Hierarchical Attention Network with Correction for Cross-Domain User Association

Wenlong Liu[1,2], Ze Wang[1,2*], Chenlong Wu[1,2], Yude Bai[1,2] Ji Zhang[3]
[1]School of Software, Tiangong University, Tianjin, China
[2]Tianjin Key Laboratory of Autonomous Intelligence Technology and Systems, Tiangong University, Tianjin, China
[3] University of Southern Queensland, Australia
{2431111311, wangze, wuchenlong, baiyude}@tiangong.edu.cn,Ji.Zhang@unisq.edu.au

Abstract

Despite the rich spatiotemporal patterns contained in trajectory data from multiple Location-Based Social Network (LBSN) platforms, heterogeneous formats, semantic inconsistencies, and unequal user scales across platforms create substantial barriers to reliable identity mapping. Furthermore, GPS drift and sparse sampling result in degraded data quality and distribution imbalance, which render existing trajectory representation methods inadequate for capturing high-order dependencies and dynamic spatiotemporal evolution patterns in heterogeneous multi-relational graphs. To this end, we propose HANCUA (**H**ierarchical **A**ttention **N**etwork with **C**orrection for **U**ser **A**ssociation), a novel framework that employs a dual-stage correction mechanism to enhance cross-domain trajectory analysis. The approach constructs hierarchical multi-relational graphs comprising location, trajectory, and correction layers to capture fine-grained mobility patterns, behavioral associations, and inter-platform distribution differences. We design relation-aware multi-head graph attention networks to model complex interactions among heterogeneous node types, which enables comprehensive spatial relationship modeling. A spatiotemporal semantic collaborative learning module integrates temporal information with mobility patterns through interaction-aware attention mechanisms, while an ensemble correction decision module incorporates ensemble learning principles to systematically correct user association biases and address distribution imbalance problems. Extensive experiments on two real-world LBSN cross-domain datasets reveals that HANCUA significantly outperforms state-of-the-art methods in user identity linking accuracy.

Introduction

Location-Based Social Networks (LBSN) gathers extensive user behavioral data, including social interactions and spatiotemporal check-ins reflecting mobility patterns. This rich data provides insights for aligning user identities across heterogeneous service platforms. From an enterprise perspective, the integration of cross-platform user data creates comprehensive behavioral profiles that enhance service quality and operational efficiency (Ma et al. 2022). This consolidated understanding enables sophisticated applications such as cross-domain personalized recommendation systems (Chen et al. 2020), which leverage unified user models to deliver more relevant and targeted content. For individual users, comprehending their digital footprints illuminates potential privacy vulnerabilities and exposure risks inherent in their multi-platform presence (Shu et al. 2017) (Zhang and Philip 2019), which empowers users to make informed decisions about their digital privacy management strategies. Cross-domain user association also contributes to effective data fusion, used by sophisticated business intelligence frameworks that drive strategic decision-making processes (Feng et al. 2019).

Traditional methods rely on user attribute data (Goga et al. 2015) and social graph structures (Korula and Lattanzi 2013) from platforms to establish identity correspondences, but face two critical limitations: (1) Data heterogeneity across service ecosystems creates structural disparities that affect cross-platform analysis. For example, e-commerce platforms prioritize transactional data while lacking social connectivity information found in social networks. (2) User-provided information is unreliable as users may supply incomplete or falsified profiles (pseudonymous names, blank fields, etc.), which introduces substantial noise.

Researchers therefore turn to trajectory representation learning to extract features from location check-ins, which creates unique behavioral fingerprints for cross-domain association. TULVAE (Zhou et al. 2018) combines Variational Autoencoders (VAE) with Recurrent Neural Networks (RNN) to capture hierarchical semantic features from check-in trajectories, enabling the model to learn both local movement patterns and higher-level abstractions. DPLink (Feng et al. 2019) adopts multi-layer joint embedding strategies to create paired trajectory representations for matching, facilitating precise similarity comparisons across domains. SML-TUL (Zhou et al. 2021b) introduces contrastive learning to encode user mobility patterns while considering spatiotemporal factors, helping the model distinguish similar but distinct behaviors. AttnTUL (Chen et al. 2024) leverages graph neural networks to model complex mobility trajectories, capturing both local transition patterns and global spatial dependencies across temporal scales.

Despite trajectory-based approaches addressing attribute unreliability issues, existing methods have three critical limitations: (1) Poor data quality reduces representation accu-

*Corresponding author.

racy due to sampling sparsity, GPS drift, and incomplete annotations (Wang et al. 2016) (Xu et al. 2018) (He et al. 2025) . While some studies use recurrent networks (Riederer et al. 2016) (Rossi and Musolesi 2014), they fail to extract implicit patterns. Manual methods (Wang et al. 2018) require extensive tuning that limits deployment. (2) Inadequate local-global feature integration prevents complex pattern generation, as methods process local transitions or global preferences separately (Gao et al. 2020) (Zhou et al. 2021c) (Zhou et al. 2021b). (3) Cross-platform distribution imbalance limits generalization, where sample differences (Ma et al. 2010) (Riederer et al. 2016) cause training dominated by high-volume platforms. For example, ISP data contains over four times the social network density (Feng et al. 2019) that results in accuracy degradation on low-resource platforms.

To address the above challenges, we propose HANCUA (**H**ierarchical **A**ttention **N**etwork with **C**orrection for **U**ser **A**ssociation). Our model specifically solves limitation (1) by constructing hierarchical multi-relational graphs that transform low-quality trajectory data into robust representations, where the location layer captures fine-grained mobility patterns despite sampling sparsity, the trajectory layer models behavioral associations to compensate for semantic loss, and the correction layer eliminates distribution biases caused by GPS drift and noise. For limitation (2), we design a spatiotemporal semantic collaborative learning module that establishes dynamic interactions between local and global features, using interaction-aware attention mechanisms to fuse local spatiotemporal patterns with global trajectory representations, which enables comprehensive multi-dimensional trajectory modeling that adapts to cross-domain behavioral heterogeneity. To resolve limitation (3), our ensemble correction decision module implements ensemble learning principles through intelligent secondary correction, which mitigates performance degradation on low-resource platforms by leveraging collective decision-making across trajectory groups to address cross-platform distribution imbalance.

The contributions of HANCUA are summarized as follows:

- We propose a Multi-level Heterogeneous Feature Fusion Module that builds hierarchical graph attention networks, which effectively transforms unreliable trajectory data into robust behavioral representations.
- We present a Spatiotemporal Semantic Collaborative Learning Module that establishes dynamic interactions between local and global trajectory features for the adaptation of cross-domain behavioral heterogeneity and evolutionary characteristics.
- We develop an Ensemble Correction Decision Module that implements a global consistency adjustment by grouping cross-domain trajectories into virtual users and applying majority-voting consensus in the embedding space to correct distribution shifts.
- Extensive experiments on two real LBSN cross-domain datasets reveal that HANCUA achieves superior performance compared to the baselines, which attains 7.15%∼15.26% improvement in ACC@1 and 8.86%∼16.45% increase in Macro-F1.

Related Work

Deep Learning-based Association Methods. Recent years show the potential of deep learning-based methods in cross-domain user identity linking. Early work, such as TULER (Gao et al. 2017), uses RNN-based models to learn sequential patterns from trajectories. However, RNNs suffer from data sparsity and lack understanding of hierarchical semantics in mobility. TULVAE (Zhou et al. 2018) combines Variational Autoencoders (VAE) with hierarchical RNNs to capture long-term dependencies and semantic features but struggles with cross-domain data distribution differences and trajectory sampling sparsity. DPLink (Feng et al. 2019) introduces multi-layer joint embedding methods to improve performance through pairwise trajectory matching analysis, yet static embeddings limit modeling of dynamic relationships (Liu et al. 2025). EgoMUIL (Huang et al. 2023) optimizes topological similarity and stay location co-occurrence but overlooks the impact of device noise and spatiotemporal misalignment. SML-TUL (Zhou et al. 2021b) applies contrastive learning to enforce spatiotemporal consistency but struggles with fine-grained local semantics, like POI type density, and global behavioral associations. These methods excel within single domains but degrade in performance with data imbalances and noise in cross-domain scenarios.

Graph Neural Network-based Association Methods. Graph Neural Networks (GNNs) have become essential for cross-domain user association due to their ability to model complex relations. GNNTUL (Zhou et al. 2021a) applies GNNs to trajectory association, aggregating neighbor node features through message passing. However, it is limited to homogeneous trajectory graphs and cannot handle semantic heterogeneity across domains. S2TUL (Deng et al. 2023) builds multi-type graph structures to capture multi-dimensional relationships, but lacks dynamic attention, leading to interference from noise edges. AttnTUL (Chen et al. 2024) combines hierarchical spatiotemporal attention to link local trajectory segments and global patterns, but fails to address cross-domain data distribution biases. TCGCN (Wang et al. 2024) proposes three-channel graph networks to fuse multi-relational features, but its static channel weight allocation struggles with dynamic semantic changes.

Despite advances in deep learning and graph neural network-based methods, three limitations remain. First of all, multi-dimensional feature representation and high-order dependency modeling hinder the effective capture of complex spatial relationships. Inadequate integration of local and global features leads to the isolated processing of spatiotemporal patterns. Furthermore, cross-platform data distribution imbalance and low-quality trajectory sensitivity degrade performance on low-resource platforms.

Problem Formulation

We formulate the cross-domain user association task as a trajectory-based identity mapping problem that operates on spatiotemporal mobility data across heterogeneous platforms.

Let $\mathcal{U} = \{u_1, u_2, \ldots, u_i\}$ represent the set of real users who maintain identities across multiple cross-domain platforms. Given a geographical location set $\mathcal{L}$ where each location $l \in \mathcal{L}$ is uniquely identified by coordinates (longitude, latitude), user mobility is captured through spatiotemporal check-in points $p = (l, t)$, where $t \in [0, 48]$ represents discretized timestamps in half-hour units. Each user $u_i \in \mathcal{U}$ generates a mobility trajectory $T(u_i) = \{p_1, p_2, \ldots, p_m\}$ consisting of ordered spatiotemporal check-in sequences, where m is the trajectory length. As the cross-domain setting, consider K heterogeneous data sources, where each platform S_k contains a virtual user set U_k and their corresponding trajectory collection $\{T_u \mid u \in U_k\}$. The cross-domain user universe is defined as $U = \bigcup_{k=1}^{K} U_k$, including all virtual user identities across platforms with their associated cross-platform trajectory sets $\{T_{u_i}^{S_1}, T_{u_j}^{S_2}, \ldots\}$.

The cross-domain Trajectory user association problem aims to construct a discriminative mapping model $\mathcal{M}$ that learns robust trajectory representations to determine whether any cross-domain trajectory pair $\left(T_x^{S_a}, T_y^{S_b}\right)$ from different platforms S_a and S_b belongs to the same real user.

Methodology

In this section, we detail the design of HANCUA, as shown in Figure 1. The model consists of three core components: (1) Multi-level Heterogeneous Feature Fusion Module, (2) Spatiotemporal Semantic Collaborative Learning Module, and (3) ensemble Correction Decision Module. These three modules form an integrated pipeline that systematically addresses cross-domain trajectory association challenges. The hierarchical feature fusion module first transforms raw trajectory data into robust multi-dimensional representations, which provides the foundation for subsequent processing. The spatiotemporal learning module then establishes dynamic interactions between local and global features, which creates comprehensive behavioral profiles that capture both detailed movement patterns and overarching user characteristics. Finally, the ensemble correction module refines these profiles through collective intelligence across trajectory groups, which ensures reliable identity associations despite data quality variations and cross-platform disparities.

Multi-level Heterogeneous Feature Fusion Module

Hierarchical Multi-relational Graph Construction. Our approach first constructs a three-tier hierarchical graph structure that transforms raw spatiotemporal sequences into robust multi-dimensional representations to effectively capture user mobility transition patterns and reveal shared behavioral characteristics across users. The graph can be presented as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ that concatenates discrete locations into semantic units. The graph architecture employs three distinct node types $\mathcal{V} = \mathcal{V}_g \cup \mathcal{V}_t \cup \mathcal{V}_c$, corresponding to the location layer (bottom), trajectory layer (middle), and correction layer (top). These layers are connected through four edge types $\mathcal{E} = \mathcal{E}_{gg} \cup \mathcal{E}_{tt} \cup \mathcal{E}_{tg} \cup \mathcal{E}_{ct}$, which enable multi-level feature propagation and relationship modeling across different granularities of mobility patterns. To enable efficient graph construction, we employ grid-based spatial discretization techniques. Given a predefined grid edge length len, we partition the entire geographical space containing trajectories into n discrete grids, enabling each trajectory $Tr = \langle (t_1, l_1), \ldots, (t_T, l_T) \rangle$ to be mapped to $Tr = \langle (t_1, g(l_1)), \ldots, (t_T, g(l_T)) \rangle$, where $g(l_i)$ is the coordinate-to-grid mapping function that maps continuous latitude-longitude coordinates to finite grids. The discretization eliminates GPS noise, reduces data sparsity through spatial aggregation, and decreases computational complexity by limiting the number of spatial nodes.

- **Location Layer** ($\mathcal{V}_g$): In this layer, each spatial grid functions as a node, initialized with one-hot encoding to ensure unique grid representation within the feature matrix. When consecutive check-ins in any two grids g_i and g_j, we establish an edge $e_{ij} \in \mathcal{E}_{gg}$ between these grids to encode spatial transition relationships. The edge weight is calculated as the total number of trajectories moving between the two spatial regions, which captures popular mobility pathways and spatial connectivity patterns.
- **Trajectory Layer** ($\mathcal{V}_t$): This layer models behavioral similarities and cross-trajectory dependencies to identify common movement characteristics. Each complete trajectory serves as a node, initialized using multi-hot encoding where spatial grids visited by the trajectory are set to 1 and unvisited grids remain 0. For trajectories t_i and t_j sharing common spatial regions, we create an edge $e_{ij} \in \mathcal{E}_{tt}$ with weight equal to the number of overlapping grids, which quantifies behavioral similarity between different users' movement patterns. Cross-layer connections between trajectory nodes t_i and grid nodes g_j are established through edges $eij \in \mathcal{E}_{tg}$ with unit weights, which associates individual behaviors with specific spatial contexts and enables the concatenation of discrete locations into meaningful semantic units (such as "commuting from home to office").
- **Correction Layer** ($\mathcal{V}_c$): This layer addresses data quality issues and cross-domain distribution disparities through trajectory group constraints. Correction nodes represent collections of trajectories known to belong to the same user entity, initialized with multi-hot encoding where grids visited by any trajectory in the group are marked as 1. During the pre-association phase, while specific user identities remain unknown, trajectory groupings within individual platforms are available. Correction nodes c_i connect to their constituent trajectory nodes t_j via edges $e_{ij} \in \mathcal{E}_{ct}$, with edge weights initialized based on the maximum pairwise similarity within the trajectory group. This constraint mechanism enables low-quality trajectories to benefit from high-quality trajectories within the same user group, which improves overall representation robustness through collective trajectory guidance.

Relation-aware Multi-head Graph Attention Networks. Graph Attention Networks (GATs) (Cen et al. 2024) are essential for our cross-domain trajectory association task because they can simultaneously handle the heterogeneous multi-relational graph (Liu et al. 2024)structure containing three node types and four edge types. However, conventional GATs are limited for our task because they assume homogeneous graphs with uniform node and edge types, which cannot handle our hierarchical structure containing distinct

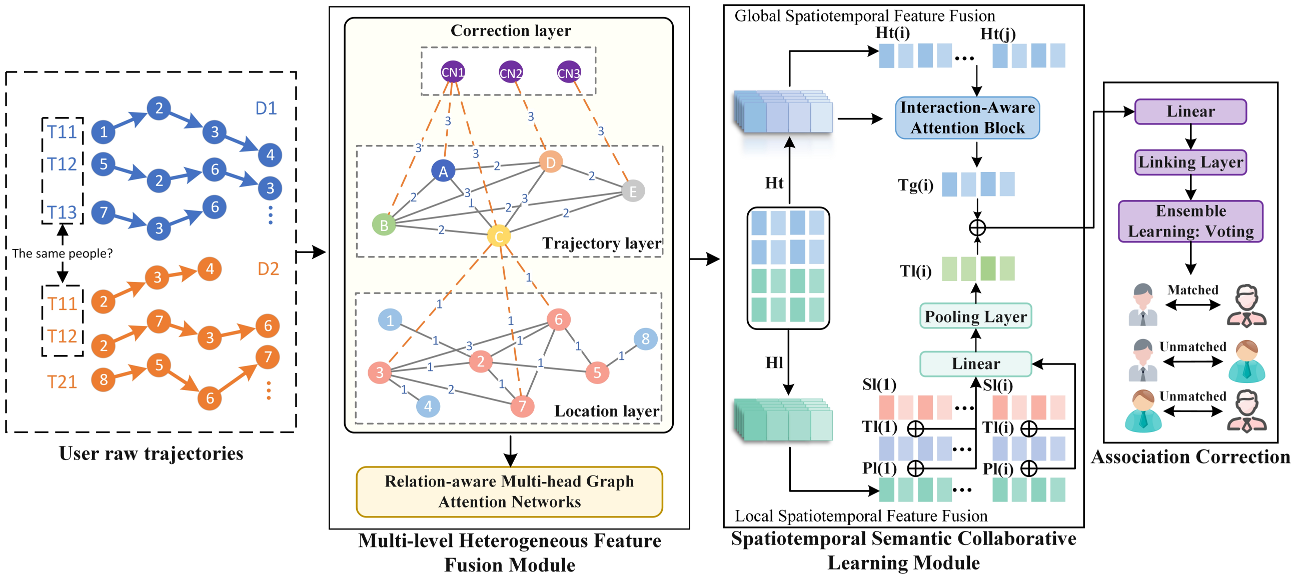

Figure 1: The overview of our HANCUA model.

node types and different edge relationships that require type-specific processing. Thus, we introduce the relation-aware attention mechanism that enables adaptive learning of which spatial transitions, behavioral similarities, and cross-domain relationships are most important for accurate user association, automatically adapting to different platform characteristics and data quality variations. It serves as the critical feature transformation bridge that converts noisy, sparse trajectory data into robust multi-dimensional representations for subsequent modules.

First, to simultaneously capture relationships among three node types and four edge types, we perform type-specific linear projections for each node type, generating learnable initial features:

$$h_i^{(0)} = \begin{cases} W_g x_i + b_g, & \forall i \in V_g \\ W_t x_i + b_t, & \forall i \in V_t \\ W_c x_i + b_c, & \forall i \in V_c \end{cases} \tag{1}$$

where $W_* \in \mathbf{R}^{d \times \text{hid}}$, $x_i \in \mathbf{R}^d$ represents original features (one-hot/multi-hot encoding), and d is the number of grids, i.e., dimension of each node.

Then, we perform relation-aware attention coefficient calculation. For the k-th head, we calculate attention coefficient $\alpha_{ij}^{(k)}$ between nodes i and j:

$$e_{ij}^{(k)} = \text{LeakyReLU}\left(a^{(k)} \phi(r_{ij}) \left[h_i^{(n-1)} \parallel h_j^{(n-1)}\right]\right) \tag{2}$$

$$\alpha_{ij}^{(k)} = \frac{\exp(e_{ij}^{(k)})}{\sum_{p \in N_i} \exp(e_{ip}^{(k)})} \tag{3}$$

where N_i is the neighbor set of node i, $\alpha^{(k)}$ is the training parameter for the k-th head, $\phi(r_{ij})$ is the initial weight between nodes, and $\parallel$ concatenates two node features from layer $n-1$. After obtaining $\alpha_{ij}^{(k)}$, we apply minimum pooling to coefficients from k heads for each edge to accelerate computation efficiency, obtaining final relation-aware attention coefficient α_{ij} for each edge.

Based on calculated attention coefficients, we aggregate neighbor information through weighted summation of neighbor node features to obtain the n-th layer feature vector $h_i^{(n)}$ for node i: $h_i^{(n)} = \sigma\left(\sum_{j \in N_i} \alpha_{ij} W h_j^{(n-1)}\right)$.

Finally, we obtain embedding matrices $H_l^{(n)}$ and $H_t^{(n)}$ for location layer grid nodes and trajectory layer trajectory nodes learned through n-layer relation-aware multi-head graph attention networks, denoted as H^l and H^t.

Spatiotemporal Semantic Collaborative Learning Module

Local Spatiotemporal Feature Fusion. Local feature processing captures the detailed spatiotemporal dynamics within individual trajectory sequences, which is crucial for distinguishing between users with similar general movement areas but different specific behavioral patterns. Beyond basic spatial grid information, we incorporate rich contextual features that characterize the temporal and kinematic aspects of user mobility (Qin et al. 2023), enabling more discriminative trajectory representations that can handle cross-domain variations in check-in behavior and platform usage patterns.

Specifically, time is discretized into 30-minute windows to capture periodic behaviors like commuting and dining patterns, avoiding noise from continuous timestamps. Speed features are calculated as the ratio of distance to time between consecutive check-ins, which distinguishes different movement modes such as walking, driving, or staying stationary.

Two embedding layers encode temporal and speed information, then combine with spatial features:

$$x_i = \tanh\left(\mathrm{FC}\left([h_i^l;\, W_t t_i + b_t;\, W_s s_i + b_s]\right)\right) \quad (4)$$

where h_i^l represents GAT-enhanced grid features, $W_t t_i + b_t$ and $W_s s_i + b_s$ encode temporal and speed information through learnable transformations, and FC creates unified location embeddings.

Max pooling extracts the most significant patterns from each trajectory sequence:

$$\mathrm{Tr}_i^l = \mathrm{Pooling}(x_i^1, x_i^2, \ldots, x_i^n) \quad (5)$$

This process generates local representations Tr_i^l that capture distinctive spatiotemporal and kinematic (speed) characteristics of individual user trajectories.

Global Spatiotemporal Feature Fusion. Global features are essential because local patterns alone cannot capture cross-trajectory behavioral similarities and long-range dependencies that are critical for cross-domain user association. Users with similar movement preferences may visit different specific locations but exhibit comparable global mobility characteristics such as travel frequency, geographic coverage, and activity timing patterns.

To this end, we design a trajectory interaction-aware attention module that selects the most relevant global information for each trajectory. The module comprises two complementary components: (1) importance - quantifies how much each trajectory contributes valuable information to the current trajectory's representation, and (2) similarity - measures direct behavioral alignment between trajectory pairs. Unlike standard sparse attention mechanisms (Martins and Astudillo 2016)(Peters, Niculae, and Martins 2019) that focus on single-modal similarity, our approach combines learned importance weights with behavioral similarity measures for heterogeneous cross-domain trajectory data.

For importance modeling, we compute trajectory interaction weights that reflect global relevance. Given trajectory embeddings $h_i^t \in \mathbf{R}^d \mid i \in \mathcal{V}_t$, we concatenate trajectory pairs and learn importance coefficients:

$$\alpha_{ij} = \mathrm{Softmax}\left(w^\top \cdot \mathrm{MLP}([h_i^t \parallel h_j^t]) + b\right) \quad (6)$$

For similarity modeling, we compute cosine similarity between trajectories to capture direct behavioral alignment, then apply sparse normalization to focus on the most relevant relationships while filtering noise. The final global representation combines both perspectives:

$$\mathrm{Tr}_i^t = (\alpha_i \cdot \beta_i) H^t \quad (7)$$

where α_i represents learned importance weights, β_i denotes sparsified similarity scores, and their element-wise multiplication creates comprehensive global trajectory representations that capture both complementary information and direct behavioral similarities.

Ensemble Correction Decision Module

Association Layer. The association layer converts trajectory representations into user identity probability distributions. It concatenates local and global trajectory features Tr_i^l and Tr_i^t, then projects them to user space $|U|$ through fully connected transformation. For similarity computation, we use Sparsemax instead of Softmax to eliminate noise from weakly-related trajectories. Given cosine similarities between trajectory i and all other trajectories:

$$\cos_i = \left\{\cos_{i,1}, \cos_{i,2}, \ldots, \cos_{i,\|Tr\|}\right\} \quad (8)$$

Sparsemax normalization produces sparse distributions that retain only significantly related trajectories:

$$\beta_i = \mathrm{Sparsemax}(\cos_i) \quad (9)$$

$$\mathrm{Sparsemax}(x) = \underset{p\in\Delta^d}{\mathrm{argmin}} \|p - x\|^2 \quad (10)$$

where $\Delta^d = \left\{p \in \mathbf{R}^d : p \geq 0, \|p\|_1 = 1\right\}$ defines the probability simplex. Unlike Softmax, Sparsemax assigns zero weights to low-similarity trajectories, reducing noise interference. The final user prediction combines local and global representations:

$$\hat{y}_i = \sigma\left(W\left[\mathrm{Tr}_i^l, \mathrm{Tr}_i^t\right] + b\right) \quad (11)$$

where $W \in \mathbb{R}^{|U|\times 2d}$ and $b \in \mathbb{R}^{|U|}$ are learnable parameters, $\sigma(\cdot)$ is the softmax function, and $\hat{y}_i$ represents the probability distribution over real users for trajectory Tr_i.

Ensemble Correction Layer. Single trajectory predictions are prone to two typical errors: local deviation due to sampling sparsity or noise, and global inconsistency where cross-domain trajectories of the same user exhibit discontinuous feature distributions, we thus propose virtual user–guided self-correction, which treats trajectories within each virtual user group as an ensemble of weak learners, and then applies group decisions to correct individual anomalies. Unlike prior approaches that perform post-hoc smoothing or rely solely on local refinement, our method uniquely integrates majority-voting consensus directly into the representation space of graph neural networks to ensure that group-level agreement drives embedding optimization.

Given trajectory Tr_i with initial prediction u_i from maximum probability of y_i, we obtain the corrected result through majority voting within the trajectory group:

$$\hat{u}_i^* = \mathrm{mode}(\overline{U}) = \arg\max_{u\in U^*} \sum_{i=1}^{n} I(u_i = u) \quad (12)$$

where $\overline{U} = u_1, \ldots, u_n$ contains predictions for all trajectories in the group, $I(\cdot)$ is the indicator function, and n is the group size. The model is trained using cross-entropy loss with L2 regularization:

$$L(\Theta) = -\frac{1}{n}\sum_{i=1}^{n} c_i\left(\sigma() + \frac{\lambda}{2}\|\Theta\|^2\right. \quad (13)$$

where c_i is the true label, n is the number of training trajectories, Θ represents all trainable parameters, and λ controls regularization strength.

Dataset	User numbers	Location numbers	Check-ins
Instagram	2505	59639	428292
Twitter	1721	46312	447972
Foursquare	2970	16689	44915
Twitter	1228	10472	53337

Table 1: Basic Information of Experimental Datasets.

Experimental Results and Analysis

Experimental Datasets. Our experimental evaluation employs two benchmark cross-domain LBSN datasets that provide complementary validation scenarios: a sparse dataset (i.e., Foursquare-Twitter (Zhang, Kong, and Yu 2014)) and a dense dataset (i.e., Instagram-Twitter (Riederer et al. 2016)). These datasets are particularly well-suited for trajectory association research because they feature authentic cross-platform user correspondences, exhibit realistic data sparsity and quality variations, and enable direct performance comparison with established baseline methods. Detailed dataset characteristics are presented in Table 1.

Baselines. We compare HANCUA against 8 carefully selected baseline algorithms that represent diverse approaches and state-of-the-art performance in cross-domain trajectory association, ranging from early RNN-based methods to recent graph neural network frameworks:

- **TULER** (Gao et al. 2017): Utilizes word2vec to embed location points and RNN to identify and link trajectories to corresponding users. Three variants: RNN with Gated Recurrent Units (TULER-G), LSTM (TULER-L), and Bidirectional LSTM (Bi-TULER).
- **T3S** (Yang et al. 2021): State-of-the-art trajectory representation learning method for trajectory similarity computation.
- **DPLink** (Feng et al. 2019): An advanced method for user account linking from heterogeneous mobile data.
- **TULAM** (Li et al. 2024): Learns sequential relationships in individual trajectories through RNN and captures contextual semantics from trajectories via multi-head attention mechanism.
- **AttnTUL** (Chen et al. 2024): Proposes hierarchical spatiotemporal attention neural networks to jointly encode local trajectory transition patterns and global spatial dependencies for trajectory user association.
- **StarNet** (Wu et al. 2025): Adaptively adjusts multi-dimensional data features through dynamic combination of multi-head attention mechanisms.

Experimental settings. For fair comparison across all learning methods, we set epochs to 80, batch size to 16, dropout to 0.5, adjust learning rate from 0.0001 to 0.01, use early stopping with patience=10 to avoid overfitting. Each experiment is repeated 10 times with average results reported to establish statistical reliability. All evaluations are conducted on a machine with Intel Xeon Gold 6126@2.60GHz 12-core CPU, 192GB memory, and NVIDIA Tesla V100-SXM2 (16GB) GPU.

Effectiveness Performance Comparison

We evaluated HANCUA's overall effectiveness on two real-world cross-domain LBSN datasets by benchmarking it against state-of-the-art methods. Table 2 reports the performance of each model, with the best results in bold and the second best underlined.

From Table 2, HANCUA achieves the highest accuracy across all metrics on both datasets. This gain arises from our hierarchical multi-relational graph construction and adapted multi-head graph-attention networks, which capture user mobility transitions, group discrete locations into semantic units, and suppress noise, thereby strengthening cross-domain generalization. The dual correction mechanisms further address trajectory quality issues and correct distribution imbalances among user trajectories.

On the sparse Foursquare–Twitter dataset, HANCUA surpasses StarNet by 15.2% in ACC@1, 20.82% in ACC@5, and 16.45% in Macro-F1, demonstrating its effectiveness in handling sparse, noisy check-in data. Even on the dense Instagram–Twitter dataset, it outperforms StarNet, delivering a 7.15% gain in ACC@1 and notable improvements in ACC@5 and Macro-F1. These results confirm HANCUA's ability to fuse multi-dimensional spatiotemporal information and deliver robust performance despite varying data sparsity.

Ablation Study

To validate each component's contribution to the performance of our model, we perform an ablation study on both datasets under six experimental settings:

- **Full Model**: The complete HANCUA, which serves as our performance benchmark.
- **w/o CM**: Removing the self-correction layer from the Ensemble Correction Decision Module.
- **w/o CL**: Omitting the correction layer from the hierarchical multi-relation graph.
- **w/o HG**: Splitting the hierarchical graph into separate location and trajectory layers.
- **w/o GAT**: Replacing the multi-head graph-attention network with a standard graph-convolutional network.
- **w/o IAA**: Excluding the trajectory interaction-aware attention module.

As seen in Table 3, compared to the **w/o CM** setting, the Full Model improves ACC@1 by 31.53% on the sparse Foursquare–Twitter dataset, which confirms that the self-correction layer effectively mitigates prediction bias and class imbalance through hard-voting consensus. Removing the correction layer from the hierarchical graph (**w/o CL**) reduces ACC@1 by 24.17%, which demonstrates that trajectory-group constraints are essential for bundling low-quality trajectories and preventing bias during user association. Under the **w/o HG** setting, ACC@1 drops by 18.42%, indicating that the hierarchical multi-relational graph-attention network establishes critical semantic links between discrete locations. Replacing GAT with GCN (**w/o GAT**) leads to a 12.38% decrease in

Methods	Foursquare-Twitter					Ins-Twitter				
	ACC@1	ACC@5	Ma-P	Ma-R	Ma-F1	ACC@1	ACC@5	Ma-P	Ma-R	Ma-F1
TULER-GRU	0.3979	0.4576	0.3552	0.3375	0.3227	0.5778	0.6416	0.5739	0.5240	0.5307
TULER-LSTM	0.3681	0.4656	0.3578	0.3125	0.3115	0.5800	0.6393	0.6201	0.5284	0.5452
TULER-BiLSTM	0.3970	0.4747	0.3700	0.3436	0.3373	0.5858	0.6471	0.6215	0.5463	0.5589
T3S	0.4191	0.4872	0.3858	0.3643	0.3481	0.5985	0.6593	0.6284	0.5520	0.5684
DPLink	0.4482	0.5342	0.3960	0.3715	0.3521	0.6029	0.6719	0.6335	0.5565	0.5768
TULAM	0.4649	0.5528	0.4098	0.3809	0.3651	0.6257	0.6805	0.6419	0.5697	0.5821
AttnTUL	0.5268	0.6182	<u>0.5144</u>	0.4084	0.4313	0.6800	0.7580	0.6937	0.6221	0.6573
StarNet	<u>0.6015</u>	<u>0.6844</u>	0.4973	<u>0.4448</u>	<u>0.4527</u>	<u>0.7123</u>	<u>0.7976</u>	<u>0.7158</u>	<u>0.6882</u>	<u>0.6918</u>
HANCUA	**0.6933**	**0.8269**	**0.5506**	**0.5200**	**0.5272**	**0.7633**	**0.9259**	**0.7488**	**0.7600**	**0.7531**
% of Improvement	**15.26%**	**20.82%**	**7.03%**	**16.9%**	**16.45%**	**7.15%**	**16.08%**	**4.61%**	**10.43%**	**8.86%**

Table 2: Effectiveness performance comparison results.

Methods	**Foursquare-Twitter**			**Instagram-Twitter**		
	ACC@1	ACC@5	Ma-F1	ACC@1	ACC@5	Ma-F1
w/o CM	0.5272	0.6954	0.4373	0.6867	0.8572	0.5832
w/o CL	0.6773	0.7921	0.4507	0.7374	0.8948	0.6956
w/o HG	0.6397	0.7996	0.5127	0.7155	0.8736	0.7198
w/o GAT	0.6171	0.8053	0.3914	0.7251	0.8404	0.6788
w/o IAA	0.6679	0.8128	0.4793	0.6971	0.8888	0.7030
Full Model	**0.6933**	**0.8269**	**0.5272**	**0.7633**	**0.9259**	**0.7531**

Table 3: The results of the ablation study.

ACC@1 on Foursquare–Twitter, which shows that multi-head attention captures heterogeneous node relationships more effectively, particularly on sparse data. Finally, excluding the interaction-aware attention module (**w/o IAA**) results in a 9.75% decline in Macro-F1, which suggests that modeling inter-trajectory similarity enhances global representation quality. Together, these findings validate our design choices and demonstrate how each module contributes to robust cross-domain user association.

Robustness Analysis

We evaluate HANCUA's resilience to increasing noise on two cross-domain benchmarks, a sparse Foursquare–Twitter dataset and a dense Instagram–Twitter dataset, by injecting 5%, 10%, 15% and 20% non–cross-domain users. Figure 2 shows that HANCUA delivers stability and consistent performance gains across all noise levels. On the sparse Foursquare–Twitter, HANCUA introduces 5% noise boosts ACC@1 from 0.6933 to 0.7015 and raises Macro-F1 to 0.5347, which indicates that moderate noise can actually boost the model's discrimination of true cross-domain associations. Even when noise increases to 20%, ACC@1 remains above 0.6700 and Macro-F1 stays above 0.5100, which demonstrats outstanding noise resistance in sparse settings. On the dense Instagram–Twitter, it sustains ACC@1 above 0.72 and Macro-F1 within 3% of its clean-data values for noise levels up to 15%. Even at 20% noise, the model records a Macro-F1 of 0.65, underlining its robust generalization.

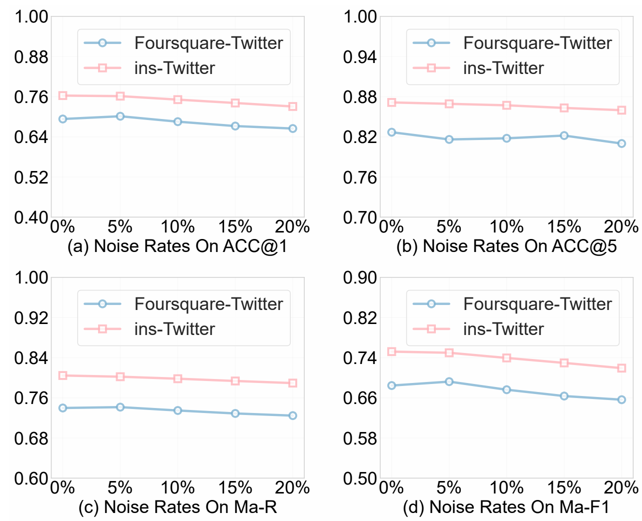

Figure 2: Robustness experimental results.

Conclusion

We proposes HANCUA, a framework that addresses issues in cross-domain user identity association. The integration of hierarchical multi-relational graphs with relation-aware attention mechanisms transforms low-quality trajectory data into robust behavioral representations, which construct identity mapping across heterogeneous service platforms. The spatiotemporal semantic collaborative learning module establishes dynamic interactions between local and global trajectory features. The ensemble correction decision module implements ensemble learning principles to mitigate distribution imbalances and enhance robustness against data quality variations. Extensive experiments on two real-world LBSN cross-domain datasets show that HANCUA achieves accurate user identify association and outperforms existing state-of-the-art baselines across all evaluation metrics. In the future, we plan to leverage richer behavioral contexts beyond spatial-temporal patterns to further enhance association accuracy.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (No. 62402338), the Key Project of Tianjin Natural Science Foundation (No. 25JCZDJC00760), and the Special Project for High-Quality Development of Manufacturing Industry in Tianjin Municipality (No. ZZY20251094).

References

Cen, K.; Yang, Z.; Wang, Z.; and Dong, M. 2024. A cross-domain user association scheme based on graph attention networks with trajectory embedding. *Machine Learning*, 113(10): 7905–7930.

Chen, H.; Yin, H.; Sun, X.; Chen, T.; Gabrys, B.; and Musial, K. 2020. Multi-level graph convolutional networks for cross-platform anchor link prediction. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, 1503–1511.

Chen, W.; Huang, C.; Yu, Y.; Jiang, Y.; and Dong, J. 2024. Trajectory-user linking via hierarchical spatio-temporal attention networks. *ACM Transactions on Knowledge Discovery from Data*, 18(4): 1–22.

Deng, L.; Sun, H.; Zhao, Y.; Liu, S.; and Zheng, K. 2023. S2tul: A semi-supervised framework for trajectory-user linking. In *Proceedings of the sixteenth ACM international conference on web search and data mining*, 375–383.

Feng, J.; Zhang, M.; Wang, H.; Yang, Z.; Zhang, C.; Li, Y.; and Jin, D. 2019. Dplink: User identity linkage via deep neural network from heterogeneous mobility data. In *The world wide web conference*, 459–469.

Gao, Q.; Zhang, F.; Yao, F.; Li, A.; Mei, L.; and Zhou, F. 2020. Adversarial mobility learning for human trajectory classification. *IEEE Access*, 8: 20563–20576.

Gao, Q.; Zhou, F.; Zhang, K.; Trajcevski, G.; Luo, X.; and Zhang, F. 2017. Identifying Human Mobility via Trajectory Embeddings. In *IJCAI*, volume 17, 1689–1695.

Goga, O.; Loiseau, P.; Sommer, R.; Teixeira, R.; and Gummadi, K. P. 2015. On the reliability of profile matching across large online social networks. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, 1799–1808.

He, W.; Huang, J.; Gu, J.; Zhang, J.; and Bai, Y. 2025. Filling the Missings: Spatiotemporal Data Imputation by Conditional Diffusion. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2025)*, 2874–2882.

Huang, H.; Ding, F.; Yin, H.; Liu, G.; Wang, C.; and Wu, D. O. 2023. EgoMUIL: Enhancing spatio-temporal user identity linkage in location-based social networks with Ego-Mo hypergraph. *IEEE Transactions on Mobile Computing*, 23(8): 8341–8354.

Korula, N.; and Lattanzi, S. 2013. An efficient reconciliation algorithm for social networks. *arXiv preprint arXiv:1307.1690*.

Li, H.; Cao, S.; Chen, Y.; Zhang, M.; and Feng, D. 2024. TULAM: trajectory-user linking via attention mechanism. *Science China Information Sciences*, 67(1): 112103.

Liu, S.; Fan, C.; Cheng, K.; Wang, Y.; Cui, P.; Sun, Y.; and Liu, Z. 2024. Inductive meta-path learning for schema-complex heterogeneous information networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Liu, S.; He, Y.; Wang, H.; Yang, W.; Wang, Y.; Cui, P.; and Liu, Z. 2025. Environment Inference for Learning Generalizable Dynamical System. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.

Ma, C. Y.; Yau, D. K.; Yip, N. K.; and Rao, N. S. 2010. Privacy vulnerability of published anonymous mobility traces. In *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, 185–196.

Ma, X.; Ding, F.; Peng, K.; Yang, Y.; and Wang, C. 2022. CP-Link: Exploiting continuous spatio-temporal check-in patterns for user identity linkage. *IEEE Transactions on Mobile Computing*, 22(8): 4594–4606.

Martins, A.; and Astudillo, R. 2016. From softmax to sparsemax: A sparse model of attention and multi-label classification. In *International conference on machine learning*, 1614–1623. PMLR.

Peters, B.; Niculae, V.; and Martins, A. F. 2019. Sparse sequence-to-sequence models. *arXiv preprint arXiv:1905.05702*.

Qin, Y.; Wu, H.; Ju, W.; Luo, X.; and Zhang, M. 2023. A diffusion model for poi recommendation. *ACM Transactions on Information Systems*, 42(2): 1–27.

Riederer, C.; Kim, Y.; Chaintreau, A.; Korula, N.; and Lattanzi, S. 2016. Linking users across domains with location data: Theory and validation. In *Proceedings of the 25th international conference on world wide web*, 707–719.

Rossi, L.; and Musolesi, M. 2014. It's the way you check-in: Identifying users in location-based social networks. In *Proceedings of the second ACM conference on Online social networks*, 215–226.

Shu, K.; Wang, S.; Tang, J.; Zafarani, R.; and Liu, H. 2017. User identity linkage across online social networks: A review. *Acm Sigkdd Explorations Newsletter*, 18(2): 5–17.

Wang, G.; Schoenebeck, S.; Zheng, H.; and Zhao, B. 2016. "Will Check-in for Badges": Understanding Bias and Misbehavior on Location-Based Social Networks. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 10, 417–426.

Wang, H.; Gao, C.; Li, Y.; Wang, G.; Jin, D.; and Sun, J. 2018. De-anonymization of mobility trajectories: Dissecting the gaps between theory and practice. In *The 25th annual network & distributed system security symposium (NDSS'18)*.

Wang, L.; Guo, D.; Wu, H.; Li, K.; and Yu, W. 2024. TC-GCN: Triple cross-attention and graph convolutional network for traffic forecasting. *Information Fusion*, 105: 102229.

Wu, C.; Wang, Z.; Cen, K.; Bai, Y.; and Hao, J. 2025. Cross-Domain Trajectory Association Based on Hierarchical Spatiotemporal Enhanced Attention Hypergraph. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 12854–12862.

Xu, F.; Zhang, G.; Chen, Z.; Huang, J.; Li, Y.; Yang, D.; Zhao, B. Y.; and Meng, F. 2018. Understanding motivations behind inaccurate check-ins. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW): 1–22.

Yang, P.; Wang, H.; Zhang, Y.; Qin, L.; Zhang, W.; and Lin, X. 2021. T3s: Effective representation learning for trajectory similarity computation. In *2021 IEEE 37th international conference on data engineering (ICDE)*, 2183–2188. IEEE.

Zhang, J.; Kong, X.; and Yu, P. S. 2014. Transferring heterogeneous links across location-based social networks. In *Proceedings of the 7th ACM international conference on Web search and data mining*, 303–312.

Zhang, J.; and Philip, S. Y. 2019. *Broad Learning Through Fusions*. Springer.

Zhou, F.; Chen, S.; Wu, J.; Cao, C.; and Zhang, S. 2021a. Trajectory-user linking via graph neural network. In *ICC 2021-IEEE International Conference on Communications*, 1–6. IEEE.

Zhou, F.; Dai, Y.; Gao, Q.; Wang, P.; and Zhong, T. 2021b. Self-supervised human mobility learning for next location prediction and trajectory classification. *Knowledge-Based Systems*, 228: 107214.

Zhou, F.; Gao, Q.; Trajcevski, G.; Zhang, K.; Zhong, T.; and Zhang, F. 2018. Trajectory-User Linking via Variational AutoEncoder. In *IJCAI*, 3212–3218.

Zhou, F.; Yin, R.; Trajcevski, G.; Zhang, K.; Wu, J.; and Khokhar, A. 2021c. Improving human mobility identification with trajectory augmentation. *GeoInformatica*, 25: 453–483.