Temporal Graph Neural Networks and Their Expressiveness

Przemysław (Przemek) Wałęga[1], *joint work with Michael Rawson*[2]

`www.walega.pl` | `p.walega@qmul.ac.uk`

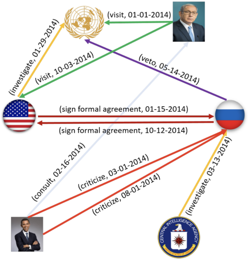
[1] Queen Mary University of London, University of Oxford,

[2] University of Southampton, TU Wien

Workshop on Foundations of Databases and AI, 30 January 2025

Message-Passing Temporal Graph Neural Networks (MP-TGNNs)

1. Take as an input a *temporal graph*:



2. Perform a fixed number of *message-passing iterations* modifying node embeddings.

3. Use *final node embeddings* to classify nodes/graph, predict missing edges, e.g. in:
   - recommender systems,
   - traffic forecasting,
   - finance networks,
   - modelling the spread of diseases,
   - ...

Research Questions

- What are the main *types of MP-TGNNs*?
- What are the *limits of their expressive power*?
- How do they *compare in terms of expressive power*?
- How does the difference in expressiveness *affect practical performance*?

Addressed in paper:
W. and Rawson, *Expressive Power of Temporal Message Passing*, AAAI 2025

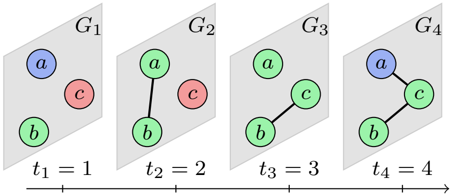
MP-TGNNs

Temporal Graph

Temporal graph is a finite sequence $TG = (G_1, t_1), \ldots, (G_n, t_n)$ of undirected node-coloured graphs stamped with real-valued t_i.

Assumptions:

- $t_1 < \cdots < t_n$,
- each G_i has the same set of vertices.



MP-TGNNs

MP-TGNNs compute *embedding* $\mathbf{h}_v^{(\ell)}(t)$ of v at time t in layer ℓ.

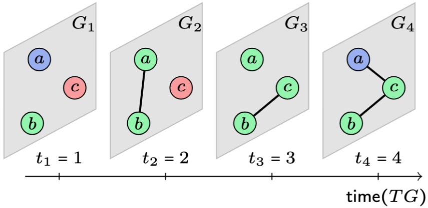
Depending on the computation we distinguish:

- *Global MP-TGNNs*
  Also called *temporal embedding* TGNN*s* [Longa et al. 2023] and used in such models as TGAT [Xu et al. 2020] or NAT [Luo and Li 2022].

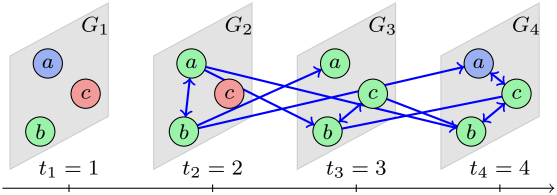
- *Local MP-TGNNs*
  Used for example in TGN [Rossi et al. 2020] and TDGNN [Qu et al. 2020].

Global MP-TGNNs

Input temporal graph:



Message-passing routes:



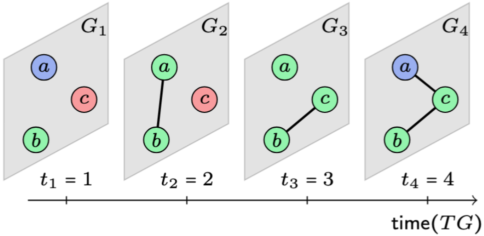
$$\mathbf{h}_v^{(\ell)}(t) = \mathsf{COM}^{(\ell)}\Big(\mathbf{h}_v^{(\ell-1)}(t), \mathsf{AGG}^{(\ell)}\Big(\{\!\{(\mathbf{h}_u^{(\ell-1)}(t'), g(t-t')) \mid (u,t') \in \mathcal{N}(v,t)\}\!\}\Big)\Big)$$

- $\mathcal{N}(v,t)$ is the *temporal neighbourhood of* (v,t):

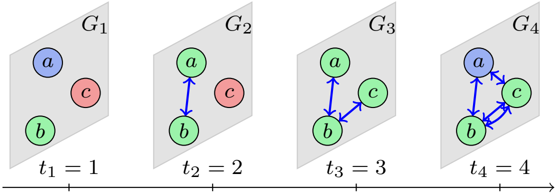
$$\mathcal{N}(v,t) = \Big\{(u,t') \mid \text{at } t' \leq t \text{ there is an edge } \{u,v\}\Big\}$$

Local MP-TGNNs

Input temporal graph:



Message-passing routes:



$$\mathbf{h}_v^{(\ell)}(t) = \mathsf{COM}^{(\ell)}\Big(\mathbf{h}_v^{(\ell-1)}(t), \mathsf{AGG}^{(\ell)}\Big(\{\!\{(\mathbf{h}_u^{(\ell-1)}(t), g(t-t')) \mid (u,t') \in \mathcal{N}(v,t)\}\!\}\Big)\Big)$$

- $\mathcal{N}(v,t)$ is the temporal neighbourhood of (v,t):

$$\mathcal{N}(v,t) = \Big\{(u,t') \mid \text{at } t' \leq t \text{ there is an edge } \{u,v\}\Big\}$$

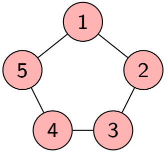
Global vs Local MP-TGNNs

- Which message-passing mechanism is *more expressive*?

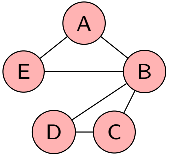
Weisfeiler-Leman Test

Weisfeiler-Leman Test

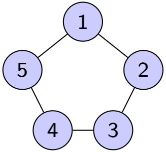
Weisfeiler-Leman test (WL) [Weisfeiler and Leman, 1968] is a heuristic for checking isomorphism of (static) graphs.



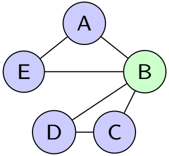
G, round 0



H, round 0



G, round 1



H, round 1

WL computes node colourings $\mathsf{wl}^{(0)}, \mathsf{wl}^{(1)}, \ldots$ as follows:

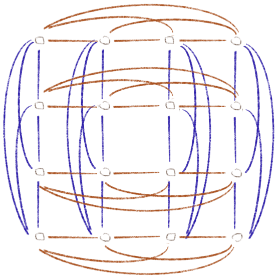
$$\mathsf{wl}^{(\ell)}(v) = HASH\Big(\mathsf{wl}^{(\ell-1)}(v), \{\{(\mathsf{wl}^{(\ell-1)}(u) \mid u \in \mathcal{N}(v)\}\}\Big),$$

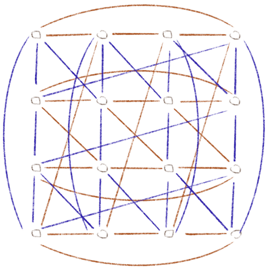
where $N(v)$ *is the neighbourhood of* v.

Weisfeiler-Leman Test

- *WL distinguishes* graphs ⇒ these graphs are *not isomorphic*.
- But the opposite implication **does not hold.**

For example 4-4 Rook's and Shirkande graphs are not isomorphic, but WL cannot distinguish them:





WL, GNNs, and Logic

The following are equivalent:

- Two graphs are *distinguishable by WL*,
- Some message-passing *GNN accepts only one* of these graphs,
- Some *formula of C_2 is true in only one* of these graphs.

Relational WL

Pablo and others, have recently considered GNNs and WL on *knowledge graphs* (multi-graphs with directed labelled edges) rwl where:

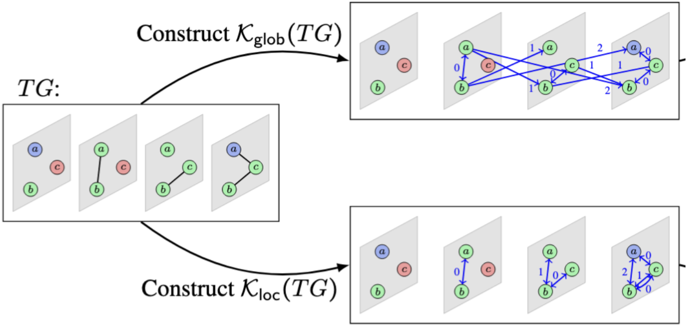
- with separate neighbourhood $\mathcal{N}_r$ for each relation r.

The following are equivalent: [Huang et al. 2023]

- Two knowledge graphs are *distinguishable by* rwl,
- Some message-passing *GNN accepts only one* of these knowledge graphs,

WL for MP-TGNNs

Convert a TG into Knowledge Graphs



Global MP-TGNNs

Theorem.

$\mathsf{rwl}^{(\ell)}(v,t) = \mathsf{rwl}^{(\ell)}(u,t')$ in $\mathcal{K}_{\text{glob}}(TG) \quad \Rightarrow \quad \mathbf{h}_v^{(\ell)}(t) = \mathbf{h}_u^{(\ell)}(t')$ in any global TGNN.

Theorem.
For each TG and $\ell \in \mathbb{N}$, there exists a global MP-TGNN $\mathcal{A}$ such that:

$\mathsf{rwl}^{(\ell)}(v,t) = \mathsf{rwl}^{(\ell)}(u,t')$ in $\mathcal{K}_{\text{glob}}(TG) \quad \Leftarrow \quad \mathbf{h}_v^{(\ell)}(t) = \mathbf{h}_u^{(\ell)}(t')$ in $\mathcal{A}$.

Local MP-TGNNs

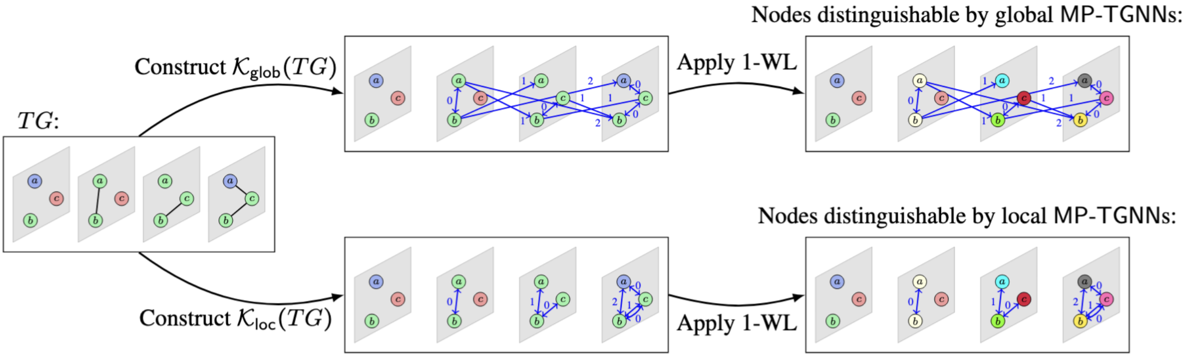
Theorem.

$\mathsf{rwl}^{(\ell)}(v,t) = \mathsf{rwl}^{(\ell)}(u,t')$ in $\mathcal{K}_{\mathsf{loc}}(TG) \quad \Rightarrow \quad \mathbf{h}_v^{(\ell)}(t) = \mathbf{h}_u^{(\ell)}(t')$ in any local TGNN.

Theorem.
For each TG and $\ell \in \mathbb{N}$, there exists a local MP-TGNN $\mathcal{A}$ such that:

$\mathsf{rwl}^{(\ell)}(v,t) = \mathsf{rwl}^{(\ell)}(u,t')$ in $\mathcal{K}_{\mathsf{loc}}(TG) \quad \Leftarrow \quad \mathbf{h}_v^{(\ell)}(t) = \mathbf{h}_u^{(\ell)}(t')$ in $\mathcal{A}$.

WL for MP-TGNNs



Temporal Isomorphisms

Pointwise Isomorphism

- GNNs can distinguish some *non-isomorphic graphs*,
- MP-TGNNs can distinguish some *non-isomorphic temporal graphs*.

What is a good notion of *isomorphism between temporal graphs*?

Pointwise Isomorphism

- GNNs can distinguish some *non-isomorphic graphs*,
- MP-TGNNs can distinguish some *non-isomorphic temporal graphs*.

What is a good notion of *isomorphism between temporal graphs*?

Definition.
$(G_1, t_1), \dots, (G_n, t_n)$ and $(G'_1, t'_1), \dots, (G'_m, t'_m)$ are *pointwise isomorphic* if:

- $t_i = t'_i$, and
- G_i and G'_i are isomorphic,

for all $i \in \{1, \dots, n\}$.

Is it a good definition?

Pointwise Isomorphism

- GNNs can distinguish some *non-isomorphic graphs*,
- MP-TGNNs can distinguish some *non-isomorphic temporal graphs*.

What is a good notion of *isomorphism between temporal graphs*?

Definition.
$(G_1, t_1), \dots, (G_n, t_n)$ and $(G'_1, t'_1), \dots, (G'_m, t'_m)$ are *pointwise isomorphic* if:

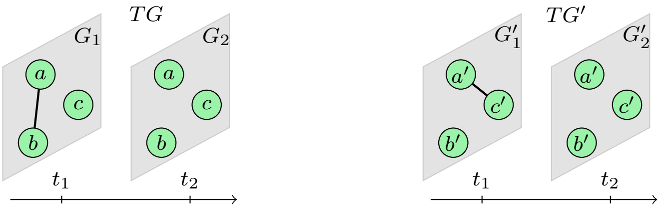
- $t_i = t'_i$, and
- G_i and G'_i are isomorphic,

for all $i \in \{1, \dots, n\}$.

Is it a good definition? NO!

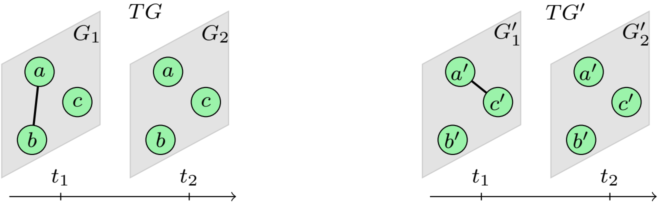
MP-TGNNs (global and local) **distinguish pointwise isomorphic temporal graphs!**

Consider *pointwise isomorphic* (a, t_2) *and* (a', t_2):

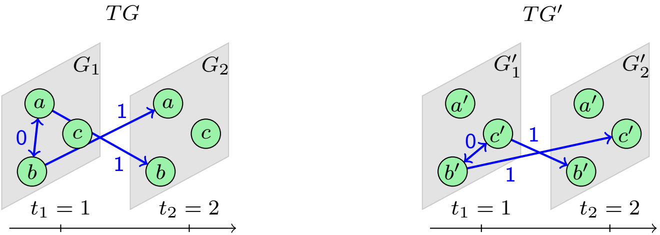


MP-TGNNs (global and local) **distinguish pointwise isomorphic temporal graphs!**

Consider *pointwise isomorphic* (a, t_2) *and* (a', t_2):



rwl$^{(1)}(a, t_2) \neq$ rwl$^{(1)}(a', t_2)$ in $\mathcal{K}_{\text{glob}}(TG)$, so **global MP-TGNNs distinguish** them:



Better Definition

Definition.
$(G_1, t_1), \ldots, (G_n, t_n)$ and $(G'_1, t'_1), \ldots, (G'_m, t'_m)$ are *timewise isomorphic* if:

- $t_{i+1} - t_i = t'_{i+1} - t'_i$, and
- there exists f which is an isomorphism between G_i and G'_i,

for every $i \in \{1, \ldots, n\}$.

Better Definition

Definition.
$(G_1, t_1), \ldots, (G_n, t_n)$ and $(G'_1, t'_1), \ldots, (G'_m, t'_m)$ are *timewise isomorphic* if:

- $t_{i+1} - t_i = t'_{i+1} - t'_i$, and
- there exists f which is an isomorphism between G_i and G'_i,

for every $i \in \{1, \ldots, n\}$.

Theorem.
If *(v, t) and (u, t') are timewise isomorphic*, then $\mathbf{h}_v^{(\ell)}(t) = \mathbf{h}_u^{(\ell)}(t')$ in any MP-TGNN (global or local) and any $\ell \in \mathbb{N}$.

Relative Expressiveness of MP-TGNNs

Results

We can show the following *expressiveness results:*

global MP-TGNNs, any TGs	$>$	global MP-TGNNs, colour-persistent TGs
$\not\leq$ $\not\geq$		$\wedge$
local MP-TGNNs, any TGs	$=$	local MP-TGNNs, colour-persistent TGs

Experiments

Experiments

Goal: Check how expressiveness results impact performance.

Setup:

- *Directly implemented* MP-TGNNs (global and local) with PyTorch,
- Tested on link prediction tasks from *TGB 2.0:*
  - `tgbl-wiki`, `tgbl-review`, `tgbl-coin`,

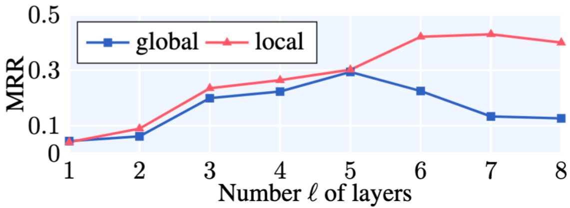
	`tgbl-wiki`	`tgbl-review`	`tgbl-coin`
nodes	9,227	352,637	638,486
edges	157,474	4,873,540	22,809,486

Results

Mean reciprocal rank (MRR) *aligns with our expressiveness results*:

	`tgbl-wiki`	`tgbl-review`	`tgbl-coin`
global	0.223	0.321	0.628
local	**0.264**	**0.359**	**0.635**

MRR on `tgbl-wiki` with *increasing number of layers*:



Conclusions

- Expressive power of MP-TGNNs can be analysed by applying *WL to corresponding knowledge graphs*,
- Expressive power of global and local MP-TGNNs is *incomparable*,
- *On colour-persistent* temporal graphs local MP-TGNNs are *strictly more expressive*,
- Experiments show *higher performance of local MP-TGNNs*.

Future work:

- If/what correspondence can we establish between TGNNs and *temporal logics*?

Thank you for your attention!

Przemysław (Przemek) Wałęga | www.walega.pl | p.walega@qmul.ac.uk

Always happy to discuss more!