Social Web in IoT: Can Evolutionary Computation and Clustering Improve Ontology Matching for Social Web of Things?

Asma Belhadi, Djamel Djenouri, Youcef Djenouri, *Senior Member, IEEE*, Ahmed Nabil Belbachir, *Member, IEEE*, Gautam Srivastava, *Senior Member, IEEE*

***Abstract*—Many Internet of Things (IoT) applications can benefit from the Social Web of Things (S-WoT) methods that enable knowledge discovery and help solving interoperability problems. The semantic modeling of the S-WoT is the main emphasis of this work where we suggested a novel solution, ECOM (Evolutionary Clustering for Ontology Matching), to explore correlations between S-WoT data using clustering and evolutionary computation methodologies. The ECOM approach uses a variety of clustering techniques to aggregate S-WoT data's strongly related ontologies into comparable categories. The principle is to match concepts of similar groups rather than the full concepts of the two ontologies, which necessitates to split the examples of each ontology into similar groups. We design two clustering algorithms for ontology matching using conventional methods, as well as sophisticated clustering techniques. Moreover, we develop an intelligent matching algorithm that uses evolutionary computation to quickly converge to (or ideally identify) the optimal matches. Numerous simulations have been conducted using various ontology databases to demonstrate the application and precision of the ECOM. The findings clearly show that ECOM beats cutting-edge ontology matching methods. The F-measure of ECOM exceeds 95% whereas it does not reach 90% for all of the baseline methods. The results also confirm ECOM scale with big data in an S-WoT environment.**

***Index Terms*—Internet of Things, Evolutionary Computation, Clustering, Social Web of Things, Ontology Matching.**

I. Introduction

According to recent market evaluations and studies, the Internet of things (IoT) real economy size is expected to grow rapidly, with the increase of more than four times from 2018 to 2023, i.e., at a compound annual growth rate of 28.4% over the forecast period [1], [2]. IoT represents a dynamic ecosystem where interconnected smart objects communicate and share data to enable a wide range of applications and services [3]–[5], with smart buildings being just one compelling example [6]. The integration of IoT technology in smart buildings enhances energy efficiency, comfort, and security, demonstrating the vast potential of IoT to transform various aspects of our daily lives. Other applications of integrating IoT technology including smart transportation [7], smart grid [8]. In the context of Social Web of Things (S-WoT), a subset smart devices ("things") acquire social behavior and form social relationships with other items to make it possible to work independently towards a common goal [9]–[11].

S-WoT boasts an immense quantity of devices, users, and diverse technologies [12]. This ecosystem facilitates extensive data generation as these devices and users communicate and producing a wealth of heterogeneous data [13]–[16]. Ontologies play a crucial role for structuring extensive datasets that are characterized by multiple relational attributes. They provide the foundational framework for a spectrum of systems, including conversational agents, recommendation engines, semantic search algorithms, and thus enhance the organization and retrieval of complex, interconnected information structures in technological domains [17], [18], and more recently in the domain of S-WoT [19]–[21].

There is a need to comprehend the various interoperability of these ontologies when concepts may be conveyed and represented in different ways by two separate entities [22]–[26]. Moreover, ontology matching becomes essential for semantic modeling and knowledge transfer among users and network devices due to the continuous grow in volume and type of heterogeneous data in S-WoT. The process of determining the semantic correlations between entities, concepts, and relations is called ontology matching. The goal is to establish a mapping between the elements of the different ontologies that are equivalent, related, or overlapping in meaning. Ontology matching is an important task in many domains, including data integration, semantic web, and knowledge management. It enables interoperability between heterogeneous systems, improves data quality, and supports knowledge reuse and discovery.

Numerous ontology matching-based S-WoT systems have been put forth. Haoyu et al. [27] defined interaction patterns between IoT devices and semantic modeling of on-device applications to derive more details about the applications on the device. Wu et al. [28] developed a space-adaptive network convolutional module that could simultaneously search the user interest and social influence propagation mechanisms from S-WoT data. Xingsi et al. [29] offered a formula to calculate the similarity of two words. The technique determines the

A. Belhadi is with OsloMet University, Oslo, Norway, asbel9617@oslomet.no.

D. Djenouri is with CSRS, Computer Science & Creative Technology Dep., University of the West of England, UK, djamel.djenouri@uwe.ac.uk.

Y. Djenouri is with University of South-Eastern Norway, Kongsberg, Norway, Norwegian Research Centre, Oslo, Norway, and IDEAS, NCBR, Warsaw, Poland, youcef.djenouri@usn.no, yodj@norceresearch.no.

A. N. Belbachir is with Norwegian Research Centre, Grimstad, Norway, nabe@norceresearch.no.

Gautam Srivastava is with the Dept. of Math and Computer Science, Brandon University, Canada, and the Research Centre for Interneural Computing, China Medical University, Taichung, Taiwan as well as Dept. of Computer Science and Math, Lebanese American University, Beirut, Lebanon (email: srivastavag@brandonu.ca)

cosine distance between two vectors and models the linguistic features of the word in vector space. Without concentrating on one particular information source, the term ”embedding strategy” also maintains the rich linguistic content importance of phrases.

He et al. [30] introduced the BERTMap ontology alignment system, which offers support for both unsupervised and semi-supervised contexts. The system employs a classifier to predict alignments by adapting the factual embedding BERT model. Subsequently, it utilizes ontology structure and logical principles to expand and refine these alignments, thereby enhancing their precision. The aforementioned solutions require similarity computation among the concepts and properties of the ontologies, which necessitate huge computation and memory resources. This calls for a thorough study of the different correlations and dependencies among the features of the ontologies using parallel computing and evolutionary approaches [31]–[33], which has the potential to intelligently and quickly explore the search space of all possible alignments [34], [35]. This represents the subject of the paper that makes use of the correlations and the evolutionary computation to improve the ontology matching process in S-WoT environment.

A. Motivation

Trivial approaches for comparing two ontology consider all their properties, which necessitates intensive computation time to handle large amount of data. Data mining are employed in order to discover and locate promise and inherent data from large databases that cannot be immediately revealed or determined, e.g., the association between alcohol and diapers. To study the relationship between different data properties, well-known data mining algorithms such as clustering [36], [37] divide the entire data set into comparable groups. Ontologies have also been subject to clustering through the use of description logic, which separate an ontology database into plentiful components created to analyze the connections between the ontologies' relevant concepts [31], [32], [38]. Since they are unable to derive even the smallest components from complex ontologies, the aforementioned approaches are unable to directly matching disparate ontologies. Furthermore, a more sophisticated and time-efficient approach– such as evolutionary computation, is required to reach a scalable solution. Inspired by the success of clustering and evolutionary computation techniques for dealing with a wide range of difficult problems, this research proposes a novel methodology for ontology matching in the context of S-WoT.

B. Contributions

To the best of our knowledge, this study is the first piece of work that examines clustering and evolutionary methods in depth for ontology matching in the context of S-WoT. Its principal contributions are listed below:

1) We create an innovative solution called Evolutionary and Clustering for Ontology Matching (ECOM) to analyze the collection of ontologies. This framework combines evolutionary and clustering techniques. The designed framework can be used to significantly improve the performances of the ontology matching process in S-WoT setting.
2) We design two clustering techniques for ontology matching that explore both traditional and advanced clustering methods, as well as an intelligent matching algorithm. Instead of retrieving all alignments between ontologies, ECOM benefits from evolutionary computation, which allows finding the best alignments in real time.
3) Extensive simulations has been carried out to evaluate ECOM. The findings demonstrated that ECOM outperformed the most complex ontology matching engines.

The remaining of this paper is structured as follows: In Section 2 the subject of ontology matching is explored. The formal formulation of the problem of ontology matching is provided in Section 3. ECOM is presented in Section 4. Section 5 describes the performance assessment of the ECOM. Section 6 anticipates the future developments of the ECOM design. Section 7 concludes the paper.

II. Related Work

A. Semantic Modeling for S-WoT

Rangra et al. [39] considered the problem of identifying communities who are vulnerable to natural disasters and identifying the optimal node where the broadcasting system may be located. They investigated social media for the purpose of spreading alarming messages, as well as linking the WoT with social networks. Zhang et al. [40] created an architecture for recommending smart objects in S-WoT. A BERT with Bi-LSTM network was used in the interest of getting the feature vector for a smart object, as well as the adequate representation of the smart object arrays. Wu et al. [28] suggested a novel item recommendation paradigm. They explored the propagation mechanism of user engagement and impact on society from S-WoT data at the same time. To merge user characterizations from both areas, a gating system is also being developed. Abdelghani et al. [41] proposed a new multilevel trust model that is dynamic and scalable for the S-WoT context. They suggested multidimensional measures to define the S-WoT entities' behaviors. The latter are gathered to use a machine learning-based system that makes it possible for user classification, attack detection, and countermeasures. In order to propagate confidence throughout the system by exploring fewer resources and keeping scalability, a hybrid solution was proposed. The work of Magdish et al. [42] denoted a thorough examination of the effectiveness of confidence attack planning when it is incorporated into the trust model, with the goal of precisely identifying node activities in order to assure safe connections of the S-WoT nodes.

Ren et al. [27] laid out an approach for scalability of on-device app management across disparate IoT devices. The authors demonstrated how the S-WoT can serve to semantically describe the functionality of every IoT system and its interaction patterns. In order to include details about a device's applications, they also established semantic modeling for item description. Corno et al. [43] created a search and recommendation system capable of offering relevant contingent rules for use in a variety of settings predicated on a conceptual user's desire. By specifying a set of fundamental S-WoT functionality, the user can communicate to a conversational agent their current personalization goal. Chen et al. [44] introduces a brand-new subsumption prediction technique called BERTSubs for OWL ontology classes. A class's contextual embeddings are computed using the pre-trained language model BERT, and special templates are suggested to take into account the class context (such as nearby classes) and the logic of the existential restriction. In addition to existential limitations from the same ontology, BERTSubs can also anticipate named subclasses from the same domain or another ontology. Ogunniye et al. [45] focused on detecting and representing confidentiality in the social web of things in a way that helps privacy assistants better understand their environment. In recent years, the focus has increasingly shifted to the technical details of privacy. However, due to the evolving privacy environment, social aspects, such as social trust, also need to be represented. The researchers explored how existing ontologies can be used to represent privacy requirements. They also talked about how these conceptual frameworks can be extended with new standards to effectively capture privacy, and they presented case studies to show how the new requirements can be applied.

B. Ontology Matching

Belhadi et al. [32] developed an ontology matching framework. The optimal qualities for matching ontologies are chosen using data mining techniques in the solution. Furthermore, by looking at various statistical qualities like the mean of the quantities for every data attribute, a novel approach of choosing pertinent elements for the matching has indeed been discovered. Djenouri et al. [46] created a clever framework for matching ontologies in applications related to smart cities. It investigates pattern extraction to unearth insights from ideas in the ontologies that need to be aligned, then unearths the pertinent data to diminish the alignment process' search space. Xue et al. [47] regarded ontology matching as a prediction problem, with the objective of combining a variety of interest group similarity measures to approximate the true matching score. To increase alignment accuracy and matching effectiveness, neural networks are applied.

Mountasser et al. [35] created a big data interoperability which made use of huge ontologies and randomness logic-based evaluation techniques. They also use methods and computing resources in combination with the multicore paradigm (Hadoop/MapReduce) to effectively undertake ontology matching in large scale data scenarios. Lv et al. [48] used distributed learning to create an entirely novel model for ontology matching. Customer feedback is more frequently considered throughout this process of continuous improvement, rather than during every new generation. To alleviate user strain, a roulette wheel technique was employed to offer only the candidate mappings which are most likely the source of problems. Fallatah et al. [49] showed how a string-based combiner and an instance-based classifier may work together. In the former, template matching is transformed into a two-sided text categorization using ontology class examples, while in the latter, pattern matching trees are combined. This method is indeed not domain-specific and therefore can handle diverse and heterogeneous ontologies. Sun et al. [50] used the knowledge graph to complement the driven flow graphs, removing the need for prior pre-processing. To minimize the amount of intermediate possibilities, a multi-label weight matrix was used while inspecting a near-optimal matching tree. This allows navigating the flow graph iteratively to accomplish all isomorphic mappings of subgraphs. Portisch et al. [51] laid out the problem of general ontology matching and examined both sources of background knowledge and methods from the literature that ultimately rely on external knowledge.

C. Discussion

The studies in the domain of ontology matching have consistently demonstrated commendable performance when applied to relatively modest datasets characterized by numerous small and medium-sized concepts, as well as in scenarios involving low-dimensional data. This performance excellence is evident in both the speed of execution (runtime) and the quality of the resultant matching outputs. However, a noticeable challenge emerges when these existing solutions are tasked with the alignment of extensive and complex ontological structures, exemplified by large-scale ontologies like S-WoT. In such cases, the scalability of conventional methods becomes a bottleneck, leading to suboptimal matching quality and significantly prolonged runtime duration. In response to this critical limitation, we propose an innovative approach that leverages a synergistic combination of evolutionary algorithms and advanced clustering techniques. This novel framework is specifically engineered to address the unique challenges posed by substantial and intricate ontological structures, thereby enabling more efficient and effective ontology matching in scenarios where existing methodologies struggle to deliver satisfactory results.

III. Problem Definition

In S-WoT, ontologies are seen as a revolutionary method of arranging and preserving information that IoT devices of the social web ecosystem have exchanged. In general, an ontology defines a domain by breaking it down into concepts and specifying the links between them. The "concept" is the most important component of an ontology in connection with the amount of knowledge. Each concept can be defined using a variety of qualities that constitute concrete data. Knowledge may be overtly kept in the shape of data values as instances. Additionally, the majority of ontologies include more information about things like data types or annotations. An ontology is properly defined as a tuple O = ¡C, R¿ such that, $C = \{C_1, C_2, \dots, C_n\}$ is a set of n concepts, and $R = \{R_1, R_2, \dots, R_m\}$ is the set of relations which connects two different concepts in C.

Example 3.1:

Figure 1 displays information on a small portion of the S-WoT that is organized according to an ontology. Each of the concepts—user, mobile, camera, and location has specific characteristics. For example, the attributes of the concept

TABLE I
TAXONOMY OF EXISTING SYSTEMS FOR MATCHING ONTOLOGIES WITH THEIR CONSTRAINTS.

Class of Models	Models	Limitation
Traditional	Li et al. [52]	Unable to deal with massive data and with a high number of features.
	Shao et al. [53]	
	Rosaci et al. [54]	
Advanced	Belhadi et al. [32]	Use an outdated matching algorithm and require a long runtime for massive data.
	Djenouri et al. [46]	
	Xue et al. [47]	
	Lv et al. [48]	
	Fallatah et al. [49]	

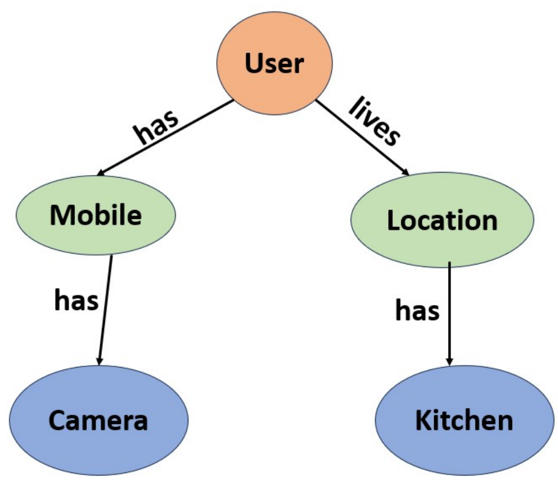

Fig. 1. Illustration of an S-WoT basic ontology. It is made of four concepts $\{user, mobile, location, camera, kitchen\}$, with four relations $\{"has", "lives", "has", "has"\}$.

"user" include things like user name, ID, and job. Four connections exist between the concepts as well. As a sketch, the word "has" links the words "user" and "mobiles," indicating that the "user" owns a mobile.

Finding an alignment between the two ontologies O_1 and O_2 is the goal of ontology matching. It means figuring out the concepts of both ontologies that have the same meaning.

Algorithm 1 ECOM algorithm

Input: ($\mathcal{C}_1$, $\mathcal{C}_2$): Sets of concepts of the ontologies $\mathcal{O}_1$, and $\mathcal{O}_2$. IMAX: maximum number of generations.
Output: $\mathcal{A}^*$: Optimal alignment of the two ontologies $\mathcal{O}_1$, and $\mathcal{O}_2$.
$P_1 \leftarrow Clustering(\mathcal{C}_1)$;
$P_2 \leftarrow Clustering(\mathcal{C}_2)$;
$\mathcal{A} \leftarrow \emptyset$;
for i = 1 to $|P_1|$ **do**
  $min_i \leftarrow \infty$;
  $index_i \leftarrow -1$;
  for j = 1 to $|P_2|$ **do**
    $sim_{ij} \leftarrow similarity(p_i^1, p_j^2)$;
    if $sim_{ij} \leq min_i$ **then**
      $min_i \leftarrow sim_{ij}$;
      $index_i \leftarrow j$;
    end if
  end for
  current_generation $\leftarrow$ 1;
  while current_generation $\leq$ IMAX **do**
    generation $\leftarrow$ initialization(p_i^1, $p_{index_i}^2$);
    extended_generation $\leftarrow$ crossover(generation);
    $extended\_generation \leftarrow extended\_generation \cup mutation(generation)$;
    $generation \leftarrow selection(extended\_generation)$;
    current_generation $\leftarrow$ current_generation + 1;
  end while
  $\mathcal{A} \leftarrow \mathcal{A} \cup \{genetation\}$;
end for
return $\mathcal{A}$.

IV. ECOM: EVOLUTIONARY CLUSTERING FOR ONTOLOGY MATCHING

In this part, the ECOM method for ontology matching and its components is presented.

A. Principle

The full set of instances for each ontology is divided by ECOM into a number of interconnected clusters. The highly connected instances in every group are then processed based on evolutionary computation. As seen in Figure 2, ECOM searches for common features among cluster of concepts. The instance set is split into various clusters during the clustering process with an appropriate number of concepts. The concepts in each cluster are highly associated with one another given the vast number of similar features shared by the concepts in each cluster. ECOM examines the concepts of the clusters to identify the matching based on the evolutionary computation. The alignment process benefits from the generated clusters, while mitigating the exploration of all concepts of both ontologies. By developing a novel alignment strategy (instead comparing two sets of concepts of the given ontologies) the alignment step builds on the clustering phase. This accelerates the matching process. By computing their shortest distances, two highly related clusters across ontologies are located using the matching procedure. As a result, each group in the first ontology aligns with the group in the second one that is the most similar. In this context, we will explore the evolutionary algorithm to quickly converge to the optimal alignments.

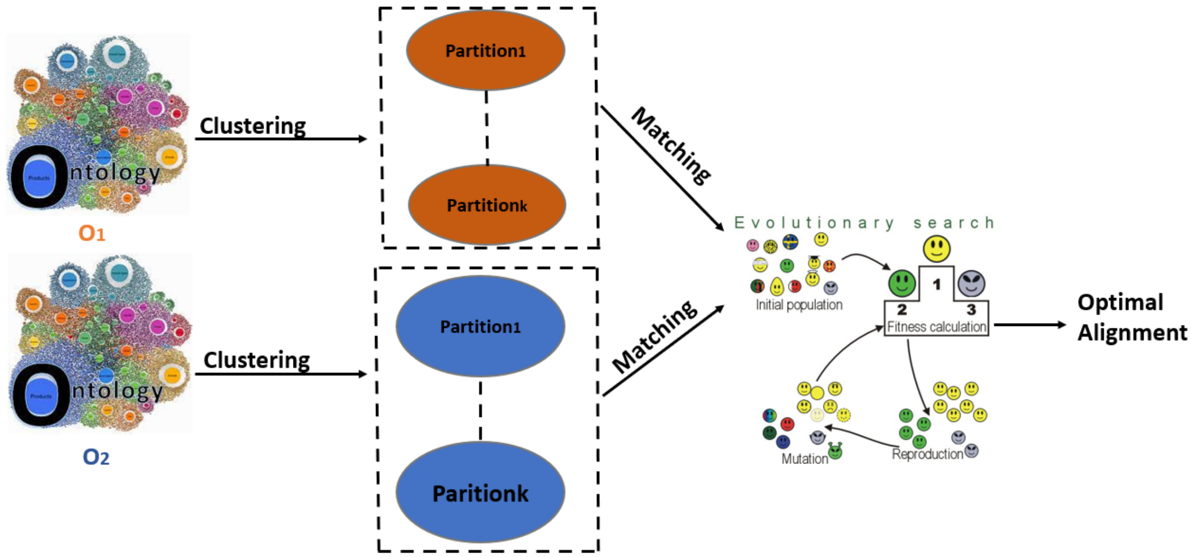

Fig. 2. ECOM framework: the concepts of both ontologies are first decomposed into similar partitions using the clustering method. The set of derived partitions is then explored using an evolutionary process to converge to the best alignments.

Algorithm 1 presents the pseudocode for ECOM a framework designed for aligning concepts from two ontologies, denoted as $\mathcal{O}_1$ and $\mathcal{O}_2$, and subsequently producing matching results. The algorithm's workflow involves several distinct steps:

1. Clustering Preparation (Lines 3-4): At the outset, the concepts from both ontologies are clustered. This clustering is executed through two distinct clustering algorithms aimed at efficiently grouping related concepts. The outcome of this phase is the formation of two sets of clusters, denoted as P_1 for concepts from ontology $\mathcal{O}_1$ and P_2 for concepts from ontology $\mathcal{O}_2$.

2. Cluster Matching (Lines 6-15): To establish correspondences between clusters from the two ontologies, the algorithm evaluates the similarity between each cluster in P_1 and its counterpart in P_2. The cluster from $\mathcal{O}_2$ that most closely resembles each cluster from $\mathcal{O}_1$ is identified and recorded.

3. Cluster Concatenation (Lines 16-24): The algorithm proceeds to match and concatenate the identified clusters. This step ensures that related clusters are aligned and combined, forming more comprehensive sets of matching clusters.

4. Final Matching (Line 26): The ultimate matching results are derived, encapsulating the aligned and concatenated clusters, signifying successful matching between the two ontologies.

In the following, we provide a detailed exposition of the clustering strategies employed in ECOM, as well as an in-depth examination of the framework's matching mechanism.

B. Clustering Step

1) K-bMOM based algorithm: We propose the adaptation of K-bMOM (k-means bootstrap Median-of-Means) [55] for ontology clustering. The method begins by evenly, independently, and with replacement sampling from the original concepts to create the set of blocks from the set of concepts. Afterwards, by tying each concept to its nearest centroid, a partition for each block is calculated. Every block's centroids are updated in accordance with its block partition, and then the empirical risk is determined. The center of the median block, which has the median empirical risk, is chosen as the current block's center. We consider the risk median, which is the empirical real-valued mean of the K-means loss derived from the concepts within every block. Therefore, the bootstrap median-of-means approach is applied. There are multiple iterations of these actions. Instead of getting the centroids of the last iteration's median block, the centroids corresponding to the most recent iterations are aggregated to produce a more accurate assessment of the centroids. This approach produces a codebook supplied consensus-based robust clustering using a set of candidates that are calculated on bootstrap sub-samples. In contrast to existing consensus clustering that aggregates the candidates in a more complicated manner by utilizing some similarity measures between various ways of clustering, we choose one of the candidates using a straightforward median criterion for one-dimension statistics. This is one of the key differences between our approach and that of consensus clustering.

2) DWMB algorithm: We propose the adaptation of DWMB (Divide Well to Merge Better) algorithm [56] for ontology clustering. It is a non-parametric technique that can discover existing clusters in concepts without concerning about the amount of preceding clusters. Similarly to hierarchical clustering, it is based on two paradigms: dividing and merging. This is with a significant distinction in the merging technique, in which the concepts clustered as subclusters in the division phase are employed. This is different from the traditional bottom-up approach that is based on each individual concept. Additionally, the density-based feature of the algorithm is used to determine whether or not to combine two clusters by computing the area where the clusters overlap. Calculating the ideal number of concept clusters for each concept size is the first step in the division process. An enhanced K-means algorithm is utilized to help figuring out the right proportion of sub-clusters for every concept size. The best amount of clusters to use during the division phase is automatically determined with the aid of an optimized version of the K-means algorithm. The division step is prevented from experiencing issues with over-splitting and under-splitting the concepts by the optimized choice of value k. To identify the current concept clusters, the sub-clusters are appraised for fusing. During this phase, all of the sub-clusters discovered during the division phase are fused with nearby sub-clusters. Only two adjacent sub-clusters are reviewed for merging at a time, and the procedure terminates when all the sub-clusters have been evaluated. The sub-clusters are either retained separate as independent clusters or fused together based on the decision of the evaluation.

C. Matching Process

The knowledge gleaned from the clustering step will be used in the matching process. The major goal of this stage is to focus on the most important concepts rather than analyzing both ontologies' whole collections of concepts. Suppose that G_1 and G_2 represent the corresponding sets of clusters for the two ontologies, O_1 and O_2. The groups of G_1 and G_2 are scanned in order to pick similar clusters, and the similarity between each pair of clusters g_i^1 and g_j^2 is determined as follows (EQ. 1):

$$distance(g_i^1, g_j^2) = |g_i^1| + |g_j^2| - |g_i^1 \cap g_j^2|, \tag{1}$$

where $|g_i^1|$, $|g_j^2|$, and $|g_i^1 \cap g_j^2|$ represent the number of properties in the clusters g_i^1 and g_j^2 and their intersection, respectively. The most similar clusters for both ontologies are selected. Let us cal $S(G_1, G_2)$ be the set of most similar clusters of ontologies O_1 and O_2. The set of most similar clusters is explored to retrieve the matching between O_1 and O_2. A naive strategy is to examine each pair of concepts of similar clusters to find the similar concepts in O_1 and O_2. This requires a large amount of time and enormous resources to make the process efficient for immediate processing.

Our conceptual framework revolves around harnessing the evolutionary process to discern analogous clusters within a given dataset. This approach hinges on a comprehensive solution space, encompassing every conceivable alignment between clusters that exhibit similarities. At the core of our methodology lies the concept of a fitness function, which operates as a measure quantifying the quality of each alignment. Our overarching objective centers on cultivating a population of alignments that collectively garners the highest possible alignment scores. The fundamental premise of our approach involves navigating and exploring the pairs of similar clusters within the dataset. To achieve this, we employ a suite of genetic operations tailored for the task, including:

1. Population Initialization: This initial phase lays the foundation for the evolutionary exploration by populating the solution space with a diverse set of candidate alignments.

2. Crossover:, This genetic operation facilitates the exchange of genetic material (in this context, alignment information) between different pairs of clusters, thus diversifying the pool of potential solutions.

3. Mutation: In this step, we introduce controlled variations into the alignments to explore potentially more favorable configurations, promoting adaptability and innovation within the population.

4. Selection: The selection process involves identifying and retaining the alignments with superior alignment scores, thereby emulating the principles of natural selection to propagate promising solutions.

Through these interplay of genetic operations, our approach orchestrates an evolutionary journey to iteratively refine and enhance the alignment of similar clusters, ultimately striving to achieve optimal alignment scores that encapsulate the essence of cluster similarity within the dataset.

D. Discussion of Complexity Reduction Benefits in ECOM

Exploring decomposition and evolutionary algorithms to increase the performance of ontology matching in real-time use offers several potential benefits:

1) Scalability: Existing ontologies on the Web are too large or too complex to be solved in a reasonable amount of time. Dividing them into smaller sub-ontologies can make them more manageable. In addition, exploring the evolutionary algorithm within each sub-ontology can help distribute the workload and reduce the overall complexity of ontology matching.
2) Parallelization: Depending on the nature of the sub-ontologies, it is possible to solve them in parallel, which can lead to significant performance improvements. This is especially true for distributed systems, where each sub-ontology can be assigned to a separate processor or node.
3) Modularity: By decomposing the ontology into smaller sub-ontologies, each sub-ontology can be approached separately and tested independently. This makes it easier to identify errors or bugs in the matching process and allows the entire development process to be more modular and iterative. In addition, the evolutionary algorithm can be easily run on each sub-ontology rather than the entire ontology. This helps in the fast convergence of the optimal alignment.
4) Reusability: If the original ontology is a frequently occurring or recurring ontology, splitting it into smaller sub-ontologies can facilitate the reuse of existing ontology matching algorithms. For example, if each sub-ontology can be solved with a known ontology matching algorithm, the overall solution can be composed from a set of existing building blocks, which can reduce the amount of code that needs to be rewritten.

V. Performance Evaluation

The proposed ECOM was validated by extensive simulations. In this study, two datasets that are commonly used in the field of ontology matching are used:

1) DBpedia[1]: It is part of Wikipedia's hub dataset. This ontology database has $2,795$ unique data concepts and $4,233,000$ occurrences.
2) SIoT dataset [57]: It is designed for modelling the SIoT interactions of more than $50,000$ users connected via different devices including smartphones, cars, tablets, smart watches and others.

A. Performance of Clustering Step

The two clustering methods K-bMoM, DWMB with the k-means algorithm are utilized for comparison. This allows validating the proposed adaptation and select the best clustering algorithm that will be used in the matching process. We used two different evaluation measures. The primary metric, denoted as "connect" (EQ. 2), is employed to quantify the intra-cluster connectivity of conceptual elements. Our objective is to maximize the "connect" metric for each algorithm. This metric uses "sim", (EQ. 3), that computes the similarity between two different partitions. The secondary metric, termed "shared" (EQ. 4), serves the purpose of assessing the degree of conceptual overlap among distinct clusters. Our purpose with the latter is to minimize the "shared" value for each algorithm. The following equations describe these measures:

$$connect(P) = \frac{\sum_{i=1}^{|P|} sim(P_i)}{|P|} \tag{2}$$

where

$$sim(P_i) = \frac{\sum_{j=1}^{|P_i|} distance(e_{ij}, p_i)}{|P_i|} \tag{3}$$

Note that e_{ij} is the j^{th} element of P_i, p_i is the centroid of the partition P_i, distance(e_{ij}, p_i) is the distance between e_{ij}, and p_i.

$$shared(P) = Max(\{share(P_i, P_j) \forall i, j \in [1..|P|^2]\}) \tag{4}$$

where share(P_i, P_j) is the number of shared concepts between P_i and P_j.

Fig. 3 shows the metrics "execution time", "connect" and "shared" when varying the number of clusters. The plots show a slight increase for all algorithms, but with slight differences between them in both datasets for runtime processing. The results also show that K-bMOM provides high connectivity between concepts within clusters and a low number of shared concepts between different clusters. For example, the connectivity of K-bMOM exceeds 27 when using 20 clusters for the DBpedia dataset, while the other algorithms (DWMB and k-means) are below 21 for the same configuration. These results demonstrate the applicability of the clustering algorithms developed in this work. We will therefore use K-bMOM with 20 clusters for the remaining experiments.

[1]http://wiki.dbpedia.org/Datasets

B. Matching Step Performance

The two matching strategies are utilized for comparison, exact vs. evolutionary. The following experiment allows to investigate these two strategies and select the one that matches better for the use in the whole ECOM pipeline. We used coverage measure that computes the coverage of alignments obtained by the evolutionary-based strategy with the alignments of the exact strategy. The aim is to minimize the "coverage" value for the evolutionary based strategy by exploring different population size and different generations. The following equation (EQ. 5) describes the coverage measure:

$$coverage(A) = \frac{|Alignment_A|}{|Alignment_{exact}|} \tag{5}$$

where A is the evolutionary strategy used for a given population size and a given number of generations. $Alignment_A$, $Alignment_{exact}$ are the sets of alignments retrieved by A, and exact strategy, respectively.

Figure 4 portrays the performance outcomes obtained from the matching phase, serving as an empirical evaluation of the algorithm's efficacy. Specifically, the following observations can be discerned from the results:

1. Runtime Behavior: Figure 4 illustrates the runtime behavior of the evolutionary strategy, which exhibits a gradual increase in execution time for both datasets under consideration. This temporal trend suggests that as the algorithm progresses, its computational demands grow at a manageable pace.

2. Matching Coverage Analysis: The matching step's coverage is meticulously assessed and displayed in Figure 4. Across all population configurations and datasets, a discernible augmentation in matching coverage is evident. This enhancement in coverage can be attributed to two primary factors: the population size and the number of generations. Collectively, these factors contribute to the variability in coverage, spanning a range from 40% to 90%, as graphically depicted.

These empirical findings substantiate the practical applicability of the evolutionary strategy, underscoring its superior scalability in comparison to the exact strategy. As a result, in the subsequent experimental investigations, we will adopt the evolutionary strategy with a specific configuration, employing 100 generations and maintaining a population size of 100 individuals as the chosen settings. This choice is motivated by the promising performance observed in this assessment.

C. ECOM vs. State-of-the-art Algorithms

The best configuration of ECOM resulting from the previous experiments (in terms of clustering algorithm and the matching strategy) is contrasted to cutting-edge matching algorithms. The F-measure is used to assess the ontology matching's quality, which produces the output of the alignment, L, and a reference alignment, L^*, as the result of the **ontology matching**. This is explained in the following equations (EQ. 6, EQ. 7 and EQ. 8):

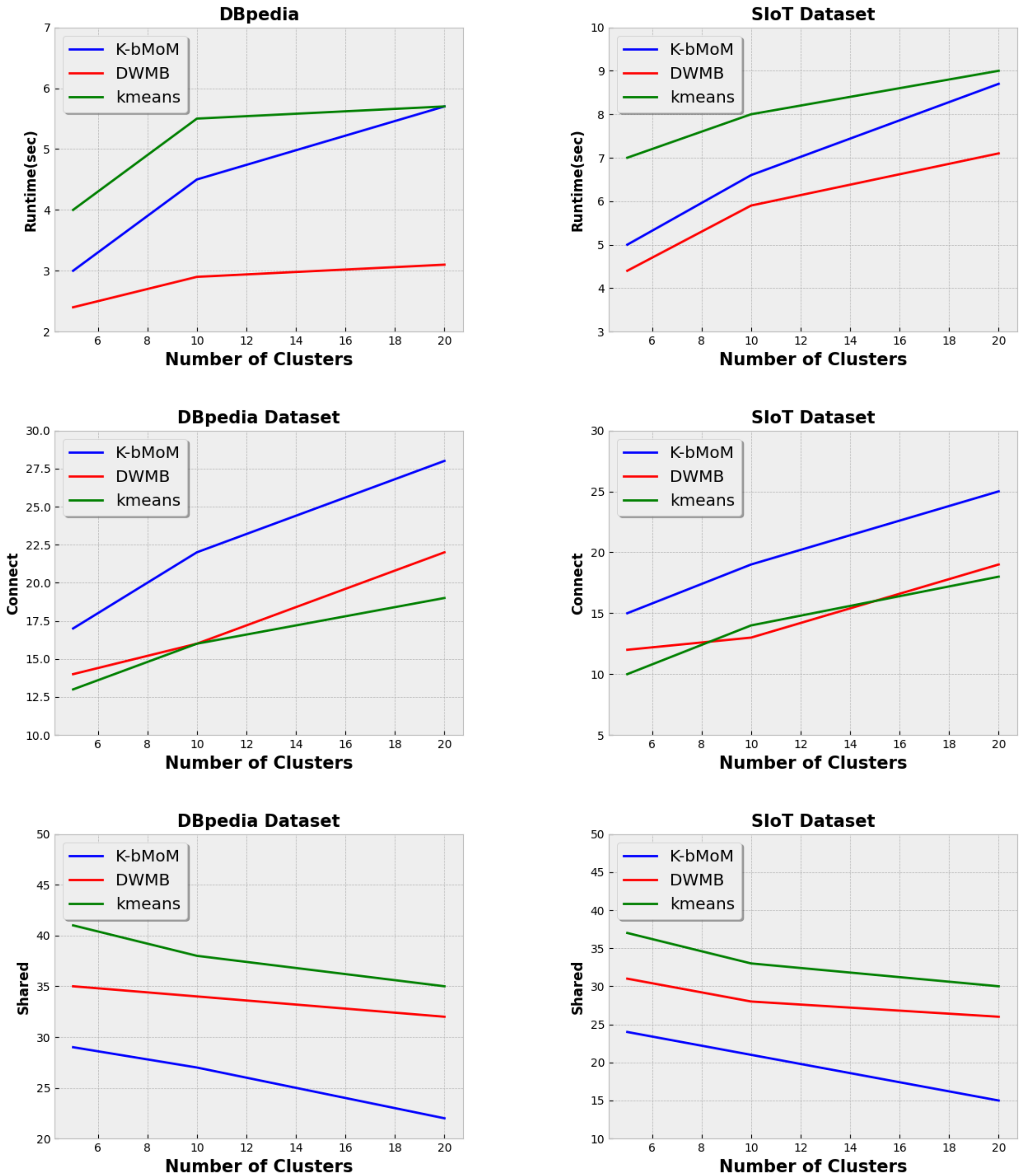

Fig. 3. Performance Results of the Clustering Step

$$F(L, L^*) = \frac{2 \times P(L, L^*) \times R(L, L^*)}{P(L, L^*) + R(L, L^*)}, \tag{6}$$

where the precision "P" is calculated as,

$$P(L, L^*) = \frac{|L^* \cap L|}{|L|}, \tag{7}$$

and the recall, "R", is calculated as,

$$R(L, L^*) = \frac{|L^* \cap L|}{|L^*|}. \tag{8}$$

Notice that the domain experts annotated the ground truth, which is a human-being process. It is represented by the best alignment. In this analysis, DMOM [32] and POMI [46] were taken into account as potential baseline methods. We also consider two variants of ECOM: COM, which examines exact matching and PCOM, which considers the use of particle swarm optimization in the matching process. The performance on both DBpedia and S-IoT was performed in the following two steps:

1) Runtime: Fig. 5 shows the processing time of ECOM, DMOM, POMI, COM, and PCOM using DBpedia and

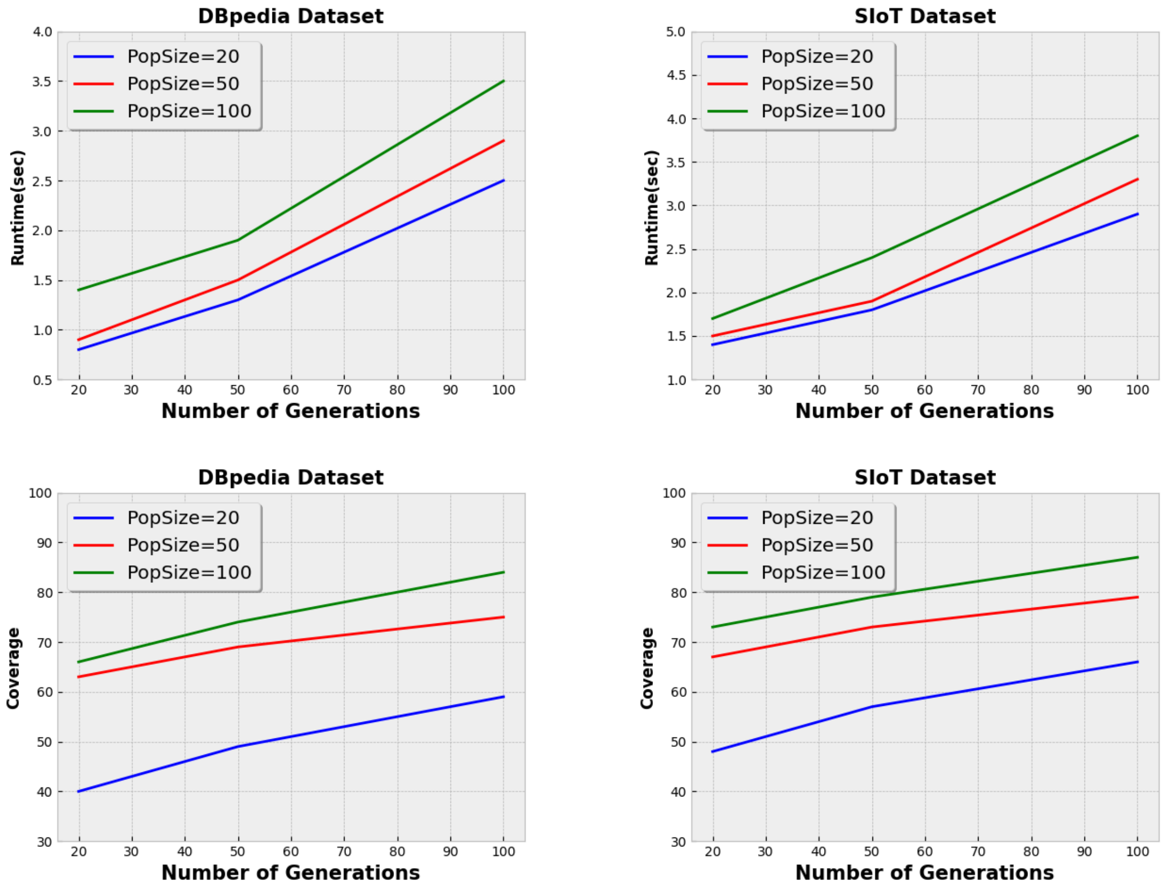

Fig. 4. Results Performance of Matching Step

S-IoT. The results confirm that ECOM performs better than DMOM, POMI, COM and PCOM with a lower runtime, especially for a large number of matches. For example, COM, DMOM and POMI required more than 80 seconds for $80,000$ or $100,000$ matches for both datasets, while ECOM and PCOM required less than 65 seconds. This confirms that by using an effective technique to analyze the information provided in each cluster of instances, the proposed approach provides a matching process that considers only closely related examples.

2) Accuracy: Fig. 5 also evaluates the accuracy in terms of the F-measure of ECOM compared to DMOM, POMI, COM and PCOM. The results show that ECOM and PCOM consistently beat the other two approaches. The results also show that the quality of ECOM and PCOM does not depend on the number of data concepts. As you can notice, the quality of DMOM, POMI and COM is limited to 87% and 90%, respectively, while the quality of ECOM and PCOM never falls below 90%. These results were made possible by using clustering and evolution algorithms to identify the most relevant ontology concepts.

TABLE II
TOP SIMILAR CONCEPTS RETRIEVED BY ECOM USING SIOT DATASET.

Topics	Ontology 1	Ontology 2	Alignment Score
Tourism	Torino City	Italian Town	0.78
	Pizza Food	Popular Meals	0.66
	Museum	Attraction	0.71
	Leonardo da Vinci	Galileo	0.69
	Cheap Hotels	three stars accommodations	0.86
Sport	NBA	Basketball	0.88
	Tony Parker	Team Player	0.91
	World Event	World Cup	0.70
	World Event	Olympic Game	0.85
	Individual Sport	Athletics	0.88
Politics	War	Military Service	0.81
	Negotiation	Parties	0.64
	Debates	Parliament	0.76
	USA	Russia	0.77
	USA	China	0.75

D. *Case Study on S-WoT*

We conducted extensive experiments with S-WIoT ontologies to analyze the results of ECOM in real-world scenarios. We created ontologies using concepts from various S-WIoT related texts. Using the Markov clustering algorithm, we were able to identify all concepts for a given topic. Table II shows some relevant concepts of the ontology matching. Based on the results, we conclude that ECOM is able to discover similar concepts in both ontologies derived from S-WIoT data. For example, ECOM finds that the Olympics is more closely

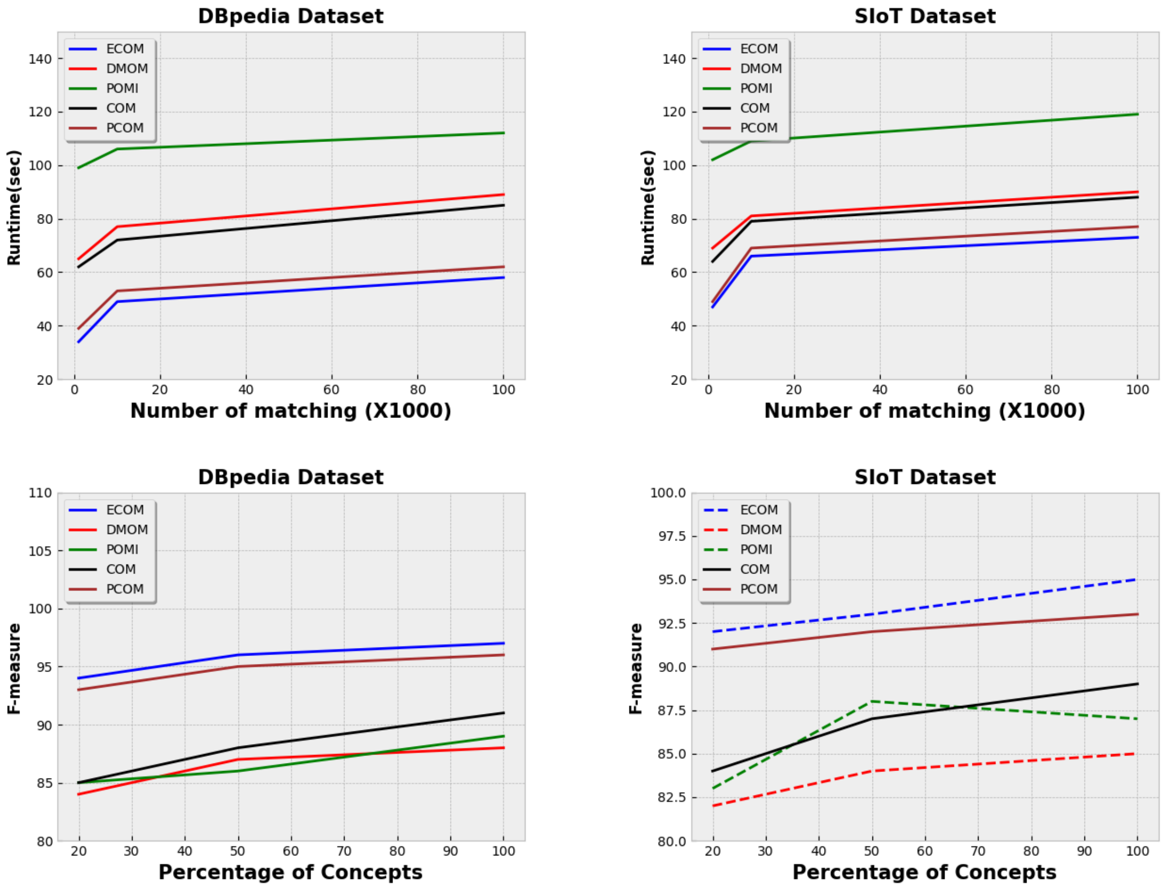

Fig. 5. ECOM Vs. State-of-the-art Algorithms

related to global events than the World Cup. The World Cup is a global event that targets a specific group of people who play a specific sport, while the Olympics is a global event that targets everyone who is interested in sports.

VI. Challenges and Perspectives

In this section, we delve into a comprehensive exploration of various critical issues and propose potential avenues for future development and refinement of the ECOM method. Our focus centers on its application in the context of aligning ontologies derived from S-WoT, which presents a unique set of challenges and opportunities. First and foremost, we observed several pertinent issues that merit careful consideration. These encompass issues related to the scalability and efficiency of the ECOM method when applied to large-scale S-WoT datasets. We delve into the intricacies of handling diverse data sources and the need for robust strategies to accommodate evolving ontological structures. Data heterogeneity within S-WoT is also another challenge of ECOM, emphasizing the importance of devising techniques to handle disparate data representations and semantics. In the following, we will go through a deep exploration of challenges and possible research directions of ECOM:

A. *Real Time Processing*

The challenge of ontology matching takes on a heightened level of complexity when applied to S-WoT, particularly in scenarios where real-time data processing is imperative. The successful alignment of ontologies in the context of S-WoT hinges upon several critical considerations. First and foremost, it demands a profound comprehension of the semantic context surrounding the social data interconnected across the web. This entails recognizing subsumption relations, where one concept encompasses another and ensuring the formal consistency of these relationships. These factors are paramount in achieving accurate and meaningful ontology alignments within the dynamic realm of S-WoT. Moreover, the process of partitioning S-WoT data into coherent and uniform clusters presents a promising directions. It is essential to investigate diverse clustering strategies [58], [59] with evolutionary decision [60]. These studies open intriguing avenues for exploration, offering potential solutions to the intricacies of data organization within S-WoT. Additionally, as S-WoT datasets can grow substantially in size, there is a compelling need to explore parallel processing solutions. The utilization of Graphics Processing Units (GPUs) [61], is a promising direction. GPUs can effectively handle the computational demands associated with large-scale ontology datasets, potentially expediting the ontology matching process in the context of S-WoT.

B. Evaluation

The assessment of core ontologies for S-WoT data has been a relatively neglected area within the research landscape. There has been a limited focus on systematically evaluating the utility and effectiveness of core ontologies in this context. To foster progress and innovation in semantic matching approaches for S-WoT environments, it is crucial to promote the creation of comprehensive evaluation datasets specifically designed for testing ontology matching solutions. Encouragingly, these datasets could serve as a catalyst for the advancement of approaches that leverage foundational ontologies. By providing a standardized benchmark for assessment, researchers and practitioners would have a common ground for evaluating and refining their semantic matching methods in the context of S-WoT data. One notable gap in the existing literature is the scarcity of publicly available alignments generated by various ontology matching techniques. This limited availability hinders the reproducible and comparison of results across different approaches. Furthermore, the formats chosen for these alignments often lack compatibility with automated processing, making it challenging to conduct systematic evaluations. In response to these challenges, there is a clear intention to contribute to the field by establishing a benchmarking framework for ontology matching in the S-WoT context. This initiative will involve the creation of evaluation datasets and the definition of specific evaluation metrics to assess the quality of matching results comprehensively.

C. ECOM Applicability

The capacity of the current smart city is being exceeded by urbanization and gentrification, which is increasingly causing environmental deterioration. Semantic analysis of smart city data and related events can be used to maximize the usage of the infrastructure. With this regard, making cities and communities ecological is aligned with the United Nations' 11th Sustainable Development Goal, and ECOM has the potential to contribute to its achievement. Energy demand is rising worldwide, which is bad for the environment and detrimental for the lives of individuals. Some potential areas where ECOM can be applied are described in the following:

1) **Urban Planning** Cities around the world are expanding quickly and their population is expected to reach 2.5 billion by 2050. To balance the competing needs for housing and to manage external shocks, it is crucial to understand urban dynamics [62], [63]. Since the amount and diversity of data from multiple sources in smart cities are expanding, ECOM has an essential function in conceptual modelling of smart cities to enhance the skill of urban planning.
2) **Smart Manufacturing** A growing number of sensors are being added to machines that may generate a significant quantity of time series data, which is a step further in the implementation of Industry 4.0. However, it is challenging to detect deficiencies in the production process, e.g., when the use of a worn tool leads to the production of defective parts or if a manufacturing procedure is prone to polluting the environment. ECOM can be used to connect and comprehend the various factors in production environments [64], [65]. It can be utilized to connect and comprehend the different production-related factors and then convey a variety of industrial production process behaviors, like alarms or sensor failures.
3) **Smart Healthcare** Health care is typically funded by governments in most countries, which requires the provision of decent treatments and services at the most affordable prices. This can only be achieved through the employment of suitable policies and technologies, particularly among health insurers [66], [67]. A precise information management system that tracks clients' healthcare demands in line with their health state can be created with the help of ECOM.

VII. Conclusion

This paper addresses the issue of ontology matching in the scope of S-WoT, and a hybrid evolutionary clustering solution has been proposed. The proposed framework (ECOM) aggregates ontologies that are strongly linked from S-WoT data into comparable categories using a number of clustering approaches. The core idea is to group the examples of each domain into related groups before matching concepts of similar groups rather than whole concepts of the two ontologies. Two clustering algorithms were created for ontology matching. These algorithms initially look into both basic and advanced clustering methods. We also developed an intelligent matching algorithm where the best alignments are optimized or idealized through an evolutionary process. The findings demonstrate that while keeping the same matching quality, ECOM beats cutting-edge ontology matching techniques in terms of computing cost. These outcomes also demonstrate ECOM's ability to manage various data types in S-WoT scenarios. In the near future, we plan to adapt ECOM for matching multiple ontologies instead of two ontologies. This might be useful to understand the behavior of multi-agent systems in the context of the social web of IoT.

Acknowledgments

This work is supported by the project ULEARN "Unsupervised Life-long Learning" and co-funded under the grant number 316080 of the Research Council of Norway.

References

[1] P. Tiwari, A. Lakhan, R. H. Jhaveri, and T.-M. Gronli, "Consumer-centric internet of medical things for cyborg applications based on federated reinforcement learning," *IEEE Transactions on Consumer Electronics*, 2023.

[2] B. Cheng, D. Zhu, S. Zhao, and J. Chen, "Situation-aware iot service coordination using the event-driven soa paradigm," *IEEE Transactions on Network and Service Management*, vol. 13, no. 2, pp. 349–361, 2016.

[3] A. Gyrard, A. Zimmermann, and A. Sheth, "Building iot-based applications for smart cities: How can ontology catalogs help?" *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3978–3990, 2018.

[4] V. M. Tayur and R. Suchithra, "Multi-ontology mapping generative adversarial network in internet of things for ontology alignment," *Internet of Things*, vol. 20, p. 100616, 2022.

[5] H. Jiang, Z. Xiao, Z. Li, J. Xu, F. Zeng, and D. Wang, "An energy-efficient framework for internet of things underlaying heterogeneous small cell networks," *IEEE Transactions on Mobile Computing*, vol. 21, no. 1, pp. 31–43, 2020.

[6] D. Djenouri, R. Laidi, Y. Djenouri, and I. Balasingham, "Machine learning for smart building applications: Review and taxonomy," *ACM Comput. Surv.*, vol. 52, no. 2, pp. 24:1–24:36, 2019.

[7] Y. Djenouri, A. Belhadi, J. C. Lin, D. Djenouri, and A. Cano, "A survey on urban traffic anomalies detection algorithms," *IEEE Access*, vol. 7, pp. 12 192–12 205, 2019.

[8] P. et al., "Federated learning for smart cities: A comprehensive survey," *Sustainable Energy Technologies and Assessments*, vol. 55, p. 102987, 2023.

[9] C. et al., "Social web of things: a survey," in *ICPDS*. IEEE, 2013, pp. 570–575.

[10] L. Li, X. Wu, M. Kong, J. Liu, and J. Zhang, "Quantitatively interpreting residents happiness prediction by considering factor–factor interactions," *IEEE Transactions on Computational Social Systems*, 2023.

[11] Q. Ni, J. Guo, W. Wu, and H. Wang, "Influence-based community partition with sandwich method for social networks," *IEEE Transactions on Computational Social Systems*, vol. 10, no. 2, pp. 819–830, 2022.

[12] B. Cheng, M. Wang, S. Zhao, Z. Zhai, D. Zhu, and J. Chen, "Situation-aware dynamic service coordination in an iot environment," *IEEE/ACM Transactions On Networking*, vol. 25, no. 4, pp. 2082–2095, 2017.

[13] S. et al., "Web of things: security challenges and mechanisms," *IEEE Access*, vol. 9, pp. 31 695–31 711, 2021.

[14] N. et al., "Semantic interoperability in the iot: extending the web of things architecture," *ACM TIOT*, vol. 1, no. 1, pp. 1–25, 2020.

[15] L. Ma, X. Wang, M. Huang, Z. Lin, L. Tian, and H. Chen, "Two-level master–slave rfid networks planning via hybrid multiobjective artificial bee colony optimizer," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 5, pp. 861–880, 2017.

[16] W. Nie, Y. Bao, Y. Zhao, and A. Liu, "Long dialogue emotion detection based on commonsense knowledge graph guidance," *IEEE Transactions on Multimedia*, 2023.

[17] C. et al., "A proposal for an adaptive recommender system based on competences and ontologies," *ESWA*, p. 118171, 2022.

[18] S. et al., "Engineering uncertain time for its practical integration in ontologies," *Knowledge-Based Systems*, p. 109152, 2022.

[19] J. et al., "Ontology-based e-learning content recommender system for addressing the pure cold-start problem," *ACM JDIQ*, vol. 13, no. 3, pp. 1–27, 2021.

[20] X. Zhou, Q. Lv, and A. Geng, "Matching heterogeneous ontologies based on multi-strategy adaptive co-firefly algorithm," *Knowledge and Information Systems*, vol. 65, no. 6, pp. 2619–2644, 2023.

[21] S. et al., "Domain-independent entity coreference for linking ontology instances," *Journal JDIQ*, vol. 4, no. 2, pp. 1–29, 2013.

[22] G. et al., "Ontology-based data quality management for data streams," *Journal JDIQ*, vol. 7, no. 4, pp. 1–34, 2016.

[23] C. Touati and A. Kemmar, "Deep reinforcement learning approach for ontology matching problem," *International Journal of Data Science and Analytics*, pp. 1–16, 2023.

[24] F. Şentürk and V. Aytac, "A graph-based ontology matching framework," *New Generation Computing*, pp. 1–19, 2023.

[25] C. et al., "Zero-shot text classification via knowledge graph embedding for social media data," *IEEE IoT*, 2021.

[26] Q. Ni, J. Guo, W. Wu, H. Wang, and J. Wu, "Continuous influence-based community partition for social networks," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 3, pp. 1187–1197, 2021.

[27] R. et al., "Towards semantic management of on-device applications in industrial iot," *ACM TOIT*, 2022.

[28] W. et al., "Eagcn: An efficient adaptive graph convolutional network for item recommendation in social internet of things," *IEEE IoT*, 2022.

[29] X. et al., "Word embedding based heterogeneous entity matching on web of things," in *Web Conference*, 2022, pp. 941–947.

[30] H. et al., "Bertmap: A bert-based ontology alignment system," in *AAAI*, vol. 36, no. 5, 2022, pp. 5684–5691.

[31] H. Belhadi, K. Akli-Astouati, Y. Djenouri, and J. C.-W. Lin, "Exploring pattern mining for solving the ontology matching problem," in *European Conference on Advances in Databases and Information Systems*. Springer, 2019, pp. 85–93.

[32] ——, "Data mining-based approach for ontology matching problem," *Applied Intelligence*, vol. 50, no. 4, pp. 1204–1221, 2020.

[33] L. Ma, Y. Liu, G. Yu, X. Wang, H. Mo, G.-G. Wang, Y. Jin, and Y. Tan, "Decomposition-based multiobjective optimization for variable-length mixed-variable pareto optimization and its application in cloud service allocation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023.

[34] H. Belhadi, K. Akli-Astouati, Y. Djenouri, J. C.-W. Lin, and J. M.-T. Wu, "Gfsom: Genetic feature selection for ontology matching," in *International Conference on Genetic and Evolutionary Computing*. Springer, 2018, pp. 655–660.

[35] M. et al., "Semantic-based big data integration framework using scalable distributed ontology matching strategy," *DPD*, vol. 39, no. 4, pp. 891–937, 2021.

[36] J. C.-W. Lin, Y. Djenouri, G. Srivastava, U. Yun, and P. Fournier-Viger, "A predictive ga-based model for closed high-utility itemset mining," *Applied Soft Computing*, vol. 108, p. 107422, 2021.

[37] Y. Peng, Y. Zhao, and J. Hu, "On the role of community structure in evolution of opinion formation: A new bounded confidence opinion dynamics," *Information Sciences*, vol. 621, pp. 672–690, 2023.

[38] J. Dong, J. Hu, Y. Zhao, and Y. Peng, "Opinion formation analysis for expressed and private opinions (epos) models: Reasoning private opinions from behaviors in group decision-making systems," *Expert Systems with Applications*, p. 121292, 2023.

[39] R. et al., "Natural disasters management using social internet of things," *Multimedia Tools and Applications*, pp. 1–15, 2022.

[40] Z. et al., "Smart objects recommendation based on pre-training with attention and the thing–thing relationship in social internet of things," *FGCS*, vol. 129, pp. 347–357, 2022.

[41] A. et al., "Dynamic and scalable multi-level trust management model for social internet of things," *The Journal of Supercomputing*, vol. 78, no. 6, pp. 8137–8193, 2022.

[42] M. et al., "A resilient trust management framework towards trust related attacks in the social internet of things," *Computer Communications*, vol. 191, pp. 92–107, 2022.

[43] C. et al., "From users intentions to if-then rules in the internet of things," *ACM TOIS*, vol. 39, no. 4, pp. 1–33, 2021.

[44] J. Chen, Y. He, Y. Geng, E. Jiménez-Ruiz, H. Dong, and I. Horrocks, "Contextual semantic embeddings for ontology subsumption prediction," *World Wide Web*, pp. 1–23, 2023.

[45] O. et al., "A survey on understanding and representing privacy requirements in the internet-of-things," *JAIR*, vol. 76, pp. 163–192, 2023.

[46] Y. Djenouri, H. Belhadi, K. Akli-Astouati, A. Cano, and J. C.-W. Lin, "An ontology matching approach for semantic modeling: A case study in smart cities," *Computational Intelligence*, 2021.

[47] X. et al., "Artificial neural network based sensor ontology matching technique," in *Web Conference*, 2021, pp. 44–51.

[48] L. et al., "A novel periodic learning ontology matching model based on interactive grasshopper optimization algorithm," *Knowledge-Based Systems*, vol. 228, p. 107239, 2021.

[49] F. et al., "A hybrid approach for large knowledge graphs matching," in *OM*. CEUR-WS, 2021.

[50] S. al., "A subgraph matching algorithm based on subgraph index for knowledge graph," *Frontiers of Computer Science*, vol. 16, no. 3, pp. 1–18, 2022.

[51] P. et al., "Background knowledge in ontology matching: A survey," *Semantic Web*, no. Preprint, pp. 1–55, 2022.

[52] L. et al., "Large scale instance matching via multiple indexes and candidate selection," *Knowledge-Based Systems*, vol. 50, pp. 112–120, 2013.

[53] S. et al., "Rimom-im: A novel iterative framework for instance matching," *Journal of computer science and technology*, vol. 31, pp. 185–197, 2016.

[54] D. Rosaci, "Finding semantic associations in hierarchically structured groups of web data," *Formal Aspects of Computing*, vol. 27, pp. 867–884, 2015.

[55] B.-S. et al., "K-bmom: A robust lloyd-type clustering algorithm based on bootstrap median-of-means," *Computational Statistics & Data Analysis*, vol. 167, p. 107370, 2022.

[56] R. et al., "Divide well to merge better: A novel clustering algorithm," *Pattern Recognition*, vol. 122, p. 108305, 2022.

[57] M. et al., "How to exploit the social internet of things: Query generation model and device profiles' dataset," *Computer Networks*, vol. 174, p. 107248, 2020.

[58] V. et al., "Density-based ifcm along with its interval valued and probabilistic extensions, and a review of intuitionistic fuzzy clustering methods," *Artificial Intelligence Review*, pp. 1–41, 2022.

[59] D. et al., "K-centroid link: a novel hierarchical clustering linkage method," *Applied Intelligence*, vol. 52, no. 5, pp. 5537–5560, 2022.

[60] J. Tan, H. Jin, H. Hu, R. Hu, H. Zhang, and H. Zhang, "Wf-mtd: Evolutionary decision method for moving target defense based on

wright-fisher process," *IEEE Transactions on Dependable and Secure Computing*, 2022.

[61] Y. Djenouri, A. Bendjoudi, M. Mehdi, N. Nouali-Taboudjemat, and Z. Habbas, "Gpu-based bees swarm optimization for association rules mining," *The Journal of Supercomputing*, vol. 71, no. 4, pp. 1318–1344, 2015.

[62] U. Ahmed, G. Srivastava, Y. Djenouri, and J. C.-W. Lin, "Knowledge graph based trajectory outlier detection in sustainable smart cities," *Sustainable Cities and Society*, vol. 78, p. 103580, 2022.

[63] Y. Djenouri, A. Belhadi, J. C.-W. Lin, and A. Cano, "Adapted k-nearest neighbors for detecting anomalies on spatio–temporal traffic flow," *IEEE Access*, vol. 7, pp. 10 015–10 027, 2019.

[64] Y. Djenouri, G. Srivastava, and J. C.-W. Lin, "Fast and accurate convolution neural network for detecting manufacturing data," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2947–2955, 2020.

[65] A. Belhadi, Y. Djenouri, G. Srivastava, and J. C.-W. Lin, "Reinforcement learning multi-agent system for faults diagnosis of mircoservices in industrial settings," *Computer Communications*, vol. 177, pp. 213–219, 2021.

[66] Y. Djenouri, A. Belhadi, G. Srivastava, and J. C.-W. Lin, "Secure collaborative augmented reality framework for biomedical informatics," *IEEE Journal of Biomedical and Health Informatics*, 2021.

[67] Y. Djenouri, A. Belhadi, A. Yazidi, G. Srivastava, and J. C.-W. Lin, "Artificial intelligence of medical things for disease detection using ensemble deep learning and attention mechanism," *Expert Systems*, p. e13093, 2022.