GADY: Unsupervised Anomaly Detection on Dynamic Graphs

Anonymous Author(s)

1 REPRODUCE OUR EXPERIMENTS

Please find the code of all experiments in this anonymous website: https://anonymous.4open.science/r/GADY

2 DATASETS

- **UCI Message**[1] [5] is a social network collected from a forum of the University of California, Irvine. It has 1,899 nodes and 13,838 edges where each node represents a student and an edge represents a message sent.
- **Email-DNC**[2] [8] is an email network from the 2016 Democratic National Committee email leak. It has 1,866 nodes and 39264 edges where each node represents a person in the Democratic Party and an edge represents an email sent from one person to another.
- **Bitcoin-OTC**[3] [3] is a network of who trusts whom among users who trade on the Bitcoin platform. It has 5,881 nodes and 35,588 edges where nodes are users and edges are ratings between users.

3 BASELINES

We compared GADY with nine competitive methods, which mainly include two categories: deep learning methods and graph encoding methods.

- **DeepWalk** [6] is a social network collected from a forum of the University of California, Irvine, where each node represents a student and an edge represents a message sent.
- **Node2vec** [2] is a social network collected from a forum of the University of California, Irvine, where each node represents a student and an edge represents a message sent.
- **Spectral Clustering** [9] is a social network collected from a forum of the University of California, Irvine, where each node represents a student and an edge represents a message sent.
- **Netwalk** [10] is a classic anomaly detection method on dynamic graphs, which generates node encodings based on random walk and dynamically updates reservoirs to model networks, and finally uses a dynamic clustering model to score edges for anomalies.
- **AddGraph** [11] is an end-to-end dynamic graph anomaly detection model. It uses GCN and GRU-attention to capture structural and temporal information respectively.
- **StrGNN** [1] is an end-to-end model for anomaly detection on dynamic graphs. It first takes h-hop subgraphs for the target edge, then uses GCN model to model the structural information within the slice, and uses GRU to model the temporal information between different slices.
- **Taddy** [4] is an end-to-end model for anomaly detection on dynamic graphs. It first samples subgraphs and obtains global encodings by diffusion, obtains local encodings based on distance on subgraphs and obtains relative time encodings according to time difference. Then it uses transformers to model dynamic graph networks and then uses a scoring module to obtain anomaly scores.

4 EXPERIMENT SETUP

For GADY(No-GAN), we generate negative samples using the same method with TADDY [4] and use BCEloss. We set E_s to 0.3, $Ratio$ to 15, and γ to 0.1 for experiments. We choose these values for hyper-parameters because of their excellent performance. For the parameters setting of the encoder part, we followed the hyper-parameters setting discussed in TGN[7]. We also follow the best parameters proposed in the papers for setting for all hyperparameter settings in the baselines. We follow [4] and use a 50%-50% (train-test) temporal split for all datasets. All experiments are completed using the same seed, and we also find that different seeds do not influence our model performance greatly.

REFERENCES

[1] Lei Cai, Zhengzhang Chen, Chen Luo, Jiaping Gui, Jingchao Ni, Ding Li, and Haifeng Chen. 2021. Structural temporal graph neural networks for anomaly detection in dynamic graphs. In *Proceedings of the 30th ACM international conference on Information & Knowledge Management*. 3747–3756.

[2] Aditya Grover and Jure Leskovec. 2016. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. 855–864.

[3] Srijan Kumar, Bryan Hooi, Disha Makhija, Mohit Kumar, Christos Faloutsos, and V. S. Subrahmanian. 2018. REV2: Fraudulent User Prediction in Rating Platforms. *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (2018).

[4] Yixin Liu, Shirui Pan, Yu Guang Wang, Fei Xiong, Liang Wang, Qingfeng Chen, and Vincent CS Lee. 2021. Anomaly detection in dynamic graphs via transformer. *IEEE Transactions on Knowledge and Data Engineering* (2021).

[5] Tore Opsahl and Pietro Panzarasa. 2009. Clustering in weighted networks. *Social networks* 31, 2 (2009), 155–163.

[6] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. 2014. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. 701–710.

[7] Emanuele Rossi, Ben Chamberlain, Fabrizio Frasca, Davide Eynard, Federico Monti, and Michael Bronstein. 2020. Temporal graph networks for deep learning on dynamic graphs. *arXiv preprint arXiv:2006.10637* (2020).

[8] Ryan Rossi and Nesreen Ahmed. 2015. The network data repository with interactive graph analytics and visualization. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 29.

[9] Ulrike Von Luxburg. 2007. A tutorial on spectral clustering. *Statistics and computing* 17 (2007), 395–416.

[10] Wenchao Yu, Wei Cheng, Charu C Aggarwal, Kai Zhang, Haifeng Chen, and Wei Wang. 2018. Netwalk: A flexible deep embedding approach for anomaly detection in dynamic networks. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 2672–2681.

[11] Li Zheng, Zhenpeng Li, Jian Li, Zhao Li, and Jun Gao. 2019. AddGraph: Anomaly Detection in Dynamic Graph Using Attention-based Temporal GCN.. In *IJCAI*, Vol. 3. 7.

[1] http://konect.cc/networks/opsahl-ucsocial
[2] http://networkrepository.com/email-dnc
[3] http://snap.stanford.edu/data/soc-sign-bitcoin-otc