Fraud Detection in Online Product Review Systems via Heterogeneous Graph Transformer

ABSTRACT

In online product review systems, users are allowed to submit reviews about their purchased items or services. However, fake reviews posted by fraudulent users often mislead consumers and bring losses to enterprises. Traditional fraud detection algorithm mainly utilizes rule-based methods, which is insufficient for the rich user interactions and graph-structured data. In recent years, graph-based methods have been proposed to handle this situation, but few prior works have noticed the camouflage fraudster’s behavior and inconsistency heterogeneous nature. Existing methods have either not addressed these two problems or only partially, which results in poor performance. Alternatively, we propose a new model named Fraud Aware Heterogeneous Graph Transformer(FAHGT), to address camouflages and inconsistency problems in a unified manner. FAHGT adopts a type-aware feature mapping mechanism to handle heterogeneous graph data, then implementing various relation scoring methods to alleviate inconsistency and discover camouflage. Finally, the neighbors’ features are aggregated together to build an informative representation. Experimental results on different types of real-world datasets demonstrate that FAHGT outperforms the state-of-the-art baselines.

**EXISTING SYSTEM**

ChebNet [14] and GCN [15] are proposed to improve efficiency by using approximation. For GNNs on spatial domain, GraphSAGE [16] samples a tree rooted at each node and computes the root’s hidden representation by hierarchically aggregating hidden node representations from the bottom to top. GAT [17] further proposes to learn in the spatial domain by computing different importance of neighbor nodes via the masked selfattention mechanism. All these methods are designed for homogeneous graphs. They cannot be directly applied to a heterogeneous graph with multiple types of entities and relations.

In recent years, lots of heterogeneous GNN based methods have been developed. HAN [18], HAHE [19], and Deep- HGNN [20] transforms a heterogeneous graph into several homogeneous graphs based on handcrafted meta-paths, applies GNN separately on each graph, and aggregates the output representations by attention mechanism. GraphInception [21] constructs meta-paths between nodes with the same object type. HetGNN [22] first samples a fixed number of neighbors via random walk strategy. Then it applies a hierarchical aggregation mechanism for intra-type and intertype aggregation. HGT [23] extends transformer architecture to heterogeneous graphs. They directly calculate attention scores for all the neighbors of a target node and perform aggregation accordingly without considering domain knowledge.

For relation-aware graph fraud detectors, their main solution is to build multiple homogeneous graphs based on edge type information of the original graph then perform type independent node level aggregation and graph level concatenation. GEM [9] learns weighting parameters for different homogeneous subgraph. Player2Vec [7] and SemiGNN [8] both adopt attention mechanism in feature aggregation and SemiGNN further leverages a structure loss to guarantee the node embeddings homophily. Some works directly aggregate heterogeneous information in the graph. For instance, under a user-review-item heterogeneous graph, GAS [3] learns a unique set of aggregators for different node types and updates the embeddings of each node type iteratively.

Disadvantages

* + In the existing work, the system did not implement Fraud Aware Heterogeneous Graph Transformer(FAHGT) to measure frauds exactly.
  + This system is less performance due to lack of META RELATION SCORING.

**PROPOSED SYSTEM**

* GraphConsis 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, the system introduces 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.

**Advantages**

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.

**SYSTEM REQUIREMENTS**

➢ **H/W System Configuration:-**

➢ Processor - Pentium –IV

➢ RAM - 4 GB (min)

➢ Hard Disk - 20 GB

➢ Key Board - Standard Windows Keyboard

➢ Mouse - Two or Three Button Mouse

➢ Monitor - SVGA

**SOFTWARE REQUIREMENTS:**

* **Operating system :** Windows 7 Ultimate.
* **Coding Language :** Python.
* **Front-End :** Python.
* **Back-End :** Django-ORM
* **Designing :** Html, css, javascript.
* **Data Base :** MySQL (WAMP Server).