceedings of the 18th International Workshop on Ontology Matching (OM 2023)*, 7 November 2023, Athens, Greece, ISSN: 1613-0073, pp. 157-158, Published by CEUR-WS.org, Available: https://ceur-ws.org/Vol-3591/oaei23_paper4.pdf.

[40] Daniel Faria, Marta C. Silva, Pedro Cotovio, Lucas Ferraz, Laura Balbi *et al.*, "Results for Matcha and Matcha-DL in OAEI 2023", in *Proceedings of the 18th International Workshop on Ontology Matching (OM 2023)*, 1 October 2023, Athens, Greece, pp. 164–169, Published by CEUR-WS.org, Available: https://ceur-ws.org/Vol-3591/oaei23_paper6.pdf.

[41] Sven Hertling and Heiko Paulheim, "OLaLa: Ontology Matching with Large Language Models", in *Proceedings of the 12th International Conference on Knowledge Capture (K-CAP 2023)*, 5-7 December 2023, Pensacola FL, USA, ISBN: 979-8-4007-0141-2, pp. 131–139, Published by ACM, DOI: 10.1145/3587259.3627571, Available: https://dl.acm.org/doi/10.1145/3587259.3627571.

[42] Jomar da Silva, Kate Revoredo, Fernanda Baião and Jérôme Euzenat, "Alin: Improving Interactive Ontology Matching by Interactively Revising Mapping Suggestions", *The Knowledge Engineering Review*, Print ISSN: 0269-8889, Online ISSN: 1469-8005, Vol. 35, 1 January 2020, Article No. e1, Published by Cambridge University Press, DOI: 10.1017/S0269888919000249, Available: https://hal.science/hal-02984949/document.

[43] Abhisek Sharma and Sarika Jain, "LSMatch and LSMatch-Multilingual Results for OAEI 2023", *in Proceedings of the 18th International Workshop on Ontology Matching (OM 2023)*, 7 November 2023, Athens, Greece, ISSN: 1613-0073, pp. 159-163, Published by CEUR-WS.org, Available: https://ceur-ws.org/Vol-3591/oaei23_paper5.pdf.

[44] Zhu Wang, "AMD Results for OAEI 2023", *in Proceedings of the 18th International Workshop on Ontology Matching (OM 2023)*, 7 November 2023, Athens, Greece, ISSN: 1613-0073, pp. 146-153, Published by CEUR-WS.org, Available: https://ceur-ws.org/Vol-3591/oaei23_paper2.pdf.

[45] Philippe Roussille and Olivier Teste, "TOMATO: Results of the 2023 OAEI Evaluation Campaign", in *Proceedings of the 18th International Workshop on Ontology Matching (OM 2023)*, 7 November 2023, Athens, Greece, ISSN: 1613-0073, pp. 191–199, Published by CEUR-WS.org, Available: https://ceur-ws.org/Vol-3591/oaei23_paper10.pdf.

[46] Guilherme Sousa, Rinaldo Lima and Cássia Trojahn, "Results of PropMatch in OAEI 2023", in *Proceedings of the 18th International Workshop on Ontology Matching (OM 2023)*, 7 November 2023, Athens, Greece, ISSN: 1613-0073, pp. 178–183, Published by CEUR-WS.org, Available: https://ceur-ws.org/Vol-3591/oaei23_paper8.pdf.