

# Enhanced Cross-Platform User Identification Using Multi-Modal Embeddings and Ensemble Learning

Anudeep  
Department of Computer Science  
Amrita Vishwa Vidyapeetham  
am.en.u4cse22315@am.students.amrita.edu

Priti Gupta  
Department of Computer Science  
Amrita Vishwa Vidyapeetham  
am.en.u4cse22365@am.students.amrita.edu

## ABSTRACT

Cross-platform user identification has become increasingly important for understanding user behavior across social media platforms. This paper presents an enhanced approach for identifying users across LinkedIn and Instagram using multi-modal embeddings and ensemble learning techniques. Our methodology combines semantic, network, temporal, and profile embeddings through advanced fusion mechanisms, followed by an ensemble of specialized matchers including Enhanced GSMUA, Advanced FRUI-P, and gradient boosting methods. Experimental results on a dataset of 147 LinkedIn and 98 Instagram users demonstrate superior performance with 87% F1-score, 89% precision, and 85% recall, significantly outperforming existing baseline methods. The proposed ensemble approach shows 11.5% improvement over the best individual matcher, highlighting the effectiveness of multi-modal feature fusion and ensemble learning for cross-platform user identification.

**Keywords:** Cross-platform user identification, multi-modal embeddings, ensemble learning, social network analysis, feature fusion

## 1. INTRODUCTION

The proliferation of social media platforms has led to users maintaining multiple accounts across different services, creating a significant challenge for understanding comprehensive user behavior patterns. Cross-platform user identification, the task of determining whether

accounts on different platforms belong to the same individual, has emerged as a critical research area with applications in recommendation systems, fraud detection, and social network analysis. Traditional approaches to cross-platform user identification often rely on simple similarity metrics or single-modal features, which fail to capture the complex relationships between user profiles across platforms. Recent advances in deep learning and representation learning have opened new possibilities for more sophisticated approaches that can leverage multiple types of information simultaneously. This paper addresses the limitations of existing methods by proposing an enhanced cross-platform user identification system that combines: (1) Multi-modal feature extraction from semantic, network, temporal, and profile data, (2) Advanced fusion techniques using cross-modal and self-attention mechanisms, (3) Ensemble learning with specialized matchers optimized for different data modalities, and (4) Comprehensive evaluation on real-world LinkedIn and Instagram datasets.

## **2. METHODOLOGY**

### **2.1 System Architecture**

Our enhanced cross-platform user identification system consists of four main components organized in a hierarchical architecture: (1) Multi-Modal Feature Extraction: Generates embeddings from semantic, network, temporal, and profile data, (2) Advanced Fusion: Combines modalities using cross-modal and self-attention mechanisms, (3) Ensemble Matching: Applies four specialized matchers for different data types, and (4) Final Prediction: Uses meta-learning for optimal combination of matcher outputs.

### **2.2 Multi-Modal Feature Extraction**

We employ both TF-IDF based approaches for efficiency and BERT-based models for semantic richness. Network structure is captured using GraphSAGE with a fallback to Graph Convolutional Networks (GCN). Temporal patterns are captured using Time2Vec combined with Transformer architectures. User profile features are extracted using learned embeddings that capture demographic and behavioral patterns.

### **2.3 Ensemble Learning**

Our ensemble consists of four specialized matchers: (1) Enhanced GSMUA: Graph-based Social Media User Alignment with multi-head attention, (2) Advanced FRUI-P: Feature-Rich User Identification with weighted propagation, (3) Gradient Boosting: LightGBM for handling non-linear feature interactions, and (4) Optimized Cosine Similarity: Baseline method with learned thresholds.

### **3. EXPERIMENTAL RESULTS**

#### **3.1 Dataset and Setup**

We evaluate our approach on a real-world dataset consisting of 147 LinkedIn user profiles, 98 Instagram user profiles, and 156 ground truth pairs (81 matches, 75 non-matches). The dataset includes diverse user types ranging from technology professionals to artists and entrepreneurs.

#### **3.2 Performance Results**

Our enhanced approach achieves superior performance across all metrics: Precision: 89%, Recall: 85%, F1-Score: 87%, AUC-ROC: 92%. This represents an 11.5% improvement in F1-score over the best baseline method (FRUI-P with 78% F1-score). The ablation study demonstrates that multi-modal fusion improves F1-score by 14.3% over single modality, cross-modal attention contributes 4.9% improvement, and ensemble learning provides additional 1.2% improvement.

### **4. CONCLUSION**

This paper presented an enhanced approach for cross-platform user identification using multi-modal embeddings and ensemble learning. Our methodology effectively combines semantic, network, temporal, and profile information through advanced fusion mechanisms and specialized ensemble matchers. Experimental results demonstrate superior performance with 87% F1-score, representing significant improvements over existing approaches. Future work will explore federated learning approaches for privacy-preserving cross-platform identification and investigate the application to additional social media platforms.

### **REFERENCES**

- [1] Y. Zhang et al., "Cross-platform identification of anonymous identical users in multiple social media networks," *IEEE Trans. Knowledge Data Eng.*, vol. 28, no. 2, pp. 411-424, 2015.
- [2] S. Liu et al., "HYDRA: large-scale social identity linkage via heterogeneous behavior modeling," *Proc. ACM SIGMOD*, pp. 51-62, 2016.
- [3] A. Vaswani et al., "Attention is all you need," *Proc. NIPS*, pp. 5998-6008, 2017.
- [4] W. Hamilton et al., "Inductive representation learning on large graphs," *Proc. NIPS*, pp. 1024-1034, 2017.
- [5] J. Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Proc. NAACL*, pp. 4171-4186, 2019.