Universal Ring-of-Abusers Detection via Multi-Modal Heterogeneous Graph Learning

Yiyue Qian ✉, Lanhao Chen, Song Cui, De Chen
Amazon, Seattle, USA
{iamyiyue, lanhaoc, songcui, dchenam} @amazon.com

ABSTRACT

As fraudulent and abusive activities performed by groups continue to plague e-commerce stores, we realize that detecting groups of abusers, or Ring-of-Abusers (RoAs), has become crucial. Unlike existing works about abuser detection on e-commerce stores that merely consider the individual features of abusers or the relationships among abusers, we design a **U**niversal **R**ing-**O**f-**A**busers **D**etection framework (abbreviated as **U-ROAD**) that integrates multi-modal features (e.g., numerical feature, text, and image) and the rich relationships among entities (i.e., seller, buyer, and product) on e-commerce stores. Especially, the U-ROAD framework designs three meta-paths to depict the high-order heterogeneous semantic relationships among entities. Then it leverages graph neural networks (GNNs) for obtaining the semantic node embeddings, which are fused with an aggregation layer and fed to an MLP classifier to detect abusers (e.g., abusive sellers and buyers) in different scenarios. As a result, (i): *Flexible*: The U-ROAD framework can be easily customized with different training labels to detect different types of abusers at e-commerce stores (e.g., rank abusers, financial abusers, and competitor abusers). (ii): *Powerful and Timely*: The framework is empirically powerful than some existing methods, which achieves an average 30% improvement in precision and 7.5% improvement in recall on four abuser (i.e., abusive seller and abusive buyer) detection tasks in two stores. Besides, our U-ROAD framework can detect abusers earlier than some existing methods on e-commerce stores. (iii): *Explainable*: The U-ROAD framework develops a visualization tool to help investigators understand the RoAs detection results.

KEYWORDS

Abuser Detection, Heterogeneous graph, Graph neural network

ACM Reference Format:
Yiyue Qian ✉, Lanhao Chen, Song Cui, De Chen. 2023. Universal Ring-of-Abusers Detection via Multi-Modal Heterogeneous Graph Learning. In *International Workshop on Resource-Efficient Learning for Knowledge Discovery (KDD '23), August 14–18, 2023, Long Beach, CA, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Workshop on Resource-Efficient Learning for Knowledge Discovery at KDD '23, August 6–10, 2023, Long Beach, CA, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

With the proliferation of e-commerce stores in recent decades, the occurrence of fraud activities orchestrated by organized groups, known as **R**ing-**o**f-**A**buser**s** (**RoAs**), has emerged as a major concern for consumers and the overall integrity of online marketplaces. These interconnected networks of abusers skillfully exploit system vulnerabilities, targeting different facets of the e-commerce ecosystem and engaging in a wide array of abusive practices including financial collusion, the proliferation of fake reviews, abuse of promotional mechanisms, and manipulation of ranking systems. For instance, RoAs create fake accounts and post positive reviews for their own products or negative reviews for their competitors, distorting the perception of product quality and influencing purchasing decisions. Another common scheme involves financial collusion, where multiple abusers including abusive sellers and buyers work together to commit financial fraud through a set of products on e-commerce stores, which is illustrated in Figure 1 where an abusive seller (identified as 4896) collaborates with a group of abusive buyers to commit financial fraud through a set of products.

To combat these activities, an increasing number of machine-learning methods have been employed on e-commerce stores to detect suspicious RoAs and take swift action against fraudulent actors. However, these methods still have several limitations: (i) They aim at detecting abusive activities either based on individual features (behavior patterns) or the relationships among entities. For instance, [19] merely leverages text information to detect fake reviews online while ignoring the rich relationships among sellers, buyers, and products. (ii) Existing works study the abusive activities merely from a single perspective, either from the seller or buyer side, but do not consider the rich connections among sellers and buyers (e.g., financial collusion among sellers and buyers in Figure 1).

To effectively detect RoAs, it is crucial to consider both multi-model features and rich relationships among sellers, buyers, and products on e-commerce stores. Unlike single types of features, multi-modal features (i.e., numerical features, text, and image content) are more informative to provide a comprehensive understanding of user intent online. For instance, some abusive sellers simply duplicate others' products to conduct financial fraud or non-fulfillment fraud. Integrating the numerical (e.g., order velocity), text (e.g., product description), and image (e.g., product image) features can comprehensively describe the behaviors of users on e-commerce stores.

Therefore, to handle the aforementioned limitations, we propose a universal RoAs Detector, abbreviated as **U-ROAD**, to detect RoAs in different scenarios (e.g., financial abuse, rank abuse, and competitor abuse). Inspired by the investigation pattern (e.g., similarity in the product description, images among products, order patterns, and listing information) of multiple entities and the relationships

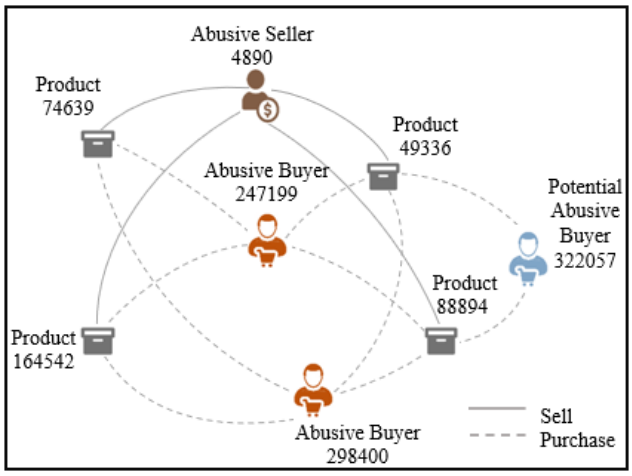


Figure 1: A showcase about financial collusion on e-commerce stores. All nodes are anonymous.

among sellers, buyers, and products, the U-ROAD framework first builds a heterogeneous graph including three types of nodes (i.e., seller, buyer, and product) and nine types of relationships (e.g., sell, purchase, etc.) to integrate various relationships among nodes and multi-modal node features. It also defines three meta-paths to depict the high-order heterogeneous semantic relationships among nodes. Afterwards, it leverages the random walk with a restart to generate meta-path guided semantic subgraphs and further designs a graph neural networks (GNNs) framework to propagate the multi-modal features among entities in each semantic subgraph. The semantic node embeddings from each subgraph are fused via an aggregation layer and are fed into the MLP classifier to detect the abusive sellers and buyers simultaneously, and further detect the RoAs. To conclude, the contributions of this work is concluded as follows:

- **Flexible**: We design a universal heterogeneous graph learning framework (the U-ROAD framework) to propagate the multi-modal features among sellers, buyers, and products to detect abusive sellers and buyers simultaneously, which can be easily customized to detect various types of abusers on e-commerce stores (i.e., rank abuse, finance abuse, competitor abuse, etc.).
- **Powerful and Timely**: The U-ROAD framework achieves an average improvement of almost 30% in precision, and 7.5% in recall. Besides, the framework is able to detect abusers much earlier than existing models, which can prevent potential damage to e-commerce stores and their customers from multiple fraudulent and abusive activities.
- **Explainable**: To aid in the understanding of RoAs detection via the U-ROAD framework, we have developed a visualization tool to display rich relationships among RoAs in an intuitive manner.

2 RELATED WORK

Abuser Detection. Most existing abuser detection tasks [7–9] focus on designing models based on either useful features or graph structure. For instance, some existing methods for fake reviewer detection leverage text information to detect fake reviews and further detect abusive reviewers [1]. However, these works underestimate the combination of multi-modal features (i.e., numerical feature, text, and image) and graph structure in detecting abusers on e-commerce stores. Moreover, existing methods study the abusive activities from a single perspective, either the seller or buyer side, but do not consider the rich connections among sellers and buyers on e-commerce stores [5, 19].

Graph Neural Networks. Existing graph neural networks (GNNs) methods [2, 4, 14, 16] aim to learn node embeddings by aggregating the features of neighbors via neural networks. For instance, GCN [4] implements the layer-wise propagation rule to learn the node embedding, and GraphSAGE[2] proposes to sample subgraphs and train a learnable aggregator for graph inductive learning. To deal with heterogeneous graphs, some heterogeneous GNN models are proposed to model the heterogeneity by using meta-path (e.g., HAN [15]) or meta-graph (Meta-GNN [11]). However, these works fail to depict the rich relationships among the real-world abuse graphs and can not be applied to real-world abuser detection tasks effectively and efficiently. Therefore, we design a universal heterogeneous graph learning framework that integrates multi-modal features and rich relationships to learn seller and buyer embeddings simultaneously for RoAs detection in different scenarios.

3 PRELIMINARY

In this section, we introduce the relevant definitions and formally define the RoAs detection problem.

Definition 3.1. **Heterogeneous Graph.** To comprehensively describe the e-commerce stores data on e-commerce stores, we build a heterogeneous graph (HG) [18], $G = (\mathcal{V}, \mathcal{E}, \mathcal{X})$, where $\mathcal{V}$ is the set of different types of nodes, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges, and $\mathcal{X}$ is the multi-modal attribute feature set. They are associated with a node type mapping function $\phi : \mathcal{V} \rightarrow \mathcal{T}$ and an edge type mapping function $\phi : \mathcal{E} \rightarrow \mathcal{R}$, where $\mathcal{T}$ and $\mathcal{R}$ are the set of node types and the set of relation types with $|\mathcal{T}| + |\mathcal{R}| > 2$. The graph schema for G is a graph with nodes from $\mathcal{V}$ and edges from $\mathcal{E}$. As shown in Figure 2.(a), we design three types of nodes and nine types of relations in our built HG.

Definition 3.2. **Meta-path.** A meta-path [13] P is a path defined on the graph schema, which is denoted in the form of $T_1 \xrightarrow{R_1} T_2 \xrightarrow{R_2} \ldots \xrightarrow{R_L} T_{L+1}$ where $R = R_1 \cdot R_2 \cdot \ldots \cdot R_L$ ($T_i \in \mathcal{T}$) is the composite relation between node types T_1 and T_{L+1}, and L is the length. In this work, we manually define three meta-paths $\mathcal{P} = \{P_1, P_2, P_3\}$ illustrated in Figure 2.(a)) to extract the semantic relations among different types of nodes based on domain knowledge.

Problem 1. ***Ring-of-Abusers Detection***. *Given data on e-commerce stores, including different types of nodes (i.e., seller, buyer, and product) with multi-modal features X, seller and buyer label Y, as well as multiple types of relationships R among nodes, we build a heterogeneous graph G to model the rich multi-modal node features and relationships among nodes and further design a universal heterogeneous graph learning framework to learn the seller and buyer embeddings simultaneously for RoAs detection in different scenarios (e.g., rank abuse and financial abuse).*

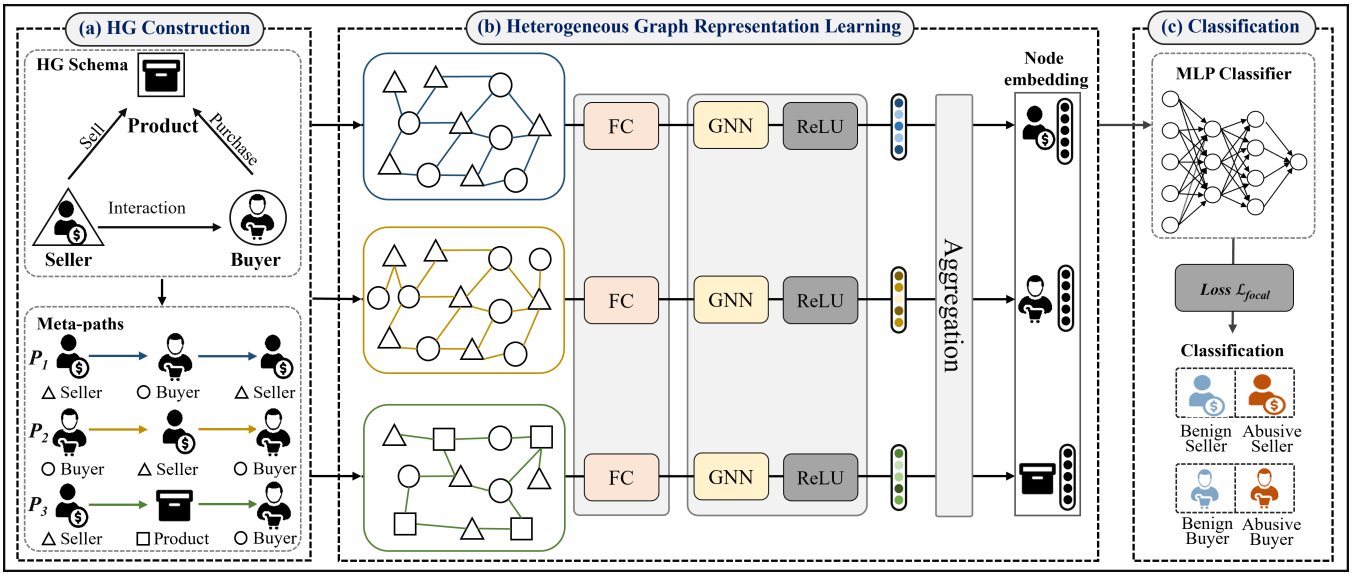


Figure 2: The framework of U-ROAD: (a) It first designs the heterogeneous graph (HG) schema to extract the rich relationships among entities (i.e., seller, buyer, and product) on e-commerce stores. Three meta-path are designed to depict the heterogeneous semantic relationships among nodes. (b). It then designs a heterogeneous graph learning framework to learn the node embeddings in meta-path guided semantic subgraphs, which are fused via an aggregation layer. (c). It last designs an MLP-based classifier optimized by the focal loss to detect abusive sellers and buyers and further detect RoAs.

4 PROPOSED MODEL

In this section, we present the details of the U-ROAD framework which includes three modules: heterogeneous graph construction (Figure 2.(a)), heterogeneous graph representation learning (Figure 2. (b)), and classification for RoAs detection (Figure 2. (c)).

4.1 Heterogeneous Graph Construction

Feature. The graph has been proven to be effective in modeling network data (e.g., e-commerce network). Thus, we propose to leverage graphs to depict both multi-modal features and various relationships within data. First of all, we extract the multi-modal features for entities (i.e., sellers, buyers, and products) on e-commerce stores, and further concatenate all features from different modalities. For instance, we apply the pre-trained language model *Sentence-Bert* [10] to get the text embedding (d_i = 384) for text information and we leverage the pre-trained image model ResNet to obtain the image feature vector (d_i = 2048) for images content. Figure 3 also illustrates an example about multi-modal content on e-commerce stores.

Relation. Besides extracting the informative multi-modal features, we also define a bunch of relation types (e.g., seller-sell-product and buyer-purchase-product) among entities to depict the rich relationships among entities on e-commerce stores. Moreover, we design three meta-paths denoted as P_1, P_2, P_3 to depict the high-order heterogeneous semantic relations among entities, which is illustrated in Figure 2.(a). To summarize, as the graph schema is shown (Figure 2.(a)), we build a HG by integrating both informative multi-modal features and rich relations among nodes. The multi-modal features are attached to all nodes in graphs. In addition, we define three meta-paths based on domain knowledge to generate the semantic-based subgraphs.

4.2 Heterogeneous Graph Representation Learning

4.2.1 Meta-path Guided Semantic Subgraph. To model the high-order semantic relationships among nodes in HG, we employ the random walk with a restart to generate the semantic subgraphs guided by meta-paths $\mathcal{P}$ (i.e., P_1, P_2, and P_3). Specifically, for each node v_i in $\mathcal{V}$, we select the meta-path from $\mathcal{P}$ to generate the random walks guided by the specific meta-path. Once the length of the walk is equal to the threshold K (i.e., 200), we stop randomly walking. After that, we obtain the semantic subgraph $G_P = (\mathcal{V}, \mathcal{E}_P, \mathcal{X})$, where P refers to the specific meta-path from $\mathcal{P}$. G_{P_1} and G_{P_2} extract all of the relationships among users, and G_{P_3} depicts all of the relationships between users and products in our graph. After obtaining all semantic subgraphs, as different types of nodes in subgraphs have different dimensions, we apply a type-specific mapping layer to transform all attribute features into the same space:

$$x_i = g(X_i) = X_i W_{T_i}, \tag{1}$$

where X_i is the original attribute feature of node v_i, x_i is the transformed node feature, and W_{T_i} is the transform weight matrix for node type T_i ($|T| = 3$).

4.2.2 Graph Representation Learning. **Graph Representation Learning on Semantic Subgraph.** The U-ROAD framework is designed as a universal framework that is applicable to most GNNs models, i.e., GCN [4], GraphSAGE [2], GAT [14], and GIN [17], to learn the node embeddings. Here we take GCN as an example to

Figure 3: A showcase of multi-modal content in an product on e-commerce stores.

introduce how we learn the node embeddings in each semantic subgraph. GCN is a layer-wise propagation rule-based model to learn the node embedding $Z_i^P \in \mathbb{R}^d$ (d: embedding dimension) corresponding to the node $v_i \in G_P$. Following the basic idea of GCN, the convolutional layer is devised as:

$$H^{P,L+1} = \sigma(\widetilde{A^P} H^{P,L} W^{P,L}), \tag{2}$$

where $H^{P,L+1}$ denotes the node embedding at $L+1$ layer and $H^{P,0} = x$ represents the transformed attribute feature. $\widetilde{A^P}$ is a symmetric normalization of A^P with self-loop, i.e., $\widetilde{A^P} = \hat{A^P}^{-\frac{1}{2}} \hat{A^P} \hat{A^P}^{-\frac{1}{2}}$ with $\hat{A^P} = A^P + I_N^P$. A^P, I_N^P, $\hat{D^P}$ are the adjacency matrix, the identity matrix, and the diagonal node degree matrix of $\hat{A^P}$, respectively. $W^{P,L}$ denotes the weight matrix at L-th layer, and σ is the activation function, such as the $\mathrm{ReLu}(\cdot) = \max(0, \cdot)$. Previous studies have shown that graph convolution is a type of Laplacian smoothing [4]. However, if we apply Laplacian smoothing many times in a deep neural network, all node embeddings will converge to similar values with their neighbors. So we suggest a three-layer GCN to detect abusers and we denote the user embedding as:

$$Z^P = \mathrm{GCN}(x, A^P) = (\widetilde{A^P}\,(\mathrm{ReLU}\,(\widetilde{A^P}\,(\mathrm{ReLU}\,(\widetilde{A^P} x\, W^{P,0}) W^{P,1}) W^{P,2}), \tag{3}$$

where $W^{P,0}, W^{P,1}, W^{P,2}$ denote the weight matrix for the input layer, the second convolutional layer, and the third convolutional layer, respectively. Z^P is the node embeddings in graph G_P.

Embedding Fusion from Multiple Sematic Subgraphs. After obtaining the node embeddings Z^P from each semantic subgraph, we design an efficient aggregation layer to fuse node embeddings:

$$\mathbf{Z}_i = \mathrm{AGG}(Z_i^P) = Z_i^{P_1} \oplus Z_i^{P_2} \cdots \oplus Z_i^{P_k}, \tag{4}$$

where $\oplus$ is the concatenate operator, $\mathbf{Z}_i$ is fused embedding of vi, and k is the number of meta-paths.

4.3 Abuser Detection Task

After obtaining the node embeddings $\mathbf{Z}$, for user nodes, we feed $\mathbf{Z}$ into a fully-connected (MLP) layer to predict the probability of a user being abusive, i.e., $\hat{Y}_i = \mathbf{Z}_i W_c$. Moreover, in order to address the class imbalance problem (i.e., abusive users and benign users), we introduce Focal Loss [6] that applies a modulating term to the cross-entropy loss to focus on hard misclassified nodes to detect abusive users. In particular, the supervised focal loss $\mathcal{L}_{\mathrm{focal}}$ in this work can be formally defined as:

$$\mathcal{L}_{\mathrm{focal}} = -\frac{1}{|\mathcal{V}_l|} \sum_{i \in \mathcal{V}_l} \alpha\,(1 - \hat{Y}_i)^\gamma \log(\hat{Y}_i), \tag{5}$$

where $\mathcal{V}_l$ is the node sets of labeled users, γ is the focusing parameter to control the rate at which easy nodes will be down-weighted, and $\alpha \in [0, 1]$ is a weighting hyper-parameter for classes. The **pseudo-code** of the U-ROAD framework is provided in Algorithm 1, which will facilitate the understanding of our designed model.

Algorithm 1: Pseudo-code of the U-ROAD Framework

Data: Graph G, Multi-modal feature X, Graph structure $\mathcal{E}$, Defined Meta-path set $\mathcal{P}$, Labeled users Y.
Result: Node abusive score $\hat{Y}$.

```
for each meta-path P ∈ 𝒫 do
    for each node v_i ∈ G do
        Perform random walk with a restart to generate walk path;
for each sematic subgraph G_P do
    Apply the type-specifical mapping layer to transform attribute features via Eq. 1;
    Apply the transformed features to the GNNs and obtain the node embeddings via Eq. 3;
Fuse semantic node embeddings from each semantic subgraph via Eq. 4;
Apply the MLP classifier to detect abusive users via Eq. 5;
Infer the potential Ring-of-Abusers based on the detection results via the visualization tool introduced in Section 5.6;
Return Abusive sellers and buyers, and the inferred RoAs.
```

5 EXPERIMENT

In this section, we first introduce two datasets from different real-world e-commerce stores (M1 and M2) to detect different types of abusers, i.e., financial abusers, and rank abusers. Then we discuss all baseline models and experimental settings in this work. Afterward, we conduct extensive experiments to evaluate our framework on four tasks, i.e., financial abusive seller detection on M1, financial abusive buyer detection on M1, rank abusive seller detection on M2, and rank abusive buyer detection on M2. Moreover, our developed visualization tool with real-world case studies are introduced in Section 5.6.

5.1 Dataset

All data we collected are from real-world e-commerce stores. To validate the effectiveness of our U-ROAD framework, we collected two datasets from different stores, i.e., M1 stores and M2 stores for financial abuser detection and rank abuser detection, respectively. Mention that all data we used in experiments are encrypted and anonymous. None of the personal and private data is involved in our experiments.

Financial Abuse on M1 Store. To study the financial abuse collusion activities on e-commerce stores, we build a graph on the M1 stores. Specifically, we collect the order data from a specific time period on M1 store to construct the HG with 337,046 nodes and 1,571,971 edges including 3,629 abusive sellers, 39,106 benign sellers, 33,137 abusive buyers, 58,130 benign buyers, and 203,044 products. The labels of seller and buyer are generated based on historical enforcement records. Labeled sellers and buyers are split into three sets, training set, validation set, and testing set for model training, validation, and testing, respectively.

Rank Abuse on M2 Store. In order to comprehensively evaluate the U-ROAD framework, we propose to collect another dataset involved with rank abuse on the M2 store. Following the three-step mechanism, we also collect the order data from a time period to construct the HG with 1,549,009 nodes and 13,589,523 edges including 2,544 abusive sellers, 171,757 benign sellers, 32,982 abusive buyers, 619,701 benign buyers, and 722,025 products.

5.2 Baseline Method

To comprehensively evaluate our U-ROAD framework, we compare it with three groups of baseline models:

G1: Feature-based model. For the abuser detection task, we apply the seller or buyer features to a 2-layer MLP classifier [12] to directly detect abusers.

G2: Graph structure-based methods. We merely consider the relationships (edges) in graphs and apply the edge to three GNNs including GIN [17], GraphSAGE [2], and GCN [4] to learn the node embedding. We generate the random features for all nodes as the feature vector for GNN models. For fair comparisons, all GNNs models have three layers. Similar to G1, we feed the user embedding to a 2-layer MLP classifier to detect abusers.

G3: Multi-modal features with graph structure methods. Except for graph structure, we also leverage the multi-modal features of nodes to learn the node embedding from GNNs models and we feed the user embeddings to a 2-layer MLP classifier for abuser detection tasks.

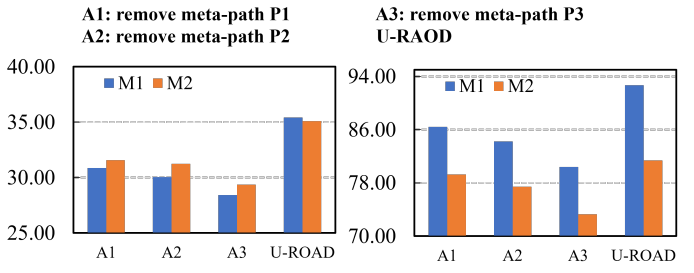
U-ROAD. U-ROAD designs three meta-paths to depict the semantic relations and leverages the multi-modal features and graph structure to learn the node embedding from each meta-path guided semantic subgraph. The fused node embeddings are fed to a 2-player MLP classifier to detect abusive users in different scenarios.

5.3 Experimental Setting

To evaluate the performance of the U-ROAD framework and all baseline methods, we adopt three widely-used metrics: precision showing the percent of abusers that are covered by the model, recall score showing the percent of abusers being detected correctly, and ROC-AUC showing the ability of the model in distinguishing abusers and benign users on e-commerce stores. Experiments are conducted under the environment of the g5.48xlarge with GPU resources. For each GNN model (i.e., GIN, GraphSAGE, and GCN), we employ three layers neural network with weight decay 1e-5 and the dimension of node embedding generated by GNNs is 200. We use Adam [3] optimizer with a learning rate of 0.001. All experiments are conducted in ten runs and the mean and stand deviation over testing data are provided.

5.4 Performance Comparison

Financial Abuser Detection on M1 Store. Table 1 shows the performances of all methods for financial abusers (i.e., abusive seller and buyer) detection on the M1 store. From Table 1, we can conclude that: (i) By comparison with G1 and G2, we find that rich graph relations among nodes increase the model performance as the performances of models with single graph structure relations in G2 are much better than the performances of the model with only user features in G1.



(a) Precision in abusive seller detection. **(b) Recall in abusive buyer detection.**

Figure 4: Performance of model variants for abuser detection on different stores.

(ii) Except for structure relationships, multi-modal features can also enhance the detection performance as the performances of graph methods with only graph structure in G2 are worse than graph methods with multi-modal features and graph structure in G3. (iii) By comparison with all baseline models, our framework

Table 1: Performance comparison for collusion financial abuse on M1 store (mean % ± std). All are conducted in ten runs to obtain mean and stand deviation. Best performances are highlighted in purple.

Setting		M1 Store-Seller			M1 Store-Buyer		
Group	Model	Recall	Precision	AUC	Recall	Precision	AUC
G1	MLP	12.04 ± 6.18	5.24 ± 5.27	41.25± 7.35	31.01 ± 10.24	32.04 ± 9.34	42.50± 7.34
G2	GIN	21.13 ± 5.24	10.84 ± 5.37	51.24 ± 6.17	55.96 ± 7.34	57.04 ± 6.58	55.25 ± 5.74
	GraphSAGE	25.39 ± 7.12	7.29 ± 6.41	49.82 ± 5.51	58.28 ± 7.51	56.86 ± 6.34	57.43 ± 6.77
	GCN	33.72 ± 5.54	12.45 ± 5.31	54.82 ± 5.72	71.11 ± 6.48	58.43 ± 6.27	58.09 ± 6.34
G3	GIN	54.31 ± 5.59	14.57 ± 5.31	56.71 ± 6.24	71.68 ± 6.34	63.90 ± 6.47	63.48 ± 6.71
	GraphSAGE	63.24 ± 13.74	12.50 ± 5.15	55.52 ± 6.11	68.95 ± 6.24	62.82± 6.31	60.65 ± 6.84
	GCN	67.37 ± 4.57	20.11 ± 4.84	68.67 ± 4.25	78.97± 5.61	62.94 ± 5.98	62.32 ± 5.78
Our	**U-ROAD**	76.42 ± 3.21	35.42 ± 4.25	80.21 ± 3.15	92.70 ± 4.14	61.77 ± 4.87	72.46 ± 4.51

Table 2: Performance comparison for collusion rank abuse on M2 store (mean % ± std). All are conducted in ten runs to obtain mean and stand deviation. Best performances are highlighted in purple.

Setting		M2 Store-Seller			M2 Store-Buyer		
Group	Model	Recall	Precision	AUC	Recall	Precision	AUC
G1	MLP	23.01 ± 8.57	5.67 ± 8.41	42.53± 8.25	21.05 ± 9.14	6.57 ± 9.21	45.21± 9.33
G2	GIN	42.34 ± 6.28	19.52 ± 6.14	57.31 ± 6.19	46.96 ± 6.45	9.48 ± 6.56	47.02 ± 6.55
	GraphSAGE	45.27 ± 6.21	25.51 ± 6.31	63.58 ± 6.05	48.38 ± 6.35	10.53 ± 6.33	48.92 ± 6.28
	GCN	50.57 ± 5.89	31.14 ± 5.92	74.46 ± 5.74	52.98 ± 6.21	13.86 ± 6.24	46.55 ± 6.11
G3	GIN	58.23 ± 5.14	21.36 ± 5.32	58.83 ± 5.41	68.33 ± 5.67	10.88 ± 5.69	53.24 ± 5.47
	GraphSAGE	60.54 ± 5.22	25.77 ± 5.13	60.30 ± 5.01	70.89 ± 5.48	12.40 ±5.46	57.72 ± 5.48
	GCN	66.59 ± 5.01	32.16 ± 5.07	82.70 ± 05.16	78.63 ± 5.34	17.27 ± 5.36	56.53 ± 5.33
Our	**U-ROAD**	83.86 ± 4.13	35.10 ± 4.24	85.10 ± 4.18	81.36 ± 5.19	24.50 ± 5.17	67.32 ± 5.14

achieves the best performance in detecting financial-abusive sellers and buyers, showing the superiority of U-ROAD for financial abuser detection on M1 store. Mention that, our U-ROAD framework has gained at least **30% improvement** in precision, **5%** improvement in recall for abusive **seller** detection, and over **25%** improvement in precision, almost **10%** improvement in recall for abusive **buyer** detection, by comparison with the model that only leveraging pure attribute features on e-commerce stores.

Rank Abuser Detection on M2 Store. Table 2 lists the performances of all methods for rank abusers detection on M2 store and we find out that all metrics increase obviously when the relationships among nodes are considered by comparing G1 with G2. Besides, multi-modal features contribute to learning better representations in GNN models. Last, the U-ROAD framework has the best performance among all methods and achieves almost **30% improvement** in precision and almost **5%** improvement in recall for abusive **seller** detection on M2 store, almost **20%** improvement in precision and **10%** improvement in recall for abusive **buyer** detection on M2, by comparison with existing models.

5.5 Ablation Study

We also conduct experiments to analyze the effectiveness of each meta-path by removing each meta-path P_1 (A1), P_2 (A2), and P_3 (A3), respectively. From Figure 4, we find out that each meta-path contributes to our designed U-ROAD framework with different levels. Meta-path P_3 designs to depict the semantic relationships among sellers and buyers via products has large contributions to the U-ROAD framework, showing that meta-path P_3 is very significant to depict the relationships among sellers, buyers, and products on real-world e-commerce stores. Besides, Meta-path P_1 and P_2 also contribute to our U-ROAD framework with obvious decreases by comparison with the U-ROAD framework in Figure 4, showing that

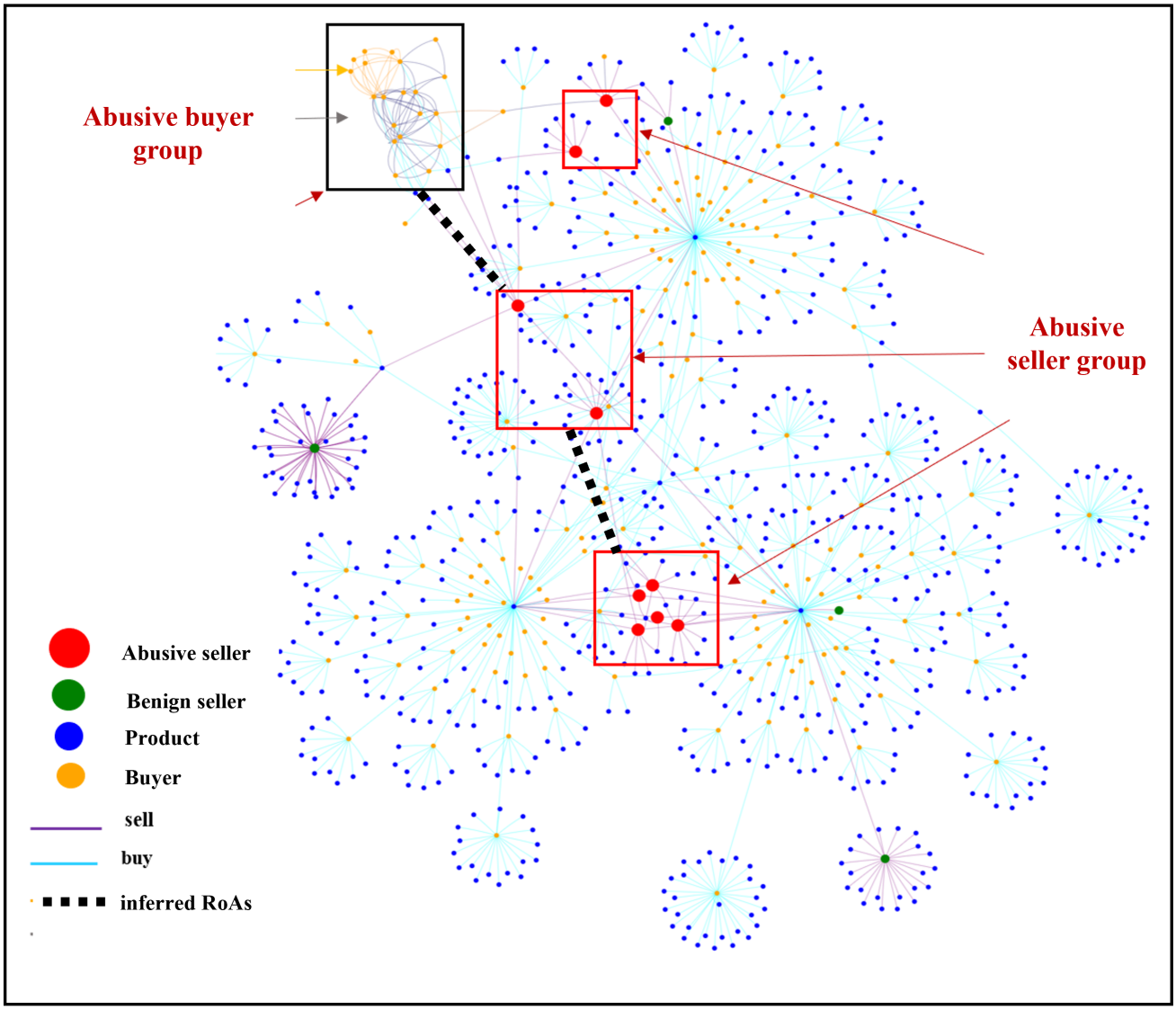


Figure 5: The visualization of node 5774 with 3-hop neighbors.

P_1 and P_2 are also useful to describe the high-order relationships among three types of entities on e-commerce stores.

5.6 Visualization Tool

Moreover, we can infer the RoAs based on the detected abusive sellers and buyers. To better visualize the RoAs detection and provide investigators with intuitive results, we develop a visualization tool to show the RoAs within multiple hops neighbors. The convenient tool is designed for deep-diving into the inferred RoAs. Investigators can easily input the node id and then would obtain all neighbor nodes connected with various relations. Next, we would show an example to illustrate how to utilize the developed tool for further investigation of RoAs.

Case Study. Figure 5 shows a financial abusive seller node 5774 (anonymized) on the M1 store with 3-hop neighbors. From Figure 5, we can find out two abuser groups, including the abusive buyer group, and the abusive seller group. Red nodes refer to abusive sellers; green nodes represent benign sellers; blue nodes stand for products and yellow nodes symbolize buyer nodes. We can easily find that abusive buyers in the black box are closely connected to each other. In addition, the abusive buyer groups connect closely and have rich connections to abusive seller groups, which can infer

a potential collusion RoAs (linked with black dash line in Figure 5) among the busive seller groups and abusive buyer groups.

6 CONCLUSION

In this paper, we design and develop a universal heterogeneous graph learning framework U-ROAD to detect different types of ring-of-abusers on e-commerce stores. The U-ROAD framework proposes to propagate the multi-modal features among sellers, buyers, and products to learn the seller and buyer embeddings, which enables it to detect abusive sellers and buyers simultaneously. Besides, the U-ROAD framework is very flexible to various types of abusers and can be easily customized with different training labels to detect different types of abusers on real-world e-commerce stores. Extensive experiments on four abuser detection tasks over two real-world stores demonstrate that the U-ROAD framework is more powerful than some existing models. Moreover, this framework develops a visualization tool to explain the RoAs detection in a more intuitive way.

REFERENCES

[1] Emerson F Cardoso, Renato M Silva, and Tiago A Almeida. Towards automatic filtering of fake reviews. *Neurocomputing*, 309:106–116, 2018.

[2] William L Hamilton, Rex Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *NeurIPS*, 2017.

[3] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *ICLR*, 2015.

[4] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *ICLR*, 2017.

[5] Zhao Li, Pengrui Hui, Peng Zhang, Jiaming Huang, Biao Wang, Ling Tian, Ji Zhang, Jianliang Gao, and Xing Tang. What happens behind the scene? towards fraud community detection in e-commerce from online to offline. In *WWW*, 2021.

[6] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *ICCV*, 2017.

[7] G Vishnu Manohar, Biplab Bhattacharjee, and Maheshwar Pratap. Preventing misuse of discount promotions in e-commerce websites: an application of rule-based systems. *International Journal of Services Operations and Informatics*, 11(1):54–74, 2021.

[8] Tuga Mauritsius, Sofia Alatas, Faisal Binsar, Riyanto Jayadi, and Nilo Legowo. Promo abuse modeling in e-commerce using machine learning approach. In *ICOT*, 2020.

[9] Mahmud Hasan Munna, Md Rifatul Islam Rifat, and ASM Badrudduza. Sentiment analysis and product review classification in e-commerce platform. In *ICCIT*, 2020.

[10] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In *EMNLP-IJCNLP*, 2019.

[11] Aravind Sankar, Xinyang Zhang, and Kevin Chen-Chuan Chang. Meta-gnn: Meta-graph neural network for semi-supervised learning in attributed heterogeneous information networks. In *ASONAM*, 2019.

[12] Donald F Specht et al. A general regression neural network. *IEEE Transactions on Neural Networks*, 1991.

[13] Yizhou Sun, Jiawei Han, Xifeng Yan, Philip S Yu, and Tianyi Wu. Pathsim: Meta path-based top-k similarity search in heterogeneous information networks. *VLDB*, 2011.

[14] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. In *ICLR*, 2018.

[15] Xiao Wang, Houye Ji, Chuan Shi, Bai Wang, Yanfang Ye, Peng Cui, and Philip S Yu. Heterogeneous graph attention network. In *WWW*, 2019.

[16] Chengfeng Xu, Pengpeng Zhao, Yanchi Liu, Victor S Sheng, Jiajie Xu, Fuzhen Zhuang, Junhua Fang, and Xiaofang Zhou. Graph contextualized self-attention network for session-based recommendation. In *IJCAI*, 2019.

[17] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In *ICLR*, 2019.

[18] Carl Yang, Yuxin Xiao, Yu Zhang, Yizhou Sun, and Jiawei Han. Heterogeneous network representation learning: A unified framework with survey and benchmark. *TKDE*, 2020.

[19] Anil R Yelundur, Vineet Chaoji, and Bamdev Mishra. Detection of review abuse via semi-supervised binary multi-target tensor decomposition. In *KDD*, 2019.