

**ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG**

Nguyễn Anh Dũng

**GỘP NHÓM CÁC GIÁ TRỊ LẬP CỦA CÁC THUỘC TÍNH
TRONG CÁC MẢNH CỦA CƠ SỞ DỮ LIỆU PHÂN TÁN**

Luận văn thạc sĩ khoa học máy tính

Thái Nguyên - 2015

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG

Nguyễn Anh Dũng

**GỘP NHÓM CÁC GIÁ TRỊ LẬP CỦA CÁC THUỘC TÍNH TRONG
CÁC MẢNH CỦA CƠ SỞ DỮ LIỆU PHÂN TÁN**

Chuyên ngành: Khoa học máy tính

Mã số: 60 48 01 01

LUẬN VĂN THẠC SỸ KHOA HỌC MÁY TÍNH

GV hướng dẫn: PGS. NCVC. TS Lê Huy Thập

Thái Nguyên - 2015

LỜI CAM ĐOAN

Tôi xin cam đoan đây là công trình nghiên cứu của tôi, được xây dựng từ những số liệu và kết quả thực tiễn khi tôi học tập tại trường Đại học Công nghệ thông tin và Truyền thông Thái Nguyên và ứng dụng dựa trên số liệu của UBND huyện Đan Phượng, chưa từng được công bố hay xuất bản trong bất cứ công trình nào khác.

Thái Nguyên, tháng 10 năm 2015

Tác giả

Nguyễn Anh Dũng

LỜI CẢM ƠN

Trước tiên tôi xin được cảm ơn thầy PGS. NCVC. TS. Lê Huy Thập, người trực tiếp giảng dạy và là người hướng dẫn tôi thực hiện luận văn này. Tôi cũng xin bày tỏ lòng biết ơn đến các thầy, cô giáo Viện Công nghệ thông tin và Khoa Công nghệ thông tin - Đại học Thái Nguyên đã tận tình giảng dạy và tạo điều kiện thuận lợi cho tôi trong suốt quá trình học tập .

Tôi xin cảm ơn lãnh đạo Phòng Văn hóa và Thông tin huyện Đan Phượng, các anh chị đồng nghiệp đã tạo điều kiện cho tôi tham gia và hoàn thành khóa học. Tôi cũng xin cảm ơn các bạn của tôi, những người luôn bên cạnh động viên, giúp đỡ và đóng góp nhiều ý kiến thiết thực trong quá trình học tập và thực hiện luận văn.

Cuối cùng tôi chân thành cảm ơn các thành viên trong gia đình tôi, đã luôn hết mình yêu thương, dìu dắt tôi trong công việc và cuộc sống.

Thái Nguyên, tháng 10 năm 2015

Sinh viên thực hiện

Nguyễn Anh Dũng

MỤC LỤC

LỜI CAM ĐOAN	i
LỜI CẢM ƠN	ii
MỤC LỤC.....	iii
DANH MỤC BẢNG BIỂU	v
DANH MỤC HÌNH	vii
DANH MỤC CÁC CHỮ VIẾT TẮT	viii
BẢNG DANH MỤC CÁC KÍ HIỆU.....	ix
LỜI NÓI ĐẦU	1
CHƯƠNG 1: CƠ SỞ LÝ THUYẾT	3
1.1. Logic.....	3
1.1.1. Mệnh đề và các phép toán mệnh đề	3
1.1.2. Các phép toán trên mệnh đề.....	3
1.1.3. Các luật logic.....	6
1.2. Cơ sở dữ liệu phân tán	9
1.2.1. Phân mảnh ngang	12
1.2.2. Phân mảnh dọc	16
1.2.3. Phân mảnh hỗn hợp.....	17
CHƯƠNG 2: CÁC THUẬT TOÁN PHÂN MẢNH THEO CÁC GIÁ TRỊ LẬP CỦA CÁC THUỘC TÍNH TRONG CƠ SỞ DỮ LIỆU PHÂN TÁN.....	18
2.1. Miền giá trị thuộc tính và khái niệm giá trị lập của thuộc tính.....	18
2.1.2. Các giá trị lập trong miền giá trị của thuộc tính	18
2.2. Các thuật toán phân mảnh theo các giá trị lập của các thuộc tính	21
2.2.1. Thuật toán phân mảnh dọc	21
2.2.2. Thuật toán phân mảnh ngang theo giá trị lập và gộp nhóm theo tiêu chuẩn giữ lại mảnh ngang.....	42
2.2.3. Thuật toán phân mảnh hỗn hợp theo giá trị lập	48

2.2.4. Thuật toán tính size của mảnh.....	50
2.3. Gắn nhãn cho mảnh hỗn hợp.....	51
2.4. Kết luận chương	51
CHƯƠNG 3: ỨNG DỤNG QUẢN LÝ VẤN ĐỀ KINH DOANH TẠI HUYỆN ĐAN PHƯỢNG	52
3.1. Thiết kế cơ sở dữ liệu quan hệ kinh doanh tại huyện Đan Phượng	52
3.2. Mệnh đề phân mảnh các giá trị lặp lại của các thuộc tính theo các tiêu chí địa điểm, hộ kinh doanh, loại hình kinh doanh,	55
3.3. Gộp nhóm các giá trị lặp của các thuộc tính trong các mảnh ngang	64
3.4. Kết luận chương	67
KẾT LUẬN	68
HƯỚNG PHÁT TRIỂN CỦA ĐỀ TÀI	69
TÀI LIỆU THAM KHẢO.....	70

DANH MỤC BẢNG BIỂU

Bảng 1.1: Bảng tổng hợp giá trị chân trị của các phép toán cơ bản.....	6
Bảng 1.2: Bảng quan hệ PROJ.....	10
Bảng 1.3: Quan hệ được phân mảnh dọc PROJ1.....	11
Bảng 1.4: Quan hệ được phân mảnh dọc PROJ2.....	11
Bảng 1.5: Quan hệ được phân mảnh ngang với BUGGET ≤ 200	11
Bảng 1.6: Quan hệ được phân mảnh ngang với BUGGET > 200	12
Bảng 2.1. Quan hệ PhanNhiem.....	19
Bảng 2.1.1. Miền giá trị khóa.....	19
Bảng 2.1.2. Giá trị lập ChucVu: Giám đốc.....	20
Bảng 2.1.3. Giá trị lập ChucVu: Phân tích và thiết kế.....	20
Bảng 2.1. 4. Giá trị lập ChucVu: Kỹ sư.....	20
Bảng 2.1.5. Giá trị lập ThoiGianLV: 24.....	20
Bảng 2.1.6 Giá trị lập ThoiGianLV: 48.....	20
Bảng 2.1.7. Giá trị lập (ChucVu , ThoiGianLV).....	21
Bảng 2.2.1. Mảnh dọc theo ThoiGianLV.....	21
Bảng 2.2.2. Mảnh dọc theo ChucVu.....	22
Bảng 2.2.3. Mảnh dọc theo ThoiGianLV.....	23
Bảng 2.2.4. Mảnh dọc theo ChucVu.....	24
Bảng 2.2.5. Toàn quan hệ PhanNhiem.....	24
Bảng 2.2.5'. Toàn quan hệ PhanNhiem.....	43
Bảng 2.2.6. Quan hệ PhanNhiem.....	45
Bảng 2.2.7. Các mảnh ngang M_{21}	47
Bảng 2.2.8. Quan hệ PhanNhiem.....	47
Bảng 2.2.9. Mảnh ngang M_{11}	47
Bảng 2.2.10. Quan hệ PhanNhiem.....	48
Bảng 2.2.11. Mảnh ngang M_{12}	48
Bảng 3.1: Dữ liệu của bảng Dang_Ky_KD.....	52
Bảng 3.2: Bảng kết quả của mảnh H_1	56

Bảng 3.3: Bảng kết quả của mảnh H ₂	57
Bảng 3.4: Bảng kết quả của mảnh H ₃	58
Bảng 3.5: Bảng kết quả của mảnh H ₄	59
Bảng 3.6: Bảng kết quả của mảnh H ₅	61
Bảng 3.7: Bảng kết quả của mảnh H ₆	62
Bảng 3.8: Bảng kết quả của mảnh H ₇	63

DANH MỤC HÌNH

Hình 1.1: Mô tả mối quan hệ giữa các quan hệ bởi các đường nối.....	14
Hình 1.2: Phân đoạn hỗn hợp	17
Hình 2.1: Ma trận mẫu về giá trị sử dụng thuộc tính.....	26
Hình 2.2: Ma trận liên đới thuộc tính.....	29
Hình 2.3: Tính toán ma trận liên đới CA.....	35
Hình 2.4: Định vị một điểm tách.....	37
Hình 3.1: Giao diện ban đầu của chương trình	64
Hình 3.2: Giao diện sau quá trình chọn button “chonDB”	65
Hình 3.3: Giao diện sau khi chọn lựa xong bảng quan hệ	65
Hình 3.4: Giao diện quá trình chọn thuộc tính lặp	66
Hình 3.5: Kết quả quá trình phân mảnh ngang	67
Hình 3.6: Giao diện dữ liệu của một mảnh ngang	67

DANH MỤC CÁC CHỮ VIẾT TẮT

CSDL	Cơ sở dữ liệu
CSDLPT	Cơ sở dữ liệu phân tán
SQL	Structured Query Language (<i>Ngôn ngữ truy vấn có cấu trúc</i>)