BỘ GIÁO DỤC VÀ ĐÀO TẠO
VIỆN HÀN LÂM KHOA HỌC VÀ CÔNG NGHỆ VIỆT NAM

VIỆN CÔNG NGHỆ THÔNG TIN

CAO TÙNG ANH

KHAI THÁC DỮ LIỆU PHÂN TÁN BẢO TOÀN TÍNH RIÊNG TƯ

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

CAO TÙNG ANH

KHAI THÁC DỮ LIỆU PHÂN TÁN BẢO TOÀN TÍNH RIÊNG TƯ

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:

1. PGS.TSKH. NGUYỄN XUÂN HUY
2. PGS.TS. NGUYỄN MẬU HÂN

HÀ NỘI - 2014

LỜI CAM ĐOAN

Tôi cam đoan đây là công trình nghiên cứu của riêng tôi. Các số liệu, kết quả trong luận án là trung thực và chưa từng công bố trong bất kỳ công trình nào khác.

Tác giả luận án

Cao Tùng Anh

LỜI CÁM ƠN

Luận án được thực hiện và hoàn thành dưới sự hướng dẫn của PGS.TSKH. Nguyễn Xuân Huy và PGS.TS. Nguyễn Mậu Hân. Trong thời gian thực hiện luận án, tác giả đã nhận được sự giúp đỡ và chỉ dẫn khoa học rất tận tình từ hai người thầy của mình để có thể hoàn thành luận án này. Nhân dịp này tác giả xin được gửi đến hai thầy: PGS.TSKH.Nguyễn Xuân Huy và PGS.TS. Nguyễn Mậu Hân lòng biết ơn sâu sắc và lời cám ơn chân thành nhất.

Tác giả cũng xin được trân trọng cảm ơn PGS.TS. Thái Quang Vinh, GS.TS. Vũ Đức Thi, PGS.TS. Đoàn Văn Ban, PGS.TS. Đặng Văn Đức, PGS.TS. Ngô Quốc Tạo, PGS.TS. Đỗ Năng Toàn, PGS.TS. Lương Chi Mai, PGS.TS.Nguyễn Thanh Tùng là những thầy (cô) của Viện Công Nghệ Thông Tin đã quan tâm chỉ bảo, động viên và giúp đỡ tác giả trong suốt quá trình học tập, nghiên cứu và hoàn thiện luận án.

Tác giả cũng xin trân trọng cảm ơn PGS.TS. Lê Hoài Bắc và các bạn đồng nghiệp trong nhóm nghiên cứu tại TP.Hồ Chí Minh đã đọc và cho những ý kiến đóng góp quý báu cho nội dung luận án.

Cuối cùng xin chân thành cảm ơn các bạn đồng nghiệp tại khoa CNTT, trường Đại học Công nghệ TP.Hồ Chí Minh đã cổ vũ, động viên, giúp đỡ về nhiều mặt cho tác giả trong thời gian thực hiện luận án.

MỤC LỤC

DANH MỤC CÁC HÌNH

DANH MỤC CÁC BẢNG

DANH MỤC TỪ VIẾT TẮT

STT	Từ viết tắt	Diễn giải tiếng Anh	Diễn giải tiếng Việt
1	CSDL	Database	Cơ sở dữ liệu
2	CSDLPT	Database distributed	Cơ sở dữ liệu phân tán
3	GTDX	Proposed protocol	Giao thức đề xuất
4	WIT-Tree	Weighted Itemset-Tidset tree	Cây tập mục-tập giao dịch có trọng số
5	TWU	Tree Weighted Utility	Cây lợi ích có trọng số
6	FI	Frequent Itemsets	Tập phổ biến
7	FP-tree	Fast Parallel tree	Cây khai thác song song nhanh
8	FDM	Fast Distributed Mining	Khai thác phân tán nhanh
9	SVM	Support Vector machines	Sử dụng vectơ trong hỗ trợ phân lớp
10	PPDM	Privacy Preserving Data Mining	Khai thác dữ liệu bảo toàn tính riêng tư
11	RSA	Revest-Shamir-Adleman	Hệ mã hóa RSA
12	SM	Safety margin	Ngưỡng an toàn
13	MST	Min support	Độ hỗ trợ tối tiểu
14	MFI	Maximal Frequent Itemset	Tập phổ biến tối đại
15	MCT	Min Confident	Ngưỡng độ tin cậy
16	TID	Transaction index	Chỉ mục của giao dịch
17	IT-Tree	Itemset Tidset tree	Cây tập mục -tập giao dịch
18	HUIs	High Utility Itemsets	Tập tiện ích cao
19	DBS	Dynamic Bit String	Chuỗi bít động
20	SH	Semi Honest	Trung thực một nửa