Alireza Limooie

# Geometric Noise Augmentation for Roof Graph Reconstruction

Date: 10/07/2025

## 1. Objective & Background

This report details the recent development and implementation of a new, stateful geometric noise augmentation pipeline for our roof graph reconstruction model. The primary objective is to create a more realistic simulation of errors found in data from automated line detectors, thereby improving the model's robustness and its ability to perform fine-grained geometric refinement.

Initial augmentation strategies, while effective, were based on a static topology approach that did not fully capture the progressive nature of data degradation. The work in this reporting period has centered on replacing that system with a dynamic, cumulative, and threshold-based noise model.

### 2. A New Paradigm: Cumulative, Threshold-Based Noise

The core of the recent work is a complete re-architecture of the on-the-fly data augmentation process. This new paradigm moves away from independent probabilities and instead simulates a gradual "destruction" process over a diffusion-like schedule (T=10).

**How It Works:**

1. **Per-Edge Effective Noise:** For each training instance, a maximum "destruction level" is randomly selected. Then, for each timestep t in the diffusion schedule, every edge calculates its own effective_noise level. This value is a function of the global destruction level, the timestep t, and a randomly assigned, per-edge sensitivity. This ensures that some edges degrade faster than others, mimicking real-world conditions.

2. **Cumulative Topology:** The state of the graph at timestep t is directly derived from the state at t-1. This is a critical departure from the previous model. If an edge is deleted or subdivided at t=5, it remains in that state for all subsequent steps (t=6, 7, ...).

3. **Threshold-Based Augmentations:** Destructive topological changes are no longer triggered by simple probabilities. Instead, they occur when an edge's effective_noise level surpasses a specific, configurable threshold. This creates a clear, causal link between the noise level and the severity of the damage.

    - **Edge Deletion:** An edge is removed from the graph if its noise level exceeds geom_noise_delete_edge_threshold.

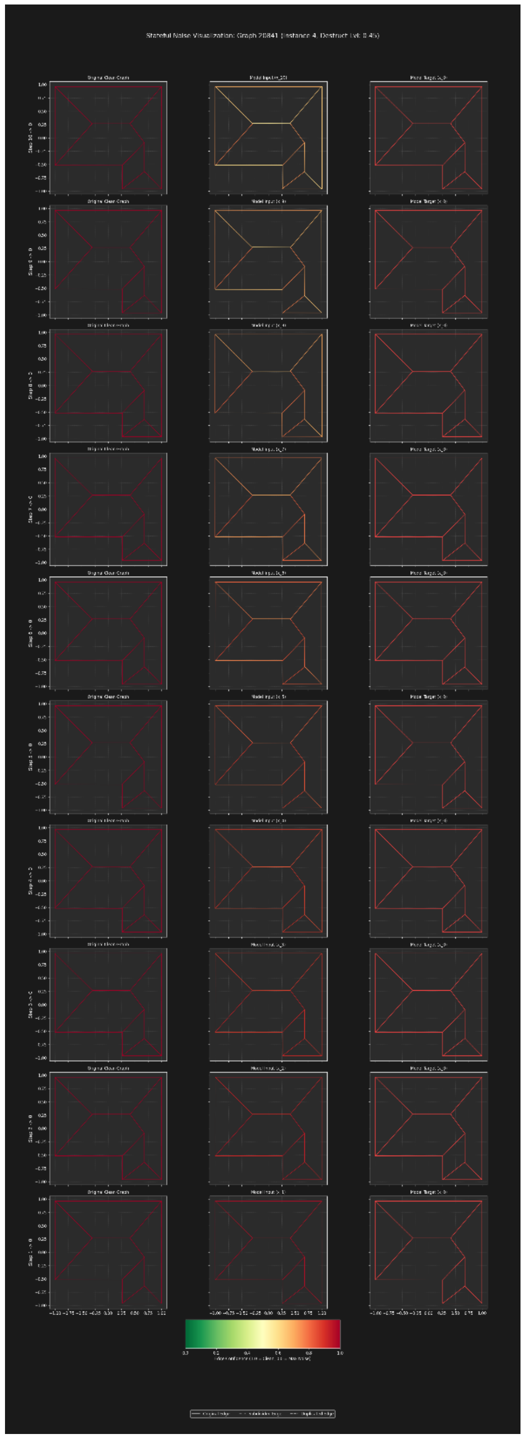
    - **Edge Subdivision (Breaking):** An edge is fragmented into multiple smaller segments if its noise level exceeds geom_noise_subdivide_edge_threshold.

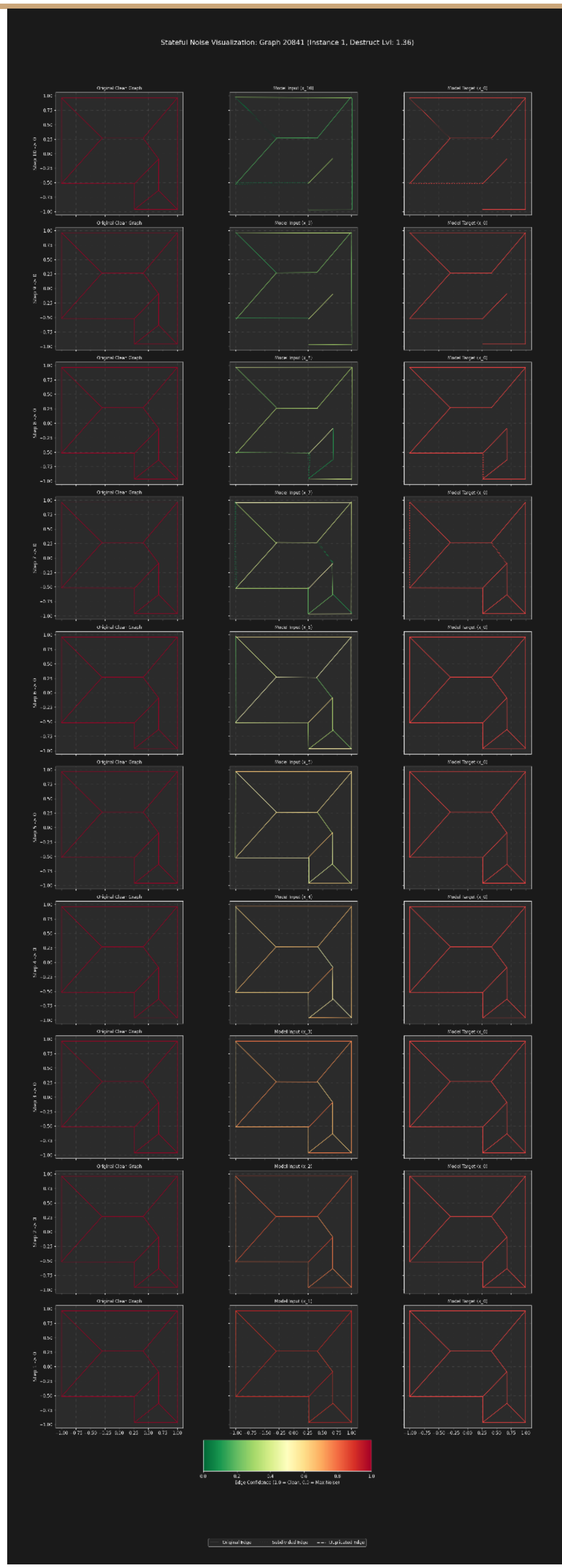
    - **Edge Duplication:** An edge is copied with a slight random offset if its noise level exceeds geom_noise_duplicate_edge_threshold.

| Feature | Previous Static Model | New Cumulative Model | Benefit of New Model |
| --- | --- | --- | --- |
| **Topology** | Decided once, upfront. Fixed for all timesteps. | Evolves at each timestep. Cumulative changes. | More realistic simulation of progressive data corruption. |
| **Trigger Mechanism** | Independent probabilities (geom_noise_p). | effective_noise crossing a set threshold (geom_noise_*_threshold). | Creates a direct, causal link between noise level and type of error. |
| **Noise Correlation** | Coordinate and topological noise were independent. | All noise types (coordinate, deletion, etc.) are driven by the same effective_noise. | More realistic, as severe coordinate noise often correlates with topological errors. |

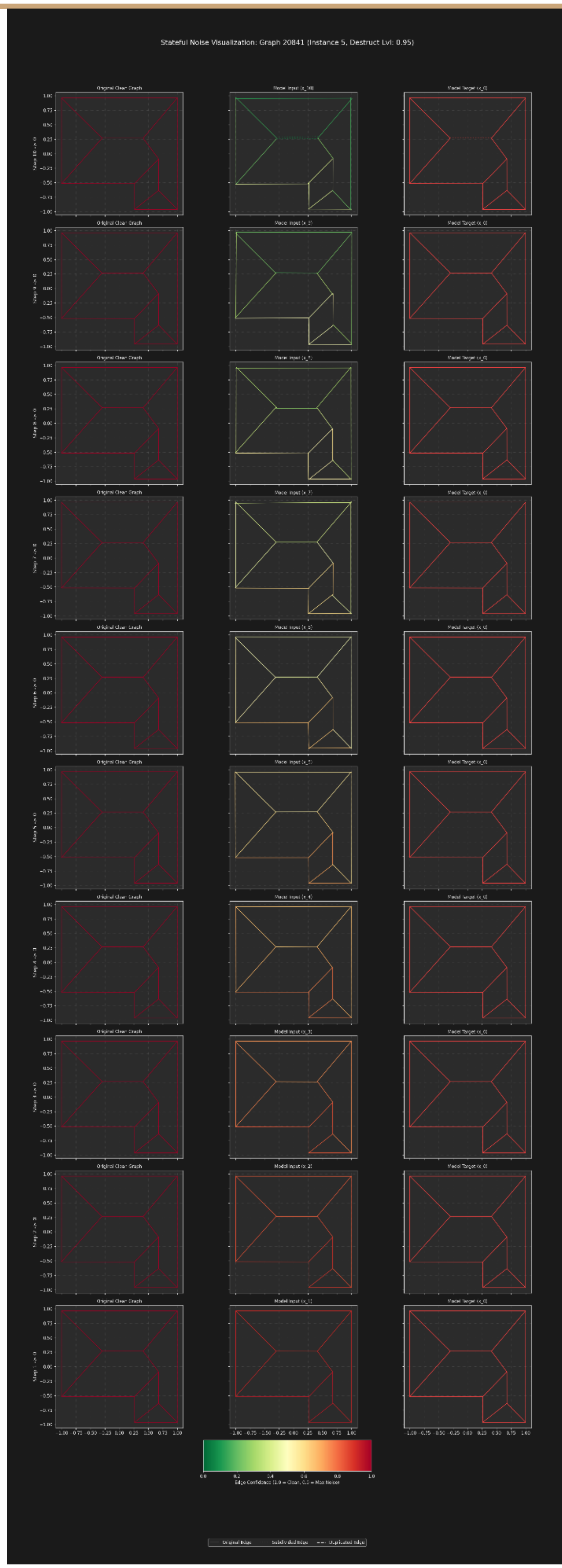
## 3. Initial Results & Observations

### Figure 1: Examples of Cumulative Noise Augmentation

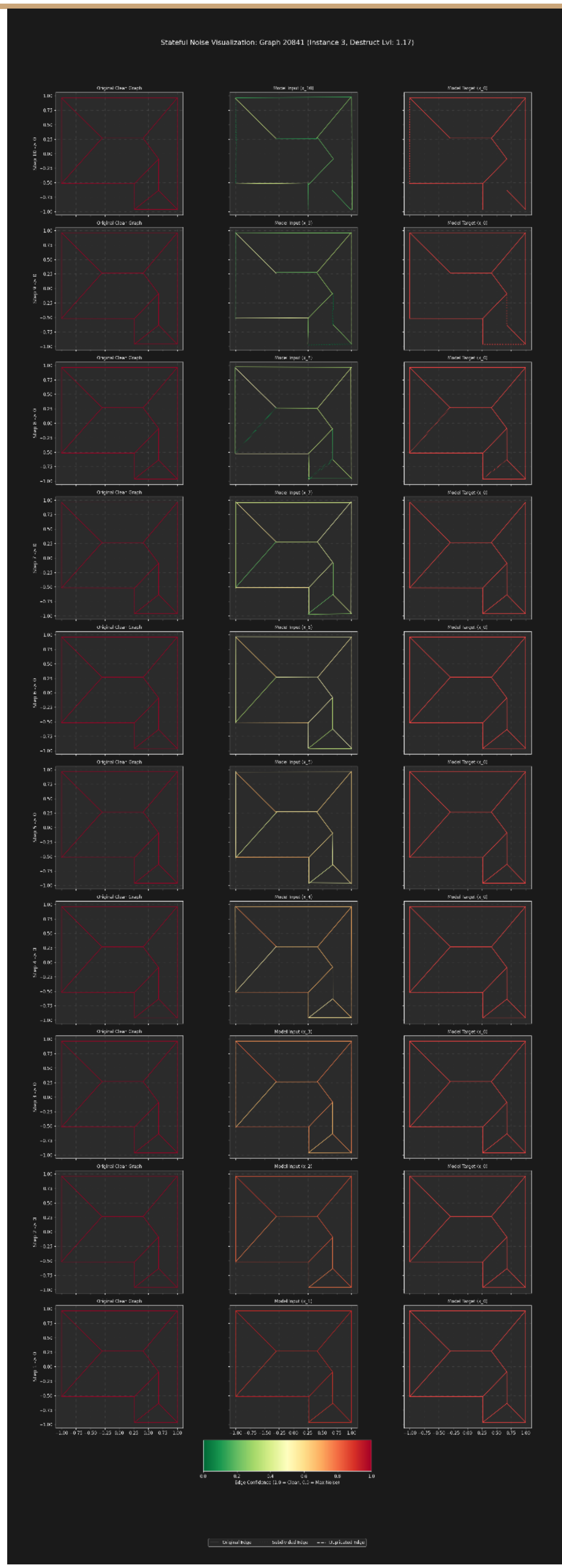




Stateful Noise Visualization: Graph 20841 (Instance 1, Destruct Lvl: 1.36)
Original Clear Graph
Model Input (x_10)
Model Target (x_0)
Edge Confidence (1.0 = Clean, 0.0 = Max Noise)
Original Edge
Subdivided Edge
Duplicated Edge



Stateful Noise Visualization: Graph 20841 (Instance 5, Destruct Lvl: 0.95)
Original Clean Graph
Model Target (x_0)
Edge Confidence (1.0 = Clean, 0.0 = Max Noise)
Original Edge
Subdivided Edge
Duplicated Edge


Stateful Noise Visualization: Graph 20841 (Instance 3, Destruct Lvl: 1.17)
Original Clean Graph
Model Target (x_0)
Edge Confidence (1.0 = Clean, 0.0 = Max Noise)
0.0
0.2
0.4
0.6
0.8
1.0
Original Edge
Subdivided Edge
Duplicated Edge

## 4. Future Work on Augmentation

Several areas for refinement and expansion have been identified:

- **Tuning Noise Thresholds:** The current thresholds for deletion, subdivision, and duplication are initial estimates. Systematic hyperparameter tuning is needed to find the optimal values that best mimic real-world detector error patterns.

- **False Positive/Negative Augmentation:** The framework can be extended to address more complex errors. Initial concepts, currently disabled, include:

  - **False Negative Repair:** Intelligently adding new line segments to fill gaps where a detector may have failed. An experimental grid-based vector system was prototyped for this purpose.

  - **False Positive Deletion:** Developing a mechanism to identify and remove entirely spurious edges that have no basis in the ground truth.

- **Integrating Advanced Augmentations:** The ultimate goal is to create a true end-to-end cleaning tool. This will involve seamlessly integrating the false positive/negative augmentations into the cumulative noise pipeline, allowing the model to learn not just how to refine geometry, but also how to correct the fundamental topology of the graph.