Enhancing Graph Neural Network-based Fraud Detectors against Camouflaged Fraudsters

Yingtong Dou (UIC), Zhiwei Liu (UIC), Li Sun (BUPT),
Yutong Deng (BUPT), Hao Peng (BUAA), Philip S. Yu (UIC)

Email: ydou5@uic.edu
Homepage: http://ytongdou.com
Project Page: https://github.com/safe-graph
Paper: https://arxiv.org/abs/2008.08692
Code: https://github.com/YingtongDou/CARE-GNN

Paper Highlight

- Comprehensive **review** of GNN-based fraud detection research.
- Introduce and summarize two **fraudster camouflaging** behaviors in the wild.
- Propose **CARE-GNN** which is efficient and adaptive to many scenarios.
- **Opensource** model code, baseline code, and new dataset.

A History of Fraud

- 1990-2000: spam email, link farming.
- 2000-2010: fake review, social bots.
- 2010-2020: fake news, deepfake.

Handcrafted
&
Human

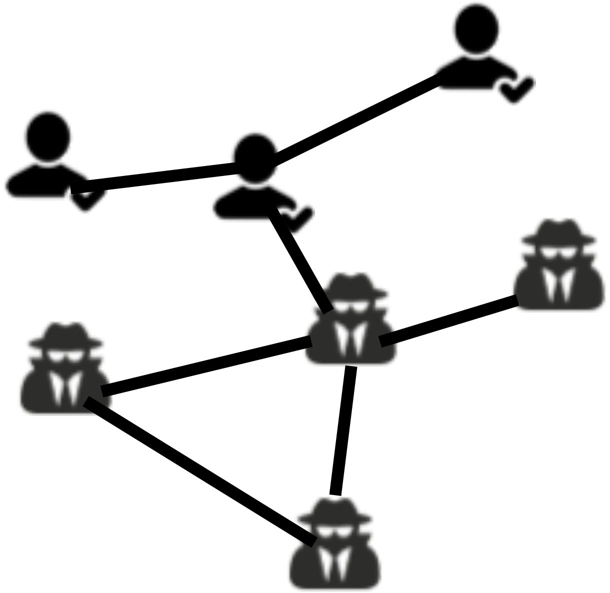
Automatic
&
Machine Learning

Ferrara, Emilio. "The history of digital spam." *Communications of the ACM* 62, no. 8 (2019): 82-91.

Graph-based Fraud Detection

Fraudster

Benign User



Graphical Model (e.g., MRF)[1].

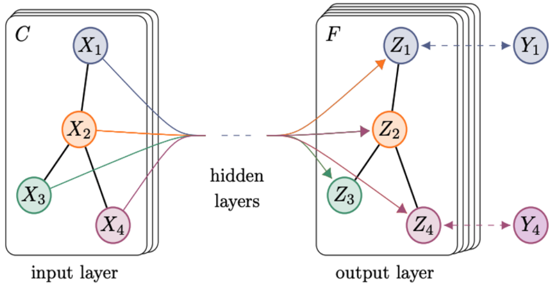
Structure-based suspicious estimation[2].

Dimension reduction (e.g., SVD) [3].

[1] Rayana, Shebuti, and Leman Akoglu. "Collective opinion spam detection: Bridging review networks and metadata." KDD. 2015.
[2] Hooi, Bryan, et al. "Fraudar: Bounding graph fraud in the face of camouflage." KDD. 2016.
[3] Shah, Neil, et al. "Spotting suspicious link behavior with fbox: An adversarial perspective." ICDM, 2014..

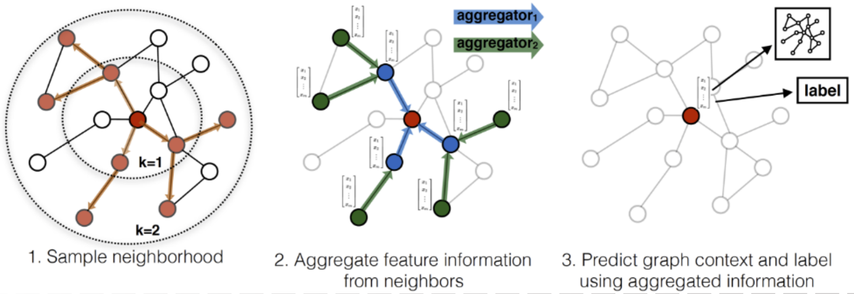
Graph Neural Network

GCN[1]



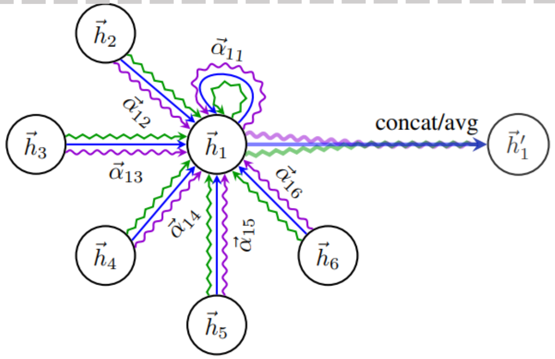
- Directly aggregate neighbors using Laplacian adjacency matrix.

GraphSAGE[2]



- Sample and aggregate neighbors.

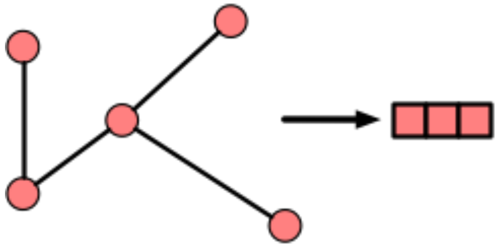
GAT[3]



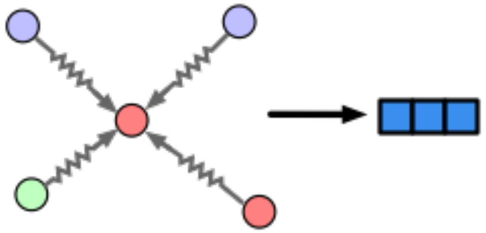
- Attentively aggregate neighbors.

[1] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[J]. arXiv preprint arXiv:1609.02907, 2016.
[2] W. Hamilton, Hamilton, William L. Ying, Rex Leskovec, Jure. Inductive Representation Learning on Large Graphs , NIPS 2017
[3] Veličković P, Cucurull G, Casanova A, et al. Graph attention networks[J]. arXiv preprint arXiv:1710.10903, 2017.

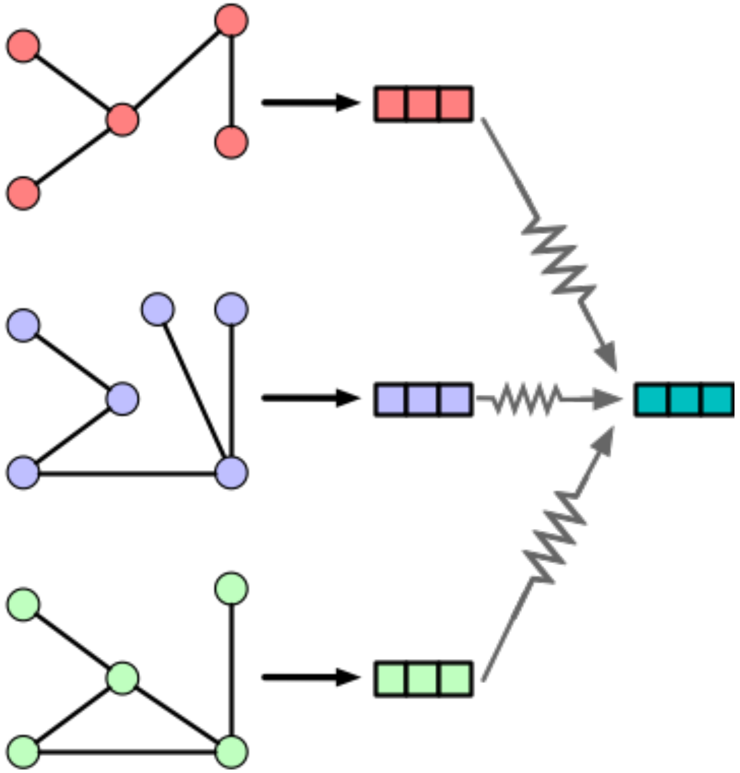
GNN-based Fraud Detectors



FdGars[1] (GCN-based)



GAS[2] (GAT-based)



Player2Vec[3] (Hybrid)

[1] Wang, J., Wen, R., Wu, C., Huang, Y. and Xion, J., 2019, May. Fdgars: Fraudster detection via graph convolutional networks in online app review system. WWW 2019.
[2] Li, A., Qin, Z., Liu, R., Yang, Y. and Li, D., 2019, November. Spam review detection with graph convolutional networks. CIKM 2019.
[3] Zhang, Y et, al. November. Key Player Identification in Underground Forums over Attributed Heterogeneous Information Network Embedding Framework. CIKM 2019

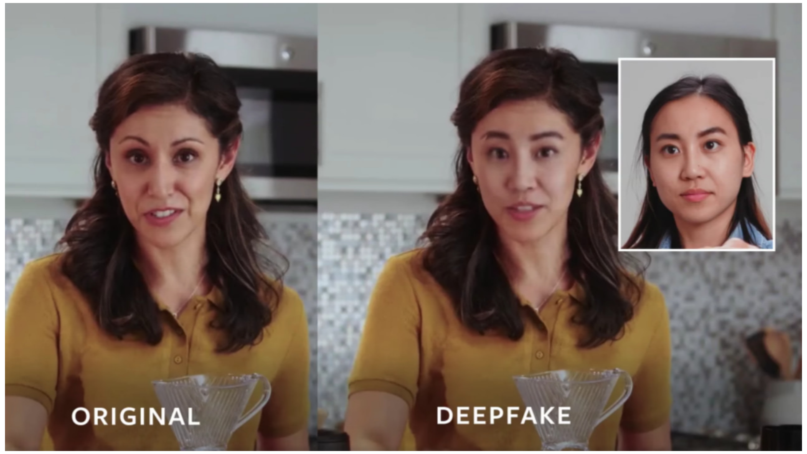
Camouflaging Behavior of Fraudsters

• Feature Camouflage

Spamouflage

Source: @benimmo

Deepfake



Source: https://elgan.com/blog/deepfakes-get-real-and-real-easy

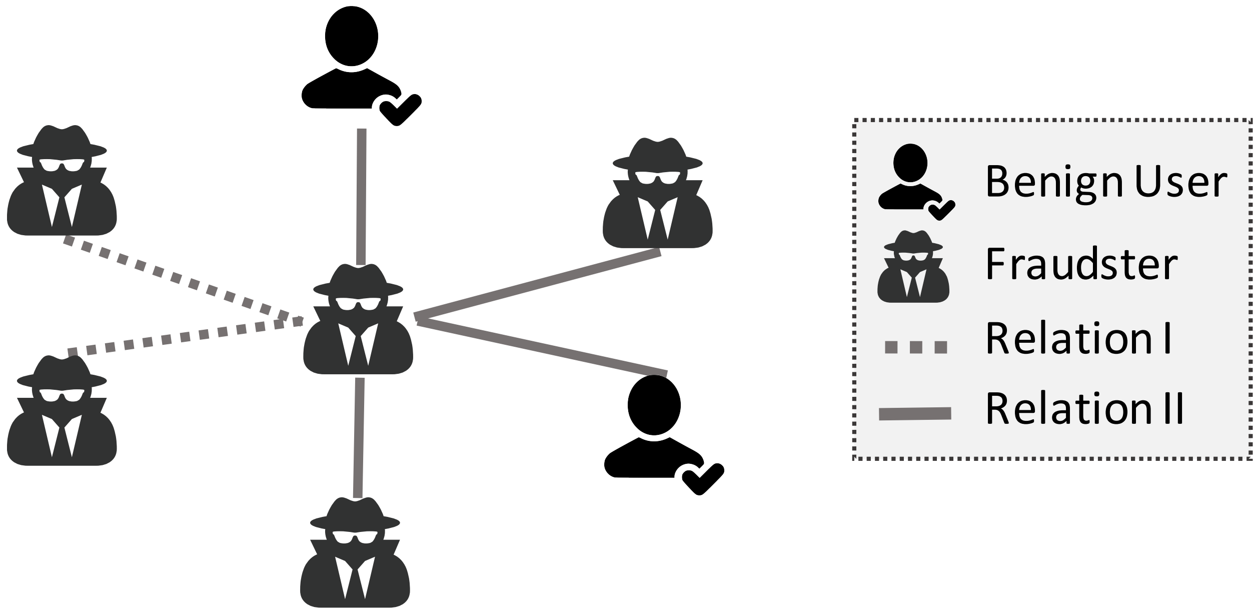
Language generation model

Generated Reviews (Yelp)
I love this place ! I 've been here several times and I 've never been disappointed . The food is always fresh and delicious . The service is always friendly and attentive . I 've been here several times and have never been disappointed .
I 've been to this location twice now and both times I 've been very impressed . I 've tried their specialty pizzas and they 're all really good . The only problem is that they 're not open on sundays . They 're not open on sundays .

Source: P. Kaghazgaran et.al. 2019. Wide-Ranging Review Manipulation Attacks: Model, Empirical Study, and Countermeasures. In CIKM.

Camouflaging Behavior of Fraudsters

- **Relation Camouflage**
  - Crafty fraudsters could connect to benign entities under a relation to alleviate its suspiciousness[1].



[1] Yang, Xiaoyu, et al. "Rumor Detection on Social Media with Graph Structured Adversarial Learning." IJCAI, 2020.

Camouflaged Fraudsters Meets GNN

Fraudster: relation camouflage.

↓

GNN: a fraudster node may have many benign neighbors.

Fraudster: feature camouflage.

↓

GNN: neighbors with similar features may have different labels.

Enhance GNN-based Fraud Detectors

- The fraudsters are smart and agile.

- It is difficult to **exactly detect** the camouflaged fraudsters.

- We propose **three neural modules** to enhance GNN.

Label-aware Similarity Measure

- Previous works use cosine similarity, Euclidean distance to measure the feature/embedding similarity.
- Unsupervised similarity measure cannot identify **feature camouflage**.
- The similarity measure must have knowledge of fraudsters.

We introduce an **MLP** to encode the label information and use its output as the similarity measure.

Similarity-aware Neighbor Selector

- For a center node, different relations may have different amount of informative neighbors.
- We proposes an adaptive neighbor selector using **reinforcement learning** to find the optimal thresholds.

The RL process is a multi-armed bandit with following rules:

- If the avg. neighbor similarity score is greater than previous epoch, we **increase** the filtering threshold;
- Else, we **decrease** the filtering threshold.

Relation-aware Neighbor Aggregator

- We need to aggregate information across different relations.

- If we have selected informative neighbors under **every** relation, the attention mechanism is useless.

We directly utilize the **neighbor filtering thresholds** as the relation aggregation weights.

Experimental Setting

- Datasets:

Table 2: Dataset and graph statistics.

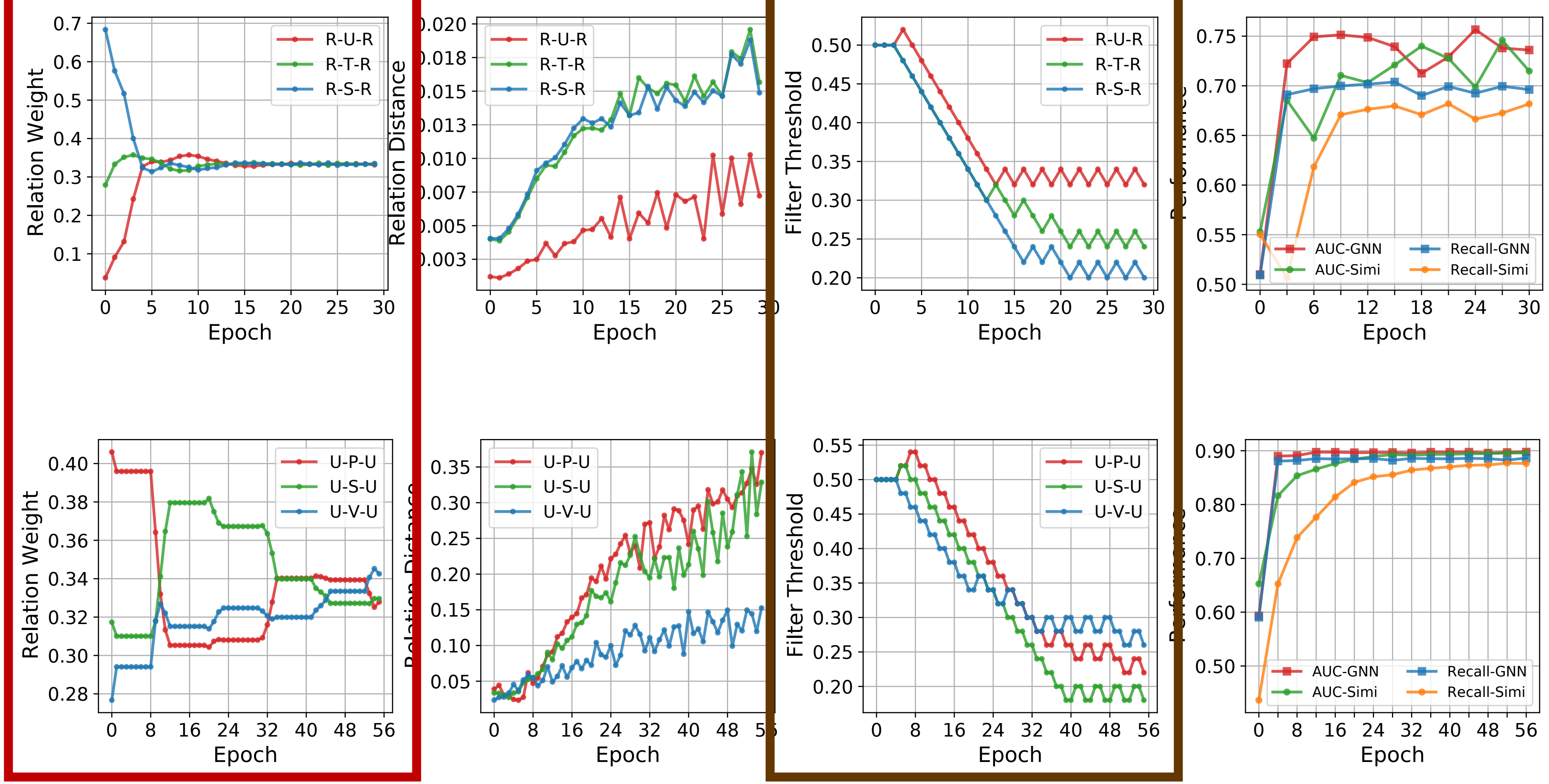
	#Nodes (Fraud%)	Relation	#Edges	Avg. Feature Similarity	Avg. Label Similarity
Yelp	45,954 (14.5%)	*R-U-R*	49,315	0.83	0.90
		R-T-R	573,616	0.79	0.05
		R-S-R	3,402,743	0.77	0.05
		ALL	3,846,979	0.77	0.07
Amazon	11,944 (9.5%)	*U-P-U*	175,608	0.61	0.19
		U-S-U	3,566,479	0.64	0.04
		U-V-U	1,036,737	0.71	0.03
		ALL	4,398,392	0.65	0.05

- Graphs: multi-relation graph with three relations.

Reinforcement Learning Process

Yelp

Amazon



Overall Evaluation

Table 3: Fraud detection performance (%) on two datasets under different percentage of training data.

	Metric	Train%	GCN	GAT	RGCN	Graph-SAGE	Genie-Path	Player-2Vec	Semi-GNN	Graph-Consis	CARE-*Att*	CARE-*Weight*	CARE-*Mean*	CARE-GNN
Yelp	AUC	5%	54.98	56.23	50.21	53.82	56.33	51.03	53.73	61.58	66.08	71.10	69.83	**71.26**
		10%	50.94	55.45	55.12	54.20	56.29	50.15	51.68	62.07	70.21	71.02	71.85	**73.31**
		20%	53.15	57.69	55.05	56.12	57.32	51.56	51.55	62.31	73.26	74.32	73.32	**74.45**
		40%	52.47	56.24	53.38	54.00	55.91	53.65	51.58	62.07	74.98	74.42	74.77	**75.70**
	Recall	5%	53.12	54.68	50.38	54.25	52.33	50.00	52.28	62.60	63.52	66.64	**68.09**	67.53
		10%	51.10	52.34	51.75	52.23	54.35	50.00	52.57	62.08	67.38	68.35	**68.92**	67.77
		20%	53.87	53.20	50.92	52.69	54.84	50.00	52.16	62.35	68.34	69.07	**69.48**	68.60
		40%	50.81	54.52	50.43	52.86	50.94	50.00	50.59	62.08	71.13	70.22	69.25	**71.92**
Amazon	AUC	5%	74.44	73.89	75.12	70.71	71.56	76.86	70.25	85.46	89.49	89.36	89.35	**89.54**
		10%	75.25	74.55	74.13	73.97	72.23	75.73	76.21	85.29	**89.58**	89.37	89.43	89.44
		20%	75.13	72.10	75.58	73.97	71.89	74.55	73.98	85.50	89.58	**89.68**	89.34	89.45
		40%	74.34	75.16	74.68	75.27	72.65	56.94	70.35	85.50	89.70	89.69	89.52	**89.73**
	Recall	5%	65.54	63.22	64.23	69.09	65.56	50.00	63.29	85.49	88.22	88.31	88.02	**88.34**
		10%	67.81	65.84	67.22	69.36	66.63	50.00	63.32	85.38	87.87	**88.36**	88.12	88.29
		20%	66.15	67.13	65.08	70.30	65.08	50.00	61.28	85.59	88.40	**88.60**	88.00	88.27
		40%	67.45	65.51	67.68	70.16	65.41	50.00	62.89	85.53	88.41	88.45	88.22	**88.48**

Model Advantage

- **Adaptability.** CARE-GNN adaptively selects best neighbors for aggregation given arbitrary multi-relation graph.
- **High-efficiency.** CARE-GNN has a high computational efficiency without attention and deep reinforcement learning.
- **Flexibility.** Many other neural modules and external knowledge can be plugged into the CARE-GNN.

SafeGraph (https://github.com/safe-graph)

- **DGFraud**: a GNN-based fraud detection toolbox.
- **UGFraud**: an unsupervised graph-based fraud detection toolbox.
- Graph-based Fraud Detection **Paper List**.

Enhancing Graph Neural Network-based Fraud Detectors against Camouflaged Fraudsters

Yingtong Dou (UIC), Zhiwei Liu (UIC), Li Sun (BUPT),
Yutong Deng (BUPT), Hao Peng (BUAA), Philip S. Yu (UIC)

Email: ydou5@uic.edu
Homepage: http://ytongdou.com
Project Page: https://github.com/safe-graph
Paper: https://arxiv.org/abs/2008.08692
Code: https://github.com/YingtongDou/CARE-GNN