

MỘT HƯỚNG TIẾP CẬN ĐỂ BỔ SUNG DỮ LIỆU THỜI GIAN VÀO HỆ THỐNG THÔNG TIN ĐANG VẬN HÀNH

LƯƠNG THÁI NGỌC*

TÓM TẮT

Bổ sung thêm dữ liệu thời gian vào hệ thống thông tin (HTTT) đang vận hành là cần thiết nhằm nâng cao tính đúng của thông tin về mặt “lịch sử”. Tuy nhiên, sử dụng phương pháp truyền thống sẽ làm thay đổi cấu trúc vật lý của các bảng (Table), đây là việc làm vô cùng đáng ngại vì nó ảnh hưởng đến hệ thống phần mềm đang hoạt động. Bài báo này trình bày giải pháp bổ sung dữ liệu thời gian nhưng vẫn đảm bảo cấu trúc ban đầu của hệ thống dựa trên mô hình thực thể - quan hệ thời gian (TimeER). Quá trình thực nghiệm được tác giả tiến hành trên HTTT Quản lý bán hàng đang vận hành nhằm đánh giá kết quả nghiên cứu.

Từ khóa: Cơ sở dữ liệu, ER, thời gian, thông tin, thực thể, quan hệ.

ABSTRACT

An approach to improve information system with time database

Adding time data to the information system is necessary to enhance the accuracy of information in terms of “history”. However, using traditional methods may modify the physical structure of tables, which is worrisome as it affects the existing software system. This article presents the solution of adding time data keeping existing system structure based on TimeER. The experiments were conducted on the information system for sale management to assess research results.

Keywords: Database, ER, time, information, entity, relation.

1. Giới thiệu

Một hệ thống thông tin tốt phải đảm bảo trả lời đúng các câu truy vấn có liên quan đến thời gian, vì vậy việc quản lý dữ liệu thời gian là cần thiết. Vấn đề này thường bị xem nhẹ, thậm chí là không được quan tâm đúng mức khi thiết kế các HTTT trước đây, dẫn đến kết quả truy vấn không đúng tại một số thời điểm nhất định.

Giả thiết, chúng ta có HTTT Quản lý bán hàng (hình 1) đang hoạt động, ta cần thực hiện một số câu lệnh truy vấn thông tin (bảng 1) có liên quan đến lịch sử. Việc trả lời các câu lệnh truy vấn này là không thể thực hiện được với HTTT hiện tại, vì vậy cần thiết phải bổ sung thêm dữ liệu thời gian. Một giải pháp để giải quyết vấn đề là bổ sung thêm các thực thể thời gian mới bằng phương pháp truyền thống như nhóm tác giả [1] đã trình bày. Tuy nhiên, giải pháp này chỉ phù hợp khi thiết kế HTTT ngay từ đầu,

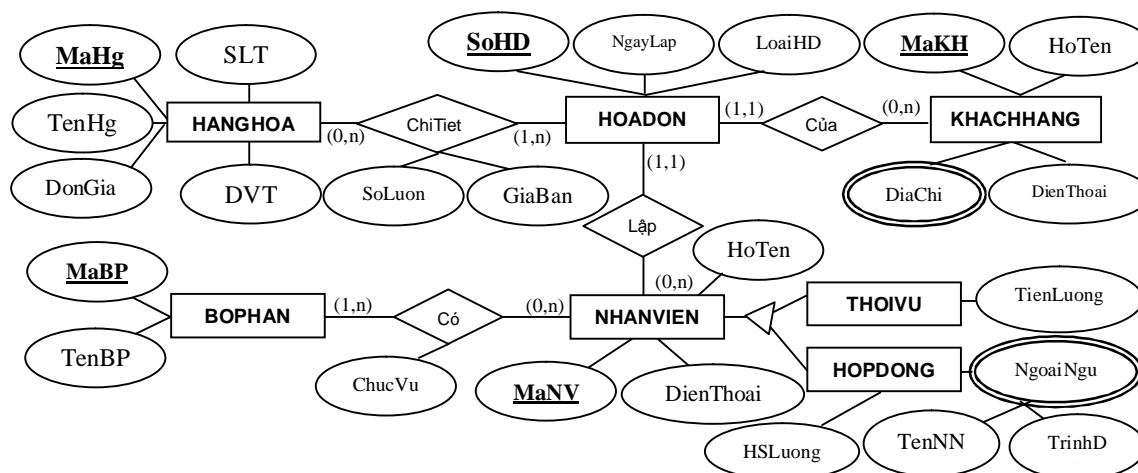
* ThS, Trường Đại học Đồng Tháp; Email: ltngoc@dtu.edu.vn

không phù hợp để áp dụng cho HTTT đang vận hành. Lí do là sử dụng giải pháp này sẽ làm thay đổi cấu trúc vật lí của các bảng (Table), điều này là một “thảm họa” cho hệ thống phần mềm đang hoạt động.

Bảng 1. Câu lệnh truy vấn thông tin thời gian

Câu	Nội dung truy vấn
a)	Cho biết đơn giá của mặt hàng H001 vào tháng 01/1014 là bao nhiêu?
b)	Cho biết lịch sử làm việc của nhân viên có mã số N005?
c)	Cho biết lịch sử số lượng tồn kho (SLT) của mặt hàng có mã số H002?
d)	Cho biết thời gian thực hiện giao tác thêm hóa đơn HD01?
e)	Cho biết lịch sử cập nhật số lượng và giá bán của các mặt hàng thuộc hóa đơn HD001?
f)	Cho biết thời gian thực hiện giao tác cập nhật địa chỉ khách hàng và lịch sử địa chỉ khách hàng K001?
g)	Cho biết lịch sử trình độ ngoại ngữ của nhân viên NV001?

Một hướng tiếp cận có thể xem là “tối ưu” cho việc bổ sung dữ liệu thời gian vào HTTT đang vận hành là mô hình thực thể, quan hệ thời gian TimeER [4, 5], giải pháp này sẽ không làm cấu trúc vật lí hiện có của HTTT đang vận hành giúp giảm thời gian và chi phí khi nâng cấp.



Hình 1. Mô hình ER của Hệ thống thông tin Quản lí bán hàng

Trong phần tiếp theo, bài báo sẽ mô tả về các loại thời gian cần quản lí trong mô hình TimeER. Mục 3 trình bày chi tiết các bước xây dựng mô hình TimeER để biểu diễn CSDL thời gian, mục 4 trình bày các bước chuyển đổi từ mô hình TimeER sang mô hình quan hệ, mục 5 trình bày thực nghiệm và cuối cùng là kết luận.

2. Các loại thời gian trong mô hình TimeER

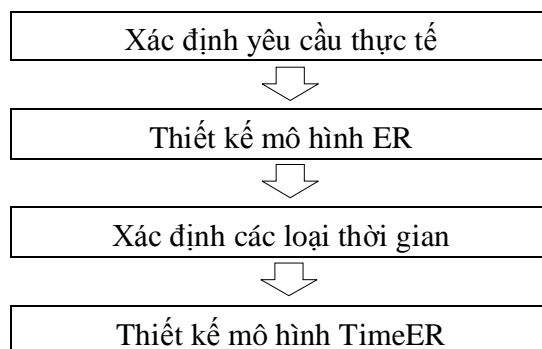
Mô hình TimeER cho phép xây dựng HTTT có yếu tố thời gian ở mức quan niệm. Mô hình này hỗ trợ 5 loại thời gian (bảng 2), trong đó thực thể hỗ trợ thời gian sống, hoặc thời gian giao tác, hoặc cả hai. Thuộc tính hỗ trợ thời gian hợp lệ, hoặc thời gian giao tác, hoặc cả hai. Mỗi quan hệ giữa các thực thể có thể xem là một thực thể hoặc một thuộc tính nên cũng hỗ trợ yếu tố thời gian. [6]

Bảng 2. Các loại thời gian trong mô hình TimeER

Loại thời gian	Đối tượng áp dụng			Kí hiệu
	Thực thể	Thuộc tính	Mối quan hệ	
Thời gian sống (LifeSpan - LS)	•		•	LS
Thời gian hợp lệ (ValidTime - VT)		•	•	VT
Thời gian giao tác (TransactionTime - TT)	•	•	•	TT
BT (BiTemporal) = VT + TT		•	•	BT
LT = LS + TT	•		•	LT

3. Xây dựng mô hình thực thể quan hệ thời gian (TimeER)

Mô hình thực thể quan hệ thời gian được xây dựng bằng cách bổ sung các kí hiệu biểu diễn thông tin thời gian vào mô hình ER truyền thống, các bước xây dựng như hình 2.



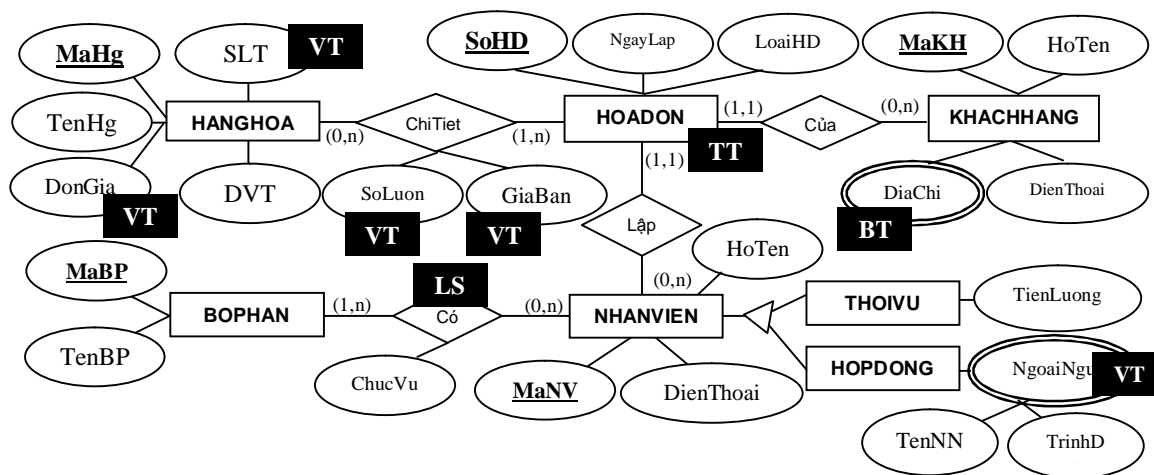
Hình 2. Các bước xây dựng mô hình TimeER

Như vậy, vấn đề quan trọng trong khi thiết kế mô hình TimeER là xác định đối tượng và loại thời gian cần quản lý (bảng 2). Ví dụ, để trả lời đúng kết quả câu truy vấn (Câu a, bảng 1) chúng ta cần quản lý chi tiết **đơn giá** các mặt hàng theo thời gian, đây là thuộc tính nên ta sử dụng thời gian hợp lệ (VT) để thiết kế. Tương tự, ta phân tích yêu cầu của các câu truy vấn khác để xác định đối tượng và loại thời gian tương ứng cần quản lý, kết quả được tổng hợp trong bảng 3.

Bảng 3. Các loại thời gian tương ứng câu lệnh truy vấn

Câu	Đối tượng áp dụng			Thời gian
	Thực thể	Thuộc tính	Mối quan hệ	
a)		DonGia		VT
b)			Có	LS
c)		SLT		VT
d)	HOADON			TT
e)		SoLuong, GiaBan		VT
f)		DiaChi		BT
g)		NgoaiNgu		VT

Sau khi xác định đối tượng và loại thời gian cần quản lí, ta tiến hành xây dựng mô hình TimeER bằng cách bổ sung các kí hiệu thời gian (bảng 3) cho đối tượng (thực thể, thuộc tính, mối quan hệ) tương ứng cần quản lí vào mô hình ER (hình 1), kết quả ta thu được mô hình TimeER như hình 3. Chi tiết quá trình chuyển đổi qua mô hình quan hệ sẽ được trình bày trong phần tiếp theo.



Hình 3. Mô hình TimeER quản lí dữ liệu thời gian

4. Chuyển đổi từ mô hình TimeER sang mô hình quan hệ

a) **Bước 1.** Loại bỏ các kí hiệu thời gian trên mô hình TimeER sẽ thu được mô hình ER truyền thống. Ta áp dụng các quy tắc chuyển đổi truyền thống trong mô hình ER để tạo ra được mô hình quan hệ. Kết quả là tạo ra mô hình quan hệ gốc (Original Relation Model - ORM) của HTTT đang vận hành. [2]

Ví dụ: Áp dụng bước này cho mô hình TimeER (hình 3) sẽ tạo ra mô hình quan hệ gốc (hình 5).

b) Bước 2. Sử dụng các quy tắc chuyển đổi mới trong mô hình TimeER, ta chuyển các thực thể, thuộc tính và mối quan hệ có yếu tố thời gian sang mô hình quan hệ. Kết quả của bước này sẽ tạo ra một tập các quan hệ thời gian (Time Relation Sets – TRS), tùy thuộc vào nhãn thời gian của đối tượng mà các quan hệ mới sinh ra có thêm các thuộc tính như hình 4.

LS	TT	VT	LT	BT
#LS _S	#TT _S	#VT _S	#LS _S	#VT _S
LS _E	TT _E	VT _E	LS _E	VT _E
			#TT _S	#TT _S
			TT _E	TT _E

Hình 4. Thuộc tính thời gian tương ứng các nhãn

Quy tắc 1. Đối với các thực thể có kí hiệu thời gian

Thực thể E có hỗ trợ yếu tố thời gian sẽ tạo ra một quan hệ thời gian TR(E) có các thuộc tính là $ID(E) \cup T$, khóa chính của TR(E) là $ID(E) \cup T'$. Trong đó T là tập các thuộc tính nhãn thời gian tương ứng trong hình 4, $T' \subset T$ là các thuộc tính có kí hiệu # phía trước, ID(E) là khóa chính của thực thể E. [2]

Ví dụ: Thực thể HOADON có thời gian giao tác TT, sẽ tạo ra một quan hệ thời gian TR(HOADON) với 3 thuộc tính: #SoHD, #TT_S, TT_E.

Quy tắc 2. Thuộc tính có kí hiệu thời gian

+ **Đơn trị:** Thuộc tính đơn trị A của thực thể E có hỗ trợ yếu tố thời gian sẽ tạo ra một quan hệ thời gian TR_A(E) có các thuộc tính là $ID(E) \cup A \cup T$, khóa chính của TR_A(E) là $ID(E) \cup T'$. Trong đó T là tập các thuộc tính nhãn thời gian tương ứng trong hình 4, $T' \subset T$ là các thuộc tính có kí hiệu # phía trước, ID(E) là khóa chính của thực thể E. [2]

Ví dụ: Thuộc tính đơn trị DonGia của thực thể HANGHOA có yếu tố thời gian hợp lệ VT, sẽ tạo ra một quan hệ thời gian TR_{DonGia}(HANGHOA) với 4 thuộc tính: #SoHD, #VT_S, VT_E, DonGia.

+ **Đa trị:** Thuộc tính đa trị A của thực thể E có hỗ trợ yếu tố thời gian sẽ tạo ra một quan hệ thời gian TR_A(E) có các thuộc tính là $ID(E) \cup A \cup T$, khóa chính của TR_A(E) là $ID(E) \cup A \cup T'$. Trong đó T là tập các thuộc tính nhãn thời gian tương ứng trong hình 4, $T' \subset T$ là các thuộc tính có kí hiệu # phía trước, ID(E) là khóa chính của thực thể E. [2]

Ví dụ: Thuộc tính đa trị ĐịaChi của thực thể KHACHHANG có yếu tố thời gian hợp lệ VT và thời gian giao tác TT, sẽ tạo ra một quan hệ thời gian $TR_{\text{ĐịaChi}}(\text{KHACHHANG})$ với 6 thuộc tính: #MaKH, #VT_S, VT_E, #TT_S, TT_E, #ĐịaChi.

+ **Đơn trị phức hợp:** Thuộc tính đơn trị phức hợp A của thực thể E có hỗ trợ yếu tố thời gian, A có các thuộc tính con đơn trị là A_1, A_2, \dots, A_n . Để chuyển sang quan hệ thời gian cho thuộc tính A ta áp dụng quy tắc 2 để xét yếu tố thời gian cho các thuộc tính đơn trị A_1, A_2, \dots, A_n . [2]

+ **Đa trị phức hợp:** Thuộc tính đa trị phức hợp A của thực thể E có hỗ trợ yếu tố thời gian, A có các thuộc tính con đơn trị là A_1, A_2, \dots, A_n . Để chuyển sang quan hệ thời gian cho thuộc tính A ta xét yếu tố thời gian cho các thuộc tính A_1, A_2, \dots, A_n . Kết quả tạo ra các quan hệ thời gian $TR_{A_i}(A)$ có thuộc tính là $ID(R(A)) \cup A_i \cup T$, khóa chính của $TR_{A_i}(A)$ là $ID(R(A)) \cup T'$. Trong đó, T là tập các thuộc tính nhãn thời gian tương ứng trong hình 4, $T' \subset T$ là các thuộc tính có kí hiệu # phía trước, $ID(R(A))$ là khóa chính của quan hệ được tạo ra từ thuộc tính đa trị phức hợp A tại bước 1. [2]

Ví dụ: Thuộc tính đa trị phức hợp NgoạiNgu của thực thể HOPDONG có yếu tố thời gian hợp lệ VT. Chúng ta tiến hành xét yếu tố thời gian VT cho 2 thuộc tính con là TenNN và TrinhDo. Kết quả tạo ra 2 quan hệ là $TR_{\text{TenNN}}(\text{NN})$ và $TR_{\text{TrinhDo}}(\text{NN})$. Trong đó quan hệ $TR_{\text{TenNN}}(\text{NN})$ có thuộc tính là: #ID_NN, #VTS, VTE, TenNN và quan hệ $TR_{\text{TrinhDo}}(\text{NN})$ có thuộc tính là: #ID_NN, #VTS, VTE, TrinhDo.

Quy tắc 3. Môi quan hệ có kí hiệu thời gian

Môi quan hệ nhị nguyên S của 2 thực thể E_1 và E_2 có hỗ trợ yếu tố thời gian sẽ tạo ra một quan hệ thời gian $TR(S)$ có các thuộc tính là $ID(E_1) \cup ID(E_2) \cup T$, khóa chính của $TR(S)$ là $ID(S) \cup T'$. [2]

Trong đó, T là tập các thuộc tính nhãn thời gian tương ứng trong hình 4, $T' \subset T$ là các thuộc tính có kí hiệu # phía trước, $ID(E_1)$ và $ID(E_2)$ lần lượt là khóa chính của thực thể E_1 và E_2 , $ID(S)$ được xác định dựa vào môi quan hệ giữa 2 thực thể E_1 và E_2 như sau:

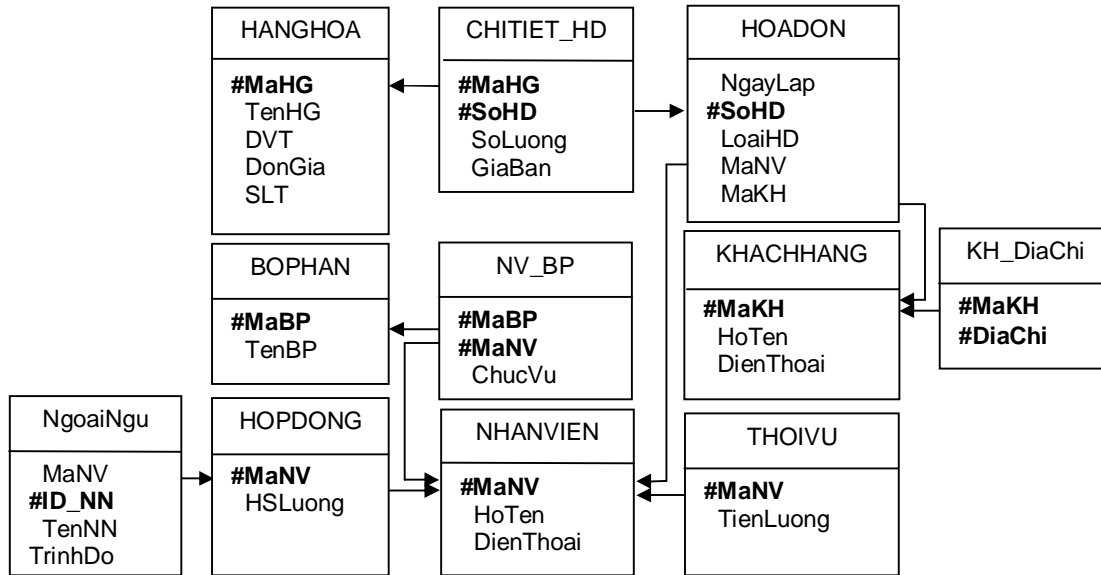
- + 1 – 1 thì $ID(S) = ID(E_1)$ hoặc $ID(S) = ID(E_2)$
- + 1 – n thì $ID(S) = ID(E_2)$, n – 1 thì $ID(S) = ID(E_1)$
- + n – n thì $ID(S) = ID(E_1) \cup ID(E_2)$

Ví dụ: Môi quan hệ “Có” giữa 2 thực thể NHANVIEN và BOPHAN có yếu tố thời gian LS, sẽ tạo ra một quan hệ $TR(\text{Có})$ có 4 thuộc tính là: #MaBP, #MaNV, #LS_S, LS_E.

5. Thực nghiệm

Từ mô hình ER của hệ thống thông tin Quản lí bán hàng (hình 1) ta có mô hình quan hệ của hệ thống đang vận hành (hình 5). Tiếp tục chuyển đổi mô hình thực thể quan hệ thời gian (hình 3), ta có mô hình quan hệ mới (hình 6) sau khi nâng cấp. Mô

hình này thực chất là mô hình ORM bổ sung thêm TRS, trong đó mỗi quan hệ giữa TR và R là mối quan hệ n-1 ($TR \in TRS$ và $R \in ORM$). Vì vậy, việc bổ sung các quan hệ thời gian mới vẫn đảm bảo trúc vật lý ban đầu của HTTT đang vận hành.

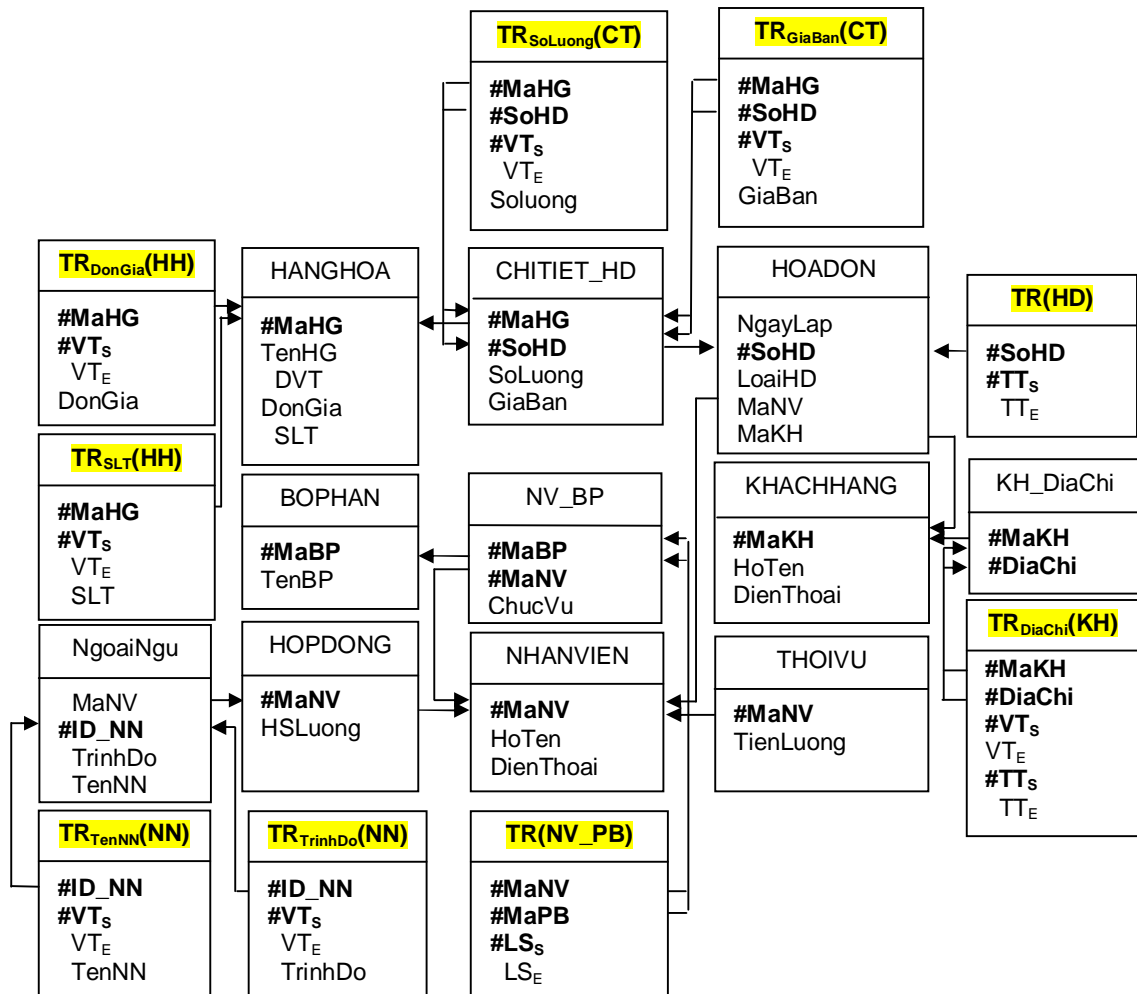


Hình 5. Mô hình quan hệ tương ứng của hệ thống đang vận hành

Sử dụng mô hình quan hệ thời gian (hình 6) sau nâng cấp, chúng ta có thể trả lời các câu lệnh truy vấn (bảng 1). Câu lệnh truy vấn được viết bằng đại số quan hệ tương ứng được trình bày trong bảng 4.

Bảng 4. Câu lệnh truy vấn được viết bằng đại số quan hệ

Câu	Biểu thức đại số quan hệ
a)	$\Pi_{DonGia}(\sigma_{VTs \leq 30/01/2014 \wedge VTe \geq 31/12/2013}(TR_{DonGia}(HH)))$
b)	$\Pi_{MaNV, MaPB, LSs, LSe}(\sigma_{MaNV = 'N005'}(TR(NV_PB)))$
c)	$\Pi_{MaHG, VTs, VTe, SLT}(\sigma_{MaHG = 'H002'}(TR_{SLT}(HH)))$
d)	$\Pi_{SoHD, TTs, TTe, Tte - TTs}(\sigma_{MaHD = 'HD01'}(TR(HD)))$
e)	$R_1 = \Pi_{MaHG, VTs, VTe, SoLuong}(\sigma_{MaHD = 'HD01'}(TR_{SoLuong}(CT)))$ $R_2 = \Pi_{MaHG, VTs, VTe, GiaBan}(\sigma_{MaHD = 'HD01'}(TR_{SoLuong}(CT)))$
f)	$\Pi_{MaKH, TTs, TTe, TTe - TTs, VTs, VTe}(\sigma_{MaKH = 'K001'}(TR_{DiaChi}(KH)))$
g)	$R_1 = \Pi_{MaNV, VTs, VTe, TeNN}(\sigma_{MaNV = 'NV001'}(TR_{TenNN}(NN) \bowtie NgoaiNgu \bowtie HOPDONG))$ $R_2 = \Pi_{MaNV, VTs, VTe, TrinhDo}(\sigma_{MaNV = 'NV001'}(TR_{TrinhDo}(NN) \bowtie NgoaiNgu \bowtie HOPDONG))$



Hình 6. Mô hình quan hệ sau bổ sung thông tin thời gian

Mô hình quan hệ thời gian (hình 6) sau nâng cấp có thể quản lý thông tin thời gian của đối tượng thông qua các thực thể thời gian trong TRS, dữ liệu trong thực thể thời gian TR ($TR \in TRS$) phải được cập nhật đồng thời khi cập nhật thông tin trên đối tượng cần quản lý thời gian tương ứng. Ví dụ: Khi cập nhật dữ liệu trên quan hệ NV_BP, ta đồng thời phải cập nhật dữ liệu thời gian vào quan hệ TR(NV_BP) như bảng 5 trong đó NOW là hằng thời gian.

Bảng 5. Dữ liệu thời gian của quan hệ TR(NV_BP)

NV	BP	LS _S	LS _E
NV001	P01	1/4/2014	5/10/2014
NV002	P02	7/5/2014	NOW
NV003	P02	15/3/2014	10/7/2014
NV003	P03	10/7/2014	7/3/2015
NV001	P03	5/10/2014	NOW
NV003	P02	7/3/2015	NOW

Quan sát dữ liệu trên quan hệ thời gian TR(NV_BP) ta thấy rằng nhân viên NV003 đã 3 lần thay đổi bộ phận làm việc với thời gian tương ứng là (15/3/2014 đến 10/7/2014; 10/7/2014 đến 7/3/2015; 7/3/2015 đến NOW).

6. Kết luận

Như vậy, bài báo đã trình bày hướng tiếp cận mô hình TimeER để bổ sung dữ liệu thời gian vào HTTT đang vận hành. Qua thực nghiệm, tác giả nhận thấy rằng hướng tiếp cận này giúp bổ sung các thực thể mới để lưu trữ dữ liệu thời gian mà vẫn đảm bảo cấu trúc vật lý trước đó của hệ thống. Đây là một giải pháp tốt, có thể ứng dụng hiệu quả vào thực tế để nâng cấp các HTTT của cơ quan, đơn vị, trường học. Việc bổ sung dữ liệu thời gian vào HTTT hướng đối tượng, phi cấu trúc (XML) sẽ là hướng nghiên cứu tiếp theo của tác giả trong tương lai.

TÀI LIỆU THAM KHẢO

1. Nguyễn Hữu Duyệt, Lương Thái Ngọc (2015), “Đề xuất giải pháp biểu diễn cơ sở dữ liệu thời gian bằng mô hình thực thể quan hệ (ER) truyền thống”, *Tạp chí Khoa học Trường Đại học Sư phạm TPHCM*, 2(67), 109-102.
2. Hoàng Quang, Nguyễn Việt Chánh (2012), “Một cách tiếp cận trong việc thiết kế cơ sở dữ liệu thời gian”, *Tạp chí Khoa học Đại học Huế*, 5(74A), 99-107.
3. Elmasri R., Navathe S.B. (2007), *Fundamentals of Database Systems*, Addison Wesley 5th Edition.
4. Gregersen H., Jensen C.S. (1999), “Temporal Entity Relationship Models - A Survey”, *IEEE*, 3(11), 464-497.
5. Gregersen H., Jensen C.S (1998), *Conceptual Modeling of Timevarying Information*, Timecenter Technical Report.
6. Jensen C.S., Snodgrass R.T.(1999), “Temporal Data Management”, *IEEE Transactions on Knowledge and Data Engineering*, 1(11), 36-44.

(Ngày Tòa soạn nhận được bài: 19-4-2015; ngày phản biện đánh giá: 08-5-2015;

ngày chấp nhận đăng: 18-5-2015)