

Article STO2Vec: A Multiscale Spatio-Temporal Object Representation Method for Association Analysis

Nanyu Chen 🔍, Anran Yang, Luo Chen *🔍, Wei Xiong and Ning Jing

College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China

* Correspondence: luochen@nudt.edu.cn

Abstract: Spatio-temporal association analysis has attracted attention in various fields, such as urban computing and crime analysis. The proliferation of positioning technology and location-based services has facilitated the expansion of association analysis across spatio-temporal scales. However, existing methods inadequately consider the scale differences among spatio-temporal objects during analysis, leading to suboptimal precision in association analysis results. To remedy this issue, we propose a multiscale spatio-temporal object representation method, STO2Vec, for association analysis. This method comprises of two parts: graph construction and embedding. For graph construction, we introduce an adaptive hierarchical discretization method to distinguish the varying scales of local features. Then, we merge the embedding method for spatio-temporal objects with that for discrete units, establishing a heterogeneous graph. For embedding, to enhance embedding quality for homogeneous and heterogeneous data, we use biased sampling and unsupervised models to capture the association strengths between spatio-temporal objects. Empirical results using real-world open-source datasets show that STO2Vec outperforms other models, improving accuracy by 16.25% on average across diverse applications. Further case studies indicate STO2Vec effectively detects association relationships between spatio-temporal objects in a range of scenarios and is applicable to tasks such as moving object behavior pattern mining and trajectory semantic annotation.

Keywords: multiscale spatio-temporal objects; association analysis; adaptive discretization; embedding

1. Introduction

Location-based sharing applications and modern location-aware devices, such as smart wearables and unmanned mobile platforms, have generated and accumulated vast amounts of spatio-temporal data. Spatio-temporal data mining (STDM) research aims to extract valuable information from these data [1]. Spatio-temporal association analysis is a commonly used method to identify groups of entities that exhibit specific spatio-temporal association relationships, such as co-occurrence, in spatio-temporal datasets [2].

Associations are universal [3], making association analysis applicable to various fields with different relationship types. Sharma et al. [2] grouped spatio-temporal associations into three types based on whether a temporal sequence was considered: sequential (e.g., analyzing event-oriented spatio-temporal association in video surveillance [4]), cascading (e.g., studying relationships between events, locations, and criminal activities in criminal geography [5]), and co-occurrences (e.g., similar associations between trajectories [6], co-location patterns between geographic entities [7], semantic annotation of trajectories [8–22], and location embedding [23–33], etc.). By comparing their frequency of co-occurrence, spatio-temporal co-occurrence-based association analysis can reveal implicit associations between entities. This facilitates spatio-temporal semantic understanding, such as urban regional functions, moving object location preferences and behavior patterns [34]. It holds important research value by being an association analysis type that this paper focuses on, and has a positive impact on city



Citation: Chen, N.; Yang, A.; Chen, L.; Xiong, W.; Jing, N. STO2Vec: A Multiscale Spatio-Temporal Object Representation Method for Association Analysis. *ISPRS Int. J. Geo-Inf.* 2023, *12*, 207. https:// doi.org/10.3390/ijgi12050207

Academic Editors: Wolfgang Kainz and Sisi Zlatanova

Received: 26 February 2023 Revised: 16 May 2023 Accepted: 19 May 2023 Published: 21 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). planning, emergency management, and resource allocation. Therefore, in this paper, "association" refers to the interaction between spatio-temporal objects due to the presence of local or global spatio-temporal co-occurrence relationships.

Relationships between spatio-temporal objects are complex and implicit [1], especially in contexts involving a large number of spatio-temporal objects with different scales. For example, as shown in Figure 1, we need to measure and discover direct or indirect associations between aircraft A and aircraft B, farmland M, lake L, airport P, and airport Q from a large number of spatio-temporal objects. These spatio-temporal objects exist in different scales in space and time. The traditional semantic trajectory achieves the matching of moving objects and geographic entities through the stops detection algorithm while ignoring the semantic information of the moving objects during the move process. As a result, they can only obtain the association information between aircraft A and the airport. Location embedding maps the spatio-temporal objects into a unified vector space. The degree of quantitative association between any spatio-temporal objects can be obtained by measuring distances in the vector space. This allows for the discovery of a richer association of relationships.



Figure 1. Illustration of Spatio-Temporal Association. Aircraft *A* takes off from airport *Q* to farmland *M* on a pesticide-spraying mission and then lands at airport *P*. There is an association between *A* and *B* due to similar local trajectories; aircraft *A* hovers repeatedly over farmland *M*, so there is a strong association between *A* and *M*. There is also an association between *A* and *L* as creek *C* passes through farmland *M* to deliver pesticides to lake *L*.

Most studies in location embedding research concentrate on small-scale regions, such as cities [5,10,29,30,35,36], where spatio-temporal objects typically have a uniform scale. As a result, there is a lack of attention paid to the diversity of spatio-temporal object scales. However, the scale attributes of spatio-temporal objects are varied in large geographic regions. Here, "scale" refers to the actual range over which an object exists spatio-temporally, while "multiscale" describes the diversity of scales resulting from differences in the spatio-temporal presence range of the object. As depicted in Figure 1, the scale of the trajectory of aircraft *A* during its flight to and from farmland *M* is much larger than the scale over the farmland *M*. Moreover, there are geographic entities such as creeks, streets, and houses at different scales at *M*. Existing spatio-temporal object-oriented [23–26] or discrete unit-oriented [27–33] embedding algorithms lose many spatial features during the computation

resulting in inaccurate association analysis results: the former may obtain stronger associations between aircraft and point of interests(POI) such as houses (rather than farmland M); meanwhile, the latter may ignore the association between A and M due to the large scale of the predefined grid and the small scale of the local trajectory in M.

It can be seen that the diverse scale characteristics of spatio-temporal objects are an important factor affecting the results of association analysis. While location embedding studies utilizing fine-grained grids have been shown to preserve more spatio-temporal features [31], sparse data present a challenge whereby computational resources are not efficiently utilized [5]. Multilevel discretization techniques can overcome this limitation by retaining association information at multiple scales. However, implementing such methods requires prior knowledge to set fixed multiple levels [32,33]. Our research addresses these shortcomings by adaptively selecting different resolution grid units for discretization based on the scale size of spatio-temporal objects. This preserves feature information across diverse scales. To describe associations between spatio-temporal objects with varying structures and grids at different levels, we integrate the approaches for spatio-temporal objects is obtained through node embedding. Specifically, the contributions of this paper are as follows:

- We propose an adaptive discretization method based on hexagons, which can decompose spatio-temporal objects of different scale sizes into a collection of grids with different resolutions, thereby conserving more of the original spatio-temporal features.
- We designed an associated heterogeneous graph model that can describe the geographic scope and frequency of co-occurrence between spatio-temporal objects based on scale differences. This model enables object embedding for association analysis.
- To improve the scalability of representation methods and the quality of representation results, we designed a biased sampling strategy that can provide richer, applicationspecific associative information for object representation.
- We constructed a multiscale spatio-temporal object representation method called STO2Vec, which is oriented towards association analysis. We performed accuracy tests on association analysis using the representation results of STO2Vec on real datasets.

The remainder of this paper is organized as follows. Section 2 introduces the related work. Section 3 presents the relevant underlying concepts and problem definitions. Section 4 describes the STO2Vec framework specifics. Section 5 presents quantitative experiments and case studies of the proposed framework, while Section 6 presents the discussion and conclusion.

2. Related Work

There are currently numerous studies analyzing the association relationships based on spatio-temporal co-occurrence from different perspectives. These studies can be classified into semantic trajectory-based approaches [8–22] and location embedding-based approaches [23–33] according to the analysis methods.

2.1. Semantic Trajectory

From the perspective of moving objects, the current main research focus is on constructing semantic trajectories by combining trajectory data with geographic information. These trajectories are analyzed to explore spatio-temporal object associations, essentially discovering the relationships between stopping points of moving objects and geographic entities from a spatio-temporal viewpoint. This approach enables a better understanding of spatio-temporal semantics of moving object trajectories.

Alvares et al. proposed the stop-move model, which converts trajectories into sequences with labels through semantic annotation, thereby mining and analyzing the interaction and association of moving objects in geographic space [11]. Based on the stop-move model, some of the research work has focused on how to better geographically associate this semantic annotation of trajectories [9,12–14]; meanwhile, many studies have constructed semantic trajectory models and designed corresponding association analysis algorithms for different application domains. For instance, Ying et al. used the frequent pattern of geography-time-semantics in semantic trajectories for location prediction of moving objects [15]. Kontarinis et al. designed an indoor semantic trajectory model to support the mining and analysis of indoor moving-object trajectories, which was used to test existing analysis algorithms and the proposed algorithm in conjunction with trajectory data inside the Louvre [10]. Noureddine et al. put forward a semantic trajectory model that covers both indoor and outdoor spaces, along with an organizational management method. This model enables associative semantic queries to comprehend people's flow patterns in urban spaces [16]. Based on this model, a graph-based semantic trajectory model was constructed by utilizing open source crowdsourcing data [8] and employing graph analysis algorithms for association analysis among objects. Choi et al. used a semantic trajectory model to mine the movement behavior pattern of pedestrians in the local area [17]. To mine the travel purpose of urban moving objects, Wan et al. developed the SMOPAT algorithm which analyzes frequent patterns in private car trajectories [18]. There are also many studies that focus on association analysis between moving objects such as the similarity metric calculation of trajectories [19,20].

Due to the aggregation of sampled features of trajectories at stopping points that can be explicitly associated with geographic entities, most semantic trajectory studies tend to focus only on the semantic information of the *STOP* phase. Lehmann observed that current semantic trajectory similarity analysis algorithms neglect the semantic information of moving objects during the *MOVE* phase [21], particularly during continuous movement scenarios in free space (e.g., drones and fishing boats), where the association semantics between the *MOVE* phase and geographic entities are abundant. Obviously, the trajectory characteristics of the *MOVE* phase are more complex and variable compared to the *STOP* phase, and this difference in semantic distribution makes the analysis of spatio-temporal object associations in free space more difficult. Xiang et al. [22] described the basic spatial associations by deriving topological, orientation, and distance relationships between the move objects and geographic entities. However, this method involves complex modeling, matching, and inference processes. It is better for analyzing topological movement in specific scenarios than it is for discovering association sin general scenarios.

2.2. Location Embedding

In some studies, geographic entities or regions are embedded using representation learning models that integrate road networks, POI, geotags, origin destination(OD) streams, and trajectory data. This approach facilitates discovering associations between spatiotemporal objects and performing tasks such as regional functional analysis and classification of POI. Based on the embedding object structure, it can be categorized into two types: spatio-temporal object-oriented embedding and discrete unit-oriented embedding.

In a spatio-temporal object-oriented approach, the embedding representation of the object itself can be obtained directly. Due to the advantages of POI data such as strong presentational and large data volume, most studies abstract geographic entities into POIs and identify urban functional areas by mining potential associations between POIs and regions [23]. Zhang et al. proposed a global vector-based POI embedding model GPTEM to mine the implicit semantic associations between POIs and urban functional types by integrating the co-occurrence information and spatial contexts of POIs [24]. Zhang et al. proposed Traj2Vec based on Word2Vec to find pedestrian-location associations in trajectory data and obtain mixed land use characteristics of urban areas [25]. Zhu et al. designed a spatial embedding algorithm Location2vec that combines the interrelated effects between urban locations and moving objects [26]. To address the problem that spatial association between regions is ignored, Sun proposed the Block2vec model, which integrates the information of association between regions based on Skip-gram [23]. In reality, geographic entities do not exist in the form of point elements, and thus these approaches lose most of the spatial information.

Discrete unit-oriented embedding offers the advantage of a consistent method for describing diverse spatio-temporal objects through cell aggregation. Different techniques can be used to discretize spatio-temporal objects, such as Zone2Vec by Du et al., which partitions a city into different regions based on its road network and applies the Skip-gram model to obtain region embedding representations from Beijing taxi trajectories. This approach supports applications such as urban region classification and region clustering [27]. In contrast to the discrete approach of Zone2Vec, Crivellari et al. proposed the Mot2Vec, which segments trajectories at equal time intervals. Trajectories are converted to ID sequences by selecting valuable points for nearest neighbor matching with trajectory segments, preserving more semantic information [28]. To obtain richer semantic information from multiple sources, Jenkins et al. divided the study area into rectangular cells according to a predetermined scale and combined multimodal data such as OD stream data, POI, and remote sensing images to construct regional embeddings [29]. This discretization method highlights background semantic information but only for specific topics. It has poor scalability and cannot handle large-scale datasets for discrete tasks. Spatial indexes based on global discrete grids, such as Geohash encoding [37], Google S2 [38], and Uber H3 [39], can be used to solve these problems, where the grid encoding has a hierarchical structure and is capable of uniquely encoding regions. Woźniak et al. used the open source platform OpenStreetMap tagging data combined with Uber H3 grid division to discover the regional functionality of cities [30].

The above studies are all oriented to local areas, which involve spatio-temporal objects at a relatively uniform scale. Some recent studies have started to focus on location embedding at large scales. Tian et al. proposed the GCN-L2V model. It constructs flow graphs based on trajectory and spatial graphs based on spatial relationships. This helps create fine-grained location embedding at large scales for fixed-level Google S2 grids [31]. In response to the difficulty of a fixed-level grid to solve the data sparsity problem caused by fine-grained embedding in a large-scale context, Shimizu et al. proposed a multilevel grid embedding model. This model obtains fine-grained grid embedding representation with proximity information by discretizing a target region with predefined grids of varied resolutions. However, this technique demands predetermined resolution for each level based on prior knowledge, making it incapable of adjusting to multiscale spatio-temporal objects with scale differences [32]. The above studies are still limited to city-wide analysis, and the extension to larger areas will involve challenges related to balancing accuracy and complexity. Yin et al. implemented a global-scale global positioning system (GPS) coding embedding based on Universal Transverse Mercator (UTM). They did not consider associations between locations based on moving objects. Instead, they emphasized the multimodal semantic information of the locations themselves. Relationships between locations according to movements of objects were not explored [33]. Furthermore, these experiments failed to consider the scale differences of spatio-temporal objects during the embedding process. As a result, their representation in vector space is relatively crude.

3. Preliminary

3.1. Spatio-Temporal Object

Entities in geographic space can be abstracted as different spatio-temporal objects [40]. Due to the complexity of the real world, there often exist significant scale differences between such objects. Therefore, we refer to spatio-temporal objects with a notable scale variance as *multi-scale spatio-temporal objects*.

Moving objects and geographic entities are two types of spatio-temporal objects that are closely related but different in structure. We refer to objects with different structures as *heterogeneous spatio-temporal objects* and those with the same structure as *homogeneous spatiotemporal objects*. For example, moving objects and geographic entities are heterogeneous spatio-temporal objects, while moving objects belong to homogeneous spatio-temporal objects among themselves. Geographic entities are the fundamental units of human cognition of the geographic world. These objects can be either natural or artificial. They have a basic stable spatial position in the world that exists independently. We use a vector data model to describe geographic entities by abstracting them into point, line, and polygon elements:

Definition 1. Geographic Entity $Eg = \{T, (a_1, a_2, ..., a_i, ..., a_n) | a_i = (lon_i, lat_i), T = (t_s, t_e)\}$, where *T* is the time period, and $(a_1, a_2, ..., a_i, ..., a_n)$ are the coordinates of Eg's spatial position in that time period. When Eg is a point element, n = 1, and when it is a polygon element, $a_n = a_1$, and n > 3.

When the geographic entity Eg can be divided into smaller geographic units, called subgeographic entities Eg_{sub} , that have independent characteristics within Eg—for example, tributaries of a river or cities within a country—we refer to them as the subgeographic entities of Eg. There is a containment relationship between Eg and Eg_{sub} , where Eg_{sub} is contained within Eg.

The trajectories generated by moving objects are called spatio-temporal trajectories. Spatio-temporal trajectory data are the raw data collected by the position sensors. In this paper, spatio-temporal trajectories are defined as follows:

Definition 2. Spatio-Temporal Trajectory $Tr = \{a_1, a_2, ..., a_i, ..., a_n | a_i = (p_i, t_i)\}$, where *n* is the number of sample points in the trajectory; a_i is a multidimensional sample point in the trajectory, also called a trajectory point; and $p_i = (x_i, y_i)$, x_i , and y_i are the spatial latitude and longitude coordinates of a_i , respectively, while t_i is the sampling time stamp.

Spatio-temporal trajectory segmentation involves dividing a trajectory into segments using various methods. These trajectory segments have similar internal structural features. The trajectory segments are defined as follows:

Definition 3. Given a spatio-temporal trajectory $Tr = \{a_1, a_2, ..., a_i, ..., a_n | a_i = (p_i, t_i)\}$, we define the segmentation of Tr as $Tre = \{e_1, e_2, ..., e_u, ..., e_m\}$, such that (i) $\forall u \ s.t. \ 1 \le u \le m$, e_u is a trajectory segment, which is a subsequence $\{a_l, a_{l+1}, ..., a_{l+k}\}$ of Tr. (ii) $\bigcup_{u=1}^m e_u = Tr$ and $e_u \cap e_v = \emptyset$, $(u \ne v)$.

The association between spatio-temporal objects often occurs locally rather than globally, and thus dividing spatio-temporal trajectories into segments with complete semantic features can better support the detection of local association relations. There is a containment relationship between Tr and e_u , where e_u is contained within Tr.

3.2. Space Discretization

There are many ways to divide space into units, which can be done manually or according to fixed methods such as administrative divisions. However, these methods cannot flexibly deal with spatio-temporal objects of different scales. Global grid encoding systems such as Google S2, Geohash, and Uber H3 can decompose spatio-temporal objects at different resolutions without being constrained by the background. S2 and Geohash use quadrilaterals for division, while Uber H3 uses hexagons for division. We compared these three grid encoding systems based on projection method, isotropy, and hierarchical coverage. The results are shown in Table 1.

In this study, the isotropy of a grid refers to the consistency of its weights with adjacent grids. As shown in Figure 2, quadrilateral grids have two types of neighboring grids, copoint and colinear neighbors, while hexagonal grids have only colinear neighbors. The isotropy of hexagonal grids is superior to that of quadrilateral grids. Since moving objects in the real world is unlikely to involve movement in a way that aligns with the grid, using Google S2 can complicate the analysis because the analyst needs to consider more different types of neighbors than Uber H3 does [39]. In terms of hierarchical coverage, quadrilateral grids can achieve precise coverage between levels, while hexagonal grids cannot. The Uber

H3 grid system uses a seven-aperture hierarchical division method, and the approximate coverage between adjacent levels is achieved by rotating the angle.

Grid Systems	Projection	Isotropy	Hierarchical Coverage
Google S2	Regular hexahedron Orthoaxial	Copoint, Colinear	Accurate
Geohash	equiangular cylindrical	Copoint, Colinear	Accurate
Uber H3	Regular icosahedron	Colinear	Approximate

Table 1. Global grid encoding systems comparison.

We chose to use Uber H3 for the discretization of spatio-temporal objects, mainly for the following three reasons: First, the smaller projection error of Uber H3 grids can more accurately achieve the discretization of spatio-temporal objects [39]. Second, the isotropy of hexagons simplifies the construction of spatial neighbor relationships in association analysis: in the calculation process, grid distance can be used instead of geographic distance. Third, although the hierarchical coverage of hexagonal grids is inferior, the aggregation range of different level neighborhood information matches the resolution of that level. As such, the precise coverage between levels is not required.



Figure 2. The difference between quadrilateral and hexagonal grids in terms of isotropy.

Definition 4. Structural relationships among geographic grids. There are structural relationships between multilevel geographic grids, including adjacency relationships between grids at the same level, and hierarchical relationships between adjacent-level grids.

As shown in Figure 3a, the grids that have adjacency relationships with the red grid on the L-level are the blue grids of the same level. The yellow grids at level L - 1 and L + 1 have hierarchical relationships with the red grid. Among them, the yellow grids at level L - 1 are the parent grids of the red grid and its neighboring grids, while the yellow grids at L + 1 are the child grids of the red grid.

Definition 5. Mapping relationship between spatio-temporal objects and grids. There is a mapping relationship between spatio-temporal objects and their discretized grids, i.e., a spatio-temporal object can be mapped to a set of geographic grids.

As shown in Figure 3b, the blue line represents the trajectory, and the red grids represent the discretization result of the trajectory at a certain level. There exists a mapping relationship between the trajectory and the grids in the figure.



Figure 3. Structural relationship and mapping relationship.

3.3. Heterogeneous Graph

Heterogeneous graphs are capable of expressing associations between different types of spatio-temporal objects by utilizing node types. This feature enables researchers to design custom sampling schemes that suit their specific requirements, making it an effective tool for fusing multiple types of nodes with association information. Some important concepts related to heterogeneous graphs and their embedding are as follows:

Definition 6. Heterogeneous Graph. The heterogeneous graph G_H can be expressed as $G_H = \{V, E, T\}$, where V is the set of all nodes in G_H , E is the set of all edges in G_H , ϕ is the node v-type mapping $\phi(v) : V \to T_V$, ψ is the edge e-type mapping $\psi(e) : E \to T_E$, T_V and T_E are the sets of node type and edge type, respectively, while $|T_V| + |T_E| > 2$.

Definition 7. Heterogeneous Graph Embedding. For a heterogeneous graph $G_H = \{V, E, T\}$ with its node attribute matrix $X_{T_{Vi}} \in \mathbb{R}^{|V_{T_{Vi}}| \times d_{T_{Vi}}}$, the goal of heterogeneous graph embedding is to obtain the node embedding representation $h_v \in \mathbb{R}^d$ for all $v \in V$ by learning so that h_v can reflect the structural and semantic information of the graph G, where $d \ll |V|$, \mathbb{R}^d denotes the *d*-dimensional Euclidean space.

Definition 8. Association Strength. For spatio-temporal objects O_a and O_b , whose embedding representations are h_{O_a} and h_{O_b} , respectively, the association strength $I(h_{O_a}, h_{O_b})$ between O_a and O_b can be expressed as Equation (1).

$$I(h_{O_{a}}, h_{O_{b}}) = \frac{h_{O_{a}}{}^{T}h_{O_{b}}}{\|h_{O_{a}}\| \cdot \|h_{O_{b}}\|}, I \in [-1, 1]$$
(1)

According to Equation (1), we define the spatio-temporal object association analysis problem as follows:

Definition 9. Spatio-Temporal Object Association Analysis. Given the set of spatio-temporal objects $Os = \{O_1, O_2, \dots, O_i, \dots, O_k\}, \forall O_i \in Os$, the purpose of spatio-temporal object association analysis is to find the set of spatio-temporal objects $Os_{sub} = \{O_{sub_1}, O_{sub_2}, \dots, O_{sub_i}\}, (O_{sub_1}, O_{sub_2}, \dots, O_{sub_i}), (O_{sub_i}, \dots, O_{sub_i}), (O$

 $Os_{sub} \subset Os$, and $O_i \notin Os_{sub}$, such that $\forall O_{sub_m} \in Os_{sub}$, $O_{sub_n} \in Os_{sub}$, n > m; we can thus obtain Equation (2):

$$I(h_{O_i}, h_{O_{sub_m}}) > I(h_{O_i}, h_{O_{sub_m}}) > \chi$$
⁽²⁾

where h is the embedding representation, and χ is the association strength threshold, $0 \leq \chi < 1$.

4. Method

In this section, we propose a new multiscale spatial-temporal object representation method, named STO2Vec, for association analysis. The overall structure of the method will be introduced in Section 4.1. The process of data preprocessing is described in Section 4.2 while the construction method of the spatio-temporal association heterogeneous graph is presented in Section 4.3. Finally, we introduce the heterogeneous graph node embedding algorithm for spatio-temporal association in Section 4.4.

4.1. Overall Framework

To address the problem of multiscale spatial-temporal object association measurement and discovery, as shown in Figure 4, we propose a multiscale spatial-temporal object representation method named STO2Vec for association analysis. The method is divided into two stages: graph construction and embedding. In the graph construction stage, we decompose spatial-temporal objects into grids at multiple levels using adaptive discretization and describe their associated relationships using a heterogeneous graph that includes discrete spatio-temporal co-occurrence and grid structural relationships. In the embedding stage, biased second-order sampling is performed to obtain node sequences based on the association between homogeneous or heterogeneous spatial-temporal objects. These sequences are combined with an embedding model to generate vector representations of the nodes, facilitating association discovery and measurement between objects.



Figure 4. Overall Framework of STO2Vec.

4.2. Data Preprocessing

We first remove the redundant values from the trajectories in the preprocessing stage. Additionally, we apply the well-established Kalman filtering algorithm to suppress noise. Given that interactions mainly occur locally and the full trajectory can potentially dilute the association information of moving objects at specific points in time, the cleaned trajectory needs to be segmented to describe the association between spatio-temporal objects in as a precise way as possible, which is a common preprocessing operation for most trajectory analysis algorithms [41].

Traditional stopping point segmentation methods are not applicable to the association analysis of moving objects such as aircraft and ships. Due to cost saving or route planning,

the movement characteristics such as heading, speed, and altitude of such moving objects tend to be kept stable in the process of transferring and cruising. Nevertheless, their movement characteristics occasionally fluctuate rapidly while interacting with geographic entities. To capture associations more precisely, we extract trajectory segments that exhibit frequent variations in movement features. Joint information entropy is a reliable metric that measures the uncertainty of multiple random events. By considering changes in mobile traits as a sequence of random events, we segment trajectories using a unified joint information entropy metric of mobile features. The algorithmic procedure is as follows.

For each point a_i in the trajectory $Tr = \{(a_1, a_2, ..., a_i, ..., a_n) \mid a_i = (p_i, t_i)\}$, the motion features of f_{d_i} (heading) and f_{v_i} (instantaneous velocity) are first calculated. Other features such as geographic altitude, barometric altitude, and attitude parameters can also be added according to the data situation and background.

Then, the motion features of *Tr* are clustered by numerical distribution. We use the widely-used DBSCAN algorithm to perform clustering. We obtain the set C_d of heading feature clusters consisting of *k* clusters and the set C_v of instantaneous velocity clusters consisting of *l* clusters, where $C_d = \{c_{d_1}, c_{d_2}, \ldots, c_{d_i}, \ldots, c_{d_k} \mid c_{d_i} \subset Tr, 0 < k < n\}$, $C_v = \{c_{v_1}, c_{v_2}, \ldots, c_{v_i}, \ldots, c_{v_k} \mid c_{v_i} \subset Tr, 0 < l < n\}$. The motion feature label of each point is $L_{a_i} = (c_{d_i}, c_{v_i}) \mid c_{d_i} \in C_d, c_{v_i} \in C_v$.

To obtain the joint information entropy of motion characteristics for trajectory points within each window, for a sliding window W with window size s_w ($0 < s_w < n$) and sliding step e_w ($0 < e_w < n$), Equation (3) is sequentially applied in order of trajectory sequence.

$$H_w(c_{d_i}, c_{v_i}) = -\sum_{c_{d_i} \in C_d} \sum_{c_{v_i} \in C_v} P(c_{d_i}, c_{v_i}) \log_2 P(c_{d_i}, c_{v_i})$$
(3)

where $P(c_{d_i}, c_{v_i})$ is the joint probability of c_{d_i} , and c_{v_i} is within the window W. We then extract the set $W_T = \{w_1, w_2, \ldots, w_j, \ldots, w_m \mid w_j \subset Tr, 0 < m < n\}$ of windows whose entropy values exceed the threshold h_w .

Finally, the consecutive adjacent windows in W_T are merged to obtain the set $S_{Tre_{ita}}$ of trajectory segments which has potential interaction semantics. The remaining trajectory segments in the original trajectory are composed into the set $S_{Tre_{mov}}$ so that the final trajectory segmentation results are obtained.

4.3. Graph Construction

4.3.1. Adaptive Discretization

Previous uniform grid division methods discretize all spatio-temporal objects with the same resolution. When the grid resolution is low, this leads to the loss of a large number of geometric features and association details. This can result in a decrease in embedding quality. Using higher-level grids to process spatio-temporal objects incurs significant computational costs [42]. Discretization methods at the same level are limited in their ability to describe features of spatio-temporal objects at multiple scales, whereas multilevel partitioning methods require prior knowledge to set the level parameters. To this end, we propose an adaptive discretization approach to differentially dissect spatio-temporal objects at different scales based on H3, preserving more detailed features while taking into account computational overhead.

For the spatio-temporal object O_i , let its discretization level be r. We examine the variables that affect the accuracy and computational cost during the discretization process.

- The number of grids $N_{O_i}(r)$: $N_{O_i}(r)$ is the total number of grids after discretization of the spatio-temporal object O_i at level r. Obviously, the accuracy of the discretization O_i increases as r increases, which also leads to an increase in $N_{O_i}(r)$ and the computational overhead.
- Discretization error $Err_{O_i}(r)$: We use $Err_{O_i}(r)$ to measure the information loss brought by discretization to the spatio-temporal object description, which is calculated as in

Equation (4). For polygon elements, inspired by the error analysis method of rasterization of vector elements [43], we choose to use the area relative error $Err_{O_i}^{poly}(r)$ to calculate (Equation (5)), where $S_{orig}^{O_i}$ is the area before discretization and $S_{disc}^{O_i}(r)$ is the area after discretization. For line elements, we generate the buffer $B_{b_l}(O_i)$ of line elements with radius b_l and then use the area relative error for calculation (Equation (6)). For point elements, a uniform grid resolution r_p is used for discretization.

$$Err_{O_i}(r) = \begin{cases} Err_{O_i}^{poly}(r), \text{ if } O_i \text{ is Polygon} \\ Err_{O_i}^{line}(r), \text{ if } O_i \text{ is Line String} \\ 0, \text{ if } O_i \text{ is Point} \end{cases}$$
(4)

$$Err_{O_i}^{poly}(r) = \left| \frac{S_{orig}^{O_i} - S_{disc}^{O_i}(r)}{S_{orig}^{O_i}} \right|, \left(S_{orig} \neq 0 \right)$$
(5)

$$Err_{O_{i}}^{line}(r) = \left| \frac{S_{orig}^{B_{b_{l}}(O_{i})} - S_{disc}^{O_{i}}(r)}{S_{orig}^{B_{b_{l}}(O_{i})}} \right|, \left(S_{orig}^{B_{b_{l}}(O_{i})} \neq 0 \right)$$
(6)

Algorithm 1: Adaptive Discretization Algorithm

Input :Spatio-Temporal Object Set Os, Level Parameter hr_s, Buffer Radius b₁ Output: Discretization Level Set R 1 /*Initialization*/ 2 $R[n][m] \leftarrow \emptyset;$ 3 /*Discretization*/ 4 foreach O_i of Os do $r_m \leftarrow <$ Level upper limit>; 5 foreach $r_i \in (r_m, r_m + hr_s)$ do 6 *countDict*[*j*][*c*] \leftarrow <The number of discretized cells under *r*_{*i*}>; 7 if type of O_i is Line then 8 $errDict[j][e] \leftarrow <$ Get the error according to Equation (6)>; 9 else 10 $errDict[j][e] \leftarrow <$ Get the error according to Equation (5)>; 11 end 12 end 13 *normMergeSet* \leftarrow <Combine *countDict* and *errDict* by normalization>; 14 $R_i \leftarrow \langle \text{Key of the minimum value in } normMergeSet \rangle;$ 15 $R[i][r] \leftarrow R_i$ 16 17 end 18 return R

Thus, the adaptive discretization problem of multiscale spatio-temporal objects can be transformed into an optimization problem with constraints of the form shown in Equation (7).

$$\arg\min_{r} \left(N_{O_i}^{norm}(r) + Err_{O_i}^{norm}(r) \right)$$
(7)

where $N_{O_i}^{norm}(r)$ and $Err_{O_i}^{norm}(r)$ are the normalized $N_{O_i}(r)$ and $Err_{O_i}(r)$, respectively. For the spatio-temporal object O_i , we only need to consider the level of the grids that are close in scale to it. Therefore, we first obtain the approximate scale range of the spatio-temporal object by calculating the convex area of the spatio-temporal object $S_{O_i}^{conv}$. We choose the level $r_{max(O_i)}$ with the smallest difference between the average area of the grid [39] and $S_{O_i}^{conv}$ as the upper limit of the search. Search down according to the parameter hr_s and let *D* be the feasible domain of Equation (7); as shown in Equation (8), there exists such a constraint.

$$D = \left\{ r \mid r \in \left[r_{\max(O_i)}, r_{\max(O_i)} + hr_s \right] \right\}$$
(8)

As the discretization level for points is the default, we only give the pseudo-code of the discretization algorithm for lines and polygons as Algorithm 1.

4.3.2. Heterogeneous Graph Model

To describe the associations between multiscale spatio-temporal objects, we map geographic entities and trajectories into a heterogeneous graph. Unlike previous studies, we use heterogeneous rather than homogeneous graphs for the following reasons: Rather than embedding towards the grid, we embed different spatio-temporal objects. We describe the spatio-temporal co-occurrence between spatio-temporal objects through the grid and use the structural relationships of the grid to associate different levels. The nodes in the heterogeneous graph have different categories, which can describe the interaction and association between different spatio-temporal objects. The association semantics between different types of nodes can be explored according to specific requirements. This enhances the generality and scalability of the method and allows for designing sampling algorithms tailored to different applications.

As shown in Figure 5, there are 5 kinds of nodes in the heterogeneous graph $G_r = \{V, E, T\}$, namely $T_V = \{M, S, H, P, G\}$, where *M* is the moving object node, *S* is the trajectory segment node, *H* is the grid node, *P* is the subgeographic entity node, and *G* is the geographic entity node. *M* and *G* are collectively referred to as spatio-temporal object nodes, while *S* and *P* are collectively referred to as subobject nodes. Since we are not concerned about the association of moving objects with geographic entities other than spatio-temporal interactions, we describe moving objects by their trajectories. When the trajectory or geographic entity does not exist for a subobject, it is represented by itself instead.



Figure 5. Illustration of the Association Heterogeneous Graph Model.

Since spatio-temporal objects are discretized into different levels of the grid, we have to consider the structural relations of the grid itself. Additionally, we need to consider the spatio-temporal co-occurrence-based associations between the spatio-temporal objects. The edge $T_E = \{R_{MS}, R_{SH}, R_{PH}, R_{GP}, R_{H_n}, R_{H_l}, \}$ in the graph consists of two parts. One part is the containment relationship of spatio-temporal objects and their mapping relationship with the geographic grid, including the edge R_{MS} between moving objects and trajectory segments, the edge R_{SH} between the trajectory segments and geographic grid, the edge R_{PH} between the subgeographic entities and geographic grid, and the edge R_{GP} between geographic entities and subgeographic entities. There is a weight λ between the edges, where $\lambda_{SH} \ge 1$ and $\lambda_{PH} \ge 1$, indicating the number of times this grid appears in the discretization result of the subobject. This value can reflect the association strength of spatio-temporal objects about a certain grid region. The weights $\lambda = 1$ for all edges except R_{SH} , R_{PH} , as shown in Figure 6. The other part is the structural relationship of the geographic grid itself, as shown in the dashed box of Figure 6, including the adjacency relationship R_{H_n} between the grid and its surrounding grids (indicated by red lines) as well as the hierarchical relationship R_{H_l} between the grid and its neighboring hierarchical grids (indicated by blue lines). For example, for grid H_2 , owing to the excellent isotropy of the hexagonal grid, the adjacency relation R_{H_n} with H_1 and H_3 , and the hierarchy relation R_{H_l} with the upper grid H_7 and the lower grid H_{10} can be obtained quickly.



Figure 6. Example of Association Heterogeneous Graph. The moving object M_a has trajectory segments S_{a_1} and S_{a_2} , where S_{a_1} is discretized with grids H_1 , H_2 , H_3 . The geographic entity G_a has subgeographic entities P_{a_1} and P_{a_2} , where P_{a_1} is discretized with grids are H_7 , H_8 .

4.4. Embedding

Unlike the previous approach of using feature engineering based on expert knowledge, we obtain the vector representation of each spatio-temporal object by node embedding. This not only captures the interaction information between different spatio-temporal objects in the embedding process but also preserves the spatial proximity information of these objects. We introduce the objective function of embedding learning in Section 4.4.1. A heterogeneous graph sampling algorithm based on biased wandering is proposed in Section 4.4.2. This algorithm is used to implement node embedding for different application contexts.

4.4.1. Objective Function

Our objective is to obtain an embedding representation for objects in a finite spatiotemporal range. The strength of association among these objects is reflected by the cosine similarity of their embedding vectors: the stronger the association is, the higher the cosine similarity. The strength of spatio-temporal association can be reflected by the spatiotemporal co-occurrence frequency [7], which is similar to the learning of word vector representation obtained by word co-occurrence laws in natural language processing. Therefore, we introduce the Skip-gram [44] model in conjunction with the heterogeneous graph constructed in Section 4.3. Given the heterogeneous graph $G_{Hr} = \{V, E, T\}$, the objective function of Skip-gram can be expressed as Equation (9):

$$\arg\max_{\theta} \sum_{v \in V} \sum_{t \in T_V} \sum_{c_t \in N_t(v)} \log p(c_t \mid v; \theta)$$
(9)

where N(v) denotes the set of neighborhood nodes of node v, and $N_t(v)$ denotes the set of type t neighborhood nodes of node v. For example, the set of type P nodes of H_4 in Figure 6

is $\{P_{a_2}\}$. $p(c_t | v; \theta)$ denotes the probability of occurrence of node c_t given node v. For $p(c_t | v; \theta)$, a softmax function is usually used to define [45]: $p(c_t | v; \theta) = \frac{e^{X_{c_t} \cdot X_v}}{\sum_{u \in V} e^{X_{u} \cdot X_v}}$, where X_v is the v-th row of the matrix X, namely the embedding vector of node v. To improve the computational efficiency, we adopt the negative sampling strategy for optimization and set the number of negative samples as K. In this case, Equation (9) can be converted into Equation (10):

$$\log \sigma(X_{c_t} \cdot X_v) + \sum_{k=1}^{K} \mathbb{E}_{u^k \sim P(u)} \log \sigma(-X_{u^k} \cdot X_v)$$
(10)

where $\sigma(x) = \frac{1}{1+e^{-x}}$, P(u) is a predefined probability distribution, and u^k is the *k*th negative sample.

4.4.2. Biased Sampling

Since random walks in heterogeneous graphs tend to capture highly visible nodes [46], a meta-path-based algorithm is usually used to constrain sampling in heterogeneous graph embeddings. This can result in sequences of nodes that more accurately reflect information about the structure of the graph. Some scenarios focus more on associations between homogeneous spatio-temporal objects, such as similarity analysis of trajectories and functional analysis of urban areas. In contrast, in environmental management and crime investigation, more attention is paid to associations between heterogeneous spatio-temporal objects. However, realistic data are usually a mixture of several different associations. Similar to the concept of structural equivalence and homogeneous community semantic embedding in node2vec [47], we propose a second-order biased sampling algorithm that uses hyperparameters and multiple meta-paths to generate corresponding sampling strategies for different spatio-temporal associations, enabling diverse association analysis.



Figure 7. Illustration of Biased Sampling.

The relationship between trajectories and trajectory segments, as well as between geographic entities and subgeographic entities, is a straightforward containment relationship. This allows us to acquire the embedded representations of moving objects and geographic entities by merging their respective subobjects. Therefore, it is only necessary to analyze within the subgraph $G_{Hr_{sub}} = \{V_{sub}, E_{sub}, T_{sub}\}$ of G_{Hr} , where $T_{subv} = \{S, H, P\}$ and $T_{sub_E} = \{R_{SH}, R_{PH}, R_{H_n}, R_{H_l}\}$. For $G_{Hr_{sub'}}$ given a starting node v and a sampling sequence length l_w , the sampling algorithm will generate a sequence starting with v and containing l_w nodes, where the *i*-th step transfer probability can be expressed as Equation (11).

$$p\left(v^{i+1} \mid v_t^i\right) = \begin{cases} \frac{\pi v_t^i v^{i+1}}{Z}, \left(v^{i+1}, v_t^i\right) \in E_{\text{sub}}\\ 0, \text{ otherwise} \end{cases}$$
(11)

where $v_t^i \in V_t$, $\pi_{v_t^i v^{i+1}}$ is the non-normalized transfer probability, and *Z* is the normalization constant.

Given that the connection among spatio-temporal objects is wide-ranging and nonexclusive, a singular meta-path may skew the embedding reflection of the actual association status. Therefore, we do not adopt predefined meta-paths directly. Instead, we utilize hyperparameters to manage sampling bias and augment the association information captured between homogeneous or heterogeneous objects within a confined range. As shown in Figure 7, assuming that the current node sampled is $v_{t_s}^s$ (i.e., node *s* in the figure), $\phi(v_{t_s}^s) = t_s$, whose last visited node is $v_{t_b}^b$ (i.e., node *b* in the figure), the next visited node is $v_{t_x}^s$. The transfer probability of visiting $v_{t_x}^x$ from $v_{t_s}^s$ is defined as $\pi_{v_{t_t}^s v_{t_x}^x}^x = \xi\left(v_{t_b}^b, v_{t_x}^x\right) \cdot \lambda_{v_{t_s}^s v_{t_x}^x}$, where $\lambda_{v_{t_s}^s v_{t_x}^x}$ is the weight and $\xi\left(v_{t_b}^b, v_{t_x}^x\right)$ is the sampling probability. When $t_s \neq H$, $\forall v_n \in N(v_{t_s}^s), \phi(v_n) = H$, where $\pi_{v_{t_t}^s v_{t_x}^s}^x = \lambda_{v_{t_s}^s v_{t_x}^x}$. In contrast, when $t_s = H$, as shown in Equation (12), we use the homogeneous parameter *m*, the heterogeneous parameter *n*, and the spatial parameter *k* to control the sampling probability.

$$\xi\left(v_{t_b}^b, v_{t_x}^x \mid \phi(v_{t_s}^s) = H\right) = \begin{cases} \frac{1}{m}, & \text{if } t_b = t_x \text{ and } t_x \neq H\\ \frac{1}{n}, & \text{if } t_b \neq t_x \text{ and } t_x \neq H\\ \frac{1}{k}, & \text{if } t_b \neq t_x \text{ and } t_x = H\\ 0, & \text{if } t_b = t_x \text{ and } t_x = H \end{cases}$$
(12)

where $(v_{t_s}^s, v_{t_x}^x) \in E_{sub}$, and the sampling probability is 1/m when $v_{t_b}^b$ and $v_{t_x}^x$ are homogeneous spatio-temporal object nodes. The sampling probability is 1/n when $v_{t_b}^b$ and $v_{t_x}^x$ are heterogeneous spatio-temporal object nodes; in order to obtain the neighborhood information and hierarchical information of the multilevel grid, we control the sampling of the wandering algorithm among the grid nodes by the spatial parameter k. When $v_{t_b}^b$ is a spatio-temporal object node and $v_{t_x}^x$ is a grid node, the sampling probability is 1/k. When $v_{t_b}^b$ is a grid node, we no longer perform the sampling of the neighboring grid so that the association aggregation around the sparse region can be controlled within a reasonable range. In addition, since we are oriented toward spatio-temporal object embedding rather than grid embedding, the grid nodes will be filtered out in the final node sequence so that only spatio-temporal object nodes are retained.

5. Experiment

We tested the effectiveness of our proposed method in discovering and measuring associations between spatio-temporal objects, both homogeneous and heterogeneous, through two quantitative experiments. To visualize the results of the association analysis, we present a case study of it in a visual way.

5.1. Experiment Setup

5.1.1. Data Preparation

We generated a 60 km buffer radius for the US mainland region as a test area for the experiment, with an area size of 9.705×10^6 km², which contains a large number of spatio-temporal objects at different scales. For the geographic entity data, we used the Natural Earth dataset [48] by selecting a part of the data as geographic entity objects, supplemented with OurAirports dataset [49] as airport data. The details are summarized in Table 2.

Geographic Entity	Count
Airport	27,675
Coastline	192
Lakes	418
Parks and protected lands	148
Railroad	873
Rivers	1391
River and lake center lines	314
Roads	39,918
State	254
Urban	1120

Table 2. Statistics of the Geographic Entity Dataset.

For moving object data, we used the open source ADS-B dataset provided by OpenskyNetwork [50], a crowdsourcing-based nonprofit receiving network that has been continuously collecting air traffic monitoring data since 2013. We selected four days of flight trajectory data in June 2022. Since trajectories with lower flight altitudes have a higher probability of being interactively associated with geographic entities, we randomly selected some of the flight trajectories with flight altitudes below 1500 m. After preprocessing according to Section 4.2, we divided the trajectories into datasets *A*, *B*, and *C*. Among these, the International Civil Aviation Organization (ICAO) 24-bit aircraft address in dataset *B* and *C* are the same, among which each aircraft has flight records in these 4 days. The preprocessed moving object dataset summary characteristics are shown in Table 3.

Table 3. Preprocessed moving object dataset summary characteristics.

Dataset	Α	В	С
Date	20-06	06-06, 13-06, 20-06	27-06
Num.Aircraft	5661	900	900
Num.Traces	5661	2753	919
Num.Traces Seg	50,321	41,846	13,166
Avg Num.Points	330.757	519	522.534
Avg Trace Length ¹	276.504	464.043	482.19

¹ Data in minutes.

5.1.2. Parameter Setting

In the association graph construction stage, we set the adaptive discrete hierarchical parameter hr_s to 3 so that for each spatio-temporal object, only 3 calculations are needed to obtain a relatively appropriate discretization level. As the point elements are all airports and aprons, whose areas ranges between approximately 0.001 km² and 10 km², we set the default level r_p for the point elements to 9, which has a mean grid area of 0.1053 km². We set the buffer radius b_l to 5 m to preserve as much of the line element geometry as possible.

In the embedding phase, we set the vector dimension *d* of the embedding to 256 in order to retain more association information. Other hyperparameters were set as follows: learning rate $lr = 2.5 \times 10^{-3}$, window size span = 5, number of sampled sequences of a single node 100, length of sampled sequences $l_w = 100$, and the number of negative samples K = 5.

5.1.3. Evaluation Metrics

We tested the association analysis in terms of both association discovery and association metrics. There was no sequential relationship between the results of association discovery. We used Hitting Ratio in Top *K* list (HR@K) [51] to evaluate the results. HR@K was defined as shown in Equation (13):

$$HR@K = \frac{1}{|S_{te}|} \sum_{\kappa \in S_{te}} \frac{|L_{\kappa}^{S}@K \cap L_{\kappa}^{K}|}{|L_{\kappa}^{R}|}$$
(13)

where κ refers to the spatio-temporal object under the test, S_{te} is the set of spatio-temporal objects under the test, L_{κ}^{R} is the list of spatio-temporal objects in the ground truth that have an association with κ , and $L_{\kappa}^{S}@K$ is the list of the top K spatio-temporal objects with the strongest association to κ output by the model.

The list of results of association metric tests can be sorted according to the strength of association. We used normalized discounted cumulative gain at a particular rank K(NDCG@K) [35] to evaluate the association metric results, which was defined as shown in Equation (14):

$$NDCG@K = \frac{DCG@K}{IDCG@K}$$
(14)

where IDCG@K is the ideal DCG@K value, which is the DCG@K value of the ground truth list, and DCG@K is defined as shown in Equation (15).

$$DCG@K = \sum_{i=1}^{K} \frac{2^{rel_i} - 1}{\log_2(i+1)}$$
(15)

where *rel_i* is the order association of test results at *i*. For this experiment, *rel_i* takes 0 and 1. For the ground truth, we will illustrate its generation in the specific experiments.

5.1.4. Baseline

To assess the validity of STO2Vec, the following three models with similar research questions were used as baseline for comparison with STO2Vec.

- Mot2vec [28]: The algorithm is based on the Word2vec model, which uses the trajectories of pedestrians moving between geographic entities to construct a behavioral representation of locations. The algorithm does not use grid partitioning but rather generates IDs of geographic entities based on point clustering and then converts the trajectories into ID sequences for embedding. According to the data characteristics, we use a time window of *timestep* = 5 *min* in the preprocessing stage to label the trajectories with geographic entities, with the minimum spatial resolution being 5 km.
- Hier [32]: This similar algorithm uses a multilevel embedding grid, which aggregates rectangular grid vectors of different levels into fine-grained grids during embedding. Due to the large study area in this paper, we used 100 km and 10 km grid cells in the first and second level, respectively. The fine-grained level uses 1 km grid cells. The embedding dimension is the same as STO2Vec, the first 48 dimensions correspond to 100 km grid, the last 80 dimensions correspond to a 10 km grid, the remaining 128 dimensions correspond to a 1 km grid.
- GCN-L2V [31]: To generate fine-grained grid embeddings, the GCN and Skip-gram models were used to construct spatial graphs and flow graphs to account for both spatial proximity and the movement patterns of moving objects. The algorithm uses a single-level Google S2 grid system. According to the scale of the study area, we mapped trajectories and geographic entities to the 11-level (the average area of each grid area is about 20.2682 km²) of Google S2 to generate edges between grids in spatial graphs with a distance threshold of 1 km.

5.2. Homogeneous Association Analysis

To evaluate the effectiveness of association analysis between homogeneous spatiotemporal objects, we examined the effect of the algorithm on the association analysis between moving objects and geographic entities. This evaluation was conducted through two applications: *Trajectory Similarity Analysis* and *Region Association Analysis*. We constructed the association heterogeneous graph based on dataset *A* and geographic entity dataset, in which there are 1,854,380 nodes and 4,986,765 edges. We then performed homogeneous biased sampling in the heterogeneous graph and set the homogeneous parameter to m = 2, the heterogeneous parameter to n = 4, and the spatial parameter to k = 4. Finally, we realized the embedding representation of nodes by embedding the model and calculated the association strength between homogeneous spatio-temporal objects.

5.2.1. Trajectory Similarity Analysis

Ground Truth Generation. For moving objects, the similar association of their trajectories is a typical spatio-temporal association. By comparing the similarity analysis results of trajectories, we could verify the effect of the algorithm in association metrics and discovery between moving objects. There are many existing mature algorithms for trajectory similarity metrics based on geometric features, which accurately calculate the similarity between trajectories by matching point by point. Therefore, we used the dynamic time warping algorithm [52] to generate the similarity matrix between trajectories on the dataset *A*. We then created the experimental ground truth according to the order of similarity, ensuring that all trajectories had a similarity greater than 0.7.

Result for Trajectory Similarity Analysis. We randomly selected 1000 trajectories from the dataset *A*, using different models to generate a list of similar trajectories for comparison with the ground truth. Since the results of the trajectory similarity metric are order sensitive, we used *NDCG@K* to evaluate the experimental results. The results are shown in Table 4.

Method	NDCG@10	NDCG@20	NDCG@30	NDCG@40	NDCG@50
Mot2Vec	0.3325	0.3741	0.4021	0.4416	0.4592
Hier	0.265	0.3327	0.3987	0.4122	0.4157
GCN-L2V	0.3755	0.4312	<u>0.498</u>	0.5152	<u>0.516</u>
STO2Vec	0.5468	0.6205	0.6562	0.671	0.6752

Table 4. Results of Trajectory Similarity Analysis.

The best results are shown in **bold** and the second best results are <u>underlined</u>.

The experimental results show that STO2Vec achieves the best performance. This is because STO2Vec segments trajectories based on the joint information entropy and discretizes trajectory segments according to different spatial resolutions, addressing the problem of uneven spatio-temporal semantic distribution and discovering local similar associations among trajectories better. Mot2Vec converts trajectories into geographic entity sequences at equal time intervals, losing some geometric feature information of the trajectories. Hier also uses a multilevel discretization strategy but needs to set a fixed hierarchical structure based on prior knowledge, without considering the scale of specific spatio-temporal objects. GCN-L2V divides the study area into uniform fine-grained grids but still cannot adaptively discretize the trajectories when facing spatio-temporal objects of multiple scales, resulting in some degree of information loss. In addition, Hier and GCN-L2V do not segment trajectories but convert them into grid sequences using stop detection algorithms. These approaches are more suitable for mining trajectories with obvious stop semantics, such as pedestrian trajectories, while ignoring the association information during the *MOVE* phase for objects such as aircraft.

5.2.2. Region Association Analysis

Ground Truth Generation. According to the *first law of geography* [3] and the semantic characteristics of geographic entities, it is known that geographic entities with spatial proximity and semantic similarity are more strongly associated [30]. Region association analysis is often used to discover the functional structure of cities to help in transportation and urban planning. Due to the large-scale of the study area, we consider geospatial entities within the same *region* as clusters that have associations. We generate a ground truth by sorting these clusters according to their distance from each other. The *region* range was generated by the Urban dataset.

Results for Trajectory Similarity Analysis. There are a total of 16,073 other geographic entities with containment relationship with the Urban dataset, from which we randomly selected 2000 for testing. Ideally, the *K* geographic entities with the greatest association strength with the sample still belong to the region where the sample is located. Moreover, the closer the entities are, the higher the association strength. For this reason, we chose NDCG@K for evaluation. The experimental results are shown in Table 5.

From the experimental results, GCN-L2V has the best performance for NDCG@10 and NDCG@20, while STO2Vec has the best performance for NDCG@30, NDCG@40, and NDCG@50. This is due to the fact that GCN-L2V constructs spatial graphs by performing a large number of distance calculations to accurately describe the distance relationships between the grids. The larger the distance threshold parameter is, the higher the computational cost. The distance threshold parameter used in this experiment was 1 km. Therefore, in NDCG@10 and NDCG@20, the geographic entities with greater association strength could be accurately captured by GCN-L2V. However, as the value of K value increases and exceeds the distance threshold, grid association information beyond the threshold becomes difficult to capture, which affects the association accuracy of GCN-L2V. STO2Vec adopts a multilevel discretization strategy, which fully utilizes the structural relationships between grids and obtains more accurate association results over larger ranges. Although Hier also uses a multilevel approach for fine-grained grid embedding, the preset grid resolution cannot match the scale of spatio-temporal objects in the study area accurately, leading to the ignoring of spatial proximity between some geographic entities. Mot2Vec focuses on the association created by human movements between different geographic entities rather than on the spatial association of the entities themselves.

Table 5. Results of Region Association Analysis.

Method	NDCG@10	NDCG@20	NDCG@30	NDCG@40	NDCG@50
Mot2Vec	0.3125	0.3853	0.4152	0.428	0.4394
Hier	0.3875	0.4521	0.4817	0.4946	0.5084
GCN-L2V	0.4133	0.497	0.5156	0.5227	<u>0.5291</u>
STO2Vec	0.4096	<u>0.4918</u>	0.5262	0.5356	0.5434

The best results are shown in **bold** and the second best results are <u>underlined</u>.

5.3. Heterogeneous Association Analysis

The stronger the association between heterogeneous spatio-temporal objects, the higher the probability of their spatio-temporal co-occurrence. The relationship between moving objects and geographic entities can be used to predict object access patterns. Geographic entities that have a strong association with moving objects are more likely to be accessed by them. Therefore, we use *Visit Prediction* to test heterogeneous spatio-temporal objects. In practical applications, *Visit Prediction* can help people to dispatch and control based on the prediction results, which is important for public security management, airspace control, etc.

Ground Truth Generation. In our experiments, we used different models to output the *K* geographic entities most associated with moving objects, to test the effectiveness of the algorithm predictions by comparing the actual visit results. The actual access results, namely ground truth, are generated by extracting the geographic entities visited by each moving object from the dataset *C*.

Result for Trajectory Similarity Analysis. To obtain more information on historical visits, we construct a heterogeneous graph oriented to spatio-temporal association based on the *B* dataset and the geographic entity dataset, containing 1, 676, 516 nodes and 4, 206, 884 edges, setting the homogeneous parameter k = 1, the heterogeneous parameter m = 2 and the spatial parameter l = 0. Since the real visited list is not order sensitive, we used the hit rate HR@K as the evaluation metric for this experiment. The experimental results are shown in Table 6.

Method	HR@10	HR@30	HR@40	HR@50
Mot2Vec	12.63%	28.7%	33.57%	39.68%
Hier	16.76%	<u>35.66%</u>	<u>42.57%</u>	46.9%
GCN-L2V	<u>17.15%</u>	34.93%	41.39%	47.23%
STO2Vec	18.08%	40.71%	47.06%	51.42%

Table 6. Results of Heterogeneous Association Analysis.

The best results are shown in **bold** and the second best results are underlined.

As shown in Table 6, STO2Vec achieved the best results, while the experimental results of GCN-L2V and Hier were not significantly different, but the results of Mot2Vec were general. STO2Vec can regulate the random walk tendency based on task requirements by adjusting sampling parameters. In this way, it obtains more association information between heterogeneous spatio-temporal objects. This helps to prevent insufficient association information from being present in the embedding results due to differences in structure between those objects. The spatial parameters can aggregate the information around the grid, thus better solving the data sparsity problem. In addition, according to the scale of spatio-temporal objects, the multilevel discretization strategy can describe the geometric features of geographic entities and trajectories more accurately. It provides a more accurate corpus for training models aimed at "spatio-temporal co-occurrence" among heterogeneous objects. Compared with STO2Vec, the discretization method used by GCN-L2V and Hier cannot accurately describe trajectories and geographic entities of different scales. As a result, much local access information is ignored. During clustering, Mot2Vec loses some historical access information of geographic entities, resulting in less information obtained during the embedding process than STO2Vec.

5.4. Case Study

We used a case study to visualize the results of the association analysis in order to better verify the effectiveness of the algorithm. We still used the spatio-temporal association graph constructed in Section 5.2 for embedding learning. To verify the effect of the sampling algorithm on the association results, we modified the sampling hyperparameters by setting the homogeneous parameter to m = 4, the heterogeneous parameter to n = 2, and the spatial parameter to k = 4, thus generating embedding representation results oriented to heterogeneous spatio-temporal object associations. We chose to visualize the results of the association analysis in two areas with significant differences in scale. First, we queried and analyzed the geographic entities associated with the moving objects in the urban area where the geographic entities were dense with a smaller scale. Then, we queried and analyzed the moving objects associated with the geographic entities in the area around the coastline where the moving objects were dense with a larger scale.

5.4.1. Urban Association Analysis

Urban security management often requires the assistance of drones and helicopters, as shown in Figure 8 where the yellow dashed line shows the trajectory of a helicopter conducting security patrols in the urban area of Los Angeles. Los Angeles is the second largest city in the United States, and geographic entities in the urban area mainly include streets, railways, and airports. We selected the 150 geographic entities most strongly associated with the helicopter for visualization. Streets and railways are marked with red lines, and airports are marked with green dots. Moreover, a darker color indicates a stronger association.

Figure 8a shows the output of STO2Vec; the comparison shows that the associated geographic entities output by STO2Vec are more concentrated in spatial distribution. The strength of association basically matches the spatial characteristics of the trajectory, which can reflect the interaction between this helicopter and various types of geographic entities in urban areas more accurately. Owing to the fact that STO2Vec discretizes the trajectories by segmenting them to different spatial resolutions, it is able to adaptively match the scale

size of the trajectories and geographic entities. In data-intensive areas such as urban areas, fine-grained geographic entities are described by a high-level grid that retains most of the spatial characteristics of urban streets, railways, airports, and helicopter trajectories. Furthermore, as can be seen from the color shades of the elements and their corresponding association strength values, STO2Vec has a more dispersed distribution of association strength values than does GCN-L2V and Hier (e.g., 0.758–0.845 for airports and 0.711–0.889 for roads and railroads in Figure 8a); this indicates that the objects are more significantly different in STO2Vec embedding space. Without changing the association structure, the parameters can be set to adjust the sampling trend to the application requirements, which allows the model to learn more information about the interaction between the helicopter and the geographic entities.



(c) Hier

(d) Mot2Vec



As shown in Figure 8b,c, the output of GCN-L2V and Hier can reflect the association between moving objects and geographic entities to some extent. Both models take into account the context of surrounding neighboring areas. However, compared with those of STO2Vec, their associated geographic entities are spatially more dispersed. Since the scale of the upper grid in the Hier model is too large for spatio-temporal objects in the region, the model aggregates information from a large area around the trajectory during training. This can solve the problem of data sparsity but also affects the accuracy of local association analysis. GCN-L2V has difficulty associating smaller-scale objects such as airports. The 11-level grid of Google S2 (with an average grid area of approximately 20.2682 km²) is still relatively coarse for a geographic entity in an urban area. The fixed grid resolution

makes it difficult to take into account different scales of geographic entities such as railways and airports; therefore, the model learns less variation in features, and the distribution of association strength values is more concentrated. As shown in Figure 8d, the output of Mot2Vec is more dispersed, as most of the spatial proximity information of geographic entities is lost in the process of converting trajectories into ID sequences. For helicopters with strong mobility, the time window size of trajectory segmentation also affects the embedding accuracy.

For the area around the coastline, such as the Martha's Vineyard coastline area shown by the yellow line in Figure 9a, the island is approximately 32 km long and 3-16 km wide. Summer is the peak tourist season for the area. Therefore, the dataset contains a large number of flight trajectories for tourists traveling to the area and helicopter flight trajectories for sightseeing around the island. The red lines in Figure 9 show the trajectories of the associated moving objects. The top 40 trajectories with the strongest association to the island's coastline were selected for visualization. As shown in Figure 9a, since STO2Vec segments the trajectory based on joint information entropy, the model can output more accurate results in the form of trajectory segments. Trajectory segments of the moving object during the move phase around the island are most strongly associated with the coastline. This is due to the ability of STO2Vec to use a multilevel grid to portray the geometry of the coastline relatively accurately during the discretization of geographic entities. Moreover, the geographic entities in the area around the coastline are more sparse. The biased sampling algorithm allows the embedding results to be unaffected by the large number of intertrajectory associations. Segments of trajectories that fly around islands and hover near coastlines are able to generate rich co-occurrence contexts with coastline entities. Other trajectory segments that have basic topological associations, such as those adjacent or crossing, exhibited a weaker association strength due to the sparse structure of the association graph.

5.4.2. Coastline Association Analysis

The output of other models can only be visualized as trajectories, as shown in Figure 9b–d. Since geographic entities are sparse and trajectories are richer, a large number of association interactions are covered by similar or topological associations between trajectories. Therefore, the three models tend to discover flying trajectory-based associations between the island, the mainland, and adjacent islands. However, due to the uneven semantic distribution of trajectories, some of the stronger local associations (e.g., local flight around the island, local hovering) are diluted by the full trajectories. As a result, the trajectory segments that are strongly associated with the coastline in STO2Vec do not fully manifest in the output results of the three comparison models. Furthermore, considering spatio-temporal objects of different scales during selection of grid resolution is challenging for Hier and GCN-L2V. Consequently, their analysis may not be accurate or complete. Mot2Vec does not use a discretization approach to consider spatial associations between geographic entities. Accordingly, when data is sparse in specific regions such as the coastline, it relies on the IDs of distributed geographic entities along the continental coast for context rather than grid cells near the coastline.



(c) Hier

(d) Mot2Vec



6. Discussion and Conclusions

In this paper, we proposed a multiscale spatio-temporal object representation method, named STO2Vec, for association analysis. The method aims to address the issue of scale differences among spatio-temporal objects affecting the accuracy of analysis in the association analysis. To this end, we decompose spatio-temporal objects of varying scales into grids of different resolutions and obtain their representations by node embedding of an associated heterogeneous graph. During the embedding process, we enhance its scalability through biased sampling. Quantitative experiments demonstrated that STO2Vec could outperform other methods in various spatio-temporal association analysis, and visit prediction. This confirms that the resulting representation of STO2Vec retains more spatio-temporal association information by reflecting the scale differences. Additionally, the case analysis results demonstrated the effectiveness of STO2Vec in accurately measuring and discovering associations. These associations occur between spatio-temporal objects that differ in scale, highlighting the ability of STO2Vec to solve problems related to multiscale spatio-temporal object association analysis.

Representing geographic entities as point features can simplify computation but loses their spatial characteristics and most association information. Dividing spatio-temporal objects into fixed grids retains some of the features but struggles to consider scale differences or nonuniform regions. Presetting multilevel grids with prior knowledge partly alleviates scale difference impacts, but discerning suitable levels is difficult. Compared to other methods, STO2Vec provides better representation of various scales of spatio-temporal objects, enabling more accurate measurement and discovery of complex spatio-temporal association relationships. However, STO2Vec has limitations: representing object results in vector form—not grids—sacrifices flexibility, hindering incremental updates or aggregation while placing timeliness demands on data.

In the next step of our research, the focus will be on combining the behavior patterns of moving objects. The aim is to identify specific association semantics that exist between spatio-temporal objects. This will involve constructing a fine-grained association graph enabling deeper association analysis applications.

Author Contributions: Conceptualization, Luo Chen, Anran Yang and Nanyu Chen; methodology, Nanyu Chen; writing—original draft, Nanyu Chen; writing—review and editing, Luo Chen, Anran Yang, Wei Xiong and Nanyu Chen; funding acquisition, Luo Chen and Ning Jing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant no. 41971362 and 42101432), the Natural Science Foundation of Hunan Province (grant no. 2022JJ40546), and the Innovation Science Fund of NUDT (grant no. 22-ZZCX-058).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are publicly available on the open data portal of The OpenSky Network (https://opensky-network.org, accessed on 15 October 2022), NaturalEarth (https://www.naturalearthdata.com/, accessed on 10 November 2022), and OurAirports (https://ourairports.com/, accessed on 23 December 2022).

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

Abbreviations

The following abbreviations are used in this manuscript:

POI	point of interest
OD	origin destination
UTM	Universal Transverse Mercator
GPS	global positioning system
ICAO	International Civil Aviation Organization
HR	hitting ratio
NDCG	normalized discounted cumulative gain

References

- Hamdi, A.; Shaban, K.; Erradi, A.; Mohamed, A.; Rumi, S.K.; Salim, F.D. Spatiotemporal Data Mining: A Survey on Challenges and Open Problems. *Artif. Intell. Rev.* 2022, 55, 1441–1488. [CrossRef] [PubMed]
- 2. Sharma, A.; Jiang, Z.; Shekhar, S. Spatiotemporal Data Mining: A Survey. *arXiv* **2022**, arXiv:2206.12753. [CrossRef]
- 3. Tobler, W.R. A Computer Movie Simulating Urban Growth in the Detroit Region. Econ. Geogr. 1970, 46, 234. [CrossRef]
- 4. Wu, C.; Zhu, Q.; Zhang, Y.T.; Du, Z.Q.; Zhou, Y.; Xie, X.; He, F. An Adaptive Organization Method of Geovideo Data for Spatio-Temporal Association Analysis. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* 2015, *II-4/W2*, 29–34. [CrossRef]
- 5. Crivellari, A.; Ristea, A. CrimeVec—Exploring Spatial-Temporal Based Vector Representations of Urban Crime Types and Crime-Related Urban Regions. *ISPRS Int. J. Geo-Inf.* 2021, *10*, 210. [CrossRef]
- Riyadh, M.; Mustapha, N.; Riyadh, D. Review of Trajectories Similarity Measures in Mining Algorithms. In Proceedings of the 2018 Al-Mansour International Conference on New Trends in Computing, Communication, and Information Technology (NTCCIT), Baghdad, Iraq, 14–15 November 2018; pp. 36–40. [CrossRef]
- Cai, J.; Deng, M.; Guo, Y.; Xie, Y.; Shekhar, S. Discovering Regions of Anomalous Spatial Co-Locations. Int. J. Geogr. Inf. Sci. 2021, 35, 974–998. [CrossRef]
- Noureddine, H.; Ray, C.; Claramunt, C. A Hierarchical Indoor and Outdoor Model for Semantic Trajectories. *Trans. GIS* 2022, 26, 214–235. [CrossRef]
- Sakouhi, T.; Akaichi, J. Dynamic and Multi-Source Semantic Annotation of Raw Mobility Data Using Geographic and Social Media Data. *Pervasive Mob. Comput.* 2021, 71, 101310. [CrossRef]
- Kontarinis, A.; Zeitouni, K.; Marinica, C.; Vodislav, D.; Kotzinos, D. Towards a Semantic Indoor Trajectory Model: Application to Museum Visits. *Geoinformatica* 2021, 25, 311–352. [CrossRef]

- Alvares, L.O.; Bogorny, V.; Kuijpers, B.; de Macedo, J.A.F.; Moelans, B.; Vaisman, A. A Model for Enriching Trajectories with Semantic Geographical Information. In Proceedings of the 15th Annual ACM International Symposium on Advances in Geographic Information Systems—GIS'07, Seattle, WA, USA, 7–9 November 2007; p. 1. [CrossRef]
- Zhao, B.; Liu, M.; Han, J.; Ji, G.; Liu, X. Efficient Semantic Enrichment Process for Spatiotemporal Trajectories. Wirel. Commun. Mob. Commun. 2021, 2021, 4488781. [CrossRef]
- Vidal-Filho, J.N.; Times, V.C.; Lisboa-Filho, J.; Renso, C. Towards the Semantic Enrichment of Trajectories Using Spatial Data Infrastructures. ISPRS Int. J. Geo-Inf. 2021, 10, 825. [CrossRef]
- 14. Ibrahim, A.; Zhang, H.; Clinch, S.; Harper, S. From GPS to Semantic Data: How and Why—A Framework for Enriching Smartphone Trajectories. *Computing* **2021**, *103*, 2763–2787. [CrossRef]
- 15. Ying, J.J.C.; Lee, W.C.; Tseng, V.S. Mining Geographic-Temporal-Semantic Patterns in Trajectories for Location Prediction. ACM Trans. Intell. Syst. Technol. 2013, 5, 1–33. [CrossRef]
- Noureddine, H.; Ray, C.; Claramunt, C. Semantic Trajectory Modelling in Indoor and Outdoor Spaces. In Proceedings of the 2020 21st IEEE International Conference on Mobile Data Management (MDM), Versailles, France, 30 June–3 July 2020; pp. 131–136. [CrossRef]
- 17. Choi, D.W.; Pei, J.; Heinis, T. Efficient Mining of Regional Movement Patterns in Semantic Trajectories. *Proc. VLDB Endow.* 2017, 10, 2073–2084. [CrossRef]
- Wan, C.; Zhu, Y.; Yu, J.; Shen, Y. SMOPAT: Mining Semantic Mobility Patterns from Trajectories of Private Vehicles. *Inf. Sci.* 2018, 429, 12–25. [CrossRef]
- Fang, Z.; Du, Y.; Zhu, X.; Chen, L.; Gao, Y.; Jensen, C.S. Deep Spatially and Temporally Aware Similarity Computation for Road Network Constrained Trajectories. *arXiv* 2022, arXiv:2112.09339. [CrossRef]
- Karami, F.; Malek, M.R. Trajectory Similarity Measurement: An Enhanced Maximal Travel Match Method. *Trans. GIS* 2021, 25, 1485–1503. [CrossRef]
- Lehmann, A.L.; Alvares, L.O.; Bogorny, V. SMSM: A Similarity Measure for Trajectory Stops and Moves. *Int. J. Geogr. Inf. Sci.* 2019, 33, 1847–1872. [CrossRef]
- 22. Xiang, L.; Wu, T.; Ettema, D. An Intersection-Based Trajectory-Region Movement Study. Trans. GIS 2017, 21, 701–721. [CrossRef]
- 23. Sun, Z.; Jiao, H.; Wu, H.; Peng, Z.; Liu, L. Block2vec: An Approach for Identifying Urban Functional Regions by Integrating Sentence Embedding Model and Points of Interest. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 339. [CrossRef]
- Zhang, C.; Xu, L.; Yan, Z.; Wu, S. A GloVe-Based POI Type Embedding Model for Extracting and Identifying Urban Functional Regions. *ISPRS Int. J. Geo-Inf.* 2021, 10, 372. [CrossRef]
- 25. Zhang, J.; Li, X.; Yao, Y.; Hong, Y.; He, J.; Jiang, Z.; Sun, J. The Traj2Vec Model to Quantify Residents' Spatial Trajectories and Estimate the Proportions of Urban Land-Use Types. *Int. J. Geogr. Inf. Sci.* **2021**, *35*, 193–211. [CrossRef]
- Zhu, M.; Chen, W.; Xia, J.; Ma, Y.; Zhang, Y.; Luo, Y.; Huang, Z.; Liu, L. Location2vec: A Situation-Aware Representation for Visual Exploration of Urban Locations. *IEEE Trans. Intell. Transport. Syst.* 2019, 20, 3981–3990. [CrossRef]
- Du, J.; Chen, Y.; Wang, Y.; Pu, J. Zone2Vec: Distributed Representation Learning of Urban Zones. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 880–885. [CrossRef]
- Crivellari, A.; Beinat, E. From Motion Activity to Geo-Embeddings: Generating and Exploring Vector Representations of Locations, Traces and Visitors through Large-Scale Mobility Data. *ISPRS Int. J. Geo-Inf.* 2019, *8*, 134. [CrossRef]
- Jenkins, P.; Farag, A.; Wang, S.; Li, Z. Unsupervised Representation Learning of Spatial Data via Multimodal Embedding. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing China, 3–7 Novembe 2019; pp. 1993–2002. [CrossRef]
- Woźniak, S.; Szymański, P. Hex2vec: Context-Aware Embedding H3 Hexagons with OpenStreetMap Tags. In Proceedings of the 4th ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery, Beijing China, 2–5 November 2021; pp. 61–71. [CrossRef]
- Tian, C.; Zhang, Y.; Weng, Z.; Gu, X.; Chan, W.K. Learning Fine-grained Location Embedding from Human Mobility with Graph Neural Networks. In Proceedings of the 2022 International Joint Conference on Neural Networks (IJCNN), Padua, Italy, 18–23 July 2022; pp. 1–8. [CrossRef]
- Shimizu, T.; Yabe, T.; Tsubouchi, K. Enabling Finer Grained Place Embeddings Using Spatial Hierarchy from Human Mobility Trajectories. In Proceedings of the 28th International Conference on Advances in Geographic Information Systems, Seattle WA, USA, 3–6 November 2020; pp. 187–190. [CrossRef]
- Yin, Y.; Liu, Z.; Zhang, Y.; Wang, S.; Shah, R.R.; Zimmermann, R. GPS2Vec: Towards Generating Worldwide GPS Embeddings. In Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Chicago, IL, USA, 5–8 November 2019; pp. 416–419. [CrossRef]
- Souza, A.P.R.; Renso, C.; Perego, R.; Bogorny, V. MAT-Index: An Index for Fast Multiple Aspect Trajectory Similarity Measuring. *Trans. GIS* 2022, 26, 691–716. [CrossRef]
- Gao, C.; Zhang, Z.; Huang, C.; Yin, H.; Yang, Q.; Shao, J. Semantic Trajectory Representation and Retrieval via Hierarchical Embedding. *Inf. Sci.* 2020, 538, 176–192. [CrossRef]
- Chu, C.; Zhang, H.; Lu, F. Inferring Consumption Behavior of Customers in Shopping Malls from Indoor Trajectories. J. Geo-Inf. Sci. 2022, 24, 1034–1046.
- 37. Niemeyer, G. Geohash. Available online: http://geohash.org/ (accessed on 16 October 2022).

- 38. GitHub Inc. S2 Geometry.. Available online: https://s2geometry.io (accessed on 21 October 2022).
- Uber Technologies Inc. H3: Uber's Hexagonal Hierarchical Spatial Index. Available online: https://eng.uber.com/h3 (accessed on 17 December 2022).
- 40. Pelekis, N.; Theodoulidis, B.; Kopanakis, I.; Theodoridis, Y. Literature Review of Spatio-Temporal Database Models. *Knowl. Eng. Rev.* 2004, *19*, 235–274. [CrossRef]
- 41. Tao, Y.; Both, A.; Silveira, R.I.; Buchin, K.; Sijben, S.; Purves, R.S.; Laube, P.; Peng, D.; Toohey, K.; Duckham, M. A Comparative Analysis of Trajectory Similarity Measures. *Glsci. Remote Sens.* **2021**, *58*, 643–669. [CrossRef]
- 42. Mai, G.; Janowicz, K.; Hu, Y.; Gao, S.; Yan, B.; Zhu, R.; Cai, L.; Lao, N. A Review of Location Encoding for GeoAI: Methods and Applications. *Int. J. Geogr. Inf. Sci.* 2022, *36*, 639–673. [CrossRef]
- 43. Chen, J.; Zhou, C.; Cheng, W. Area Error Analysis of Vector to Raster Conversion of Areal Feature in GIS. *Acta Geod. Cartogr. Sin.* **2007**, *36*, 344–350.
- Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.S.; Dean, J. Distributed Representations of Words and Phrases and Their Compositionality. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 5–10 December 2013; Burges, C.J.; Bottou, L.; Welling, M.; Ghahramani, Z.; Weinberger, K.Q., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2013; Volume 26.
- Dong, Y.; Chawla, N.V.; Swami, A. Metapath2vec: Scalable Representation Learning for Heterogeneous Networks. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017; pp. 135–144. [CrossRef]
- 46. Sun, Y.; Han, J.; Yan, X.; Yu, P.S.; Wu, T. PathSim: Meta Path-Based Top-K Similarity Search in Heterogeneous Information Networks. *Proc. VLDB Endow.* 2011, *4*, 992–1003. [CrossRef]
- 47. Grover, A.; Leskovec, J. Node2vec: Scalable Feature Learning for Networks. arXiv 2016, arXiv:1607.00653. [CrossRef]
- 48. Kelso, N.V.; Patterson, T. Natural Earth. 2022. Available online: https://www.naturalearthdata.com/ (accessed on 10 November 2022).
- 49. Megginson, D. Ourairports. 2022. Available online: https://ourairports.com (accessed on 23 December 2022).
- Schafer, M.; Strohmeier, M.; Lenders, V.; Martinovic, I.; Wilhelm, M. Bringing up OpenSky: A Large-Scale ADS-B Sensor Network for Research. In Proceedings of the IPSN-14 Proceedings of the 13th International Symposium on Information Processing in Sensor Networks, Berlin, Germany, 15–17 April 2014; pp. 83–94. [CrossRef]
- Han, P.; Wang, J.; Yao, D.; Shang, S.; Zhang, X. A Graph-based Approach for Trajectory Similarity Computation in Spatial Networks. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, Singapore, 14–18 August 2021; pp. 556–564. [CrossRef]
- Keogh, E.J.; Pazzani, M.J. Scaling up Dynamic Time Warping for Datamining Applications. In Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Boston MA, USA, 20–23 August 2000; pp. 285–289. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.