Data Science and Engineering (2023) 8:98–111
https://doi.org/10.1007/s41019-023-00207-w

RESEARCH PAPER

Memory-Enhanced Transformer for Representation Learning on Temporal Heterogeneous Graphs

Longhai Li[1] · Lei Duan[1,3] · Junchen Wang[1] · Chengxin He[1] · Zihao Chen[1] · Guicai Xie[1] · Song Deng[2] · Zhaohang Luo[4]

Received: 4 December 2022 / Revised: 7 February 2023 / Accepted: 1 March 2023 / Published online: 22 March 2023

Abstract
Temporal heterogeneous graphs can model lots of complex systems in the real world, such as social networks and e-commerce applications, which are naturally time-varying and heterogeneous. As most existing graph representation learning methods cannot efficiently handle both of these characteristics, we propose a Transformer-like representation learning model, named THAN, to learn low-dimensional node embeddings preserving the topological structure features, heterogeneous semantics, and dynamic patterns of temporal heterogeneous graphs, simultaneously. Specifically, THAN first samples heterogeneous neighbors with temporal constraints and projects node features into the same vector space, then encodes time information and aggregates the neighborhood influence in different weights via type-aware self-attention. To capture long-term dependencies and evolutionary patterns, we design an optional memory module for storing and evolving dynamic node representations. Experiments on three real-world datasets demonstrate that THAN outperforms the state-of-the-arts in terms of effectiveness with respect to the temporal link prediction task.

Keywords Temporal heterogeneous graphs · Graph neural networks · Graph representation learning · Transformer

1 Introduction

Graph representation learning, as an important task in machine learning, has significant practical value in areas such as social networks and recommendation systems. Existing graph representation learning methods usually take static graphs as the input to obtain low-dimensional embeddings by encoding local non-Euclidean structures and have achieved extensive excellent performance in downstream tasks such as link prediction [1–3], node classification [4, 5], and graph classification [6, 7].

However, most graphs in the real world are naturally heterogeneous and dynamic, which cannot be accurately represented by static homogeneous graphs. Several studies incorporate heterogeneous data models into a unified graph model [8], promoting the research of graph data. Taking the example of a user-item interaction network in e-commerce scenarios, illustrated in Fig. 1a, there are two types of nodes (i.e., *user* and *item*) and three types of interactions (i.e., *favorite*, *browse* and *buy*). Additionally, each interaction is associated with a continuous timestamp to indicate when it occurred. In this paper, we define such interaction sequences

✉ Lei Duan
leiduan@scu.edu.cn

Longhai Li
lilonghai@stu.scu.edu.cn

Junchen Wang
wangjunchen@stu.scu.edu.cn

Chengxin He
hechengxin@stu.scu.edu.cn

Zihao Chen
chenzihao@stu.scu.edu.cn

Guicai Xie
guicaixie@stu.scu.edu.cn

Song Deng
dengsong@njupt.edu.cn

Zhaohang Luo
luozhaohang@qq.com

1 School of Computer Science, Sichuan University, Chengdu, China

2 Institute of Advanced Technology, Nanjing University of Posts & Telecommunications, Nanjing, China

3 Med-X Center for Informatics, Sichuan University, Chengdu, China

4 Nuclear Power Institute of China, Chengdu, China