

Representation Learning with Efficient Transformers

Maciej Besta, Afonso Claudino Catarino, Lukas Gianinazzi, Nils Blach, Piotr Nyczyk, Hubert Niewiadomski, Torsten Hoefler

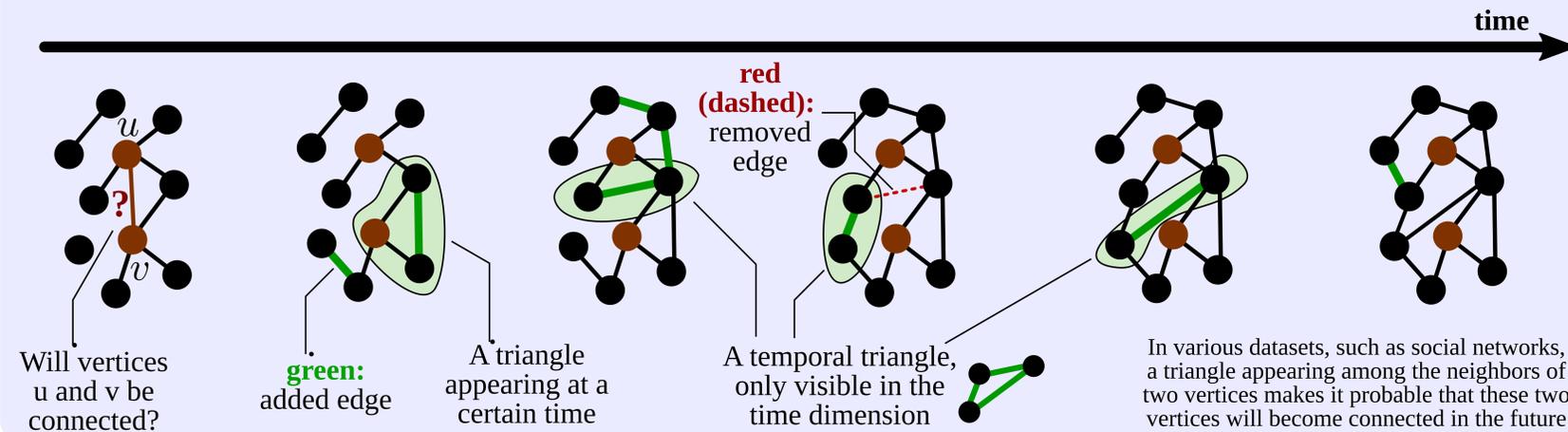
Department of Computer Science, ETH Zurich



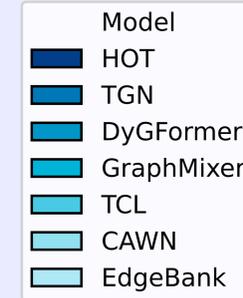
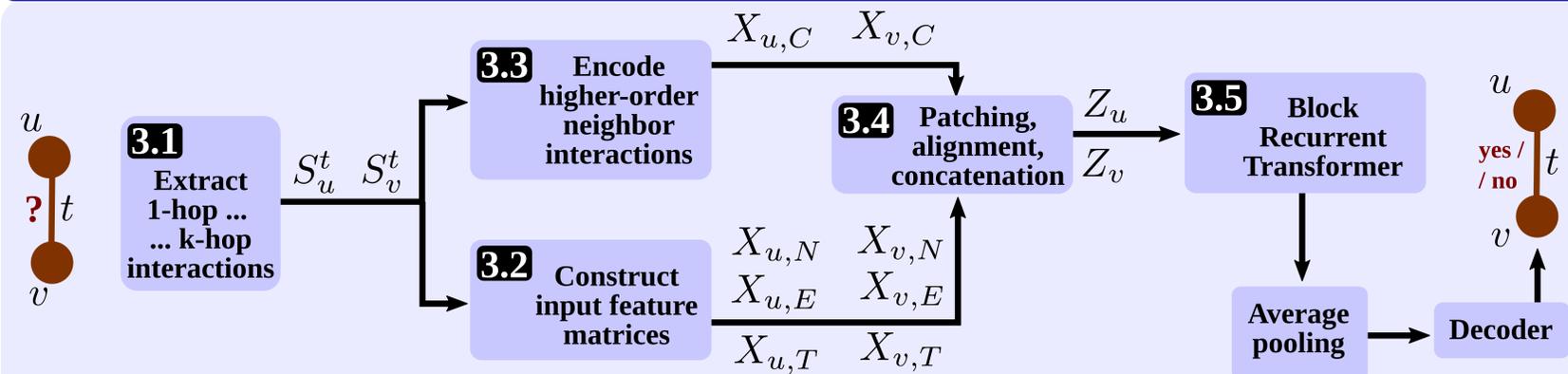
Motivation

A growing amount of graph representation learning (GRL) workloads are dynamic, with millions of edges added or removed per second. A fundamental workload in such a setting is dynamic link prediction: using a history of graph updates to predict whether a given pair of vertices will become connected. Recent schemes for link prediction in such dynamic settings employ Transformers, modeling individual graph updates as single tokens. In this work, we propose HOT: a model that enhances this line of works by harnessing higher-order (HO) graph structures.

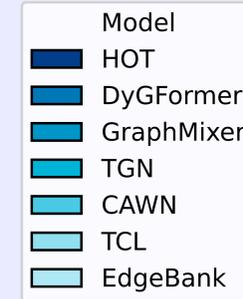
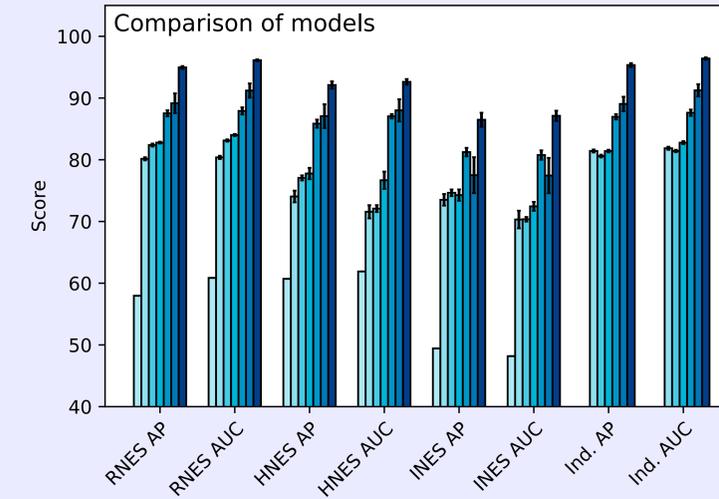
Temporal Higher-Order Example



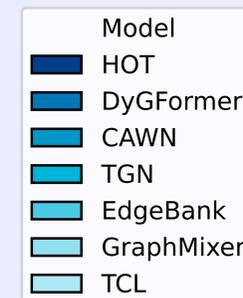
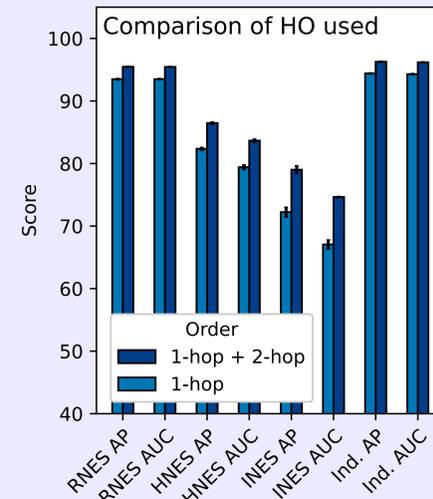
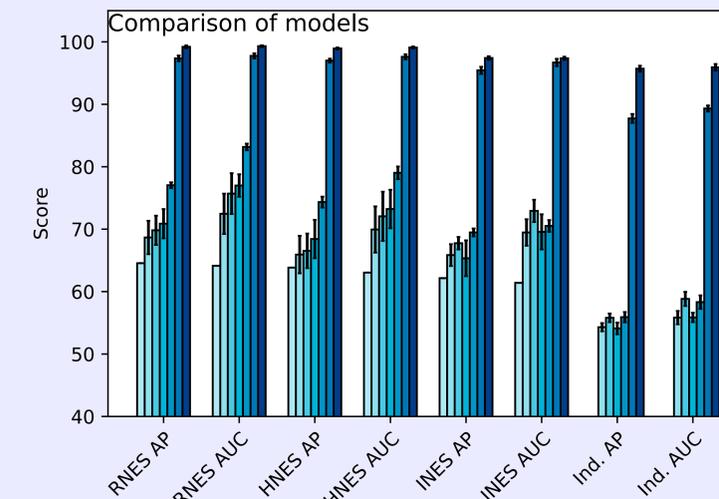
The Design of HOT



The MOOC graph dataset



The LastFM graph dataset



The CanParl graph dataset

