Grab

Simultaneously Detecting
Node-Level and Edge-Level Anomalies on Heterogeneous Attributed Graphs

Rizal Fathony*
Grab, Indonesia

Jenn Ng
Grab, Singapore

Jia Chen
Grab, Singapore

*Presenter

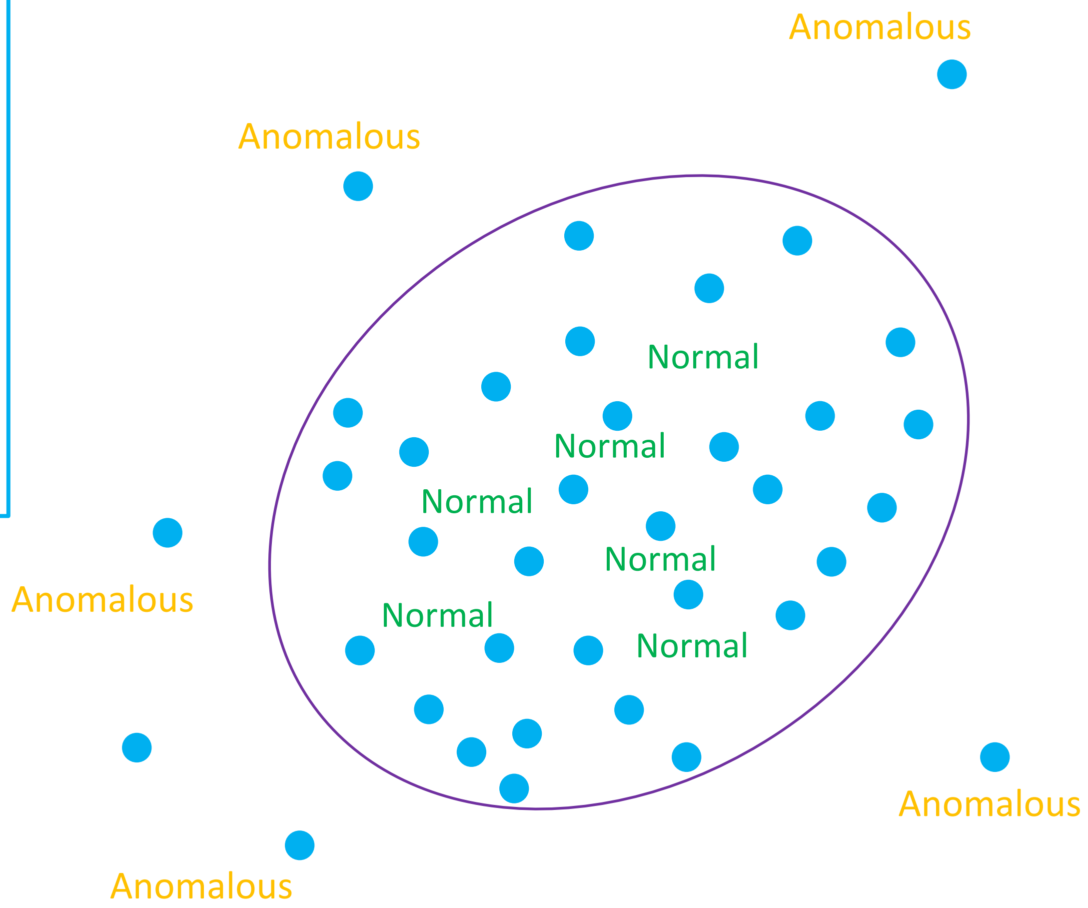
Introduction

Motivations of our study

What is Anomaly Detection?

→ Anomaly Detection

is the process of identifying unexpected observations in datasets, which deviate significantly from the majority of the data.



Why Anomaly Detection is Important?

Anomalous behaviors

Serious implications

Financial Institutions

Anomalous Transactions →

- Stolen Credit Cards
- Money Laundering

Computer Networks

Anomalous Traffic →

- Security Breach
- Network Intrusions

E-Commerce

Anomalous Purchase →

- Fraudulent Transactions
- Fake Reviews

Why **Unsupervised** Anomaly Detection?

Anomaly Detection (AD)

Usually done without label supervision
(unsupervised learning)

Label Availability Issues

- Anomaly events are rare.
- Labeling is sometimes expensive (domain expert are needed).

Adversarial Fraudsters

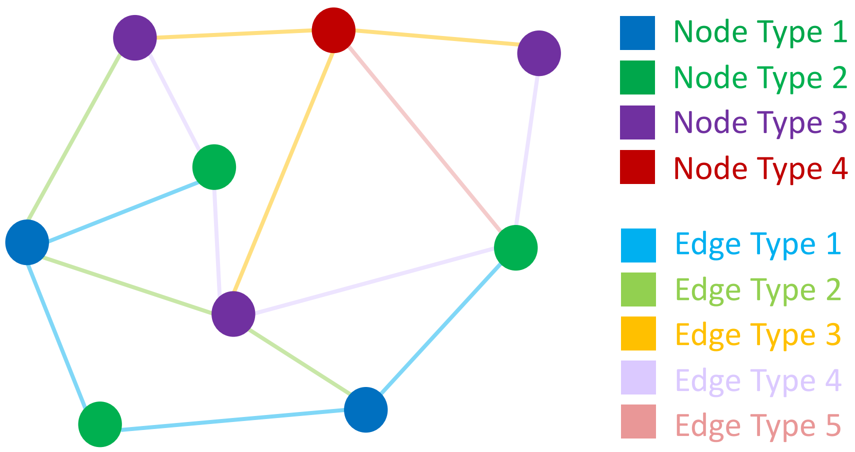
- Fraudsters are incentivized to adversarially innovate their methods of conducting fraud.
- Supervised learning model that rely on historical labels unable to detect new types of fraud.

Why Heterogeneous Graphs?

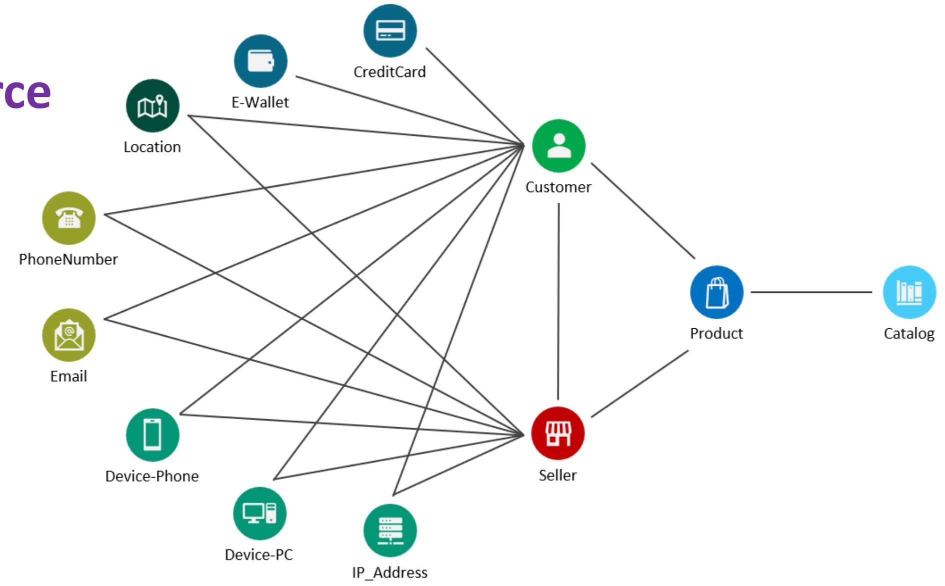
Heterogeneous Graphs

Main abstraction for modeling complex interactions among multiple groups.

Real-world, industrial level, interaction data:
complex and involve multiple entity types



E-commerce



Node Type:
- Customer
- Seller
- Product
- Catalog
- CreditCard
- Location
- Email
- etc…

Edge Type:
- Customer—Buy--Product
- Customer--BuyFrom--Seller
- Seller--Has--Product
- Customer--Use--CreditCard
- etc…

Social Media

Node Type:
- User
- Group
- Page
- Post

Edge Type:
- Like
- Comment
- Share
- Friendship

Why Node and Edge Attributes?

Real World Heterogeneous Graphs:

→ **Rich of information**

in both the entity (**node**)

and the interaction (**edge**)

→ **Example:**

e-commerce graph

customer node:
- costumer profile
- historical preference, etc.

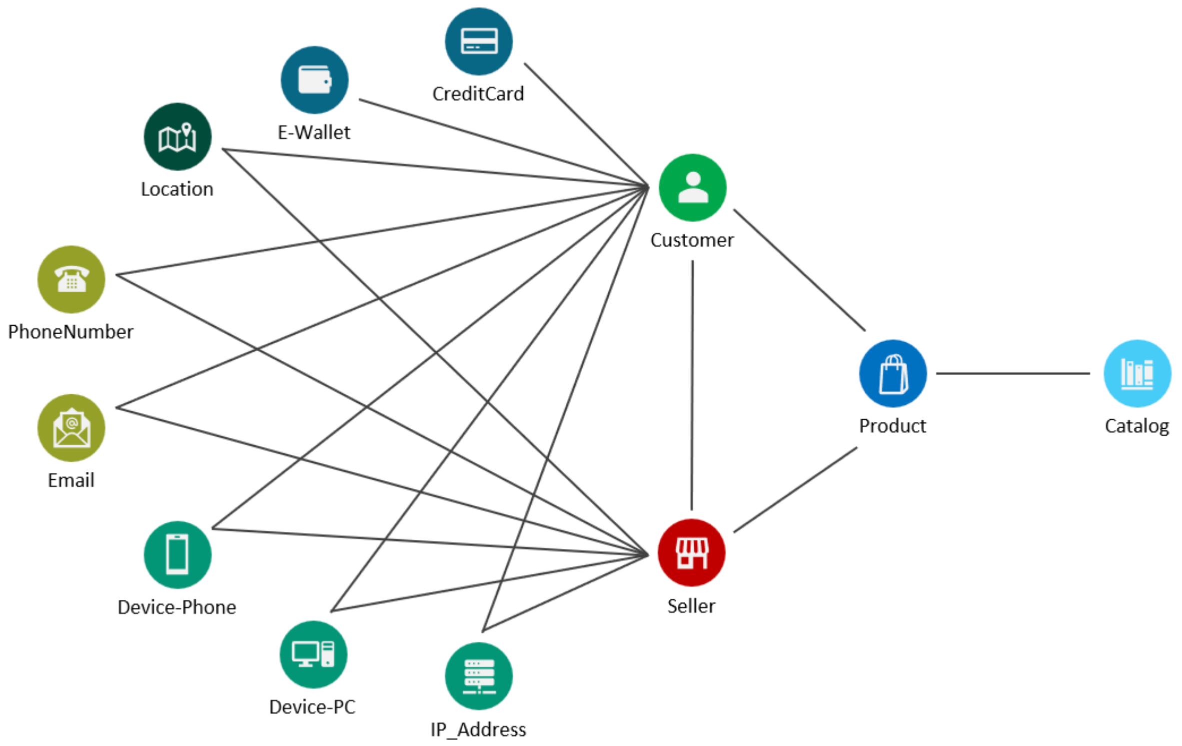
product node:
- product description
- product category, etc.

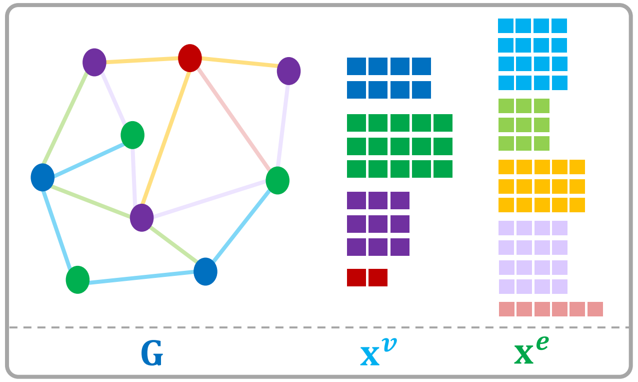
seller node:
- seller profile
- seller location
- previous customers stats, etc.

customer-buy-product:
- price paid
- payment method
- rating, review, etc.

customer-buyfrom-seller:
- how many products
- average prices, etc.

customer-use-creditcard:
- average transaction amounts
- transaction frequency, etc.





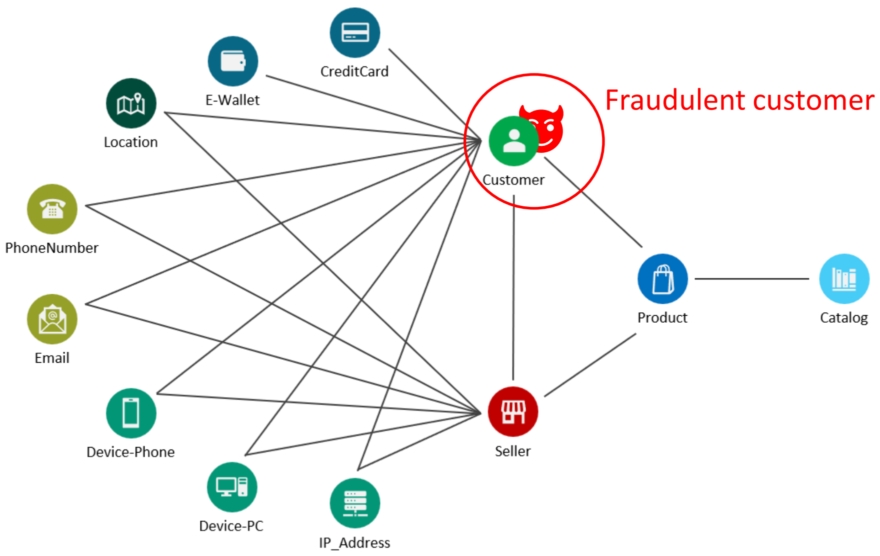
Why Node-Level and Edge-Level Anomaly Detection?

Node-Level Anomaly

Node-level anomaly suggests abnormal behavior from a specific entity.

➜ **Example:** *e-commerce graph*

A customer might start exhibiting unexpected activity that could indicate fraudulent behavior.

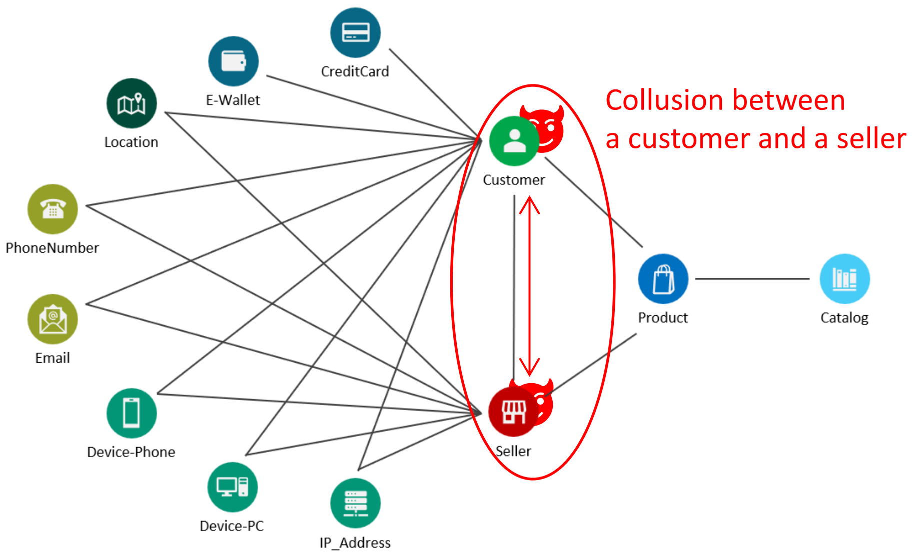


Edge-Level Anomaly

Edge-level anomaly indicates unusual interactions or relationships.

➜ **Example:** *e-commerce graph*

The frequency of interactions between a customer and a seller might change unexpectedly, suggesting **collaborative** fraudulent patterns, such as collusion.



Anomaly Detection on Heterogeneous Graphs

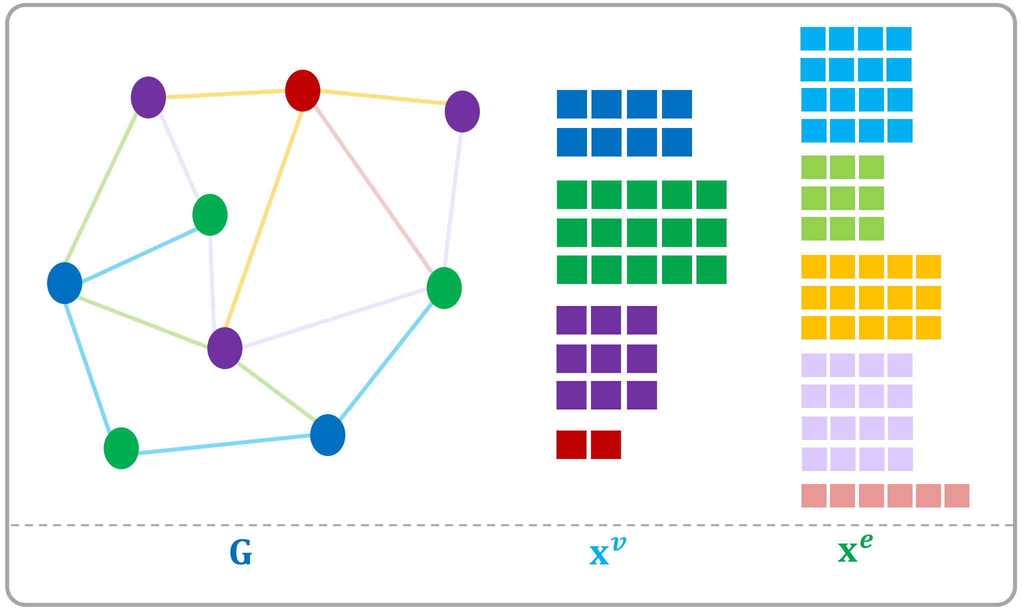
High-performing
anomaly detection system

→ All rich information (node and edge attributes) in the graph need to be considered by the model

→ The model need to be capable to detect **entity** (node-level) anomaly, and **interaction** (edge-level) anomaly

→ Our work!

Heterogeneous Attributed Graph



Related Works

Unsupervised Anomaly Detection
using Graph Neural Networks (GNN)

Anomaly Detection using Graph Neural Networks (GNN)

GNN models for unsupervised anomaly detection on attributed graphs

→ Motivated by the success of GNN architectures for supervised and semi-supervised learning

Method	Accept Node Features	Accept Edge Features	Node-Level Anomaly Detection	Edge-Level Anomaly Detection	Support Homogeneous Graphs	Support Bipartite Graphs	Support Heterogeneous Graphs
Reconstruction-based Homogeneous GNN - DOMINANT [Ding, et.al; 2019] - AnomalyDAE [Fan et al., 2020] - GAD-NR [Roy et al., 2023], etc...	Yes	No	Yes	No	Yes	No	No
Contrastive-based Homogeneous GNN - CoLA [Liu et al., 2021], - CONAD [Xu et al., 2022] - ANEMONE [Jin et al., 2021], etc…	Yes	No	Yes	No	Yes	No	No
GraphBEAN [Fathony et al., 2023]	Yes	Yes	Yes	Yes	No	Yes	No
AHEAD [Yang et al., 2022]	Yes	No	Yes	No	Yes	Yes	Yes
Our Method	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Our Approach

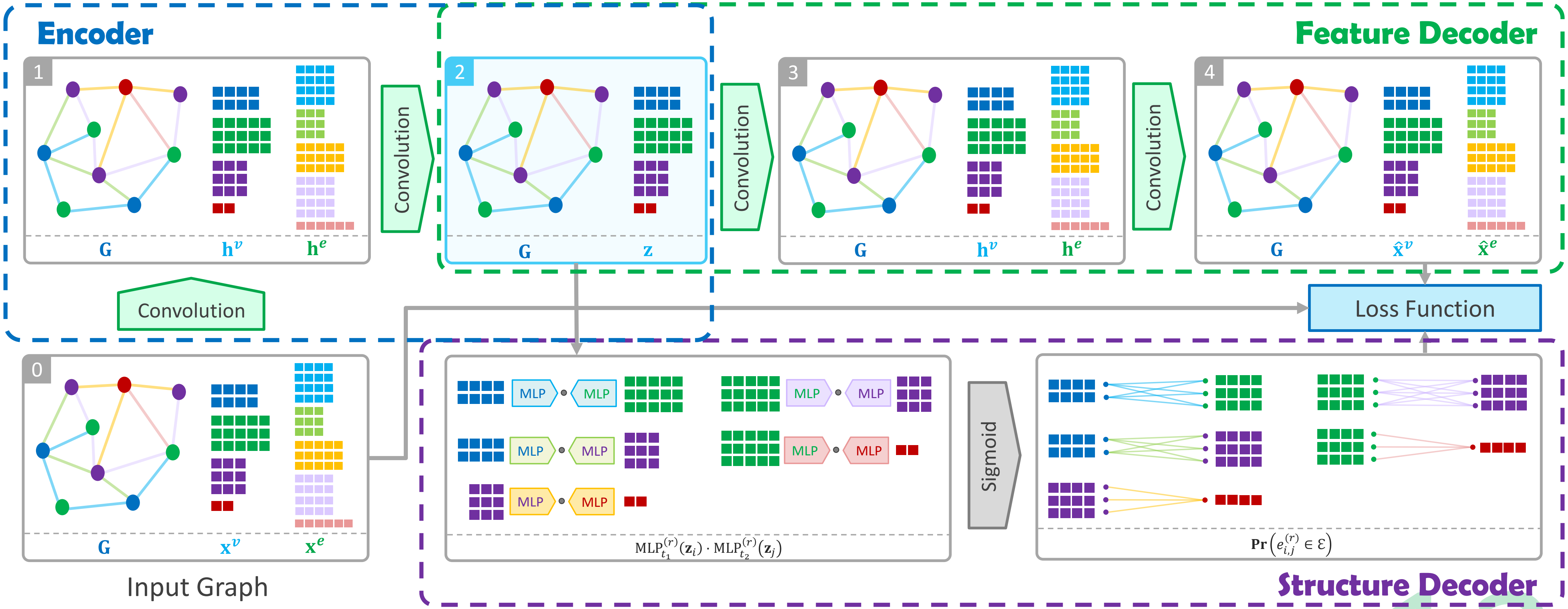
Our GNN architecture

Architecture Overview

HeagNet

Heterogeneous Node-and-**E**dge-**A**ttributed **G**raph Neural **Net**works

An autoencoder-like model

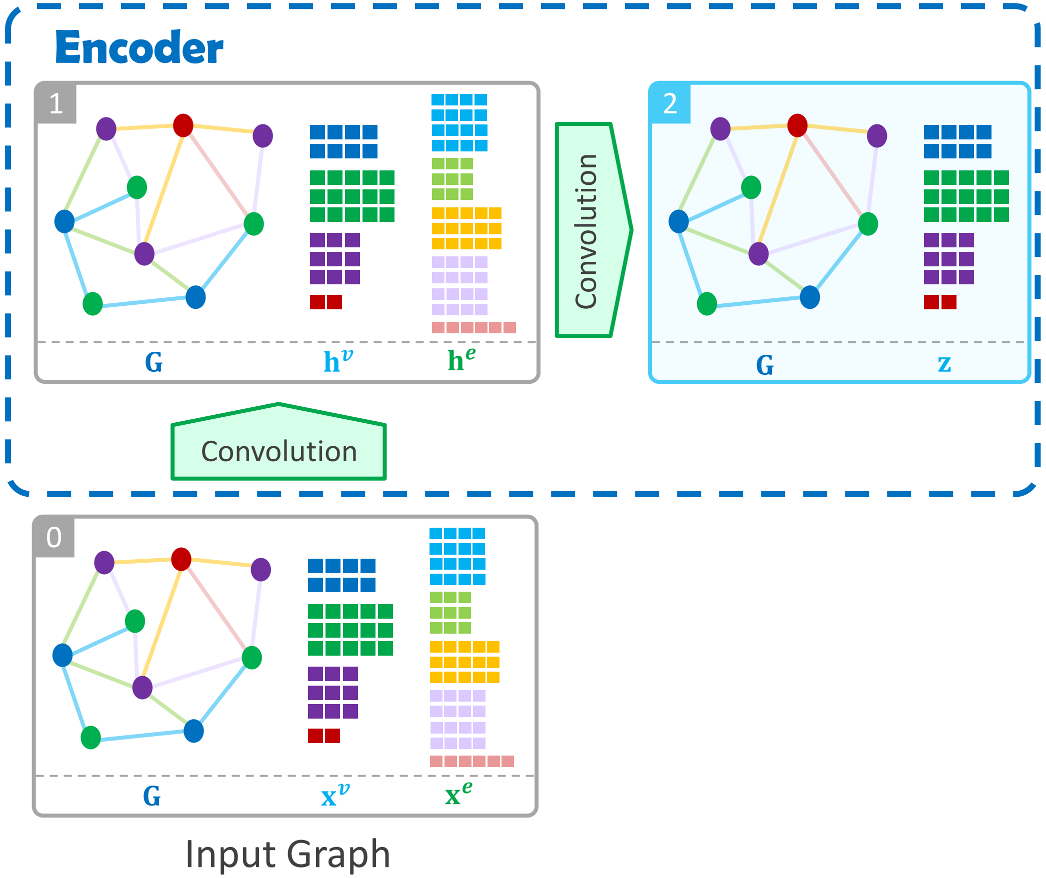


HeagNet with 4 layers

Encoder – Graph Convolution

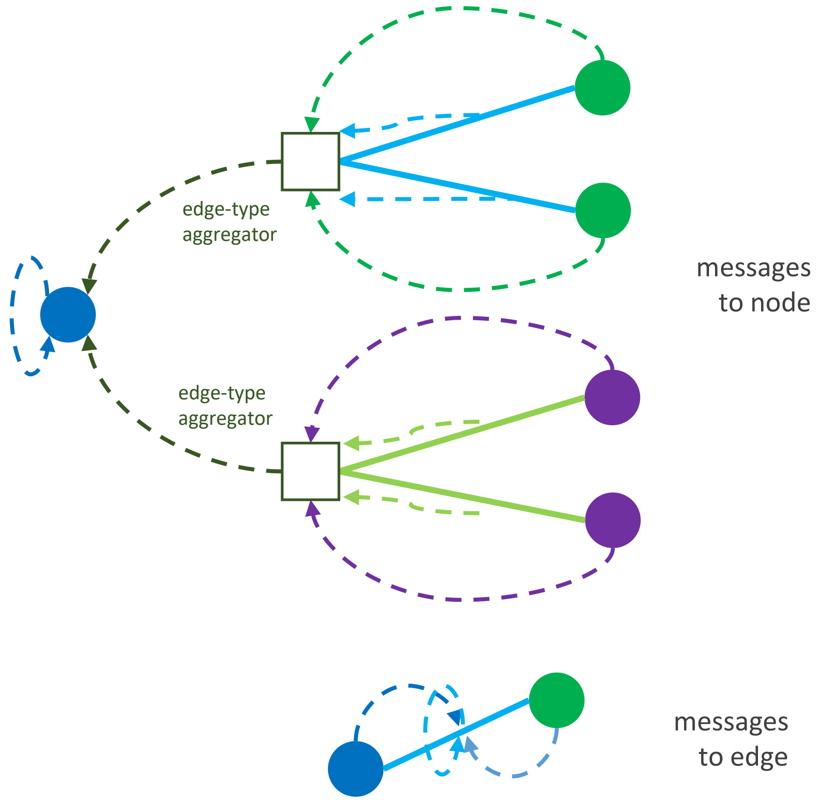
HeagNet

Heterogeneous Node-and-**E**dge-**A**ttributed **G**raph Neural **Net**works



Encoder

Message Passing Flow:



HeagNet-C

Simple aggregation (averaging) over all edge-types.

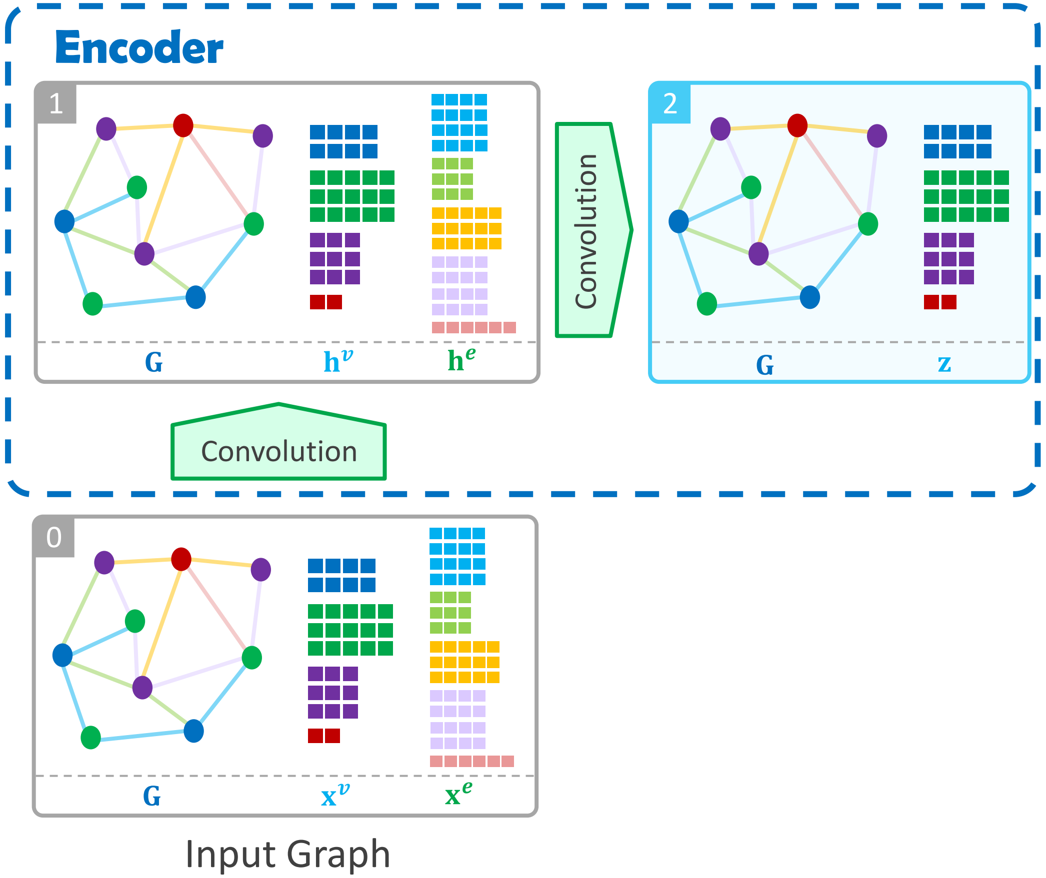
HeagNet-A

Attention based aggregation over all edge-types.

Encoder – Latent Representation

HeagNet

Heterogeneous Node-and-**E**dge-**A**ttributed **G**raph Neural **Net**works



Latent representation (node only)
∀ node types

No edge latent

Encoder

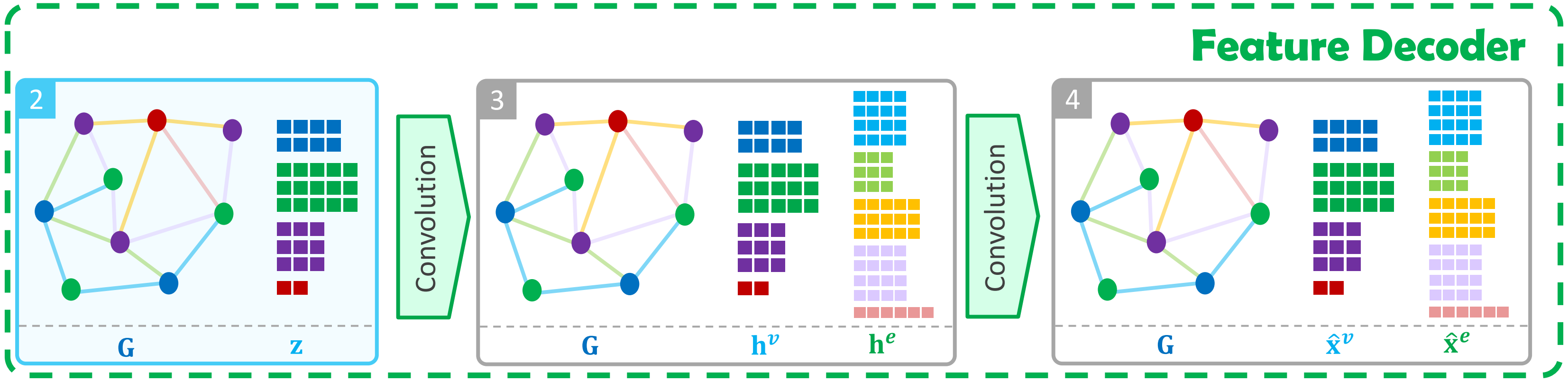
Feature Decoder

HeagNet

Heterogeneous Node-and-**E**dge-**A**ttributed **G**raph Neural **Net**works

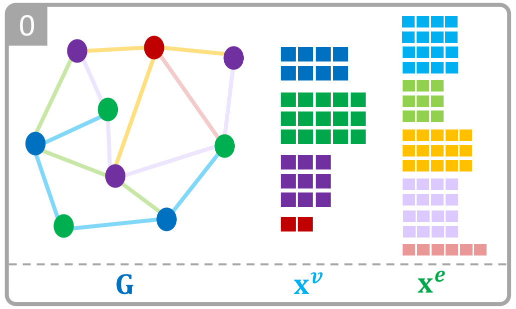
Latent representation
(node only)
∀ node types

No edge latent



Reconstructed graph with
node features and edge features

∀ node types
∀ edge types



Input Graph

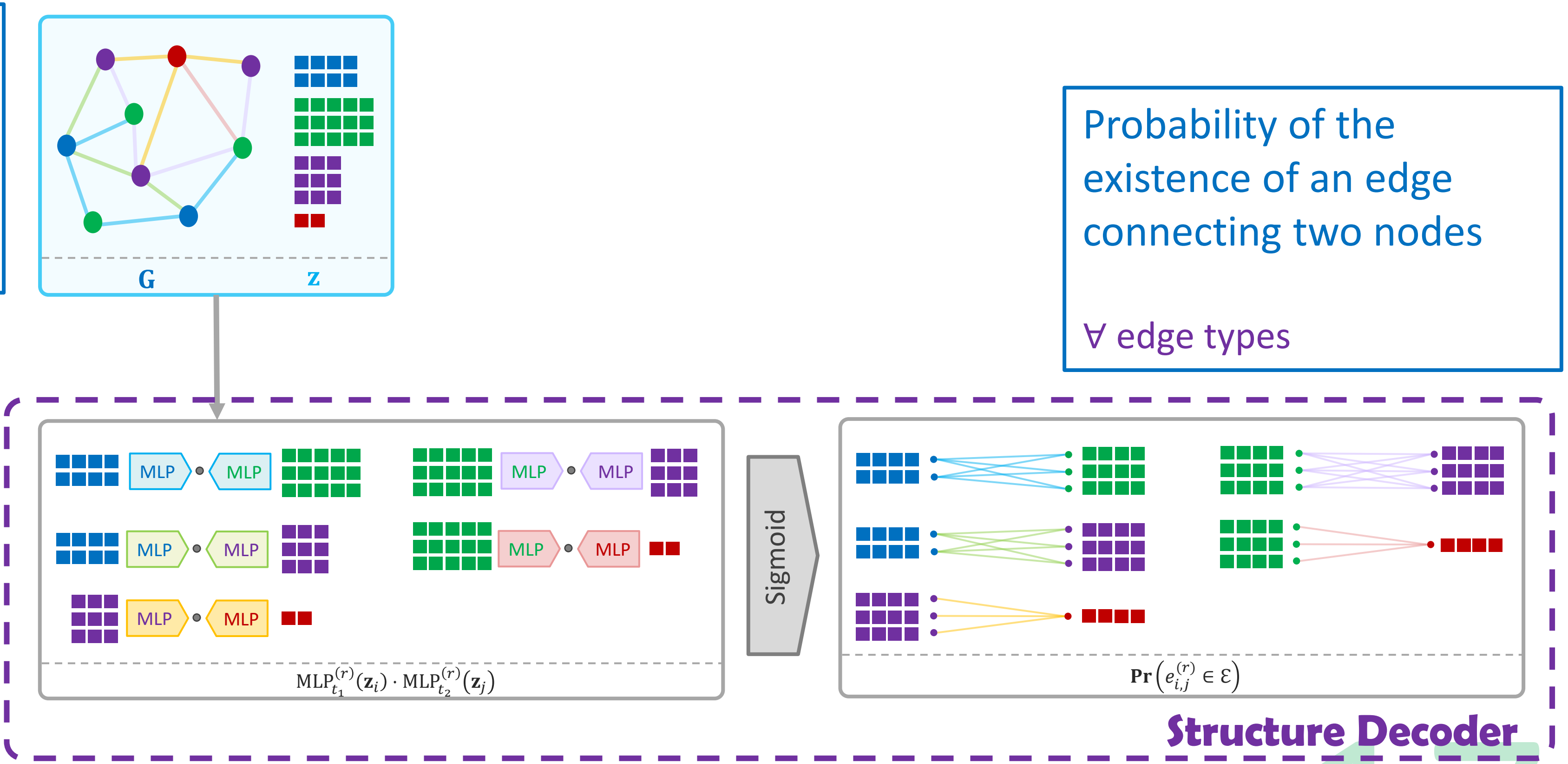
Structure Decoder

HeagNet

Heterogeneous Node-and-**E**dge-**A**ttributed **G**raph Neural **Net**works

Latent representation
(node only)
∀ node types

No edge latent



Probability of the existence of an edge connecting two nodes

∀ edge types

Structure Decoder

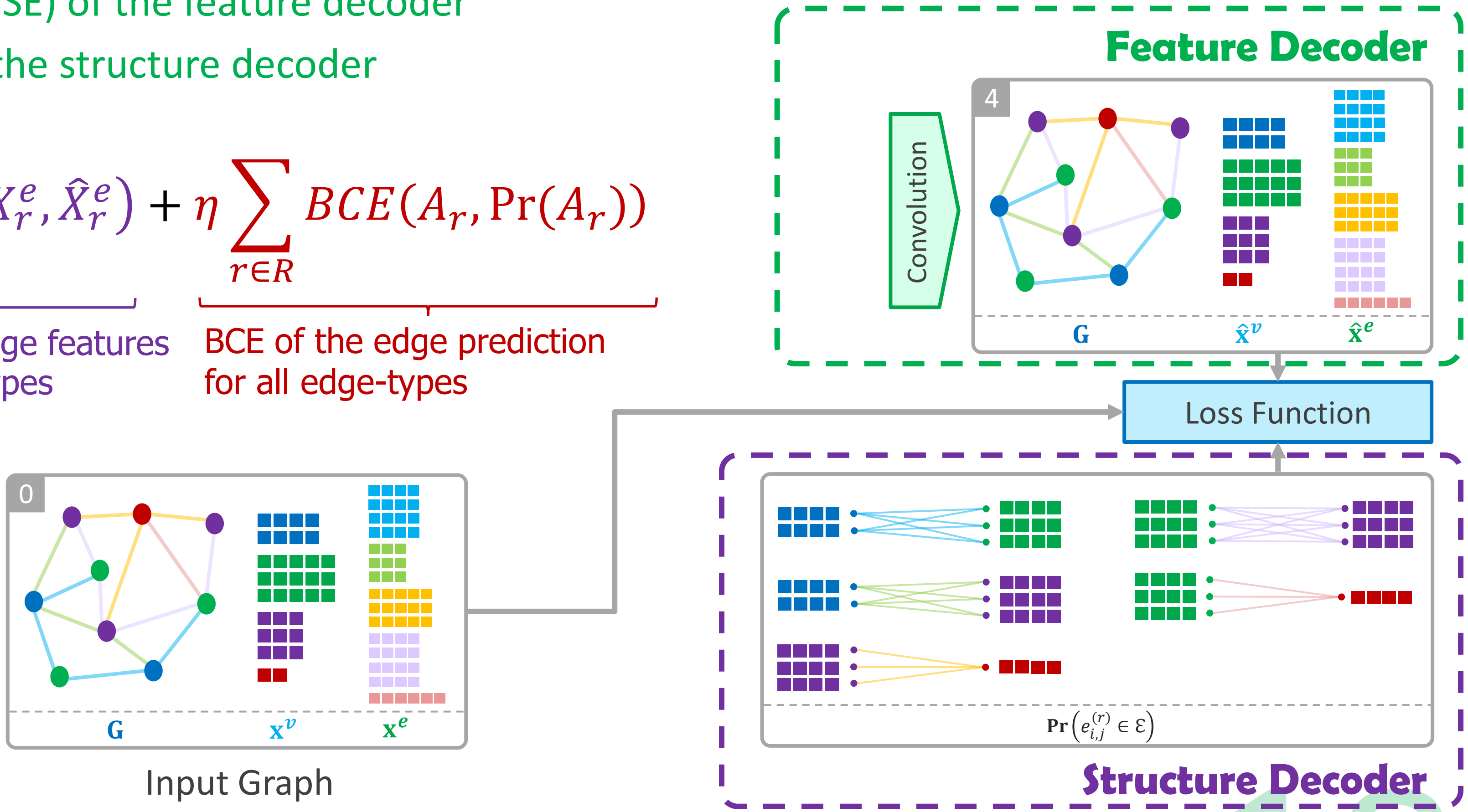
Loss Function

Our optimization objective:

→ Reconstruction Loss

- feature reconstruction error (MSE) of the feature decoder
- edge prediction error (BCE) of the structure decoder

$$\min \underbrace{\sum_{t\in T} MSE\left(X_t^v, \hat{X}_t^v\right)}_{\text{MSE of nodes features for all node-types}} + \underbrace{\sum_{r\in R} MSE\left(X_r^e, \hat{X}_r^e\right)}_{\text{MSE of the edge features for all edge-types}} + \underbrace{\eta \sum_{r\in R} BCE(A_r, \Pr(A_r))}_{\text{BCE of the edge prediction for all edge-types}}$$



Anomaly Score Construction

Reconstruction-based anomaly score

Normal behaviors : common ➜ can be easily reconstructed

Anomalous behaviors : rare ➜ cannot be reconstructed easily

$$\mathrm{score}_e^{(r)}(e_{i,j}^{(r)}) = \mathrm{MSE}\left(\mathbf{x}_{i,j}^{e^{(r)}}, \hat{\mathbf{x}}_{i,j}^{e^{(r)}}\right) + \eta \cdot \mathrm{BCE}\left(e_{i,j}^{(r)}\right)$$

edge type r

Edge-Level Anomaly Score

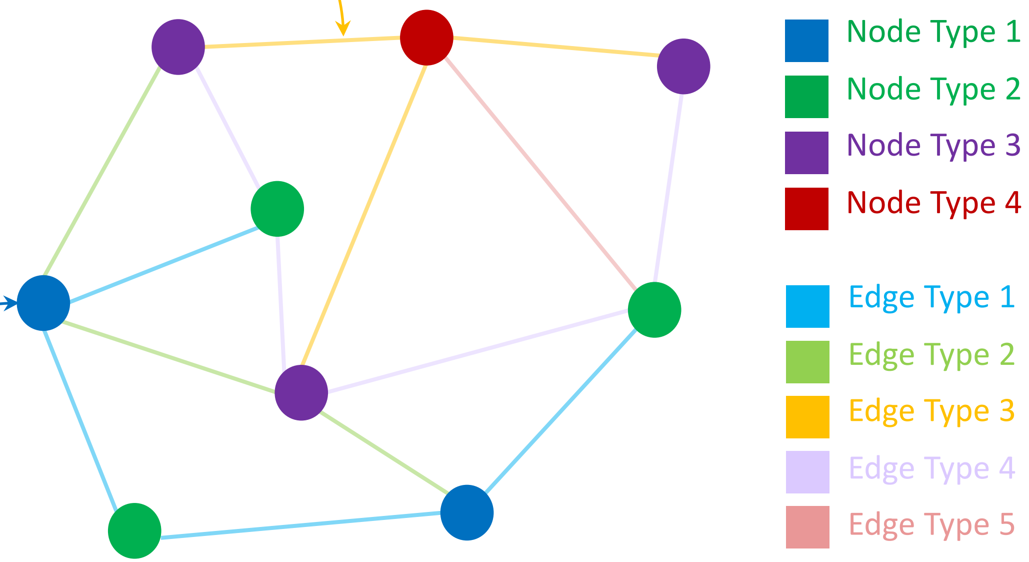
edge reconstruction error

Node-Level Anomaly Score

node features reconstruction error

+ aggregate over edge scores

in all edge types



Aggregate operator:
max or **mean**

$$\mathrm{score}_v^{(t)}(v_i) = \mathrm{MSE}\left(\mathbf{x}_i^{v^{(t)}}, \hat{\mathbf{x}}_i^{v^{(t)}}\right) + \underset{\substack{r \in \mathcal{R}(v_i) \\ e_{i,j}^{(r)} \in \mathcal{M}^{(r)}(v_i)}}{\mathrm{Agg}} \mathrm{score}_e^{(r)}(e_{i,j}^{(r)})$$

node type t

Experiments

Model evaluation

Datasets

Telecom
relationship of users & behaviors in a telecommunication network
node types: user, package, app, cell

Reddit
user interactions on the Reddit forum
node types: different groups of users and subreddits

Brightkite | Gowalla
user interactions on location-based social networks
node types: different groups of users and clusters of geohash locations

Datasets with various characteristics

Regular size datasets

Large size dataset

Anomaly ratio: 0.2% - 4.7%

Injection: - topological structure anomaly
- attributes anomaly

TABLE I: Dataset properties.

Dataset	#(node, edge) type	#node	#edge	avg deg.	avg (node, edge) dim.	node +ratio	edge +ratio
Telecom-Small	(4, 3)	80,380	890,000	11.1	(370, 50)	0.012	0.005
Reddit	(4, 4)	64,180	76,193	1.2	(384, 384)	0.010	0.011
Brightkite	(5, 4)	125,467	608,466	4.8	(10, 8)	0.026	0.047
Gowalla	(5, 4)	282,812	2,092,019	7.4	(10, 8)	0.018	0.012
Telecom-Large	(4, 3)	170,380	8,900,000	52.2	(370, 50)	0.017	0.002

Baselines & Evaluation Metric

BASELINES

Homogeneous GNN Models

DOMINANT (Ding et.al; 2019)

AnomalyDAE (Fan, et.al; 2020)

CONAD (Xu et.al; 2022)

Convert heterogeneous graphs into homogeneous graph

Edge anomaly score = average of the connected node scores

Features Only

Isolation Forest (Liu et.al; 2008)

Classical, non graph model

One model for each node type and each edge type

Heterogeneous GNN Model

AHEAD (Yang et.al; 2022)

node-level only *anomaly detection on heterogeneous graph*

An HGT-based model. It does not accept edge features

Edge anomaly score = average of the connected node scores

EVALUATION METRIC

Area under the Precision Recall Curve (AUC-PR)

Suitable for a very imbalance dataset
like in the anomaly detection task.

Overall Results

AUC-PR for node and edge anomaly detections

Edge-Level Anomaly: HeagNet significantly outperforms the baselines, often by a considerably large margin.

Node-Level Anomaly: HeagNet also maintain a relatively significant lead over all baselines.

HeagNet is the only model capable to utilize node and edge features and natively perform node-level and edge level detection

Large Size Datasets: HeagNets are **scalable** to the datasets, whereas most of the GNN baselines are not.

TABLE II: The mean (and stdev.) of the Average-AUCPR metrics over multiple experiment runs in each dataset.

Model / Dataset	IsoForest Node	IsoForest Edge	DOMINANT Node	DOMINANT Edge	AnomalyDAE Node	AnomalyDAE Edge	CONAD Node	CONAD Edge	AHEAD Node	AHEAD Edge	HeagNet-C Node	HeagNet-C Edge	HeagNet-A Node	HeagNet-A Edge
Telecom-Small	0.924 (0.06)	0.556 (0.04)	0.428 (0.05)	0.196 (0.06)	0.132 (0.04)	0.036 (0.04)	0.427 (0.05)	0.196 (0.06)	0.942 (0.05)	0.597 (0.06)	**0.970** (0.02)	**0.715** (0.07)	0.965 (0.04)	0.711 (0.07)
Reddit	0.955 (0.05)	0.770 (0.17)	0.644 (0.15)	0.705 (0.08)	0.533 (0.10)	0.567 (0.07)	0.644 (0.15)	0.705 (0.08)	0.949 (0.05)	0.545 (0.09)	**0.968** (0.03)	0.788 (0.17)	0.963 (0.02)	**0.791** (0.15)
Brightkite	0.893 (0.07)	0.547 (0.15)	0.731 (0.12)	0.569 (0.05)	0.185 (0.04)	0.235 (0.09)	0.719 (0.12)	0.568 (0.05)	0.616 (0.11)	0.202 (0.09)	**0.928** (0.05)	**0.590** (0.12)	0.907 (0.05)	0.534 (0.10)
Gowalla	0.845 (0.05)	0.246 (0.06)	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	**0.952** (0.03)	**0.445** (0.09)	0.930 (0.03)	0.310 (0.07)
Telecom-Large	0.945 (0.07)	0.493 (0.10)	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	**0.964** (0.05)	**0.642** (0.11)	0.961 (0.05)	0.616 (0.10)

Overall Results

AUC-PR for node and edge anomaly detections

Edge-Level Anomaly: HeagNet significantly outperforms the baselines, often by a considerably large margin.

Node-Level Anomaly: HeagNet also maintain a relatively significant lead over all baselines.

HeagNet is the only model capable to utilize node and edge features and natively perform node-level and edge-level detection

Large Size Datasets: HeagNets are **scalable** to the datasets, whereas most of the GNN baselines are not.

Attention mechanism in HeagNet: HeagNet-C perform better than HeagNet-A, indicating that attention mechanism does not contribute much to the model performance.

TABLE II: The mean (and stdev.) of the Average-AUCPR metrics over multiple experiment runs in each dataset.

Model / Dataset	IsoForest Node	IsoForest Edge	DOMINANT Node	DOMINANT Edge	AnomalyDAE Node	AnomalyDAE Edge	CONAD Node	CONAD Edge	AHEAD Node	AHEAD Edge	HeagNet-C Node	HeagNet-C Edge	HeagNet-A Node	HeagNet-A Edge
Telecom-Small	0.924 (0.06)	0.556 (0.04)	0.428 (0.05)	0.196 (0.06)	0.132 (0.04)	0.036 (0.04)	0.427 (0.05)	0.196 (0.06)	0.942 (0.05)	0.597 (0.06)	**0.970** (0.02)	**0.715** (0.07)	0.965 (0.04)	0.711 (0.07)
Reddit	0.955 (0.05)	0.770 (0.17)	0.644 (0.15)	0.705 (0.08)	0.533 (0.10)	0.567 (0.07)	0.644 (0.15)	0.705 (0.08)	0.949 (0.05)	0.545 (0.09)	**0.968** (0.03)	0.788 (0.17)	0.963 (0.02)	**0.791** (0.15)
Brightkite	0.893 (0.07)	0.547 (0.15)	0.731 (0.12)	0.569 (0.05)	0.185 (0.04)	0.235 (0.09)	0.719 (0.12)	0.568 (0.05)	0.616 (0.11)	0.202 (0.09)	**0.928** (0.05)	**0.590** (0.12)	0.907 (0.05)	0.534 (0.10)
Gowalla	0.845 (0.05)	0.246 (0.06)	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	**0.952** (0.03)	**0.445** (0.09)	0.930 (0.03)	0.310 (0.07)
Telecom-Large	0.945 (0.07)	0.493 (0.10)	OOM	OOM	OOM	OOM	OOM	OOM	OOM	OOM	**0.964** (0.05)	**0.642** (0.11)	0.961 (0.05)	0.616 (0.10)

Precision-Recall Curve

Precision Recall Trade-off
at any given point in PR Curve

at almost all thresholding points **HeagNet-C** outperforms all the baselines, sometimes by a **significant** margin.

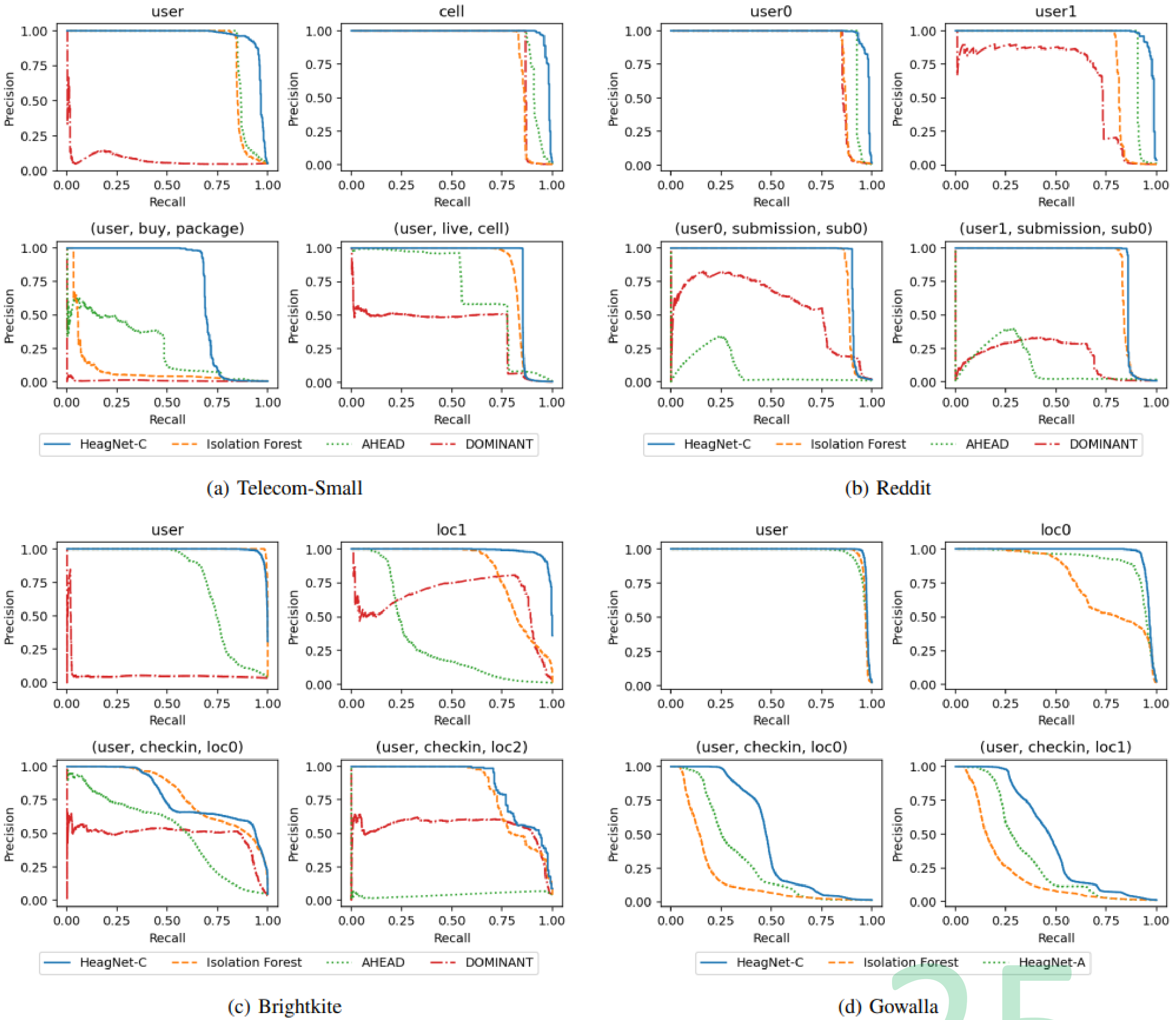


Fig. 3: Precision-recall curves of the node-level and edge-level anomaly detection on each dataset.

Conclusions

Conclusions and Remarks

Conclusions

Heterogeneous Graphs

All available information need to be considered

to build a high-performing anomaly detection model

Our proposed model

HeagNet is effective in detecting

both node-level and edge-level anomalies

on heterogeneous graphs

Grab

Open-Source Implementation

https://github.com/grab/HeagNet/

PyTorch

PyG

Thank You