Detecting Fraud With Oversampling Techniques and Sparsity Constraints

Prabina Pokharel
p2pokharel@ucsd.edu

Yandong (Dennis) Xiang
yaxiang@ucsd.edu

Jingyu Zhang
jiz036@ucsd.edu

Mentor: Gal Mishne
gmishne@ucsd.edu

Mentor: Yusu Wang
yusuwang@ucsd.edu

UC San Diego
HALICIOĞLU DATA SCIENCE INSTITUTE

MOTIVATION

Fraud detection is increasingly crucial due to the rise of online platforms, yet existing techniques often struggle to address imbalanced data in fraudulent activities. Our research aims to fill this gap by proposing a model that combines oversampling techniques and sparsity constraints to effectively detect and mitigate fraudulent activities, even in the face of imbalanced datasets.

DATASET

- Utilizes Amazon, Yelp, and Reddit fraudulent datasets

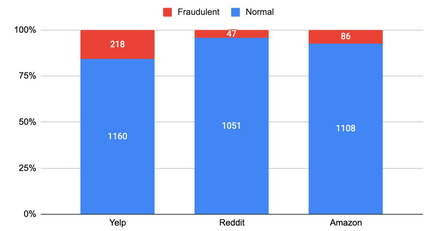
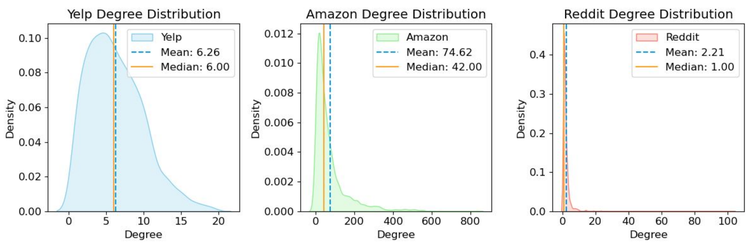


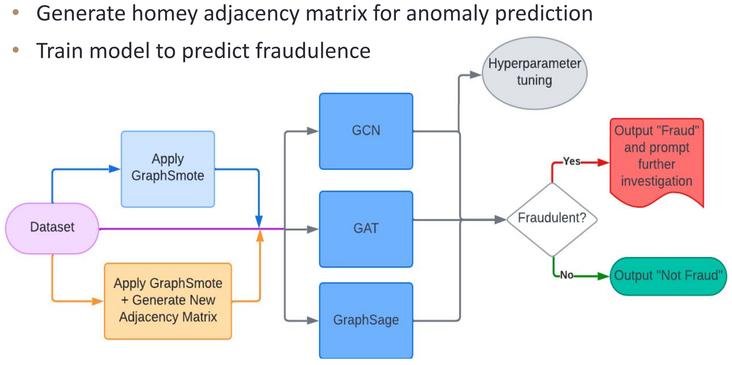
Figure represents imbalanced class structures, with Reddit having the highest imbalance (4.28%) and Yelp having the lowest (15.82%)



- There is a significant difference in degree distribution among the three datasets, with Reddit having the lowest and Amazon having the highest median number of degrees, causing different model performance on these three datasets

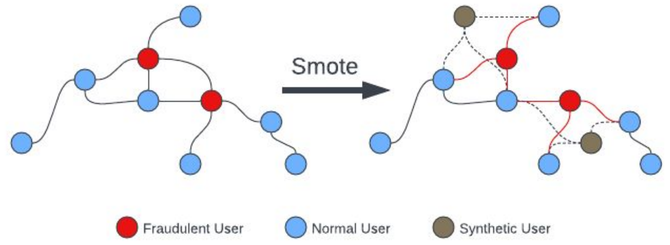
DATA WORKFLOW

- Input datasets from DGL (Deep Graph Library)
- Balance class distribution using GraphSmote
- Generate homey adjacency matrix for anomaly prediction
- Train model to predict fraudulence

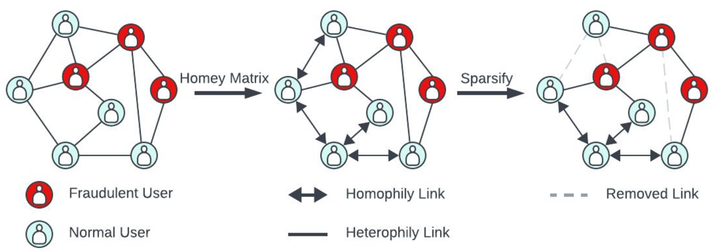


METHODS

- Combination of GraphSmote and (principles of) SparseGAD
- GraphSmote for oversampling minority class nodes
- SparseGAD for anomaly detection with sparsity constraints



- Generates synthetic nodes for minority class
- Upsampling method "replicates" nodes and connections
- Maintains heterogeneity in links
- Adds minor differences for distinction



- Fraudulent users tend to link with heterophilic users
- Adjusts adjacency matrix to separate homophily and heterophily links
- Filters unnecessary elements in homey adjacency matrix
- Uses threshold delta and KNN to remove unnecessary links

TRAINING

- Pretrain Model
  - Focuses on reconstructing the graph and learning general features
  - Loss is based on how well the model reconstructs the original graph
  - Updated the encoder
- Finetune Model
  - Loss is task-specific, focusing on classification accuracy
  - Both encoder and classifier optimizers are updated

RESULTS

Yelp

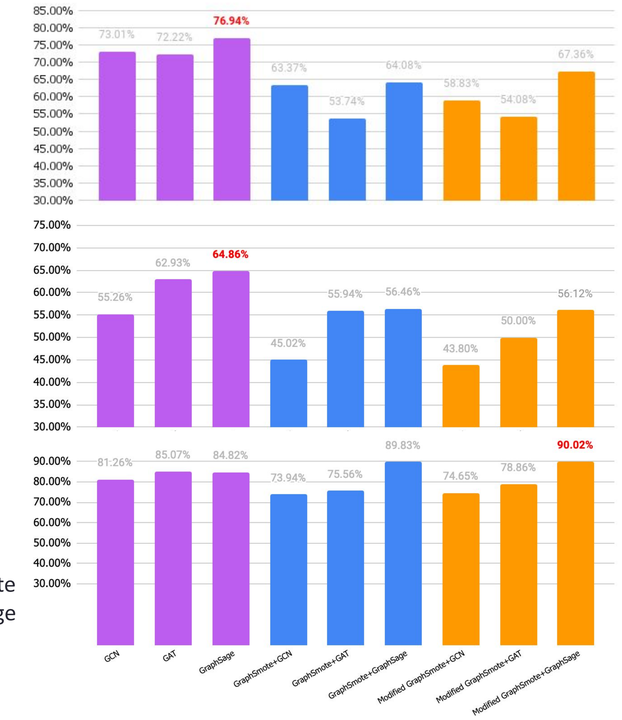
- Best result with GraphSage

Reddit

- Best result with GraphSage

Amazon

- Best result with Mod. GraphSmote + GraphSage



- GraphSage outperforms GCN and GAT because we find that similarities among a node and its neighbors is the most important standard to detecting fraud. GraphSage samples and aggregates features from a node's neighborhood to generate its representation
- Our improved model along with Graphsage performs the best on Amazon dataset for detecting the fraudulent users as expected
- In the Yelp and Reddit datasets, the median number of degrees is lower, and there are more nodes with only self-loops. This lack of connectivity limits the potential for synthetic nodes generated by the SMOTE method to establish meaningful connections within the graph. As a result, the ability of the SMOTE method to effectively balance the class distribution and enhance minority class representation may be constrained in these datasets

FUTURE WORK

- Refine the model to accommodate variability in datasets (degrees of connectivity and num. of connected nodes)
- Seek partnerships for deployment and cybersecurity collaboration

Website

Code

References