

THÔNG TIN LUẬN ÁN

Tên đề tài luận án: Dự đoán sự kiện trên đồ thị tri thức thời gian

Ngành: Khoa học máy tính

Mã số ngành: 9480101

Họ tên nghiên cứu sinh: Nguyễn Ngọc Trung

Khóa đào tạo: 2020

Người hướng dẫn khoa học: GS.TS. Bùi Xuân Hải, TS. Bùi Xuân Bình Minh

Cơ sở đào tạo: Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM

1. TÓM TẮT NỘI DUNG LUẬN ÁN:

Đồ thị tri thức thời gian (Temporal Knowledge Graph – TKG) mở rộng đồ thị tri thức truyền thống bằng cách gắn thông tin thời gian vào các quan hệ và sự kiện, cho phép mô hình hóa sự tiến hóa của tri thức theo thời gian. Luận án nghiên cứu bài toán dự đoán sự kiện trên TKG, gồm hai lớp bài toán con có liên hệ chặt chẽ: nội suy – hoàn chỉnh các sự kiện hoặc liên kết còn thiếu trong khoảng thời gian đã quan sát; và ngoại suy – dự báo các sự kiện xảy ra ngoài biên thời gian đã biết. Trên cơ sở đó, luận án đề xuất ba mô hình học biểu diễn theo một mạch thống nhất, đi từ nội suy đến ngoại suy: (i) mô hình khai thác không gian số phức kép và cơ chế quay để biểu diễn đồng thời thực thể, quan hệ và thời gian; (ii) mô hình dựa trên phân rã tensor kết hợp cơ chế trọng số thời gian linh hoạt; và (iii) mô hình dựa trên kiến trúc Transformer kết hợp tích chập đồ thị để khai thác và ngoại suy các sự kiện lịch sử phục vụ dự báo tương lai. Các mô hình được đánh giá trên nhiều bộ dữ liệu chuẩn theo các độ đo Mean Reciprocal Rank (MRR), Hits@k và Mean Absolute Error (MAE).

2. NHỮNG KẾT QUẢ MỚI CỦA LUẬN ÁN:

- **Mô hình TBicomR:** mô hình nhúng đồ thị tri thức thời gian trên không gian số phức kép, biểu diễn đồng thời thực thể, quan hệ và thời gian bằng cơ chế quay, hợp nhất phép quay tròn và phép quay hyperbol trong một biểu diễn. Mô hình biểu diễn được các mẫu quan hệ đối xứng, phản đối xứng, nghịch đảo và quan hệ chu kỳ theo thời gian với độ phức tạp thời gian tuyến tính $O(d)$; được đánh giá trên các bộ dữ liệu ICEWS14, ICEWS05-15, GDELT, YAGO11k và Wikidata12k.

- **Mô hình FTPComplex:** mô hình nhúng đồ thị tri thức thời gian dựa trên phân rã tensor, bổ sung cơ chế trọng số thời gian linh hoạt với hai tham số điều khiển góc nhìn thời gian, cho phép điều chỉnh mức ảnh hưởng của thông tin thời gian theo đặc điểm từng bộ dữ liệu, trong khi giữ độ phức tạp tương đương mô hình nền TPComplex. Trên bộ dữ liệu YAGO15k, mô hình cải thiện MRR khoảng 9% và Hits@1 khoảng 11% so với mô hình nền dưới cùng một thiết lập đánh giá.
- **Mô hình HGCT (History Graph Convolution Transformer):** mô hình suy luận đồ thị tri thức thời gian cho bài toán ngoại suy, kết hợp Fact Graph Transformer (khai thác cấu trúc đồ thị lịch sử) với Temporal Convolution Transformer (mã hóa vị trí theo thời gian, nắm bắt phụ thuộc ngắn hạn và dài hạn) và bộ giải mã Query-ConvTransE. Trên sáu bộ dữ liệu ICEWS14, ICEWS18, ICEWS05-15, GDELT, YAGO và WIKI theo thiết lập lọc có nhận biết thời gian, mô hình cải thiện Hits@k khoảng 4,33% và MRR khoảng 3,84% so với phương pháp tiên tiến.

3. CÁC ỨNG DỤNG/KHẢ NĂNG ỨNG DỤNG TRONG THỰC TIỄN HAY NHỮNG VẤN ĐỀ CÒN BỎ NGỎ CẦN TIẾP TỤC NGHIÊN CỨU:

Các mô hình đề xuất có khả năng ứng dụng trong các lĩnh vực có dữ liệu động và nhu cầu dự báo sự kiện theo thời gian như tài chính, y tế, mạng xã hội, thương mại điện tử, an ninh mạng và hỗ trợ ra quyết định, phục vụ các tác vụ như dự báo xu hướng, phát hiện bất thường, phân tích hành vi và suy luận tri thức từ dữ liệu lịch sử. Các hướng nghiên cứu tiếp theo bao gồm: xử lý dữ liệu quy mô lớn, dữ liệu nhiều hoặc thiếu; học biểu diễn trong môi trường thời gian liên tục; nâng cao khả năng giải thích của mô hình; tối ưu chi phí tính toán; và kết hợp với dữ liệu đa phương thức cũng như mô hình ngôn ngữ lớn.

THESIS INFORMATION

Thesis title: Event Prediction on Temporal Knowledge Graphs

Speciality: Computer Science

Code: 9480101

Name of PhD Student: Nguyen Ngoc Trung

Academic year: 2020

Supervisor: Prof. Dr. Bui Xuan Hai, Dr. Bui Xuan Binh Minh

At: VNUHCM – University of Science

1. SUMMARY:

Temporal Knowledge Graphs (TKGs) extend traditional knowledge graphs by attaching temporal information to relations and events, enabling the modeling of how knowledge evolves over time. This dissertation studies event prediction on TKGs, which comprises two closely related sub-problems: interpolation — completing missing events or links within the observed time range; and extrapolation — forecasting events beyond the known temporal boundary. On this basis, the dissertation proposes three representation-learning models forming a coherent progression from interpolation to extrapolation: (i) a model that leverages the bicomplex space and a rotation mechanism to jointly represent entities, relations, and time; (ii) a model based on tensor factorization augmented with a flexible temporal-weighting mechanism; and (iii) a model based on a Transformer architecture combined with graph convolution to exploit and extrapolate historical events for future forecasting. The models are evaluated on several benchmark datasets using Mean Reciprocal Rank (MRR), Hits@k, and Mean Absolute Error (MAE).

2. NOVELTY OF THESIS:

- **TBicomR:** a temporal knowledge graph embedding model in the bicomplex space that jointly represents entities, relations, and time through a rotation mechanism, unifying circular and hyperbolic rotations within a single representation. The model captures symmetric, antisymmetric, inverse, and cyclic temporal relational patterns with linear time complexity $O(d)$; it is

evaluated on ICEWS14, ICEWS05-15, GDELT, YAGO11k, and Wikidata12k.

- **FTPCmplEx:** a tensor-factorization-based temporal knowledge graph embedding model that introduces a flexible temporal-weighting mechanism with two time-perspective control parameters, allowing the influence of temporal information to be adjusted according to the characteristics of each dataset while keeping computational complexity comparable to the TPCmplEx baseline. On the YAGO15k dataset, the model improves MRR by about 9% and Hits@1 by about 11% over the baseline under the same evaluation setting.
- **HGCT (History Graph Convolution Transformer):** a temporal knowledge graph reasoning model for the extrapolation task that combines a Fact Graph Transformer (to exploit historical graph structure) with a Temporal Convolution Transformer (temporal position encoding capturing short- and long-term dependencies) and a Query-ConvTransE decoder. On six datasets — ICEWS14, ICEWS18, ICEWS05-15, GDELT, YAGO, and WIKI — under the time-aware filtered setting, the model improves Hits@k by about 4.33% and MRR by about 3.84% over the state-of-the-art.

3. APPLICATIONS / APPLICABILITY / PERSPECTIVE:

The proposed models are applicable to domains involving dynamic data and temporal event forecasting, such as finance, healthcare, social networks, e-commerce, cybersecurity, and decision support, serving tasks such as trend forecasting, anomaly detection, behavior analysis, and knowledge reasoning from historical data. Future research directions include handling large-scale, noisy, or incomplete data; representation learning in continuous-time settings; improving model interpretability; optimizing computational cost; and integrating multimodal data and large language models.

DANH MỤC CÔNG TRÌNH CỦA NGHIÊN CỨU SINH

Tạp chí quốc tế

- [CT01]. Nguyen, Ngoc-Trung, Tran, Chi, & Le, Thanh (2024). TBicomR: Event Prediction in Temporal Knowledge Graphs with Bicomplex Rotation. Knowledge-Based Systems, 112711. <https://doi.org/10.1016/j.knosys.2024.112711>. Q1, IF 7.6.
- [CT02]. Nguyen, Ngoc-Trung, Ngo, Thuc, Hoang, Nguyen, & Le, Thanh (2024). FTPComplEx: A flexible time perspective approach to temporal knowledge graph completion. Engineering Applications of Artificial Intelligence, 139, 109717. <https://doi.org/10.1016/j.engappai.2024.109717>. Q1, IF 8.0
- [CT03]. Hoa Dao, Nguyen Phan, Thanh Le, Ngoc-Trung Nguyen, HGCT: Enhancing temporal knowledge graph reasoning through extrapolated historical fact extraction, Knowledge-Based Systems 316 (2025) 113358. <https://doi.org/10.1016/j.knosys.2025.113358>. Q1, IF 7.6

Hội nghị quốc tế

- [CT04]. Nguyen, Ngoc-Trung, Vu, Thanh, & Le, Thanh (2024). T-WinG: Windowing for Temporal Knowledge Graph Completion. ESANN 2024 Proceedings, 393–398. <https://doi.org/10.14428/esann/2024.ES2024-166>. Rank B.