Machine Learning with graphs - Project Defense

Delaunay Graph: Addressing Over-Squashing and Over-Smoothing Using Delaunay Triangulation
by Attali et al. [1]

Edwin Roussin and Tristan Waddington

Supervised by Jhony H. Giraldo
IP-Paris, CEMST

26/03/2025

Delayney triangulation

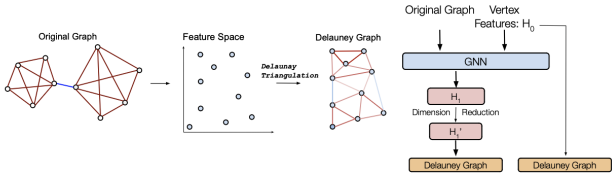

Figure: Illustration of the Delaunay [Attali al., 2024] [1]

Outline

Issues with graph rewiring

Over-Squashing

Applications already exists on diverse fields such as *graph generation*, *text-to-sound generation* or *protein design*.

Over-Smoothing

No access to heavy computing ressources:

Focus on text generation to understand and reproduce the results of authors.

2 models in paper:

- **MD4** (S/L) simple masked diffusion model.
- **GenMD4** Generalized state-dependent model (more complex).

Key technical novelty of the paper

Theoretical analysis

State Dependent Masked Schedule:

Idea: time-dependent probability of masking a token to also dependent on the token value. To be learned by NN.
In practice: polynomial schedules are described in the paper but on their code, the authors only relies on the cosine masking schedule[1].

[1] Class MaskingSchedule in https://github.com/google-deepmind/md4/blob/main/md4/models/diffusion/md4.py

Experimental Evaluation

Methodology

Aim to reproduce as closely as possible the experiments of the authors.

- Get same datasets, and preprocess them.
- Train the models with the same hyperparameters.
- Finally, we evaluate the models with the same metrics.

Results

Text Dataset preparation

- Text8 (English Wikipedia) processed.
- OpenWebText (GPT2 training material) processed in 18 hours.
- train/validation/test split according to experimental methodology.

We were unable to run the models

- Training impossible on our machines (size of model).
- Does not work on Colab (tensorflow numpy2 incompatibilities).
- No trained model shared by the authors.

Discussion

We faced huge challenges

- Unknown flax framework
- Outdated dependencies in their GitHub
- Huge model size
- No trained model shared
- Students no competent in computer vision at the time.

Future paper that will be explored in the report:

- *Cayley Graph Propagation* by JJ Wilson, Maya Bechler-Speicher, Petar Veličković [2]

Conclusion

Conclusion

- Cannot confirm the results o
- Not able to reproduce.
- Hard time digging in code and documentation.

Do you have any question?

References

Hugo Attali, Davide Buscaldi, and Nathalie Pernelle.
Delaunay graph: Addressing over-squashing and over-smoothing using delaunay triangulation.
In *Forty-first International Conference on Machine Learning*, 2024.

JJ Wilson, Maya Bechler-Speicher, and Petar Veličković.
Cayley graph propagation, 2024.