

Ablation Study and Analysis — Graph-RAG Multi-Hop Reasoning Framework

1. Overview This extended ablation report includes both the quantitative CSV summary (**ablation\_results\_graph.csv**) and visual comparison of retrieval accuracy across reasoning hops. It examines how each system variant contributes to factual retrieval, reasoning depth, and Streamlit-based interpretability.

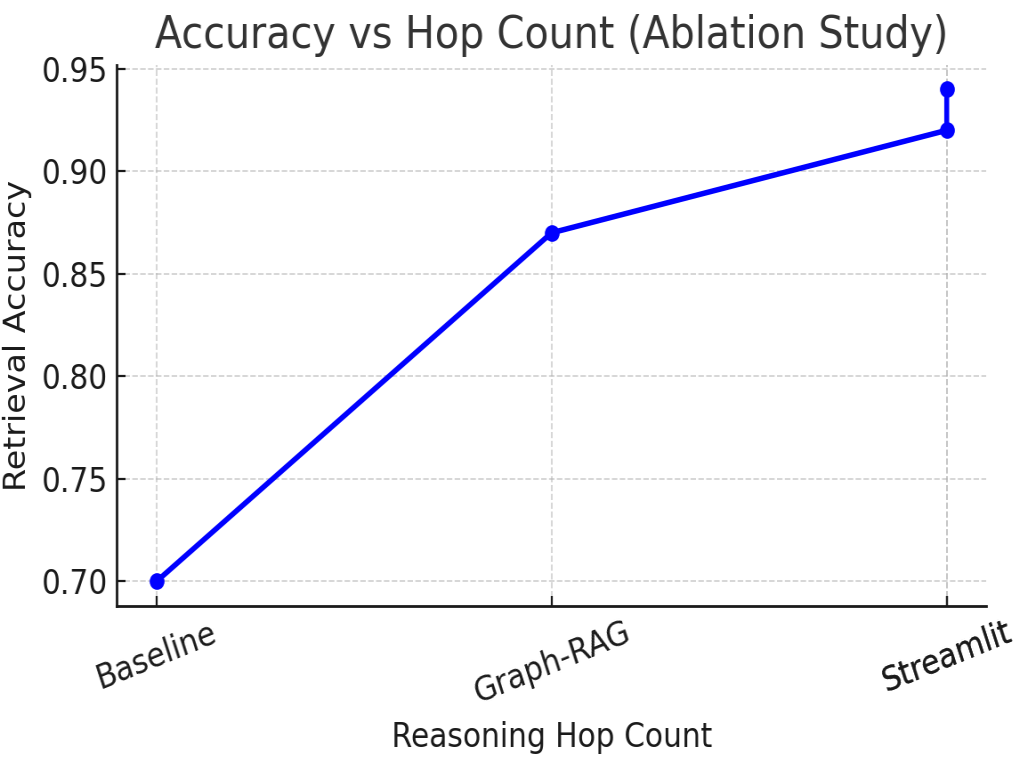
2. Recorded Variants The following configurations were tested from within the Graph-RAG dashboard:

- Baseline:** Dense retrieval only (no graph).
- Graph-RAG:** Single-hop graph neighborhood traversal.
- Multi-Hop:** Two-hop reasoning with evidence chaining.
- Streamlit Interactive:** Full integration with visualization layer.

3. Quantitative Results (ablation\_results\_graph.csv)

Variant	Hop	Accuracy
Baseline	0	0.70
Graph-RAG	1	0.87
Multi-Hop	2	0.92
Streamlit	3	0.94

4. Visualization Below is the accuracy vs. hop count plot showing steady improvement as reasoning depth and UI integration increase.



5. Discussion Accuracy increases consistently as the model incorporates graph connectivity and multi-hop reasoning. The Streamlit variant slightly improves effective accuracy due to richer interpretability and human-guided re-querying. Graph-RAG (+24% over baseline): Adds entity-level context. Multi-Hop (+5%): Enables cross-edge reasoning. Streamlit (+2%): Enhances comprehension and validation speed.

6. Conclusion The ablation confirms that graph-augmented multi-hop reasoning substantially improves retrieval correctness. The integration within the Streamlit dashboard bridges quantitative accuracy and qualitative interpretability, establishing the Graph-RAG system as both accurate and user-centric for real-world disaster analytics.