Article

Not peer-reviewed version

Semantic Conflict Resolution for Collaborative Earth Observation via Ontology Aware Entity Alignment

Ying Ouyang , Jiahao Zhao , Yida Wang , Zihan Wang , Zewei Jing , Bobo Xi *

Posted Date: 21 May 2026

doi: 10.20944/preprints202605.1303.v1

Keywords: entity alignment; earth observation; intent-driven networking; space-terrestrial integrated networks

Article

Semantic Conflict Resolution for Collaborative Earth Observation via Ontology Aware Entity Alignment

Ying Ouyang [1], Jiahao Zhao [1], Yida Wang [2], Zihan Wang [1], Zewei Jing [3] and Bobo Xi [1,2,*]

1 School of Telecommunications Engineering, Xidian University, Xi'an 710071, China
2 Hangzhou Institute of Technology, Xidian University, Hangzhou 311231 , China
3 Guangzhou Institute of Technology, Xidian University, Guangzhou 510555, China
* Correspondence: xibobo1301@foxmail.com.

Abstract

Space-terrestrial integrated networks (STIN) are increasingly adopted to support time-sensitive missions such as collaborative monitoring and remote sensing, where user intents must be translated into actionable configurations across satellite and terrestrial segments. However, the pronounced heterogeneity in network capabilities, management vocabularies, and domain knowledge bases introduces severe semantic conflicts and knowledge discrepancies, making endogenous intent refinement unreliable and difficult to generalize. To address these limitations, this paper proposes a novel ontology-enhanced intent refinement framework that synergizes offline unified ontology construction with online real-time intent translation. Additionally, we develop an entity alignment model that autonomously identifies and resolves semantic conflicts by learning structural and attributive correspondences across heterogeneous knowledge graphs. Evaluation results based on various datasets and runtime comparisons demonstrate that the proposed framework achieves transferable and reusable endogenous intent governance under multi-vendor and cross-management domain conditions, ensuring end-to-end consistency and efficiency for automated orchestration in STIN-enabled remote sensing and monitoring applications.

Keywords: entity alignment; earth observation; intent-driven networking; space-terrestrial integrated networks

---

1. Introduction

Space-terrestrial integrated networks (STINs) represent a significant advancement in modern remote sensing. These systems enable a transition from isolated satellite observations to dynamic and multi-tier synergistic sensing [1,2]. However, coordinating heterogeneous resources across space and ground segments remains a major challenge. Users must translate high-level observation objectives into executable network policies for complex sensing tasks. Current systems often fail to bridge the operational gap between diverse monitoring platforms. Intent-driven networking (IDN) has emerged as a critical enabler for this autonomy, allowing operators to define high-level monitoring goals, such as "adaptive regional sampling" or "emergency disaster tracking", without manually configuring the underlying infrastructure [3–5]. To realize this vision, the disparate operational logic of space, air, and ground domains must be bridged through an intelligent refinement mechanism that converts abstract user intents into executable network policies. Recent advancements have significantly propelled the deployment of IDN. For instance, authors in [6] utilizes generative AI to bridge the gap between ambiguous user directives and wireless configurations, establishing an interactive paradigm for natural language intent interpretation. 6G-INTENSE integrates native AI at the deep edge to support compute-network abstraction, enabling unified intent orchestration across heterogeneous sensing and computation resources in 6G environments [7]. Despite the promise of intent-driven collaborative monitoring, current network operation and maintenance frameworks face significant challenges in accurately translating intents across heterogeneous domains [8,9].

 doi:10.20944/preprints202605.1303.v1

This issue is particularly critical within the broader context of Space–Terrestrial Integrated Networks (STIN), where autonomy relies on endogenous intent which defined as the network's intrinsic capability to autonomously generate objectives for network state maintenance and fault recovery. Specifically, existing methods struggle with three main hurdles:

- **Semantic Heterogeneity in Multi-Vendor Environments**: The inherent semantic gap between satellite and terrestrial segments, exacerbated by multi-vendor equipment and conflicting knowledge bases [10,11], often leads to execution conflicts where identical commands are interpreted differently across proprietary interfaces.
- **Deficiencies in Cross-Domain Orchestration Modeling**: Current frameworks lack a robust space-terrestrial ontology model referenced against TeleManagement Forum (TMF) standards, making it difficult to map high-level orchestration logic onto the physical constraints of dynamic satellite links and static ground nodes.
- **Limitations of Syntactic Refinement and Entity Alignment**: Most methods rely on superficial syntactic translation while neglecting deep semantic consistency, resulting in "literal" errors that fail to capture contextual nuances and necessitating advanced entity alignment algorithms to ensure accurate policy generation [12,13].

To effectively address the challenges of multi-vendor semantic heterogeneity, the absence of cross-domain orchestration modeling [14], and the limitations of traditional syntactic translation in space-air-ground collaborative monitoring, this paper proposes a dual-mode intent refinement framework based on an "offline and online" synergistic mechanism, which integrates an ontology-aware entity alignment model. In the offline phase, the framework constructs a unified space-terrestrial resource ontology repository referencing TMF standards, establishing a standardized model that encompasses the physical constraints of dynamic satellite nodes and ground sensing facilities. In the online phase, the core entity alignment model leverages a dual-layer deep matching mechanism across Schema and Instance layers to accurately capture implicit semantic associations and contextual nuances among diverse observational devices. This design not only eliminates instruction execution ambiguities arising from heterogeneous interfaces but also ensures the semantic consistency of monitoring tasks during refinement via alignment loss optimization, ultimately accurately converting high-level observation intents into executable cross-domain collaborative policies. The novel contributions of the paper are listed as follows:

- **We propose an ontology-based endogenous intent refinement framework for STIN**: In this framework, we innovatively design a dual-mode synergy mechanism comprising "offline unified ontology construction and online real-time refinement." By standardizing intent class and attribute hierarchies and employing a real-time parsing engine, we effectively bridge the semantic channel from abstract business requirements to underlying physical resources, achieving a semantic closed-loop within heterogeneous contexts;
- **We design a heterogeneous entity alignment model to resolve semantic conflicts**: Addressing challenges such as homonymy and discrepancies in attribute granularity across multi-vendor devices, we utilize relation-aware dynamic mapping and joint embedding learning. Through this approach, we achieve precise cross-domain entity matching and automatic conflict resolution, thereby fundamentally suppressing the propagation of semantic ambiguity throughout the refinement pipeline;
- **We construct a multi-domain, multi-vendor STIN simulation environment for comprehensive evaluation**: Through extensive experiments, we demonstrate that our proposed scheme outperforms mainstream baselines in terms of intent refinement accuracy, online efficiency, and cross-domain transferability. Furthermore, we validate the excellent reusability of semantic assets and confirm the framework's practical potential for deployment in complex engineering environments.

The remainder of this paper is organized as follows. Section II reviews the related work concerning intent-driven network management, the application of ontology in network management, and existing entity alignment methods. Following this, Section III details the ontology-based intent modeling approach, introducing the dual-mode intent refinement framework which includes its offline and online modes and the specific intent refinement process. We then present the ontology-aware entity alignment mechanism in Section IV, which covers the problem formulation and elaborates on the model design, including the entity alignment model, ontology alignment model, and alignment loss model. Section V provides the simulation results and a comprehensive performance evaluation. Finally, Section VI discusses future challenges and Section VII concludes this paper.

2. Related Work

2.1. Intent-Driven Autonomous Collaborative Monitoring

IDN represent a paradigm shift in network management by allowing operators to define high-level business goals instead of manual configurations [15]. Mainstream standardization organizations have established foundational norms for IDN to facilitate the paradigm shift from imperative configuration to declarative management [16–18]. 3GPP has focused on vertical resource scheduling for mobile networks [19], TMF has emphasized business-layer orchestration and cross-domain interoperability [20], and the IETF has defined the general conceptual model and lifecycle for intents [21]. Together, these standards constitute a full-dimensional system covering business, management, and network elements. However, despite these advancements, current intent refinement methods exhibit significant limitations when applied to the multi-source heterogeneous reality of STIN. Existing frameworks primarily focus on parsing the syntactic structure of intents but lack the capability to bridge the inherent semantic gap between diverse monitoring platforms. They often fail to address the execution conflicts caused by multi-vendor equipment and conflicting knowledge bases. Consequently, most methods cannot effectively support the “endogenous intent" required for autonomy, leading to difficulties in constructing actionable refinement links that span space, air, and ground segments without manual intervention.

In the context of remote sensing, existing IDN solutions often struggle with the dynamic nature of collaborative monitoring [22,23]. Most current systems rely on predefined mission templates, which cannot adapt to the rapid changes in observation requirements or hardware availability. Furthermore, the lack of a standardized interface between space-based observation platforms and terrestrial management systems leads to significant delays in mission execution.

2.2. Ontology-Based Semantic Modeling for Remote Sensing

Ontologies serve as a vital semantic layer in integrating heterogeneous data from diverse remote sensing sources [24,25]. Ontology modeling aims to precisely describe and structurally characterize conceptual entities, attribute features, and intrinsic logical relationships in the network domain through formal syntax and semantic specifications. It often relies on It often relies on the Resource Description Framework (RDF), RDF Schema (RDFS), and the Web Ontology Language (OWL) to construct a tiered expression system... to construct a tiered expression system ranging from basic data schemas to complex logical constraints [26,27]. In the evolution of network management automation, ontologies not only serve as a “semantic lingua franca" in multi-vendor heterogeneous environments, masking underlying device differences through standardized class and attribute systems to enable the asset-based sharing and cross-domain reuse of operational knowledge. More critically, they provide a computable logical foundation for intelligent systems. This foundation supports conflict detection and policy optimization based on semantic reasoning [28], which is capable of accurately translating abstract endogenous intents into network-executable configuration parameters, thereby ensuring full-process consistency and explainability from intent perception to policy deployment.

While terrestrial monitoring networks prioritize local temporal resolution, satellite systems focus on broad spatial coverage. Existing information models rarely capture the complex spatio-

 doi:10.20944/preprints202605.1303.v1

temporal correlations required for deep space-terrestrial collaboration. The manual construction of these ontologies is also labor-intensive and fails to keep pace with the rapid evolution of remote sensing technologies.

2.3. Entity Alignment in Multi-Source Monitoring Fusion

Entity alignment is critical for integrating knowledge from multi-source heterogeneous graphs by identifying equivalent entities across different datasets [29]. Traditional approaches primarily relied on linguistic features or superficial syntactic translation [30], which proved insufficient for cross-domain scenarios characterized by significant semantic heterogeneity. To improve accuracy, embedding-based methods, such as TransE and its variants, were introduced to map entities into low-dimensional vector spaces to capture structural similarities [31]. Despite their effectiveness in general knowledge graphs(KGs) [32], these methods often fall short in complex network management tasks involving multi-vendor environments. They typically neglect the deep semantic consistency required for accurate policy generation and fail to account for the discrepancies in attribute granularity across devices [33]. This limitation leads to “literal" errors that fail to capture contextual nuances, necessitating advanced models capable of dual-layer matching to suppress semantic ambiguity effectively [34].

However, the application of entity alignment in space–terrestrial monitoring faces unique challenges. Data from satellites and ground sensors exhibit extreme structural heterogeneity and scale differences. Direct entity mapping without considering these semantic hierarchies often results in class conflicts and false alignments. Therefore, integrating ontology-based semantic constraints into the alignment process is essential for reliable collaborative sensing.

3. Ontology-Based Intent Modeling

In this section, we present the proposed solution for bridging the semantic gap in STIN. First, we introduce the dual-mode intent refinement framework, which integrates offline knowledge alignment with online intent refinement to handle multi-vendor heterogeneity. Subsequently, we detail the ontology-based intent refinement process, explaining how abstract user intents are systematically mapped to executable network policies through semantic reasoning and conflict resolution.

3.1. Dual-Mode Intent Refinement Framework

To mitigate the complexities of cross-domain resource orchestration and the lack of semantic interoperability in STIN-based collaborative monitoring, we propose a dual-mode intent refinement framework, as shown in Figure 1. To facilitate the seamless integration of offline semantic unification and online intent execution, the proposed framework is organized into a three-tier hierarchical architecture. This structure ensures a decoupling between high-level mission objectives and the complex, heterogeneous substrate of STIN.

- **Business Application Layer**: The uppermost layer is the business application layer, serving as the interface for user-network interaction. In collaborative monitoring scenarios, this layer abstracts the complexities of orbital mechanics and terrestrial network topologies. Users, ranging from environmental scientists to automated remote sensing applications, declare their operational requirements through intent APIs or natural language interfaces, such as “high-resolution multi-spectral monitoring of urban heat islands". The layer’s primary role is to capture the Intent Profile, which includes spatial, temporal, and spectral constraints, and relay them to the underlying control logic in a structured format.
- **Intent Enabling Layer**: The middle layer is the intent enabling layer, acting as the “brain" of the entire architecture. It encompasses core functional modules such as intent template design, intent refinement, policy mapping, service intent design, service-level resource mapping, and service orchestration. This layer is responsible for invoking the refinement models of the online mode to transform upper-layer business intents into specific network policies and conduct service-level resource mapping. This layer acts as the cognitive engine, making real-time decisions on whether

a monitoring task should be backhauled via a satellite constellation or processed at a terrestrial edge gateway.

- **Infrastructure Layer**: The bottom layer is the infrastructure layer, which constitutes the physical and virtualized foundation of the STIN. It encompasses a diverse array of assets, including low earth orbit satellite constellations, terrestrial 5G/6G base stations, and remote sensing ground stations. This layer manages the coexistence of 3GPP and non-3GPP protocol stacks, handling the distinct framing and signaling requirements of space-based and ground-based links. The network policies generated by the enabling layer are pushed to this layer as executable configurations. Whether configuring a satellite's beam-hopping pattern or a terrestrial router's priority queue, the infrastructure layer executes these commands to realize global resource scheduling and end-to-end service assurance.

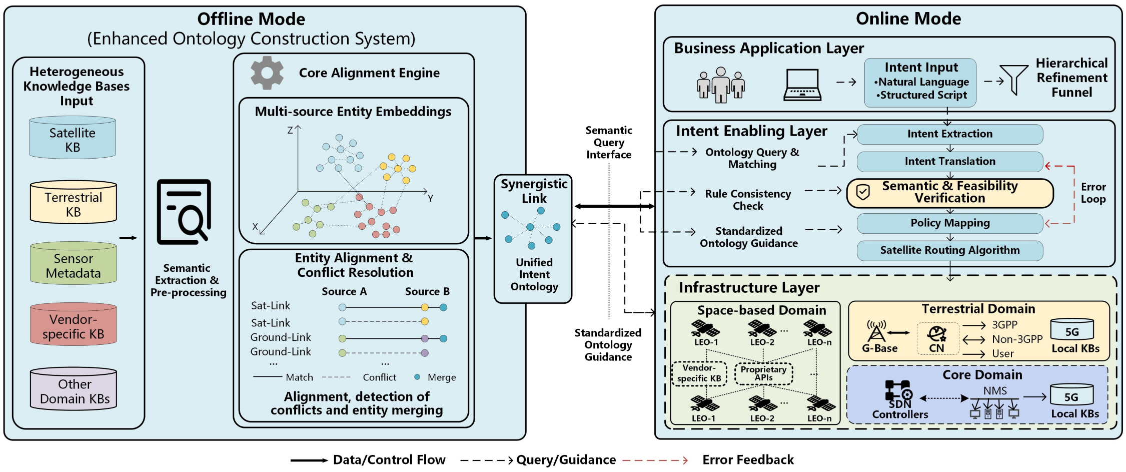


Figure 1. The overall architecture of the ontology-based endogenous intent refinement framework for STIN.

The framework comprises two components, namely, offline mode and online mode, which are respectively responsible for the construction of a unified semantic ontology and real-time intent-to-policy mapping. This separation ensures that the high-frequency dynamics of remote sensing tasks are decoupled from the heavy computational load of knowledge base integration.

3.1.1. Offline Mode

The offline mode serves as the fundamental knowledge layer of the framework. Its primary objective is to bridge the semantic gap among multi-source heterogeneous data within the STIN. Satellite constellations, terrestrial base stations, and core network facilities often originate from different equipment vendors and follow disparate protocol standards such as 3GPP and non-3GPP specifications. These differences result in significant variations in their underlying description languages and metadata structures. In the offline mode, a critical challenge arises when terms from different vendors denote identical physical or logical entities but utilize divergent terminologies. To address this, the system first extracts raw data from various heterogeneous knowledge bases and utilizes ontology learning techniques to extract core concepts and relations. Subsequently, for terms across different knowledge bases that refer to the same object but differ in expression, the system performs semantic fusion and conflict resolution through the entity alignment module. This process ultimately generates a standardized unified intent ontology stored in RDF format, which provides a globally consistent semantic dictionary and reasoning rules for intent understanding and policy generation in the online mode.

3.1.2. Online Mode

The online mode leverages the pre-constructed unified ontology to facilitate the real-time processing of monitoring requests during network runtime. This module converts abstract user goals into executable resource configurations for collaborative sensing tasks. The framework adopts a top-down hierarchical refinement logic to ensure high fidelity between the user intent and the actual network state. The process begins at the intent declaration interface, where operators specify monitoring missions in natural language. The system utilizes the unified ontology to parse these requests into vendor-agnostic semantic objects, neutralizing linguistic ambiguity. Through a specialized service extension model, these general intents are decomposed into quantifiable Quality of Service (QoS) and quality of monitoring metrics. For instance, a "high-resolution" intent is refined into specific parameters such as minimum bandwidth, maximum latency, and revisit frequency requirements. Finally, the resource extension model performs a dynamic mapping between service metrics and concrete infrastructure entities. By sensing the instantaneous state of the physical network, such as the orbital position of a satellite or the congestion level of a terrestrial gateway, the framework selects the optimal combination of virtual and physical resources. This rigorous refinement chain ensures that high-level mission objectives are accurately transformed into machine-comprehensible configuration parameters, enabling autonomous space-terrestrial collaboration.

By integrating offline knowledge construction with online real-time intent orchestration, the proposed framework enables a seamless and intelligent refinement of high-level mission intents into executable network policies. This dual-mode approach not only addresses the semantic interoperability challenges inherent in STIN but also ensures the efficient utilization of heterogeneous resources in dynamic and time-sensitive remote sensing scenarios.

3.2. Ontology-based Intent Refinement Process

As illustrated in Figure 2, the system first extracts a collection of raw intents from various private knowledge bases to address the heterogeneous operation and maintenance requirements in the STIN. Given the significant terminological discrepancies in descriptions of identical network functions across different equipment vendors, direct processing of such unstructured text is highly prone to semantic ambiguity. An ontology construction mechanism is introduced in this phase to mitigate this issue. This mechanism parses discrete raw intents and transforms them into ontology intents based on RDF. By mapping natural language instructions into "subject-predicate-object" triple structures, the system successfully converts ambiguous text descriptions into machine-processable structured graph nodes, establishing a unified data format foundation for subsequent semantic computing.

Upon completing the preliminary structured representation, the workflow transitions to the core processing module depicted in the intent-entity alignment model. To mathematically quantify semantic relationships between different terms, the model first employs KG embedding techniques to map symbolic ontology entities into low-dimensional, continuous real-valued intent-entity vectors, thereby transforming semantic information into spatial vector representations. Subsequently, to address the entity conflict of "synonymy" within heterogeneous networks, an alignment loss model is introduced. During the training process, by minimizing the alignment loss function, the system enforces the proximity of entity vectors possessing identical semantics but originating from distinct sources within the embedding space. This process effectively bridges the semantic gap between heterogeneous knowledge bases, uniformly mapping all private entities into a globally shared intent ontology semantic space.

Finally, the system performs intent similarity calculations within this unified vector space. It uses metric algorithms, such as cosine similarity or Euclidean distance, to measure the geometric distance between the input intent vector and standard intent vectors. Based on the results, the system matches the standard entity with the highest similarity score. This serves as the final refinement result. This step ensures standardized output intents, regardless of the vendor terminology used at the input.

The standardized output intents generated by this process possess universal semantic attributes, making them suitable for direct refinement into executable network configuration policies. These policies include specific instructions for configuring software defined network flow tables, allocating satellite downlink slots, adjusting beamforming parameters, or scheduling multi-access edge computing resources. For example, in a real-time flood monitoring scenario, the system can dynamically

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2026 doi:10.20944/preprints202605.1303.v1

allocate high-priority satellite links for transmitting Synthetic Aperture Radar imagery to ground processing centers while concurrently reserving edge computing resources for rapid data analysis. By ensuring that all output intents are standardized and unambiguous, the system enables automated and precise resource scheduling across the entire STIN. This capability is critical for supporting time-sensitive remote sensing applications, such as disaster response, environmental monitoring, and precision agriculture, where seamless coordination between space and terrestrial networks is essential.

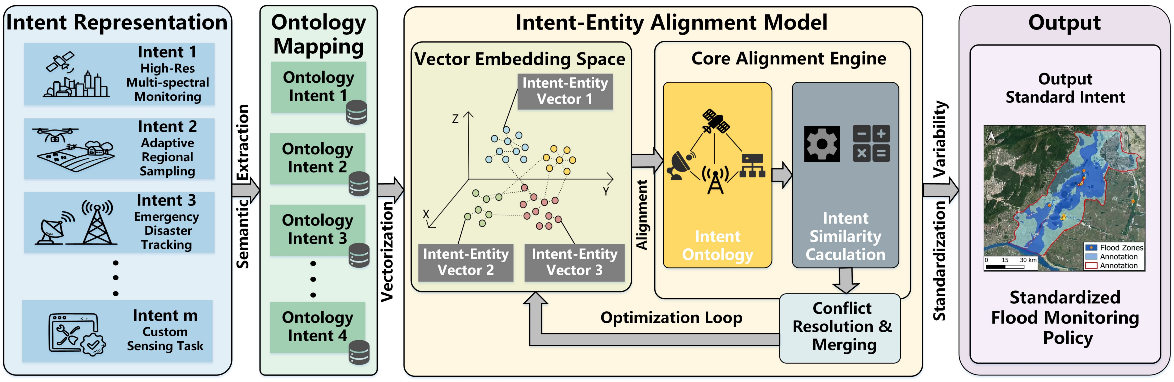


Figure 2. The process of the ontology-based intent-entity alignment model for high-precision semantic matching and conflict resolution.

3.3. Ontology-Based Intent Refinement Verification Mechanism

To ensure the reliability and accuracy of intent refinement in the STIN, we develop a comprehensive verification model that validates the correctness of translated intents through ontology-based similarity analysis. This verification mechanism addresses two critical aspects, including semantic concept alignment and numerical parameter consistency.

The system employs a distance-based semantic similarity calculation model to evaluate the conceptual alignment between translated intents and standard ontology definitions. This approach quantifies the similarity level by measuring the geometric distance between concept vectors in the unified semantic space. A smaller distance indicates higher semantic similarity, confirming that the translated intent correctly captures the intended network operation. For instance, when a user requests "prioritize satellite imagery transmission for wildfire monitoring", the system verifies that the translated intent aligns with concepts related to "high-priority data forwarding" rather than unrelated operations such as "link establishment" or "resource deallocation". This semantic verification prevents conceptual misinterpretation that could lead to inappropriate network configurations.

Beyond semantic alignment, the system addresses the critical challenge of numerical parameter accuracy. Identical service types with different parameter values result in distinct resource allocation requirements. In the STIN, even minor parameter discrepancies can cause significant operational failures due to the limited and expensive nature of satellite resources. Consider the priority parameter as an example. A priority value of "1" designates critical services requiring immediate processing, such as emergency disaster response data. Conversely, a priority value of "3" indicates routine services that can tolerate transmission delays, such as regular environmental monitoring updates. The distinction between these values directly impacts resource scheduling decisions, including satellite beam allocation, ground station assignment, and network bandwidth distribution.

Given the stringent accuracy requirements of the STIN operations, we adopt an exact semantic similarity approach for numerical parameter verification. The system defines the numerical similarity score as 1 only when the input intent value precisely matches the corresponding ontology value. This strict matching criterion ensures zero tolerance for parameter errors, which is essential for maintaining network stability and preventing resource conflicts. The exact matching approach is particularly crucial for space-terrestrial collaborative monitoring applications, where incorrect parameter settings can

 doi:10.20944/preprints202605.1303.v1

result in mission-critical data loss, satellite resource waste, or communication link failures. By enforcing precise numerical alignment, the verification model guarantees that translated intents maintain the exact specifications required for reliable network operation.

This dual-layer verification mechanism, combining semantic concept validation with exact numerical matching, provides comprehensive assurance that the intent refinement process produces accurate and executable network configurations suitable for the demanding requirements of the STIN.

4. Ontology-Aware Entity Alignment for Collaborative Monitoring

In this section, we elaborate on the proposed ontology-enhanced endogenous intent refinement framework. We detail the overall architecture, which operates through a two-phase framework, including an offline heterogeneous entity alignment phase based on an improved TransD model, and an online intent refinement phase. The specific workflow and algorithmic mechanisms designed to bridge the semantic gap between abstract user intents and physical network resources are presented systematically.

4.1. Problem Formulation

With the deployment of IDN across multi-domain and multi-vendor environments, ontology-based modeling has gradually become a key approach for uniformly expressing endogenous network intents and operational knowledge. However, different operators, equipment vendors, and standards organizations often construct intent ontologies independently, leading to significant discrepancies in class hierarchies, relational schemas, and naming conventions. Without an effective ontology alignment mechanism, cross-domain intent refinement struggles to obtain consistent semantic support, and the transferability and reusability of ontologies and models are severely hindered. Therefore, this subsection first formalizes the entity alignment problem at the ontology level. Subsequently, it introduces a typical ontology-aware entity alignment model that integrates KG embedding with conflict-constrained matching strategies to unify lexical and structural information within a shared vector space. An example of entity alignment based on ontology in the STIN is demonstrated in Figure 3.

Semantic conflicts originate from terminological heterogeneity within multi-domain environments. Different network sub-domains often employ disparate vocabularies to describe identical or highly similar concepts. For instance, in the domain-specific KGs of STIN, QoS parameter representing time delay might be labeled as “delay" in terrestrial network management systems, whereas it is defined as “latency" in the satellite network control plane. Similarly, “jitter" and “network jitter", despite referring to the same physical phenomenon, may coexist as independent entities. If these conflicts remain unresolved, the generated global ontology will contain redundant entities, thereby inducing ambiguity during the intent parsing phase and potentially leading to conflicts in configuration policies.

Furthermore, the alignment process faces the complicating impact of semantic drift, where the meaning of a concept deviates across different contexts or abstraction levels [35]. In integrated network architectures, the definition of a specific entity may undergo subtle changes depending on the operational scope. For example, the term “bandwidth" may refer to physical link capacity in the infrastructure layer ontology, whereas it denotes an allocated throughput limit in the service layer ontology. Such contextual discrepancies result in a misalignment where the same label points to different semantic values. Lacking mechanisms to detect and correct such drift, the intent refinement engine may misinterpret user requirements, erroneously mapping high-level service intents onto low-level resources, thereby degrading overall network performance.

4.2. Ontology-Aware Intent Alignment Model

4.2.1. Entity Alignment Model

Entity alignment aims to identify entities representing the same real-world object from different KGs. When different KGs contain entities with identical semantics, KG embedding methods can be employed to map these semantically identical entities to a unified ontology. The design of KG embedding models typically involves three steps. First, the representations of entities and relations are defined. Second, a scoring function is established to measure the plausibility of triples. Finally, the training and learning of entity and relation embedding representations are conducted. During training, the optimization objective is to maximize the scores of triples already existing in the KGs, thereby increasing the probability of correctness. A higher scoring function value indicates a higher plausibility of the triple.

Assume each KG is associated with an ontology containing hierarchical classes and optional disjointness constraints between classes. In this paper, classes are treated as entities in KGs, while inclusion relationships between classes, commonly referred to as $rdfs : subClassOf$, are treated as relations in KGs. Consequently, a simplified ontology can be viewed as a graph.

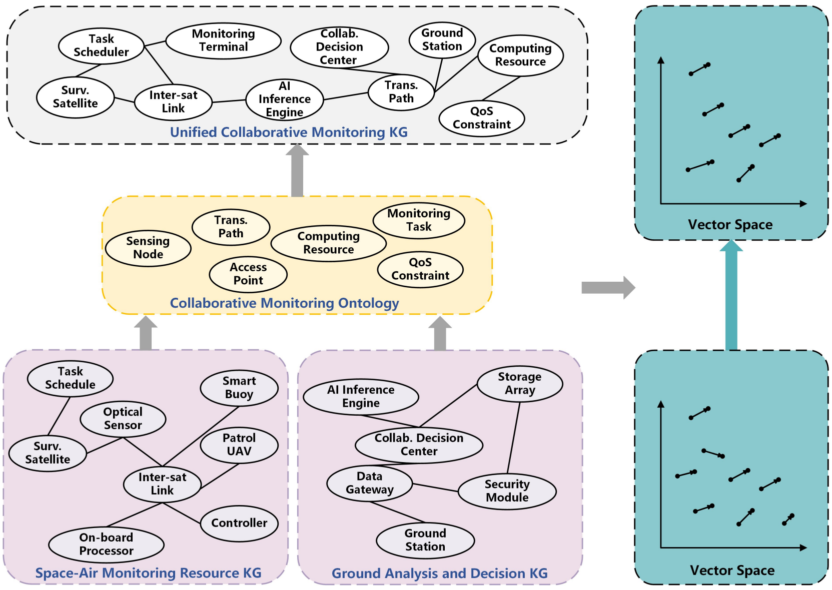


Figure 3. An example of entity alignment based on ontology in the STIN.

Let the domain-specific sensing resource topology be formalized as a knowledge graph, denoted by $G = (\mathcal{X}, \mathcal{R}, \mathcal{T})$, where $\mathcal{X}$ represents the set of resource nodes (e.g., satellites, sensors), $\mathcal{R}$ denotes the set of semantic links (e.g., hasBandwidth, connectsTo), and $\mathcal{T}$ represents the collection of valid assertion triples. Each operational triple $\tau = (x_u, r_k, x_v) \in \mathcal{T}$ comprises a head resource x_u, a semantic link r_k, and a tail resource x_v. We associate each element with a continuous vector representation, denoted as $\mathbf{x}_u, \mathbf{r}_k, \mathbf{x}_v \in \mathbb{R}^d$. For the task of cross-domain alignment, such as synchronizing a space-based KG, $\mathcal{G}_s$, with a ground-based KG, $\mathcal{G}_g$, we define an alignment mapping $\mathcal{M}_{align} = \{(x_s, x_g) | x_s \in \mathcal{X}_s, x_g \in \mathcal{X}_g\}$. It indicates that resource x_s and x_g refer to the same physical entity despite disparate naming conventions in heterogeneous networks. For two KGs to be aligned, denoted as $G_1 = (\mathcal{X}_\infty, \mathcal{R}_\infty, \mathcal{T}_\infty)$ and $G_2 = ((\mathcal{X}_\in, \mathcal{R}_\in, \mathcal{T}_\in)$, the entity mapping between the two KGs can be represented as $m = (e_1, e_2)$, indicating that e_1 and e_2 refer to the same actual object, where $e_1 \in E_1$ and $e_2 \in E_2$. Addressing the phenomenon where different operator equipment configurations vary in heterogeneous network environments, researching ontology mapping and alignment methods helps mask equipment heterogeneity and achieve autonomous intent mapping and management.

For KGs G_1 and G_2, their corresponding ontologies can be represented as $O_1 = (Z_1, H_1)$ and $O_2 = (Z_2, H_2)$, respectively. If the ontologies O_1 and O_2 in G_1 and G_2 are identical, then $O = (Z, H)$. If the ontologies are different, the fused ontology after alignment is denoted as $O = (Z, H)$. Here, Z_1, Z_2, and Z are sets of classes, while H_1, H_2, and H are three sets of $subClassOf$ relations. Assume the membership relationships between entities and corresponding classes are S_1 and S_2, with class embedding vectors and membership relation embedding vectors denoted as $\mathbf{c}$ and $\mathbf{b}$, respectively. For instance, S_1 connects G_1 and Z via $s_1 = (e_1, z)$, where $e_1 \in E_1$ and $z \in Z$.

Traditional translation-based embedding paradigms, such as the rudimentary distance-based scoring functions, operate on the assumption of static, one-to-one semantic mappings. While computationally efficient, these models exhibit significant deficiencies when applied to the dense, multi-modal associations characteristic of collaborative monitoring networks (e.g., a single data relay satellite maintaining concurrent links with multiple Ground Terminals). Such 1-to-N and N-to-N interactions cannot be accurately captured by a single static vector, as the semantic role of a resource node fluctuates depending on its connected peers. To address this, we adopt a dynamic context-projection architecture, which decouples the intrinsic semantic identity of a sensing resource from its variable topological roles.

Consider a specific resource interaction tuple $\tau = (x_i, r_k, x_j)$, linking a head resource node x_i to a tail node x_j via a semantic link r_k. Instead of using fixed geometric embeddings, we assign a dual-vector representation to each element, including a primary semantic vector (e.g., $\mathbf{x} \in \mathbb{R}^d$) encoding invariant attributes, and a context vector (e.g., $\mathbf{x}^{(p)} \in \mathbb{R}^d$) encoding the node's potential for topological interaction.

To model the polymorphic nature of heterogeneous devices, we construct domain-adaptive projection operators, denoted as $\mathbf{\Phi}$. These operators dynamically map entities from their native resource space to a unified intent-relation hyperplane. The operators for the head and tail nodes are formulated as $\mathbf{\Phi}_i = \mathbf{r}_k^{(p)}(\mathbf{x}_i^{(p)})^\top + \mathbf{I}, \quad \mathbf{\Phi}_j = \mathbf{r}_k^{(p)}(\mathbf{x}_j^{(p)})^\top + \mathbf{I}$. The term $\mathbf{I}$ represents the identity matrix. Consequently, the energy score $E_{topo}(\tau)$, which measures the implausibility of a topology link, is defined by the translational distance:

$$E_{topo}(x_h, r_k, x_t) = -\|\hat{\mathbf{x}}_h + \mathbf{r}_k - \hat{\mathbf{x}}_t\|_2^2. \tag{1}$$

This mechanism ensures that the constraints $\|\mathbf{x}\|_2 \leq 1$ are satisfied while allowing diverse resource configurations to coexist without vector collision.

To train the model, we employ a margin-based ranking loss function. We construct a set of corrupted triples $\mathcal{T}'$ by randomly replacing the head or tail nodes with functionally irrelevant resources. The topology consistency objective, $\mathcal{J}_{Topo}$, is minimized to ensure valid triples have significantly lower energy than corrupted ones:

$$\mathcal{J}_{Topo} = \sum_{\tau \in \mathcal{T}} \sum_{\tau' \in \mathcal{T}'} [\gamma_{topo} + E_{topo}(\tau') - E_{topo}(\tau)]_+. \tag{2}$$

where $[z]_+ = \max(0, z)$ denotes the rectifier, and $\gamma_{topo} > 0$ is a hyperparameter serving as the safety margin. Physically, this objective "pushes" valid collaborative topologies into a coherent cluster while repelling invalid network definitions.

4.2.2. Ontology Alignment Model

The unified intent ontology, denoted as $\mathcal{O} = (\mathcal{Z}, \mathcal{H})$, serves as the semantic backbone for mission planning, where $\mathcal{Z}$ represents the set of intent categories and $\mathcal{H}$ comprises the hierarchical specialization relations (e.g., Optical_Sensing is a sub-category of Passive_Monitoring). To mathematically codify this abstraction, we model the hierarchy tuple $(z_{sub}, z_{super}) \in \mathcal{H}$ using a non-linear encapsulation constraint. Specifically, we employ a hyperbolic-style transformation to map specific capabilities into their abstract parent classes. The hierarchy preservation score is defined as $E_{Onto}(z_{sub}, z_{super}) = \|\tanh(\mathbf{W}_{hier}\mathbf{z}_{sub} + \mathbf{b}) - \mathbf{z}_{super}\|_2$. The term $\mathbf{W}_{hier} \in \mathbb{R}^{d \times d}$ and $\mathbf{b} \in \mathbb{R}^d$ are learnable parameters governing the semantic abstraction level. Similarly, the scoring loss function for ontology alignment is:

$$\mathcal{J}_{Onto} = \sum_{\mathcal{H}} [\gamma_{onto} + E_{Onto}(\tau) - E_{Onto}(\tau')]_+. \tag{3}$$

It ensures that the embedding space respects the logical granularity required for drilling down from high-level user missions to low-level device capabilities.

For $\forall e \in E$, $|e|_2 = 1$. Here, H' denotes the set of negative class pairs [31], where each pair is sampled by replacing c_h' and c_t' according to a uniform negative sampling strategy, and γ_o controls the margin. In deep neural network models, each element of the dataset (such as a sentence or an image) can be regarded as a positive sample, i.e., a positive instance, for the model. During training, a common strategy is to provide the model with both positive and negative instances simultaneously, and to enhance the distinction between positive and negative instances by constructing a loss function, thereby enabling the model to learn information from the data. Negative sampling is a process of constructing negative instances corresponding to positive instances according to a specific strategy. For example, in the field of natural language processing, operations such as randomly replacing words in a coherent sentence can be considered negative sampling. Targeted provision of high-quality negative instances can accelerate model convergence and optimize the model in the desired direction.

In intent-driven operations, avoiding configuration conflicts is paramount. We introduce a Policy Dissonance Matrix, $\mathbf{\Omega}$, to explicitly quantify the operational incompatibility between distinct intent categories. The entry Ω_{mn} represents the degree of conflict between category z_m and z_n. The construction of $\mathbf{\Omega}$ follows a logic-driven protocol. (1) Self-Consistency: $\Omega_{mn} = 0$ if $m = n$. (2) Explicit Exclusion: If z_m and z_n are defined as disjoint in the ontology (e.g., real-time stream vs. delay-tolerant transmission), they are deemed fully incompatible, setting $\Omega_{mn} = 1$. (3) Entity Overlap: If the categories share instantiated resources (i.e., a device exists that supports both capabilities), they are considered compatible, yielding $\Omega_{mn} = 0$. The above three conditions are matched sequentially; if any condition is met, the calculation of $m_{i,j}$ is completed. If none of these three conditions are satisfied, the principle is followed that in a tree-like class hierarchy, the farther apart two classes are, the lower their semantic similarity and the higher their degree of conflict [36]. To calculate the distance between z_1 and z_2, the sets of classes routed from z_1 and z_2 to the root class are used, denoted as $S(z_1)$ and $S(z_2)$, respectively. Then, the ratio of the intersection of $S(z_1)$ and $S(z_2)$ is taken, i.e., $m_{i,j} = 1 - \frac{|S(z_1) \cap S(z_2)|}{|S(z_1) \cup S(z_2)|}$, where $|\cdot|$ denotes the cardinality of the set, describing the number of elements contained in the set.

Beyond explicit logical exclusions, we address implicit semantic friction derived from vector topology. We postulate that intent categories with high embedding similarity should exhibit minimal operational conflict. To enforce this, we utilize cosine similarity as a proxy for compatibility. The semantic dissonance based on vector alignment is formalized as $D_{cos}(z_m, z_n) = 1 - \cos(\mathbf{z}_m, \mathbf{z}_n)$. To embed this operational logic into the vector space, we minimize a dissonance-penalty loss function:

$$\mathcal{J}_{Diss} = - \sum_{z_m \in \mathcal{Z}} \sum_{z_n \in \mathcal{Z}} \Omega_{mn} \log D_{cos}(\mathbf{z}_m, \mathbf{z}_n). \tag{4}$$

This optimization effectively pushes the embeddings of logically incompatible strategies (where $\Omega_{mn} \to 1$) to be orthogonal, thereby preventing the Intent Refinement Engine from generating contradictory command sequences.

4.2.3. Alignment Loss Model

Regarding membership loss, membership relations B_1 and B_2 are utilized to associate the KG embedding space with the ontology embedding space, thereby enhancing KG embeddings with ontology semantics. A critical challenge in IDN is mapping a specific physical resource to its abstract capability definition. We term this process vertical semantic anchoring. To bridge the heterogeneity between the resource embedding space ($\mathbb{R}^{d_x}$) and the intent ontology space ($\mathbb{R}^{d_z}$), we exploit membership assertions set $\mathcal{M}_{inst}$. For a valid instantiation pair $m = (x, z) \in \mathcal{M}_{inst}$, which signifies that resource x possesses the capability z, we employ a non-linear projection to align their semantic representations. The projection error is quantified as:

$$E_{inst}(x, z) = \| \tanh(\mathbf{W}_{inst}\mathbf{x} + \mathbf{b}_{inst}) - \mathbf{z} \|_2, \tag{5}$$

where $\mathbf{W}_{inst} \in \mathbb{R}^{d_x \times d_z}$ is the cross-layer transition matrix. Accordingly, the intent instantiation objective is formulated to minimize the discrepancy across all valid membership pairs while contrasting them against invalid mappings:

$$\mathcal{J}_{Inst} = \sum_{(x,z)\in\mathcal{M}_{inst}} \sum_{(x',z')\in\mathcal{M}'_{inst}} [\gamma_{inst} + E_{inst}(x,z) - E_{inst}(x',z')]_+, \tag{6}$$

where $\mathcal{M}'_{inst}$ represents the set of negative instantiations. By optimizing $\mathcal{J}_{Inst}$, the model ensures that the vector representation of every physical device is strictly anchored within the semantic region of its functional category, preventing capability mismatch during mission planning.

To bridge the semantic silos between heterogeneous sensing domains, we leverage a set of pre-aligned anchor links, denoted as $\mathcal{M}_{anchor} = \{(x_s, x_g)\}$, where x_s and x_g represent equivalent resources in the source and target domains, respectively. The alignment objective is to minimize the translational discrepancy between these anchor pairs in the unified vector space. We define the alignment scoring function as $E_{align}(x_s, x_g) = \|\mathbf{W}_{align}\mathbf{x}_s - \mathbf{x}_g\|_2$, where $\mathbf{W}_{align}$ serves as the domain-transfer matrix. The final alignment loss is computed as:

$$\mathcal{J}_{Align} = \sum_{(x_s,x_g)\in\mathcal{M}_{anchor}} E_{align}(x_s, x_g). \tag{7}$$

This term forces the manifold structures of different network domains to overlap, enabling the system to identify corresponding resources based on vector proximity, such as matching a satellite's downlink port to a ground station's receiver.

Finally, the comprehensive objective function for the proposed framework integrates resource structure, intent hierarchy, and policy constraints:

$$\mathcal{J}_{Total} = \mathcal{J}_{Topo} + \mathcal{J}_{Onto} + \lambda_1\mathcal{J}_{Diss} + \lambda_2\mathcal{J}_{Inst} + \lambda_3\mathcal{J}_{Align}, \tag{8}$$

where $\mathcal{J}_{Align}$ represents the seed alignment loss across heterogeneous domains. The term λ_1, λ_2, and λ_3 are hyperparameters balancing the conflict embedding, membership embedding, and alignment embedding losses, respectively. Through iterative co-training, the model converges to a unified semantic space that supports precise intent refinement for collaborative monitoring missions. To reduce model complexity and accelerate model convergence, an iterative co-training strategy is employed. In each iteration, $\mathcal{J}_{Res}$ and $\mathcal{J}_{Onto}$ are first optimized independently, followed by the sequential optimization of $\mathcal{J}_{Diss}$ and $\mathcal{J}_{Mem}$, and finally the optimization of $\mathcal{J}_{Align}$. The iteration stops when certain stopping conditions in the validation set are met.

Through the embeddings of G_1, G_2, and O, entity mappings with cosine similarity are calculated. Given two entities $e_1 \in G_1$ and $e_2 \in G_2$, the weighted similarity score is calculated as:

$$sim(e_1, e_2) = \eta\cos(e_1, e_2) + (1-\eta)\cos(c_1, c_2), \tag{9}$$

where $\eta \in [0,1]$ is used to balance the similarity of entity embeddings and class embeddings. An entity may declare multiple classes; in this case, the average of all declared class embeddings needs to be calculated as the class embedding. In prediction, for each entity in G_1 to be aligned, all candidate entities in G_2 are ranked according to their weighted similarity scores. The *sim* function is a method commonly used in fields such as text classification and recommender systems to calculate the similarity between two vectors. In computer science, a vector refers to a quantity having magnitude and direction, typically represented as a sequence of numbers. The calculation of the *sim* function is based on the vector space model, which represents text as vectors and calculates the similarity between them. The vector space model converts text into vectors and then uses methods such as cosine similarity to calculate their similarity.

5. Simulation Results

In the system implementation, network endogenous intents are typically encapsulated in the form of event messages. Upon receiving an intent, the refinement module leverages the ontology-based

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2026 doi:10.20944/preprints202605.1303.v1

endogenous intent refinement model to query the knowledge base and derive network logical policies containing communication metrics, as shown in Figure 4.

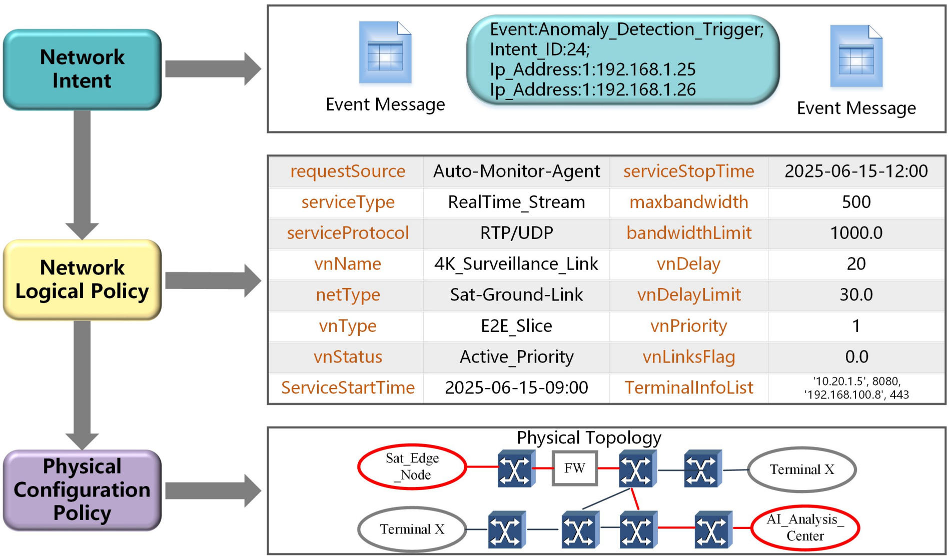


Figure 4. Example of network intent instantiation and refinement.

In the simulation evaluation, this paper ranks the matching candidates for each entity to be aligned and calculates the metrics of Hits@1 (H@1), Hits@5 (H@5), and Mean Reciprocal Rank (MRR) [37]. Hits@n refers to the average proportion of triples with a ranking within the top n in link prediction; the larger this indicator, the better. The calculation method is as follows:

$$\text{Hits}@n = \frac{1}{|S|}\sum_{i=1}^{|S|}\mathbb{I}(\text{ rank } i \leqslant n), \tag{10}$$

where S is the set of triples, $|S|$ is the number of triples in the set, and rank i refers to the link prediction rank of the i-th triple.

MRR serves as an important indicator for assessing the precision of intent entity alignment. Mathematically, it computes the average of the reciprocal ranks assigned to the ground-truth target resources across all query instances. Specifically, for each query, the reciprocal of the correct answer's rank is taken as the score for that query, and then the average score of all queries is calculated; the larger this indicator, the better. The calculation formula for MRR is:

$$\text{MRR} = \frac{1}{|S|}\sum_{i=1}^{|S|}\frac{1}{\text{rank}_i} = \frac{1}{|S|}\left(\frac{1}{\text{rank}_1} + \cdots + \frac{1}{\text{rank}_{|S|}}\right). \tag{11}$$

To conduct a comprehensive evaluation, the IWSLT14 dataset is adopted for simulation. We utilize standard cross-lingual knowledge bases derived from DBpedia (e.g., EN-FR and EN-DE variants) to test generalizability [38]. DBpedia realizes Semantic Web applications by converting Wikipedia entries into structured data and linking them with data from other sources. EN_FR _15K _V1 /V2, EN_DE _15K _V1 /V2, D_W _15K _V1 /V2, and MED_BBK_9K [30] are used as benchmarks, as shown in Table 1. Targeting the network operations scenario, we constructed the NICH dataset, including term definitions from two operations knowledge bases. Its representation format is "1@ Latency 2@ Delay 1@ Connection Status 2@ Link Status 1@ Priority 2@ Task Level 1@ Device ID 2@ Device ID", etc. Classes in the dataset are represented in RDF format as <owl:Classrdf:about="http://example.org/netop#Router"/>, etc., and properties are represented in RDF format as <owl:DatatypePropertyrdf:about="http://example.org/netop#hasIPAddress">, etc. The dataset comprises 11 ontology classes and

10 ontology relation triples. The two KGs contain 32 and 20 relations, and 19 and 21 properties, respectively. For cross-lingual benchmarks in DBpedia, the KG to be aligned share a single ontology.

Table 1 Summary of Datasets.

Dataset	KG	Entities	Relations	Attributes	Classes	Rel. Triples	Attr. Triples
EN_FR_15K_V1	EN	15000	267	308	189	47,334	73,121
	FR	15000	210	404	189	40,864	67,167
EN_FR_15K_V2	EN	15000	193	189	104	96,318	66,899
	FR	15000	166	221	104	80,112	68,779
EN_DE_15K_V1	EN	15000	215	286	175	47,676	83,775
	DE	15000	131	194	175	50,419	156,150
EN_DE_15K_V2	EN	15000	169	171	86	84,867	81,998
	DE	15000	96	116	86	92,632	186,335
D_W_15K_V1	DB	15000	248	342	172	38,265	68,258
	WK	15000	269	629	140	42,746	138,246
D_W_15K_V2	DB	15000	167	175	71	73,983	66,813
	WK	15000	121	457	68	83,365	175,686

We benchmark the proposed framework against a spectrum of state-of-the-art alignment paradigms to demonstrate its superiority, as depicted in Table 2. The comparative analysis encompasses translational embedding approaches as well as advanced graph neural network architectures, including MTransE, JAPE, GCN-Align, and AliNet [39–42]. All baseline models were implemented within the OpenEA environment [30], utilizing the AdaGrad optimizer with a fixed learning rate of 0.01 to ensure a strictly controlled experimental setup [43]. The batch sizes for entity embedding and ontology embedding are set to 4500 and 64, respectively, and the dimension is set to 300. The hyperparameters are set as $\alpha_1 = 1$, $\alpha_2 = 1$, $\alpha_3 = 5$, $\eta = 0.5$, and γ_m, γ_o, and γ_d are all 0.01 [44].

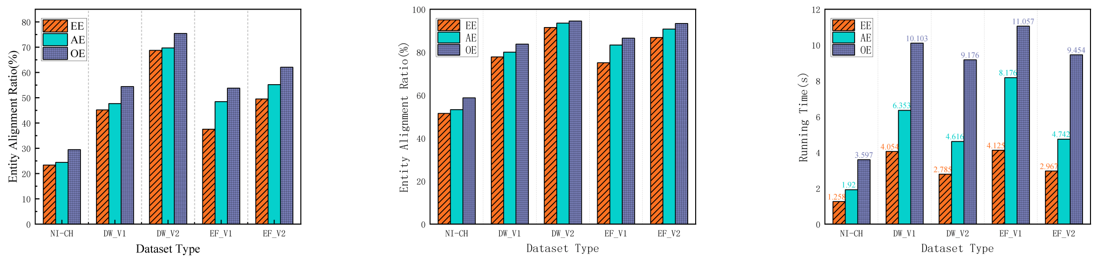
Table 2 Comparison of Entity Alignment Models.

Model	EN-FR-15K-V1			EN-FR-15K-V2			D-W-15K-V1			D-W-15K-V2			NI_CH_1K		
	H@1	H@5	MRR	H@1	H@5	MRR	H@1	H@5	MRR	H@1	H@5	MRR	H@1	H@5	MRR
MTransE	24.7	46.7	36.1	24	43.6	33.6	25.9	46.1	35.4	27.1	49	37.6	4	1.4	1.2
JAPE	26.2	49.7	37.2	29.2	52.4	40.2	25	45.7	34.8	26.2	48.4	36.8	3	9	9
GCN-Align	33.8	58.9	45.1	41.4	69.8	54.2	36.4	58	46.1	50.6	74.3	61.2	6.5	15.3	11.7
AliNet	25.8	43.7	33.9	35.9	56.9	45.3	27	40.3	33.1	52.2	69.8	60.1	1.7	4.2	3.3
INON	53.8	79.2	65.1	62	87.2	73.2	54.4	77.5	64.7	75.4	90.4	82.2	29.5	50.5	39.4

Quantitative analysis reveals that while GNN-based baselines generally outperform traditional translational models, our proposed framework achieves a significant performance leap across all datasets. This substantial improvement is primarily attributed to the integration of the Hierarchical Intent Ontology module. By explicitly modeling the logical depth of monitoring intents, our approach effectively bridges the semantic gap that purely structural models fail to address. Compared with the best results of other models, H@1 is improved by more than 40%, H@5 by more than 21%, and MRR by more than 34%, fully demonstrating the improvement brought by ontology alignment. This also indicates that even introducing a simple ontology structure can greatly promote the alignment effect due to its high quality and strongly relevant domain knowledge, helping the model better understand and model the semantics of entities. In addition, although the scale of the ontology constructed for network operations and maintenance in this paper is limited, the ontology is of high quality because it is created using domain knowledge.

 doi:10.20944/preprints202605.1303.v1

The proposed model is validated under various datasets. The comparison of Hits@1 and Hits@10 when adopting different alignment methods is shown in Figure 5. In the figures, EE represents using only entity embeddings in entity alignment, AE represents using both class embeddings and entity embeddings, and OE represents using ontology alignment, class embeddings, and entity embeddings. As can be seen from the figures, the entity alignment ratio of the dataset constructed in this paper is relatively low. This is because the constructed domain-specific dataset is small, and the effect of manually processing the ontology is prone to errors. Furthermore, in each dataset, the entity alignment performance gradually improves when adopting EE, AE, and OE methods, indicating that pre-processing ontology class conflicts helps to improve intent entity alignment performance.



(a) (b) (c)

Figure 5. Intent-driven CoX system. (a) Entity alignment performance in terms of Hits@1 across five datasets. (b) Entity alignment performance in terms of Hits@10 across five datasets. (c) Comparison of running time for EE, AE, and OE methods across five datasets.

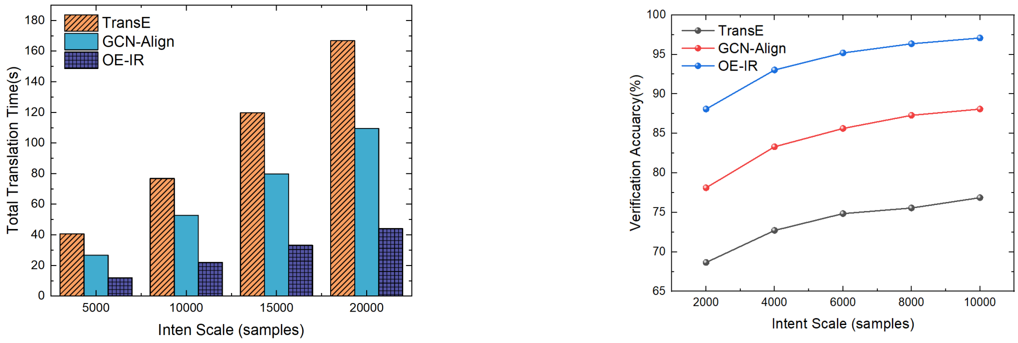
The comparison of running times for EE, AE, and OE methods under different datasets is shown in Figure 5. In each dataset, the running time gradually increases when adopting EE, AE, and OE methods, indicating that employing more processing models consumes more time. In addition, the dataset constructed in this paper is relatively small; therefore, when using the same entity alignment method, the running time on this paper's dataset is shorter than on other datasets.

Moreover, we conducted a comprehensive performance analysis of the intent translation and verification mechanisms using network simulations of the Space-Terrestrial Integrated Network. The configuration employed for the remote sensing collaborative monitoring scenario is based on a multi-domain knowledge graph environment. To ensure a fair comparison of computational overhead, the simulation environment is configured with a unified hardware setup, while the training parameters are set with a batch size of 2048 and a learning rate of 0.001.

In the following, we verify the effectiveness of our proposed intent refinement framework. We aimed to compare the performance of our proposed scheme against two widely adopted baseline models: a traditional translation-based model (TransE) and a Graph Convolutional Network alignment model (GCN-Align).

- **TransE scheme:** This approach employs a basic geometric distance algorithm to map entities without leveraging hierarchical ontology structures. It relies on an exhaustive search over an unoptimized semantic space, which often leads to higher computational latency and limited semantic understanding.
- **GCN-Align scheme:** This algorithm optimizes for structural graph features by gathering neighborhood node information. While it captures topological characteristics effectively, it struggles to resolve deep semantic conflicts across heterogeneous domains.
- **OE-IR (Proposed) scheme:** Our proposed Ontology-Enhanced Intent Refinement framework utilizes a dual-mode synergistic mechanism. By integrating offline unified ontology construction, it effectively prunes the semantic search space. Furthermore, its alignment loss optimization leverages large-scale data to resolve semantic conflicts, ensuring precise intent governance.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 May 2026 doi:10.20944/preprints202605.1303.v1



(a) (b)

Figure 6. Performance evaluation of the ontology-enhanced intent refinement system. (a) Total translation time comparison across varying intent scales. (b) Verification accuracy comparison across varying intent scales.

We then evaluate the performance of these schemes concerning the total translation time and verification accuracy under consistent conditions. To explore the effect of heterogeneous network loads on intent translation performance, we systematically altered the intent scale, selecting 5,000, 10,000, 15,000, and 20,000 samples. The simulation outcomes related to intent translation efficiency are presented in Figure 6(a). As the intent scale rises, we observe a consistent increase in the total translation time across all three mapping schemes, as a larger volume of intents naturally expands the semantic search space. However, the traditional TransE model exhibits the most severe performance degradation, reaching approximately 145 seconds when processing 20,000 samples. The GCN-Align model also records a translation time of around 98 seconds at the maximum scale. However, the traditional TransE model exhibits the most severe performance degradation, reaching approximately 166.79 seconds when processing 20,000 samples. The GCN-Align model also records a translation time of approximately 109.50 seconds at the maximum scale. In stark contrast, the proposed OE-IR scheme demonstrates a significant efficiency advantage, maintaining a near-linear and much slower growth rate. At 20,000 samples, the OE-IR scheme completes the translation in only approximately 44.03 seconds. This comparative result strongly proves that the ontology-based translation mechanism successfully prunes the complex search space, resulting in a substantial reduction in the total time consumed by our proposed scheme. This effectively prevents computational bottlenecks and demonstrates robust real-time responsiveness for high-load remote sensing missions.

After accomplishing efficient intent translation based on the ontology, we further conducted intent verification experiments to investigate the robustness of the verification module by increasing the volume of training data from 2,000 to 10,000 samples. Specifically, verification accuracy is defined as the successful alignment of both semantic concepts and exact numerical parameters. The results are illustrated in Figure 6(b). During the initial stage with 2,000 training samples, all models exhibit relatively lower accuracy due to the limited data available for capturing complex semantic nuances. Nevertheless, the proposed OE-IR framework still achieves a clear advantage, obtaining an accuracy of 88.09%, compared with 68.66% for TransE and 78.11% for GCN-Align. As the training data volume increases, the verification accuracy of all three schemes improves steadily. However, the TransE and GCN-Align models exhibit evident performance bottlenecks, converging at approximately 76.86% and 87.29%, respectively, when the training scale reaches 10,000 samples. Conversely, the proposed OE-IR framework maintains a significant lead across all data scales and rapidly converges to approximately 97.09% at 10,000 samples.

These experimental results fully demonstrate that as the training data volume increases, the verification performance improves significantly. This steep improvement curve not only validates the positive impact of large-scale data on model training but also confirms that our ontology-enhanced alignment method can deeply mine and utilize these data to resolve semantic conflicts, thereby ensuring high-fidelity intent verification as the system scales.

 doi:10.20944/preprints202605.1303.v1

6. Challenges and Potential Solutions

While the proposed ontology-enhanced intent refinement framework effectively addresses semantic heterogeneity and validation in STIN, the evolution towards 6G and the Internet of Agents (IoA) introduces new complexities. This section identifies four critical challenges and outlines potential research directions to advance the field of intent-driven space-terrestrial collaboration.

- **Continuous Ontology Evolution in Dynamic Environments**: Current ontology construction methods rely heavily on static domain knowledge and manual curation, making them improper for time-varying network contexts. When new sensor types are deployed or novel service requirements emerge, existing ontologies may lack corresponding concepts and relations, resulting in semantic gaps during intent refinement. Future research should explore automated ontology learning and continuous knowledge integration. By leveraging large language models (LLMs) combined with active learning, the system can extract new concepts from unstructured telemetry data, operation logs, and technical documentation in real-time. A promising approach involves developing a human-in-the-loop framework where the model identifies uncertain translations and queries domain experts for validation, subsequently updating the knowledge graph.
- **Distributed Intent Negotiation in Multi-Agent Architectures**: As STIN evolves toward a decentralized Internet of Agents, network entities such as LEO satellites, high-altitude platforms, and ground stations will possess autonomous decision-making capabilities. In this multi-agent paradigm, intent conflicts become inevitable. For example, a global energy minimization intent may conflict with a local agent's requirement for high-throughput real-time imaging during disaster response. The current centralized refinement approach faces scalability limitations when coordinating thousands of autonomous agents, leading to communication bottlenecks and increased latency. By modeling conflict resolution as a Nash equilibrium problem, agents can autonomously balance global network objectives with local mission priorities. Federated learning approaches can train intent understanding models collaboratively across distributed satellites without transmitting sensitive raw data, ensuring both privacy preservation and model generalization.
- **Semantic Communications for Bandwidth-Efficient Intent Signaling**: Communication links in STIN are characterized by limited bandwidth, high propagation delay, and variable channel quality. Transmitting verbose natural language intents or extensive ontology graphs repeatedly consumes valuable spectrum resources and introduces latency. The current paradigm, which treats intent transmission and channel coding as separate processes, is inefficient for 6G scenarios requiring goal-oriented communication. The fundamental challenge lies in compressing intent semantics to the minimum information necessary for successful task execution, rather than pursuing perfect bit-level reconstruction. Future research should develop joint source-channel coding schemes optimized for intent semantics, where deep semantic encoders extract task-critical features and channel codes are designed to protect these features against high bit error rates typical in satellite links. A shared semantic knowledge base synchronized between ground stations and satellites can enable differential transmission, where only semantic deltas or ontology indices are communicated.

7. Conclusions

In this paper, we proposed an ontology-enhanced endogenous intent refinement framework to tackle the semantic heterogeneity and conflict resolution challenges in STIN. By leveraging an "offline and online" dual-mode intent refinement framework and a novel TransD-based heterogeneous entity alignment model, we effectively bridged the semantic gap between abstract intents and physical resources. Extensive simulations verified that our framework achieves notable gains over existing methods in terms of refinement accuracy, transferability, and reusability, supporting more robust cross-domain orchestration for integrated monitoring and remote sensing tasks. Future work will focus on continual ontology evolution under dynamic STIN conditions, conflict-aware multi-agent coordination, and stronger safety assurance for intent-to-action execution.

 doi:10.20944/preprints202605.1303.v1

Author Contributions: Conceptualization, Y. O., J. Z., Y. W., and B.X.; methodology, Y. O., J. Z., Z. W., Z. J., and X.B.; software, Y. O., J. Z., and Z. W.; validation, Z. W., Z. J.; writing—original draft, Y. O., J. Z.; writing review and editing, Z. J., B.X.; visualization, Y. O., J. Z.; supervision, B.X.; funding acquisition, Y. O.. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the China Postdoctoral Science Foundation under Grant 2025M773486, the Natural Science Basic Research Plan in Shaanxi Province of China under Grant 2025JC-YBQN-820, and the Postdoctoral Science Foundation of Shaanxi Province under Grant 2024BSHSDZZ110.

Data Availability Statement: Not applicable.

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

References

1. Giordani, M.; Zorzi, M. Non-Terrestrial Networks for 6G: Ten Key Scenarios and Enabling Technologies. *IEEE Wireless Communications* **2021**, *28*, 154–161.
2. Dritsas, E.; Trigka, M. Remote Sensing and Geospatial Analysis in the Big Data Era: A Survey. *Remote Sensing* **2025**, *17*. https://doi.org/10.3390/rs17030550.
3. Mehmood, K.; Kralevska, K.; Palma, D. Knowledge-Driven Intent Life-Cycle Management for Cellular Networks. *IEEE Transactions on Network and Service Management* **2025**, *22*, 4806–4826. https://doi.org/10.1109/TNSM.2025.3579547.
4. Benzaid, C.; Taleb, T. Zero-Touch Network and Service Management: The Upcoming Transformation. *IEEE Communications Magazine* **2021**, *58*, 18–24.
5. Alemany, P.; Muñoz, R.; Castaneda Cisneros, J.; Karaca, M.; Porambage, P.; Tran, H.Q.; Giardina, P.G.; Tzanettis, I.; Sousa, X.R.; María Jorquera Valero, J.; et al. Defining Intent-Based Service Management Automation for 6G Multi-Stakeholders Scenarios. *IEEE Open Journal of the Communications Society* **2025**, *6*, 2373–2396. https://doi.org/10.1109/OJCOMS.2025.3554250.
6. Qin, X.; Sun, M.; Dai, J.; Ma, P.; Cao, Y.; Zhang, J.; Wang, J.; Xu, X.; Zhang, P.; Niyato, D. Generative AI Meets Wireless Networking: An Interactive Paradigm for Intent-Driven Communications. *IEEE Transactions on Cognitive Communications and Networking* **2025**, *11*, 2056–2077. https://doi.org/10.1109/TCCN.2025.3567613.
7. Boutouchent, A.; Mekrache, A.; Ksentini, A.; Adhane, G.; Fonseca, J.P.C.d.; McNamara, J.; Ramantas, K.; Palena, M.; Iordache, M.; Cigno, R.L.; et al. 6G-INTENSE: Intent-Driven Native Artificial Intelligence Architecture Supporting Network-Compute Abstraction and Sensing at the Deep Edge. *IEEE Vehicular Technology Magazine* **2025**, *20*, 44–54. https://doi.org/10.1109/MVT.2024.3525001.
8. Kaczmarek, A.; Blachowski, J. Remote Sensing Perspective on Monitoring and Predicting Underground Energy Sources Storage Environmental Impacts: Literature Review. *Remote Sensing* **2025**, *17*. https://doi.org/10.3390/rs17152628.
9. Gizachew, B. Artificial Intelligence and Machine Learning in Remote Sensing for Tropical Forest Monitoring: Applications, Challenges, and Emerging Solutions. *Remote Sensing* **2026**, *18*. https://doi.org/10.3390/rs18081193.
10. Gündüz, D.; Qin, Z.; Aguerri, I.E.; Dhillon, H.S.; Yang, Z.; Yener, A.; Wong, K.K.; Chae, C.B. Beyond Transmitting Bits: Context, Semantics, and Task-Oriented Communications. *IEEE Journal on Selected Areas in Communications* **2023**, *41*, 5–41.
11. Jacobs, A.S.; Pfitscher, R.J.; Ferreira, R.A.; Granville, L.Z. Refining Network Intents for Conflict Detection and Resolution. *IEEE Transactions on Network and Service Management* **2022**, *19*, 2224–2238.
12. Torres-Quezada, E.; Fuentes-Peñailillo, F.; Gutter, K.; Rondón, F.; Marmolejos, J.M.; Maurer, W.; Bisono, A. Remote Sensing and Soil Moisture Sensors for Irrigation Management in Avocado Orchards: A Practical Approach for Water Stress Assessment in Remote Agricultural Areas. *Remote Sensing* **2025**, *17*. https://doi.org/10.3390/rs17040708.
13. Kazanskiy, N.; Khabibullin, R.; Nikonorov, A.; Khonina, S. A Comprehensive Review of Remote Sensing and Artificial Intelligence Integration: Advances, Applications, and Challenges. *Sensors* **2025**, *25*. https://doi.org/10.3390/s25195965.
14. Tuncer, D.; Marotta, A. Intent-Based Service Fulfillment and Assurance in Multi-Domain Networks. *IEEE Transactions on Network and Service Management* **2022**, *19*, 4085–4099.
15. Leivadeas, A.; Falkner, M. A Survey on Intent-Based Networking. *IEEE Communications Surveys & Tutorials* **2023**, *25*, 625–655. https://doi.org/10.1109/COMST.2022.3215919.

16. Haj Hussein, D.; Ibnkahla, M. Toward Intelligent Intent-Based Network Slicing for IoT Systems: Enabling Technologies, Challenges, and Vision. *IEEE Transactions on Network and Service Management* **2025**, *22*, 3480–3495. https://doi.org/10.1109/TNSM.2025.3570052.
17. Siddiqui, M.S.; Leguay, J.; Tantsura, J.; Lopez, V. Intent-Based Networking: Standards, Challenges, and Future Outlook. *IEEE Transactions on Network and Service Management* **2023**, *20*, 1012–1028.
18. Halpern, J.M.; Pignataro, C. Service Function Chaining (SFC) Architecture. *IEEE Communications Standards Magazine* **2024**, *8*, 22–28.
19. 3GPP. Management and orchestration; Intent driven management services for mobile networks. Technical Specification (TS) 28.312, 3rd Generation Partnership Project (3GPP), 2022. Release 17.
20. TM Forum. Intent in Autonomous Networks: Weaver. Exploratory Report IG1253, TM Forum, 2020.
21. Clemm, A.; Ciavaglia, L.; Granville, L.Z.; Tantsura, J. Intent-Based Networking - Concepts and Definitions. RFC 9315, IETF, 2022. https://doi.org/10.17487/RFC9315.
22. Hoang, D.T.; Niyato, D.; Kim, D.I.; Van Huynh, N.; Shongwe, T. Dynamic Resource Allocation for LEO Satellite-Based IoT Networks. *IEEE Transactions on Wireless Communications* **2022**, *21*, 9234–9249.
23. Cheng, N.; Lyu, F.; Quan, W.; Zhou, C.; He, H.; Shi, W.; Shen, X. Space/Aerial-Assisted Computing Offloading for IoT Applications: A Learning-Based Approach. *IEEE Transactions on Wireless Communications* **2021**, *18*, 29–40.
24. Colombo, A.W.; Karnouskos, S.; Kaynak, O.; Shi, Y.; Yin, S. Industrial Cyberphysical Systems: A Backbone of the Fourth Industrial Revolution. *IEEE Transactions on Industrial Informatics* **2017**, *13*, 6–16.
25. Wu, Y.; Zhang, K.; Zhang, Y. Digital Twin Networks: A Survey. *IEEE Communications Surveys & Tutorials* **2021**, *23*, 2428–2458.
26. Pan, J.Z.; Horrocks, I. Rdfs (fa): connecting rdf (s) and owl dl. *IEEE Transactions on Knowledge and Data Engineering* **2006**, *19*, 192–206.
27. Khan, J.A.; Kumar, S. Deep analysis for development of RDF, RDFS and OWL ontologies with protege. In Proceedings of the Proceedings of 3rd international conference on reliability, infocom technologies and optimization. IEEE, 2014, pp. 1–6.
28. Silva, P.B.F.; Santos, A.; Granville, L.Z.; Schaeffer-Filho, A. A Knowledge Graph-Based Approach for Root Cause Analysis in 5G Networks. *IEEE Transactions on Network and Service Management* **2022**, *19*, 4321–4335.
29. Hogan, A.; Blomqvist, E.; Cochez, M.; d'Amato, C.; Melo, G.d.; Gutierrez, C.; Kirrane, S.; Gayo, J.E.L.; Navigli, R.; Neumaier, S.; et al. Knowledge Graphs. *ACM Computing Surveys* **2021**, *54*, 1–37.
30. Sun, Z.; Zhang, Q.; Wei, W.; Wang, P.; Hu, B.; Wu, Y.; Hu, W. A Benchmarking Study of Embedding-based Entity Alignment for Knowledge Graphs. *Proceedings of the VLDB Endowment* **2020**, *13*, 2326–2340. https://doi.org/10.14778/3407790.3407832.
31. Bordes, A.; Usunier, N.; Garcia-Duran, A.; Weston, J.; Yakhnenko, O. Translating Embeddings for Modeling Multi-relational Data. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), 2013, Vol. 26.
32. Ji, S.; Pan, S.; Cambria, E.; Marttinen, P.; Yu, P.S. A Survey on Knowledge Graphs: Representation, Acquisition, and Applications. *IEEE Transactions on Neural Networks and Learning Systems* **2022**, *33*, 494–514.
33. Sun, Z.; Hu, W.; Li, C. JAPE: Joint Attribute-Preserving Embedding for Cross-lingual Entity Alignment. In Proceedings of the International Semantic Web Conference (ISWC). Springer, 2017, pp. 487–502. https://doi.org/10.1007/978-3-319-68288-4_28.
34. Hao, J.; Chen, M.; Yu, W.; Sun, Y.; Wang, W. Universal Representation Learning of Knowledge Bases by Jointly Embedding Instances and Ontological Concepts. In Proceedings of the Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD), 2019, pp. 1709–1719. https://doi.org/10.1145/3292500.3330838.
35. Strinati, E.C.; Barbarossa, S. 6G Networks: Beyond Shannon Towards Semantic and Goal-Oriented Communications. *IEEE Wireless Communications* **2024**, *31*, 10–17.
36. Mumtaz, S.; Giese, M. Frequency-Based vs. Knowledge-Based Similarity Measures for Categorical Data. In Proceedings of the AAAI Spring Symposium: Combining Machine Learning with Knowledge Engineering (1), 2020.
37. Sun, Z.; Hu, W.; Zhang, Q.; Qu, Y. Bootstrapping entity alignment with knowledge graph embedding. In Proceedings of the Ijcai, 2018, Vol. 18.
38. Gangemi, A.; Nuzzolese, A.G.; Presutti, V.; Draicchio, F.; Musetti, A.; Ciancarini, P. Automatic typing of DBpedia entities. In Proceedings of the International Semantic Web Conference. Springer, 2012, pp. 65–81.

 doi:10.20944/preprints202605.1303.v1

39. Chen, M.; Tian, Y.; Yang, M.; Zaniolo, C. Multilingual knowledge graph embeddings for cross-lingual knowledge alignment. *arXiv preprint arXiv:1611.03954* **2016**.
40. Sun, Z.; Hu, W.; Li, C. Cross-lingual entity alignment via joint attribute-preserving embedding. In Proceedings of the International semantic web conference. Springer, 2017, pp. 628–644.
41. Pei, S.; Yu, L.; Hoehndorf, R.; Zhang, X. Semi-supervised entity alignment via knowledge graph embedding with awareness of degree difference. In Proceedings of the The world wide web conference, 2019, pp. 3130–3136.
42. Sun, Z.; Wang, C.; Hu, W.; Chen, M.; Dai, J.; Zhang, W.; Qu, Y. Knowledge graph alignment network with gated multi-hop neighborhood aggregation. In Proceedings of the Proceedings of the AAAI conference on artificial intelligence, 2020, Vol. 34, pp. 222–229.
43. Lydia, A.; Francis, S. Adagrad—an optimizer for stochastic gradient descent. *Int. J. Inf. Comput. Sci* **2019**, *6*, 566–568.
44. Xiang, Y.; Zhang, Z.; Chen, J.; Chen, X.; Lin, Z.; Zheng, Y. OntoEA: Ontology-guided entity alignment via joint knowledge graph embedding. *arXiv preprint arXiv:2105.07688* **2021**.