Deep Generative Graph Models & Adversarial Robustness

I. Makarov & V. Pozdnyakov & D. Kiselev

BigData Academy MADE from Mail.ru Group

Graph Neural Networks and Applications

Topics

1. Random Graph Models
2. Generative GNN
3. Robustness to Adversarial Attacks

Random Graph Models

Random graph models comparison

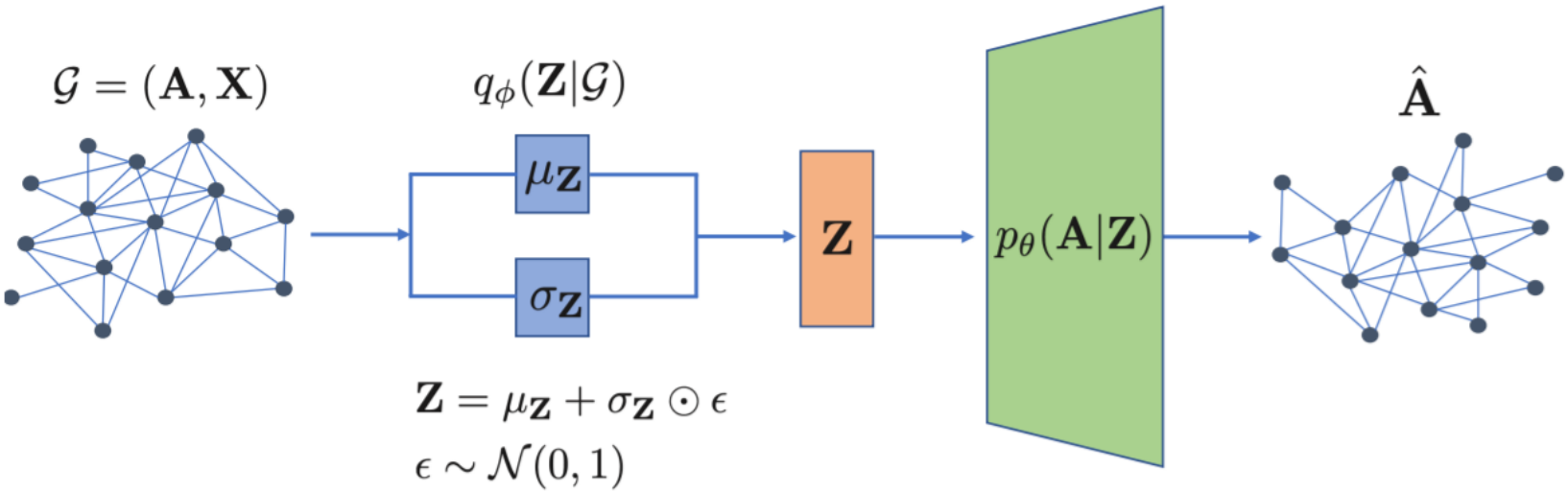
	Random	BA model	WS model	Empirical networks
P(k)	$\frac{\lambda^k e^{-\lambda}}{k!}$	k^{-3}	poisson like	power law
C	$\langle k \rangle / N$	$N^{-0.75}$	const	large
$\langle L \rangle$	$\frac{\log(N)}{\log(\langle k \rangle)}$	$\frac{\log(N)}{\log\log(N)}$	$\log(N)$	small

Many more models, e.g., Stochastic Block Model:
Two parameters $p >> q$, p stands for edge probability inside community, q stands for intra-cluster edges probability

Graph Generative Models

Variational Graph Auto-Encoder

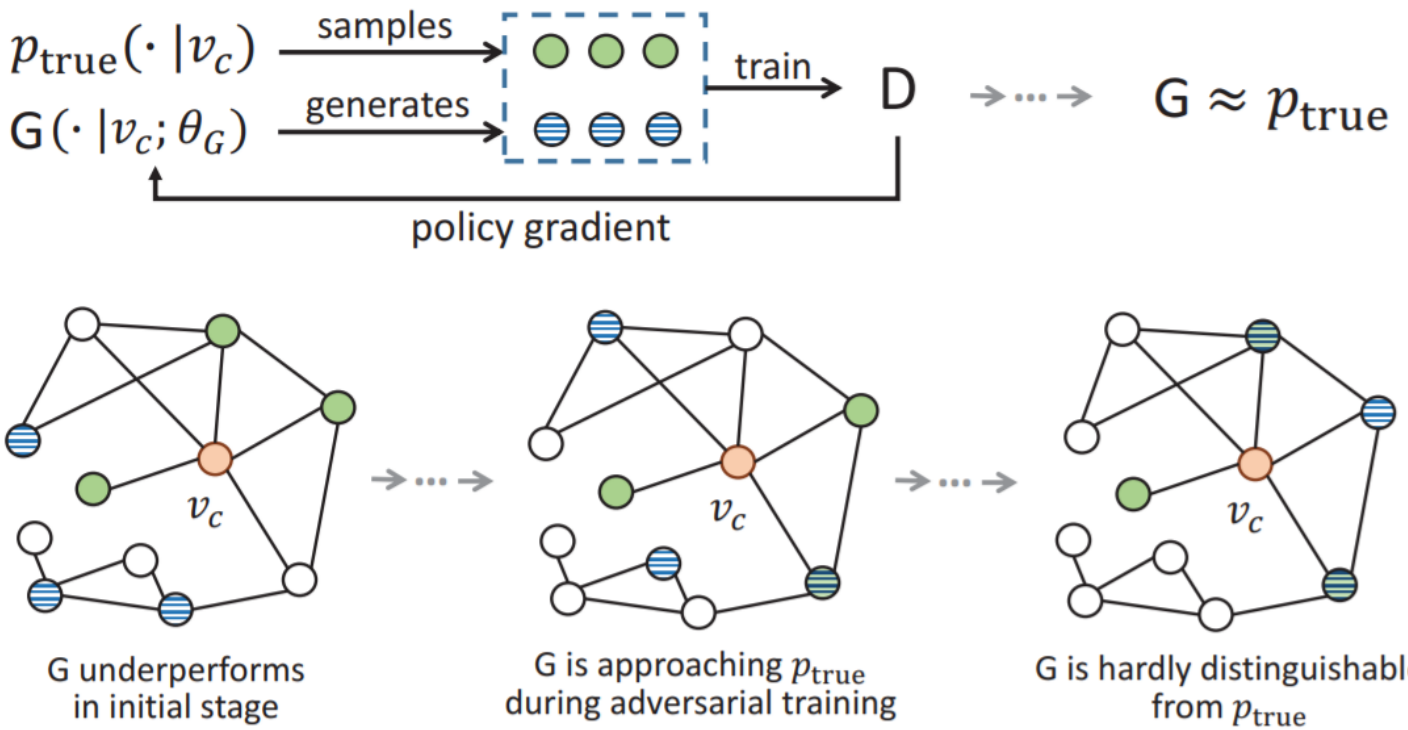
1. An encoder neural network maps the input graph $G = (A, X)$ to a posterior distribution $q_\phi(Z \mid G)$ over latent variables Z.
2. Given a sample from this posterior, the decoder model $p_\theta(A \mid Z)$ attempts to reconstruct the adjacency matrix.
3. See also Adversarial VGAE in Pechenizkiy et al. 2021



from Hamilton et al., 2020

Learning graph representation with generative adversarial nets (GraphGAN)

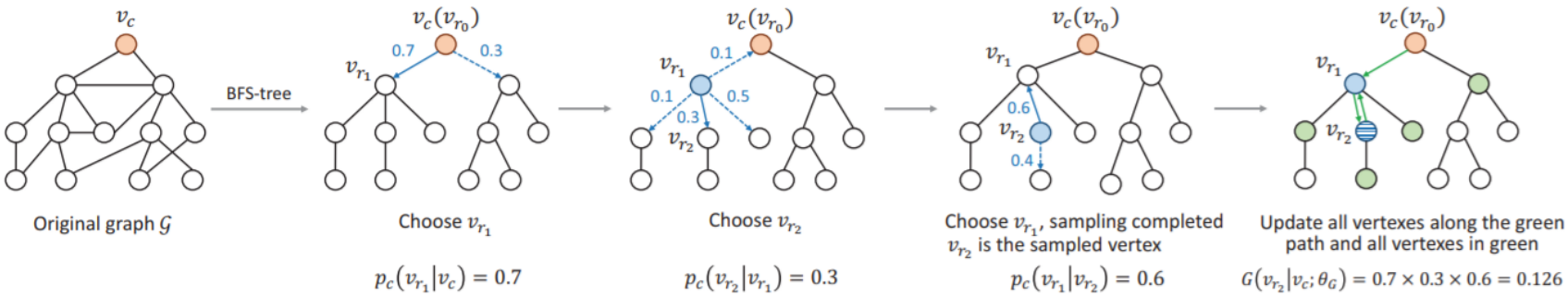
1. Classic GAN training for node-level distributions with negative sampling for policy optimization
2. Graph SoftMax with efficient, distance-aware probabilistic framework



from Guo et al., 2019

Learning graph representation with generative adversarial nets (GraphGAN)

1. Blue digits are the relevance probability, and the blue solid arrow indicates the direction that G chooses to move in.
2. Upon completion of sampling, the vertex with blue stripes is the sampled one, and all colored vertices in the rightmost tree require updating accordingly
3. Gaussian Graph Processes in Kyrki et al., 2021



from Guo et al., 2019

Graph Recurrent Attention Network (GraphRNN)

1. From $P(G|z) = \prod_{(u,v)} P(A[u,v]|z)$ to

$$P(G|z) = \prod_i P(L[v_i,:]|L[v_1,:],...,L[v_{i-1},:],z),$$

where L is low-triangular part of A.

2. Hiearchical RNN to model the edge dependencies.
3. The first graph-level RNN models hidden state h_i, which is updated after generating each row of the adjacency matrix $L[v_i,:]$ via $h_{i+1} = RNNgraph(h_i, L[v_i, L])$.
4. $h_0 = 0$ or could be learned by a graph encoder model or sampled from a latent space in a VAE-style approach.

from Leskovec et al., 2018

Graph Recurrent Attention Network (GraphRNN)

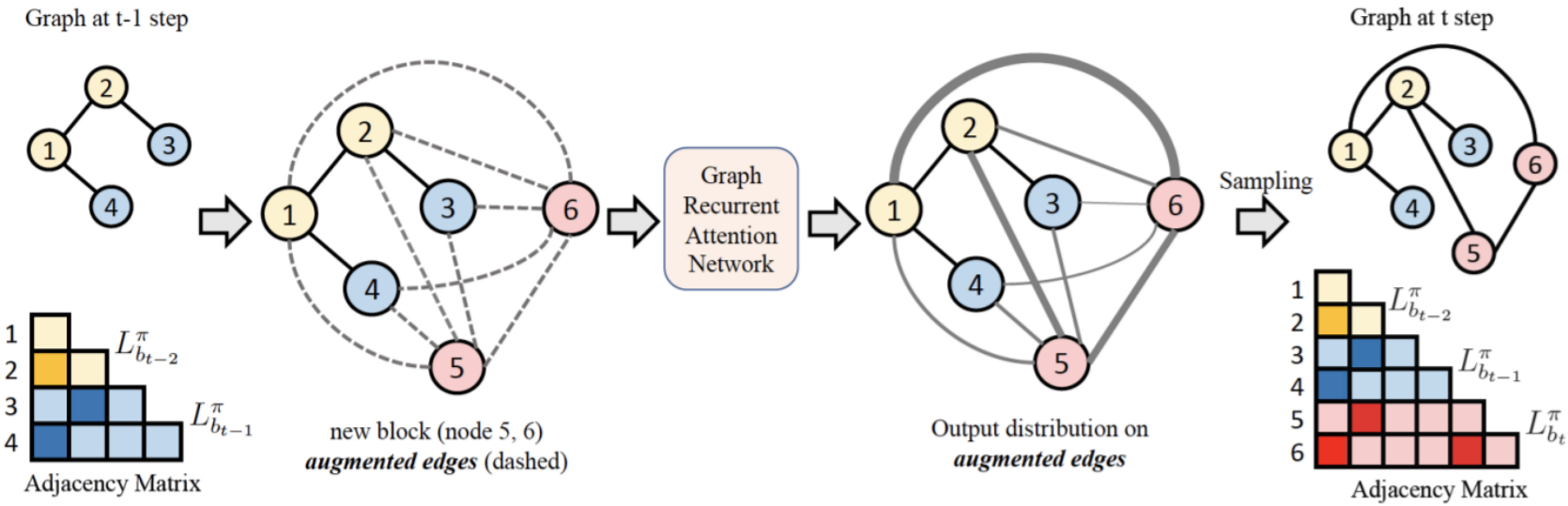
1. The second node-level RNN generates the entries of L[vi, :] in an autoregressive manner, assuming a conditional Bernoulli distribution for each entry.
2. Both the graph-level RNNgraph and the node-level RNNnode can be optimized to maximize the likelihood the training graphs using the teaching forcing strategy with the ground truth values of L to update the RNNs during training.
3. To control the size of the generated graphs, the RNNs are also trained to output end-of-sequence tokens.
4. GraphRNN model still generates unrealistic long chains and is hard to train due to vanishing gradient

from Leskovec et al., 2018

Graph Recurrent Attention Network (GRAN)

1. GRAN maintains the autoregressive decomposition of the generation process, but instead of using RNNs to model the autoregressive generation process, GRAN uses GNNs.
2. The key idea in GRAN is that we can model the conditional distribution of each row of the adjacency matrix by running a GNN

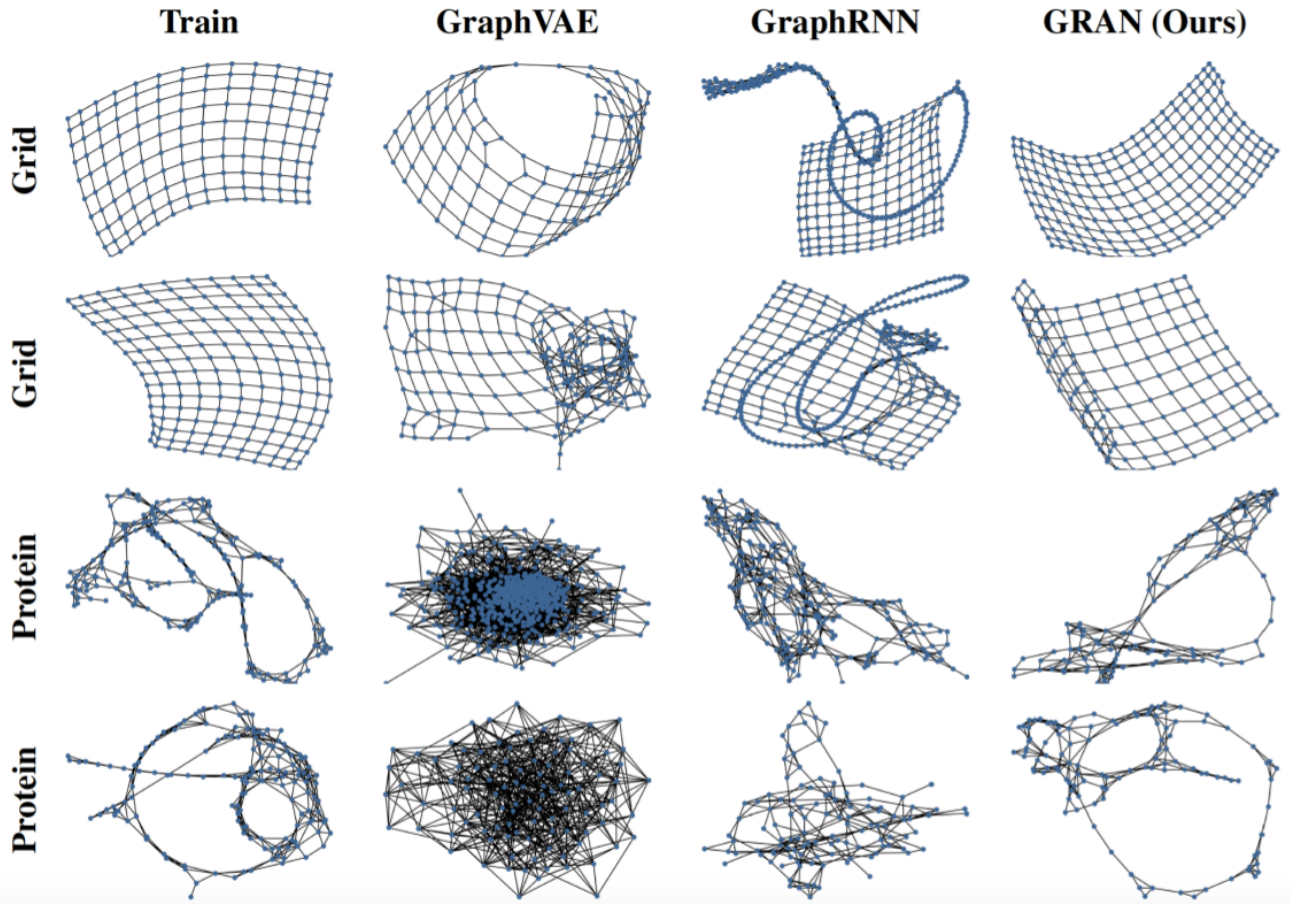
$$P(L[v_i, :]|L[v_1, :], ..., L[v_{i-1}, :], z) \equiv GNN(L[v_1 : v_{i-1}, :], X)$$



from Zemel et al., 2019

Graph Generation Evaluation

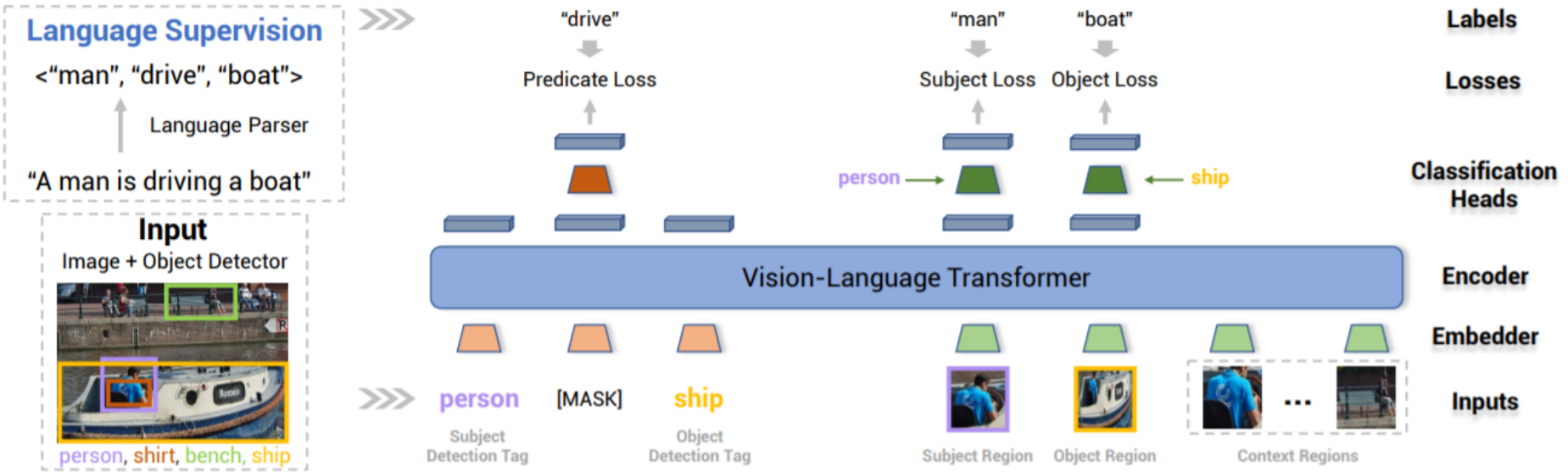
1. Selecting specific statistics or validating domain knowledge



from Zemel et al., 2019

Graph Generation via Related Task Supervision

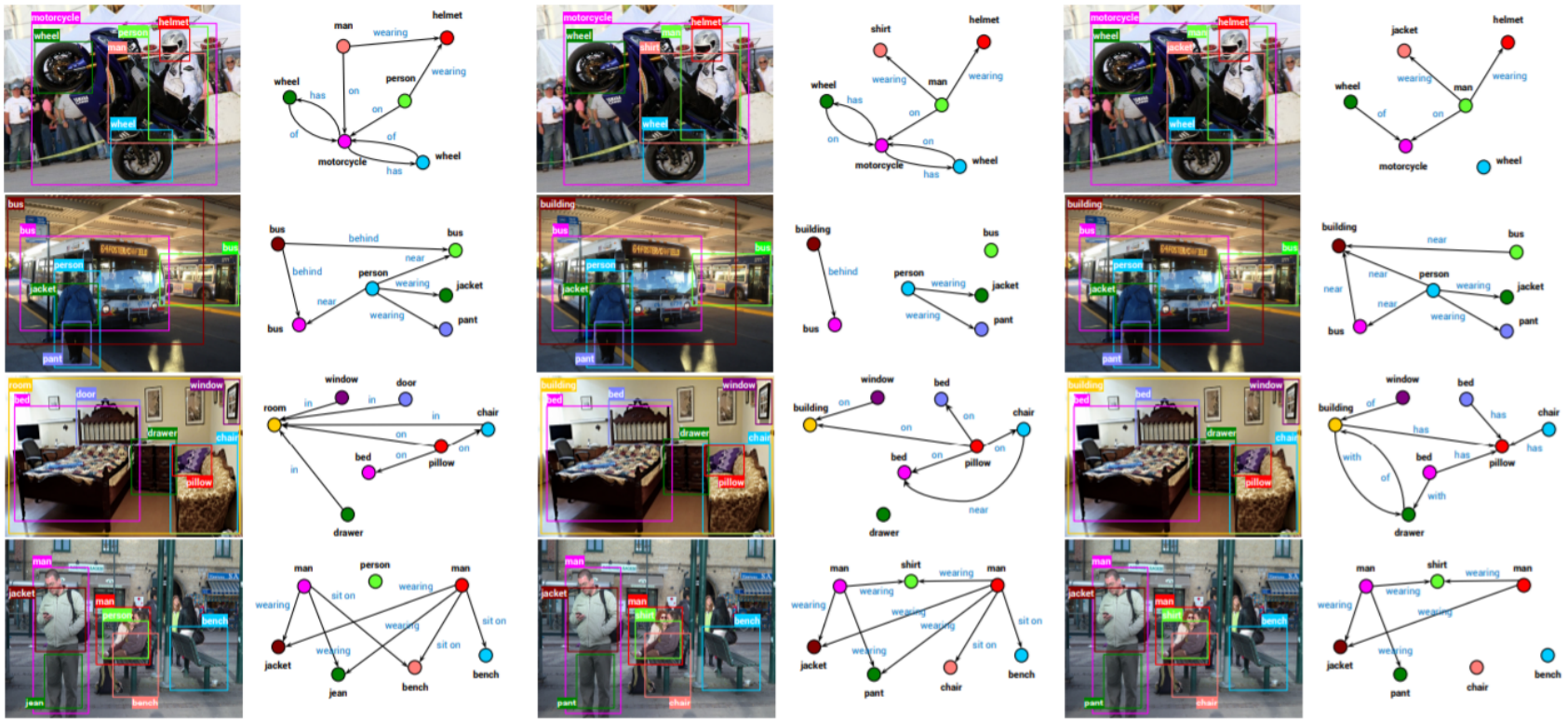
1. Language supervised scene graph generation.
2. Multi-modal NLP+CV model for extracting Visual Knowledge Graph



from Li et al., 2021

Graph Generation via Related Task Supervision

1. Graph generation is defined by object detector
2. Relations are extracted from visual and textual encoders via transformer.

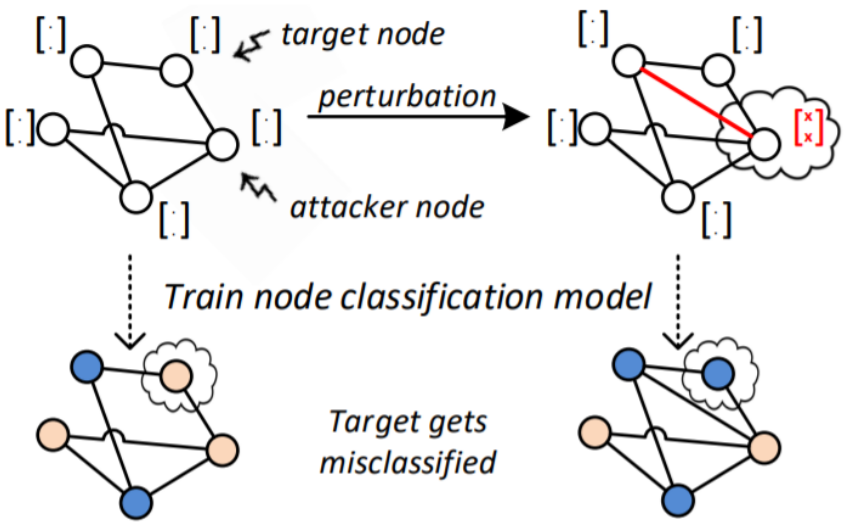


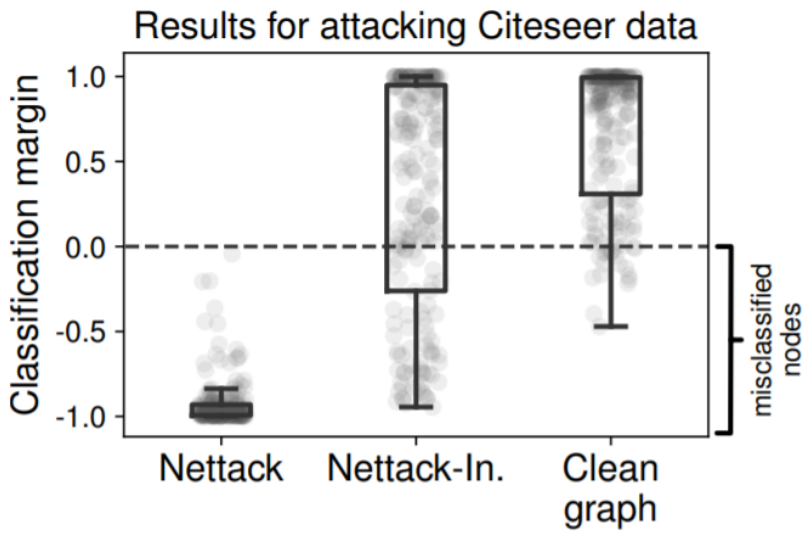
from Li et al., 2021

Graph Anomaly Detection and Adversarial Attacks

Adversarial Attacks on Neural Networks for Graph Data

1. Connecting to node neighborhood to change its embedding

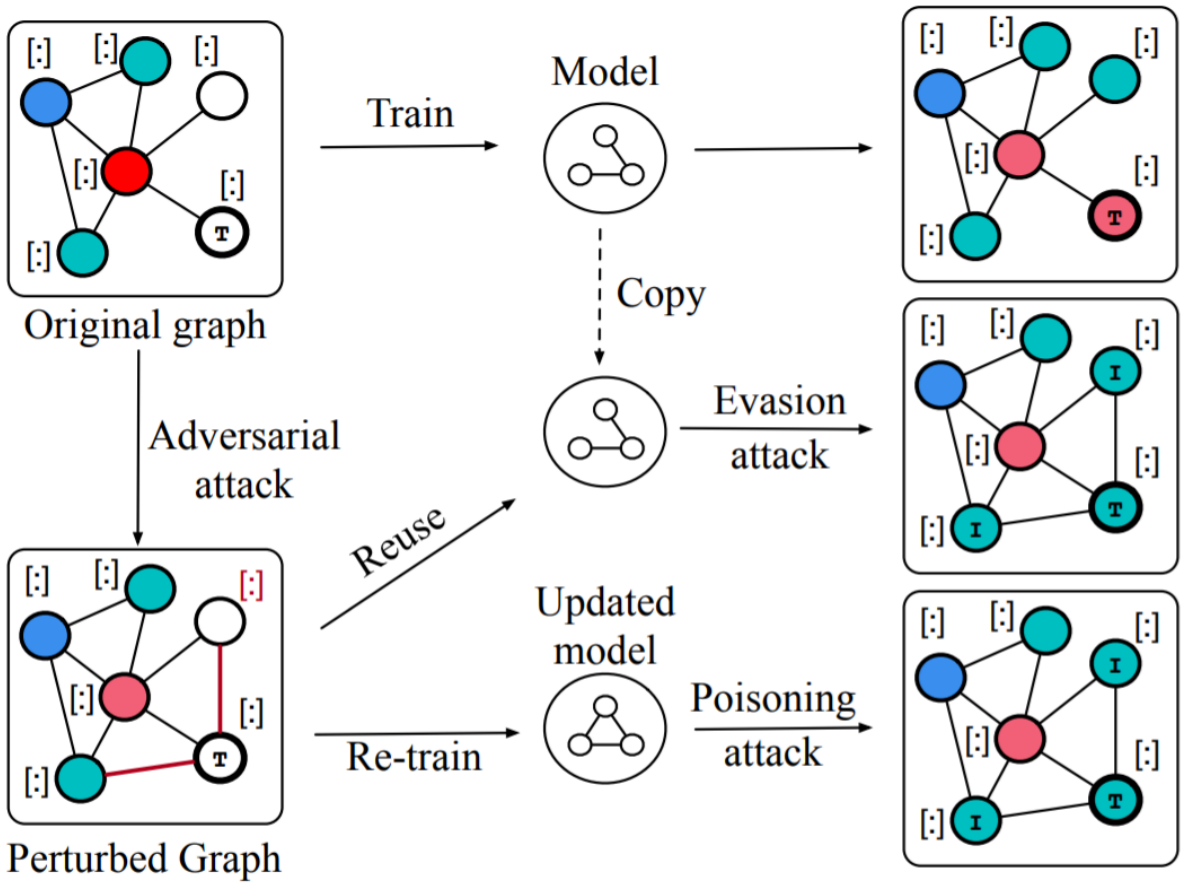




from Günnemann et al., 2018

Adversarial Attacks on Neural Networks for Graph Data

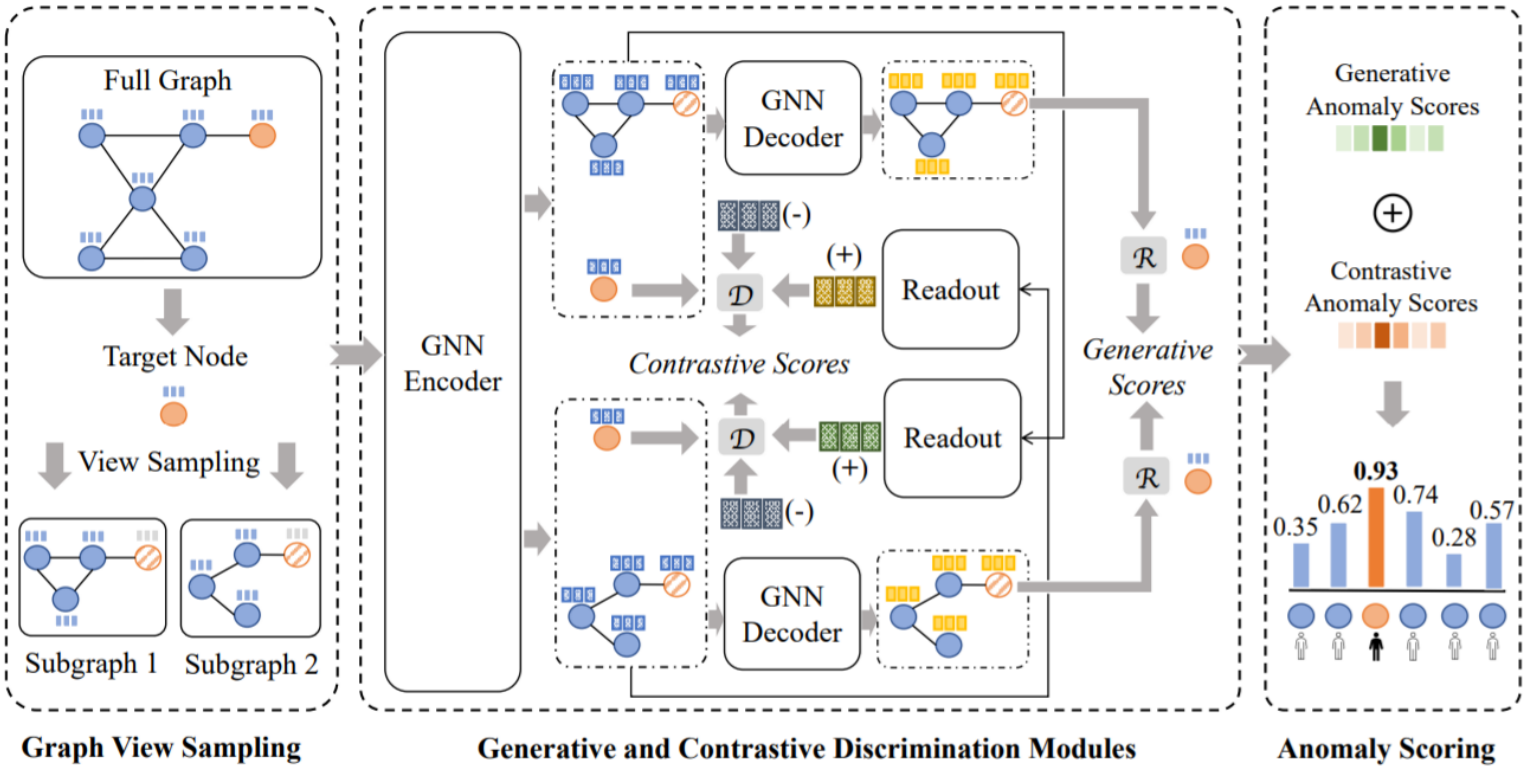
1. An example of the evasion attack and the poisoning attack.
2. Small perturbations affect node embeddings



from Zheng et al., 2020; Günnemann et al., 2018

Generative and Contrastive Self-Supervised Learning for Graph Anomaly Detection

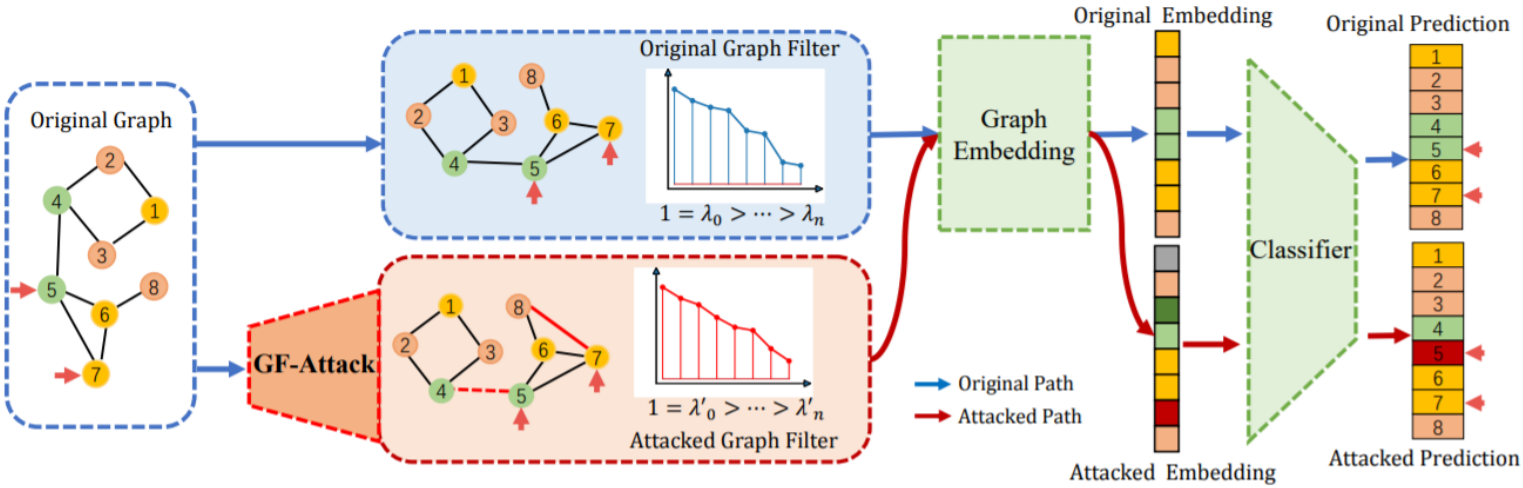
1. Consistent representations over sampling subgraphs
2. Matching distributions from attributes regression and contrastive learning



from Chen et al., 2015

A restricted black-box adversarial framework towards attacking graph embedding models

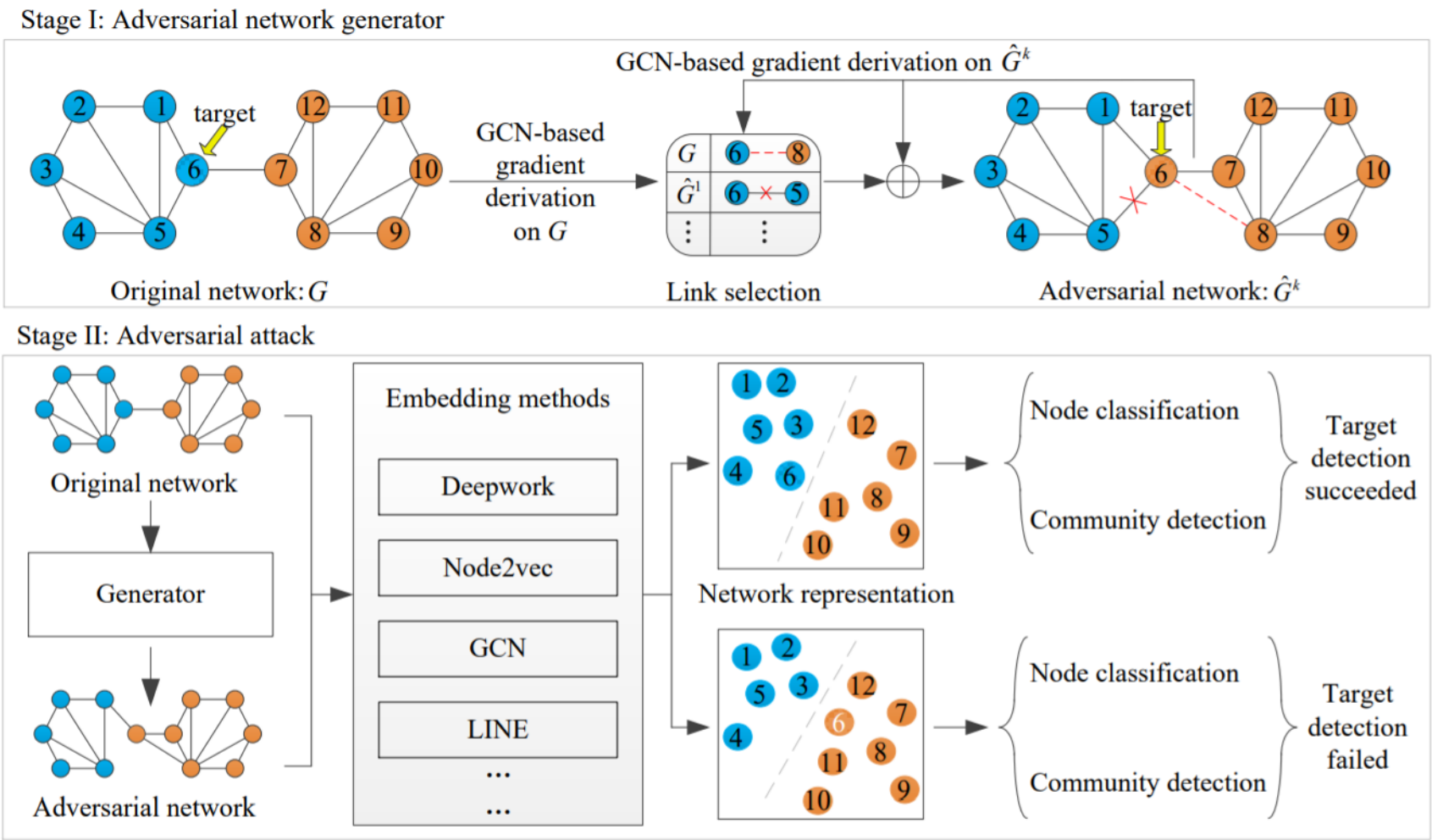
1. Attack particular nodes by edge perturbations
2. Fit graph filters



from Huang et al., 2020

Fast Gradient Attack on Network Embedding

1. Change gradient direction in graph structure



from Xuan et al., 2018

Fast Gradient Attack on Network Embedding

1. Adversarial Network Attack Generator
2. Gradient and Laplacian regularization may help, see Tang et al., 2021

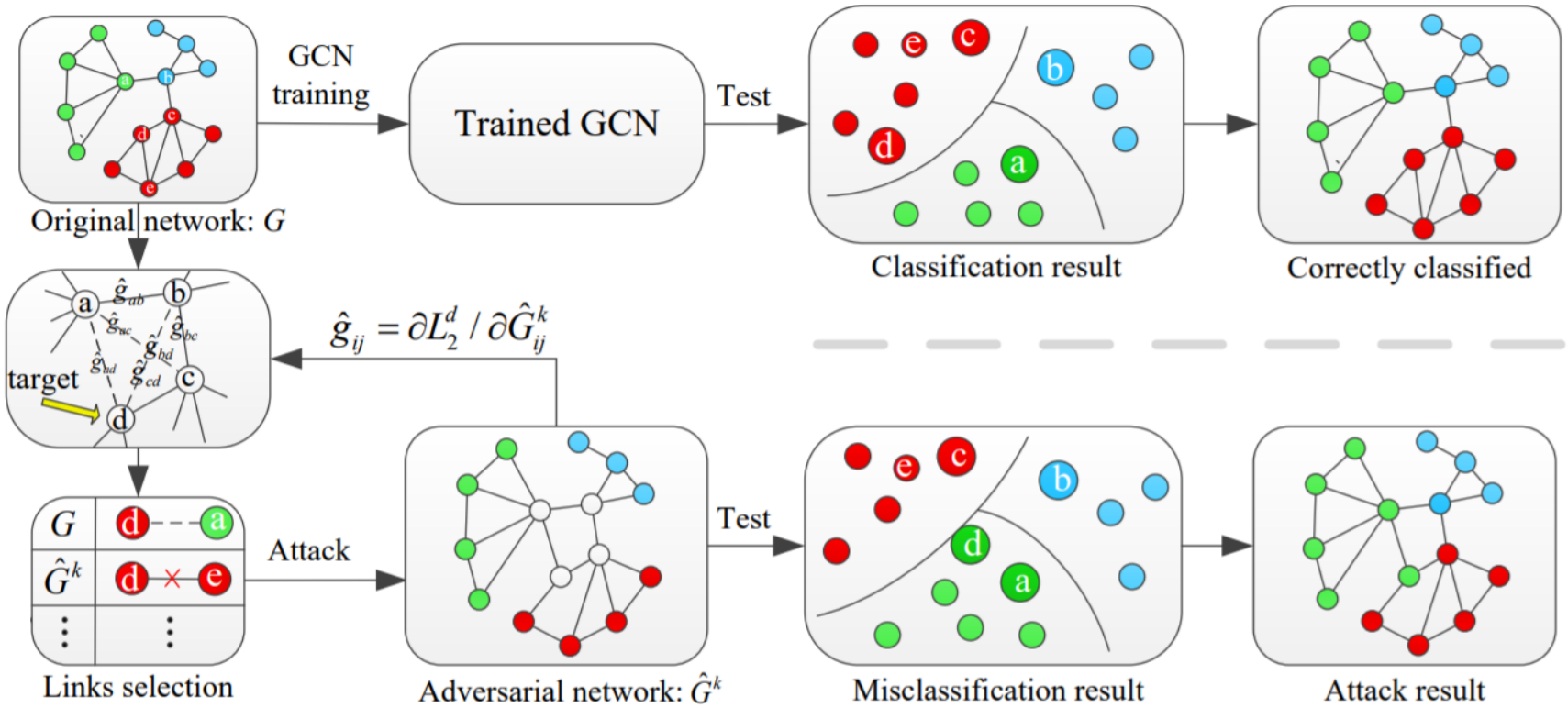
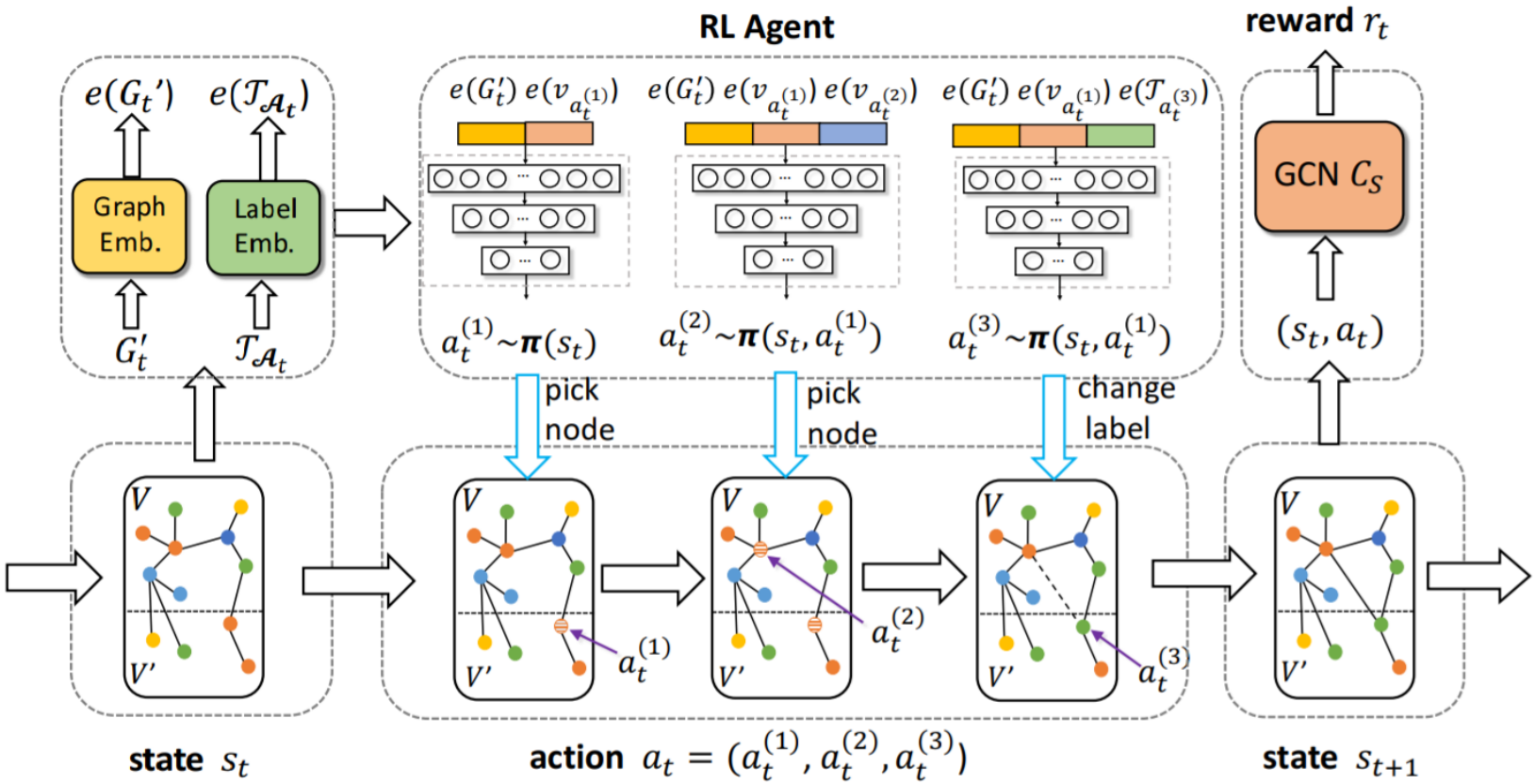


Fig. 2: Adversarial network attack generator via GCN.

from Xuan et al., 2018

Node Injection Attacks on Graphs via Reinforcement Learning

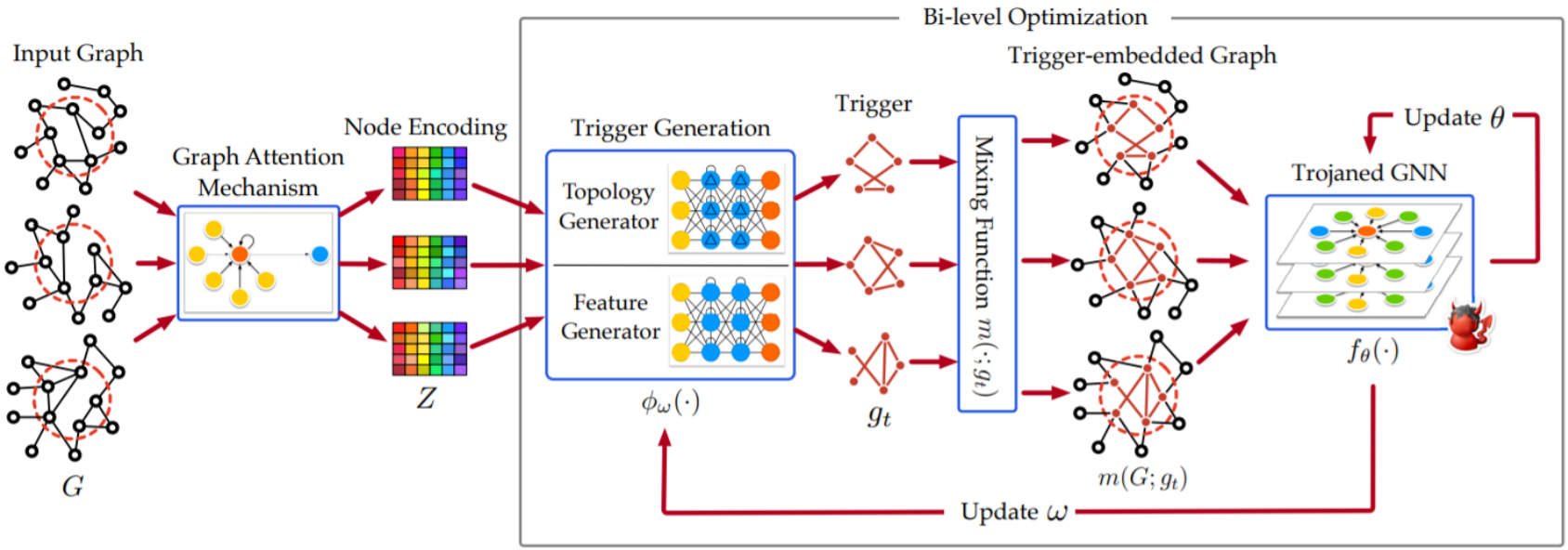
1. Adversarial Network Attack for Node Injection
2. Label poisoning via group attack



from Honavar et al., 2019

Graph Backdoor

1. Triggers as specific subgraphs, including both topological structures and descriptive features
2. Dynamically adapts triggers to individual graphs, thereby optimizing both attack effectiveness and evasiveness
3. Downstream model-agnostic and works for inductive settings

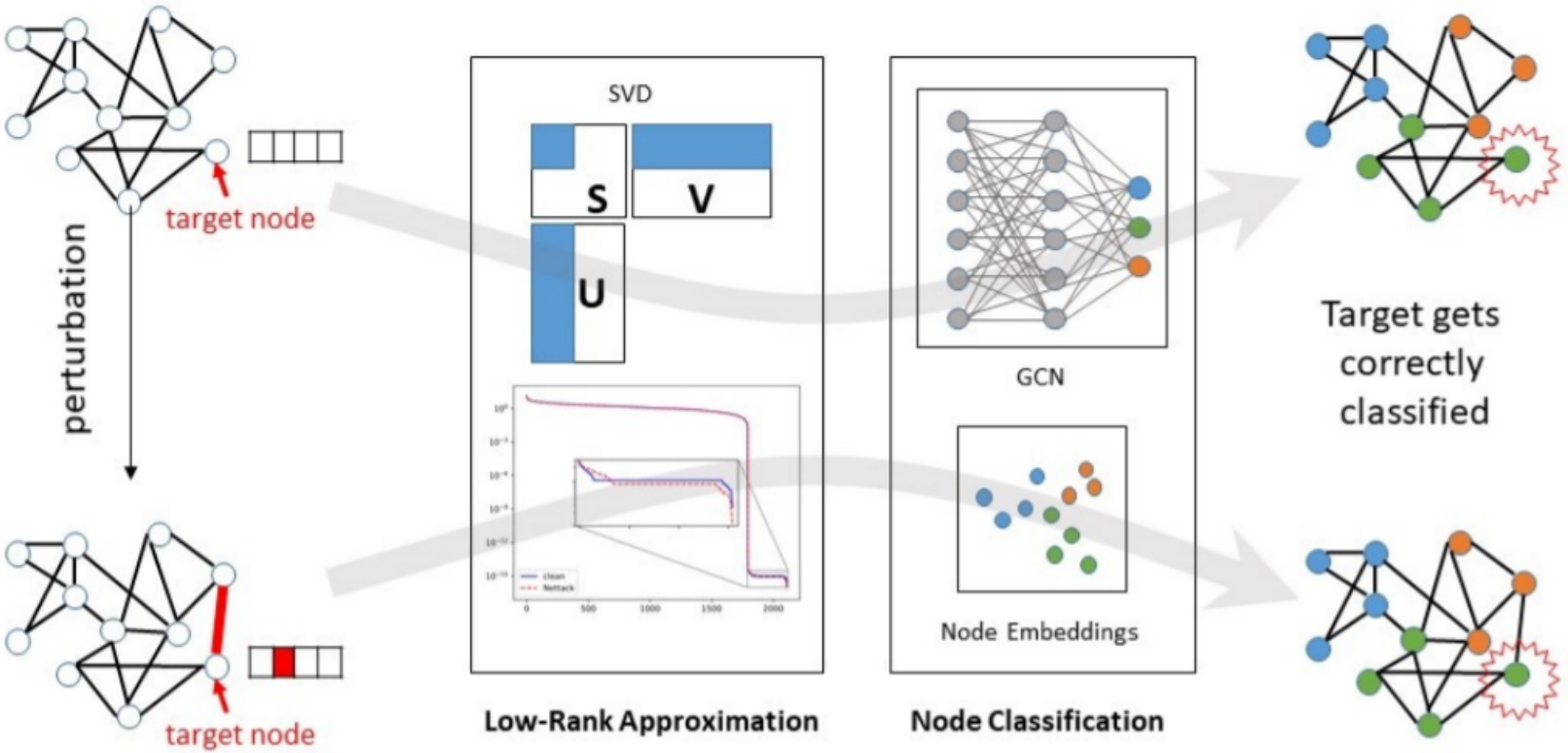


from Wang et al., 2020

Defense against Adversarial Attacks on Graphs

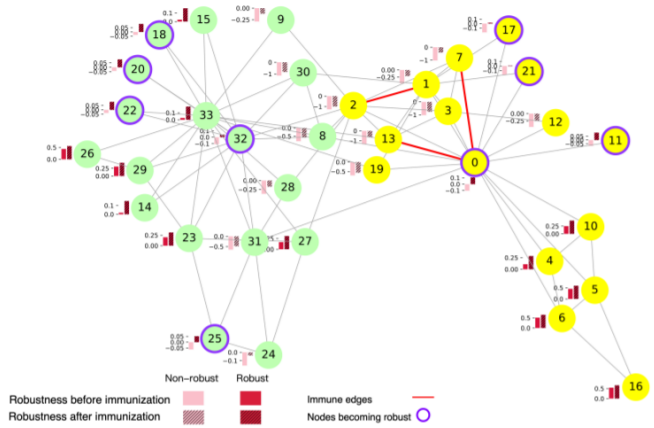
All You Need Is Low (Rank): Defending Against Adversarial Attacks on Graphs

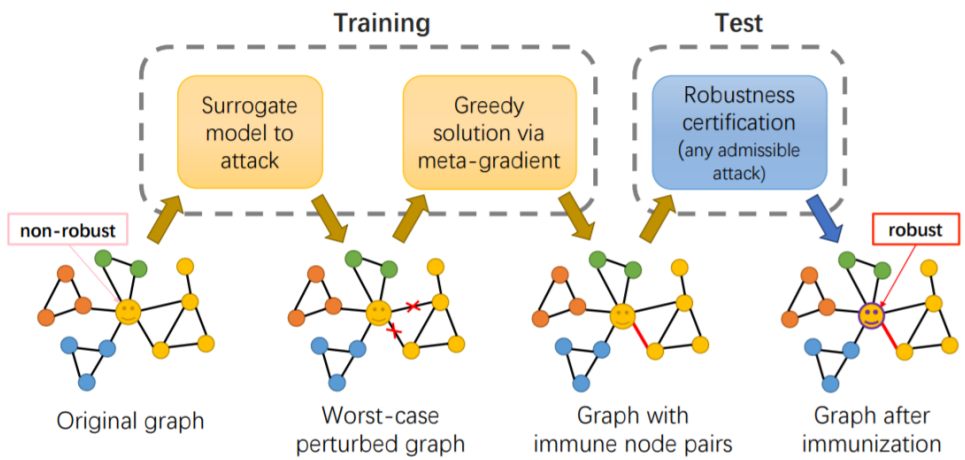
1. Low-rank approximation of graph structure and feature matrices to vaccinate the node classification method and discard perturbations.
2. Difference in eigen-values in log-scale is essential to defense



Adversarial Immunization for Certifiable Robustness on Graphs

1. The node is certified as robust (red), when its robustness ¿ 0, otherwise as non-robust (pink). Purple circle indicates the node that becomes robust through immunization.
2. The red edges are immune (musthave) edges.

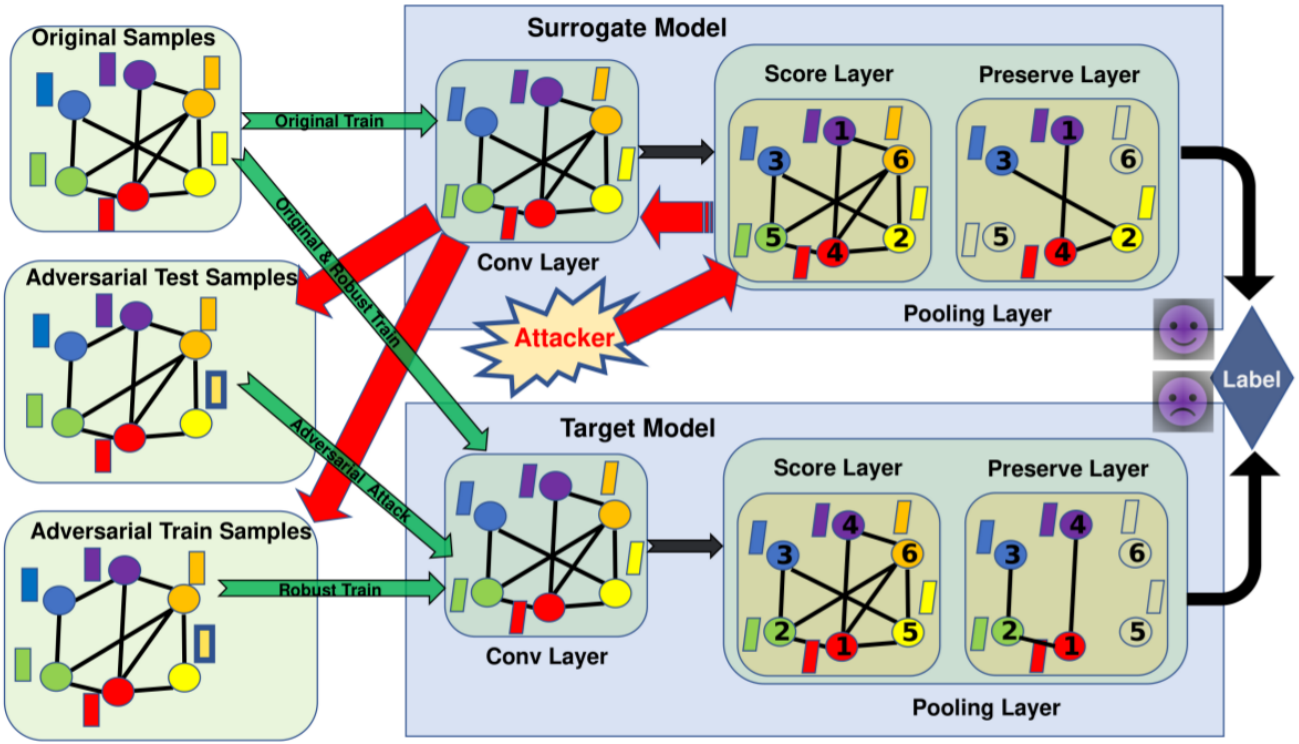




from Cheng et al., 2021

Adversarial Attack on Hierarchical Graph Pooling Neural Networks

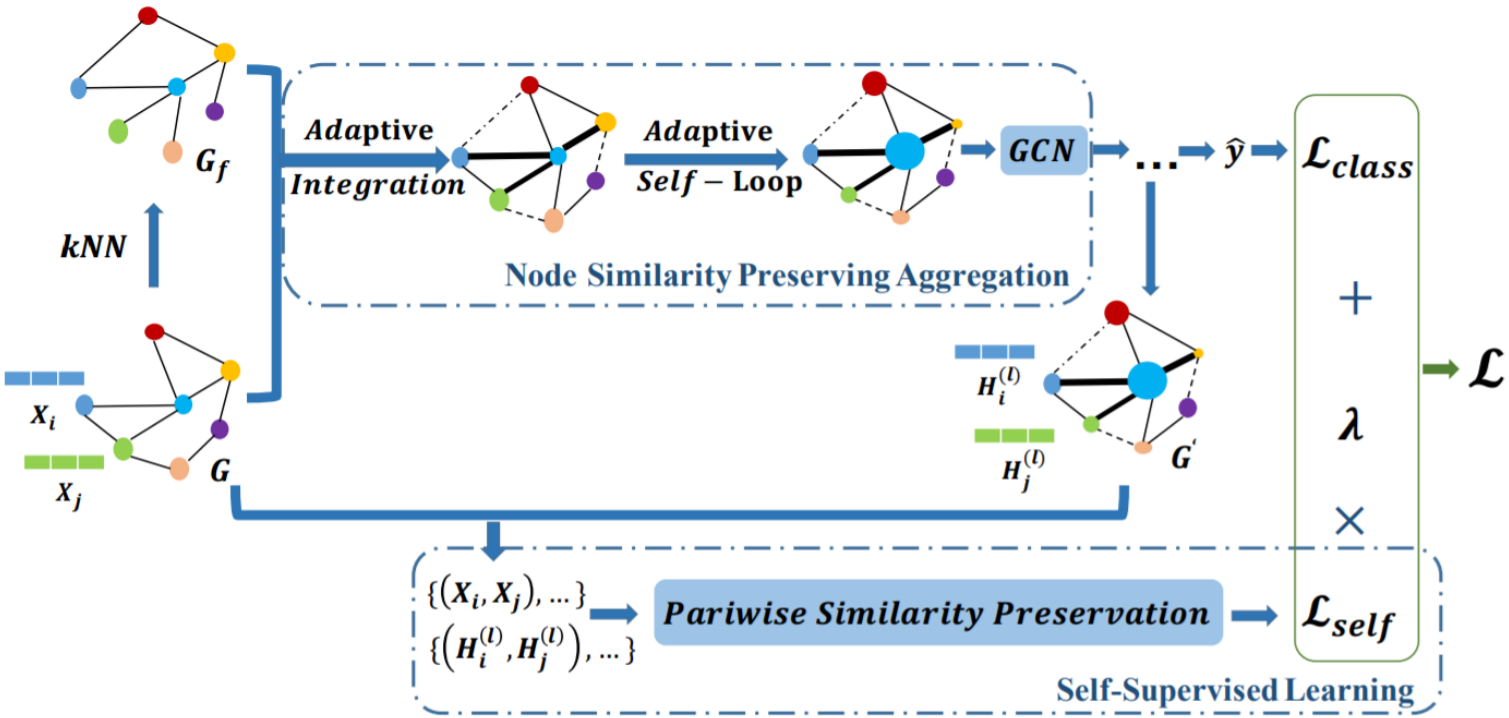
1. Attack on Graph Classification.
2. Red arrow represents the adversarial samples' generating. Green arrow is the workflow and black arrow is the neural network propagation.



from Zhan et al., 2020

Node Similarity Preserving Graph Convolutional Networks

1. Robust distributional semantics
2. Robust performance but computational complexity si drawback



from Tang et al., 2021

References

- Hamilton, W.L., 2020. Graph representation learning. Synthesis Lectures on Artifical Intelligence and Machine Learning, 14(3), pp.1-159.
- You, Jiaxuan, Rex Ying, Xiang Ren, William Hamilton, and Jure Leskovec. "Graphrnn: Generating realistic graphs with deep auto-regressive models." In International conference on machine learning, pp. 5708-5717. PMLR, 2018.
- Liao, R., Li, Y., Song, Y., Wang, S., Nash, C., Hamilton, W.L., Duvenaud, D., Urtasun, R. and Zemel, R.S., 2019. Efficient graph generation with graph recurrent attention networks. arXiv preprint arXiv:1910.00760.
- Wang, H., Wang, J., Wang, J., Zhao, M., Zhang, W., Zhang, F., Li, W., Xie, X. and Guo, M., 2019. Learning graph representation with generative adversarial nets. IEEE Transactions on KDE.

References

- Huang, T., Pei, Y., Menkovski, V. and Pechenizkiy, M., 2021, September. On Generalization of Graph Autoencoders with Adversarial Training. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases (pp. 367-382). Springer, Cham.
- Blanco-Mulero, D., Heinonen, M. and Kyrki, V., 2021. Evolving-Graph Gaussian Processes. arXiv preprint arXiv:2106.15127.
- Zhong, Y., Shi, J., Yang, J., Xu, C. and Li, Y., 2021. Learning to generate scene graph from natural language supervision. In Proceedings of the IEEE/CVF ICCV (pp. 1823-1834).
- Zheng, Y., Jin, M., Liu, Y., Chi, L., Phan, K.T. and Chen, Y.P.P., 2021. Generative and Contrastive Self-Supervised Learning for Graph Anomaly Detection. IEEE Transactions on KDE.
- https://github.com/ChandlerBang/awesome-graph-attack-papers
- https://github.com/gitgiter/Graph-Adversarial-Learning

References

- Chang, H., Rong, Y., Xu, T., Huang, W., Zhang, H., Cui, P., Zhu, W. and Huang, J., 2020, April. A restricted black-box adversarial framework towards attacking graph embedding models. In Proceedings of the AAAI (Vol. 34, No. 04, pp. 3389-3396).
- Chen, J., Wu, Y., Xu, X., Chen, Y., Zheng, H. and Xuan, Q., 2018. Fast gradient attack on network embedding. arXiv preprint arXiv:1809.02797.
- Zügner, D., Akbarnejad, A. and Günnemann, S., 2018, July. Adversarial attacks on neural networks for graph data. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2847-2856).
- Entezari, N., Al-Sayouri, S.A., Darvishzadeh, A. and Papalexakis, E.E., 2020, January. All you need is low (rank) defending against adversarial attacks on graphs. In Proceedings of the 13th International Conference on Web Search and Data Mining (pp. 169-177).

References

- Tao, S., Shen, H., Cao, Q., Hou, L. and Cheng, X., 2021, March. Adversarial Immunization for Certifiable Robustness on Graphs. In Proceedings of the 14th ACM International Conference on Web Search and Data Mining (pp. 698-706).
- Tang, H., Ma, G., Chen, Y., Guo, L., Wang, W., Zeng, B. and Zhan, L., 2020. Adversarial attack on hierarchical graph pooling neural networks. arXiv preprint arXiv:2005.11560.
- Liu, X., Jin, W., Ma, Y., Li, Y., Liu, H., Wang, Y., Yan, M. and Tang, J., 2021, July. Elastic graph neural networks. In International Conference on Machine Learning (pp. 6837-6849). PMLR.
- Jin, W., Derr, T., Wang, Y., Ma, Y., Liu, Z. and Tang, J., 2021, March. Node similarity preserving graph convolutional networks. In Proceedings of the 14th ACM International Conference on Web Search and Data Mining (pp. 148-156).

References

- Sun, Y., Wang, S., Tang, X., Hsieh, T.Y. and Honavar, V., 2019. Node injection attacks on graphs via reinforcement learning. arXiv preprint arXiv:1909.06543.
- Jin, W., Li, Y., Xu, H., Wang, Y., Ji, S., Aggarwal, C. and Tang, J., 2020. Adversarial Attacks and Defenses on Graphs: A Review, A Tool and Empirical Studies. arXiv preprint arXiv:2003.00653.
- Chen, L., Li, J., Peng, J., Xie, T., Cao, Z., Xu, K., He, X. and Zheng, Z., 2020. A survey of adversarial learning on graphs. arXiv preprint arXiv:2003.05730.
- Sun, L., Dou, Y., Yang, C., Wang, J., Yu, P.S., He, L. and Li, B., 2018. Adversarial attack and defense on graph data: A survey. arXiv preprint arXiv:1812.10528.
- Xi, Z., Pang, R., Ji, S. and Wang, T., 2021. Graph backdoor. In 30th USENIX Security Symposium (USENIX Security 21).