Association-Focused Path Aggregation for Graph Fraud Detection

Tian Qiu, Wenda Li, Zunlei Feng, Jie Lei, Tao Wang,
Yi Gao, Mingli Song, Yang Gao
https://github.com/horrible-dong/GPA

ZHEJIANG UNIVERSITY

NEURAL INFORMATION PROCESSING SYSTEMS

Highlights

- This paper proposes a novel fraud detector based on Graph Path Aggregation (GPA), which goes beyond the local view of neighbors and explores hidden and complex fraud associations at the path level.
- Extensive experiments across datasets in multiple fraud scenarios demonstrate that the proposed GPA outperforms mainstream fraud detectors by up to +15% in Average Precision (AP).

Motivation

Existing graph fraud detectors are limited by their narrow "receptive fields", as they focus only on the relations between an entity and its neighbors while neglecting longer-range structural associations hidden between entities.

Therefore, it is imperative to develop a global-aware technique that can effectively capture long-range structural dependencies in graphs to unveil hidden fraudulent patterns within massive data.

Comparison

Table 1: Model performance comparison in various fraud scenarios. "/" denotes "out of memory".

Method	Scenario	Internet		Finance		Social		Reviews			
	Dataset	G-Internet		Elliptic		T-Finance		T-Social		YelpChi		Amazon	
	Metric	AUC	AP	AUC	AP	AUC	AP	AUC	AP	AUC	AP	AUC	AP
Non-GNN	MLP	61.0	4.7	88.3	43.8	92.2	74.2	73.1	9.7	81.6	47.7	96.9	87.3
	KNN	51.7	3.0	88.0	61.0	92.7	75.0	77.6	36.3	84.6	54.4	94.9	84.4
Standard	GCN	98.5	81.9	81.7	25.4	94.6	78.2	96.6	76.4	58.6	20.9	85.2	45.7
	GraphSAGE	98.6	78.9	87.6	57.8	95.6	84.7	95.7	75.3	82.9	46.6	90.9	82.5
	GAT	94.2	57.9	86.3	27.5	95.8	82.7	90.3	32.1	79.1	43.6	97.1	87.9
Spectral	BernNet	91.5	57.6	87.5	38.3	96.7	89.2	93.7	44.3	83.5	51.9	95.8	84.9
	AMNet	90.8	56.0	88.9	69.5	96.4	88.9	92.5	37.7	81.9	46.9	96.9	88.4
	BWGNN	95.8	72.7	89.6	48.4	96.9	89.4	96.9	78.9	87.1	61.5	98.3	91.5
	GHRN	94.3	66.1	90.0	55.2	96.5	87.6	97.1	86.8	84.6	55.4	**98.3**	89.5
	AHFAN	87.6	72.8	87.5	72.7	86.8	73.3	87.1	73.5	86.8	72.7	98.2	88.2
Spatial	GAS	96.7	72.9	86.7	29.8	96.5	86.0	95.0	62.4	76.4	35.1	92.7	81.4
	DCI	82.4	23.6	85.7	27.4	87.9	63.7	84.0	13.0	78.4	39.9	95.4	85.2
	CARE-GNN	70.2	15.5	87.8	37.2	90.0	61.8	78.3	41.2	84.0	53.0	97.0	85.6
	PC-GNN	79.6	22.4	86.5	42.7	94.0	83.3	96.9	80.3	80.8	44.5	98.0	89.3
	RioGNN	72.8	15.1	86.4	29.1	91.3	62.6	81.7	17.6	85.7	56.5	96.9	87.6
	H^2-FDetector	83.7	27.9	63.2	10.5	/	/	/	/	89.9	57.5	96.1	84.9
	GDN	90.7	45.7	88.7	65.2	95.8	85.7	88.4	52.3	90.4	67.4	97.3	86.8
	GFCN	85.9	43.4	85.8	45.5	92.6	82.1	86.1	47.4	87.6	68.4	95.5	83.8
	HedGe	91.2	49.7	88.9	64.5	96.5	89.0	96.8	83.7	91.3	70.7	**98.3**	92.3
Path	GPA	**99.8**	**97.6**	**91.3**	**76.1**	**97.3**	**89.6**	**99.6**	**96.0**	**91.8**	**73.7**	98.1	**92.5**

Method

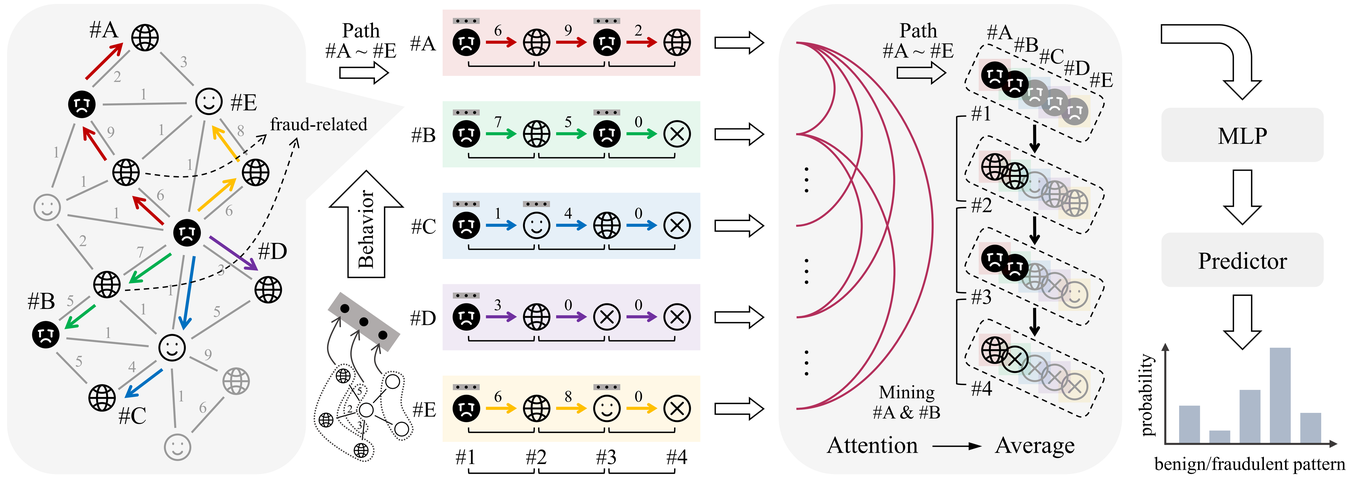

Figure 1: An illustration of the proposed fraud detector based on Graph Path Aggregation (GPA) under the internet fraud scenario.

Interpretability

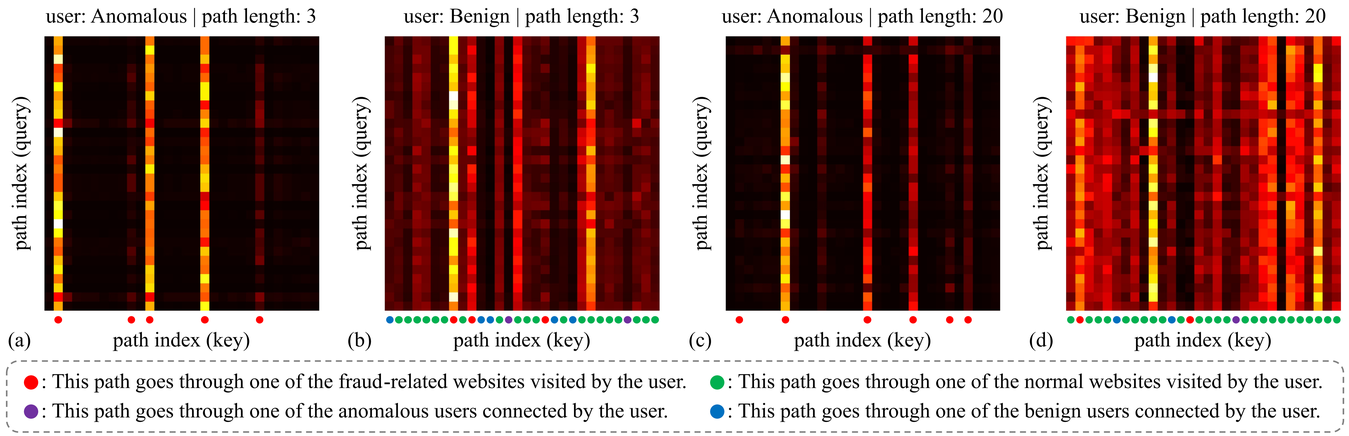

Figure 2: Attention map between all paths starting from a user in G-Internet during path interaction.