TemporalGAT: Attention-Based Dynamic Graph Representation Learning

Ahmed Fathy and Kan Li

School of Computer Science and Technology,

Beijing Institute of Technology

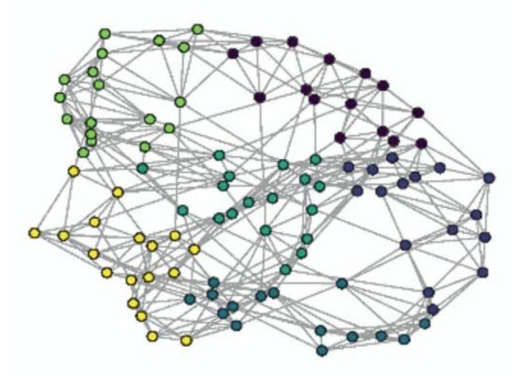
北京理工大學
BEIJING INSTITUTE OF TECHNOLOGY

Ahmed Fathy and Kan Li. “TemporalGAT: Attention-Based Dynamic Graph Representation Learning”.

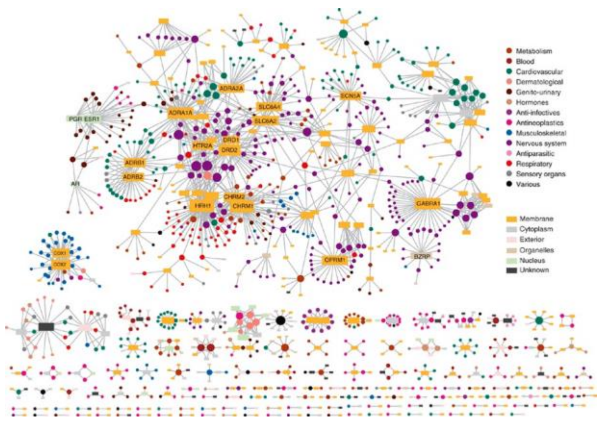
Real-world Graphs are in Continuous Evolution

- Real-world dynamic graphs are continuously evolved.
- New nodes and edges are added or removed during graph evolution.

Social networks



Collaboration networks



Drug-protein networks

Dynamic Graph Representation Learning

- Finds the most informative *low-dimensional representation* of the data that preserves the structural, temporal and other network information.

Why is it important?

- Dynamic Information networks are becoming more prevalent in our daily lives such as social networks and citation networks, etc.
- It has extensive real-world applications such as friendship prediction, and user recommendation.

Research Questions

- How to integrate network **topology structure**, **node features**, and **temporal information** in the representation learning process?

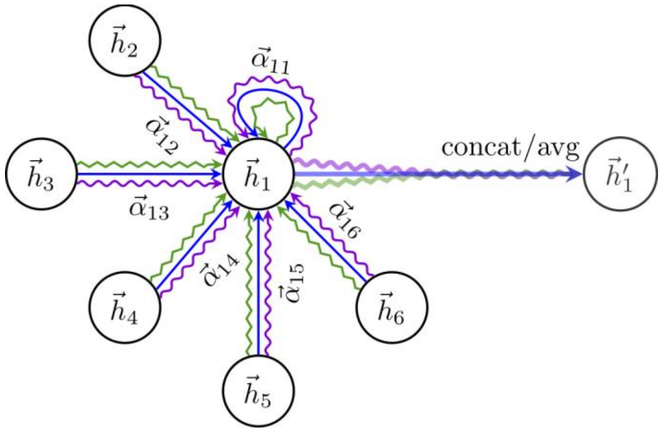
- How to learn meaningful graph representations using state-of-the-art **graph neural networks** and **deep learning** methods?

Related Work: Static Graph Representation Learning

Learn effective node representations for static graphs

- Network Structure:
  - DeepWalk (Perozzi et al., 2014)
  - LINE (Tang et al., 2015)
  - …

- Network Structure and Node Features:
  - Graph Convolutional Networks – GCN (Kipf et al., 2016)
  - Graph Attention Networks – GAT (Veličković et al., 2018)
  - …



Graph Attention Network

Pros and Cons of Static Graph Representation Learning Methods

Pros

- Learn superior node representations for static graphs
- Effective in several network analytic tasks such as:
  - Node classification
  - Link prediction
  - Network visualization

Cons

- Unable to model the evolving nature of dynamic graphs
- Model training and inference on dynamic graphs are difficult, due to the evolving nature of graph structures

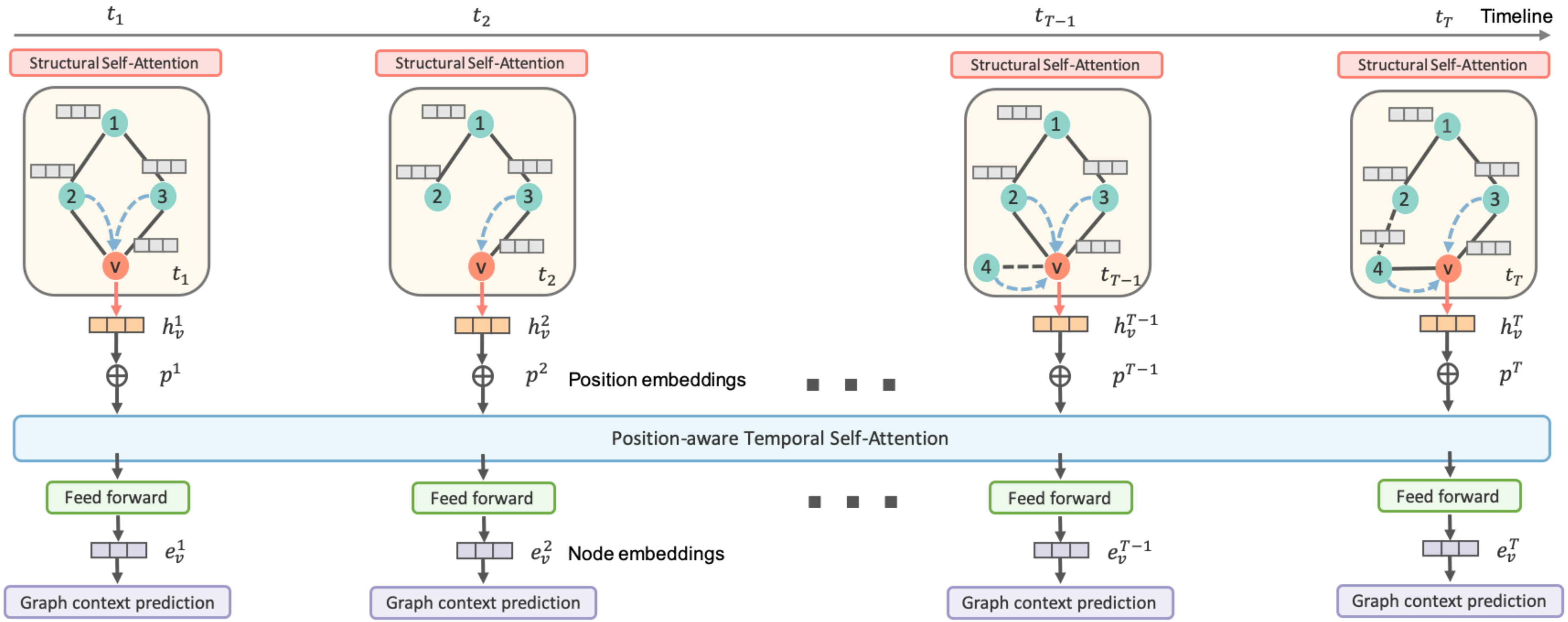
Related Work: Dynamic Graph Representation Learning

Examples:

- DANE (Li et al., 2017)
  - Updates the eigenvectors of graph Laplacian matrix over time series.
- DynGEM (Goyal et al., 2017)
  - Applies an autoencoder model to learn incremental node representations across time steps.
- DySAT (Sankar et al., 2018)
  - Captures structural properties separately using GAT, and applies temporal attention layers to capture the variations in the generated representations.

Related Work:

DySAT: Dynamic graph representation learning via self-attention networks



Pros and Cons of Dynamic Graph Representation Learning Methods

Pros

- Inspired by static graph representation methods.
- Capture the temporal evolutionary patterns of dynamic graphs.

Cons

- Several methods be considered as an *extension* of static methods.
- Tend to learn representation for each graph snapshot separately, and update the embedding weights according to snapshot position.

Can we do better?

TemporalGAT: Attention-Based Dynamic Graph Representation Learning

- Inspired by recent work, DySAT (Sankar et al., 2018).
- Towards combining structural and temporal information.
- Applies self-attention mechanism along structural neighborhoods over temporal dynamics.
- Leverage **temporal convolutional network** (TCN).

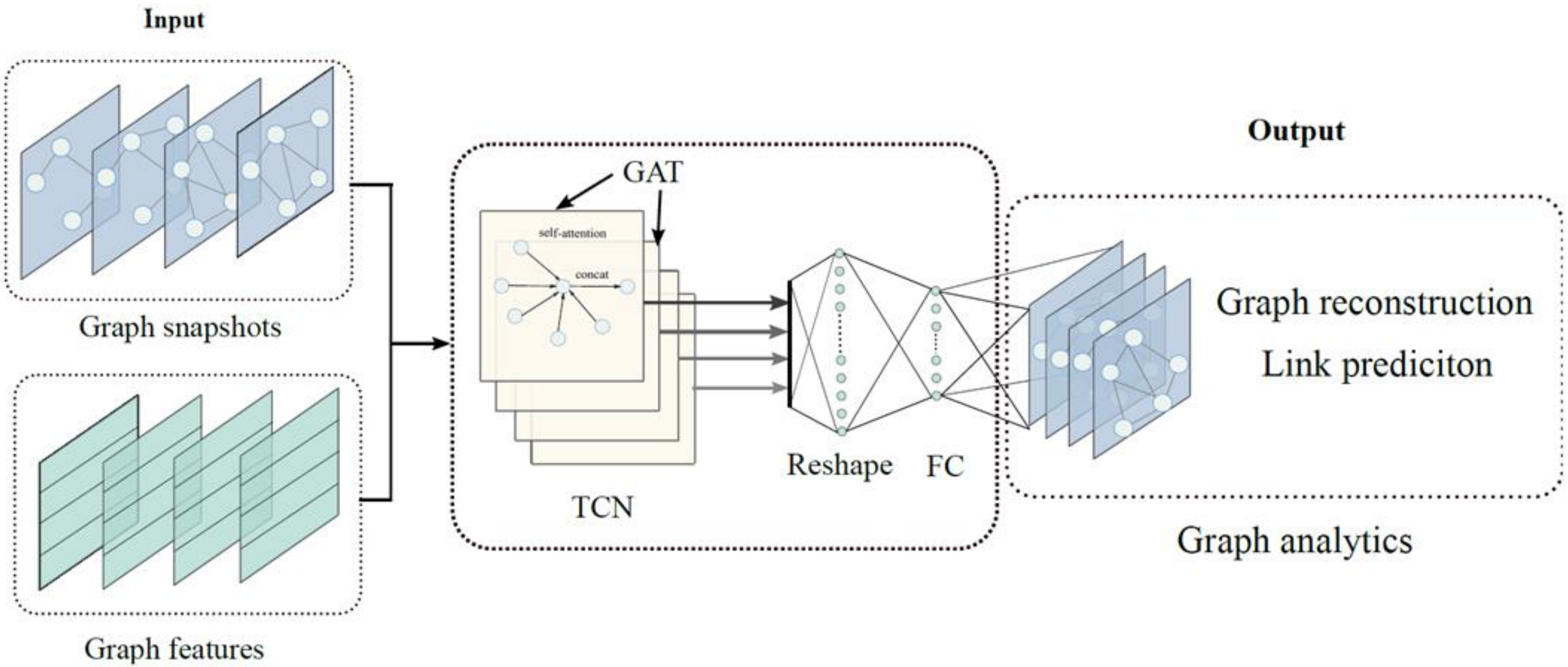
TemporalGAT: Method

- Apply multi-head graph attentions and TCNs as a special recurrent structure to improve model efficiency.
- The proposed GAT layer has **dilated convolutions** to ensure that ***no information leakage*** from future to past graph snapshots.

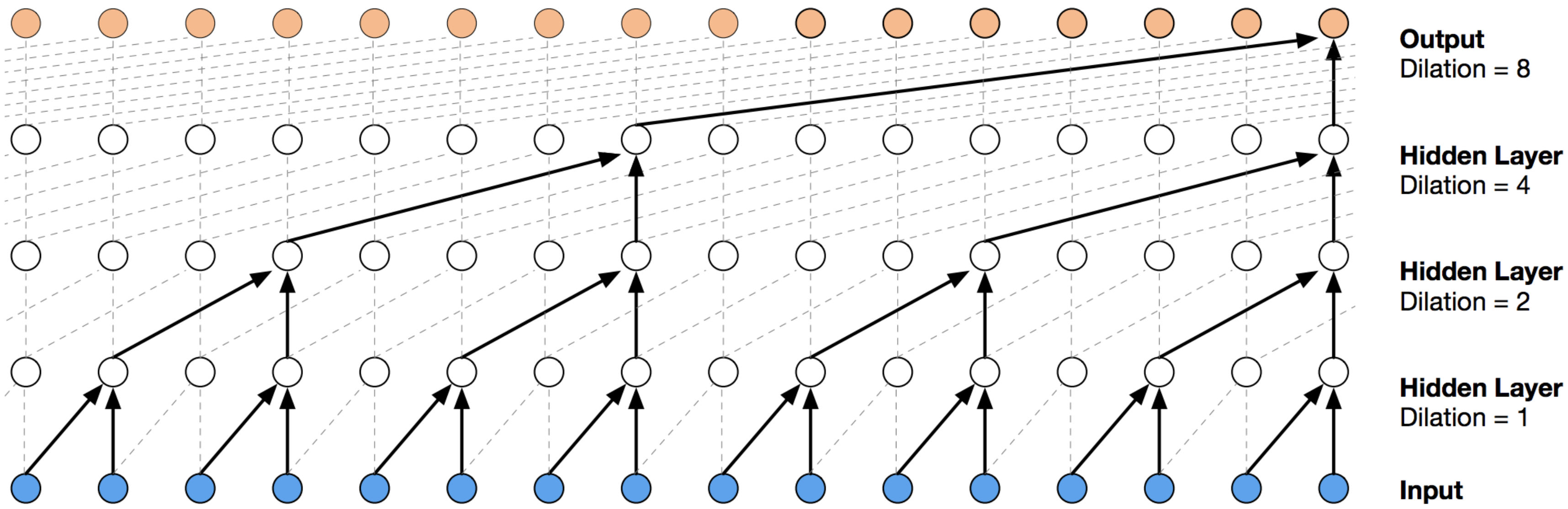
$$Conv_d(u) = (x *_d f)(u) = \sum_{i=0}^{k-1} f(i) \cdot x_{u-d\cdot i}$$

- Model weights are shared across graph snapshots.

TemporalGAT: Framework



TemporalGAT: Method



Stack of dilated causal convolutional layers (Wavenet, 2016)

Applications: Link Prediction

- Leverage structural and temporal information up to time step t and predict the existence of an edge between a pair of vertices at $t + 1$.
- Dynamic graph datasets used for performance evaluation:

Dataset	# Nodes	# Edges	# Time steps	Category
Enron	143	2,347	10	Communication
UCI	1,809	16,822	13	
Yelp	6,569	95,361	12	Rating

Results of Link Prediction

- State-of-the-art performance

Algorithm	Enron		UCI		Yelp	
	Micro	Macro	Micro	Macro	Micro	Macro
Node2Vec	83.7±0.7	83.1±1.2	80.0± 0.4	80.5±0.6	67.9±0.2	65.34±0.2
G-SAGE	82.5± 0.6	81.9±0.5	79.2± 0.4	82.9 ± 0.2	61.0±0.1	58.56±0.2
G-SAGE + GAT	72.5±0.4	73.3±0.6	74.0±0.4	79.8±0.2	66.2±0.1	65.1±0.2
GCN-AE	81.6±1.5	81.7±1.5	80.5±0.3	83.5±0.5	66.7±0.2	65.8±0.2
GAT-AE	75.7±1.1	76.0±1.4	80.0±0.2	81.9±0.3	65.9±0.1	65.4±0.1
DynamicTriad	80.3±0.8	79.0±0.9	77.6±0.6	80.3±0.5	63.5±0.3	62.7±0.3
Know-Evolve	61.6±1.1	62.3±1.5	71.2±0.5	80.9±0.2	56.9±0.2	59.7±0.2
DynGEM	67.8±0.6	69.7±1.3	77.5±0.3	79.8±0.5	66.0±0.2	66.0±0.2
DySAT	85.7±0.3	86.6±0.2	81.0±0.2	**85.8±0.1**	70.2±0.1	69.9±0.1
TemporalGAT	**86.4±0.4**	**86.8±0.3**	**82.7±0.2**	85.2±0.2	**71.9±0.3**	**70.3±0.2**

Summary

- Dynamic graph representation learning is a challenging research problem.
- TemporalGAT: towards combining structural information, node features and temporal evolutionary patterns during model training.
- Incorporating temporal information during training stage improves model capacity.
- State-of-the-art results on link prediction task.

Thank you