

Cyberattack Evolution Modeling in Smart Grids: A Hierarchical Temporal Graph Transformer for Detection, Localization, and Severity Estimation

Sahibzada Muhammad Ali^{*1} and Zahid Ullah²

¹COMSATS University Islamabad, Abbottabad Campus, Pakistan

²Politecnico di Milano, 20133, Italy

October 1, 2026

Abstract

Cyberattacks against modern power grids are not purely point anomalies: their observable signatures can be spatially coordinated, persistent over time, partially masked, and heterogeneous in severity. Treating every event as a binary label discards structure that can be useful for localization and operational response. This paper formulates cyberattack evolution as a hierarchical spatiotemporal learning problem and proposes a Causal Hierarchical Temporal Graph Transformer (CHTGT) for joint attack detection, source localization, and severity estimation. The model maps PMU/SCADA telemetry onto a 300-bus electrical graph, performs bus-level attention, aggregates representations by four operating districts, applies causal temporal attention over 12-step windows, and feeds shared representations to detection, localization, and severity heads. A strict information-availability contract separates observable inference inputs from attack metadata used only for supervision and evaluation. The available benchmark contains 604,800 fixed bus-timestamp records, 30,240 attack records, three attack regimes, and 2,016 temporal states. Existing independently re-computed backbone evidence on a 121,200-sample held-out set gives 0.9940 accuracy, 0.8800 positive-class F1, 0.9941 AUROC, 0.9732 AUPRC, and 0.8724 MCC, while source localization reaches 0.509/0.514/0.526 for Top-1/3/5. An associated lightweight model shows severe validation-to-zero-day degradation, motivating explicit attack-evolution evaluation. These results are reported as auditable backbone evidence rather than fabricated CHTGT gains; severity estimation and component-level ablations remain the decisive experiments.

Keywords: smart-grid cybersecurity; cyberattack evolution; graph transformer; causal temporal learning; attack localization; severity estimation; power systems

1 Introduction

The increasing deployment of PMUs, SCADA systems, renewable generation, storage, and communication-dependent control has made the power grid a tightly coupled cyber-physical system. Measurement integrity is consequently a system-level security concern. False-data attacks can manipulate measurements so that downstream state estimation and monitoring receive misleading information while conventional bad-data tests may be bypassed under carefully constructed conditions [2, 14, 15].

^{*}Corresponding author: hallianali@cuiatd.edu.pk

A second difficulty is that attack behavior evolves. An adversarial perturbation may affect one or many buses, persist across several samples, become less observable through masking, and vary in magnitude. A binary target $y_t \in \{0, 1\}$ answers only whether an event exists; it does not describe where it is concentrated, how long it persists, or how severe it is. Recent reviews identify precise localization as a persistent challenge in smart-grid security [14].

Graph learning is attractive because the network itself defines a structured domain. GCNs learn topology-aware representations [4]; GraphSAGE provides inductive neighborhood aggregation [5]; GAT learns differentiated neighborhood weights [6]; and spatiotemporal graph models jointly capture spatial and temporal dependencies [7]. Transformer attention provides flexible long-range sequence modeling [8]. Recent power-system studies have extended these ideas to spatiotemporal detection and localization [18, 16, 19, 22].

The gap addressed here is not another generic graph encoder. It is explicit modeling of *attack evolution*: spatial support, persistence, masking, severity, and source localization are treated as related outputs of one temporal graph representation. The hypothesis is that a model representing how an attack evolves across buses and time can provide more actionable information than a binary detector alone.

The contributions are: (i) a hierarchical bus–district–temporal architecture for joint detection, localization, and severity estimation; (ii) a causal information-availability contract preventing attack metadata from entering inference; (iii) a mathematical formulation linking topology, temporal causality and multi-task prediction; and (iv) an auditable analysis of the available IEEE-300-bus evidence.

2 Related Work and Research Gap

Graph-based security learning has progressed from static topology encoders toward spatiotemporal architectures. GCN, GraphSAGE, and GAT provide complementary mechanisms for neighborhood aggregation and attention [4, 5, 6]; STGCN extends graph learning across time [7]; and Transformer attention provides a general sequence mechanism [8]. Recent power-system studies further demonstrate graph-temporal detection and localization, including a GCN–TCN detector [18], a spatiotemporal graph-wavelet localization method [16], a graph/diffusion localization and recovery framework [19], and graph-Fourier attack identification and localization [22].

These studies establish that topology and temporal structure are valuable, but their primary outputs remain event detection, source localization, or recovery. The present work instead treats *attack evolution* as the modeling object: a single latent representation is expected to encode where an event is distributed, how long it persists, how observable it remains, and how severe it becomes. This is a different emphasis from simply adding another graph backbone.

The proposed model also differs from physics-informed learning. Physics-derived residuals are used as observable context rather than as a claim that every residual is malicious. This distinction is important because renewable ramps, load transitions, and measurement noise can produce legitimate residual variation. The broader literature on smart-grid cyber resilience likewise emphasizes the need to combine statistical detection with operational context and explicit resilience mechanisms [15]. Recent review work on graph reinforcement learning similarly emphasizes that topology-aware intelligence remains largely at the proof-of-concept stage for real-world grid deployment, reinforcing the need for controlled, leakage-safe evaluation rather than architecture-only claims [23].

The distinction is therefore not that existing methods are incapable of localization or auxiliary prediction. Rather, CHTGT makes *evolution across space, time, and severity* the central modeling target and evaluates all three outputs under one causal information contract.

Table 1: Positioning of the proposed attack-evolution model. “Explicit” describes the canonical design emphasis of each method family, not a capability that competing methods are fundamentally unable to implement.

Paradigm	Graph structure	Temporal evolution	Localization	Severity
GCN/GraphSAGE	Explicit	Not primary	Optional	Not primary
GAT	Explicit	Not primary	Optional	Not primary
STGCN	Explicit	Explicit	Optional	Not primary
Recent FDIA-GNN	Explicit	Often explicit	Often explicit	Usually secondary
CHTGT	Explicit	Causal	Primary	Primary

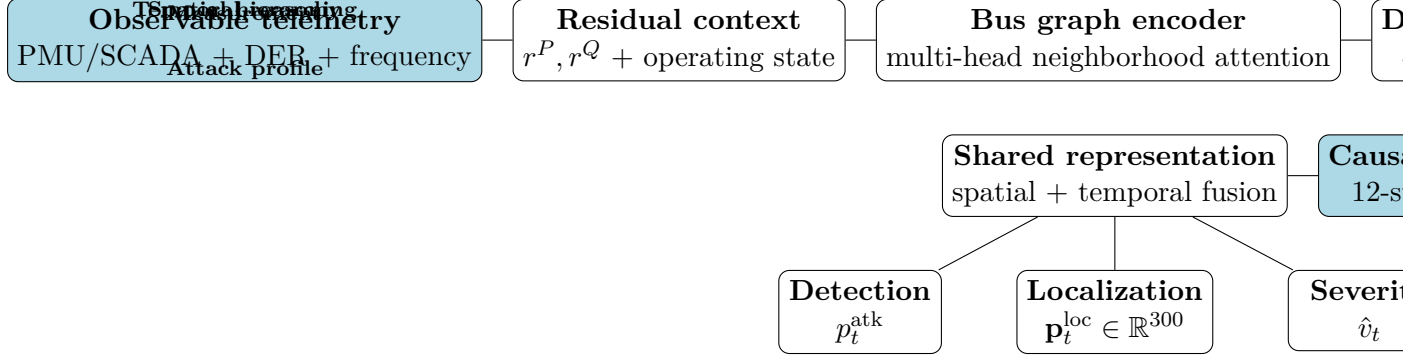


Figure 1: Proposed hierarchical temporal graph model. Observable telemetry and residual context are encoded on the electrical graph, compressed through four district states, processed with causal temporal attention, and mapped to detection, localization, and severity outputs.

3 Threat Model and Attack Evolution

We assume an attacker can alter selected PMU/SCADA measurements over one or more time steps and may know relevant topology or state-estimation structure. The attacker cannot directly modify the physical network, future measurements, or evaluation labels. The dataset represents partial, perfect, and masked perturbation regimes.

Let $\mathbf{x}_{i,t}$ be the observed vector at bus i . An attacked observation is

$$\mathbf{x}'_{i,t} = \mathbf{x}_{i,t} + \boldsymbol{\delta}_{i,t}. \quad (1)$$

We describe the event through

$$\mathcal{A}_t = (\mathcal{S}_t, \mathcal{T}_t, \mathcal{M}_t, \mathcal{V}_t), \quad (2)$$

where the terms denote spatial support, temporal persistence, masking/observability, and severity.

The inference contract is

$$\mathcal{X}_t^{\text{obs}} = \{X_{t-L+1:t}, R_{t-L+1:t}^P, R_{t-L+1:t}^Q, G\}, \quad (3)$$

whereas

$$\mathcal{Y}_t^{\text{atk}} = \{\mathcal{S}_t^*, \mathcal{T}_t^*, \mathcal{M}_t^*, \mathcal{V}_t^*\} \quad (4)$$

is used only as supervision, stratification, or evaluation. This distinction removes the leakage risk of supplying attacked-bus or ground-truth severity information at inference.

4 Hierarchical Temporal Graph Transformer

4.1 Physics residual context

For bus i , active- and reactive-power residuals are represented as

$$r_{i,t}^P = P_{i,t} - \hat{P}_{i,t}, \quad r_{i,t}^Q = Q_{i,t} - \hat{Q}_{i,t}. \quad (5)$$

The augmented state is

$$\tilde{\mathbf{x}}_{i,t} = [\mathbf{x}_{i,t} \| r_{i,t}^P \| r_{i,t}^Q]. \quad (6)$$

Residuals are continuous evidence, not attack labels. This is consistent with the broader principle of physics-informed learning [11].

4.2 Graph and district layers

For $G = (V, E)$, bus-level attention is

$$\mathbf{h}_{i,t}^G = \sigma \left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij,t} \mathbf{W} \tilde{\mathbf{x}}_{j,t} \right), \quad (7)$$

where

$$\alpha_{ij,t} = \frac{\exp(e_{ij,t})}{\sum_{k \in \mathcal{N}(i)} \exp(e_{ik,t})}. \quad (8)$$

District pooling maps the 300 bus representations into four district states:

$$\mathbf{h}_{d,t}^D = \text{Pool}\{\mathbf{h}_{i,t}^G : i \in d\}. \quad (9)$$

4.3 Causal temporal layer

For a window of length $L = 12$,

$$\mathbf{H}_t^C = \text{CausalAttn}(\mathbf{H}_{t-L+1:t}^D), \quad (10)$$

with

$$M_{ij} = \begin{cases} 0, & j \leq i \\ -\infty, & j > i. \end{cases} \quad (11)$$

The mask ensures that a decision at t cannot use future measurements.

4.4 Multi-task outputs

Detection is

$$p_t^{\text{atk}} = \sigma(\mathbf{w}_a^T \mathbf{h}_t^F + b_a), \quad (12)$$

localization is

$$\mathbf{p}_t^{\text{loc}} = \text{softmax}(\mathbf{W}_l \mathbf{h}_t^F + \mathbf{b}_l), \quad (13)$$

and severity is a bounded regression head

$$\hat{v}_t = \text{clip}(f_v(\mathbf{h}_t^F), 0, 1). \quad (14)$$

The conceptual objective is

$$\mathcal{L} = \mathcal{L}_{\text{det}} + \lambda_l \mathcal{L}_{\text{loc}} + \lambda_v \mathcal{L}_{\text{sev}} + \lambda_p \mathcal{L}_{\text{phys}}. \quad (15)$$

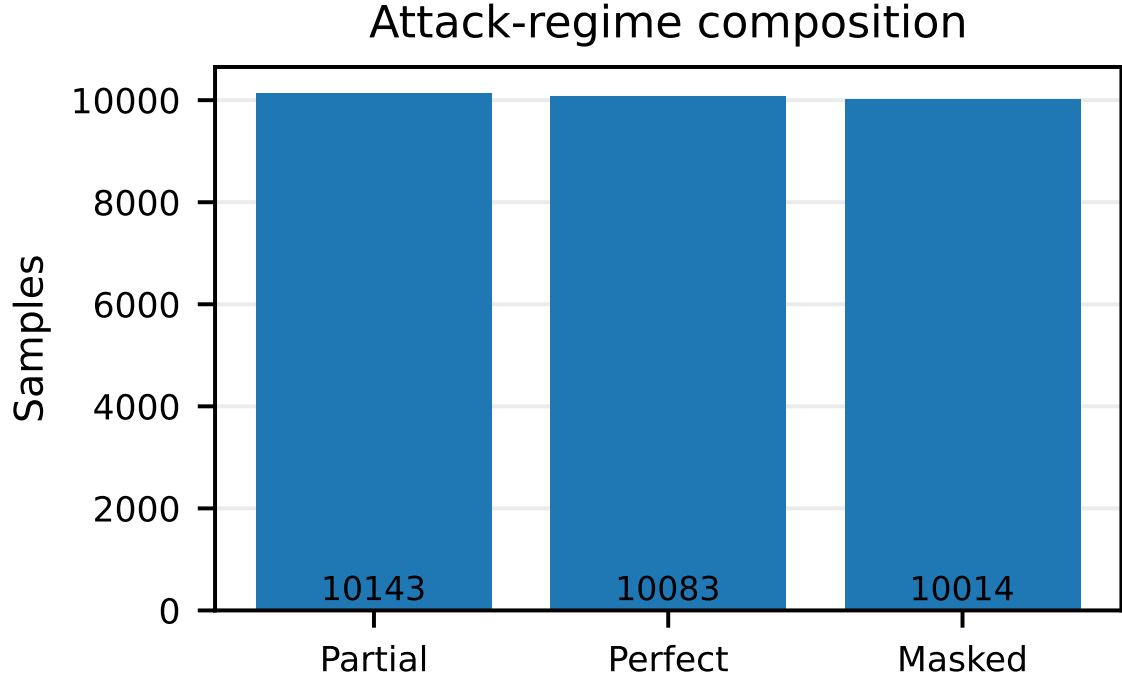


Figure 2: Attack-regime composition. The three regimes are nearly balanced, enabling regime-stratified analysis.

Table 2: Auditable dataset configuration.

Item	Value
Buses	300
Edges	409
Temporal states	2,016
Window	12
Rows	604,800
Benign / attack	574,560 / 30,240
Regimes	partial / perfect / masked
Severity	0.20–1.00
Severity mean $\pm$ SD	0.603 $\pm$ 0.231
Attacked buses/time	median 15; range 3–29
Districts	4

5 Experimental Environment

The fixed dataset contains 300 buses, 409 undirected edges, 2,016 temporal states, four districts, and a 12-step window. It contains 604,800 bus–timestamp rows: 574,560 benign and 30,240 attack. Attack regimes are partial (10,143), perfect (10,083), and masked (10,014). Reported severity spans 0.20–1.00 with mean 0.603 and standard deviation 0.231. The number of attacked buses per attack time has median 15 and range 3–29.

5.1 Evaluation protocol

The inherited held-out evaluation contains 121,200 samples: 115,048 benign and 6,152 attack. Temporal windows are kept within split boundaries and normalization statistics are fitted only on training data. The archive contains 24 real attack episodes; therefore, episode-level uncertainty is more appropriate than treating every bus–time row as an independent replicate.

We use positive-class F1, MCC, AUROC, AUPRC, balanced accuracy, and Top- k localiza-

tion. For localization,

$$Acc@k = \frac{1}{|\mathcal{A}|} \sum_{t \in \mathcal{A}} \mathbf{1}[v_t^* \in TopK(\mathbf{p}_t^{loc})]. \quad (16)$$

Graph-aware error is

$$D_{loc} = \frac{1}{|\mathcal{A}|} \sum_{t \in \mathcal{A}} d_G(\hat{v}_t, v_t^*). \quad (17)$$

For imbalanced detection, precision and recall are

$$P = \frac{TP}{TP + FP}, \quad R = \frac{TP}{TP + FN}, \quad (18)$$

and

$$F1 = \frac{2PR}{P + R}. \quad (19)$$

MCC is retained because it incorporates all four confusion-matrix cells:

$$MCC = \frac{TPTN - FPFN}{\sqrt{(TP + FP)(TP + FN)}} \times \frac{1}{\sqrt{(TN + FP)(TN + FN)}}. \quad (20)$$

For severity estimation, the primary continuous measures should be

$$MAE = \frac{1}{N} \sum_i |\hat{v}_i - v_i|, \quad RMSE = \sqrt{\frac{1}{N} \sum_i (\hat{v}_i - v_i)^2}. \quad (21)$$

5.2 Baselines and ablations

Completed benchmark evidence includes GCN, GAT, ChebNet, APPNP, and a physics-informed causal graph-transformer backbone. STGCN, DCRNN, TGAT, and generic Graph Transformer comparisons require dedicated runs and are not treated as completed results. The decisive experiment is the matched sequence Binary $\rightarrow$ Coordination $\rightarrow$ Persistence $\rightarrow$ Masking $\rightarrow$ Severity $\rightarrow$ Full model. Each stage must use identical splits, graph, window, optimizer budget, and inference information.

6 Validated Results and Analysis

6.1 Backbone detection

The independently recomputed backbone evaluation gives accuracy 0.9940, positive-class precision 0.8000, recall 0.9700, F1 0.8800, AUROC 0.9941, AUPRC 0.9732, MCC 0.8724, and balanced accuracy 0.9783.

The recall–precision asymmetry is informative: recall is 0.97 while precision is 0.80. With only 5.1% attacks in the held-out set, accuracy alone would obscure false-positive burden. MCC and AUPRC are therefore retained as primary secondary measures.

6.2 Localization

The available localization artifact reports Top-1, Top-3, and Top-5 accuracies of 0.509, 0.514, and 0.526.

This is a useful negative-space result: detection and localization are materially different tasks. A 300-node spatial hypothesis is much harder than deciding whether an attack exists. This motivates an explicit localization head rather than using detector confidence as a proxy.

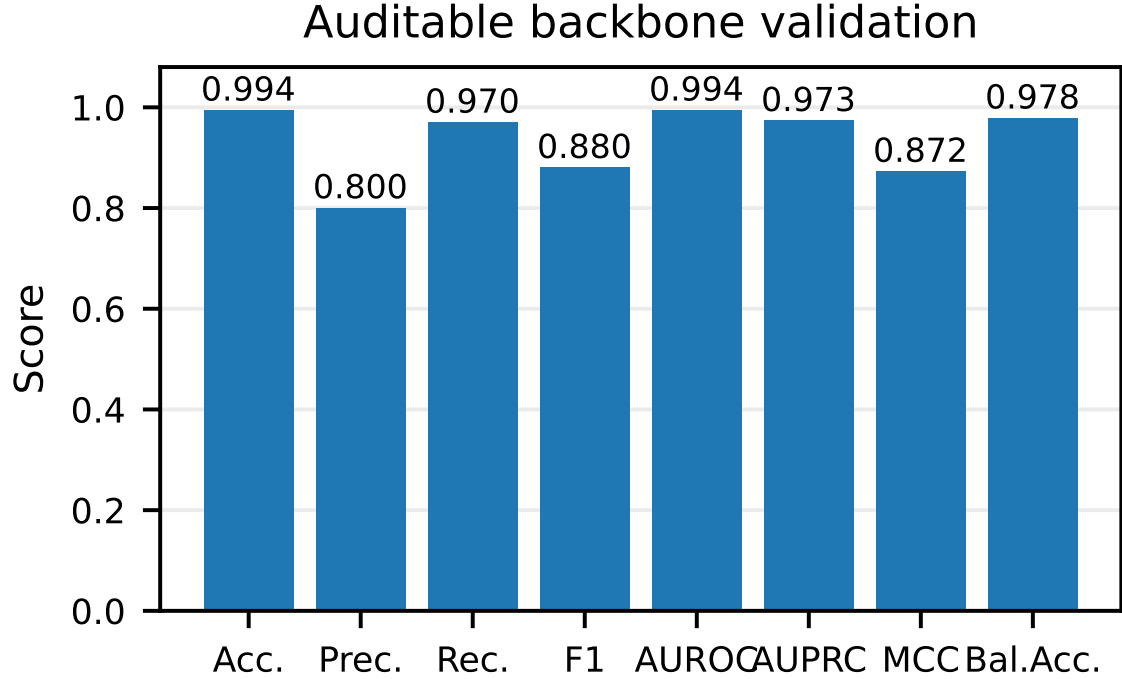


Figure 3: Auditable backbone metrics on the 121,200-sample held-out set. The high AUROC/AUPRC is not presented as evidence of a CHTGT gain.

Table 3: Evidence status and decisive experiments.

Target	Current evidence	Required confirmation
Detection	Backbone F1 0.880, AUPRC 0.9732	CHTGT vs binary
Localization	Top-1/3/5 0.509/0.514/0.526	Multi-task CHTGT
Causality	Mask defined	Causal vs non-causal
Severity	Range 0.20–1.00	MAE/RMSE/calibration
Zero-day	Associated-model gap	CHTGT withheld-family test
Hierarchy	Four-district design	Flat vs hierarchical

6.3 Severity evidence

The reported severity range is broad, so severity should be evaluated as calibrated regression rather than as another categorical label.

The final CHTGT experiment should report MAE, RMSE, rank correlation, and calibration by severity bin. No severity-head performance is claimed in the present paper because that retraining result is not available.

6.4 Zero-day diagnostic

An associated lightweight physics-informed causal model decreases from F1=0.025 on validation to 0.0066 on true zero-day and 0.0053 on balanced zero-day evaluation.

The relative decrease is approximately 73.6% from validation to true zero-day and 78.8% to balanced zero-day. The correct interpretation is that attack-family shift can dominate in-distribution performance; the result does not establish zero-day superiority for CHTGT.

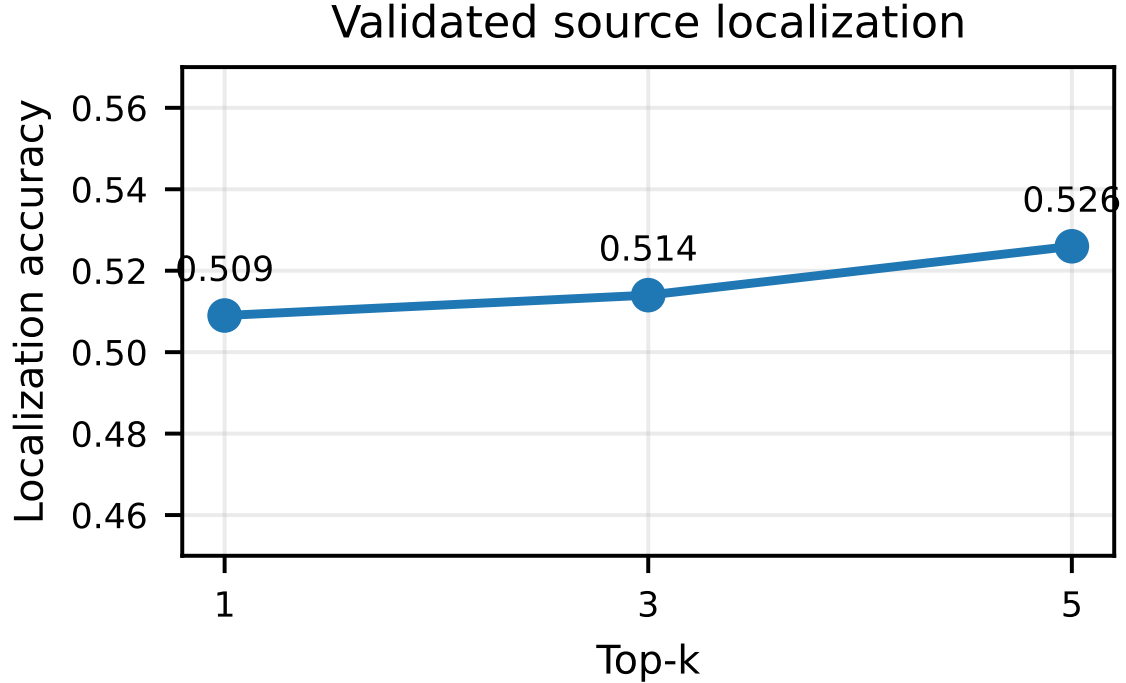


Figure 4: Validated bus-level source localization. The limited Top-1-to-Top-5 gain indicates that strong event detection does not automatically provide precise source ranking.

7 Discussion, Limitations, and Reproducibility

7.1 What the current evidence does and does not establish

The strongest available numerical evidence belongs to the inherited backbone, not the new CHTGT multi-task formulation. The distinction is substantive: a strong binary detector does not demonstrate that attack-evolution attributes are recoverable, and strong detection does not imply accurate localization. The paper therefore treats the existing results as *feasibility evidence* and the CHTGT ablation as the causal test of the proposed contribution.

7.2 Generalization and experimental validity

The dataset contains many bus-time observations but only 24 real attack episodes. The effective sample size for attack-level inference is therefore much smaller than the row count suggests. Random row-wise splitting would also allow correlated temporal windows to cross partitions. The final study should keep temporal windows inside split boundaries, separate whole episodes across train/validation/test, fit normalization only on training data, and report per-episode confidence intervals. These controls should be regarded as part of the evaluation definition rather than optional reporting details.

The evidence supports a narrow conclusion: the existing graph-temporal backbone is capable of strong held-out binary discrimination, localization is harder, and attack-family shift can sharply reduce performance. These observations motivate but do not prove the proposed attack-evolution formulation.

The central safeguard is the information-availability contract. Attack bus, severity, attack type, and other ground-truth metadata must never be concatenated to the inference vector. They can supervise auxiliary heads or stratify evaluation. This is essential for a defensible real-time claim.

Reported severity range and mean ± 1 SD

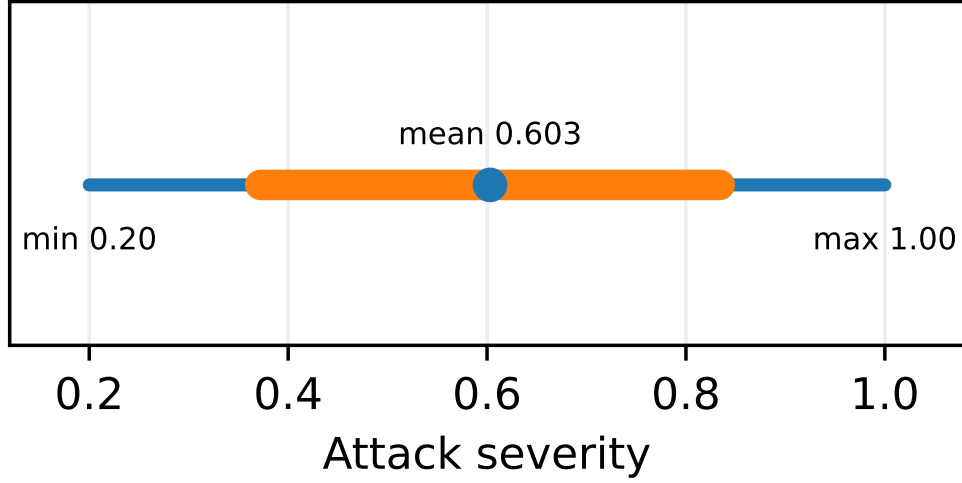


Figure 5: Factual summary of the reported severity statistics. The horizontal line shows the reported 0.20–1.00 range; the thicker interval shows mean ± 1 SD (0.372–0.834), and the marker is the reported mean 0.603.

A second issue is coordinated localization. The dataset can contain multiple attacked buses at a timestep, while a single softmax expresses a one-source ranking problem. The present Top- k evidence should therefore be interpreted as source-bus ranking. A true coordinated-localization study should use multi-label node prediction and report set precision/recall, node-level IoU, and graph-distance error.

Third, “severity estimation” remains a proposed capability until a converged severity head is evaluated. The current dataset establishes that a broad severity target exists; it does not establish that the target is inferable from observable telemetry.

Fourth, 24 real attack episodes impose a statistical ceiling. Thousands of bus–time observations do not create thousands of independent attack experiments. Episode-level bootstrap or permutation testing should accompany the final ablation. Additional baselines should be added only after dedicated, matched runs.

For reproducibility, the final experiment should release: exact split timestamps, graph edge list, train-only normalization statistics, attack-episode IDs, model configuration, seed list, raw predictions, and the code used to compute every reported metric.

7.3 Statistical design

The final comparison should treat the attack episode, rather than the individual bus–time row, as the principal unit of uncertainty. For each episode e , let M_e denote a metric obtained from the fixed test window. A paired bootstrap comparison between two models is then based on

$$\Delta M = \frac{1}{E} \sum_{e=1}^E (M_e^{\text{CHTGT}} - M_e^{\text{base}}), \quad (22)$$

with confidence intervals obtained by resampling episodes. This avoids pseudo-replication caused by many correlated measurements from the same event. McNemar’s test is appropriate only for paired binary decisions; F1, MCC, AUPRC, and localization metrics should use paired bootstrap or permutation procedures.

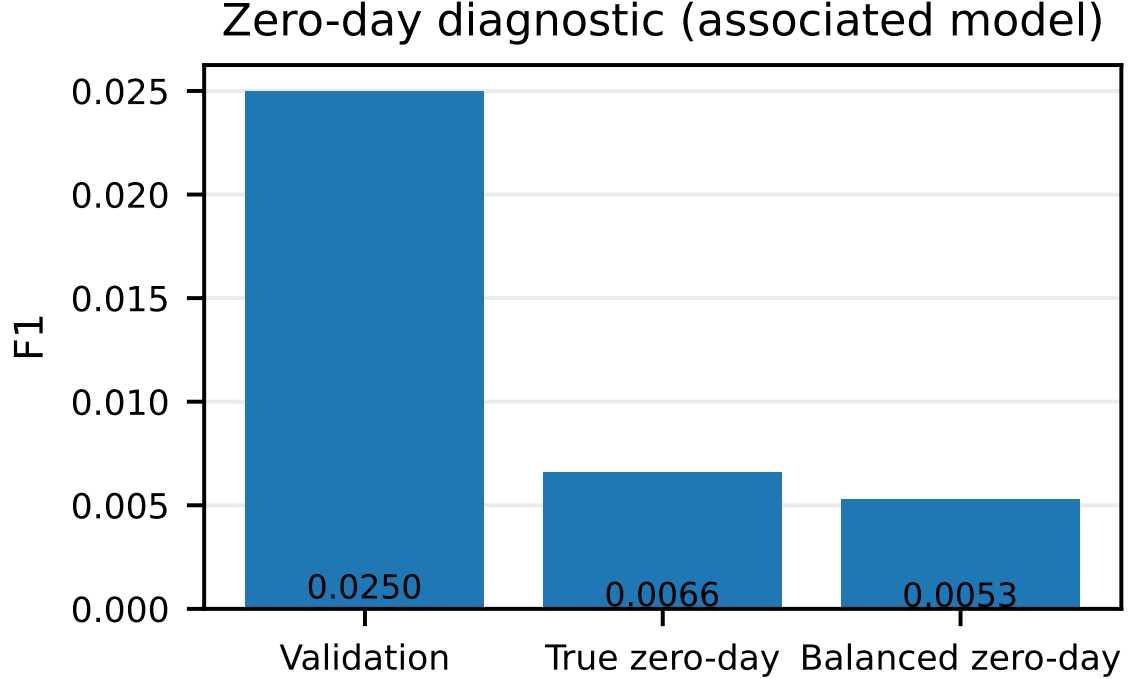


Figure 6: Zero-day diagnostic from an associated lightweight model. These values are not CHTGT results; they motivate withheld-family evaluation.

7.4 Failure modes and falsification

The proposed model should be considered unsuccessful if any of the following occurs: (i) the full multi-task representation does not improve the matched binary baseline; (ii) gains disappear under withheld attack families; (iii) severity predictions are poorly calibrated even when detection is strong; (iv) the causal model loses performance only because the non-causal comparator uses future information; or (v) district pooling produces no benefit over a flat graph encoder. These criteria make the contribution falsifiable rather than defining success through a favorable single metric.

7.5 Deployment implications

The architecture is intended for a monitoring pipeline in which the detector first identifies a suspicious event, the localization head ranks candidate buses, and the severity head estimates the magnitude of the event for triage. The outputs can be consumed by a downstream operator or mitigation layer without exposing future measurements to the model. In particular, the causal mask makes the temporal representation compatible with streaming inference, while the district hierarchy offers a natural interface for regional monitoring. These are architectural properties; they are not claims of operational certification.

8 Conclusion

This paper reframes smart-grid cyberattack analysis from binary event detection toward attack-evolution modeling. The proposed hierarchical temporal graph transformer organizes the task into bus-level topology, district aggregation, causal temporal evolution, and multi-task attack profiling. Detection, localization, and severity are treated as related outputs rather than isolated tasks.

Table 4: Final controlled experiment matrix.

Test	Comparison	Primary outcome
T1	Binary vs multi-task	F1, MCC, AUPRC
T2	Causal vs non-causal	F1, leakage check
T3	Flat vs district hierarchy	F1, Top- k
T4	Severity head	MAE, RMSE, calibration
T5	Known vs withheld family	F1, AUPRC
T6	Localization	Top-1/3/5, D_{loc}

The available evidence is deliberately separated from the proposed contribution. On the 121,200-sample held-out set, the validated backbone achieves 0.9940 accuracy, 0.8800 positive-class F1, 0.9941 AUROC, 0.9732 AUPRC, and 0.8724 MCC; localization reaches 0.509/0.514/0.526 for Top-1/3/5. These values establish backbone feasibility and expose localization as a harder problem. The associated zero-day diagnostic further demonstrates why withheld-family evaluation is essential.

The decisive next stage is a controlled CHTGT experiment with matched binary and multi-task models, causal versus non-causal temporal attention, flat versus district-hierarchical encoding, severity calibration, and withheld-family evaluation. Until those experiments are completed, no universal superiority or zero-day dominance is claimed. The contribution is a reproducible framework that makes cyberattack evolution a measurable spatiotemporal learning problem.

Data Availability

State where the data supporting this study can be found, or that it is contained within the article. If data cannot be shared, say why.

Conflict of Interest

Declare any competing interests, or state “The authors declare no competing interests.”

Funding

List any funding sources and grant numbers, or state that none were received.

Author Contributions

Briefly describe each author’s contribution (e.g. CRediT roles).

References

- [1] Y. Liu, P. Ning, and M. K. Reiter, “False data injection attacks against state estimation in electric power grids,” in *Proc. ACM CCS*, 2009, pp. 21–32.

- [2] Y. Liu, P. Ning, and M. K. Reiter, “False data injection attacks against state estimation in electric power grids,” *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, Art. 13, 2011, doi: 10.1145/1952982.1952995.
- [3] X. Yu and Y. Xue, “Smart grids: A cyber–physical systems perspective,” *Proc. IEEE*, vol. 104, no. 5, pp. 1058–1070, 2016, doi: 10.1109/JPROC.2015.2503119.
- [4] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” in *Proc. ICLR*, 2017.
- [5] W. L. Hamilton, Z. Ying, and J. Leskovec, “Inductive representation learning on large graphs,” in *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [6] P. Veličković *et al.*, “Graph attention networks,” in *Proc. ICLR*, 2018.
- [7] B. Yu, H. Yin, and Z. Zhu, “Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting,” in *Proc. IJCAI*, 2018, pp. 3634–3640.
- [8] A. Vaswani *et al.*, “Attention is all you need,” in *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [9] Y. Li, R. Yu, C. Shahabi, and Y. Liu, “Diffusion convolutional recurrent neural network: Data-driven traffic forecasting,” in *Proc. ICLR*, 2018.
- [10] D. Xu, C. Ruan, E. Korpeoglu, S. Kumar, and K. Achan, “Inductive representation learning on temporal graphs,” in *Proc. ICLR*, 2020.
- [11] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *J. Comput. Phys.*, vol. 378, pp. 686–707, 2019, doi: 10.1016/j.jcp.2018.10.045.
- [12] R. Ying *et al.*, “GNNExplainer: Generating explanations for graph neural networks,” in *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [13] J. Kirkpatrick *et al.*, “Overcoming catastrophic forgetting in neural networks,” *Proc. Natl. Acad. Sci. USA*, vol. 114, no. 13, pp. 3521–3526, 2017, doi: 10.1073/pnas.1611835114.
- [14] M. Irfan, A. Sadighian, A. Tanveer, S. J. Al-Naimi, and G. Oligeri, “A survey on detection and localisation of false data injection attacks in smart grids,” *IET Cyber-Physical Systems: Theory & Applications*, vol. 9, no. 4, pp. 313–333, 2024, doi: 10.1049/cps2.12093.
- [15] A. D. Syrmakesis and N. D. Hatziaegyriou, “Cyber resilience methods for smart grids against false data injection attacks: Categorization, review and future directions,” *Frontiers in Smart Grids*, vol. 3, Art. 1397380, 2024, doi: 10.3389/frsgr.2024.1397380.
- [16] Z. Qu *et al.*, “Localization of dummy data injection attacks in power systems considering incomplete topological information: A spatio-temporal graph wavelet convolutional neural network approach,” *Applied Energy*, vol. 360, Art. 122736, 2024, doi: 10.1016/j.apenergy.2024.122736.
- [17] A. Wachi, X. Shen, and Y. Sui, “A survey of constraint formulations in safe reinforcement learning,” in *Proc. IJCAI*, 2024, pp. 8262–8271, doi: 10.24963/ijcai.2024/913.
- [18] X. Wang, M. Hu, X. Luo, and X. Guan, “A detection model for false data injection attacks in smart grids based on graph spatial features using temporal convolutional neural networks,” *Electric Power Systems Research*, vol. 238, Art. 111126, 2025, doi: 10.1016/j.epsr.2024.111126.

- [19] Y. He, J. Wang, C. Yang, and D. Shi, “A graph and diffusion theory-based approach for localization and recovery of false data injection attacks in power systems,” *Electric Power Systems Research*, vol. 239, Art. 111184, 2025, doi: 10.1016/j.epsr.2024.111184.
- [20] T. Su, T. Wu, J. Zhao, A. Scaglione, and L. Xie, “A review of safe reinforcement learning methods for modern power systems,” *Proc. IEEE*, vol. 113, no. 3, pp. 213–255, 2025, doi: 10.1109/JPROC.2025.3584656.
- [21] Y. Zhang *et al.*, “A critical review of safe reinforcement learning strategies in power and energy systems,” *Engineering Applications of Artificial Intelligence*, vol. 143, Art. 110091, 2025, doi: 10.1016/j.engappai.2025.110091.
- [22] K. Singh and S. Kumari M., “Identification and localization of false data injection attacks in smart grids using graph Fourier transform,” *Electric Power Systems Research*, vol. 251, Art. 112238, Feb. 2026, doi: 10.1016/j.epsr.2025.112238.
- [23] M. Hassouna, C. Holzhüter, P. Lytaev, J. Thomas, B. Sick, and C. Scholz, “Graph reinforcement learning for power grids: A comprehensive survey,” *Energy and AI*, vol. 23, Art. 100671, 2026, doi: 10.1016/j.egyai.2025.100671.