Topology-Aware Visual Analysis of GNN Explanations in Dynamic Graphs

Juanpablo Heredia[1] and Jorge Poco[1]

[1]Fundação Getulio Vargas, EMAp, Rio de Janeiro, Brazil

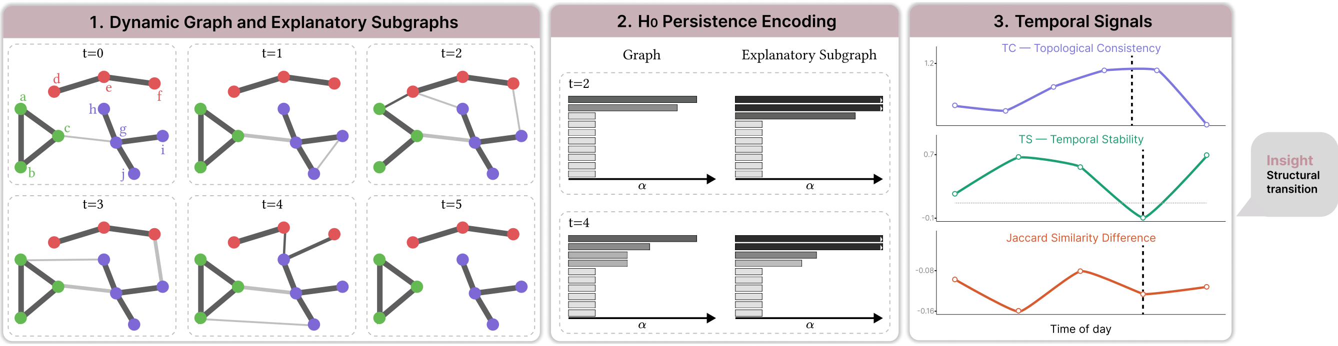

Figure 1: *Overview on a controlled toy benchmark. (1) A 10-vertex dynamic graph $\{G_t\}$ with three node classes over six timesteps; dark edges denote the explanation E_t and light edges the remaining edges in G_t. Edge thickness encodes connection strength. (2) H_0 persistence barcodes for G_t and E_t at two timesteps. (3) Temporal signals: topological consistency $TC(t)$, signed temporal stability $TS_G - TS_E$, and Jaccard similarity difference. The dashed marker at $3 \rightarrow 4$ indicates a structural reorganization in E_t.*

Abstract

Graph neural networks (GNNs) are increasingly applied to dynamic graphs, where explanations should remain coherent as graph structure evolves. Existing fidelity metrics provide scalar summaries and do not capture structural alignment or temporal behavior. This work introduces a persistent-homology-based visualization approach that treats persistence-distance distortions between data graphs and explanations as temporal signals. We introduce topological consistency *(TC), measuring per-step structural deviation, and* temporal stability *(TS), capturing and attributing changes over time. We demonstrate the approach on a controlled toy example and a high-school contact network, showing that topology-aware signals reveal structural behaviors not captured by set-based similarity measures.*

CCS Concepts

• ***Networks*** $\rightarrow$ ***Network algorithms;*** • ***Human-centered computing*** $\rightarrow$ ***Visualization systems and tools;***

1. Introduction

GNN explainability methods such as GNNExplainer [YBY*19], PGExplainer [LCX*20], and attention-based approaches [VCC*18] produce subgraphs intended to capture the structure most relevant for a prediction. Standard evaluation reduces explanation quality to scalar fidelity scores [YHS*19, AYZ*22, AQLZ23], offering limited insight into how explanations relate structurally to the underlying graph or evolve. This is particularly relevant for dynamic graphs, where topology changes [ZYW25] and explanations should remain faithful and structurally coherent.

From a visualization perspective, capturing structural evolution remains challenging. Surveys highlight the difficulty of representing evolving relational structure [BBDW17, FFKS21], while visual analytics approaches support exploring large temporal networks, temporal resolutions, and dynamic graph patterns [vdEHBvW16, LPP*23, TPL*25, JSJ*24]. In the GNN setting, visual analytics tools support inspection, error diagnosis, and interactive analysis of graph-based models and spatiotemporal graph behaviors [LWBM22, JWW*23, SM25, HERMCP25]. Yet explanation behavior over time remains mostly scalar or set-based, leaving topology unexamined.

We use persistent homology (PH) [EH10] to compare data graphs and explanation subgraphs via distances between persistence diagrams [HWSR18, CSEH07]. Since G_t and $E_t \subseteq G_t$ are not independent objects but a graph and its edge sub-selection, we treat persistence distances as explanation-induced *distortion signals*, bounded by persistence stability [CSEH07, CdSGO16].

We focus on connected components (H_0) as a first-order descriptor. Explanation subgraphs are typically sparse and fragmented; in preliminary experiments, higher-order features (e.g., H_1) produced near-constant signals. Persistence distances are interpreted as *temporal signals* [HWSR18, MMK19] capturing how the data–explanation structural relationship evolves. **Topological consistency (TC)** measures per-step structural deviation. **Temporal stability (TS)** captures when changes occur and attributes them. **Jaccard similarity (JS)** serves as a set-based baseline. Together, these signals link temporal peaks to local graph and explanation changes. We illustrate the approach on a controlled toy example and a high-school contact network [FB14]. The pipeline currently operates on pairwise graphs; higher-order representations [FFKS21, BCI*20, PET*14] are a natural direction.

2. Approach

Let $G = \{G_t\}_{t=0}^{T-1}$ be a dynamic graph on a fixed vertex set $\mathcal{V}$, with explanation subgraphs $E_t \subseteq G_t$ at each timestep. We compute H_0 persistence diagrams $D_0(G_t)$ and $D_0(E_t)$ under a weighted shortest-path filtration [EH10, HWSR18], and compare them using Wasserstein-2 or Bottleneck distances [CSEH07]. Here, E_t is an edge selection within G_t, so distances between $D_0(G_t)$ and $D_0(E_t)$ capture induced structural distortion and are bounded by the stability theorem [CSEH07, CdSGO16].

TC is the persistence distance between $D_0(G_t)$ and $D_0(E_t)$, measuring structural deviation of the explanation from the data. **TS** attributes change between timesteps as $\mathrm{TS}_G - \mathrm{TS}_E$, where TS$G$ and TSE are persistence distances between consecutive G_t and E_t diagrams. Positive values indicate that change originates primarily from the data graph; negative values, from the explanation. **JS** over consecutive explanation edge sets [PCW25] ($|\mathcal{E}_t \cap \mathcal{E}_{t+1}|/|\mathcal{E}_t \cup \mathcal{E}_{t+1}|$) is a set-based baseline, measuring which edges change, not how they are organized. Signals are shown in two views: a temporal line chart (TC, TS, JS) and a graph-overlap view of G_t and E_t at selected timesteps.

3. Results

Toy example. A 10-vertex dynamic graph with three node classes over six timesteps (Fig. 1) is constructed with controlled disruptions at t=2 and t=4. TC shows a pronounced peak at t=4, capturing the main structural reorganization, while TS detects both events, attributing the change at t=2 to the data graph and a stronger transition at t=3 $\rightarrow$ 4 to the explanation, remaining near zero otherwise and indicating overall stability. In contrast, JS remains oscillatory, weakly highlighting transitions around t=1 $\rightarrow$ 2 and t=2 $\rightarrow$ 3 without distinguishing their structural origin.

High-school contact network. We train a per-snapshot GAT [VCC*18] encoder with a GRU temporal head on a one-week proximity network [FB14], following [WQF*22]; explanations retain edges above the 80th percentile of mean attention. Figure 2 shows TC, TS, and JS for the first school day. Explanations remain sparse and structurally stable in H_0, while TS shows a clear spike at t=5 $\rightarrow$ 6, indicating a significant change in the data graph (Fig. 2E–F). This transition corresponds to the splitting of a large connected component, which is not reflected by JS, as edge-set similarity remains high and is insensitive to this topological change. This highlights TS's diagnostic role: separating explanation instability from genuine graph changes.

Figure 2: *Temporal profiles for the first school day on the high-school contact network.* ***(A)*** *TC;* ***(B)*** *signed TS;* ***(C)*** *JS;* ***(D)*** *disaggregated TS with* TS_G *(green) and* TS_E *(orange). The TS spike at* t=5 $\rightarrow$ 6 *is driven by* G_t *splitting from one to two connected components (E–F), while the explanation remains fragmented. JS does not distinguish this transition.*

4. Discussion and Ongoing Work

Framing persistence distances as distortion signals grounded in the stability theorem provides a principled lens for asking *when* and *why* explanation structure changes. The signed $\mathrm{TS}_G - \mathrm{TS}_E$ disaggregation enables *attribution* of detected transitions to data or explanation—a distinction JS cannot make, since it operates exclusively on explanation edge sets. We intentionally begin with H_0 because explanation subgraphs are often sparse and fragmented, making connected-component structure a robust first-order signal. The pipeline remains restricted to pairwise graphs, leaving cyclic and group-level structure out of reach. Many dynamic interaction datasets, including the one used here, exhibit genuine higher-order interactions [FFKS21, BCI*20, PET*14]; lifting E_t and G_t to simplicial complexes or hypergraphs is a priority direction. We position topology not as a replacement for existing metrics, but as a complementary tool that exposes structural behaviors invisible to scalar or set-based evaluation.

Acknowledgments

The authors acknowledge using LLM-based tools (Claude and Grammarly) for text polishing. This work was supported by the National Council for Scientific and Technological Development (CNPq, #311144/2022-5), São Paulo Research Foundation (FAPESP, #2021/07012-0 and #2023/04868-7), and the School of Applied Mathematics at Fundação Getulio Vargas.

Proceedings published by Eurographics - The European Association for Computer Graphics.

References

[AQLZ23] Agarwal C., Queen O., Lakkaraju H., Zitnik M.: Evaluating explainability for graph neural networks. *Scientific Data 10*, 1 (2023), 144. doi:10.1038/s41597-023-01974-x. 1

[AYZ*22] Amara K., Ying Z., Zhang Z., Han Z., Zhao Y., Shan Y., Brandes U., Schemm S., Zhang C.: GraphFramEx: Towards systematic evaluation of explainability methods for graph neural networks. In *Proceedings of the First Learning on Graphs Conference (LoG)* (2022), vol. 198 of *Proceedings of Machine Learning Research*, PMLR, pp. 44:1–44:23. 1

[BBDW17] Beck F., Burch M., Diehl S., Weiskopf D.: A taxonomy and survey of dynamic graph visualization. *Computer Graphics Forum 36*, 1 (2017), 133–159. doi:10.1111/cgf.12791. 1

[BCI*20] Battiston F., Cencetti G., Iacopini I., Latora V., Lucas M., Patania A., Young J.-G., Petri G.: Networks beyond pairwise interactions: Structure and dynamics. *Physics Reports 874* (2020), 1–92. doi:10.1016/j.physrep.2020.05.004. 2

[CdSGO16] Chazal F., de Silva V., Glisse M., Oudot S.: Persistence modules. In *The Structure and Stability of Persistence Modules*. Springer International Publishing, 2016, pp. 15–29. doi:10.1007/978-3-319-42545-0_2. 2

[CSEH07] Cohen-Steiner D., Edelsbrunner H., Harer J.: Stability of persistence diagrams. *Discrete & Computational Geometry 37*, 1 (2007), 103–120. doi:10.1007/s00454-006-1276-5. 2

[EH10] Edelsbrunner H., Harer J.: *Computational Topology: An Introduction*. American Mathematical Society, 2010. doi:10.1090/mbk/069. 1, 2

[FB14] Fournet J., Barrat A.: Contact patterns among high school students. *PLOS ONE 9*, 9 (2014), e107878. doi:10.1371/journal.pone.0107878. 2

[FFKS21] Fischer M. T., Frings A., Keim D. A., Seebacher D.: Towards a survey on static and dynamic hypergraph visualizations. In *IEEE Visualization Conference (VIS)* (2021), IEEE, pp. 81–85. doi:10.1109/VIS49827.2021.9623305. 1, 2

[HERMCP25] Heredia J., Estrada-Rayme L., Matos-Cangalaya J., Poco J.: Interactive exploration and explanation of spatio-temporal anomalies with graph-LLM integration. In *38th Conference on Graphics, Patterns and Images (SIBGRAPI)* (2025), IEEE, pp. 1–6. doi:10.1109/SIBGRAPI67909.2025.11223398. 1

[HWSR18] Hajij M., Wang B., Scheidegger C., Rosen P.: Visual detection of structural changes in time-varying graphs using persistent homology. In *2018 IEEE Pacific Visualization Symposium (PacificVis)* (2018), IEEE, pp. 125–134. doi:10.1109/PacificVis.2018.00024. 2

[JSJ*24] Jung S., Shin D., Jeon H., Choe K., Seo J.: MoNetExplorer: A visual analytics system for analyzing dynamic networks with temporal network motifs. *IEEE Transactions on Visualization and Computer Graphics 30*, 10 (2024), 6725–6739. doi:10.1109/TVCG.2023.3337396. 1

[JWW*23] Jin Z., Wang Y., Wang Q., Ming Y., Ma T., Qu H.: GNNLens: A visual analytics approach for prediction error diagnosis of graph neural networks. *IEEE Transactions on Visualization and Computer Graphics 29*, 6 (2023), 3024–3038. doi:10.1109/TVCG.2022.3148107. 1

[LCX*20] Luo D., Cheng W., Xu D., Yu W., Zong B., Chen H., Zhang X.: Parameterized explainer for graph neural network. In *Advances in Neural Information Processing Systems (NeurIPS)* (2020), vol. 33, Curran Associates, Inc., pp. 19620–19631. 1

[LPP*23] Linhares C. D. G., Ponciano J., Pedro D., Rocha L., Traina A., Poco J.: Largenetvis: Visual exploration of large temporal networks based on community taxonomies. *IEEE Transactions on Visualization and Computer Graphics* (2023). doi:10.1109/TVCG.2022.3209477. 1

[LWBM22] Liu Z., Wang Y., Bernard J., Munzner T.: Visualizing graph neural networks with CorGIE: Corresponding a graph to its embedding. *IEEE Transactions on Visualization and Computer Graphics 28*, 6 (2022), 2500–2516. doi:10.1109/TVCG.2022.3148197. 1

[MMK19] Myers A., Munch E., Khasawneh F. A.: Persistent homology of complex networks for dynamic state detection. *Physical Review E 100*, 2 (2019), 022314. doi:10.1103/PhysRevE.100.022314. 2

[PCW25] Peters H. A. B., Ceria A., Wang H.: Higher-order temporal network prediction and interpretation. *PLOS ONE 20*, 5 (2025), e0323753. doi:10.1371/journal.pone.0323753. 2

[PET*14] Petri G., Expert P., Turkheimer F., Carhart-Harris R., Nutt D., Hellyer P. J., Vaccarino F.: Homological scaffolds of brain functional networks. *Journal of The Royal Society Interface 11*, 101 (2014), 20140873. doi:10.1098/rsif.2014.0873. 2

[SM25] Singh T. C., Mukherjea S.: InteractiveGNNExplainer: A visual analytics framework for multi-faceted understanding and probing of graph neural network predictions. In *29th International Conference Information Visualisation (IV)* (2025), IEEE, pp. 187–192. 1

[TPL*25] Tinarrage R., Ponciano J. R., Linhares C. D. G., Traina A. J. M., Poco J.: Zigzagnetvis: Suggesting temporal resolutions for graph visualization using zigzag persistence. *IEEE Transactions on Visualization and Computer Graphics* (2025). doi:10.1109/TVCG.2025.3528197. 1

[VCC*18] Veličković P., Cucurull G., Casanova A., Romero A., Liò P., Bengio Y.: Graph attention networks. In *International Conference on Learning Representations (ICLR)* (2018). 1, 2

[vdEHBvW16] van den Elzen S., Holten D., Blaas J., van Wijk J. J.: Reducing snapshots to points: A visual analytics approach to dynamic network exploration. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 1–10. doi:10.1109/TVCG.2015.2468078. 1

[WQF*22] Wang S., Qiao L., Fang W., Jing G., Sheng V. S., Zhang Y.: Air pollution prediction via graph attention network and gated recurrent unit. *Computers, Materials & Continua 73*, 1 (2022), 673–687. doi:10.32604/cmc.2022.028411. 2

[YBY*19] Ying Z., Bourgeois D., You J., Zitnik M., Leskovec J.: GNNExplainer: Generating explanations for graph neural networks. In *Advances in Neural Information Processing Systems (NeurIPS)* (2019), vol. 32, Curran Associates, Inc., pp. 9240–9251. 1

[YHS*19] Yeh C.-K., Hsieh C.-Y., Suggala A. S., Inouye D. I., Ravikumar P.: On the (in)fidelity and sensitivity of explanations. In *Advances in Neural Information Processing Systems (NeurIPS)* (2019), vol. 32, Curran Associates, Inc., pp. 10967–10978. 1

[ZYW25] Zheng Y., Yi L., Wei Z.: A survey of dynamic graph neural networks. *Frontiers of Computer Science 19*, 6 (2025), 196323. doi:10.1007/s11704-024-3853-2. 1