Building the Foundations of Temporal Graph Learning: Visualization, Evaluation, and Applications[*]

Reihaneh Rabbany[1,2,*,†]

[1]*School of Computer Science, McGill University*
[2]*Mila, Quebec Artificial Intelligence Institute*

Temporal graphs provide a natural framework for modeling complex, evolving systems, yet their analysis and predictive modeling remain challenging due to temporal dependencies, structural heterogeneity, and diverse evaluation protocols. In this talk, I will present a cohesive set of contributions from our group aimed at advancing both methodology and evaluation in temporal graph learning.

We begin with TGX [1], a modular toolkit for temporal graph exploration that supports scalable preprocessing, rich statistical analyses, and interactive visualizations of evolving network structure. These visualizations have been instrumental for hypothesis generation and error analysis, and builds upon the insights introduced in our Towards Better Evaluation for Dynamic Link Prediction work [2], where we argued for more transparent and interpretable model assessment.

Building on this foundation, we turn to evaluation and benchmarking. I will discuss the Temporal Graph Benchmark (TGB) [3] and its expanded version TGB 2 [4], part of an ongoing open-science initiative providing standardized datasets, evaluation pipelines, and leaderboards for a wide range of temporal graph tasks. These resources aim to enable reproducible and comprehensive comparisons across methods, while lowering barriers for entry into temporal graph research.

Finally, I will discuss applications and recent trends on temporal graphs. These include change point detection in dynamic and multi-view networks using spectral methods (LAD/MultiLAD [5]), epidemic forecasting through sparse static approximations of dynamic contact networks (EdgeMST/DegMST [6]), and our recent exploration of large language models for temporal graph tasks [7], which examines their ability to reason over evolving network structures. Across these efforts, a common theme emerges: the integration of transparent evaluation, rich visualization, and high-quality benchmarks to move the field from isolated experiments towards cumulative, transferable knowledge in temporal graph learning

References

[1] R. Shirzadkhani, S. Huang, E. Kooshafar, R. Rabbany, F. Poursafaei, Temporal graph analysis with tgx, in: Proceedings of the 17th ACM International Conference on Web Search and Data Mining (WSDM '24), ACM, Merida, Mexico, 2024. URL: https://github.com/ComplexData-MILA/TGX. doi:10.1145/3616855.3635694, short paper.

[2] F. Poursafaei, S. Huang, K. Pelrine, R. Rabbany, Towards better evaluation for dynamic link prediction, in: NeurIPS 2022 Datasets and Benchmarks Track, 2022. URL: https://github.com/fpour/DGB.git, introduces TEA/TET visualizations, EdgeBank, and harder negative sampling.

[3] S. Huang, F. Poursafaei, J. Danovitch, M. Fey, W. Hu, E. Rossi, J. Leskovec, M. Bronstein, G. Rabusseau, R. Rabbany, Temporal graph benchmark for machine learning on temporal graphs, in: NeurIPS 2023 Datasets and Benchmarks Track, 2023. URL: https://tgb.complexdatalab.com/, large, diverse datasets and standardized evaluation for temporal graphs.

[4] J. Gastinger, S. Huang, M. Galkin, E. Loghmani, A. Parviz, F. Poursafaei, J. Danovitch, E. Rossi, I. Koutis, H. Stuckenschmidt, R. Rabbany, G. Rabusseau, Tgb 2.0: A benchmark for learning on temporal knowledge graphs and heterogeneous graphs, in: NeurIPS 2024 Datasets and Benchmarks Track, 2024. Eight new multi-relational temporal datasets and a reproducible evaluation pipeline.

[5] S. Huang, S. Coulombe, Y. Hitti, R. Rabbany, G. Rabusseau, Laplacian change point detection for single and

multi-view dynamic graphs, ACM Transactions on Knowledge Discovery from Data 18 (2024). doi:`10.1145/3631609`, lAD/MultiLAD.

[6] R. Shirzadkhani, S. Huang, A. Leung, R. Rabbany, Static graph approximations of dynamic contact networks for epidemic forecasting, Scientific Reports 14 (2024). doi:`10.1038/s41598-024-62271-0`, proposes EdgeMST and DegMST.

[7] S. Huang, A. Parviz, E. Kondrup, Z. Yang, Z. Ding, M. Bronstein, R. Rabbany, G. Rabusseau, Are large language models good temporal graph learners?, arXiv preprint arXiv:2506.05393, 2025. URL: https://arxiv.org/abs/2506.05393.