

TRÍCH XUẤT MỐI QUAN HỆ PHẢN XẠ CÓ YẾU TỐ THỜI GIAN CỦA MÔ HÌNH TIME-ER TỪ MỘT CƠ SỞ DỮ LIỆU QUAN HỆ

Hoàng Quang

Trường Đại học Khoa học, Đại học Huế

Email: hquang@husc.edu.vn

Ngày nhận bài: 16/6/2022; ngày hoàn thành phản biện: 17/6/2022; ngày duyệt đăng: 22/6/2022

TÓM TẮT

Bài toán trích xuất một mô hình cơ sở dữ liệu thời gian ở mức khái niệm từ một mô hình quan hệ cho trước được xem là một ánh xạ ngược của bài toán chuyển đổi thuận trong quy trình thiết kế cơ sở dữ liệu thời gian. Trên thực tế, việc giải quyết bài toán này sẽ cho phép điều tra được mô hình theo thời gian ở mức khái niệm, để từ đó có thể nâng cấp một hệ thống thông tin có yếu tố thời gian trên cơ sở kế thừa dữ liệu đã có. Nghiên cứu này nhằm bổ sung các quy tắc nhận dạng mối quan hệ phản xạ có yếu tố thời gian của mô hình TimeER từ một mô hình quan hệ cho trước.

Từ khóa: mô hình cơ sở dữ liệu mức khái niệm, mô hình TimeER, mối quan hệ phản xạ có yếu tố thời gian, mô hình cơ sở dữ liệu quan hệ.

1. GIỚI THIỆU

Để giải quyết vấn đề thiết kế các cơ sở dữ liệu (CSDL) thời gian ở mức khái niệm, đã có nhiều đề xuất về các mô hình ER thời gian, như các mô hình sau: TERM, RAKE, MOTAR, TEER, STEER, ERT, TimeER [6][7]. Trong đó mô hình TimeER (Time-Extended-EER) được xem là một mở rộng của mô hình EER [3] bằng cách cho phép hỗ trợ đầy đủ các yếu tố thời gian hơn so với các mô hình khác. Trên cơ sở đó, người ta có thể thiết kế các mô hình dữ liệu thời gian mức logic. Liên quan đến mô hình TimeER, đã có những đề xuất về các phương pháp chuyển đổi từ mô hình TimeER sang mô hình quan hệ [1][2][5][7].

Vấn đề trích xuất một mô hình mức khái niệm từ một mô hình mức logic được gọi là bài toán “chuyển đổi ngược” của bài toán chuyển đổi truyền thống trong quy trình thiết kế cơ sở dữ liệu được đề cập ở trên. Trên thực tế, việc giải quyết bài toán chuyển đổi ngược này trên các CSDL quan hệ có yếu tố thời gian sẽ cho phép chúng ta

điều tra được mô hình theo thời gian ở mức khái niệm. Từ đó có thể nâng cấp một hệ thống thông tin có yếu tố thời gian trên cơ sở kế thừa những dữ liệu đã có.

Bên cạnh đó, việc giải quyết bài toán chuyển đổi ngược này còn góp phần giải quyết bài toán chuyển đổi mô hình quan hệ sang các mô hình dữ liệu khác, như các mô hình XML và Ontology có yếu tố thời gian. Lúc đó phương pháp chuyển đổi từ mô hình quan hệ có yếu tố thời gian sang các mô hình này có thể sử dụng mô hình TimeER như là một kết quả chuyển đổi trung gian [12][13][14].

Quang H. và cộng sự đã xây dựng một thuật toán nhằm trích xuất mô hình TimeER từ một mô hình quan hệ cho trước có yếu tố thời gian [7][10]. Cách tiếp cận của phương pháp trích xuất được đề xuất trong nghiên cứu này là dựa vào đặc điểm của tập các thuộc tính, khoá chính, và tập các khoá ngoài của mỗi lược đồ quan hệ trong mô hình quan hệ theo thời gian. Từ đó xây dựng các quy tắc chuyển đổi liên quan đến quá trình nhận dạng các thành phần tồn tại bên trong mô hình TimeER tương ứng.

Tuy nhiên, nghiên cứu ở trên là chưa đề cập đến việc nhận dạng mối quan hệ phản xạ có yếu tố thời gian của mô hình TimeER. Trên cơ sở của phương pháp chuyển đổi thuận đối với mối quan hệ phản xạ có yếu tố thời gian của mô hình TimeER [11], bài báo đề xuất các quy tắc nhằm trích xuất mối quan hệ phản xạ có yếu tố thời gian của mô hình TimeER từ một mô hình quan hệ cho trước. Kết quả đề xuất như là một bổ sung cho bộ quy tắc trích xuất cho bài toán chuyển đổi ngược này.

Bài báo này được tổ chức như sau. Mục tiếp theo giới thiệu tổng quan phương pháp chuyển đổi ngược đã được nghiên cứu trước đây. Mục 3 trình bày các quy tắc nhằm nhận dạng mối quan hệ phản xạ có yếu tố thời gian của mô hình TimeER từ một mô hình quan hệ có yếu tố thời gian cho trước. Mục cuối cùng là phần kết luận.

2. GIỚI THIỆU PHƯƠNG PHÁP TRÍCH XUẤT MÔ HÌNH TIME-ER TỪ MÔ HÌNH QUAN HỆ

Bài toán trích xuất một mô hình TimeER từ một mô hình quan hệ cho trước được xác định như sau.

- Vào: Mô hình quan hệ theo thời gian (tập các quan hệ DB), trong đó mỗi quan hệ R của DB ta xác định được: tập các thuộc tính U_R , khóa chính PK_R , và tập các khóa ngoài FK_R .
- Ra: Mô hình TimeER. Mô hình TimeER cần trích xuất được gọi là thỏa mãn yêu cầu của thuật toán “chuyển đổi ngược”, nếu từ mô hình này, bằng cách sử dụng thuật toán “chuyển đổi thuận” ta thu được tập các quan hệ DB [9].

Phương pháp trích xuất được thực hiện thông qua việc sử dụng lần lượt các quy tắc nhằm nhận dạng những thành phần trong mô hình TimeER dựa vào đặc điểm của tập các thuộc tính, khoá chính, và tập các khoá ngoài của mỗi quan hệ trong DB. Các thành phần đó bao gồm: các kiểu thực thể, yếu tố thời gian của kiểu thực thể, thuộc tính thời gian của một kiểu thực thể, mối quan hệ giữa các kiểu thực thể, yếu tố thời gian của một mối quan hệ, thuộc tính thời gian và phi thời gian của một mối quan hệ.

Việc đề ra một quy tắc nhằm nhận dạng một trong những thành phần này được thực hiện theo một nguyên tắc chung như sau. Xây dựng một điều kiện nhằm nhận dạng thành phần nào đó sao cho nó thỏa mãn thuật toán “chuyển đổi thuận” đối với thành phần đó, nhưng không thỏa mãn đối với các thành phần còn lại. Điều này sẽ cho phép ta chứng minh tính đúng đắn của các quy tắc này theo phương pháp loại trừ.

Để tiện cho việc xây dựng các quy tắc nhận dạng, trước tiên ta thực hiện việc phân nhóm các quan hệ trong DB như sau. Gọi DB_1 là tập tất cả các quan hệ R thuộc DB mà R có không chứa các thuộc tính nhãn thời gian. Ngược lại, gọi DB_2 là tập tất cả các quan hệ R thuộc DB mà R có chứa các thuộc tính nhãn thời gian. Rõ ràng: $DB_1 \cup DB_2 = DB$.

Tiếp đến, ta chuyển đổi mỗi quan hệ R của DB thành một kiểu thực thể tạm thời tương ứng có cùng tên và cùng tập thuộc tính U_R , ký hiệu là $E(R)$.

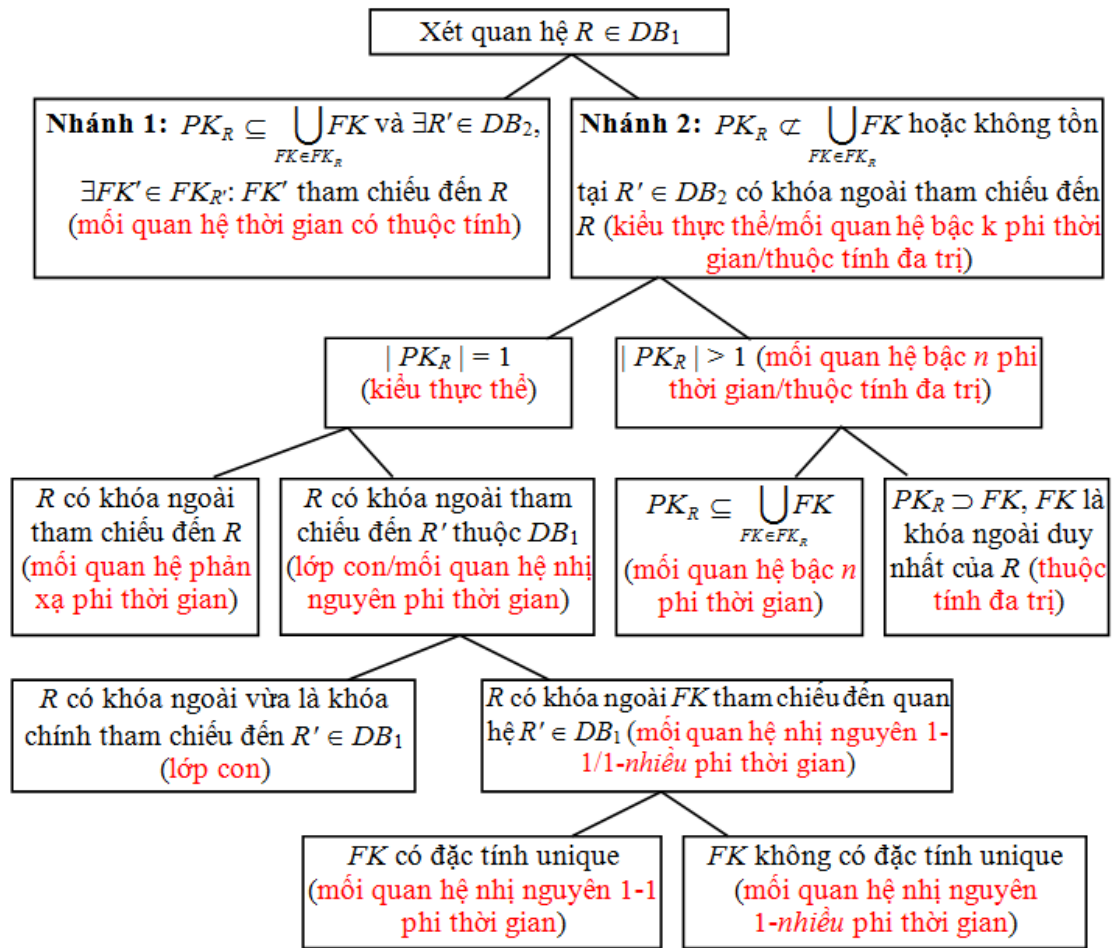
Giải thích: Theo thuật toán “chuyển đổi thuận”, mỗi kiểu thực thể trong mô hình TimeER đều tương ứng với một quan hệ R chứa tất cả các thuộc tính đơn trị phi thời gian của kiểu thực thể đó. Vì vậy, trong thuật toán “chuyển đổi ngược”, mỗi quan hệ R được ánh xạ đến một kiểu thực thể tạm thời. Tuy nhiên, ta sử dụng thuật ngữ “kiểu thực thể tạm thời” là do trong các quan hệ trên DB có những quan hệ mà sau này sẽ được nhận dạng như là các thành phần khác của mô hình TimeER (chẳng hạn: mối quan hệ, thuộc tính đa trị, ...).

Khi đó, việc trích xuất mô hình TimeER từ DB_1 và DB_2 được thực hiện lần lượt dựa vào các quy tắc trích xuất đã được nêu trong [10] nhằm nhận dạng các thành phần sau:

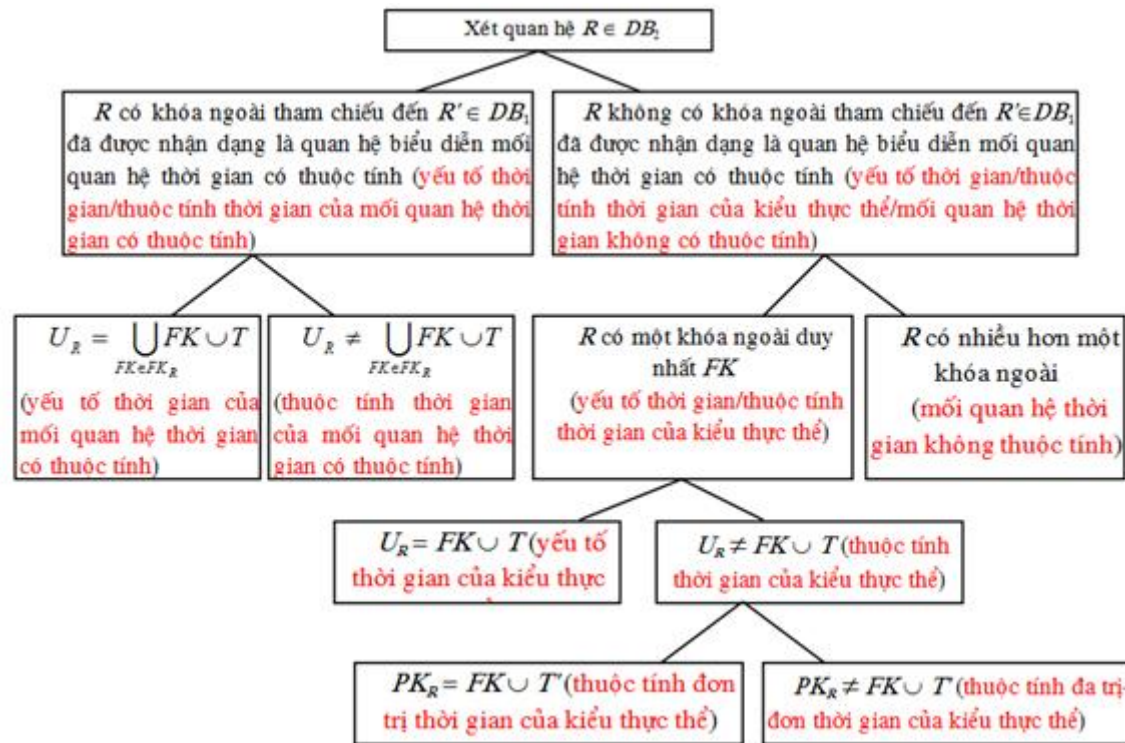
- Mối quan hệ thời gian có thuộc tính
- Yếu tố thời gian của mối quan hệ thời gian có thuộc tính
- Thuộc tính thời gian của mối quan hệ thời gian có thuộc tính
- Yếu tố thời gian của một kiểu thực thể
- Thuộc tính đơn trị có yếu tố thời gian của một kiểu thực thể
- Thuộc tính đa trị - đơn có yếu tố thời gian của kiểu thực thể

- Mỗi quan hệ thời gian không có thuộc tính
- Mỗi quan hệ lớp cha/lớp con
- Thuộc tính đa trị
- Mỗi quan hệ phản xạ phi thời gian
- Mỗi quan hệ bậc n phi thời gian
- Mỗi quan hệ nhị nguyên 1-1/1-nhiều phi thời gian

Tóm lại, các quy tắc để nhận dạng các thành phần của một mô hình TimeER từ một mô hình quan hệ thời gian cho trước có thể biểu diễn bởi hai sơ đồ phân tích các khả năng nhận dạng như sau.



Hình 1. Sơ đồ phân tích các khả năng nhận dạng đối với các quan hệ không có các thuộc tính nhân thời gian



Hình 2. Sơ đồ phân tích các khả năng nhận dạng
đối với các quan hệ có các thuộc tính nhãn thời gian

Trong cây phân tích ở Hình 1, ta thấy trong một phân nhánh con của Nhánh 2 có đề cập đến tiêu chuẩn “R có khóa ngoài tham chiếu đến chính R” thì mỗi quan hệ phản xạ phi thời gian sẽ được trích xuất. Tuy nhiên, điều này chỉ trích xuất được mỗi quan hệ phản xạ phi thời gian 1-1 hoặc 1-nhiều. Về nguyên tắc, chúng ta còn phải xét đến khả năng để trích xuất cho đối với mỗi quan hệ phản xạ phi thời gian nhiều-nhiều và các trường hợp khác của mỗi quan hệ phản xạ có yếu tố thời gian. Trên cơ sở lý thuyết “chuyển đổi thuận” của mỗi quan hệ phản xạ phi thời gian và có yếu tố thời gian [11], các quy tắc trích xuất mỗi quan hệ này trên mô hình TimeER được đề cập trong mục tiếp theo.

3. CÁC QUY TẮC TRÍCH XUẤT MỐI QUAN HỆ PHẢN XẠ CỦA MÔ HÌNH TIME-ER

Việc trích xuất các mối quan hệ phản xạ có trong mô hình TimeER từ một mô hình cơ sở dữ liệu quan hệ có thể giải quyết dựa vào ba quy tắc sau.

3.1. Quy tắc trích xuất mối quan hệ phản xạ phi thời gian 1-1/1-nhiều

Quy tắc 1: Xét quan hệ R trên DB_1 .

Nếu: Tồn tại khóa ngoài FK của quan hệ R tham chiếu đến khóa chính của chính quan hệ R ;

Thì: Khóa ngoài FK được thiết kế để biểu diễn mối quan hệ r là mối quan hệ phản xạ phi thời gian 1-1/1-nhiều trên tập thực thể $E(R)$. Cụ thể, nếu FK có ràng buộc là UNIQUE (trong SQL, ràng buộc này nhằm ngăn cản việc hai bản ghi có giá trị giống nhau trên cột dữ liệu đó), thì r là mối quan hệ phản xạ 1-1; ngược lại, r là mối quan hệ phản xạ 1- nhiều.

Giải thích: Sở dĩ trong trường hợp này, ta trích xuất được mối quan hệ phản xạ phi thời gian 1-1/ 1-nhiều, là vì trong tất cả các quy tắc chuyển đổi của phương pháp chuyển đổi thuận thì trường hợp khóa ngoài tự tham chiếu lên khóa chính trên cùng một quan hệ là chỉ xuất hiện trong quy tắc chuyển đổi mối quan hệ phản xạ phi thời gian 1-1/1-nhiều.

3.2. Quy tắc trích xuất mối quan hệ phản xạ có yếu tố thời gian 1-1/1-nhiều

Quy tắc 2: Xét quan hệ R trên DB_1 .

Nếu: Tồn tại quan hệ R' thuộc DB_2 có hai khóa ngoài cùng tham chiếu đến khóa chính của quan hệ R và chỉ một trong hai khóa ngoài này là tham gia vào khóa chính của R' ;

Thì: Quan hệ R' được nhận dạng là để biểu diễn mối quan hệ r là mối quan hệ phản xạ có yếu tố thời gian 1-1/1-nhiều trên tập thực thể $E(R)$.

Giải thích: Trong quy tắc này, ta trích xuất được mối quan hệ phản xạ có yếu tố thời gian 1-1/ 1-nhiều, là vì trong tất cả các quy tắc chuyển đổi của phương pháp chuyển đổi thuận thì trường hợp này là chỉ xuất hiện trong hai quy tắc chuyển đổi mối quan hệ phản xạ có yếu tố thời gian 1-1 và 1-nhiều [11].

3.3. Quy tắc trích xuất mối quan hệ phản xạ nhiều-nhiều

Quy tắc 3: Xét quan hệ R trên DB_1 .

Nếu: Tồn tại quan hệ R' thuộc DB_1 có hai khóa ngoài cùng tham chiếu lên khóa chính của quan hệ R ;

Thì: Quan hệ R' được nhận dạng là để biểu diễn mối quan hệ r là mối quan hệ phản xạ nhiều-nhiều trên tập thực thể $E(R)$. Cụ thể, nếu tiếp tục tồn tại quan hệ R'' thuộc DB_2 có khóa ngoài tham chiếu đến khóa chính của quan hệ R' , thì r là mối quan hệ phản xạ có yếu tố thời gian; ngược lại, r là mối quan hệ phản xạ phi thời gian.

Giải thích: Khác với đầu vào của Quy tắc 2, đó là quan hệ R' được xét trên DB_2 , còn trong quy tắc này, quan hệ R' là được xét trên DB_1 , nên ta chỉ cần thêm điều kiện hai

khóa ngoài của R' cùng tham chiếu lên khóa chính của quan hệ R là đủ để có thể trích xuất được mỗi quan hệ phản xạ nhiều-nhiều. Ngoài ra, theo quy tắc chuyển đổi thuận mỗi quan hệ phản xạ có yếu tố thời gian nhiều-nhiều [11], thì điều kiện tồn tại quan hệ thời gian có khóa ngoài tham chiếu đến khóa chính của quan hệ R' là điều kiện cần thiết để trích xuất mỗi quan hệ r là mỗi quan hệ có yếu tố thời gian.

4. KẾT LUẬN

Trong bài báo này chúng tôi đã đề xuất các quy tắc nhằm trích xuất các mối quan hệ phản xạ của mô hình TimeER từ một mô hình quan hệ cho trước dựa vào đặc điểm của khóa chính và các khóa ngoài của mỗi lược đồ quan hệ có trong mô hình này. Cách tiếp cận này là có tính thực tiễn đối với các CSDL quan hệ có yếu tố thời gian đang tồn tại, do nó chỉ dựa vào phần siêu dữ liệu được xác định bởi ngôn ngữ định nghĩa dữ liệu trên mô hình quan hệ (các câu lệnh CREATE TABLE, ALTER TABLE trong SQL).

Các quy tắc chuyển đổi này đã được cài đặt thành công với đầu vào là các file script được sử dụng trên hệ quản trị CSDL Microsoft SQL Server và đầu ra là sử dụng các tài liệu XML để biểu diễn mô hình TimeER.

Vấn đề hoàn thiện phương pháp trích xuất mô hình TimeER từ mô hình quan hệ có thể ứng dụng trong việc chuyển đổi mô hình quan hệ thời gian sang các mô hình dữ liệu khác (như mô hình Temporal XML hay Ontology có yếu tố thời gian), bằng cách sử dụng mô hình TimeER như là một kết quả chuyển đổi trung gian.

TÀI LIỆU THAM KHẢO

- [1]. H. Quang, H.T. Thanh (2009). Extension of Method for Converting TimeER Model to Relational Model, *Journal of Computer Science and Cybernetics*, vol. 25, no. 3, pp. 246-257.
- [2]. H. Quang, H.T. Thanh (2008). A Mapping Algorithm from TimeER Model to Relational Model, *Proceedings of the Second Hanoi Forum on Information – Communication Technology*, Hanoi, December 11-13, 2008, pp. 37-45.
- [3]. Elmasri R., Navathe S. B. (2016). “Chapter 26: Enhanced Data Models: Introduction to Active, Temporal, Spatial, Multimedia, and Deductive Databases”, *Fundamentals of Database Systems*, 7th Edition, Addison Wesley, tr. 961-1020.
- [4]. Jensen C. S., Snodgrass R. T. (1999). Temporal Data Management, *IEEE Transactions on Knowledge and Data Engineering*, vol. 11, no. 1, pp. 36-44.
- [5]. C. S. Jensen, Temporal Database Management, Dr.techn. thesis, Aalborg University, 2000 (<http://www.cs.auc.dk/~csj/Thesis/>).

- [6]. Gregersen H. and Jensen C. S. (1999). Temporal EntityRelationship Models – a Survey, *IEEE Transactions on Knowledge and Data Engineering*, vol. 11, no. 3, pp. 464–497.
- [7]. Quang Hoang, Toan Van Nguyen (2011), Extraction of TimeER Model from a Relational Database, *Proceedings of ACIIDS 2011, Lecture Notes in Artificial Intelligence* 6591 (Springer), pp. 57-66.
- [8]. Torp K., Snodgrass R. T., and Jensen C. S. (2000). Effective Timestamping in Databases, *VLDB Journal*, vol. 8, no. 40, pp. 1263–1313.
- [9]. R.H.L. Chiang, T.M. Barron, V.C. Storey (1994), Reverse Engineering of Relational Databases: Extraction of an EER Model from a Relational Database, *Data & Knowledge Engineering*, vol. 12, no. 2, pp. 107-142.
- [10]. Quang Hoang, Toan Van Nguyen (2013). Extraction of a Temporal Conceptual Model from a Relational Database, *International Journal of Intelligent Information and Database Systems*, vol. 7, no. 4, pp. 340-355.
- [11]. Hoàng Quang (2021). Chuyển đổi mối quan hệ phản xạ có yếu tố thời gian của mô hình TimeER sang mô hình quan hệ, *Tạp chí Khoa học và Công nghệ, Trường Đại học Khoa học – Đại học Huế*, tập 19, số 1, trang 47-55.
- [12]. Quang Hoang, Van Tinh Nguyen, Hoang Lien Minh Vo, Thi Nhu Thuy Truong (2016). A Method for Transforming TimeER Model-based Specification into Temporal XML, *Proceedings of ICCSAMA 2016*, Laxenburg (Austria), Advanced Computational Methods for Knowledge Engineering 453 (Springer), pp. 59-73.
- [13]. Toan Van Nguyen, Hoang Lien Minh Vo, Quang Hoang, Hanh Huu Hoang (2016). A New Method for Transforming TimeER Model-Based Specification into OWL, *Proceedings of ACIIDS 2016*, Đà Nẵng, Recent Developments in Intelligent Information and Database Systems 642 (Springer), pp. 111-121.
- [14]. Minh Hoang Lien Vo, Quang Hoang (2018). Transforming Extended Entity-Relationship model into OWL in Temporal databases, *Journal of Computer Science and Cybernetics*, vol. 34, no. 1, pp. 77-96.

EXTRACTION OF A TEMPORAL RECURSIVE RELATIONSHIPS OF THE TIME-ER MODEL FROM A RELATIONAL DATABASE

Hoang Quang

Faculty of Information Technology, University of Sciences, Hue University

Email: hquang@husc.edu.vn

ABSTRACT

The problem of extracting a temporal conceptual model from a given relational model is considered as a inverse mapping of the traditional conversion problem in the temporal database design process. In fact, solving this problem will facilitate an upgraded temporal information system. This study aims to supplement rules for identifying temporal recursive relationships of the TimeER model from a relational model.

Keywords: conceptual database model, TimeER model, temporal recursive relationship, relational database model.



Hoàng Quang sinh ngày 01/01/1962 tại Thừa Thiên Huế. Năm 1983, ông tốt nghiệp cử nhân chuyên ngành Toán ứng dụng tại Trường Đại học Tổng hợp Huế. Năm 2004, ông nhận bằng Tiến sĩ chuyên ngành Đảm bảo toán học cho các hệ thống tính toán tại Viện Công nghệ thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam. Hiện nay, ông công tác tại Khoa Công nghệ thông tin, Trường Đại học Khoa học, Đại học Huế.

Lĩnh vực nghiên cứu: Cơ sở dữ liệu hướng đối tượng, Cơ sở dữ liệu thời gian, Cơ sở dữ liệu không gian, Cơ sở dữ liệu bán cấu trúc XML, Thiết kế Ontology.

