FRAUD DETECTION IN ONLINE PRODUCT REVIEW SYSTEMS VIA HETEROGENEOUS GRAPH TRANSFORMER

A.NAGESHWAR RAO[1], E.TUNIKI RISHIKESH[2], SOHAIL KHAN[3], MOHD ARBAZ SHAREEF[4]

ASSISTANT PROFESSOR[1], UG SCHOLAR[2,3&4]

DEPARTMENT OF CSE, CMR INSTITUTE OF TECHNOLOGY, KANDLAKOYA VILLAGE, MEDCHAL RD, HYDERABAD, TELANGANA 501401

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.

INDEX TERMS- Online Product Review Systems, Fake Reviews,Fraud Detection,Graph-Based Methods,Heterogeneous Graph Data.

I.INTRODUCTION

The internet has revolutionized communication, commerce, and entertainment, but it has also created opportunities for fraudsters to disguise themselves as regular users to post spam or collect sensitive user information. The connections between various entities on the internet, each with different relationships, make detecting fraudulent activity challenging. Traditional machine learning methods often struggle to handle this complexity, especially in heterogeneous graph data. To address this, researchers have turned to graph neural networks (GNNs), which have been successfully applied in fraud detection across different domains like e-commerce, social platforms, and financial services. However, most GNN-based approaches do not fully account for the heterogeneous nature of the data and the camouflaging behavior of fraudsters, which can disguise their fraudulent activities by interacting with reputable entities or using sophisticated methods like generative language models.

To overcome these issues, the paper introduces the Fraud Aware Heterogeneous Graph Transformer (FAHGT), which incorporates heterogeneous mutual attention and a label-aware neighbor selector to address inconsistency and camouflage behaviors in fraud detection. FAHGT effectively handles multi-relation and multi-node types in heterogeneous graphs without needing manual meta-path design. By leveraging a novel "score head mechanism," it attentively selects neighbors for feature aggregation and fraud detection, improving the model's adaptability and efficiency. FAHGT also ensures scalability with a low computational complexity through a parallelizable multi-head mechanism. Experimental results on real-world datasets show that FAHGT outperforms existing state-of-the-art GNN-based fraud detection models, providing significant improvements in key metrics like KS and AUC. The system’s flexibility allows it to incorporate domain knowledge, enhancing its overall effectiveness in detecting fraud.

II.LITERATURE SURVEY

A**)** J. Wang, R. Wen, and C. Wu, “Fdgars: Fraudster detection via graph convolutional networks in online app review system,” in WWW Workshops, 2019

The paper "FDGARS: Fraudster Detection via Graph Convolutional Networks in Online App Review Systems" presents a method for detecting fraudsters in online app review systems by leveraging Graph Convolutional Networks (GCNs). Online app review systems are commonly targeted by fraudsters who manipulate reviews to mislead users or promote certain products. Traditional fraud detection techniques often fail to handle the complex relationships and interactions between users, apps, and reviews, which are inherently structured as graphs. The authors propose FDGARS, a fraud detection model that utilizes GCNs to effectively capture and exploit these relationships to detect fraudulent activities.

FDGARS constructs a heterogeneous graph that includes users, apps, and reviews as nodes, with edges representing interactions between them. The GCN-based model is designed to learn node embeddings that represent the characteristics of these entities, considering both direct and indirect relationships. By doing so, FDGARS can identify anomalous behaviors and flag suspicious users or reviews. The model is evaluated using real-world data, and the results show that FDGARS outperforms traditional fraud detection methods, demonstrating the potential of GCNs in improving the accuracy and effectiveness of fraud detection in online app review systems.

B)] A. Li, Z. Qin, R. Liu, Y. Yang, and D. Li, “Spam review detection with graph convolutional networks,” in CIKM, 2019.

The paper "Spam Review Detection with Graph Convolutional Networks" presents a novel approach for detecting spam reviews in online platforms by utilizing Graph Convolutional Networks (GCNs). Spam reviews, which are often fabricated or manipulated to deceive users and promote certain products, have become a significant problem for e-commerce and online review platforms. Traditional methods typically rely on feature engineering or supervised learning techniques, which can be insufficient in capturing the complex and interrelated structure of online review systems. The authors propose using GCNs to better model the relationships between users, reviews, and products as a heterogeneous graph. By leveraging the graph structure, the model can incorporate both local and global contextual information, allowing it to detect spam reviews with higher accuracy. The proposed method involves constructing a graph where nodes represent users, reviews, and products, and edges represent interactions, such as reviews written by users or products reviewed. Through graph convolution, the model learns node representations and identifies fraudulent reviews by detecting anomalies in these representations. The results of experiments conducted on real-world datasets show that the proposed approach outperforms traditional methods and significantly improves spam review detection performance.

C) Z. Liu, Y. Dou, P. S. Yu, Y. Deng, and H. Peng, “Alleviating the inconsistency problem of applying graph neural network to fraud detection,” in SIGIR, 2020.

The paper titled "Alleviating the Inconsistency Problem of Applying Graph Neural Network to Fraud Detection" addresses the challenges posed by the inconsistency problem in fraud detection when using Graph Neural Networks (GNNs). GNNs have been widely applied in fraud detection tasks due to their ability to model complex relationships in graph-structured data. However, traditional GNN-based methods face difficulties in handling inconsistencies in the graph, where nodes (such as users or transactions) with different attributes or behaviors are connected, leading to poor performance in detecting fraudulent activities.The authors propose a novel approach to mitigate the inconsistency problem by introducing a mechanism that improves the ability of GNNs to distinguish between legitimate and fraudulent activities. The proposed solution modifies the aggregation process in GNNs, allowing the model to more effectively handle nodes with distinct characteristics or behaviors. This improvement enhances the model’s capacity to detect fraud by ensuring that the graph's inherent inconsistencies do not negatively impact the learning process. Experimental results on real-world datasets show that the proposed method significantly outperforms existing GNN-based fraud detection techniques, demonstrating improved accuracy and robustness in identifying fraudulent activities.

IMPLEMENTATION

Modules
Service Provider

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Login, Browse Data Sets and Train & Test, View Trained and Tested Accuracy in Bar Chart, View Trained and Tested Accuracy Results, View All Antifraud Model for Internet Loan Prediction, Find Internet Loan Prediction Type Ratio, View Primary Stage Diabetic Prediction Ratio Results, Download Predicted Data Sets, View All Remote Users.

View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER AND LOGIN, PREDICT PRIMARY STAGE DIABETIC STATUS, VIEW YOUR PROFILE.

CONCLUSION

In this paper, we introduce FAHGT, a novel heterogeneous graph neural network designed to improve fraudulent user detection in online review systems. To address the challenge of inconsistent features across the graph, we incorporate heterogeneous mutual attention, which enables the automatic construction of meta paths. For detecting camouflage behaviors, we propose a label-aware scoring mechanism that effectively filters out noisy neighbors. These two neural modules are integrated into a unified framework known as the "score head mechanism," which contributes to the computation of edge weights during the final feature aggregation. Experimental results on real-world business datasets demonstrate the superior performance of FAHGT in fraud detection, showcasing its ability to alleviate inconsistencies and detect camouflage behaviors. Additionally, sensitivity analysis of hyperparameters and visual assessments confirm the stability and efficiency of our model. In conclusion, FAHGT sets a new benchmark in fraud detection, achieving state-of-the-art results in various scenarios. Future work will focus on adapting the model to handle dynamic graph data and exploring its application to other domains, such as robust item recommendation in e-commerce and loan default prediction in financial services**.**

REFERENCES

1. J. Wang, R. Wen, and C. Wu, "Fdgars: Fraudster detection via graph convolutional networks in online app review system," in WWW Workshops, 2019.
2. A. Li, Z. Qin, R. Liu, Y. Yang, and D. Li, "Spam review detection with graph convolutional networks," in CIKM, 2019.
3. Z. Liu, Y. Dou, P. S. Yu, Y. Deng, and H. Peng, "Alleviating the inconsistency problem of applying graph neural network to fraud detection," in SIGIR, 2020.
4. Y. Dou, Z. Liu, L. Sun, Y. Deng, H. Peng, and P. S. Yu, "Enhancing graph neural network-based fraud detectors against camouflaged fraudsters," in CIKM, 2020.
5. R. Wen, J. Wang, C. Wu, and J. Xiong, "Asa: Adversary situation awareness via heterogeneous graph convolutional networks," in WWW Workshops, 2020.

6. Y. Zhang, Y. Fan, Y. Ye, L. Zhao, and C. Shi, “Key player identification in underground forums over attributed heterogeneous information network embedding framework,” in CIKM, 2019.
7. D. Wang, J. Lin, P. Cui, Q. Jia, Z. Wang, Y. Fang, Q. Yu, and J. Zhou, “A semisupervised graph attentive network for fraud detection,” in ICDM, 2019.
8. Z. Liu, C. Chen, X. Yang, J. Zhou, X. Li, and L. Song, “Heterogeneous graph neural networks for malicious account detection,” in CIKM, 2018.
9. Y. Dou, G. Ma, P. S. Yu, and S. Xie, “Robust spammer detection by nash reinforcement learning,” in KDD, 2020.
10. P. Kaghazgaran, M. Alfifi, and J. Caverlee, “Wide-ranging review manipulation attacks: Model, empirical study, and countermeasures,” in CIKM, 2019.
11. Z. Zhang, P. Cui, andW. Zhu, “Deep learning on graphs: A survey,” TKDE,2020