**INTRODUCTION**

Internet services have brought human beings with ecommerce, social networking, and entertainment platforms, which not only facilitate information exchange but also provide chances to fraudsters. Fraudsters disguise themselves as ordinary users to publish spam information [1] or collect user privacy, compromising the interest of both platforms and users. In addition, multiple entities on the Internet are connected with multiple relationships. Traditional machine learning algorithms cannot handle this complicated heterogeneous graph data well. The current approach is to model the data as a heterogeneous information network so that similarities in characteristics and structure of fraudsters can be discovered. Due to the effectiveness in learning the grap representation, graph neural networks (GNNs) have already been introduced into fraud detection areas including product review [2]–[5], mobile application distribution [6], cyber crime identification [7] and financial services [8], [9]. However, most existing GNN based solutions just directly apply homogeneous GNNs, ignoring the underlying heterogeneous graph nature and camouflage node behaviors. This problem has drawn great attention with many solutions proposed [4], [5], [10]. Graph Consis [4] found that there are three inconsistency problems in fraud detection and CAREGNN [5] further proposed two camouflage behaviors. These problems could be summarized as follows: \_

Camouflage: Previous work showed that crowd workers could adjust their behavior to alleviate their suspicion via connecting to benign entities like connecting to highly reputable users, disguise fraudulent URLs with special characters [3], [6], or generate domain-independent fake reviews via generative language model [11] to conceal their suspicious activities.

\_ Inconsistency: Two users with distinct interests could be connected via reviewing a common product such as food or movies. Direct aggregation makes GNNs hardly distinguish the unique semantic user pattern. Also, if a Use r is suspicious, then the other one should be more likely to be distrustful if they are connected by common activity relation since fraudulent users tend to post many fraudulent reviews in the same short period.

To address the above two problems, many methods have been proposed. Graph Consis addresses the inconsistency problem by computing the similarity score between node embeddings, which cannot distinguish nodes with different types. CAREGNN enhances GNN-based fraud detectors against camouflaged fraudsters by reinforcement learning based neighbor selector and relation aware aggregator. Its performance still suffers from the heterogeneous graph. In this paper, we introduce the Fraud Aware Heterogeneous Graph Transformer(FAHGT), where we propose heterogeneous mutual attention to address the inconsistency problem and design a label-aware neighbor selector to solve the camouflage problem. Both are implemented in a unified manner called the “score head mechanism”. We demonstrate the effectiveness and efficiency of FAHGT on many real world datasets. Experimental results suggest that FAHGT can significantly improve KS and AUC over state-of-the-art GNNs as well as GNN-based fraud detectors.

The advantages of FAHGT can be summarized as follows: \_

Heterogeneity: FAHGT is able to handle heterogeneous graphs with multi-relation and multi-node type without designing meta-path manually \_ Adaptability: FAHGT attentively selects neighbors given a noise graph from real-world data. The selected neighbors are either informative for feature aggregation or risky for fraud detection.

\_ Efficiency: FAHGT admits a low computational complexity via a parallelizable multi-head mechanism in relation scoring and feature aggregation.

\_ Flexibility: FAHGT injects domain knowledge by introducing a flexible relation scoring mechanism. The score of a relation connecting two nodes not only comes from direct feature interaction but is also constrained by domain knowledge.