

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

LÊ ANH DŨNG

**GIẤU TIN TRONG FILE ÂM THANH BẰNG
CÁC PHÉP BIẾN ĐỔI RỜI RẠC**

LUẬN VĂN THẠC SỸ: KHOA HỌC MÁY TÍNH

THÁI NGUYÊN, NĂM 2015

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

LÊ ANH DŨNG

**GIẤU TIN TRONG FILE ÂM THANH BẰNG CÁC PHÉP
BIẾN ĐỔI RỜI RẠC**

Chuyên ngành: Khoa học máy tính

Mã số: 60 48 0101

LUẬN VĂN THẠC SĨ: KHOA HỌC MÁY TÍNH

HƯỚNG DẪN KHOA HỌC: TS TRỊNH THANH LÂM

THÁI NGUYÊN, NĂM 2015

LỜI CAM ĐOAN

Tôi xin cam đoan luận văn “*Giấu tin trong file âm thanh bằng các phép biến đổi rời rạc*” là sản phẩm của riêng cá nhân, không sao chép lại của người khác. Trong toàn bộ nội dung của luận văn, những điều được trình bày hoặc là của cá nhân hoặc là được tổng hợp, nghiên cứu từ nhiều nguồn tài liệu. Tất cả các tài liệu tham khảo đều có xuất xứ và trích dẫn rõ ràng.

Tôi xin hoàn toàn chịu trách nhiệm và chịu mọi hình thức kỷ luật theo quy định cho lời cam đoan của mình.

Thái Nguyên, ngày 15 tháng 05 năm 2015

Học viên

Lê Anh Dũng

LỜI CẢM ƠN

Lời đầu tiên, tôi xin bày tỏ lòng biết ơn đến thầy TS Trịnh Thanh Lâm - ĐHQG Hà Nội, người đã tận tình hướng dẫn, chỉ bảo và giúp đỡ tôi trong suốt quá trình nghiên cứu và hoàn thành luận văn này.

Tôi xin chân thành cảm ơn các thầy cô giáo trường Đại học Công nghệ Thông tin và Truyền thông - Đại học Thái Nguyên đã giảng dạy và cung cấp cho chúng tôi những kiến thức rất bổ ích trong thời gian học cao học, giúp tôi có nền tảng tri thức để phục vụ nghiên cứu khoa học sau này.

Tôi cũng xin cảm ơn Lãnh đạo và đồng nghiệp tại đơn vị đã tạo điều kiện và giúp đỡ tôi trong suốt quá trình nghiên cứu và hoàn thành luận văn. Tôi cũng xin bày tỏ lòng cảm ơn đến gia đình và bạn bè, những người luôn quan tâm, động viên và khuyến khích tôi trong quá trình học tập.

Thái Nguyên, ngày 15 tháng 05 năm 2015

Lê Anh Dũng

MỤC LỤC

	Trang
LỜI CAM ĐOAN	i
LỜI CẢM ƠN	ii
MỤC LỤC	iii
DANH MỤC CÁC HÌNH ẢNH	vi
DANH MỤC CÁC BẢNG BIỂU	vii
MỞ ĐẦU	1
CHƯƠNG 1. TỔNG QUAN VỀ GIẤU TIN VÀ ÂM THANH SỐ	5
1.1. Giới thiệu chung về giấu tin	5
1.1.1. Mã hóa và giấu tin	5
1.1.2. Phân loại kỹ thuật giấu tin	6
1.2. Các đối tượng của một bài toán giấu tin	8
1.2.1. Thông tin mật	8
1.2.2. Đối tượng chứa	8
1.2.3. Đối tượng đã nhúng	9
1.2.4. Khoá mật	9
1.3. Mô hình kỹ thuật giấu tin	9
1.4. Các tiêu chí đánh giá bài toán giấu tin	10
1.4.1. Khả năng không bị phát hiện	10
1.4.2. Tính bền vững	11
1.4.3. Khả năng lưu trữ	11
1.4.4. Tính vô hình	12
1.4.5. Độ phức tạp của thuật toán	12
1.5. Một số ứng dụng cụ thể	12
1.6. Các tấn công trên các hệ giấu tin	15
1.7. Âm thanh số	16
1.7.1. Khái niệm về âm thanh và âm thanh số	17
1.7.2. Một số định dạng file âm thanh trên máy tính	18
1.7.3. Cấu trúc file âm thanh dạng WAV	21

1.8. Một số kỹ thuật giấu tin trong file âm thanh.....	23
CHƯƠNG 2. KỸ THUẬT GIẤU TIN BẰNG CÁC PHÉP BIẾN ĐỔI RỜI RẠC.....	26
2.1. Các phép biến đổi từ miền không gian sang miền tần số.....	26
2.1.1. Phép biến đổi Fourier.....	26
2.1.2. Phép biến đổi Cosin rời rạc.....	27
2.1.3. Phép biến đổi Wavelet.....	29
2.2. Một số kỹ thuật giấu tin dựa trên biến đổi khối bit nhị phân.....	30
2.2.1. Mã hóa LSB (Least Significant Bit).....	31
2.2.2. Mã hóa Parity (Parity Coding).....	32
2.3. Thuật toán giấu tin bằng các phép biến đổi rời rạc trên số nguyên.....	34
2.3.1. Một số phép biến đổi rời rạc trên số nguyên.....	34
2.3.2. Thuật toán Wu-Lee.....	35
2.3.3. Thuật toán Wu-Lee cải tiến.....	38
2.3.4. Thuật toán giấu một chuỗi bit trong một khối tin.....	40
CHƯƠNG 3. TRIỂN KHAI CHƯƠNG TRÌNH THỬ NGHIỆM.....	48
3.1. Mục đích, yêu cầu.....	48
3.2. Yêu cầu về cấu hình hệ thống.....	48
3.3. Lựa chọn định dạng file âm thanh trong thực nghiệm.....	48
3.4. Sơ đồ chương trình.....	49
3.5. Thuật toán giấu tin và trích rút tin theo kỹ thuật đề xuất.....	50
3.5.1. Giấu tin.....	50
3.5.2. Trích rút tin mật.....	52
3.5.3. Một số hàm và thủ tục giấu tin.....	53
3.6. Kết quả thực nghiệm.....	54
3.7. Đánh giá kết quả thực nghiệm.....	64
3.8. Các khả năng ứng dụng.....	64
KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN.....	67
TÀI LIỆU THAM KHẢO.....	69

DANH MỤC CÁC TỪ VIẾT TẮT TRONG LUẬN VĂN

- AAC - Định dạng âm thanh chuẩn (*Advanced Audio Coding*)
A/D D/A - Biến đổi tương tự/số và ngược lại (*Analog/Digital*)
AIFF - Định dạng không mất thông tin (*Audio Interchange File Format*)
DCT - Phép biến đổi Cosin rời rạc (*Discrete Cosine Transform*).
DES - Hệ mật mã chuẩn (*Data Encryption Standard*)
DSP - Xử lý tín hiệu kỹ thuật số (*Digital signal processing*)
FLAC - Nén âm thanh không mất dữ liệu (*Free Lossless Audio Codec*),
FT - Biến đổi Fourier (*Fourier Transform*)
HAS - Hệ thống thính giác (*Human Auditory system*)
HVS - Hệ thống thị giác (*Human Vision System*)
IDE - Môi trường phát triển tích hợp (*Integrated Development Environment*)
IFT - Biến đổi Fourier ngược (*Inverse Fourier Transform*)
LSB - Bít ít quan trọng nhất (*Least Significant Bit*)
MP3 - Định dạng nén âm thanh (*Movie Picture Experts Group-Layer 3*)
PCM - Điều biến mã xung (*Pulse Code Modulation*)
RSA - Mã hóa công khai RSA (*Rivest, Shamir và Adleman*)
WAV - Định dạng âm thanh dạng sóng (*Waveform Audio Format*)
WMA - Định dạng âm thanh của Microsoft (*Windows Media Audio*)

DANH MỤC CÁC HÌNH ẢNH

Trang

Hình 1.1. Mô hình mã hóa thông tin	5
Hình 1.2. Một cách phân loại các kỹ thuật giấu tin	7
Hình 1.3. Lược đồ chung cho quá trình giấu tin	9
Hình 1.4. Lược đồ chung cho quá trình trích rút thông tin	10
Hình 1.5. Môi quan hệ giữa các yếu tố trong bài toán giấu tin.....	12
Hình 1.6. Ảnh gốc Lena và logo của viện khoa học Ấn Độ	13
Hình 1.7. Ảnh Lena đã được nhúng thủy vân hiển	14
Hình 1.8. Thông tin bị xuyên tạc	14
Hình 1.9. Tín hiệu âm thanh.....	17
Hình 1.10. Số hóa tín hiệu âm thanh.....	18
Hình 1.11. Mô tả định dạng của file.wav.....	21
Hình 1.12. Mô tả 72 byte của một file âm thanh WAV.....	23
Hình 2.1. Minh họa kỹ thuật LSB.....	31
Hình 2.2. Minh họa kỹ thuật mã hóa Parity	33
Hình 3.1. Sơ đồ chương trình thử nghiệm	49
Hình 3.2. Phổ biên độ và phổ pha của file chưa trước khi giấu tin	57
Hình 3.3. Phổ biên độ và phổ pha của file sau khi giấu tin	57
Hình.3.4. Trích đoạn các byte của file Sony.wav sau khi nhúng tin mật	63

DANH MỤC CÁC BẢNG BIỂU

	Trang
Bảng 1.1. So sánh giấu thông tin mật và giấu thông tin thủy vân	8
Bảng 1.2. Một số định dạng file âm thanh trên máy tính	21
Bảng 1.3. Phần định dạng kiểu RIFF	22
Bảng 1.4. Phần định dạng thông tin âm thanh	22
Bảng 1.5. Phần dữ liệu âm thanh	23
Bảng 3.1. Một số phần mềm giấu tin	49

MỞ ĐẦU

1. Đặt vấn đề

Ngày nay, Internet là môi trường phổ biến cho việc trao đổi thông tin giữa các nhà cung cấp và người sử dụng. Do đó, vấn đề an toàn dữ liệu trên mạng luôn luôn là một thách thức đối với các nhà quản lý và các nhà nghiên cứu. Các thông tin trên Internet có thể dễ dàng bị làm giả mạo, sai lệch và bị đánh cắp bởi hacker trong quá trình truyền tải dữ liệu. Thông tin của cá nhân, tổ chức hoặc quốc gia đứng trước nguy cơ bị xâm nhập bất cứ lúc nào. Cùng với nó là vấn nạn ăn cắp bản quyền, xuyên tạc thông tin,... ngày càng gia tăng. Vì vậy, vấn đề đặt ra làm thế nào để đảm bảo được sự an toàn, và toàn vẹn thông tin trong quá trình truyền tải trên Internet. Hai giải pháp cho vấn đề này là mã hóa và giấu thông tin có vai trò quan trọng trong việc bảo vệ quá trình truyền tải thông tin mật. Sự xác thực và bản quyền trong môi trường trao đổi công cộng. Việc tìm giải pháp cho những vấn đề này giúp ta hiểu thêm về một công nghệ đang phát triển và còn tạo ra những cơ hội mới [1].

Trong những giải pháp đã và đang được triển khai thì giấu tin (Data Hiding) là một trong những giải pháp được các nhà nghiên cứu và phát triển coi đó là một hướng đi có nhiều triển vọng. Giấu thông tin là kỹ thuật nhúng một lượng thông tin số nào đó vào trong một đối tượng thông tin số khác mà các đối tượng đó thường là một tài liệu, hình ảnh, âm thanh hoặc video. Các kỹ thuật giấu tin có thể chia ra làm hai nhóm. Nhóm thứ nhất là các phương pháp che giấu thông tin trực tiếp. Nhóm này thường sử dụng các bit ít quan trọng nhất của một khối bit nhị phân được sửa đổi để giấu thông tin. Nhóm thứ hai lại che giấu thông tin thông qua các phép biến đổi chẳng hạn như phép biến đổi Cosin hay wavelet rời rạc được sử dụng rộng rãi [4].

Sau khi tiến hành nghiên cứu các tài liệu liên quan đến lĩnh vực giấu tin trong đa phương tiện và nhận thấy các kỹ thuật trên đều cho kết quả tốt với