**TRANG THÔNG TIN LUẬN ÁN**

Tên đề tài luận án: Dự đoán liên kết trong đồ thị tri thức sử dụng nhúng dịch chuyển và mạng tích chập.

Ngành: Khoa học máy tính

Mã số ngành: 9480101

Họ tên nghiên cứu sinh: Lê Ngọc Thành

Khóa đào tạo: 2020

Người hướng dẫn khoa học: GS.TS. Lê Hoài Bắc

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

**1. TÓM TẮT NỘI DUNG LUẬN ÁN**:

 Vào nửa đầu năm 2012, Google đã giới thiệu một loại dữ liệu mới mang tên Đồ thị Tri thức. Đây là một cấu trúc đặc biệt, được thiết kế để lưu trữ thông tin ngữ nghĩa trong các đồ thị đa quan hệ. Nó đóng vai trò quan trọng trong việc thực hiện nhiều nhiệm vụ như dự đoán liên kết, phân lớp các đỉnh và cạnh, cũng như phát hiện cộng đồng. Google khẳng định rằng với kiến trúc này, công cụ tìm kiếm của họ đã trở nên thông minh hơn 1000 lần. Tuy nhiên, đồ thị tri thức được xây dựng từ nhiều nguồn thông tin khác nhau, bao gồm cả dữ liệu mở, dẫn đến các vấn đề như thông tin không chính xác, nhiễu và mâu thuẫn. Đây là một đặc tính phổ biến của các tập dữ liệu đồ thị tri thức hiện nay. Chính vì thế, các nghiên cứu nhằm hoàn thiện đồ thị tri thức đang ngày càng được phát triển, trong đó có bài toán dự đoán liên kết. Bài toán này có nhiều ứng dụng thực tế, từ hệ thống gợi ý, truy vết đối tượng có nguy cơ lây nhiễm sinh học, dự đoán hành động của các hội nhóm, đến hệ thống trả lời câu hỏi.

 Luận án đặt mục tiêu cải thiện độ chính xác trong dự đoán liên kết trên đồ thị tri thức. Qua quá trình nghiên cứu và khảo sát các phương pháp liên quan, luận án đã xác định được rằng các phương pháp dựa trên nhúng đồ thị cho thấy tiềm năng lớn. Trong ba hướng nhúng khác nhau phục vụ cho bài toán dự đoán liên kết, nhóm dịch chuyển và mạng nơ-ron nổi bật với nhiều ưu điểm trong việc thiết kế kiến trúc cải tiến. Nhóm dịch chuyển có tính trực quan và khả năng giải thích quá trình liên kết, trong khi nhóm mạng nơ-ron chứng minh khả năng mô hình hóa với kết quả tốt hơn so với các phương pháp khác. Từ những nhận định này, luận án đã đề xuất bốn giải pháp cải tiến: 1) Áp dụng mạng HyperNetwork trong mô hình dựa trên mạng nơ-ron để tạo các bộ lọc tốt cho quá trình khai thác tương tác giữa thực thể và quan hệ, đồng thời khởi tạo trọng số giúp mạng tìm kiếm nhanh hơn và tăng chiều sâu của mạng; 2) Thực hiện chiếu thực thể trên siêu phẳng để xác định quan hệ, kết hợp với phép xoay phức; 3) Khai thác không gian Quaternion và lý thuyết nhóm trong phép xoay kép và xoay đẳng hướng để xác định thực thể đuôi; 4) Kết hợp phép tích chập với phép xoay để tăng cường thông tin trong quá trình biến đổi. Mỗi đề xuất này đã được thử nghiệm trên các bộ dữ liệu chuẩn và đánh giá từ nhiều khía cạnh để chứng minh hiệu quả dự đoán của mô hình.

**2. NHỮNG KẾT QUẢ MỚI CỦA LUẬN ÁN**:

Qua quá trình thực hiện, các đóng góp mới về mặt hiệu suất trong các mô hình đề xuất gồm:

* **Mô hình ConvHyper:** Đây là một cải tiến của mô hình ConvKB, áp dụng HyperNetwork để tạo ra các bộ lọc xác định quan hệ nhằm tăng khả năng nắm bắt tương tác giữa các thực thể. Mạng này còn được dùng để tạo bộ trọng số giúp giảm thời gian tìm kiếm trọng số và tăng độ sâu của mạng. Hiệu suất cải thiện từ 0.8% đến 5.5% trên các tập dữ liệu ở các độ đo chuẩn.
* **Mô hình RotatPRH:** Phương pháp biểu diễn thực thể và quan hệ trong không gian số phức cùng với phép xoay trên từng siêu phẳng xác định mối quan hệ đã giúp tăng khả năng biểu diễn, từ đó cải thiện hiệu suất của mô hình dự đoán. Bên cạnh đó, mô hình còn áp dụng các ràng buộc mềm để đảm bảo việc huấn luyện không làm sai các tính chất của không gian nhúng. Bằng phương pháp này, hiệu suất của mô hình tăng từ 1.4% đến 2% trên bộ dữ liệu không bị rò rĩ kiểm tra. Đặc biệt trên bộ kích thước lớn như YAGO3-10, mô hình cải thiện lên đến 15% trên độ đo Hits@1.
* **Mô hình Rotate4D:** Dựa trên thành công của khai thác không gian số phức, mô hình Rotate4D thực hiện nhúng thực thể và quan hệ vào không gian phức cao chiều hơn, cụ thể là không gian Quaternion và thực hiện phép xoay để dự đoán liên kết. Mô hình triển khai với hai phiên bản: xoay kép và xoay đẳng hướng trên hai mặt phẳng riêng biệt, cả hai đều bảo tồn tính chất của đồ thị và mô hình hóa hiệu quả các loại quan hệ. Phiên bản xoay kép nổi bật với tính tổng quát cao hơn. Lý thuyết nhóm, đặc biệt là các nhóm trực giao, được sử dụng để đảm bảo tính đúng đắn trong việc biểu diễn các phép xoay. Thêm vào đó, phép co giãn Quaternion sau xoay giúp xử lý quan hệ phân cấp. Kết quả cải thiện từ 1.53% đến 3.49% khi so với mô hình cơ sở. Nếu chỉ xét trên quan hệ phân cấp, mô hình cải thiện lên đến 12%.
* **Mô hình ConvRot:** Tích chập có khả năng khai thác tốt tương tác, trong khi đó phép xoay mô hình hóa phép biến đổi từ thực thể đầu đến thực thể đuôi thông qua quan hệ. Bằng cách kết hợp điểm mạnh của hai hướng này, mô hình đã đạt kết quả cao hơn so với các đề xuất trước đó, mở ra hướng nghiên cứu về cách thiết kế mô hình với sự tổng hợp thế mạnh của mỗi nhóm để tăng hiệu suất nhưng không ảnh hưởng quá nhiều đến thời gian huấn luyện hay thực thi. Mô hình ConvRot giúp cải thiện đến 7% so với kết quả của các công trình trước đây.

**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 được đề xuất trong luận án có tiềm năng trong ứng dụng thực tế, từ việc cải thiện dữ liệu trong đồ thị đến phát triển hệ thống gợi ý và phân tích tương tác giữa các phần tử trong hợp chất. Một số ứng dụng bước đầu đã triển khai nền tảng lý thuyết và kiến trúc của các mô hình này gồm hệ thống Chatbot tiếng Việt sử dụng truy vấn dựa trên bản nhúng đồ thị và hệ thống phát hiện thuốc trong đơn thuốc. Tuy nhiên, vẫn còn các thách thức cần giải quyết để ứng dụng khả thi hơn trong thực tế bao gồm việc triển khai mô hình trên các dạng đồ thị có yếu tố thời gian, khai thác tối đa các tính chất trong không gian đa tạp, cung cấp giải thích cho các mô hình và kết quả dự đoán trong mạng nơ-ron, cùng với việc tối ưu hóa hiệu suất để xử lý dữ liệu đồ thị lớn một cách hiệu quả.

|  |  |
| --- | --- |
| **TẬP THỂ CÁN BỘ HƯỚNG DẪN**(Ký tên, họ tên) | **NGHIÊN CỨU SINH**(Ký tên, họ tên) |

 **Lê Hoài Bắc Lê Ngọc Thành**

**XÁC NHẬN CỦA CƠ SỞ ĐÀO TẠO**

**HIỆU TRƯỞNG**

**THESIS INFORMATION**

Thesis title: Link Prediction on Knowledge Graphs using Transition-based and Convolutional Network-based Embedding

Speciality: Computer Science

Code: 9480101

Name of PhD Student: Le Ngoc Thanh

Academic year: 2020

Supervisor: Prof. Dr. Le Hoai Bac

At: VNUHCM - University of Science

**1. SUMMARY**:

 In the first half of 2012, Google introduced a new type of data called the Knowledge Graph. This is a unique structure designed to store semantic information in multi-relational graphs. It plays a crucial role in performing various tasks such as link prediction, classification of vertices and edges, as well as community detection. Google asserts that with this architecture, their search engine has become 1000 times smarter. However, the knowledge graph is built from various information sources, including open data, leading to issues like inaccurate information, noise, and contradictions. This is a common characteristic of current knowledge graph datasets. Therefore, research aimed at improving the knowledge graph is increasingly being developed, including the problem of link prediction. This issue has many practical applications, from recommendation systems, tracing potentially infected biological subjects, predicting the actions of groups, to question-answering systems.

 The thesis aims to improve accuracy in predicting links on knowledge graphs. Through the research and examination of related methods, the thesis has identified that methods based on graph embedding show great potential. Among three different embedding directions serving the link prediction problem, the translation group and neural networks stand out with many advantages in designing improved architectures. The translation group is intuitive and capable of explaining the linking process, while the neural network group demonstrates better modeling capabilities with superior results compared to other methods. From these observations, the thesis has proposed four improvement solutions: 1) Applying HyperNetwork in neural network-based models to create better filters for the process of mining interactions between entities and relations, while also initiating weights to help the network search faster and increase network depth; 2) Implementing entity projection on hyperplanes for relation determination, combined with complex rotation; 3) Exploiting Quaternion space and group theory in double and isotropic rotations to identify tail entities; 4) Integrating convolution with rotation to enhance information in the transformation process. Each proposal has been tested on standard datasets and evaluated from multiple aspects to demonstrate the model's predictive effectiveness.

**2. NOVELTY OF THESIS**:

Throughout its implementation, the new contributions in terms of performance in the proposed models include:

- **ConvHyper model**: This is an improvement of the ConvKB model, applying HyperNetwork to create relationship-defining filters to enhance the ability to capture interactions between entities. This network is also used to create a set of weights that help reduce the weight search time and increase the depth of the network. The performance improvement ranges from 0.8% to 5.5% on datasets with standard metrics.

- **RotatPRH model**: The method of representing entities and relations in complex space, along with rotation on each hyperplane to define relationships, has helped enhance representation capabilities, thereby improving the performance of the prediction model. In addition, the model applies soft constraints to ensure that training does not violate the properties of the embedding space. With this method, the model's performance has increased from 1.4% to 2% on datasets without test leakage. Particularly on large datasets like YAGO3-10, the model shows an improvement of up to 15% in the Hits@1 metric.

- **Rotate4D model**: Building on the success of exploiting complex space, the Rotate4D model embeds entities and relations into a higher-dimensional complex space, specifically Quaternion space, and performs rotations to predict links. The model is implemented in two versions: double rotation and isotropic rotation on two separate planes, both preserving the properties of the graph and effectively modeling various types of relationships. The double rotation version stands out for its higher generality. Group theory, especially orthogonal groups, is used to ensure accuracy in representing rotations. Additionally, Quaternion scaling after rotation helps in handling hierarchical relationships. The results show an improvement of 1.53% to 3.49% compared to the baseline model. When focusing solely on hierarchical relationships, the model improves by up to 12%.

- **ConvRot model**: Convolution excels in extracting interactions, while rotation effectively models the transformation from the head entity to the tail entity through relationships. By combining the strengths of these two approaches, the model has achieved higher results than previous proposals, opening a research direction on designing models that synthesize the strengths of each group to increase efficiency without significantly impacting training or execution time. The ConvRot model helps improve performance by up to 7% compared to the results of previous works.

**3**. **APPLICATIONS/ APPLICABILITY/ PERSPECTIVE**

 The models proposed in the dissertation have potential in practical applications, ranging from improving data in graphs to developing recommendation systems and analyzing interactions among elements in compounds. Some initial applications that have implemented the theoretical foundations and architectures of these models include a Vietnamese Chatbot system using graph-embedding-based queries and a system for detecting drugs in prescriptions. However, there are still challenges to be addressed for more feasible real-world applications. These include deploying the model on time-sensitive graphs, maximizing the properties in manifold spaces, providing explanations for neural network models and their predictions, and optimizing performance to efficiently handle large graph data.

|  |  |
| --- | --- |
|  **SUPERVISOR** | **PhD STUDENT** |

 **Le Hoai Bac Le Ngoc Thanh**

**CONFIRMATION UNIVERSITY OF SCIENCE**

**PRESIDENT**