Inferring Hierarchical Structures in Human Networks Using a Graph Neural Network-based Unsupervised Learning Framework

Yeri Gu
R&D Department
UniSoul Friends. Co., Ltd.
Chuncheon, Korea
yerigu@dongsimwoo.com

Jion Kim
Byeongkwan Woo
Department of Forensic Information Science and Technology
Hallym University
Chuncheon, Korea
{jion972, byeongkwan.woo}@hallym.ac.kr

Myung-Sun Baek, *Member IEEE*
Department of Artificial Intelligence and Information Technology
Sejong University
Seoul, Korea
msbaek@sejong.ac.kr

***Abstract*— Cybercrime, characterized by anonymity and decentralized operations, presents major challenges in identifying suspects, uncovering hidden accomplice ties, and dismantling organizational structures. Traditional investigative methods often fail when explicit connections are absent, and early analytical approaches based on simple similarity metrics lead to significant information loss. To overcome these limitations, we propose a novel unsupervised learning framework centered on Graph Neural Networks (GNNs). Our approach begins by constructing a 3-layer heterogeneous graph that models entities as distinct node types: persons, cases, and identifiers (e.g., bank accounts, IP addresses), preserving the rich context of investigative data. The GNN then directly learns deep relational patterns from the graph's structure. This enables: (1) latent link prediction to uncover hidden relationships, (2) community detection on learned embeddings to identify criminal organizations, and (3) a comprehensive risk assessment that integrates network centrality with GNN-derived node importance scores. By effectively modeling the intrinsic complexity of the data, our framework provides investigators with objective, quantitative evidence to prioritize targets and enhance strategic decision-making.**

Keywords*— *Cybercrime, Network Analysis, Unsupervised Learning, Graph Neural Network (GNN), Criminal Investigation, Heterogeneous Graph

I. Introduction

The rapid growth of cybercrime—such as online fraud, phishing, and cryptocurrency laundering—is driven by the anonymity of cyberspace. Criminals frequently operate in decentralized, cell-based structures, using multiple alias accounts and disposable communication tools, making it exceedingly difficult for investigators to reconstruct their hierarchies [1]. Conventional investigative strategies, which follow explicit evidence trails, often falter when links are missing or intentionally obfuscated. This results in prolonged investigations and unresolved cases. While data-driven analysis offers a solution, early methods relying on simple attribute similarity compress diverse, meaningful relationships into a single score, leading to a critical loss of contextual information.

To address this gap, we introduce an unsupervised learning framework based on a Graph Neural Network (GNN). Unlike prior approaches, our framework operates without labeled data and is designed to capture the intrinsic structure of cybercrime networks. Our contributions are threefold:

1. Minimized Information Loss: We construct a heterogeneous 3-layer graph (Human–Case–Identifier) that preserves the nuanced meaning of different evidence types, moving beyond simplistic similarity scores.
2. Deep Relational Learning: We employ a GNN to directly learn the complex, underlying patterns of relationships from the graph's structure, enabling more accurate discovery of latent ties and organizations.
3. Multi-Dimensional Risk Assessment: We introduce a sophisticated risk model that fuses traditional network centrality with GNN-derived node importance and domain-based severity metrics for robust suspect prioritization.

II. Related Work

Early research on criminal network analysis predominantly applied Social Network Analysis (SNA) to visualize structures and identify central actors [2]. While effective for small-scale analysis, SNA struggles with heterogeneous data and non-explicit ties. With the rise of machine learning, semi-supervised learning (SSL) frameworks have been applied to suspect prediction. Kim et al. [3] used label propagation to predict suspects in hierarchical networks, while Jhee et al. [4] introduced a latent network to improve inference speed in large-scale graphs.

However, SSL approaches require labeled data, limiting applicability in early-stage investigations. Moreover, most prior studies compress relationships into single similarity scores, leading to loss of semantic and structural information. Our framework addresses this by treating investigative identifiers as independent nodes and using an unsupervised GNN to learn their unique relational patterns directly from the data.

III. Methodology for Inferring Hierarchical Structures in Human Networks

The proposed methodology comprises three stages: (1) Heterogeneous Graph Construction, (2) GNN-based Relationship Learning and Organization Analysis, and (3) Risk Assessment.

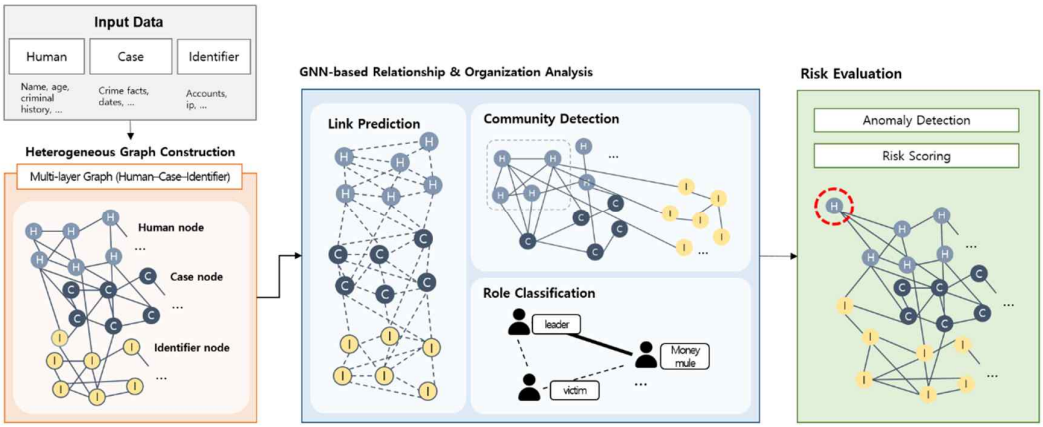

Fig. 1 Conceptual diagram of the framework for inferring hierarchical structures in human networks

A. Heterogeneous Graph Construction

We integrate structured and unstructured data into a heterogeneous graph $G = (V, E)$. To preserve the structural context and minimize information loss, we construct a 3-layer heterogeneous graph by defining three distinct node types and the edges that connect them:

- Human nodes (V_H) : All individuals involved in cases, such as suspects, victims, and persons of interest, with each node containing attributes like name and contact information.
- Case nodes (V_C): Individual criminal incidents, with attributes like case number and crime type.
- Identifier nodes (V_I) : Mediating entities that link humans and cases, such as bank account numbers, IP addresses, and crypto wallet addresses. These nodes preserve the explicit evidence connecting other entities.
- Edges (E): The relationships between nodes, such as a person owning an account (Human-Identifier) or being involved in a case (Human-Case).

1) Illustrative Example – Voice Phishing Case:

Consider a case where Human A (organization leader) instructs Human B (money mule) to withdraw funds from Victim C. The stolen money is transferred through Bank Account X (Identifier1), and communication occurs via a disposable phone (Identifier2) and a Telegram alias (Identifier3). Human A issues instructions through an IP address (Identifier4).

In this scenario, our framework models:

- Human–Identifier edges (e.g., Human A ↔ IP address, Human B ↔ bank account),
- Case–Identifier edges (e.g., Case ↔ disposable phone), and
- Human–Case edges (e.g., Victim C ↔ Case).

This explicit modeling allows investigators to capture latent ties between Human A and Human B, even if they never directly appear in the same evidence record.

B. GNN-based Relationship Learning and Organization Analysis

Using the constructed heterogeneous graph, a GNN learns the complex patterns of different relationship types and generates a low-dimensional vector representation (embedding) for each node.

1) Latent Link Prediction and Community Detection

The similarity between node embeddings is used to predict unobserved latent links (potential accomplices), following the classical formulation of the link prediction problem [5]. Applying community detection algorithms such as the Louvain method [6] allows for more accurate identification of potential criminal organizations than methods based on simple similarity scores.

2) Role Classification

A classifier can be built upon the learned embeddings to automatically categorize the functional roles of individuals (e.g., leader, money launderer), providing actionable insights into the group's operational structure.

C. Risk Assessment

To prioritize investigative resources, we quantitatively assess the risk level of each individual. The risk score $R_{(v)}$ is defined as a weighted combination of three components:

$$R_{(v)} = w_1 C_{struct}\ (v) + w_2 C_{GNN}(v) + w_3 E_S(v)$$

- C_{struct} : The structural centrality of a node, calculated from metrics like Degree and Betweenness Centrality.
- C_{GNN} : The GNN-based node importance, derived from attention scores, which captures latent influence missed by traditional metrics.
- $E_S(v)$: A score reflecting the average severity of the cases in which the individual is involved, based on domain knowledge.

The final risk score provides an objective, multi-faceted ranking to guide the allocation of limited investigative resources.

IV. Conclusion and Future Work

This paper introduced an unsupervised GNN-based framework designed to address the limitations of traditional approaches in analyzing complex cybercrime networks. By constructing a multi-layer heterogeneous graph and leveraging GNNs to capture deep relational patterns, the proposed methodology provides a systematic means of uncovering hidden hierarchical structures, identifying latent organizations, and generating objective risk assessments of key individuals and groups.

The principal contribution of this work lies in presenting a conceptual framework that advances the integration of heterogeneous evidence sources with unsupervised learning techniques. Moving forward, an important direction is the empirical validation of the framework using large-scale, real-world cybercrime datasets to assess its robustness and operational effectiveness. Future research will also focus on: (1) experimenting with diverse GNN architectures to enhance inference accuracy, (2) developing data-driven methods to optimize the weighting scheme for risk assessment, and (3) incorporating Explainable AI (XAI) techniques to provide transparent and interpretable justifications for the model's reasoning. Collectively, these extensions aim to establish the framework as a practical and trustworthy tool to support law enforcement in cybercrime investigations.

Acknowledgement

This research was supported in part by the Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded by the Korea Government [Ministry of Science and ICT (MSIT)], Republic of Korea, under the Metaverse Support Program to Nurture the Best Talents under Grant IITP-2024-RS-2023-00254529; in part by Development of an Integrated Analysis and Inference System for Cybercrime Investigation Clues program through the Korea Institutes of Police Technology(KIPoT) funded by the Korean National Police Agency (RS-2025-02218280).

References

[1] Cyber Safety Bureau, Korean National Police Agency, *Analysis of cybercrime trends 2023*, Seoul, Korea, 2024.

[2] M. K. Sparrow, "The application of network analysis to criminal intelligence: An assessment of the prospects," *Social Networks*, vol. 13, no. 3, pp. 251–274, 1991.

[3] M. Kim, D.-G. Lee, and H. Shin, "Semi-supervised learning for hierarchically structured networks," *Pattern Recognition*, vol. 95, pp. 191–200, Nov. 2019.

[4] J. H. Jhee, M. J. Kim, M. Park, J. Yeon, and H. Shin, "Fast prediction for criminal suspects through neighbor mutual information-based latent network," *International Journal of Intelligent Systems*, Article ID 9922162, 2023.

[5] D. Liben-Nowell and J. Kleinberg, "The link-prediction problem for social networks," *J. Am. Soc. Inf. Sci. Technol.*, vol. 58, no. 7, pp. 1019–1031, 2007.

[6] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," J. Stat. Mech.: Theory Exp., vol. 2008, no. 10, P10008.