

**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**VIỆN HÀN LÂM KHOA HỌC**

**VÀ CÔNG NGHỆ VIỆT NAM**

**VIỆN CÔNG NGHỆ THÔNG TIN**

**LƯU THỊ BÍCH HƯƠNG**

**NGHIÊN CỨU VÀ PHÁT TRIỂN**  
**KỸ THUẬT THỦY VĂN CƠ SỞ DỮ LIỆU QUAN HỆ**

**LUẬN ÁN TIẾN SĨ TOÁN HỌC**

**HÀ NỘI – 2014**

**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**VIỆN HÀN LÂM KHOA HỌC**

**VÀ CÔNG NGHỆ VIỆT NAM**

**VIỆN CÔNG NGHỆ THÔNG TIN**

---

**LƯU THỊ BÍCH HƯƠNG**

**NGHIÊN CỨU VÀ PHÁT TRIỂN  
KỸ THUẬT THỦY VÂN CƠ SỞ DỮ LIỆU QUAN HỆ**

**Chuyên ngành: BẢO ĐẢM TOÁN HỌC CHO MÁY TÍNH  
VÀ HỆ THỐNG TÍNH TOÁN**

**Mã số: 62.46.35.01**

**LUẬN ÁN TIẾN SĨ TOÁN HỌC**

**NGƯỜI HƯỚNG DẪN KHOA HỌC:**

**PGS.TS Bùi Thế Hồng**

**HÀ NỘI – 2014**

## LỜI CẢM ƠN

Để hoàn thành luận án này, tôi đã nhận được sự giúp đỡ rất tận tình các Thầy, Cô giáo trong Viện Công nghệ thông tin - Viện Hàn Lâm Khoa học và Công nghệ Việt Nam và trường ĐHSP Hà Nội 2. Tôi xin gửi lời cảm ơn các Thầy, Cô giáo trong Viện Công nghệ thông tin và trường ĐHSP Hà Nội 2 đã tạo điều kiện học tập, nghiên cứu, giúp đỡ tôi rất nhiều trong quá trình làm luận án. Đặc biệt tôi xin cảm ơn PGS.TS. Bùi Thế Hồng đã tận tình hướng dẫn chỉ bảo cho tôi trong toàn bộ quá trình học tập, nghiên cứu đề tài và giúp tôi hoàn thành bản luận án này.

Hà Nội, ngày tháng năm 2014

Nghiên cứu sinh

**Lưu Thị Bích Hương**

## **LỜI CAM ĐOAN**

Tôi xin cam đoan đây là công trình nghiên cứu của tôi dưới sự hướng dẫn khoa học của PGS.TS. Bùi Thế Hồng. Các kết quả được viết chung với các đồng tác giả đã được sự chấp thuận của các tác giả trước khi đưa vào luận án.

Các số liệu, kết quả nêu trong luận án là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

Tác giả luận án

**Lưu Thị Bích Hương**

# MỤC LỤC

LỜI CẢM ƠN .....	ii
LỜI CAM ĐOAN .....	ii
MỤC LỤC .....	iii
Bảng các ký hiệu, chữ viết tắt .....	v
Danh sách bảng .....	vii
Danh sách hình vẽ .....	viii
MỞ ĐẦU .....	1
Chương 1. THỦY VÂN CƠ SỞ DỮ LIỆU QUAN HỆ .....	9
1.1. Các khái niệm về cơ sở dữ liệu .....	9
1.1.1. Cơ sở dữ liệu .....	9
1.1.2. Mô hình dữ liệu quan hệ .....	9
1.1.3. Thuộc tính, miền thuộc tính và kiểu thuộc tính .....	10
1.1.4. Quan hệ, lược đồ quan hệ .....	10
1.1.5. Khoá của quan hệ .....	11
1.2. Một số khái niệm về thủy vân cơ sở dữ liệu quan hệ .....	12
1.2.1. Thủy vân .....	12
1.2.2. Thủy vân cơ sở dữ liệu quan hệ .....	12
1.2.3. Khóa thủy vân .....	13
1.2.4. Lược đồ thủy vân .....	14
1.2.5. Sự cần thiết của các kỹ thuật thủy vân cơ sở dữ liệu quan hệ .....	15
1.3. Các yêu cầu của thủy vân trên cơ sở dữ liệu quan hệ .....	17
1.3.1. Khả năng có thể phát hiện .....	17
1.3.2. Tính bền vững và dễ vỡ .....	18
1.3.3. Khả năng cập nhật dữ liệu .....	18
1.3.4. Tính ẩn và hiện .....	18
1.3.5. Phát hiện mù .....	19
1.4. Ứng dụng của thủy vân cơ sở dữ liệu quan hệ .....	19
1.4.1. Bảo vệ bản quyền .....	19
1.4.2. Đảm bảo sự toàn vẹn .....	20
1.4.3. Giấu vân tay .....	21
1.5. Những tấn công trên thủy vân cơ sở dữ liệu quan hệ .....	21
1.5.1. Cập nhật thông thường .....	21
1.5.2. Tấn công có chủ đích .....	22
1.6. Các lược đồ thủy vân cơ sở dữ liệu quan hệ .....	23
1.6.1. Bảo vệ bản quyền cơ sở dữ liệu quan hệ .....	23
1.6.2. Đảm bảo sự toàn vẹn của cơ sở dữ liệu quan hệ .....	27
1.7. Kết luận chương 1 .....	30
Chương 2. PHÁT TRIỂN LƯỢC ĐỒ THỦY VÂN BẢO VỆ BẢN QUYỀN CHO CƠ SỞ DỮ LIỆU QUAN HỆ .....	31
2.1. Xây dựng lược đồ thủy vân dựa vào việc chèn thêm ảnh nhị phân ..	31
2.1.1. Xây dựng lược đồ thủy vân .....	33

2.1.2.	Đánh giá độ phức tạp.....	36
2.1.3.	Chứng minh tính đúng đắn.....	36
2.1.4.	Đánh giá thử nghiệm.....	38
2.1.5.	Kết luận.....	40
2.2.	Phát triển lược đồ thủy vân dựa vào bit ý nghĩa nhất (MSB).....	40
2.2.1.	Cải tiến lược đồ thủy vân.....	42
2.2.2.	Tính bền vững và chi phí về thời gian và bộ nhớ.....	46
2.2.3.	Đánh giá thử nghiệm.....	48
2.2.4.	Kết luận.....	50
2.3.	Kết luận chương 2.....	50
<b>Chương 3. XÂY DỰNG LƯỢC ĐỒ THỦY VÂN ĐẢM BẢO SỰ TOÀN VẬN CỦA CƠ SỞ DỮ LIỆU QUAN HỆ</b> .....		<b>51</b>
3.1.	Phân nhóm quan hệ.....	51
3.2.	Phát triển lược đồ thủy vân với thuộc tính phân loại.....	53
3.2.1.	Cải tiến lược đồ thủy vân.....	54
3.2.2.	Đánh giá độ phức tạp.....	58
3.2.3.	Chứng minh tính đúng đắn.....	58
3.2.4.	Cân đối giữa số bộ trong quan hệ và số nhóm.....	60
3.2.5.	Đánh giá thử nghiệm.....	63
3.2.6.	Kết luận.....	65
3.3.	Thủy vân với dữ liệu kiểu số.....	66
3.3.1.	Lược đồ thủy vân.....	66
3.3.2.	Khoanh vùng các giả mạo.....	69
3.3.3.	Khôi phục dữ liệu gốc.....	69
3.3.4.	Chứng minh tính đúng đắn của thuật toán khôi phục.....	71
3.3.5.	Kết luận.....	73
3.4.	Xây dựng lược đồ thủy vân với dữ liệu kiểu văn bản.....	73
3.4.1.	Một số định nghĩa.....	73
3.4.2.	Tư tưởng.....	74
3.4.3.	Xây dựng lược đồ thủy vân.....	75
3.4.4.	Phân tích tính đúng đắn.....	81
3.4.5.	Đề xuất lược đồ thủy vân để khoanh vùng giả mạo.....	84
3.4.6.	Đánh giá thử nghiệm.....	88
3.4.7.	Kết luận.....	88
3.5.	Kết luận chương 3.....	89
Kết luận và hướng phát triển.....		90
Danh mục các công trình của tác giả.....		91
Tài liệu tham khảo.....		92

## Bảng các ký hiệu, chữ viết tắt

<i>Ký hiệu</i>	<i>Ý nghĩa của ký hiệu</i>
$R$	Lược đồ quan hệ
$r$	Quan hệ thuộc lược đồ $R$
$\gamma$	Số thuộc tính của quan hệ
$\omega$	Số bộ của quan hệ
$g$	Số nhóm của quan hệ
$r_i$	Bộ thứ $i$ trong quan hệ $r$
$r_i.A_j$	Giá trị thuộc tính thứ $j$ của bộ thứ $i$
$K$	Khóa thủy văn
$G_k$	Nhóm thứ $k$
$q_k$	Số bộ trong nhóm $G_k$
$P$	Thuộc tính khóa chính của quan hệ
$A_w$	Thuộc tính kiểu văn bản có thể chứa nhiều từ
$H_i$	Thuộc tính kiểu văn bản có tác động cao thứ $i$
$L_i$	Thuộc tính kiểu văn bản có tác động thấp thứ $i$
$H(K \parallel r_i.A_1 \parallel r_i.A_2 \parallel \dots \parallel r_i.A_\gamma)$	Giá trị băm khóa $K$ cùng với các giá trị thuộc tính của bộ $r_i$
$r_w$	Quan hệ thủy văn được tạo ra trong quá trình thủy văn
$\eta$	Tham số tạo thủy văn
$\tau$	Tham số phát hiện thủy văn
$W^l_j$	Thủy văn được nhúng vào thuộc tính thứ $j$ của tất cả các bộ trong một nhóm (thủy văn thuộc tính/cột)
$W^2_i$	Thủy văn được nhúng vào tất cả các thuộc tính của bộ thứ $i$ trong một nhóm (thủy văn bộ/dòng)
$W^{*l}_j$	Thủy văn được trích từ thuộc tính thứ $j$ của tất cả các bộ trong một nhóm đã thủy văn
$W^{*2}_i$	Thủy văn được trích từ tất cả các thuộc tính của bộ thứ $i$ trong một nhóm đã thủy văn
$V^l_j$	Kết quả xác nhận thủy văn đối với $W^l_j$

$V_i^2$	Kết quả xác nhận thủy vân đối với $W_i^2$
$n$	Số thuộc tính kiểu văn bản có tác động thấp trong quan hệ
$m$	Số thuộc tính kiểu văn bản có tác động cao trong quan hệ
$e_i$	Giá trị thứ $i$ trên đường chéo chính của ma trận thủy vân
$W_j$	Ký tự thủy vân thứ $j$
$ATOC()$	Hàm chuyển mã Unicode thành ký tự
$Converter()$	Hàm chuyển từ dạng số sang dạng nhị phân
$Substring(x,p,q)$	Hàm lấy ra $q$ ký tự của $x$ từ vị trí thứ $p$
$t_H$	Chi phí sinh một số ngẫu nhiên của hàm băm
$t_{mod}$	Chi phí của phép mod
$t_{if}$	Chi phí của phép if
$t_{delA}$	Chi phí cho phép xóa một thuộc tính
$t_{bit}$	Chi phí cho việc gán/so sánh một bit
$t_{count}$	Chi phí gán/cập nhật một con đếm
$t_{sort}$	Chi phí cho việc đổi chỗ hai bộ
$m_{count}$	Số bit cần thiết để ghi một con đếm
$m_{tuple}$	Số bit để ghi một bản sao của một bộ
$m_{wkey}$	Số bit ghi khóa thủy vân
$m_{pkey}$	Số bit ghi giá trị khóa chính
LSB	Bit ít ý nghĩa nhất ( <b>L</b> east <b>S</b> ignificant <b>B</b> it)
MSB	Bit ý nghĩa nhất ( <b>M</b> ost <b>S</b> ignificant <b>B</b> it)
MAC	Mã chứng thực thông điệp ( <b>M</b> essage <b>A</b> uthentication <b>C</b> ode)
CA	Cơ quan đăng ký bản quyền ( <b>C</b> ertificate <b>A</b> uthority)
MD5	Thuật toán MD5 ( <b>M</b> essage <b>D</b> igest algorithm <b>5</b> )



## Danh sách bảng

<b><i>Bảng 1.1.</i></b> Biểu diễn quan hệ r.....	11
<b><i>Bảng 3.1.</i></b> Tỷ lệ phát hiện đối với các tần công trên một bộ giá trị .....	64
<b><i>Bảng 3.2.</i></b> Kết quả thử nghiệm .....	88

## Danh sách hình vẽ

<b>Hình 1.</b> Phân loại các kỹ thuật giấu tin.....	3
<b>Hình 2.</b> Thủy vân trên đồng dolla của Mỹ .....	4
<b>Hình 1.1.</b> Sơ đồ mô tả lược đồ thủy vân cơ sở dữ liệu quan hệ cơ bản.....	15
<b>Hình 2.1:</b> (a) Ảnh nhị phân và giá trị thập phân tương ứng. (b) Thuộc tính văn bản sau khi được thủy vân, trong đó các chỉ số là số thứ tự các dấu cách đơn và DS là dấu cách đúp.....	32
<b>Hình 2.2.</b> Ảnh nhị phân sử dụng để thủy vân. (a) ảnh IOIT 12x4 (b) ảnh Smiley 8x8.....	38
<b>Hình 2.3.</b> Kết quả tấn công thêm.....	39
<b>Hình 2.4.</b> Kết quả tấn công xóa .....	39
<b>Hình 2.5.</b> Kết quả tấn công thay đổi dữ liệu.....	40
<b>Hình 2.6.</b> Tấn công thêm bộ đối với $\tau$ .....	48
<b>Hình 2.7.</b> Tấn công sửa bộ đối với $\tau$ .....	49
<b>Hình 2.8.</b> Tấn công xóa bộ đối với $\tau$ .....	49
<b>Hình 3.1.</b> Tỷ lệ phát hiện đối với các tấn công thêm nhiều bộ.....	64
<b>Hình 3.2.</b> Tỷ lệ phát hiện đối với các tấn công xóa nhiều bộ.....	65
<b>Hình 3.3.</b> Tỷ lệ phát hiện đối với các tấn công sửa nhiều bộ .....	65