Proceedings of Machine Learning Research 304, 2025 ACML 2025

High-Order Consistency-Guided User Identity Linkage with Large Language Model

Tianpeng Li LTPNIMEIA@TJU.EDU.CN
School of New Media and Communication, Tianjin University, Tianjin, China

Xiangjun Pei XPEI30@GATECH.EDU
Georgia Tech Shenzhen Institute, Tianjin University, shenzhen, China

Yinghui Wang WANGYINGHUI@TJU.EDU.CN
National Key Laboratory of Information Systems Engineering, Beijing, China

Qiyao Peng* QYPENG@TJU.EDU.CN
School of New Media and Communication, Tianjin University, Tianjin, China

Wenjun Wang WJWANG@TJU.EDU.CN
College of Intelligence and Computing, Tianjin University, Tianjin, China and Yazhou Bay Innovation Institute, Hainan Tropical Ocean University, China.

Editors: Hung-yi Lee and Tongliang Liu

Abstract

With the rapid expansion of the Internet, people commonly maintain multiple account identities across different online platforms, creating latent cross-network associations. User Identity Linkage (UIL), which seeks to identify and associate multiple accounts belonging to the same individual across platforms, has emerged as a vital research direction with broad applications in cross-platform recommendation, unified user profiling, and so on. However, existing methods face two major challenges in real-world environments: cross-platform feature heterogeneity and attribute-structure representation fusion. To address these challenges, this paper propose a Multi-View Feature High-Order Consistency-Guided User Identity Linkage method UIL-HC-MV. Our approach mitigates cross-network heterogeneity by deeply integrating multi-view features and mining consistency in shared thematic information among users and their relational networks. We decompose cross-platform feature heterogeneity into two subproblems: attribute heterogeneity and structural heterogeneity. We first fuse attribute and structural views by coupling nodes' random-walk sequences with neighborhood sampling to jointly extract node attributes and topological context. We then employ a Large Language Model to capture deep semantic information and contextual relationships across multiple text segments, distilling unified themes or high-order community features from the combined attribute-structure representation. Finally, we fine-tune a BERT model on the extracted high-order information to reinforce feature consistency and enable transfer learning for improved generalization. Extensive comparative experiments on real-world datasets demonstrate significant performance improvements over existing mainstream methods, validating the effectiveness of high-order information in alleviating cross-network heterogeneity and confirm the contribution of each component within our deeply integrated multi-view feature learning framework.

Keywords: User Identity Linkage(UIL), Graph Alignment, Graph Information Fusion, Graph High-Order Information Extraction, Cross-platform Heterogeneity modeling,LLM enhanced topic generation

* Corresponding Author

1. Introduction

In today's digital era, the rapid expansion of the Internet has fundamentally transformed how people access information and interact online, leading to an explosive growth of networked data. Online platforms—now an indispensable part of daily life—offer differentiated services that precisely address diverse user needs. For example, users may turn to a visual-centric lifestyle-sharing app (e.g., Instagram) to post daily snapshots, a microblogging service (e.g., Twitter) to join real-time discussions, and a film-and-book review community (e.g., Goodreads) to record their viewing and reading experiences. As a result, individuals often maintain multiple accounts across platforms, exhibiting distinct "identities" that carry latent cross-network associations.

Recent studies have examined how the same user's behavior manifests both similarities and divergences across different platforms Zhou et al. (2024). Service-specific features drive pronounced behavioral differentiation: on a lifestyle-sharing app, posts focus on personal moments. On a microblog, content centers on current events. And on a review community, entries reflect cinematic or literary critique. Likewise, a user's follower network varies by context—following influencers on one platform, news outlets on another, and critics on yet another—yielding diverse cross-platform interaction patterns. At the same time, users exhibit underlying behavioral consistencies: people tend to repost about the same topic across channels to amplify reach Zhang et al. (2021a), and core interests (e.g., artificial intelligence) often guide whom they follow on each service. These observations motivate cross-platform user modeling—commonly called User Identity Linkage (UIL)—as a vital research direction in network data mining Huang et al. (2022).

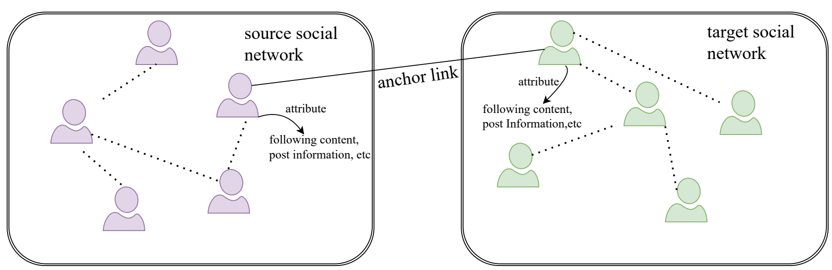

Figure 1: An example of the User Identity Linkage. Solid black lines connect accounts belonging to the same individual across two social networks—the "anchor link" join paired "anchor nodes".

User Identity Linkage (UIL) Senette et al. (2024) seeks to identify and associate the multiple account identities of a single person across different platforms. Mainstream approaches abstract this as an Anchor Link Prediction or Network Alignment problem. As illustrated in Figure 1, given a source network and a target network, an "anchor link" connects two nodes representing the same user. Those nodes are called "anchor nodes." Beyond social media, UIL methods apply in academic networks, e-commerce sites, and other domains—enabling cross-platform recommendation, unified user profiling, malicious-entity detection, and comprehensive assessment of researchers' contributions for fostering collaboration Zhang et al. (2023). Owing to its theoretical richness and broad applicability, UIL has become a focal point in complex-network science, big-data analytics, and data-mining research Li et al. (2021).

In summary, existing user identity linkage methods have made significant progress and generally rely on the consistency assumption—the premise that the same user should exhibit similar characteristics across different platforms—when exploring feature similarities between anchor nodes. However, in complex real-world network environments, these methods still exhibit two main limitations. This paper focuses on addressing the following challenges:

Cross–platform Feature Heterogeneity. We decompose this problem into two subproblems: attribute heterogeneity and structural heterogeneity. Here, "heterogeneity" refers specifically to the differentiated features of an anchor node (i.e., the same user) across platforms, which differs fundamentally from (a) traditional heterogeneous networks—where multiple node and edge types coexist-and (b) intra-network heterogeneity within a single homogeneous graph. Current mainstream approaches Chen and Chen (2022) often overlook the dual challenges posed by diverse content semantics and user-relationship patterns. While high-order information, capturing global user and community consistency, can help mitigate these heterogeneity effects, existing methods Wang et al. (2024) fall short in extracting such thematic features.

Attribute–structure Representation Fusion. In many fields, multi-view embedding fusion is critical for enhancing analytical effectiveness and model performance Li et al. (2019). Existing techniques Zhang et al. (2021b) typically apply naive concatenation, failing to account for the relative importance and intricate dependencies between attribute and structural embeddings. This leads to suboptimal fusion and underutilization of complementary information.

By analyzing how high-order information can alleviate heterogeneity, this work proposes a **U**ser **I**dentity **L**inkage method that guided with **H**igh-Order **C**onsistency and **M**ulti-**V**iew Features(UIL-HC-MV), which deeply integrates multi-view features to mitigate cross-network heterogeneity by mining the consistency of shared thematic information among users and their relational networks. Specifically, we first fuse attribute and structural views by coupling nodes' random-walk sequences with neighborhood sampling to jointly extract node attributes and topological context. We then employ a Large Language Model to capture deep semantic information and contextual relationships across multiple text segments, distilling unified themes or high-order community features from the combined attribute–structure representation. Finally, we fine-tune a BERT model Devlin et al. (2019) on the extracted high-order information to reinforce feature consistency and enable transfer learning for improved generalization. Extensive comparative experiments on real-world datasets demonstrate the effectiveness of each component and the overall benefit of deeply integrated multi-view feature learning within the proposed framework. The main contributions of this work are summarized as follows:

- **Graph Information Transformation for Cross-platform Heterogeneity.** We propose a strategy that deeply fuses node attributes and structural context—combining random-walk sequences with neighborhood sampling—to mitigate cross-platform feature heterogeneity.

- **High-Order Thematic Extraction via LLMs.** We leverage a Large Language Model to distill deep semantic and community-level themes from the fused multi-view representation, effectively addressing the challenge of attribute–structure embedding fusion.

- **BERT-Based Fine-Tuning and Transfer Learning.** We design a fine-tuning protocol and transfer-learning mechanism on BERT using extracted high-order information, thereby enhancing generalization performance and cross-domain adaptability.

- **Empirical Validation of Heterogeneity Mitigation.** We conduct extensive experiments on real-world datasets to demonstrate how high-order information alleviates cross-network heterogeneity and improves linkage accuracy.

2. Related Work

According to the types of features they exploit, current user identity linkage methods can be broadly classified into three categories:

Attribute-based methods. Early studies focused on comparing profile attributes across platforms, analyzing features such as username, gender, birthdate, interests, occupation, and location. Matching was typically performed via text-similarity measures (e.g., TF-IDF, edit distance) to identify accounts of the same person Mu et al. (2016). However, privacy concerns often lead users to conceal or falsify profile information, limiting the reliability of these approaches. To overcome this, researchers turned to publicly visible user-generated content—particularly text—to link accounts. Li et al. Li et al. (2018) develop a supervised framework that analyzes spatiotemporal and topical similarities in user content to identify identical accounts, while Chen et al. Chen et al. (2021) introduce a time-aware attention model that jointly considers temporal patterns and content similarity, addressing challenges of heterogeneous multimodal data and dataset sparsity. These methods assume that users maintain consistent posting patterns across platforms. When users deliberately vary their content, however, such assumptions break down, and deep thematic preferences remain difficult to capture through surface-level content analysis alone.

Structure-based methods. In addition to content, the structural properties of user relationships provide durable and accessible signals for cross-platform linkage. Researchers have increasingly leveraged network representation learning to exploit structural consistency. Huang et al. Huang et al. (2024) proposed the EgoMUIL model, which constructs a heterogeneous hypergraph incorporating both social topology and mobility behavior features, then learns user representations through Graph Convolutional Networks. The method introduces cross-domain ego networks to quantify relationship strength between users via topological similarity and stay-region similarity, thereby enhancing the model's linking performance in data-scarce scenarios. Some work—such as Wang et al.'s variational autoencoder approach with Laplacian regularization and heterogeneity constraints Wang et al. (2022) addresses structural incongruities, yet most still overlook the complementary signals in user-generated content and deep community themes. Zheng et al. Zheng et al. (2024) proposed the JORA model, which employs an inductive graph convolutional neural network to model user nodes, thereby learning their low-dimensional representations while effectively preserving users' core characteristics. The model simultaneously integrates two strategies: representation learning and alignment learning. Specifically, it maintains intra-network structural similarity through representation learning, while introducing both hard alignment and attention-based soft alignment mechanisms to mitigate errors caused by predefined similarity metrics in traditional approaches.

Methods Integrating Attribute and Structural Information. To enhance linkage robustness, advanced approaches fuse both content and structure. Wang et al. Wang et al. (2019) model interest themes by projecting structural and content features into a shared vector space, capturing cross-network co-occurrence and designing transition mechanisms between users and topics. Kaushal et al. Kaushal et al. (2020) proposed the NeXLink model, which generates local node representations by preserving the local structural information of nodes within a single network platform. For user nodes involved in cross-platform linking, it retains their shared structural information to learn global node representations. Finally, the model integrates these multi-level representations to perform user identity matching. Zhang et al. Zhang et al. (2020) introduce a multi-resolution transformation model for unified integration of structure and attributes. These methods generally assume attribute–structure consistency and employ low-rank matrix approximations for efficiency.

3. METHODOLOGY

As shown in Figure 2, the model mainly consists of the following five parts ,where parts one, two, and three are the model construction and training process, and parts four and five are the inference and matching process.

First, based on node attributes and network structure, attribute text chains and two types of text chains that integrate attributes and structure are constructed respectively. Subsequently, the model designs specific Prompt Engineering strategies for each text chain, and extracts high-order topic information from the text chains through a large language model. Then, the obtained topic text chains are used to fine-tune the BERT model. After that, the three types of topic text chains are input into the fine-tuned BERT model for topic representation learning to obtain the corresponding topic representations. Among them, the topic representation of the attribute text chain and the original attribute representation of the node are connected by a residual connection to obtain the semantic alignment representation of the node with enhanced high-order information, and the topic representations of the two types of text chains that integrate attributes and structure form the structural alignment representation of the node with integrated multi-perspective high-order information. Finally, the above representations are integrated into a complete node representation for the user identity linkage task.

3.1. Problem Definition

A graph (or network) consists of user nodes, inter-user relationships, and user-generated content, and can be represented as an attributed graph structure $G = (V, E, A)$, where V denotes the set of users, E represents user relationships, and A represents user attributes (e.g., textual content). In a cross-platform scenario, the same user exists in different networks and is regarded as an anchor user. The objective of this task is to infer all unknown matching user identity pairs given a set of known anchor user pairs. The main notations and their definitions used in this paper are listed in Supplementary Material Appendix A.

3.2. Text Chain Design

Text chain is a chain formed by splicing the text information in node attributes based on specific rules, and its form can be understood by an LLM. We proposes a graph information transformation strategy that integrates attribute and structure information to construct text chains, providing input for subsequent LLM inference.

Specifically, we constructs two types of text chains for each node v: attribute text chain and attribute and structure integrated text chain. Among them, the attribute text chain is based on the text attribute A_t of the node, and the attribute and structure text chain is constructed by capturing neighborhood information and random walk sequences and combining the node attributes in the neighborhood or sequence.

Attribute Text Chain. For a node v, its text attribute is represented as A_{t}, that is, the text information of the node v itself. It is directly converted into a text chain:

$$T_v^{\mathrm{a}} = f_{\mathrm{attr}}(A_{\mathrm{t}}) \ , \tag{1}$$

where $f_{\mathrm{attr}}(\cdot)$ represents the mapping function that converts the text attribute of the node into a text chain, and finally the text chain T_v^{a} is generated, which is used to represent the attribute information of the node.

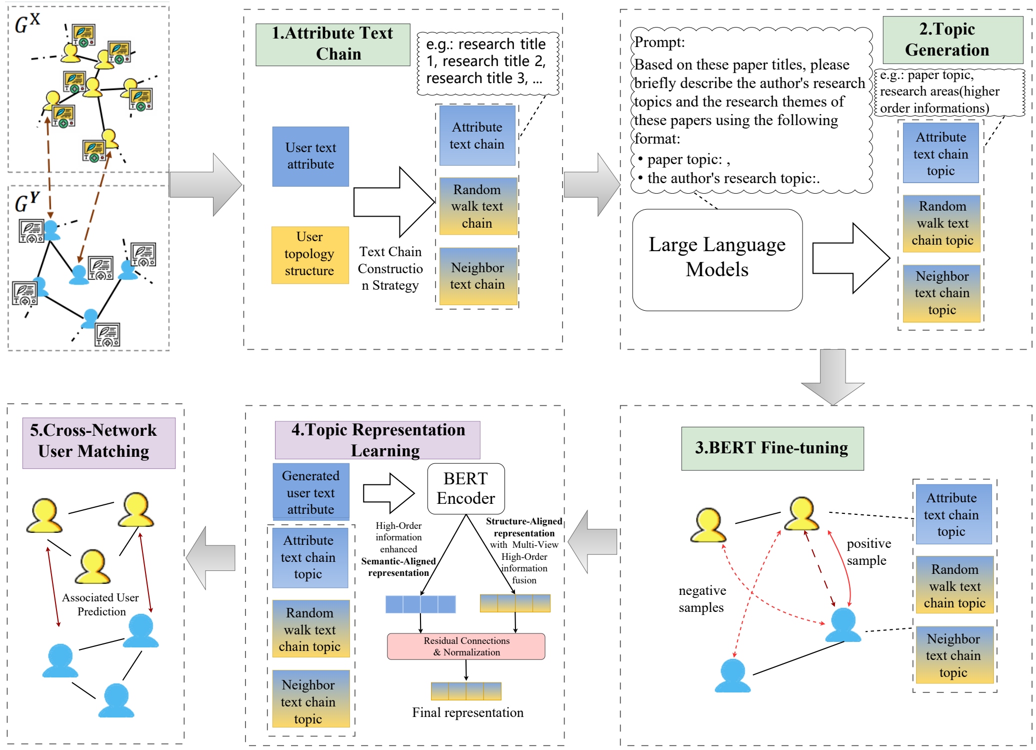

Figure 2: Overall Framework of UIL-HC-MV

Attribute and Structure Integrated Text Chain. The construction of the attribute and structure integrated text chain includes two parts: the neighbor text chain representing local neighborhood information and the random walk text chain representing global topological information.

(1) Neighbor Text Chain: Let the neighbor set of node v be $\mathcal{N}_v = \{u_1, u_2, ..., u_k\}$. We constructs the text chain T_v^{n} of its neighborhood using the co-authorship priority strategy:

$$T_v^{\mathrm{n}} = f_{\mathrm{neighbor}}(\{A_{\mathrm{t}}^{(u)} | u \in \mathcal{N}_v, \text{co-authorship priority}\}) \;, \tag{2}$$

where $f_{\mathrm{neighbor}}(\cdot)$ is the information conversion mapping function of global topological information. For specific scenarios such as a co-authorship network, the model will preferentially select the co-authorship content of neighbor nodes (such as the titles of jointly published papers). If the total length of the co-authorship content text does not reach the limited length of the text chain, the text attributes of other neighbor nodes will be further supplemented to ensure that the text chain has sufficient information.

(2) Random Walk Text Chain: We further adopts a random walk strategy based on the weights of node degrees and text attributes. Specifically, a node sequence $\mathcal{W}_v$ is generated starting from node v, and the generated text chain is the splicing of the node attributes in the sequence. During the walking process, the transition probability of a node is jointly determined by the degree information of the node and the length of the neighbor text attribute information:

$$P(u|v) = w_1 \cdot f_{\mathrm{d}}(d_u) + w_2 \cdot f_{\mathrm{a}}(A_{\mathrm{t}}^{(u)}) \;, \tag{3}$$

where d_u represents the degree of the neighbor node u of the current node, and $f_{\mathrm{d}}(\cdot)$ is used for degree normalization. $A_{\mathrm{t}}^{(u)}$ is the text attribute of the neighbor node, and $f_{\mathrm{a}}(\cdot)$ is used to calculate the amount of attribute information. w_1 and w_2 are weight parameters used to control the preference of the random walk for the degree and the length of the text information. The termination condition of the random walk process is that the length of the generated text sequence reaches the limited length of the text chain, and finally the random walk text chain is obtained by:

$$T_v^{\mathrm{r}} = f_{\mathrm{rw}}(\{A_{\mathrm{t}}^{(v_i)} | v_i \in \mathcal{W}_v\}) , \tag{4}$$

where $f_{\mathrm{rw}}(\cdot)$ is the information conversion mapping function of the random walk text chain, v_i is a node in the node sequence $\mathcal{W}_v$, and $A_{\mathrm{t}}^{(v_i)}$ is the text attribute of v_i.

In conclusion, this module obtains three text chains: the attribute text chain T_v^{a} represents the semantic information of the node itself. The neighbor text chain T_v^{n} captures local neighborhood information. The random walk text chain T_v^{r} considers the degree and text attribute information and walks on the graph to capture global information. The above text chains provide structured input for subsequent large language model inference, and the latter two text chains achieve the transformation of graph information that integrates attributes and structure.

3.3. Topic Generation of Text Chains

This part introduce an LLM-driven method for generating topics of text chains. By using the powerful reasoning and summarization capabilities of the LLM, the information of the text chains is aggregated and inducted, and redundant information is removed, so as to extract the essential topics in the text chains, and finally extract the more consistent high-order information in the attribute and structure features. This method does not require fine-tuning the LLM or extracting its intermediate representations, and can complete the task only relying on text input and output. Specifically, we uses ByteDance's Doubao LLM[1]. And through carefully designed Prompt Engineering, the LLM is effectively guided to extract high-order topic information. The Doubao has good generation ability and response stability, and at the same time has high accessibility and ease of use, so it is suitable as a basic model for performing complex language understanding tasks.

Topic Extraction of Attribute Text Chains. For the attribute text chain T_v^{a} of each node v, by designing a specific prompt template, the large language model is guided to extract the core topic of this text chain. The specific formal expression is as follows:

$$Q_{\mathrm{a}} = f_{\mathrm{prompt}}(T_v^{\mathrm{a}}) , \tag{5}$$

where $f_{\mathrm{prompt}}(\cdot)$ represents the preset prompt template. For example, in the academic co-authorship dataset, the prompt can be designed as:

> *"Based on these paper titles, please briefly describe the research topics of this author and the research topics of these papers in the following format:*
>
> *Paper Topic: ,*
>
> *Author's Research Topic: ."*

Topic Extraction of Attribute and Structure Integrated Text Chains. For the attribute and structure integrated text chains T_v^{n} (neighbor text chain) and T_v^{r} (random walk text chain), a special prompt template is designed to extract high-order structure information. The specific formalization is as follows:

1. https://www.volcengine.com/product/doubao

For the topic extraction of the neighbor text chain:

$$Q_{\mathrm{n}} = f_{\mathrm{prompt}}(T_v^{\mathrm{n}}) \ , \tag{6}$$

For the topic extraction of the random walk text chain:

$$Q_{\mathrm{r}} = f_{\mathrm{prompt}}(T_v^{\mathrm{r}}) \ , \tag{7}$$

For the design of the prompt for the attribute and structure integrated text chain, it is necessary to particularly emphasize the correlation between the initial node and the text chain to deeply excavate the network topology information and effectively extract the common concerned topics of the node users in the sequence and the initial user. Taking the academic co-authorship dataset as an example, the prompt design is as follows:

> *"I have collected the paper titles of multiple authors, and these authors belong to the academic co-authorship network around a specific author. Based on these paper titles, please identify the topics in the papers of other authors that are closest to or have a significant overlap with the research field of this specific author. Only focus on those fields that are highly relevant to the research topic of the specific author or directly built on the basis of his research (that is, do not list all the topics in the papers of other authors). Please briefly summarize your findings in the following format:*
>
> *Paper Topic: ,*
>
> *Author's Research Topic: ."*

Through the above method, we can effectively use the large language model to extract the heterogeneous text information in the text chains, so as to extract more consistent high-order semantic information, and at the same time retain the key information in the network topology structure, providing more effective information for the subsequent user identity linkage task.

3.4. BERT Fine-tuning Strategy Based on Text Chain Topics

Before the topic representation learning, this part introduce a BERT Vaswani et al. (2017) fine-tuning strategy based on text chain topics to further enhance the consistency of the features of anchor nodes across networks and improve the performance of the model in handling cross-network user identity linkage tasks. The fine-tuning strategy in this work is mainly as follows:

To fine-tune the BERT model, this paper first constructs three groups of positive sample data based on the cross-network anchor node training set, and each group of samples is composed of paired text chain topics. For each pair of anchor nodes $(v_i^{\mathrm{X}}, v_j^{\mathrm{Y}})$, this paper uses their three types of text chain topics: the attribute text chain topics $Q_{\mathrm{a}}^{(v_i^{\mathrm{X}})}, Q_{\mathrm{a}}^{(v_j^{\mathrm{Y}})}$, the neighbor structure text chain topics $Q_{\mathrm{n}}^{(v_i^{\mathrm{X}})}, Q_{\mathrm{n}}^{(v_j^{\mathrm{Y}})}$, and the random walk structure text chain topics $Q_{\mathrm{r}}^{(v_i^{\mathrm{X}})}, Q_{\mathrm{r}}^{(v_j^{\mathrm{Y}})}$. The goal is to annotate whether the pair of anchor nodes belongs to the same identity through the text chain topics. Specifically, if the anchor nodes $(v_i^{\mathrm{X}}, v_j^{\mathrm{Y}})$ represent the same identity in different networks G_{X} and G_{Y}, the label of their text chain topic samples (for example, $Q_{\mathrm{a}}^{(v_i^{\mathrm{X}})}$ and $Q_{\mathrm{a}}^{(v_j^{\mathrm{Y}})}$) is set to 1 (positive sample).

To construct negative samples, for each type of text chain topic of each pair of anchor nodes $(v_i^{\mathrm{X}}, v_j^{\mathrm{Y}})$, for example, for the text chain topic of the anchor node v_i^{X} (for example, $Q_{\mathrm{a}}^{(v_i^{\mathrm{X}})}$), we randomly selects a non-anchor node v_k^{Y} from the set of non-anchor nodes in G_{Y}, and selects its corresponding text chain topic (that is, $Q_{\mathrm{a}}^{(v_k^{\mathrm{Y}})}$) to form a negative sample. Similarly, for the topic of the other anchor node v_j^{Y} (for example, $Q_{\mathrm{a}}^{(v_j^{\mathrm{Y}})}$), a non-anchor node v_l^{X} is randomly selected from the set of non-anchor nodes in G_{X}, and its corresponding text chain topic (that is, $Q_{\mathrm{a}}^{(v_l^{\mathrm{X}})}$) is selected to form another negative sample, and the label value of this negative sample is 0.

In the optimization process, the model uses the CrossEntropyLoss function to measure and predict the difference between the label and the true label:

$$\mathcal{L}_{\text{finetune}} = -\left(y \cdot \log(p) + (1-y) \cdot \log(1-p)\right) \ , \tag{8}$$

where y is the true label of the node pair (1 indicates the same identity, 0 indicates a different identity), and p is the probability value output by the BERT model. This loss function aims to maximize the accuracy of the model in predicting node pairs with the same identity and minimize the misjudgment rate of node pairs with different user identities.

Transfer Learning Strategy. To improve the model's cross-domain ability and reduce the training cost, this work adopts a Zero-shot Transfer strategy. Specifically, first use the training set of the source dataset to fine-tune the BERT model, enabling it to learn the high-order semantic information of cross-network nodes. Since the fine-tuned BERT model has captured the cross-domain general features of the data, it can be directly transferred to the target dataset for the user identity linkage task, avoiding the computational overhead caused by repeated fine-tuning. During the transfer process, the target dataset uses the same text chain input form as the source dataset (node attribute text chain, neighborhood structure text chain, random walk structure text chain) for inference, and generates the final representation of the user. Through this transfer learning strategy, the model can adapt to the target dataset and improve the generalization ability of the model across datasets.

3.5. Topic Representation Learning

Based on the topic information extracted by the large language model, we uses the fine-tuned BERT to encode various types of topic texts to learn user feature representations. The representations obtained from the topic representation learning in this work are divided into: Enhanced High-level Information Semantic Alignment Representation (EHISA) and Multi-view High-level Information Structural Alignment Representation (MHISA).

Enhanced High-level Information Semantic Alignment Representation. For the topic Q_a of the user node attribute text chain, first perform word segmentation processing, and add special markers $[CLS]$ and $[SEP]$ before and after the sequence. Then input the sequence into the fine-tuned BERT to obtain the word-level representation matrix $\mathbf{H}_q^a$ of the last projection layer:

$$\mathbf{H}_q^a = \text{BERT}([\text{[CLS]}, w_1, w_2, \ldots, w_n, \text{[SEP]}]) \ , \tag{9}$$

Then, the topic representation of the node attribute text chain is obtained by performing average pooling on this matrix:

$$\mathbf{h}_v^a = \text{MeanPool}(\mathbf{H}_q^a) \ , \tag{10}$$

For each node attribute (denoted as A_t), we uses the same BERT encoding strategy as described above for processing, and performs a residual connection between the node attribute representation

$\mathbf{e}_v^a$ and the node attribute text chain topic representation $\mathbf{h}_v^a$ to obtain the final Enhanced High-level Information Semantic Alignment Representation $\mathbf{h}_v^{\text{attr}}$:

$$\mathbf{h}_v^{\text{attr}} = \mathbf{e}_v^a + \mathbf{h}_v^a \,. \tag{11}$$

This representation combines the original attribute information and the topic information through a residual connection, thereby enhancing the high-order semantic information of the node and forming a complete node attribute representation for the user identity matching task.

Structural Alignment Representation Integrating Multi-perspective High-order Information. For the topic Q_s of the neighborhood structure text chain and the topic Q_r of the random walk structure text chain, we also adopts a processing method similar to the semantic alignment representation with enhanced high-order information, and generates two types of text chain topic representations integrating attributes and structures respectively: $\mathbf{h}_v^n$ (representing the local topological structure of node v) and $\mathbf{h}_v^r$ (representing the global topological structure of node v). Subsequently, these two structural representations are fused through a residual connection, and finally the structural alignment representation $\mathbf{h}_v^{\text{str}}$ integrating multi-perspective high-order information is obtained:

$$\mathbf{h}_v^{\text{str}} = \mathbf{h}_v^n + \mathbf{h}_v^r \,, \tag{12}$$

This representation fuses the high-order information of attributes and structures, thereby obtaining the structural representation of the node.

Joint Representation of Attributes and Structures. Finally, the attribute representation $\mathbf{h}_v^{\text{attr}}$ and the structural representation $\mathbf{h}_v^{\text{str}}$ are fused through a residual connection and layer normalization (Add + Norm) to obtain the final representation of node v:

$$\mathbf{h}_v = \text{AddNorm}(\mathbf{h}_v^{\text{attr}}, \mathbf{h}_v^{\text{str}}) \,, \tag{13}$$

In this way, the node attribute information supplemented with high-order information and the high-order structural information (neighborhood structure and global structure) integrating structural and attribute information are integrated to form a complete node representation.

3.6. Cross-Network User Identity Matching

After completing the unified processing of the representations, we uses the distance measurement method to achieve cross-network user identity matching. Specifically, given nodes from two different networks with their corresponding final representations $\mathbf{h}_i^X$ and $\mathbf{h}_j^Y$, the Euclidean distance between them is calculated, that is, $D(\mathbf{h}_i^X, \mathbf{h}_j^Y) = |\mathbf{h}_i^X - \mathbf{h}_j^Y|_2^2$. The smaller the distance value, the more likely there is an anchor link between the two nodes, that is, the higher the possibility that they correspond to the same natural person. The training and implementation process of the model is in Supplementary Materials Appendix B.

4. Experiments

4.1. Experimental Setup

The dataset employed in this chapter is the DBLP_1 dataset Chen and Chen (2022) and the DBLP_2 dataset. Specifically: (1) The training set of DBLP_1 is used to fine-tune the BERT model to adapt it for the UIL-HC-MV model. (2) The BERT model fine-tuned on DBLP_1 is directly transferred and applied to the DBLP_2 dataset.

For the selection of baseline methods, we compares with several existing mainstream user identity linkage approaches: MAUIL Chen and Chen (2022), GAlign Trung et al. (2020), CENALP Du et al. (2022), Grad-Align Park et al. (2022), NeXtAlign Zhang et al. (2021b), and NetTrans Zhang et al. (2020). The aforementioned methods all propose solutions for the user identity linkage task by considering the fusion of attributes and structure. Additionally, this chapter uses the top@K (K=1,5,10) and MRR metrics to compare with other methods. The implementation details are introduced in the Supplementary Material Appendix C.

4.2. Analysis of Experimental Results

Table 1: Comparison of Experimental Results on Two Datasets

Method	DBLP_1				DBLP_2			
	top@1	top@5	top@10	MRR	top@1	top@5	top@10	MRR
UIL-HC-MV	**0.8188**	**0.8661**	**0.8854**	**0.8510**	**0.4283**	0.5769	0.6412	**0.5302**
Grad-Align	0.6151	0.7327	0.7807	0.6710	0.3696	0.5739	0.6756	0.4669
CENALP	0.2861	0.2951	0.2997	0.3172	0.2948	0.3254	0.3383	0.2446
GAlign	0.7612	0.8659	0.8847	0.7614	0.1849	0.4310	0.5221	0.1964
NeXtAlign	0.5035	0.6534	0.7094	0.5731	0.3439	**0.6264**	**0.7238**	0.4662
NetTrans	0.6360	0.7303	0.7562	0.6429	0.0884	0.1643	0.2004	0.1260
MAUIL	0.7388	0.8391	0.8756	0.8052	0.3354	0.5803	0.6995	0.5050

Table 1 presents the comprehensive comparison results of the UIL-HC-MV model against existing methods on two benchmark datasets. From the table, it can be observed that baseline models based on graph convolutional neural networks (e.g., Grad-Align, NeXtAlign, and GAlign) perform well on both datasets, indicating that such models can effectively capture graph structural information. However, the UIL-HC-MV model still outperforms these methods, demonstrating that its graph information transformation strategy not only integrates attribute information but also effectively captures structural information from the graph.

Moreover, UIL-HC-MV shows significant improvements in both evaluation accuracy (MRR) and top-K accuracy (top-K hit rate), with the MRR metric increasing by approximately 4-5% compared to the best-performing baseline model. The superior performance of UIL-HC-MV primarily stems from two aspects: (1) The semantically aligned representations enhanced by higher-order information effectively supplement node attribute information affected by heterogeneity, and the incorporated higher-order information significantly improves user identity linkage accuracy. (2) The structurally aligned representations that fuse multi-perspective higher-order information successfully integrate attribute and structural information while capturing higher-order local and global structural information beyond attributes.

Notably, when transferred to DBLP_2, the proposed model did not fine-tune the BERT model again using DBLP_2 data, yet it still outperformed other models in MRR and top@1. This indicates that the model's attribute-structure fusion strategy and higher-order information extraction method endow it with strong transferability and generalization capability, enabling it to adapt to the characteristics of different datasets and enhance cross-domain performance.

4.3. Ablation Study

This part conducts ablation experiments on the modules of the UIL-HC-MV model and components of the MHISA representation that integrates attribute and structural higher-order information. Notably, the Computational efficiency experiment and the analysis are introduced in the Supplementary Material Appendix D.

Ablation Study on UIL-HC-MV Model. First, ablation experiments are performed on the UIL-HC-MV model to analyze the impact of the Enhanced Higher-order Information Semantic Alignment (EHISA) representation and the Multi-perspective Higher-order Information Structural Alignment (MHISA) representation on experimental results, with top@1 as the evaluation metric. The following model variants are designed:

INITIAL: The baseline model, which only uses the original BERT model to represent user node attributes and performs user identity matching based on this representation. **EHISA**: Only the EHISA module is used for user identity matching, i.e., augmenting the INITIAL model with thematic representations from node attribute text chains. **MHISA**: Only the MHISA module is retained for user identity matching, while the rest remains the same as the INITIAL model.

Table 2: Ablation Results of the UIL-HC-MV Model

Dataset	INITIAL	EHISA	MHISA	UIL-HC-MV
DBLP_1	0.6871	0.7565	0.7553	**0.8188**
DBLP_2	0.3399	0.4067	0.4049	**0.4283**

The results in Table 2 reveal the following key findings: **INITIAL** performs the worst across all datasets, indicating that attribute heterogeneity negatively impacts alignment accuracy. **EHISA** and **MHISA** exhibit similar improvements, suggesting that although their information sources differ, both representations effectively provide higher-order information. The full model consistently outperforms variants using only a single representation, proving that they offer complementary higher-order information from different perspectives. The complete **UIL-HC-MV** model achieves the best performance on all datasets due to: (1) mitigating cross-platform heterogeneity by leveraging more consistent higher-order information, and (2) enhancing cross-network node representation consistency through deep integration of attribute and structural representations.

Ablation Study on MHISA Representation. For the MHISA representation, further ablation experiments are conducted to analyze the contributions of Neighbor Chain (NC) and Random Walk Chain (RWC) higher-order information, with top@1 as the evaluation metric. The tested variants are:

NC: Only uses neighbor text chain thematic representations for user identity matching. **RWC**: Only retains random walk text chain thematic representations for user identity matching. **MHISA**: Employs the full MHISA representation integrating multi-perspective higher-order information.

Table 3: Ablation Study on MHISA Representation

Dataset	NC	RWC	MHISA
DBLP_1	0.3859	0.2341	**0.5071**
DBLP_2	0.3003	0.2011	**0.3093**

Table 3 shows that **RWC** performs the worst on both datasets, indicating that global/long-range higher-order semantics are more dispersed, whereas **NC** shows better performance due to concentrated neighborhood higher-order semantics. Besides, **MHISA** achieves the best results across all datasets, demonstrating that neighborhood information from NC and global information from RWC are complementary.

4.4. Transfer Effect Analysis

This part focuses on analyzing the effectiveness of the model's text chain topic information application in the DBLP_1 dataset and its transfer performance on the DBLP_2 dataset. The study aims to evaluate the impact of three types of higher-order information on experimental results: attribute text chain topic information for BERT fine-tuning (Attribute Chain, AC), neighbor text chain topic information integrating attributes and structures (Neighbor Chain, NC), and random walk text chain topic information (Random Walk Chain, RWC). Using MRR as the evaluation metric, the experimental variants are:

BASE: The baseline model without BERT fine-tuning, which only constructs text chains, extracts higher-order information using large language models, and performs direct user matching. **AC**: Incorporates only attribute text chain topic information (AC higher-order information) for BERT fine-tuning, with other components identical to BASE. **NC**: Incorporates only neighbor text chain topic information (NC higher-order information) for BERT fine-tuning, with other components identical to BASE. **RWC**: Incorporates only random walk text chain topic information (RWC higher-order information) for BERT fine-tuning, with other components identical to BASE. **AC+NC**: Combines attribute text chain and neighbor text chain topic information for BERT fine-tuning, with other components identical to BASE. **AC+RWC**: Combines attribute text chain and random walk text chain topic information for BERT fine-tuning, with other components identical to BASE. **NC+RWC**: Combines neighbor text chain and random walk text chain topic information for BERT fine-tuning, with other components identical to BASE.

As shown in Table 4, applying text chain topic information significantly improves model prediction accuracy. Specifically, on DBLP_1, the UIL-HC-MV method achieves a 9% higher MRR than the BASE method without text chain optimization. When transferred to DBLP_2, the fine-tuned model still shows a 5% improvement, demonstrating that text chain topic information enhances both original dataset performance and cross-dataset generalization.

Table 4: Transfer Experiment Results, here UIL represents UIL-HC-MV model.

Dataset	Topic Information Type							
	BASE	AC	NC	RWC	AC+NC	AC+RWC	NC+RWC	UIL
DBLP_1	0.7592	0.8372	0.8336	0.8317	0.8275	0.8419	0.8321	**0.8510**
DBLP_2	0.4821	0.5124	0.4843	0.4786	0.4966	0.5044	0.4889	**0.5302**

Besides, Single text chain topic methods (AC/NC/RWC) improve upon BASE in the original dataset (DBLP_1) but show unstable transfer performance (DBLP_2), indicating limited generalization capability with isolated information sources. Combined methods (AC+NC/AC+RWC/NC+RWC) generally outperform single-source approaches across datasets, though not all combinations consistently surpass single-method results, suggesting complementary but suboptimally combined information. UIL-HC-MV achieves optimal performance by holistically integrating all text chain topic information, effectively enhancing representation consistency for maximal performance gains.

4.5. Analysis of Higher-order Information Effects

We evaluates the role of higher-order information through anchor node similarity, defined as:

$$\text{Sim} = \frac{\sum_{i=1}^{n} v_i \cdot u_i}{\sqrt{\sum_{i=1}^{n} v_i^2} \cdot \sqrt{\sum_{i=1}^{n} u_i^2}}, \tag{14}$$

where v_i and u_i represent components of cross-network node representations ($\tilde{\mathbf{h}}_i^X$ and $\tilde{\mathbf{h}}_j^Y$). Higher similarity indicates stronger information consistency.

Table 5: Analysis of Higher-order Information Effects

Dataset	INITIAL	AC	NC+RWC
DBLP_1	0.5348	0.6035	**0.6863**
DBLP_2	0.4685	0.5087	**0.5267**

The analysis compares three representation types: initial attribute-based (INITIAL), attribute chain (AC) topic information, and fused attribute-structure (NC+RWC) topic information. As Table 5 shows, both topic information extraction methods demonstrate higher anchor node similarity than raw attributes, confirming that text chain-derived higher-order information provides more consistent cross-network representations for user identity linkage.

5. Conclusion

In this paper, we proposes a high-order consistency-guided approach that effectively addresses the heterogeneity challenges in cross-platform user identity linkage by extracting high-order feature information. By analyzing how high-order information helps mitigate cross-network heterogeneity issues, the method implements a graph-based feature transformation strategy that deeply integrates attribute and structural information. The approach leverages the powerful semantic understanding capabilities of LLMs to mine high-order information from both attribute and structural features. It further enhances the consistency between attribute and structural features by utilizing this high-order information to fine-tune BERT, while also improving the model's generalization ability through transfer learning. Experimental results on two real-world datasets demonstrate that the method achieves deeper integration of attribute and structural representations. By leveraging more consistent high-order information, it effectively alleviates cross-network heterogeneity issues and overall improves model performance on the cross-platform user identity linkage task.

Acknowledgments

This paper was supported by the National Key R&D Plan of China(Grant No.2022YFC2602305)

References

Baiyang Chen and Xiaoliang Chen. Mauil: Multilevel attribute embedding for semisupervised user identity linkage. *Information Sciences*, 593(C):527–545, 2022.

Xiaolin Chen, Xuemeng Song, Siwei Cui, Tian Gan, Zhiyong Cheng, and Liqiang Nie. User identity linkage across social media via attentive time-aware user modeling. *IEEE Transactions on Multimedia*, 23(1):3957–3967, 2021.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 14th Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4171–4186, 2019.

Xingbo Du, Junchi Yan, Rui Zhang, and Hongyuan Zha. Cross-network skip-gram embedding for joint network alignment and link prediction. *IEEE Transactions on Knowledge and Data Engineering*, 34(3):1080–1095, 2022.

Haojun Huang, Fengxiang Ding, Hao Yin, Gaoyang Liu, Chen Wang, and Dapeng Oliver Wu. 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, 2024.

Xiaowen Huang, Jitao Sang, and Changsheng Xu. Image-based personality questionnaire design. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 18(4):1–20, 2022.

Rishabh Kaushal, Shubham Singh, and Ponnurangam Kumaraguru. Nexlink: Node embedding framework for cross-network linkages across social networks. In *Proceedings of the 6th NetSci-X International Winter School and Conference on Network Science*, pages 61–75, 2020.

Xiaoxue Li, Yanan Cao, Qian Li, Yanmin Shang, Yangxi Li, Yanbing Liu, and Guandong Xu. Rlink: Deep reinforcement learning for user identity linkage. *World Wide Web*, 24(1):85–103, 2021.

Yingming Li, Ming Yang, and Zhongfei Zhang. A survey of multi-view representation learning. *IEEE Transactions on Knowledge and Data Engineering*, 31(10):1863–1883, 2019.

Yongjun Li, Zhen Zhang, You Peng, Hongzhi Yin, and Quanqing Xu. Matching user accounts based on user generated content across social networks. *Future Generation Computer Systems*, 83(1): 104–115, 2018.

Xin Mu, Feida Zhu, Ee-Peng Lim, Jing Xiao, Jianzong Wang, and Zhi-Hua Zhou. User identity linkage by latent user space modelling. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1775–1784, 2016.

Jin-Duk Park, Cong Tran, Won-Yong Shin, and Xin Cao. Grad-align: Gradual network alignment via graph neural networks (student abstract). In *Proceedings of the 36th AAAI Conference on Artificial Intelligence*, volume 36, pages 13027–13028, 2022.

Caterina Senette, Marco Siino, and Maurizio Tesconi. User identity linkage on social networks: A review of modern techniques and applications. *arXiv preprint arXiv:2409.08966*, 2024.

Huynh Thanh Trung, Tong Van Vinh, Nguyen Thanh Tam, Hongzhi Yin, Matthias Weidlich, and Nguyen Quoc Viet Hung. Adaptive network alignment with unsupervised and multi-order convolutional networks. In *Proceedings of the 36th IEEE International Conference on Data Engineering*, pages 85–96, 2020.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pages 6000–6010, 2017.

Chenxu Wang, Peijing Jiang, Xiangliang Zhang, Pinghui Wang, Tao Qin, and Xiaohong Guan. Gtcalign: Global topology consistency-based graph alignment. *IEEE Transactions on Knowledge and Data Engineering*, 36(5):2009–2025, 2024.

Yaqing Wang, Chunyan Feng, Ling Chen, Hongzhi Yin, Caili Guo, and Yunfei Chu. User identity linkage across social networks via linked heterogeneous network embedding. *World Wide Web*, 22 (1):2611–2632, 2019.

Yinghui Wang, Qiyao Peng, Wenjun Wang, Xuan Guo, Minglai Shao, Hongtao Liu, Wei Liang, and Lin Pan. Network alignment enhanced via modeling heterogeneity of anchor nodes. *Knowledge-Based Systems*, 250(C):109116, 2022.

Fanjin Zhang, Xiao Liu, Jie Tang, Yuxiao Dong, Peiran Yao, Jie Zhang, Xiaotao Gu, Yan Wang, Evgeny Kharlamov, Bin Shao, Rui Li, and Kuansan Wang. Oag: Linking entities across large-scale heterogeneous knowledge graphs. *IEEE Transactions on Knowledge and Data Engineering*, 35 (9):9225–9239, 2023.

Peng Zhang, Baoxi Liu, Xianghua Ding, Tun Lu, Hansu Gu, and Ning Gu. Studying and understanding characteristics of post-syncing practice and goal in social network sites. *ACM Transactions on the Web*, 15(4):1–26, 2021a.

Si Zhang, Hanghang Tong, Yinglong Xia, Liang Xiong, and Jiejun Xu. Nettrans: Neural cross-network transformation. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 986–996, 2020.

Si Zhang, Hanghang Tong, Long Jin, Yinglong Xia, and Yunsong Guo. Balancing consistency and disparity in network alignment. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 2212–2222, 2021b.

Conghui Zheng, Li Pan, and Peng Wu. Jora: Weakly supervised user identity linkage via jointly learning to represent and align. *IEEE Transactions on Neural Networks and Learning Systems*, 35 (3):3900–3911, 2024.

Qi Zhou, Peng Zhang, Hansu Gu, Tun Lu, and Ning Gu. Exploring cross-site user modeling without cross-site user identity linkage: A case study of content preference prediction. *ACM Transactions on Information Systems*, 43(1):16:1–16:28, 2024.