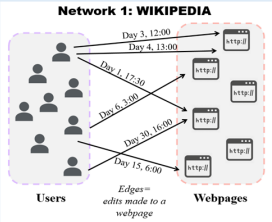
Improving Temporal Graph Network Messaging

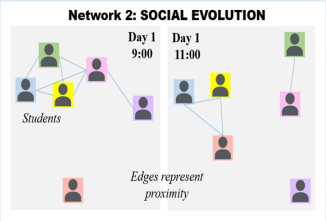
Altering TGN Algorithm to incorporate additional neighbourhood information in networks that change over time

Introduction

Task — *Link Prediction in Temporal Graphs*

- Predict next edge between nodes over time
- Can model processes such as recommendation systems, social networks, protein-protein interactions, etc.



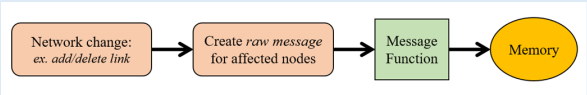


Algorithm - *Temporal Graph Network (TGN)* (Rossi et al. 2020)

- Graph Neural Network algorithm TGN-attn significantly outperforms recent algorithms such as JODIE, TGAT, DyRep for link prediction

How does it work?

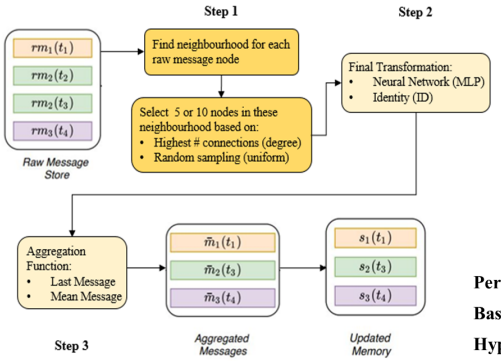
TGN works by storing a *memory for each node* which stores changes that affect it over time.



Question — *Room for improvement?*

- Can we improve link predictions by taking into account neighbourhood changes in the message function of TGN-attn.?

Methodology



Variants Tested:

Variant Name	*# Neighbours*	*Nbhd Sample*	*Msg Transformation*
0-Rec-ID	0	Recent	ID
5-Rec-ID	5	Recent	ID
10-Rec-ID	10	Recent	ID
5-Deg-ID	5	Degree	ID
10-Deg-ID	10	Degree	ID
5-Uni-ID	5	Uniform	ID
10-Uni-ID	10	Uniform	ID
5-Deg-MLP	5	Degree	MLP
10-Deg-MLP	10	Degree	MLP
5-Uni-MLP	5	Uniform	MLP
10-Uni-MLP	10	Uniform	MLP

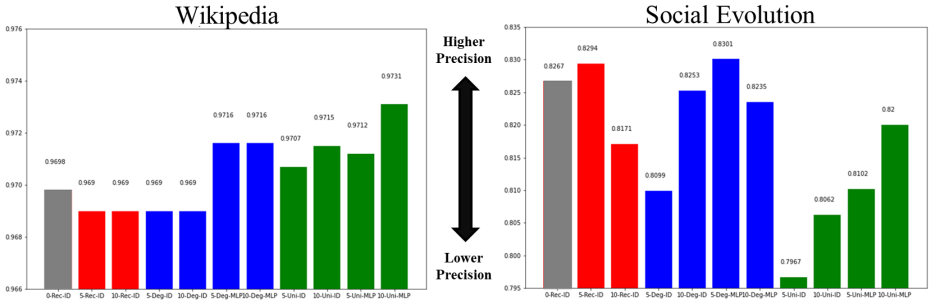
Performance Metric: Average Precision (AP) over 4 runs

Baseline Comparison Variant: 0-Rec-ID

Hyperparameters: Same as in baseline variant

Results

Average Precision — Grouped by Neighbourhood Sample Type



Gray = Baseline, Red = Most Recent, Blue = Highest # Connections (Degree), Green = Random (Uniform)

Best Performing Variants Overall:

- *Wikipedia:* 10-Uni-MLP
- *Social Evolution:* 5-Rec-ID/5-Deg-MLP

Discussion and Conclusions

Very small improvement in Average Precision for addition of neighbourhood information in message function

- Graph attention likely captures enough message sharing, so additional neighbourhood information makes little difference

Neighbourhood Sample: Different neighbourhood samples performed best in both networks, reflecting their very different structure

Neighbourhood Effect: No clear pattern for either 5 or 10

- It is possibly more dependent on the data and the other variant options

Transformation: Neural Network (MLP) was best

Time: Addition of neighbourhood information and MLP transformation make runs take 4-5 times longer, but can still be considered as trade-off

Future Directions

- Use alternate embedding (not graph attention) with neighbourhood additions to message function
- Compare results on more temporal network data with similar structures to Wikipedia and Social Evolution
- Consider 2nd neighbourhood, different number of neighbours

Thank you for your interest in my project! Feel free to reach out if you have further questions.

Contact Info

McGill University

Celeste Groux – celeste.groux@mail.mcgill.ca

McGill BSc Honours Applied Mathematics, Minor Computer Science

https://www.linkedin.com/in/celeste-groux/