

Towards Robust Dynamic Network Embedding

Chengbin Hou^{1,2} and Ke Tang¹

¹Guangdong Provincial Key Laboratory of Brain-Inspired Intelligent Computation, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China

²School of Computer Science, University of Birmingham, Birmingham, United Kingdom
 chengbin.hou10@foxmail.com, tangk3@sustech.edu.cn

Abstract

Dynamic Network Embedding (DNE) has recently drawn much attention due to the dynamic nature of many real-world networks. Comparing to a static network, a dynamic network has a unique character called the degree of changes, which can be defined as the average number of the changed edges between consecutive snapshots spanning a dynamic network. The degree of changes could be quite different even for the dynamic networks generated from the same dataset. It is natural to ask whether existing DNE methods are effective and robust w.r.t. the degree of changes. Towards robust DNE, we suggest two important scenarios. One is to investigate the robustness w.r.t. different slicing settings that are used to generate different dynamic networks with different degree of changes, while another focuses more on the robustness w.r.t. different number of changed edges over timesteps.

1 Introduction

A network or graph has been widely used to represent a complex system in sociology, biology, the Internet, etc. Network Embedding that embeds a static network into a low dimensional vector space [Cui *et al.*, 2019], as a kind of topological feature extraction, has become a well-studied technique.

Recently, the research interest has moved to dynamic networks, as real-world networks are often dynamic by nature [Barros *et al.*, 2021; Cui *et al.*, 2019]. As a result, *Dynamic Network Embedding* (DNE) has drawn much attention. It aims to embed a dynamic network into a low dimensional vector space promptly at each timestep and meanwhile preserve some network topologies. A dynamic network \mathcal{G} can be represented as a sequence of snapshots $\{G^0, G^1, \dots, G^t, \dots\}$ taken at each timestep. We could define the *degree of changes* as the average number of the changed edges between consecutive snapshots spanning a dynamic network.

There are two straightforward ways to generate a dynamic network as shown in Figure 1. First, the dynamic network \mathcal{G}_1 is generated by slicing with a fixed number of edges [Du *et al.*, 2018; Zhang *et al.*, 2018]. Second, the dynamic network \mathcal{G}_2 is generated by slicing with a fixed time interval [Hou *et al.*, 2020; Singer *et al.*, 2019]. For the both cases, the degree

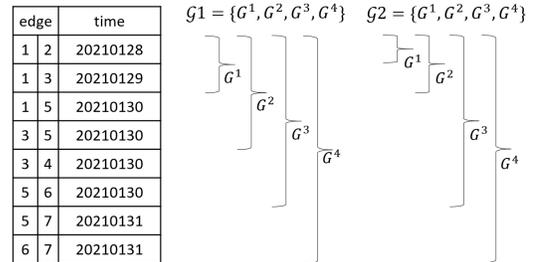


Figure 1: A dataset generates two dynamic networks \mathcal{G}_1 and \mathcal{G}_2 by slicing with a fixed number of edges and a fixed time interval. The degree of changes for \mathcal{G}_1 and \mathcal{G}_2 is 2 and $(1+4+2)/3$ respectively.

of changes could be quite different even for the dynamic networks generated from the same dataset. According to these two cases, we discuss the following two scenarios.

- The choice of a fixed number of edges for slicing might be difficult to justify, because of lacking obvious real-world significance. One may thus adopt different slicing settings. Are existing DNE methods effective and robust w.r.t. the different slicing settings that are used to generate different dynamic networks with different degree of changes from the same dataset?
- It might be easier to choose a fixed time interval for slicing, e.g., one day, owing to the obvious connection to real-world time. However, the number of changed edges might vary a lot over timesteps, due to different rate of streaming edges over time. Are existing DNE methods effective and robust w.r.t. different degree of changes while slicing through different slicing settings? More importantly, are they effective and robust w.r.t. different number of changed edges over timesteps?

We found that existing DNE methods are not robust in the former scenario. A robust DNE method via ensembles was therefore proposed to address this issue [Hou *et al.*, 2021]. For the latter scenario, it is still an open problem.

2 Our Research

The degree of changes is a unique character of dynamic networks compared to static networks. It could be affected by various factors such as the slicing setting and the rate of streaming edges. It is natural to ask if existing DNE methods

are effective and robust w.r.t. the degree of changes, which motivates us to conduct the following research works.

For the *finished work*, we started with the simpler scenario where the number of changed edges is identical over timesteps. A dynamic network is generated by slicing with a fixed number of edges as \mathcal{G}_1 illustrated in Figure 1. In this scenario, each dataset generates several different dynamic networks using different choices of a fixed number of edges. We then investigated whether existing DNE methods can robustly obtain a good performance over these different dynamic networks with different degree of changes.

We compared six existing DNE methods on five real-world datasets, and evaluated the resulting embeddings by three types of downstream tasks. Comprehensive experiments demonstrated that the existing DNE methods are not robust in this scenario. To address this issue, we proposed a robust DNE method which has been shown to be effective and robust. Concretely, the proposed method employs an incremental Skip-Gram embedding model as the base learner. Multiple base learners with a novel diversity enhancement strategy then form the ensembles to capture different levels of local-global topology, so as to learn node embeddings at each timestep [Hou *et al.*, 2021]. The source code is available at <https://github.com/houchengbin/SG-EDNE>

For the *future work*, we would generate a dynamic network by slicing with a fixed time interval. This scenario is more challenging, because the number of changed edges may vary a lot over timesteps, e.g., much more friendships than usual would be built in a social network during a social event. This scenario tries to study the performance of a DNE method at each timestep, while the previous one focuses more on the average performance over all timesteps. Ideally, a desirable DNE method should be effective and robust w.r.t. not only different degree of changes by different slicing settings, but also different number of changed edges over timesteps.

We would benchmark existing DNE methods under this new scenario. After that, a more robust DNE method for this scenario would be proposed. Specifically, we would again exploit the notion of ensembles, while other details such as the choice of base learners and how base learners cooperate with each other are not decided yet.

It is promising to employ ensembles to improve the model robustness, since *ensembles* could provide a *redundant design* to alleviate the *uncertainties* in generating dynamic networks or in dynamic networks themselves. Furthermore, ensembles have been successfully applied in traditional streaming data [Sun *et al.*, 2018] and graph streaming data [Hou *et al.*, 2021] to improve the model robustness. If the aforementioned desirable DNE method is achieved, we would not only improve the average performance over all timesteps, but also apply it in industrial applications such as anomaly detection [Zheng *et al.*, 2019] to avoid occasionally seriously degraded performance at some snapshots or timesteps.

3 Contributions

The contributions are as follows. First, we investigate the robustness of DNE w.r.t. the degree of changes, which is the first time to be formally studied. This is important to DNE

community, as robustness itself is a critical topic in dynamic environment. Second, we suggest two important scenarios. One is to investigate the robustness w.r.t. different slicing settings that are used to generate different dynamic networks with different degree of changes, while another would focus more on the robustness w.r.t. different number of changed edges over timesteps. Third, we have proposed a robust DNE method based on ensembles (for the first time) for the former scenario, and would propose another robust DNE method also based on ensembles for the latter scenario.

4 Conclusion

Though we have suggested the two important scenarios, there could be other complicated scenarios by other preprocessing approaches on the original dataset. It is impossible to list all. In the future, we should rethink of how to represent a dynamic network in a more proper way to reduce the uncertainties in generating dynamic networks, so that we can then propose a robust DNE method for broader applications.

References

- [Barros *et al.*, 2021] Claudio DT Barros, Matheus RF Mendonça, Alex B Vieira, and Artur Ziviani. A survey on embedding dynamic graphs. *arXiv preprint arXiv:2101.01229*, 2021.
- [Cui *et al.*, 2019] Peng Cui, Xiao Wang, Jian Pei, and Wenwu Zhu. A survey on network embedding. *IEEE Transactions on Knowledge and Data Engineering*, 31(5):833–852, 2019.
- [Du *et al.*, 2018] Lun Du, Yun Wang, Guojie Song, Zhicong Lu, and Junshan Wang. Dynamic network embedding: An extended approach for skip-gram based network embedding. In *IJCAI*, pages 2086–2092, 2018.
- [Hou *et al.*, 2020] Chengbin Hou, Han Zhang, Shan He, and Ke Tang. GloDyNe: Global topology preserving dynamic network embedding. *IEEE Transactions on Knowledge and Data Engineering*, pages 1–1, 2020.
- [Hou *et al.*, 2021] Chengbin Hou, Guoji Fu, Peng Yang, Shan He, and Ke Tang. Robust dynamic network embedding via ensembles. *arXiv preprint arXiv:2105.14557*, 2021.
- [Singer *et al.*, 2019] Uriel Singer, Ido Guy, and Kira Radinsky. Node embedding over temporal graphs. In *IJCAI*, pages 4605–4612, 2019.
- [Sun *et al.*, 2018] Yu Sun, Ke Tang, Zexuan Zhu, and Xin Yao. Concept drift adaptation by exploiting historical knowledge. *IEEE transactions on neural networks and learning systems*, 29(10):4822–4832, 2018.
- [Zhang *et al.*, 2018] Ziwei Zhang, Peng Cui, Jian Pei, Xiao Wang, and Wenwu Zhu. TIMERS: error-bounded SVD restart on dynamic networks. In *AAAI*, pages 224–231, 2018.
- [Zheng *et al.*, 2019] Li Zheng, Zhenpeng Li, Jian Li, Zhao Li, and Jun Gao. Addgraph: Anomaly detection in dynamic graph using attention-based temporal gcN. In *IJCAI*, pages 4419–4425, 2019.