TRƯỜNG ĐẠI HỌC HÀNG HẢI VIỆT NAM

KHOA CÔNG NGHỆ THÔNG TIN

THUYẾT MINH
ĐỀ TÀI NCKH CẤP TRƯỜNG

ĐỀ TÀI

NGHIÊN CỨU KỸ THUẬT GIẤU THÔNG TIN TRONG AUDIO - ỨNG DỤNG CHO BẢN QUYỀN AUDIO SỐ VÀ TRUYỀN TIN MẬT BẰNG AUDIO

Chủ nhiệm đề tài: TS. HỒ THỊ HƯƠNG THƠM

Thành viên tham gia: THS. TRỊNH THỊ NGỌC HƯƠNG

THS. NGUYỄN THỊ THU HÀ

Hải Phòng, tháng 5 / 2015

MỤC LỤC

DANH MỤC CÁC CỤM TỪ VIẾT TẮT 4

DANH MỤC CÁC HÌNH VẼ 5

LỜI MỞ ĐẦU 6

Chương I. TỔNG QUAN 7

1.1. Giấu thông tin là gì ? 7

1.2. Phân loại các kỹ thuật giấu tin 9

1.2.1. Phân loại theo phương tiện chứa tin: 9

1.2.2. Phân loại theo cách thức tác động lên các phương tiện 11

1.2.3. Phân loại theo các mục đích sử dụng: 12

1.3. Một số ứng dụng và xu hướng phát triển 13

1.4. Âm thanh, đặc tính và cấu trúc lưu trữ của âm thanh 14

1.4.1. Sóng âm và cảm giác âm 14

1.4.2. Mã hóa âm thanh wave 14

1.4.3. Cấu trúc tập tin âm thanh 16

1.5. Sơ đồ giấu tin và tách tin trong dữ liệu âm thanh 17

1.6. Tăng độ an toàn cho thông tin đem giấu 18

1.7. Đánh giá chất lượng âm thanh sau khi giấu tin 19

Chương II. KỸ THUẬT GIẤU TIN TRONG ÂM THANH 20

2.1. Kỹ thuật giấu tin trên bit có trọng số thấp LSB (Least Significant Bit) 20

2.1.1. Thuật toán giấu tin: 22

2.1.2. Thuật toán tách tin: 22

2.2. Kỹ thuật giấu tin ảnh xám trên tín hiệu audio (LSB nâng cao) 23

2.2.1. Thuật toán giấu tin: 23

2.2.2. Thuật toán tách tin: 24

2.3. Đề xuất phương pháp giấu văn bản kết hợp MSB và LSB của tín hiệu audio 25

2.3.1. Phương pháp 1 26
2.3.2. Phương pháp 2 27
2.4. Kỹ thuật thủy vân số trên miền biến đổi của tín hiệu audio 28
2.4.1. Thuật toán giấu tin: 31
2.4.2. Thuật toán tách tin: 32
Chương III. CÀI ĐẶT, THỬ NGHIỆM VÀ ĐÁNH GIÁ 34
3.1. Môi trường cài đặt và một số giao diện chương trình 34
3.1.1. Môi trường cài đặt 34
3.1.2. Giao diện chương trình 34
3.2. Thử nghiệm 39
3.2.1. Tập dữ liệu thử nghiệm 39
3.2.2. Thực hiện thử nghiệm 40
3.3. Nhận xét và đánh giá 42
KẾT LUẬN 44
TÀI LIỆU THAM KHẢO 45

DANH MỤC CÁC CỤM TỪ VIẾT TẮT

AES	Advanced Encryption Standard
BPF	Band Pass Filter
DES	Data Encryption Standard
EOF	End of file
HVS	Human Vision System
HAS	Human Auditory System
LSB	Least Significant Bit
LPF	Low Pass Filter
MSB	Most Significant Bit
MOS	Mean Opinion Score
SNR	Signal-To-Noise Ratio
RSA	Ron Rivest, Adi Shamir và Len Adleman
PCM	Pulse Code Modulation

DANH MỤC CÁC HÌNH VẼ

Hình 1.1. Kỹ thuật giấu tin công bố trong giai đoạn 1992 – 2007[3] 8
Hình 1.2. Tỉ lệ phương tiện được lựa chọn để giấu tin năm 2008 [3] 8
Hình 1.3: Sơ đồ phân loại các kỹ thuật giấu thông tin 13
Hình 1.4: Định dạng tệp wave 15
Hình 1.5: Trích mẫu 15
Hình 1.6. Sơ đồ tổng quát giấu tin và tách tin trong audio.....18
Hình 2.1. Giá trị 8 bit của tín hiệu A = 218.....20
Hình 2.2. Giá trị 8 bit của tín hiệu B=219 sau khi giấu bit 1 vào LSB của A.....21
Hình 2.3. Giá trị 8 bit của tín hiệu C=219 sau khi giấu bit 0 vào LSB của A.....21
Hình 2.4. Sơ đồ giấu tin trên 8 bit LSB của tín hiệu audio cơ sở.....21
Hình 2.5. Minh họa kỹ thuật giấu ảnh trên tín hiệu audio.....25
Hình 2.6. Minh họa kỹ thuật giấu văn bản dựa vào7 bit MSB và 4 bit LSB của tín hiệu audio 27
Hình 2.7. Minh họa kỹ thuật giấu văn bản dựa trên 4 bit MSB và 3 bit LSB của tín hiệu audio 28
Hình 2.8. Sơ đồ tổng quát giấu tin và tách tin trên miền tần số [9].....29
Hình 2.9. Biến đổi tần số Fourier của tín hiệu cơ sở và tín hiệu thông điệp.....30
Hình 2.10. a) Tín hiệu cơ sở sử dụng bộ lọc LPF, b) tín hiệu thông điệp sử dụng BPF 30
Hình 2.11. Tín hiệu đã mang thông tin giấu.....31
Hình 2.12. Tín hiệu thông điệp sau khi tách ra 31
Hình 3.1. Giao diện chính của chương trình 34
Hình 3.2. Giao diện chính với module chức năng giấu tin mật (Stegangraphy).....35
Hình 3.3. Giao diện chính với module chức năng Thủy vân số (Watermarking) 35
Hình 3.4. Giao diện giấu tin văn bản trên LSB của Audio.....36
Hình 3.5. Giao diện tách tin văn bản đã giấu trên LSB của Audio 36
Hình 3.6. Giao diện thủy vân một ảnh nhị phân trên LSB của Audio 37
Hình 3.7. Giao diện tách thủy vân một ảnh nhị phân trên LSB của Audio.....37
Hình 3.8. Giao diện thủy vân một audio trên tín hiệu biến đổi của Audio gốc.....38
Hình 3.9. Giao diện tách thủy vân là một audio trên tín hiệu biến đổi của Audio.....38
Hình 3.10. Giao diện tính thay đổi nhiễu PSNR của audio gốc và audio giấu tin 39
Hình 3.11. ảnh nhị phân cần giấu trong tập dữ liệu audio.....41

LỜI MỞ ĐẦU

Giấu thông tin (Hiding information) là kỹ thuật giấu thông tin quan trọng vào đối tượng dữ liệu số mà không làm ảnh hưởng trực giác đến chất lượng ban đầu của dữ liệu số. Dữ liệu số dùng để che giấu tin có thể là ảnh số (image), âm thanh số (audio), phim hoặc đoạn clip (video)... Giấu tin có hai mục đích chính:

- Thứ nhất, giấu tin nhằm mục đích bảo vệ cho chính tài liệu số dùng để bao che thông tin giấu, đây chính là hình thức dùng để bảo vệ bản quyền hoặc chống xuyên tạc nội dung ... Và người ta gọi hình thức giấu này là thủy vân số (Watermarking). Thủy vân số chỉ cần giấu với một lượng thông tin rất nhỏ nhưng đủ mạnh (bền vững) để có thể bảo vệ tài liệu số có mang thông tin.
- Thứ hai, giấu tin nhằm mục đích trao đổi thông tin mật đến một đối tượng đồng minh mà không muốn đối tượng thứ ba (không mong muốn) có thể phát hiện ra hay nghi ngờ. Người ta gọi hình thức giấu này là giấu tin mật (Steganography), các kỹ thuật giấu theo hình thức này thường cố gắng nghiên cứu giấu được càng nhiều thông tin vào dữ liệu số càng tốt nhưng vẫn đảm bảo chất lượng theo trực giác của dữ liệu bao tin.

Lĩnh vực giấu tin đã và đang thu hút sự quan tâm của nhiều nghiên cứu trên thế giới, đặc biệt trong lĩnh vực an ninh, quốc phòng, thương mại, …. Tuy nhiên ở Việt Nam giấu tin trong ảnh đã có rất nhiều công trình nghiên cứu và công bố, còn giấu tin trong âm thanh số vẫn còn rất hạn chế, chưa có nhiều công bố rộng rãi trong lĩnh vực này. Do đó trong đề tài nghiên cứu này nghiên cứu và tìm hiểu một số kỹ thuật giấu tin trong âm thanh với cả hai hình thức Watermarking và Steganography. Dựa vào cơ sở nghiên cứu được xây dựng chương trình giấu tin trong tín hiệu âm thanh với nhiều kiểu dữ liệu thông tin có thể giấu trong audio: văn bản, ảnh nhị phân, ảnh cấp xám hay một đoạn audio.

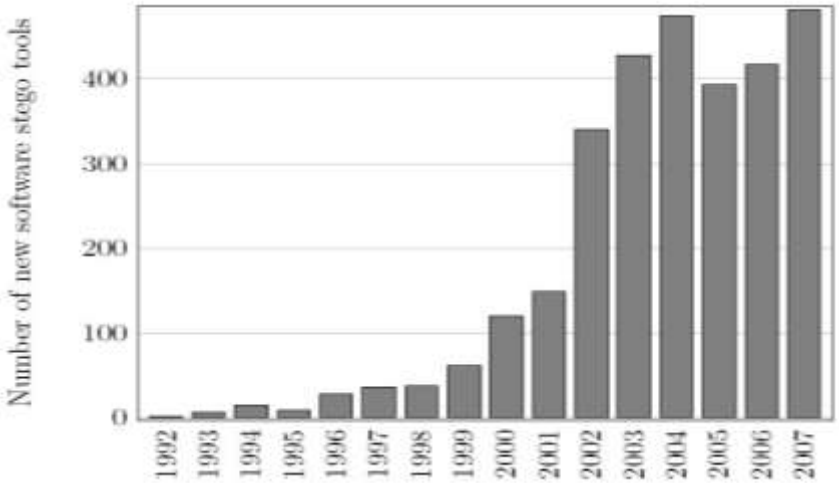
Chương I. TỔNG QUAN

1.1. Giấu thông tin là gì ?

"Giấu thông tin" (Steganography) là kỹ thuật liên lạc mật dựa trên hình thức giấu thông tin quan trọng vào đối tượng khác. Từ thời kỳ cổ đại người ta đã sử dụng phương pháp này để liên lạc mật cho nhau. Một ví dụ cổ điển hình về giấu tin [1] (485-525 trước công nguyên) là câu chuyện của một người tên là Histaiæus muốn gửi thông tin quan trọng về "Kế hoạch ủng hộ cuộc nổi dậy chống lại đức vua Ba tư Xerxes" cho nhà nhiếp chính thành phố Miletus bằng cách xăm thông tin lên da đầu người nô lệ tin cậy của mình, cho đến khi tóc mọc dài trở lại ông ta cử người nô lệ đó đến gặp nhà nhiếp chính, trên đường đi không bị lộ thông tin, có thể vượt qua mọi cuộc kiểm soát, khi đến gặp nhà nhiếp chính, nhà nhiếp chính chỉ việc cạo trọc đầu người nô lệ là có thể đọc được thông tin. Hay một phát minh khác của Pliny T. Elder (23-79 sau công nguyên) về mực "không màu" chính là sữa của một số loại động vật, khi mực này viết trên giấy để khô khó phát hiện ra, và chỉ khi giấy đó được hơ nóng các vết mực sẽ chuyển sang nâu, có thể đọc các dòng chữ đã viết. Vào thời kỳ phục hưng, năm 1518 Johannes Trithemius viết cuốn sách về mã hóa "Polygraphia". Trong cuốn sách này người ta thấy xuất hiện đầu tiên thuật ngữ "Steganographia", đây là một từ ghép bắt nguồn từ ngôn ngữ Hy lạp *steganos* nghĩa là bao bọc "cover" và *graphia* nghĩa là *bản viết* "writing" [2].

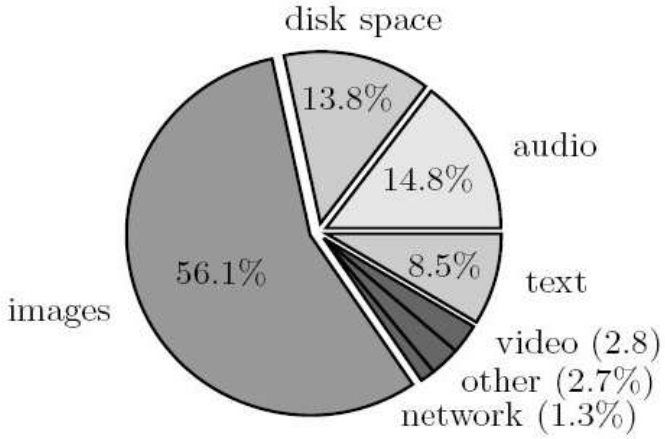
Trải qua nhiều thời kỳ biến động của xã hội loài người, ngày nay khi mà kỹ thuật số bùng nổ, con người cũng "số hoá" lĩnh vực giấu tin phục vụ cho cuộc sống hiện đại. Trong [3] người ta đã thông kê sự phát triển mạnh mẽ của các công cụ giấu tin trong giai đoạn 1992 đến 2007.

Giấu tin có một ưu điểm mà mật mã học (Cryptography) còn hạn chế đó là có thể "bảo vệ được bản quyền số, hay khi giữa các đối tượng liên lạc mật với nhau trên các kênh thông tin công cộng mà ít bị nghi ngờ". Lý do vì bản quyền số đã mã hóa sau khi được giải mã thì khó có thể giữ được bản quyền, hay thông tin mật cần trao đổi giữa các bên, sau khi được mã hóa sẽ làm cho người khác biết rõ là các bên có trao đổi thông tin mật nào đó cho nhau. Giấu tin trong dữ liệu đa phương tiện chính là những "bổ khuyết" cho các vấn đề trên của mật mã học.



Hình 1.1. Kỹ thuật giấu tin công bố trong giai đoạn 1992 – 2007[3]

Về nguyên lý, giấu tin trong dữ liệu video, dữ liệu âm thanh hay trong dữ liệu ảnh số không khác gì nhiều. Giấu tin trong ảnh dễ thực hiện hơn, giấu được nhiều thông tin hơn, và ảnh là đối tượng được sử dụng khá phổ biến trên Internet hiện nay, nên kỹ thuật giấu tin trong ảnh chiếm tỉ lệ nhiều nhất trong các loại dữ liệu đa phương tiện (hình 1.2).



Hình 1.2. Tỉ lệ phương tiện được lựa chọn để giấu tin năm 2008 [3]

Giấu tin trong audio dựa vào hệ thống thính giác của con người nên thường khó hơn trong các dữ liệu số khác do hệ thống thính giác của con người khá nhạy cảm với các nhiễu, tuy nhiên audio cũng là loại dữ liệu được sử dụng khá phổ biến trên internet hiện nay vì vậy nó vẫn được quan tâm và nghiên cứu rộng rãi thứ 2 theo [3] đưa ra (hình 1.2) cho nhiều mục đích khác nhau: bảo vệ bản quyền, chống sao chép, xuyên tạc nội dung...

1.2. Phân loại các kỹ thuật giấu tin

Có nhiều cách để tiến hành phân loại các phương pháp giấu thông tin thông qua các tiêu chí khác nhau, như theo phương tiện chứa tin, các phương pháp tác động lên các kỹ thuật các phương tiện chứa tin, hay phân loại dựa theo các ứng dụng cụ thể...

1.2.1. Phân loại theo phương tiện chứa tin:

1.2.1.1. Giấu thông tin trong ảnh:

Thông tin sẽ được giấu cùng với dữ liệu ảnh nhưng chất lượng ảnh ít thay đổi và không ai biết được đằng sau ảnh đó mang những thông tin có ý nghĩa. Ngày nay, khi ảnh số đã được sử dụng rất phổ biến, thì giấu thông tin trong ảnh đã đem lại rất nhiều những ứng dụng quan trọng trên nhiều lĩnh vực trong đời sống xã hội. Ví dụ như đối với các nước phát triển, chữ kí tay đã được số hoá và lưu trữ sử dụng như là hồ sơ cá nhân của các dịch vụ ngân hàng và tài chính, nó được dùng để nhận thực trong các thẻ tín dụng của người tiêu dùng. Phần mềm WinWord của MicroSoft cũng cho phép người dùng lưu trữ chữ kí trong ảnh nhị phân rồi gắn vào vị trí nào đó trong file văn bản để đảm bảo tính an toàn của thông tin. Tài liệu sau đó được truyền trực tiếp qua máy fax hoặc lưu truyền trên mạng. Theo đó, việc nhận thực chữ kí, xác thực thông tin đã trở thành một vấn đề cực kì quan trọng khi mà việc ăn cắp thông tin hay xuyên tạc thông tin bởi các tin tặc đang trở thành một vấn nạn đối với bất kì quốc gia nào, tổ chức nào. Thêm vào đó, lại có rất nhiều loại thông tin quan trọng cần được bảo mật như những thông tin về an ninh, thông tin về bảo hiểm hay các thông tin về tài chính, các thông tin này được số hoá và lưu trữ trong hệ thống máy tính hay trên mạng. Chúng rất dễ bị lấy cắp và bị thay đổi bởi các phần mềm chuyên dụng. Việc nhận thực cũng như phát hiện thông tin xuyên tạc đã trở nên vô cùng quan trọng, cấp thiết. Và một đặc điểm của giấu thông tin trong ảnh đó là thông tin được giấu trong ảnh một cách vô hình, nó như là một cách mà truyền thông tin mật cho nhau mà người khác không thể biết được bởi sau khi giấu thông tin thì chất lượng ảnh gần như không thay đổi đặc biệt đối với ảnh mầu hay ảnh xám.

1.2.1.2. Giấu tin trong audio

Giấu thông tin trong audio mang những đặc điểm riêng, khác với giấu thông tin trong các đối tượng đa phương tiện khác. Một trong những yêu cầu cơ bản của giấu tin là đảm bảo tính chất ẩn của thông tin được giấu đồng thời không làm ảnh hưởng đến chất lượng của dữ liệu gốc. Để đảm bảo yêu cầu này, kỹ thuật giấu thông tin trong ảnh phụ thuộc vào hệ thống thị giác của con người - HVS (Human Vision System) còn kỹ thuật giấu thông tin trong audio lại phụ thuộc vào hệ thống thính giác HAS (Human Auditory System). Và một vấn đề khó khăn ở đây là hệ thống thính giác của con người nghe được các tín hiệu ở các giải tần rộng và công suất lớn nên đã gây khó dễ đối với các phương pháp giấu tin trong audio. Nhưng thật may là HAS lại kém trong việc phát hiện sự khác biệt các dải tần và công suất điều này có nghĩa là các âm thanh to, cao tần có thể che giấu được các âm thanh nhỏ thấp một cách dễ dàng. Các mô hình phân tích tâm lí đã chỉ ra điểm yếu trên và thông tin này sẽ giúp ích cho việc chọn các audio thích hợp cho việc giấu tin. Vấn đề khó khăn thứ hai đối với giấu thông tin trong audio là kênh truyền tin. Kênh truyền hay băng thông chậm sẽ ảnh hưởng đến chất lượng thông tin sau khi giấu. Ví dụ để nhúng một đoạn java applet vào một đoạn audio (16 bit, 44.100 Hz) có chiều dài bình thường thì các phương pháp nói chung cũng cần ít nhất là 20 bit/s. Giấu thông tin trong audio đòi hỏi yêu cầu rất cao về tính đồng bộ và tính an toàn của thông tin. Các phương pháp giấu thông tin trong audio đều lợi dụng điểm yếu trong hệ thống thính giác của con người.

1.2.1.3. Giấu thông tin trong video

Cũng giống như giấu thông tin trong ảnh hay trong audio, giấu tin trong video cũng được quan tâm và được phát triển mạnh mẽ cho nhiều ứng dụng như điều khiển truy cập thông tin, nhận thực thông tin và bảo vệ bản quyền tác giả. Ta có thể lấy một ví dụ là các hệ thống chương trình trả tiền xem theo đoạn với các video clip (pay per view application). Các kỹ thuật giấu tin trong video cũng được phát triển mạnh mẽ và cũng theo hai khuynh hướng là thuỷ vân số và giấu tin mật. Nhưng phần giới thiệu này chỉ quan tâm tới các kỹ thuật giấu tin trong video. Một phương pháp giấu tin trong

video được đưa ra bởi Cox là phương pháp phân bố đều. ý tưởng cơ bản của phương pháp là phân phối thông tin giấu dàn trải theo tần số của dữ liệu chứa gốc. Nhiều nhà nghiên cứu đã dùng những hàm cosin riêng và các hệ số truyền sóng riêng để giấu tin. Trong các thuật toán khởi nguồn thì thường các kỹ thuật cho phép giấu các ảnh vào trong video nhưng thời gian gần đây các kỹ thuật cho phép giấu cả âm thanh và hình ảnh vào video. Như phương pháp của Swanson đã sử dụng phương pháp giấu theo khối, phương pháp này đã giấu được hai bit vào khối 8*8. Hay gần đây nhất là phương pháp của Mukherjee là kỹ thuật giấu audio vào video sử dụng cấu trúc lưới đa chiều...

1.2.1.4. Giấu thông tin trong văn bản dạng text

Giấu thông tin vào các văn bản dạng text khó thực hiện hơn do có ít các thông tin dư thừa, để làm được điều này người ta phải khéo léo khai thác các dư thừa tự nhiên của ngôn ngữ. Một cách khác là tận dụng các định dạng văn bản (mã hóa thông tin và khoảng cách giữa các từ hay các dòng văn bản). Từ nội dung của thông điệp cần truyền đi, người ta cũng có thể sử dụng văn phạm phi ngữ cảnh để tạo nên các văn bản "phương tiện chứa" rồi truyền đi.

1.2.2. Phân loại theo cách thức tác động lên các phương tiện

1.2.2.1. Phương pháp chèn dữ liệu

Phương pháp này tìm các vị trí trong tệp (file) dễ bị bỏ qua và chèn các dữ liệu cần giấu vào đó, cách giấu này không làm ảnh hưởng gì tới sự thể hiện của các file dữ liệu ví dụ như dữ liệu được giấu sau các ký tự EOF (End of file).

1.2.2.2. Phương pháp thay thế

Thực hiện việc thay thế các phần tử không quan trọng của phương tiện chứa bằng các dữ liệu của thông điệp cần chuyển đi. Vì thay thế vào các phần tử không quan trọng của phương tiện chứa nên dễ đánh lừa được cảm nhận của con người. Phương pháp thay thế có nhiều cách thực hiện như: thay thế các bit ít quan trọng, thay thế trong miền tần số, các kỹ thuật trải phổ, thống kê.

1.2.2.3. Phương pháp tạo các phương tiện chứa

Từ các thông điệp cần chuyển sẽ tạo ra các phương tiện chứa để phục vụ cho việc truyền thông tin đó, tại phía người nhận dựa trên các phương tiện chứa này để tái tạo lại các thông điệp.

1.2.3. Phân loại theo các mục đích sử dụng

1.2.3.1.Giấu thông tin mật

Đây là ứng dụng phổ biến nhất từ trước tới nay, đối với loại này người ta quan tâm chủ yếu tới các mục tiêu: độ an toàn của thông điệp và lượng thông điệp tối đa có thể giấu mà vẫn đảm bảo an toàn, độ bảo mật của thông tin trong trường hợp giấu tin bị phát hiện.

1.2.3.2.Giấu thông tin thuỷ vân

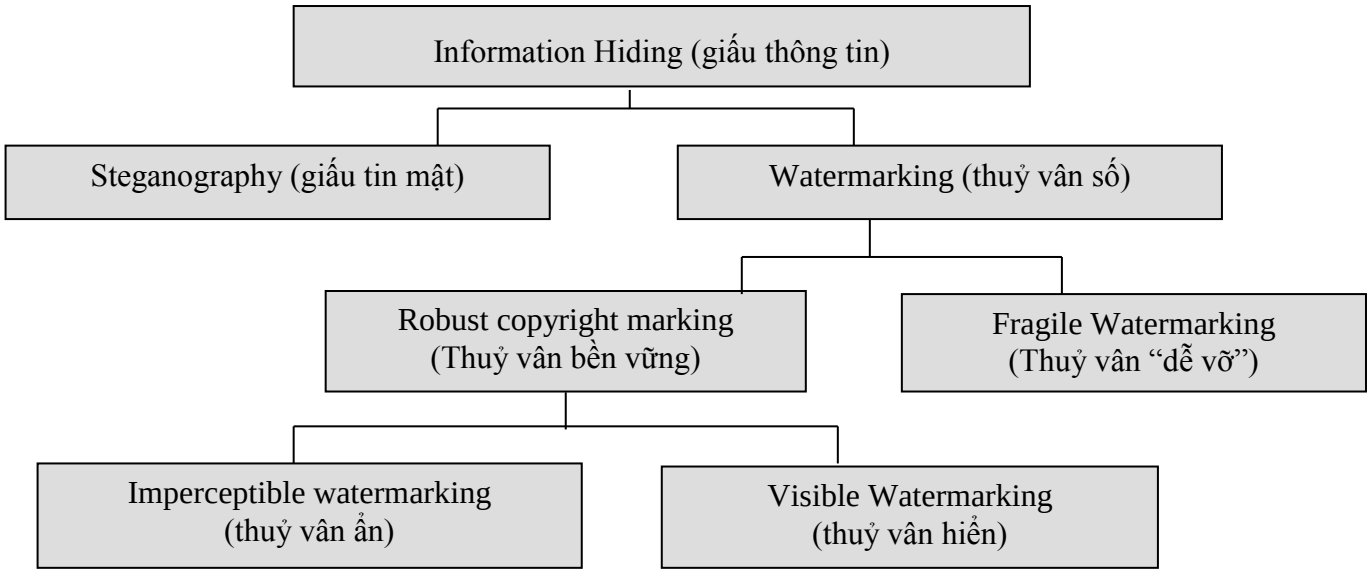
Do yêu cầu bảo vệ bản quyền, xác thực và chống xuyên tạc... nên giấu tin thuỷ vân có yêu cầu khác với giấu tin bí mật. Yêu cầu đầu tiên là các dấu hiệu thuỷ vân phải đủ bền vững trước các tấn công vô tình hay cố ý gỡ bỏ nó. Thêm vào đó các dấu hiệu thuỷ vân phải có ảnh hưởng tối thiểu (về mặt cảm nhận) đối với các phương tiện chứa. Như vậy các thông tin cần giấu sẽ càng nhỏ càng tốt. Phân biệt giấu thông tin bí mật và thuỷ vân có thể mô tả tóm lược trong các bảng 1.1 sau [1]:

Bảng 1.1. So sáng sự khác biệt giữa giấu tin mật và thủy vân số

	Giấu thông tin bí mật (steganography)	**Thuỷ vân sô (watermark)**
Mục tiêu	Tàng hình các phiên liên lạc, để bảo mật thông tin Dùng trong các liên lạc xác đinh	Chủ yếu phục vụ cho mục đích bảo vệ bản quyền Chủ yếu dùng trong các hoạt động xuất bản
Cách thực hiện	Không làm "thay đổi" phương tiện chứa	Có thể tác động nhỏ về cảm nhận tới phương tiện chứa
Yêu	Giấu được nhiều thông tin nhất	Chỉ cần nhúng ít dữ liệu

cầu	Không cần quan tâm tới độ bền của phương tiện chứa Không thể quan sát được việc giấu thông tin Không kiểm tra được nếu không có khoá thích hợp	Dữ liệu nhúng cần phải mạnh Đảm bảo trước các phương pháp nén dữ liệu Dữ liệu nhúng có thể nhận thấy hay không nhận thấy Không kiểm tra được nếu không có khoá thích hợp.

Hình 1.3 là sơ đồ phân loại giấu tin được Fabien A.P.Petitcolas đề xuất năm 1999 [5].



Hình 1.3: Sơ đồ phân loại các kỹ thuật giấu thông tin

1.3. Một số ứng dụng và xu hướng phát triển

Che giấu thông tin nói chung có rất nhiều ứng dụng tuỳ theo từng hoàn cảnh cụ thể. Giấu thông tin bí mật góp phần thực hiện "tàng hình" các phiên liên lạc, một bổ sung lý tưởng cho công tác bảo mật thông tin. Dưới đây là một số ứng dụng đang được triển khai trên thế giới:

- Bảo vệ bản quyền tác giả (copyright protection)
- Nhận thực thông tin hay phát hiện xuyên tạc thông tin (authentication and temper detection)

- Giấu vân tay hay dán nhãn (fingerprinting and labeling)
- Điều khiển sao chép (copy control)
- Truyền thông tin mật (steganography)

1.4. Âm thanh và cấu trúc lưu trữ của âm thanh

1.4.1. Sóng âm và cảm giác âm

Khi một vật dao động về một phía nào đó, nó làm cho các lớp không khí liền trước bị nén lại, và lớp không liền sau dãn ra. Sự nén và dãn không khí như vậy lặp đi lặp lại một cách tuần hoàn nên đã tạo ra trong không khí một sóng đàn hồi. Sóng này truyền tới tai, nén vào màng nhĩ khiến cho màng nhĩ cũng dao động với cùng tần số. Khi màng nhĩ dao động, các vị trí phân biệt của màng nhĩ trên bề mặt giống như nó chuyển động về trước hay sau đáp ứng với các sóng âm vào. Khi cùng một thời điểm, ta nghe thấy nhiều âm, thì mọi âm thanh phân biệt này được trộn với nhau một cách tự nhiên trong tai giống như một hình mẫu đơn của áp suất không khí thay đổi. Tai và óc làm việc cùng nhau để phân tích tín hiệu này ngược lại thành những cảm giác về âm riêng biệt.

1.4.2. Mã hóa âm thanh wave

Dữ liệu âm thanh trong tập tin WAV là dạng dữ liệu âm thanh không nén (RAW data) dựa trên định dạng mã hóa PCM (Pulse Code Modulation). Định dạng WAV có ưu điểm là cấu trúc đơn giản , chất lượng âm thanh được bảo toàn, nhưng nhược điểm là dùng lượng file khá lớn. Nếu được lấy mẫu với tần số 44.1 kHz (44100 lần/giây), độ phân giải 16 bit (tương đương với chất lượng CD) thì 1 phút âm thanh sẽ tiêu tốn tới 10 MB, nghĩa là một bài hát khoảng 5 phút sẽ mất dung lượng 50MB ổ cứng. Phổ biến hơn là các chuẩn nén âm thanh không bảo toàn nội dung (Lossy) như MP3, WMA, AC-3 … mặc dù âm thanh không còn nguyên vẹn, nhưng chất lượng cũng tương đối tốt, và đặc biệt chỉ chiếm khoảng 5MB cho một bài hát khoảng 5 phút (bằng 10% so với chuẩn WAV).

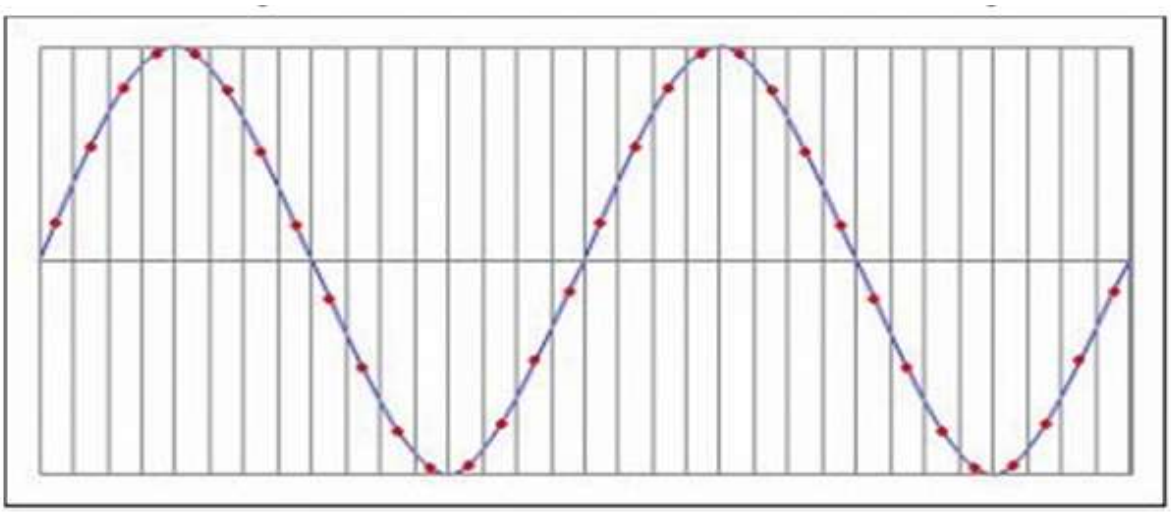
Trong phạm vi của ứng dụng này, vì một số hạn chế, ứng dụng chỉ hỗ trợ file WAV mà thôi. Trong tương lai, nhóm phát triển sẽ tiếp tục mỡ rộng để hỗ trợ các chuẩn nén thông dụng hiện nay như MP3, WMA…

Hình 1.4: Định dạng tệp wave

Trích mẫu

Ta đã biết, sóng âm là một loạt các dao động của vật chất trong môi trường truyền âm. Vào mỗi một thời điểm nhất định, vật chất có một biên độ nhất định. Tưởng tượng như trong dao động của con lắc đồng hồ. Mỗi lúc con lắc có một độ lệch nào đó so với vị trí cân bằng ở chính giữa.

Khi muốn đưa âm thanh vào mã hóa trên máy tính, thì ta cần phải xác định biên độ dao động của sóng âm vào các thời điểm khác nhau. Việc đó gọi là trích/lấy mẫu. Trong một giây phát ra âm thanh, người ta trích lấy một số mẫu biên độ mà đưa vào dữ liệu. Con số ấy gọi là tần số trích mẫu (Sample rate). Thí dụ, ta mở phần thuộc tính của một tập tin âm thanh và thấy đề Sample rate = 44.1 kHz (44100 Hz) nghĩa là trong một giây trích mẫu 44100 lần. Một điều dễ nhận thấy là tần số trích mẫu càng cao thì âm thanh được mã hóa càng chính xác.



Hình 1.5: Trích mẫu

Hình 1.5 minh họa việc trích mẫu. Đường màu xanh là đồ thị sóng âm. Mỗi khung (lát cắt) màu xám là một mẫu âm thanh. Các chấm đỏ là các giá trị (biên độ dao động) sẽ được trích lấy.

Chiều sâu bit

Một yếu tố khác cũng tác động tới độ chính xác của việc trích mẫu âm thanh. Đó là chiều sâu bit, trên bảng thuộc tính tập tin WAV ghi là Audio sample size, thường là 8 hoặc 16 bits. Ta biết, 8 bits là 1 byte, lưu được 256 giá trị (từ 0 tới 255). Còn 16 bits là 2 bytes tức 1 word, lưu được 65536 giá trị (từ -32768 tới 32767). Số bit càng lớn thì âm thanh lấy mẫu càng chính xác. Giống như thang điểm 100 sẽ chấm đúng (chi tiết) hơn thang điểm 10 vậy.

Kênh âm thanh (Channels)

Âm thanh chia ra 2 kênh trái-phải (ở đây không bàn tới âm thanh 4.1, 5.1 hay 7.1 gì đó...) kêu bằng Stereo. Cũng có âm thanh chỉ một kênh (cùng một luồng dữ liệu đi ra hai loa) là Mono.

Đối với âm thanh hai kênh, các mẫu trích (sample) cũng phân ra hai kênh.

Kích thước mẫu trích

Công thức kích thước mẫu trích (tính bằng byte) là:

LengthOfSample = Channels * AudioSampleSize / 8

1.4.3. Cấu trúc tập tin âm thanh

Thuộc họ RIFF, một tập tin WAV được chia ra thành các phần gọi là chunk.

Phần dưới đây liệt kê các chunk và nội dung của chúng theo thứ tự: vị trí , kích thước và mô tả.

❖ ***RIFF chunk***

00 - 03	4	Chuỗi "RIFF".
04 - 07	4	Số byte theo sau con số này, tức là *kích thước tập tin - 8*.
08 - 11	4	Chuỗi "WAVE".

❖ ***FORMAT chunk***

12 - 15 | 4 | Chuỗi "fmt " (ký tự cuối là dấu khoảng trắng, mã ASCII 32).
16 - 19 | 4 | Kích thước FORMAT chunk, mặc nhiên là 16.
20 - 21 | 2 | Định dạng mã hóa âm thanh, thường là 1 (PCM).
22 - 23 | 2 | Số kênh, 1 (Mono) hay 2 (Stereo).
24 - 27 | 4 | Tần số trích mẫu, tính bằng Hz (mẫu/giây).
28 - 31 | 4 | Số byte dữ liệu mỗi giây.

BytesPerSecond = SampleRate * Channels * AudioSampleSize / 8.

32 - 33 | 2 | Số byte trong một mẫu trích. **BytesPerSample = Channels * AudioSampleSize / 8.**
34 - 35 | 2 | Chiều sâu bit (AudioSampleSize), là 8 hoặc 16.

❖ ***DATA chuck***

36 - 39 | 4 | Chuỗi "data".
40 - 43 | 4 | Kích thước dữ liệu âm thanh. **DataSize = Samples * Channels * AudioSampleSize / 8.** *Samples là tổng số mẫu trích (ThờiLượng_Giây * TầnSốTrichMẫu).*
44 - cho tới hết || Dữ liệu âm thanh.

Nếu âm thanh là 2 kênh và 16 bits thì mỗi mẫu trích có kích thước 4 bytes: 2 bytes đầu cho kênh trái, 2 bytes sau cho kênh phải. Nếu 2 kênh và 8 bits thì mỗi kênh 1 byte, mẫu trích 2 bytes. Nếu 1 kênh và 16 bits thì mẫu trích cũng 2 bytes. Nếu chiều sâu bit là 8 thì kiểu dữ liệu là BYTE không dấu (0 tới 255). Nếu là 16 thì kiểu dữ liệu là WORD (2bytes) có dấu (-32768 tới 32767).

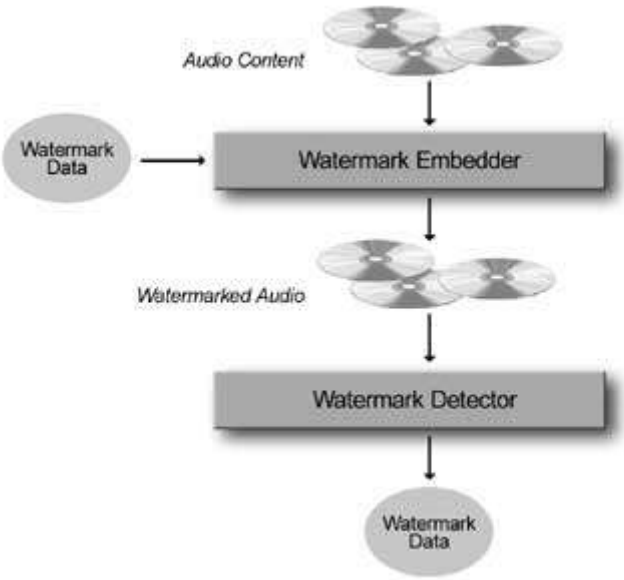
❖ ***Thứ tự Byte***

Dữ liệu trên tập tin WAVE tuân theo quy tắc Little-endian.

Thí dụ, ta có tần số trích mẫu bằng 44100 Hz. Giá trị 4 bytes, hệ thập lục phân của nó là 0x0000AC44 (mỗi cặp số là một byte). Trên tập tin, con số này sẽ được lưu theo thứ tự: 44 AC 00 00. Như vậy gọi là Little-endian. (Kiểu sắp xếp bình thường gọi là Big-endian.)

1.5. Sơ đồ giấu tin và tách tin trong dữ liệu âm thanh

Sơ đồ tổng quát quá trình giấu tin và tách tin trong audio được mô tả tóm tắt theo hình 1.6.



Hình 1.6. Sơ đồ tổng quát giấu tin và tách tin trong audio

Trong đó:

- Audio content: là dữ liệu âm thanh số đầu vào ban đầu dùng để bao thông tin
- Watermark Data: là dữ liệu quan trọng cần giấu trong các dữ liệu âm thanh
- Watermark Embedder: bộ chương trình giấu tin
- Watermarked Audio: là dữ liệu âm thanh số đã giấu tin
- Watermark Detector: bộ chương trình tách thông tin đã giấu trong audio

Watermark Data đem giấu có thể là một đoạn văn bản, một ảnh logo, một đoạn mã ID định danh nào đó liên quan đến bản dữ liệu số che giấu nó hoặc có thể lại là một đoạn âm thanh số ngắn nào đó ... trong trường hợp muốn sử dụng dữ liệu âm thanh số để giấu tin mật ta có thể giấu một văn bản mật (vài trang giấy đến vài chục trang giấu) để trao đổi mật với đối tác đồng minh mà không muốn đối tác thứ ba can thiệp vào.

1.6. Tăng độ an toàn cho thông tin đem giấu

Mặc dù giấu tin trong audio là vô hình so với các phương pháp an toàn bảo mật khác. Tuy nhiên trong trường hợp nghi ngờ người ta có thể sử dụng phương pháp thông kê tần suất hoặc quan sát quy luật tự nhiên của các tín hiệu audio để tách ra được thông tin đã giấu. Trong trường hợp này để tăng tính an toàn cho thông tin đem giấu người ta có thể sử dụng phương pháp mã hóa cho thông tin mật trước khi giấu sử

dụng các kỹ thuật mã hóa như RSA, Elgama, AES hoặc DES... Cũng có thể sử dụng phương pháp mã hóa đơn giản đó là chuyển thông tin giấu sang chuỗi bit nhị phần, sau đó sự dụng phép toán XOR của chuỗi bit thông điệp với chuỗi khóa.

1.7. Đánh giá chất lượng âm thanh sau khi giấu tin

Để đánh giá chất lượng của tín hiệu âm thanh ở đầu ra của bộ mã hoá, người ta thường sử dụng hai tham số: Sai số bình phương trung bình – MSE (*Mean Square Error*) và phương pháp hệ số tỷ lệ tín hiệu / tín hiệu nhiễu PSNR (*Peak Signal to Noise Ratio*).

MSE giữa tín hiệu gốc và tín hiệu đã giấu tin được tính như sau:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2$$

Ở đây:

x_i biểu thị giá trị tín hiệu gốc

y_i biểu thị giá trị tín hiệu đã bị biến đổi

N là độ dài của tín hiệu âm thanh.

PSNR, đơn vị: *deciben* (dB), thường được sử dụng trong xử lý tín hiệu số:

$$\text{PSNR} = 10*\log10\left(\frac{\max(x_i)^2}{MSE}\right)$$

Thông thường, nếu PSNR > 35dB thì hệ thống mắt người gần như không phân biệt được sự khác biệt giữa tín hiệu gốc và tín hiệu bị biến đổi để giấu tin. PSNR càng cao thì chất lượng của tín hiệu càng ít bị thay đổi so với gốc. Khi hai tín hiệu giống hệt nhau, MSE sẽ bằng 0 và PSNR đi đến vô hạn.

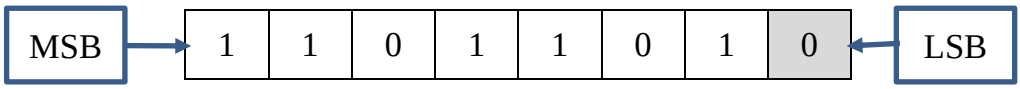
Chương II. KỸ THUẬT GIẤU TIN TRONG ÂM THANH

Trong chương này trình bày một số kỹ thuật giấu tin mà tác giả đã tìm hiểu bao gồm kỹ thuật giấu tin trên bit có trọng số thấp nhất LSB (Least Significant Bit) của tín hiệu audio và kỹ thuật thủy vân trên miền biến đổi tần số cho phép giấu văn bản, ảnh nhị phân, ảnh xám hay một đoạn audio nào đó. Trong chương này nhóm tác giả cũng đề xuất một phương pháp giấu tin văn bản trên tín hiệu audio dựa vào sự kết hợp của 7 bit MSB (Most significant bit) và 4 bit LSB, nhưng chỉ làm thay đổi 1 đến 2 bit LSB của tín hiệu gốc không gây ra quá nhiều thay đổi giá trị của tín hiệu gốc.

2.1. Kỹ thuật giấu tin trên bit có trọng số thấp LSB (Least Significant Bit)

Phương pháp giấu tin trên bit có trọng số thấp hay còn gọi là phương pháp mã hóa LSB (Least Significant Bit) là phương pháp nhúng bit thông tin vào các bit có trọng số thấp của dữ liệu audio.

Giả sử một tín hiệu A = 218 ứng với giá trị 8 bit như hình 2.1, khi đó bit bên trái nhất (có giá trị 1) được gọi là bit có trọng số lớn nhất MSB (Most significant bit), và bit bên phải nhất (có giá trị là 0) được gọi là bit có trọng số thấp nhất LSB (Least significant bit).



Hình 2.1. Giá trị 8 bit của tín hiệu A = 218

Với bit có trọng số thấp nhất khi thay đổi giá trị từ 0 sang 1 hoặc từ 1 sang 0 sẽ không làm thay đổi nhiều giá trị gốc ban đầu. Do vậy khi nhúng thông tin mật vào tín hiệu audio người ta thường nhúng vào bit có trọng số thấp nhất này để không làm ảnh hưởng đến trực giác của người nghe.

Giả sử muốn giấu một bit có giá trị 1 vào tín hiệu A thì ta được tín hiệu mới B = 219 ứng với giá trị 8 bit như hình 2.2. Ngược lại muốn giấu một bit có giá trị 0 vào tín hiệu A thì ta được tín hiệu mới C =218 (giữ nguyên giá trị ban đầu) ứng với 8 bit như hình 2.3, điều này trong giấu tin được hiểu là ngầm định đã giấu.

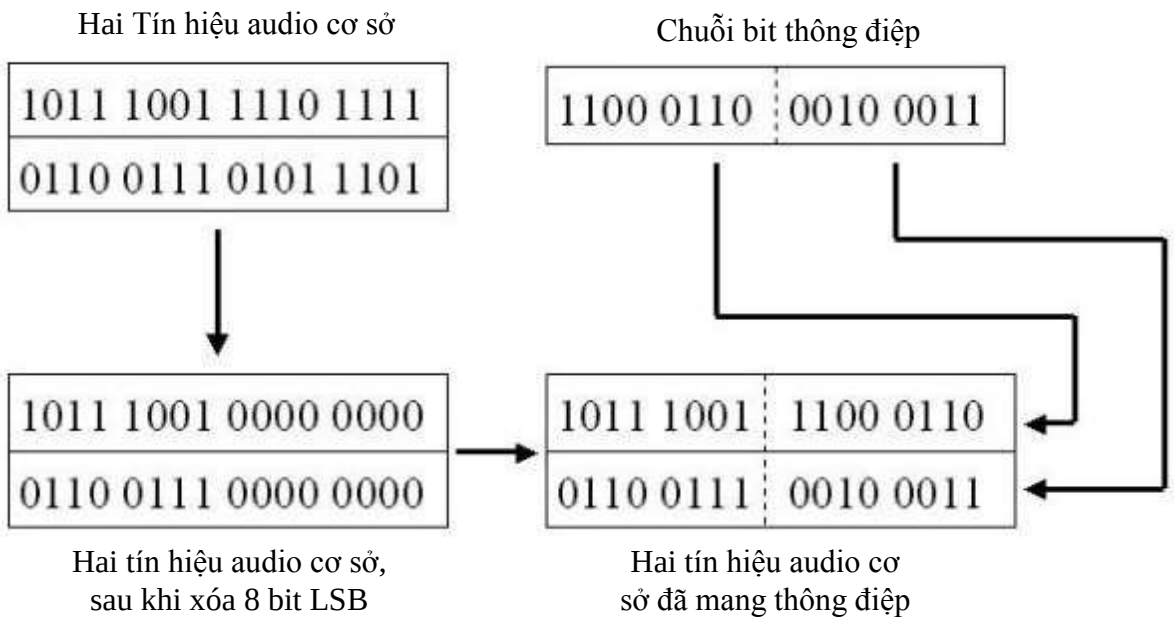
1	1	0	1	1	0	1	1

Hình 2.2. Giá trị 8 bit của tín hiệu B=219 sau khi giấu bit 1 vào LSB của A

1	1	0	1	1	0	1	0

Hình 2.3. Giá trị 8 bit của tín hiệu C=219 sau khi giấu bit 0 vào LSB của A

Trong trường hợp tín hiệu audio được lấy mẫu với tần số lấy mẫu là 44100Hz thì tín hiệu audio có thể biểu diễn dưới dạng 16 bit, khi đó người ta có thể giấu thông tin đến 8 bit có trọng số thấp thay vì 1 bit có trọng số thấp như sơ đồ trong hình 2.4 [9].



Hình 2.4. Sơ đồ giấu tin trên 8 bit LSB của tín hiệu audio cơ sở

Với trường hợp này có thể giải thích vì sao nó vẫn có thể được sử dụng phổ biến trong trao đổi thông tin mật vì do thị giác của con người khó phát hiện ra sự khác biệt với các tín hiệu có tần số âm thanh thấp (nhỏ hơn <250Hz) [6-8], do vậy người ta có thể lựa chọn giấu trên nhiều bit LSB để có thể trao đổi nhiều thông tin mật trong một tệp audio nào đó trong quá trình truyền đi.

Để tăng độ an toàn cho quá trình giấu tin và tách tin người ta có thể chọn số lượng bit LSB dùng để giấu tin sao cho phù hợp nhất mà không ảnh hưởng đến chất lượng âm thanh ban đầu.

Thuật toán giấu tin và tách tin trên k bit LSB của tín hiệu thời gian thực (tín hiệu cơ sở) trình bày chi tiết dưới đây.

2.1.1. Thuật toán giấu tin:

Đầu vào: Audio gốc A có độ dài tín hiệu L, chuỗi tin cần giấu M.

Đầu ra: Audio đã giấu tin.

Các bước thực hiện:

Bước 1: Đọc audio vào A, dựa vào tần số lấy mẫu và các thông số liên quan đến cấu trúc lưu trữ của tệp audio ta được vector giá trị của tín hiệu mẫu lưu vào mảng một chiều để thực hiện giấu tin.

Bước 2: Thực hiện chuyển đổi chuỗi tin cần giấu M sang chuỗi bit nhị phân để có thể giấu vào audio, tính độ dài số bit thông điệp lưu vào L.

Bước 3: Chọn giá trị k phù hợp nhất (tức là chọn số bit LSB của tín hiệu audio sẽ giấu tin)

Bước 4. Dựa vào k được chọn ở bước 3, thực hiện giấu L (độ dài bit thông điệp) vào LSB của ba tín hiệu đầu tiên hoặc cuối cùng của tín hiệu audio để phục vụ tách tin.

Bước 5: Dựa vào k đã chọn và độ dài L của thông điệp ta thực hiện chia chuỗi bit thông điệp thành các chuỗi con có độ dài k bit. Mỗi chuỗi con này sẽ được thay thế vào k bit LSB của L/k tín hiệu audio để có thể giấu đủ L bit thông điệp.

Bước 6: Lưu lại các tín hiệu audio vào tệp audio kết quả ta được audio đã giấu tin **S**.

2.1.2. Thuật toán tách tin:

Đầu vào: Audio đã giấu tin S.

Đầu ra: Thông điệp đã giấu M.

Các bước thực hiện:

Bước 1: Đọc audio vào S, dựa vào tần số lấy mẫu và các thông số liên quan đến cấu trúc lưu trữ của tệp audio ta được vector giá trị của tín hiệu mẫu lưu vào mảng một chiều để thực hiện tách tin.

Bước 2: Cho biết giá trị k (số bit LSB đã giấu tin).

Bước 3: Tách ra độ dài bit L đã giấu trên ba tín hiệu đầu tiên hoặc cuối cùng của tín hiệu audio.

Bước 4: Thực hiện tách k bit LSB của L/k tín hiệu đã giấu tin ghép lại thành chuỗi bit, ta được chuỗi bit đã giấu.

Bước 5: Chuyển đổi chuỗi bit đã tách về dạng ban đầu ta được thông điệp cần tách.

Thông điệp ban đầu cần giấu có thể là văn bản, dữ liệu ảnh hoặc là một đoạn audio nào đó. Trong hình 2.4 là một minh họa cho trường hợp số bit LSB của tin hiệu audio dùng để giấu tin ứng với k=8.

2.2. Kỹ thuật giấu tin ảnh xám trên tín hiệu audio (LSB nâng cao)

Đây là kỹ thuật được tác giả Pradeep Kumar Singh và cộng sự đề xuất năm 2010 dựa trên phương pháp cải tiến phương pháp giấu trên LSB của tín hiệu âm thanh [10]. Phương pháp này áp dụng để giấu ảnh xám trên tín hiệu audio bằng cách sử dụng 7 bit quan trọng nhất MSB của các tín hiệu audio để giấu các điểm ảnh, quét lần lượt các tín hiệu audio so sánh chuỗi bit nhị phân của điểm ảnh với 7 bit MSB của tín hiệu audio:

+ Nếu giá trị một điểm ảnh cần giấu (chuyển sang chuỗi nhị phân) trùng với 7 bit quan trọng nhất MSB của tín hiệu audio thì viết vị trí tương ứng vào 3 bit LSB của tín hiệu audio là 1 để đánh dấu đã giấu một điểm ảnh vào tín hiệu này.

+ Nếu giá trị điểm ảnh cần giấu không trùng với 7 bit MSB thì thay thế 3 bit LSB của tín hiệu audio đều bằng 0 để đánh dấu không giấu điểm ảnh ở tín hiệu này.

Quá trình giấu điểm ảnh sẽ thực hiện tiếp cho các điểm ảnh tiếp theo trên các tín hiệu audio cho đến khi giấu hết các điểm ảnh vào audio. Nếu không quét đến hết audio mà vẫn còn điểm ảnh chưa giấu thì thông báo ảnh quá lớn để giấu vào audio.

Dưới đây trình bày chi tiết thuật toán giấu tin và tách tin.

2.2.1. Thuật toán giấu tin:

Đầu vào: Audio gốc A có độ dài tín hiệu L, ảnh xám cần giấu I.

Đầu ra: Audio đã giấu tin.

Các bước thực hiện:

Bước 1: Đọc ảnh cần giấu và chuyển đổi giá trị các điểm ảnh sang chuỗi bit nhị phân. Lưu số hàng và số cột của ảnh theo phương pháp nào đó để phục vụ khi khôi phục lại ảnh (có thể sử dụng các tín hiệu đầu tiên hoặc cuối cùng của audio để giấu giá trị số hàng và số cột của ảnh).

Bước 2: Đọc audio dùng để che giấu ảnh.

Bước 3: Chọn các tín hiệu audio và so sánh 7 bit MSB với 7 bit điểm ảnh cần giấu, nếu trùng thì đánh giấu vào 3 bit LSB của tín hiệu audio bằng 1 và chuyển sang giấu điểm ảnh tiếp theo, nếu không trùng thì đánh giấu vào 3 bit tín hiệu audio bằng 0 (nghĩa là không thể giấu vào) chuyển sang tín hiệu tiếp theo đề tìm tín hiệu thảo mãn.

Bước 4: Lặp lại bước 3 cho đến khi giấu hết các điểm ảnh, nếu sau khi duyệt hết các tín hiệu mà vẫn còn điểm ảnh chưa giấu thì thông báo "ảnh không thể giấu vào audio này" ngược lại thông báo "Giấu thành công".

Bước 5: Nếu quá trình "Giấu thành công" lưu tín hiệu audio vào tệp audio ta được audio đã giấu tin.

2.2.2. Thuật toán tách tin:

Đầu vào: Audio đã giấu tin S.

Đầu ra: Ảnh đã giấu I.

Các bước thực hiện:

Bước 1: Đọc audio vào S, tách ra giá trị số hàng và số côt đã giấu vào audio theo bước 1 của quá trình giấu tin.

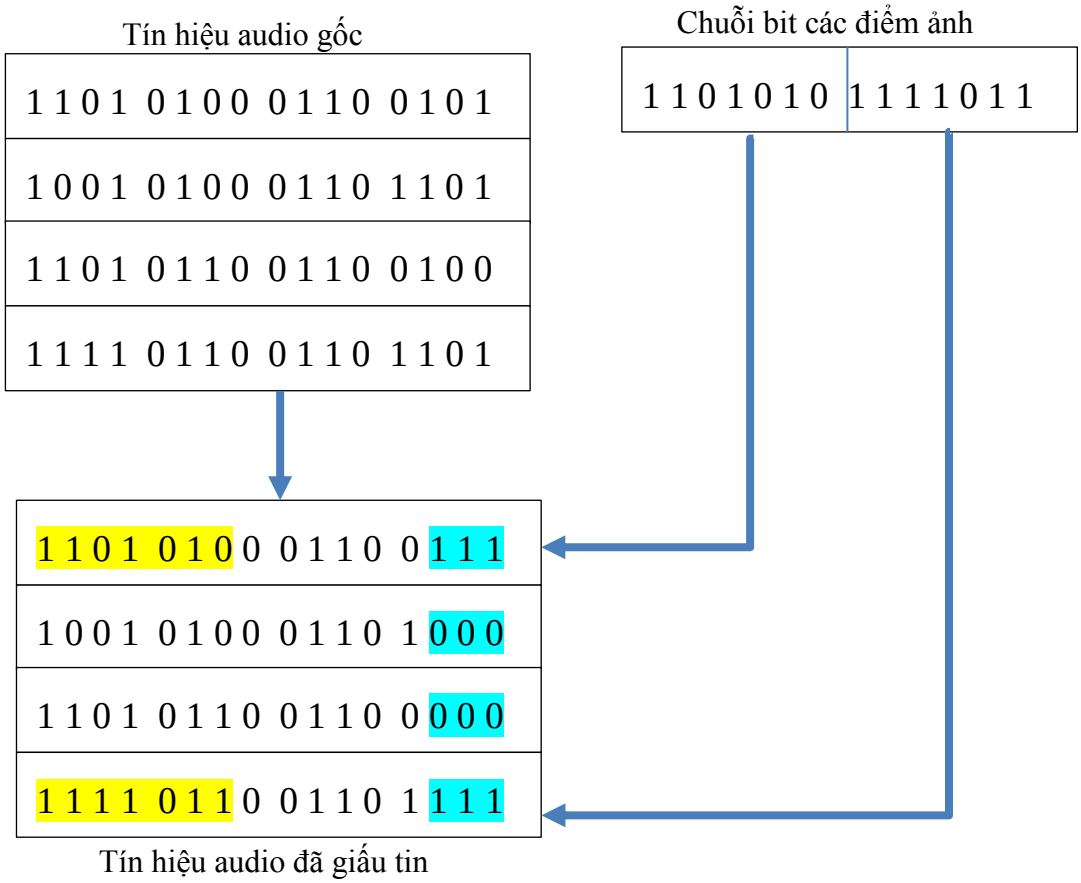
Bước 2: Kiểm tra lần lượt các tín hiệu audio, nếu 3 bit LSB của tín hiệu đều khác 0 thì tách ra 7 bit MSB của tín hiệu ta được điểm ảnh đã giấu.

Bước 3: Thực hiện lặp lại bước 2 cho đến khi tách hết các điểm ảnh đã giấu.

Bước 4: Dựa vào số hàng và số cột của ảnh ta có thể khôi phục lại được ảnh ban đầu đã giấu vào trong audio.

Kỹ thuật giấu tin và tách tin ảnh giấu trên audio có thể minh họa theo hình 2.5. sau đây.

Theo nhận định của nhóm nghiên cứu chúng tôi thì với thuật toán LSB nâng cao này của Pradeep Kumar Singh và cộng sự chúng ta có thể giấu các điểm ảnh xám vào các tín hiệu audio nhưng chỉ làm thay đổi đến 3 bit LSB của tín hiệu. Tuy nhiên phụ thuộc vào ảnh cần giấu và tín hiệu audio có độ dài khác nhau mà không phải ảnh nào cũng có thể giấu vào trong audio.



Hình 2.5. Minh họa kỹ thuật giấu ảnh trên tín hiệu audio

2.3. Đề xuất phương pháp giấu văn bản kết hợp MSB và LSB của tín hiệu audio

Dựa vào kỹ thuật giấu ảnh xám của tác giả Pradeep Kumar Singh và cộng sự đề xuất năm 2010 trình bày trong mục 2.2. Tác giả của đề tài nghiên cứu này thấy rằng có thể cải tiến áp dụng cho phương pháp giấu thông tin văn bản vào tin hiệu audio bằng cách kết hợp MSB và LSB chi tiết như sau đây.

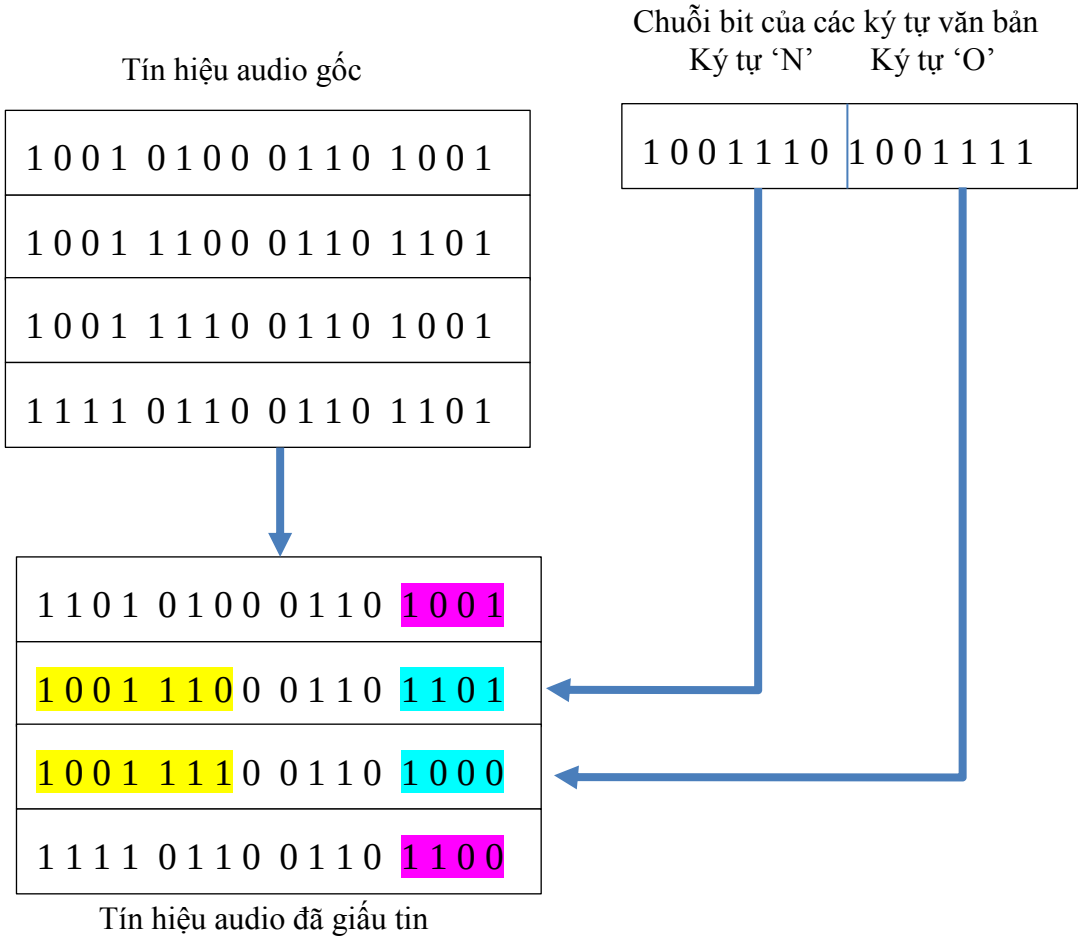
Như chúng ta đã biết thông tin văn bản là dạng thông tin có nội dung toàn văn bản ký tự và số đơn thuần, trong nội dung không chứa các thông tin bảng biểu hay đồ hoặc... với những thông tin này khi lưu trữ trong máy tính chúng ta có thể lưu trong các tệp text của NotePad (*.txt) hoặc WordPad (*.rtf). Mỗi ký tự lưu trong tệp ứng với một mã ASCII có giá trị là 1 byte (giá trị từ 0 đến 255) nhưng chỉ có 128 ký tự đầu là hay dùng, còn lại là các ký tự mở rộng. Các ký tự có mã từ 0 đến 31 gọi là các ký tự điều khiển, không in ra được, được dùng để điều khiển các thiết bị ngoại vi, chẳng hạn ký tự có mã là 7 dùng để tạo một tiếng kêu bip, ký tự có mã là 13 dùng để chuyển con trỏ màn hình xuống đầu dòng dưới…do đó để truyền thông tin mật dưới dạng văn bản đặc biệt văn bản tiếng Anh ta chỉ cần sử dụng các ký tự từ 32 đến 127 là có thể biểu diễn đủ nội dung cần trao đổi.

2.3.1. Phương pháp 1

Với nhận xét trên tác giả xin đề xuất phương pháp giấu văn bản trong tín hiệu audio bằng cách sử dụng 7 bit/tín hiệu để ẩn dữ liệu mà ít có sự thay đổi của tín hiệu audio gốc. Cụ thể đoạn văn bản được chuyển đổi thành mã ASCII, sử dụng 7 bit mã ASCII của mỗi tin so sánh với 7 bit cao MSB của các tín hiệu gốc, nếu gặp tín hiệu nào trùng khớp, ta sẽ đánh dấu sự có mặt của dữ liệu trong một tín hiệu cụ thể bằng cách sử dụng 4 bit LSB, theo nguyên tắc điều chỉnh số bit 1 trong 4 bit LSB là số lẻ. Trong trường hợp 7 bit cao MSB của tín hiệu gốc không trùng khớp với 7 bit của tin giấu ta sẽ điều chỉnh số bit 1 trong 4 bit LSB của tín hiệu là số chẵn để đánh dấu tín hiệu này không giấu tin. Quá trình giấu sẽ lặp lại cho đến khi giấu hết các ký tự của thông điệp.

Quá trình tách tin được thực hiện bằng cách kiểm tra 4 bit LSB của từng tín hiệu audio, nếu số bit 1 của 4 bit này là lẻ thì tách ra 7 bit MSB được ký tự đã giấu, ngược lại nếu số bit 1 là chẵn thì bỏ qua và thực hiện kiểm tra tiếp các tín hiệu tiếp theo.

Hình 2.6 dưới đây minh họa phương pháp giấu ký tự ‘N’ và ‘O’ trong tín hiệu audio bằng cách so khớp 7 bit MSB và điều chỉnh 4 bit LSB khi giấu vào tín hiệu.

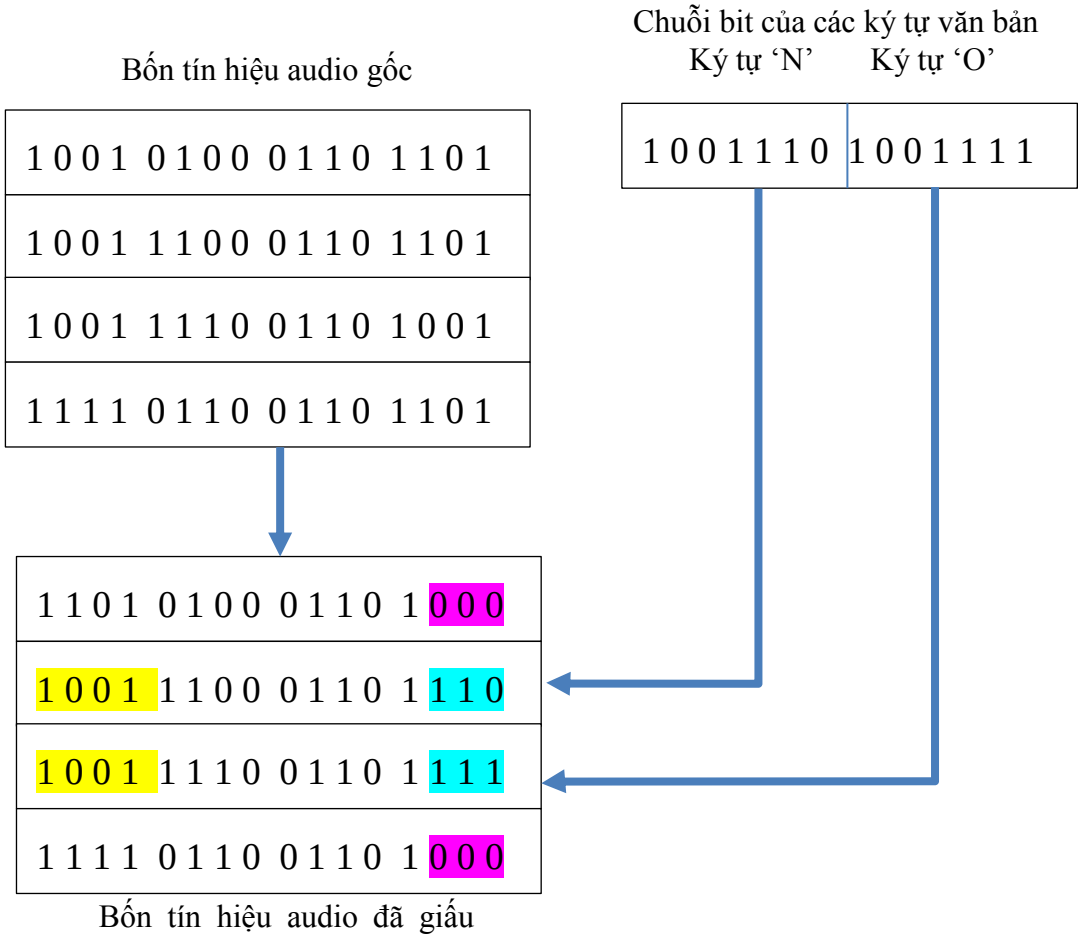


Hình 2.6. Minh họa kỹ thuật giấu văn bản dựa vào7 bit MSB và 4 bit LSB của tín hiệu audio

Với phương pháp giấu tin này chúng ta không làm thay đổi quá nhiều số bit LSB như kỹ thuật giấu ảnh cấp xám trong audio đã trình bày trong 2.2, với phương pháp giấu đó có thể dễ dàng bị tấn công bằng phương pháp thống kê LSB. Ở đây khi cài đặt chúng ta có thể điều chỉnh sử dụng 2 hoặc 4 bit LSB để kiểm tra tính chẵn lẻ của bit 1 trong chuỗi.

2.3.2. Phương pháp 2

Phương án khác trong trường hợp số lượng ký tự cần giấu vào audio quá nhiều trong khi số các tín hiệu có 7 bit MSB trùng khớp với 7bit của ký tự không đủ thì chúng ta có thể sử dụng phương án chỉ so khớp 4 bit của ký tự với 4 bit MSB của tín hiệu audio nếu trùng khớp ta thay thế 3 bit cuối của tín hiệu bằng 3 bit cuối của ký tự. Còn trong trường hợp không trùng khớp ta điều chỉnh 3 bit LSB của tín hiệu đều bằng 0 để đánh dấu tín hiệu này không giấu thông tin. Do đó trong trường hợp tách tin chúng ta chỉ việc kiểm tra 3 bit LSB nếu đều bằng 0 là không giấu tin, ngược lại ta tách tin giấu bằng cách lấy 4 bit MSB ghép với 3 bit LSB được ký tự đã giấu.



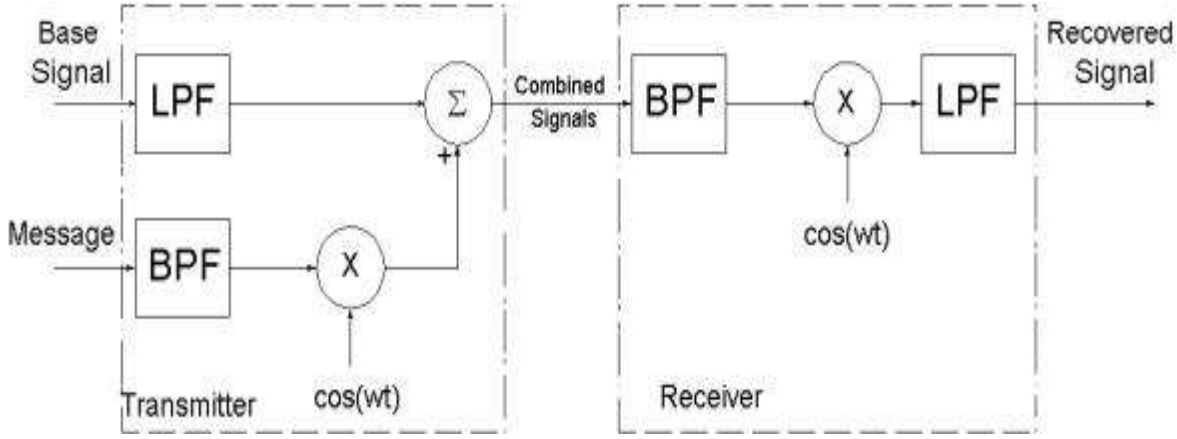
Hình 2.7. Minh họa kỹ thuật giấu văn bản dựa trên 4 bit MSB và 3 bit LSB của tín hiệu audio

Tùy vào trường hợp với số lượng ký tự giấu cụ thể mà chúng ta có thể giấu dựa vào 4 đến 7 bit MSB để có thể giấu thông tin văn bản. Thuật toán giấu tin tương tự như đã trình bày trong mục 2.2 nên ở mục này tác giả không trình bày chi tiết thuật toán giấu tin và tách tin.

2.4. Kỹ thuật thủy vân số trên miền biến đổi của tín hiệu audio

Đây là kỹ thuật do nhóm tác giả C.J. Ganier, Randall Holman, Julie Rosser và Erik Swanson đề xuất. Theo nhóm tác giả này thì hệ thống thính giác của người trưởng thành có thể phân biệt được tín hiệu âm thanh trong khoảng từ 4kHz đến 18kHz, các tín hiệu thấp nhỏ hơn 4kHz tai người khó có thể nhận biết. Trong thực tế các tín hiệu thấp nhỏ hơn 4kHz không "được sử dụng", do đó nhóm tác giả sử dụng tín hiệu này để nhúng thủy vân [9].

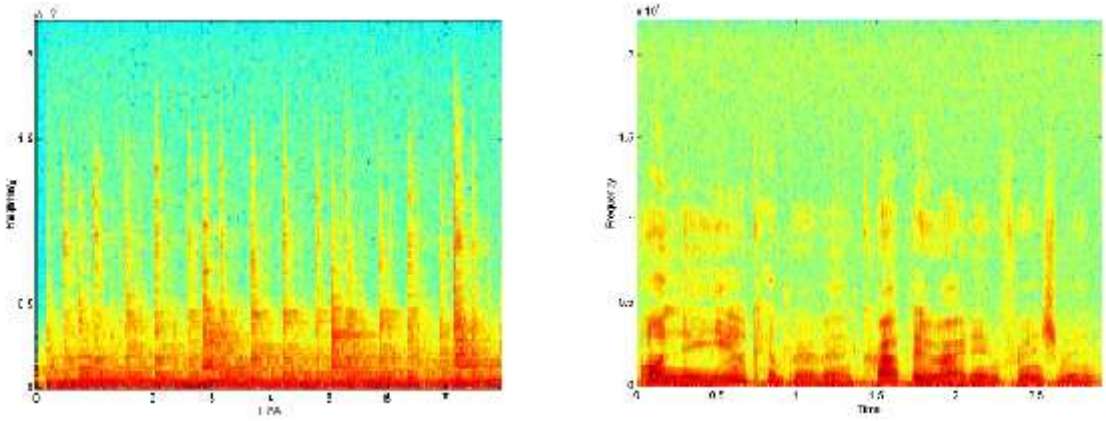
Kỹ thuật thủy vân số trên miền biến đổi là phương pháp thay vì giấu trực tiếp trên các tín hiệu miền thời gian thực thì các tín hiệu sẽ được biến đổi sang miền tần số sau đó mới chèn thông tin cần giấu vào. Tổng quát của phương pháp có thể mô tả qua sơ đồ trong hình 2.8.



Hình 2.8. Sơ đồ tổng quát giấu tin và tách tin trên miền tần số [9]

Theo sơ đồ 2.8 thì tín hiệu vào dùng để che giấu thông tin gọi là *tín hiệu cơ sở* (Base Signals), tín hiệu sau khi đã giấu thông tin gọi là tín hiệu mang tin (Combined Signals). Bộ LPF (low pass filter) là bộ lọc thông thấp, là bộ lọc chỉ cho thành phần tần số thấp hơn tần số cắt đi qua, thành phần tần số cao thì bị loại bỏ, BPF (Band pass filter) là bộ lọc thông dải, là bộ lọc chỉ cho các thành phần có tần số trong một dải đi qua thôi, các thành phần bé hơn và lớn hơn thì loại bỏ. Bộ cos(wt) là bộ điều chỉnh tín hiệu sử dụng phép toán cosine. Bộ Σ là bộ kết hợp tín hiệu cơ sở với tín hiệu thông điệp.

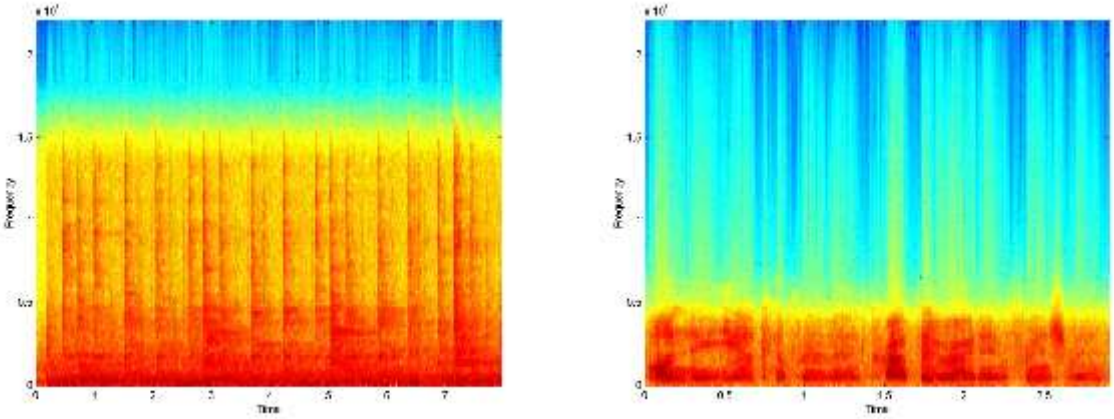
Quy trình giấu tin được mô tả tóm tắt như sau: giả sử có tín hiệu vào để che giấu thông tin (gọi là tín hiệu cơ sở) và thông điệp (giả sử cũng là một đoạn tín hiệu audio nào đó) cần được che giấu. Khi đó biến đổi tần số Fourier của hai tín hiệu này được biểu diễn lần lượt trong hình 2.9 a) và b).



a) b)

Hình 2.9. Biến đổi tần số Fourier của tín hiệu cơ sở và tín hiệu thông điệp

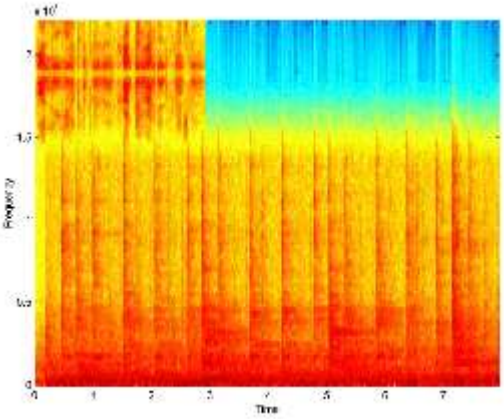
Tiếp theo chúng ta thực hiện lọc sử dụng bộ lọc thông thấp LPF đến 18kHz cho tín hiệu cơ sở, sẽ loại bỏ trên 4kHz (kết quả như hình 2.10 a). Sử dụng bộ lọc thông giải BPF cho tín hiệu thông điệp với giải lọc từ 300Hz đến 3.3 kHz (kết quả như hình 2.10b).



a) b)

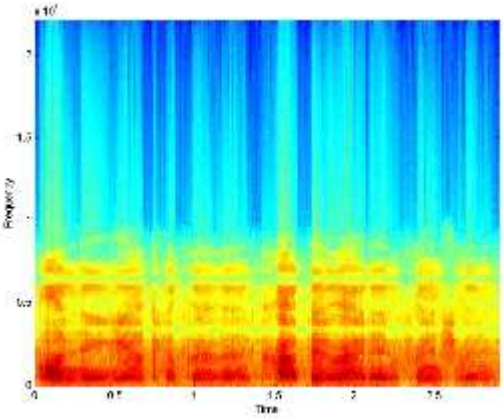
Hình 2.10. a) Tín hiệu cơ sở sử dụng bộ lọc LPF, b) tín hiệu thông điệp sử dụng BPF

Bây giờ chúng ta thực hiện điều chỉnh tín hiệu thông điệp (đã lọc) sử dụng biến đổi cosine với tần số mang 20kHz. Kết hợp tín hiệu cơ sở (đã lọc) với tín hiệu thông điệp sau khi điều chỉnh chúng ta nhận được tín hiệu mới có giấu thông tin (hình 2.11).



Hình 2.11. Tín hiệu đã mang thông tin giấu.

Quá trình tách tin được thực hiện ngược lại, sử dụng lọc thông giải BPF cho tín hiệu đã mang tin để tách ra giải thông của tin hiệu thông điệp đã gộp trong tin hiệu cơ sở. Kết quả sau bộ lọc BPF sẽ được điều chỉnh sử dụng bộ cos(wt) bằng phép cosine ta được tín hiệu sau điều chỉnh, tín hiệu này sẽ được sử dụng bộ lọc thông thấp LPF để được tín hiệu thông điệp đã giấu, kết quả hình 2.12.



Hình 2.12. Tín hiệu thông điệp sau khi tách ra

Thuật toán giấu tin và tách tin trên miền biến đổi như trình bày sau đây.

2.4.1. Thuật toán giấu tin:

Đầu vào: Audio gốc A có độ dài tín hiệu L, audio cần giấu M.

Đầu ra: Audio đã giấu tin.

Các bước thực hiện:

Bước 1: Đọc audio vào A, dựa vào tần số lấy mẫu và các thông số liên quan đến cấu trúc lưu trữ của tệp audio ta được vector giá trị của tín hiệu

mẫu, biến đổi Fourier cho tin hiệu vào (không biến đổi 3 tín hiệu audio cuối cùng để giấu độ dài thông điệp)

Bước 2: Đọc audio cần giấu M, tính độ dài số tín hiệu của M được giá trị L, lưu vào 3 tín hiệu cuối cùng của audio vào A để có thể tách ra M trong quá trình tách tin. Thực hiện biến đổi Fourier cho tin hiệu thông điệp M.

Bước 3: Thực hiện lọc sử dụng bộ lọc thông thấp LPF đến 18kHz cho tín hiệu cơ sở.

Bước 4: Sử dụng bộ lọc thông giải BPF cho tín hiệu thông điệp với giải lọc từ 300Hz đến 3.3 kHz.

Bước 5: thực hiện điều chỉnh tín hiệu thông điệp (đã lọc) sử dụng biến đổi cosine với tần số mang 20kHz. Kết hợp tín hiệu cơ sở (đã lọc) với tín hiệu thông điệp sau khi điều chỉnh chúng ta nhận được tín hiệu mới có giấu thông tin.

Bước 6: Lưu lại các tín hiệu audio đã giấu tin vào tệp audio kết quả ta được audio đã giấu tin **S**.

2.4.2. Thuật toán tách tin:

Đầu vào: Audio đã giấu tin S.

Đầu ra: Audio thông điệp đã giấu M.

Các bước thực hiện:

Bước 1: Đọc audio vào S, dựa vào tần số lấy mẫu và các thông số liên quan đến cấu trúc lưu trữ của tệp audio ta được vector giá trị của tín hiệu để thực hiện tách tin.

Bước 2: Tách ra độ dài tín hiệu L đã giấu trên ba tín hiệu đầu tiên hoặc cuối cùng của tín hiệu audio.

Bước 3: Sử dụng lọc thông giải BPF cho tín hiệu đã mang tin để tách ra giải thông của tin hiệu thông điệp đã giấu.

Bước 4: Điều chỉnh tín hiệu sử dụng bộ cos(wt) bằng phép cosine ta được tín hiệu sau điều chỉnh, tín hiệu này sẽ được sử dụng bộ lọc thông thấp LPF ta được tín hiệu thông điệp đã giấu.

Bước 5: Kết hợp với độ dài tín hiệu L đã giấu ta được audio thông điệp.

Thuật toán trong kỹ thuận này phù hợp với phương pháp thủy vân số dùng để giấu một đoạn audio vào audio cơ sở mục đích bảo vệ bản quyền số đối với dữ liệu âm thanh.

Chương III. CÀI ĐẶT, THỬ NGHIỆM VÀ ĐÁNH GIÁ

3.1. Môi trường cài đặt và một số giao diện chương trình

3.1.1. Môi trường cài đặt

Chương trình giấu tin trong audio được thực hiện trên môi trường Matlap phiên bản 2008b. Môi trường này hỗ trợ nhiều tính năng liên quan đến đọc và lưu dữ liệu audio vì vậy giúp giảm thiểu việc xử lý vào ra đối với các tệp audio.

Hệ điều hành tối thiểu để có thể sử dụng Matlab phiên bản 2008 là hệ điều hành Win7, Window Vista hoặc WinNT.

3.1.2. Giao diện chương trình

Sau đây là một số cửa sổ giao diện chính của chương trình. Cửa số chính của chương trình (hình 3.1) có 2 module chính là: module các kỹ thuật giấu tin mật và module các kỹ thuật thủy vân số. Các hình tiếp theo là ví dụ minh họa chi tiết thực hiện của một số chức năng giấu tin, tách tin cụ thể.

Hình 3.1. Giao diện chính của chương trình

Nhóm chức năng Giấu tin mật (Steganography) gồm:

+ Kỹ thuật giấu tin văn bản: Gồm có chức năng giấu nội dung dạng văn bản có thể gõ trực tiếp trên giao diện hoặc đọc vào từ một tệp *.txt lưu trong máy tính;

+ Kỹ thuật giấu tin một tệp tin bất kỳ: Cho phép giấu thông tin từ một tệp bất kỳ lưu trong máy tính.

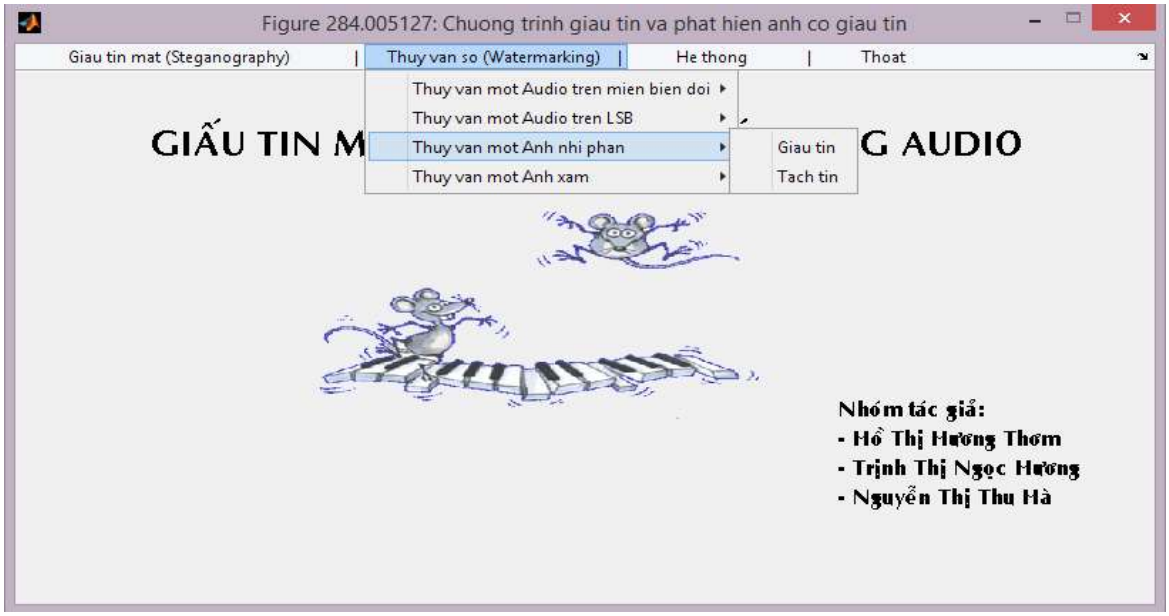
Hình 3.2. Giao diện chính với module chức năng giấu tin mật (Stegangraphy)

Nhóm chức năng Thủy vân số (Watermarking) gồm:

+ Thủy vân một audio trên audio: có thể giấu trên miền biến đổi Fourier hoặc giấu trên LSB của miền thời gian thực.

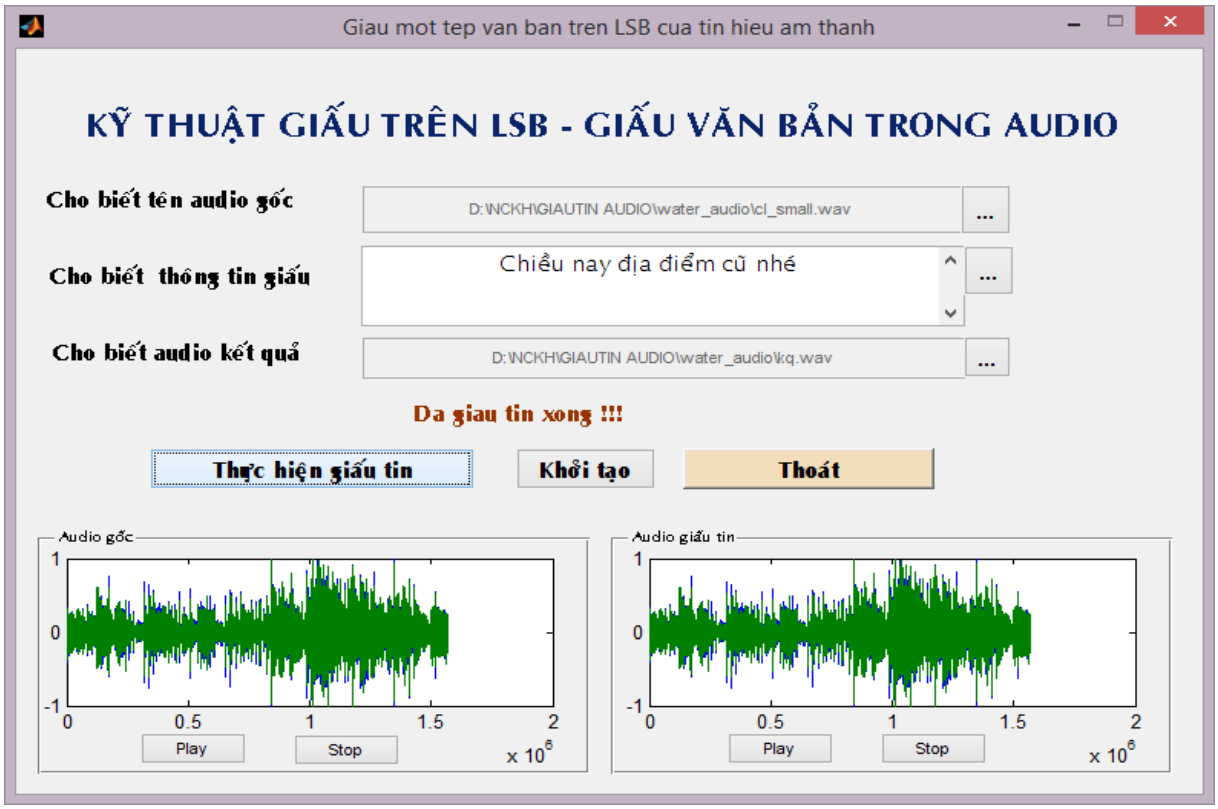
+ Thủy vân một ảnh nhị phân trên audio

+ Thủy vân một ảnh xám trên audio

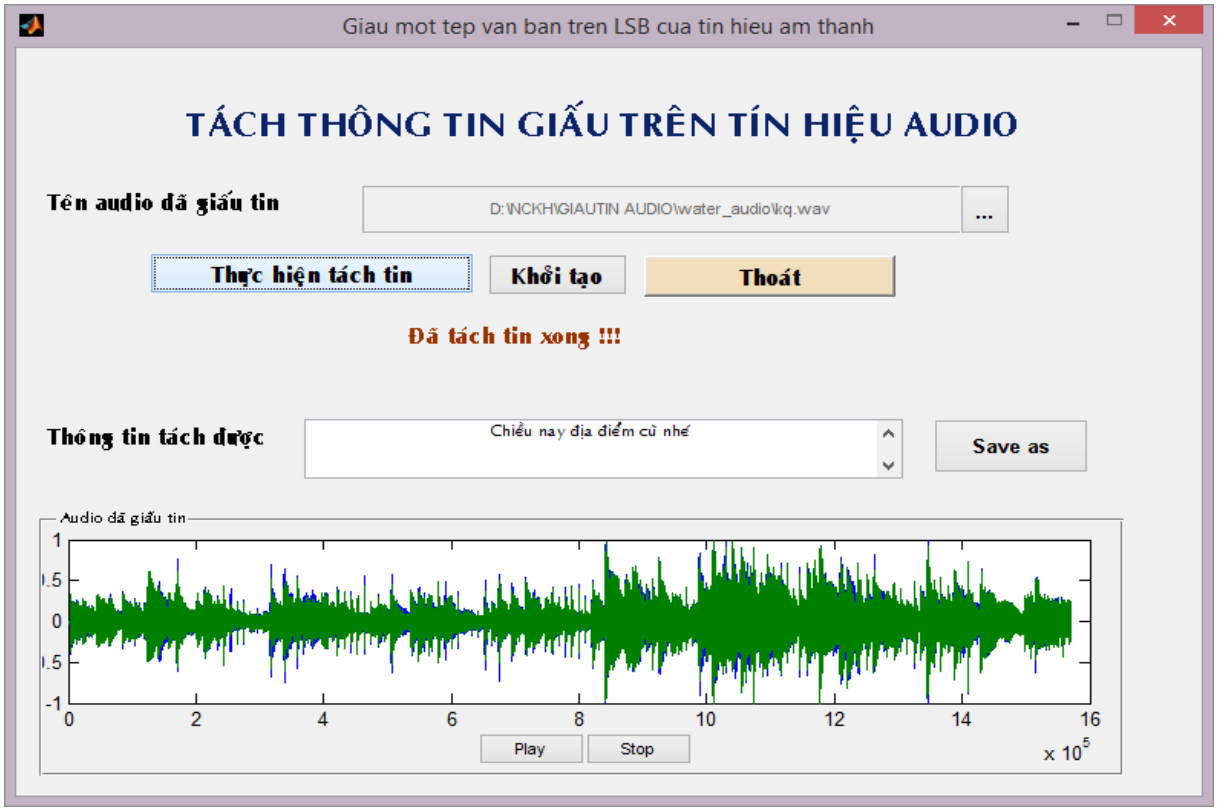


Hình 3.3. Giao diện chính với module chức năng Thủy vân số (Watermarking)

Dưới đây là giao diện giấu tin văn bản và tách tin trên LSB của tín hiệu audio.

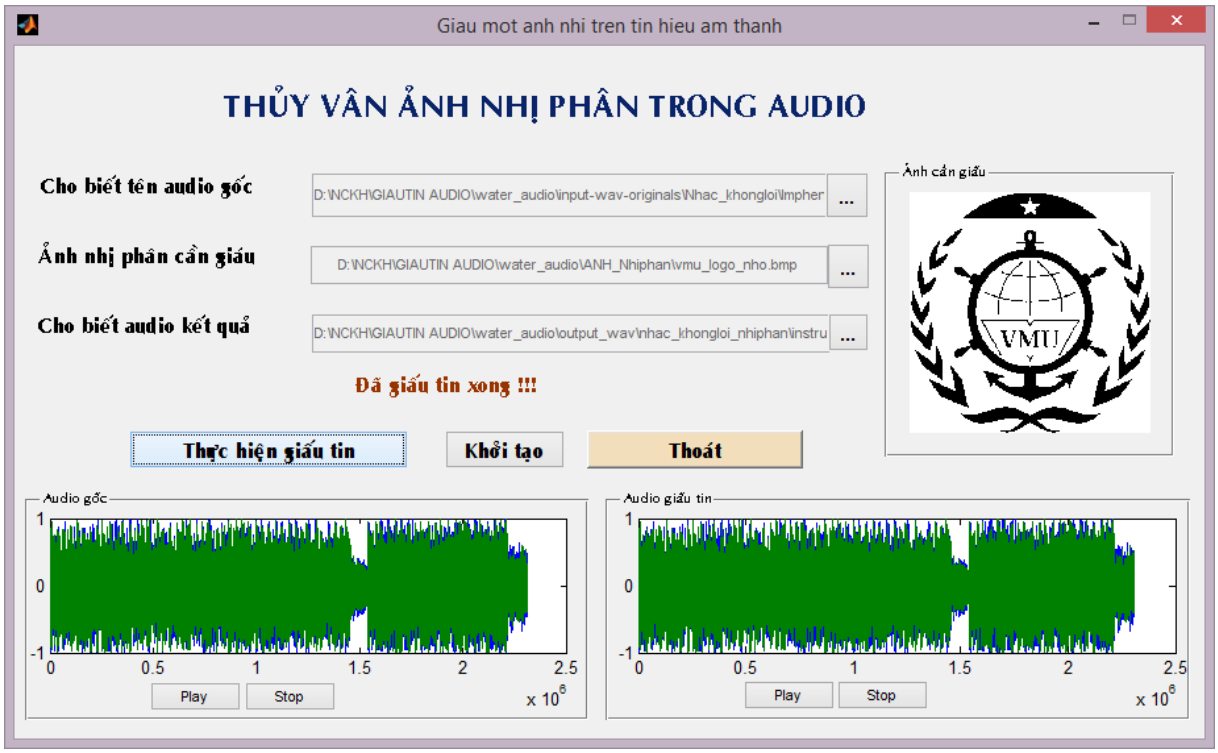


Hình 3.4. Giao diện giấu tin văn bản trên LSB của Audio



Hình 3.5. Giao diện tách tin văn bản đã giấu trên LSB của Audio

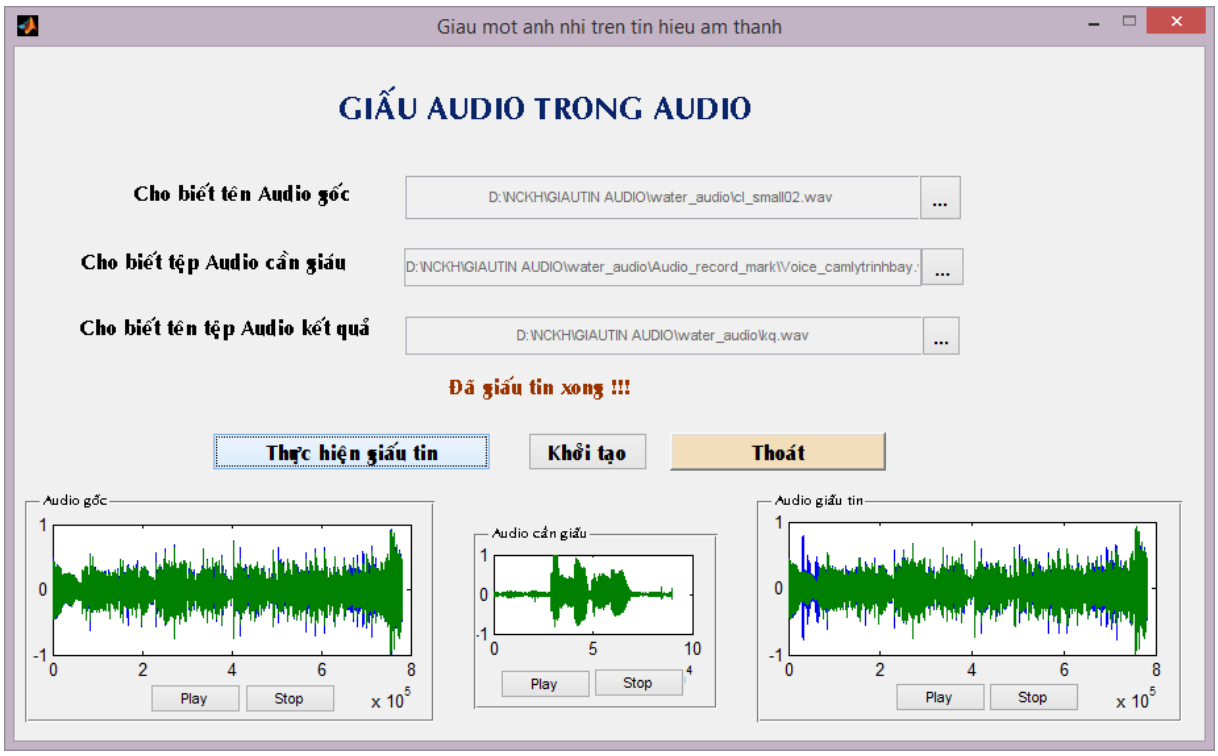
Dưới đây là giao diện thủy vân và tách thủy vân trên LSB của tín hiệu audio



Hình 3.6. Giao diện thủy vân một ảnh nhị phân trên LSB của Audio

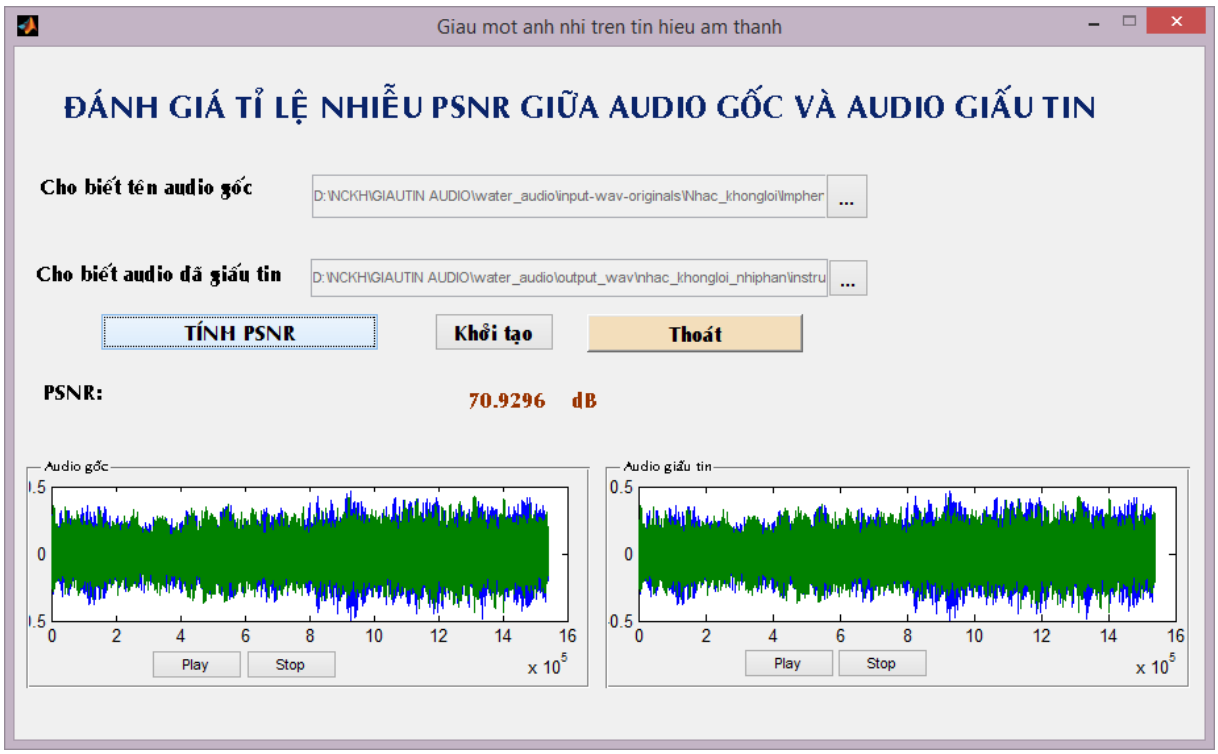
Hình 3.7. Giao diện tách thủy vân một ảnh nhị phân trên LSB của Audio

Dưới đây là giao diện thủy vân và tách thủy vân là một tín hiệu audio tín hiệu biến đổi của audio gốc.



Hình 3.8. Giao diện thủy vân một audio trên tín hiệu biến đổi của Audio gốc

Hình 3.9. Giao diện tách thủy vân là một audio trên tín hiệu biến đổi của Audio



Hình 3.10. Giao diện tính thay đổi nhiễu PSNR của audio gốc và audio giấu tin

3.2. Thử nghiệm

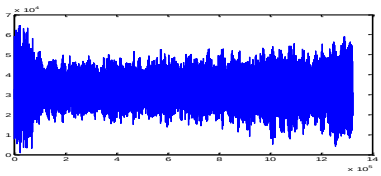
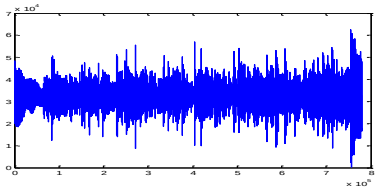
3.2.1. Tập dữ liệu thử nghiệm

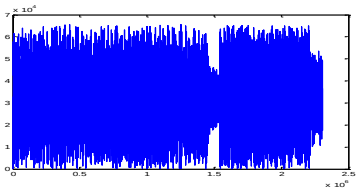
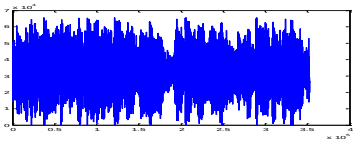
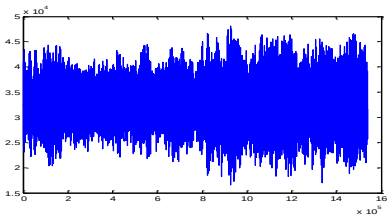
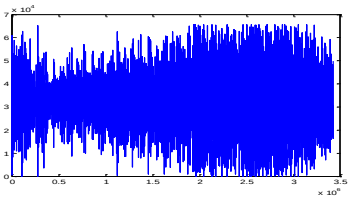
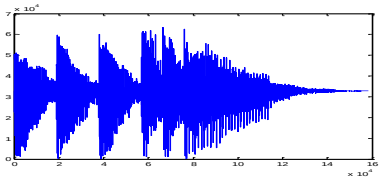
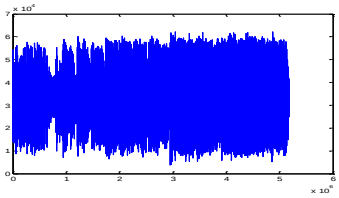
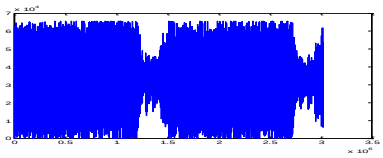
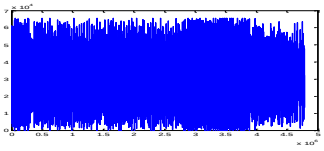
Tập dữ liệu âm thanh dùng để thử nghiệm bao gồm 10 tệp âm thanh trong đó:

+ 5 tệp âm thanh loại nhạc không lời,

+ 5 tệp âm thanh loại nhạc có lời

Bảng 3.1. Biểu đồ tín hiệu của các âm thanh đầu vào

STT	Nhạc không lời	Nhạc có lời
1		

2		
3		
4		
5		

Dự liệu thông điệp gồm: đoạn âm thanh nào đó, tệp văn bản bất kỳ, ảnh nhị phân, ảnh cấp xám.

3.2.2. Thực hiện thử nghiệm

Tiến hành thử nghiệm để giấu tin cho tập dữ liệu tín hiệu âm thanh (gồm 10 tệp), cho từng trường hợp sau đó đánh giá chất lượng của tín hiệu âm thanh sau khi giấu tin bằng PSNR ta được các kết quả cho từng trường hợp như sau:

+ Trường hợp 1: giấu ảnh nhị phân (có kích cỡ 337 x 336 pixel ứng với 113 232 bit, hình 3.11) trong tập audio sử dụng kỹ thuật giấu LSB (trình bày trong mục 2.1), cho phép thay đổi 4 bit LSB ta được bảng kết quả 3.2.

Hình 3.11. ảnh nhị phân cần giấu trong tập dữ liệu audio

***Bảng 3.2**. Kết quả đánh giá PSNR chất lượng của audio sau khi giấu ảnh nhị phân*

STT	**PSNR (Decibel – dB)**	
	Nhạc không lời	**Nhạc có lời**
1	73.3162 dB	79.0625 dB
2	75.4257 dB	77.6028 dB
3	70.9296 dB	79.4459 dB
4	63.4935 dB	78.3756 dB
5	76.4811 dB	78.9611 dB

+ Trường hợp 2: Cùng giấu audio (có độ dài 111 132 tín hiệu) trong tập dữ liệu audio sử dụng kỹ thuật giấu trên miền biến đổi tần số Fourier trình bày trong mục 2.4 ta được bảng kết quả 3.3.

***Bảng 3.3**. Kết quả đánh giá PSNR chất lượng của audio sau khi giấu tin hiệu audio*

STT	**PSNR (Decibel – dB)**	
	Nhạc không lời	**Nhạc có lời**
1	35.5063 dB	26.0642 dB
2	29.7468 dB	27.1257 dB
3	36.8012 dB	28.356 dB
4	30.768 dB	25.6758 dB

5	27.1032 dB	25.3917 dB

+ Trường hợp 3: giấu ảnh cấp xám (kích cỡ) trong tập 10 tệp audio sử dụng kỹ thuật giấu trên LSB cải tiến bày trong mục 2.2 được bảng kết quả 3.4.

***Bảng 3.4**. Kết quả đánh giá PSNR chất lượng của audio sau khi giấu ảnh cấp xám*

STT	**PSNR (Decibel – dB)**	
	Nhạc không lời	**Nhạc có lời**
1	80.0608 dB	81.3532 dB
2	79.6676 dB	84.0324 dB
3	78.4592 dB	82.0447 dB
4	81.6744 dB	86.4382 dB
5	77.3498 dB	88.5673 dB

+ Trường hợp 4: giấu văn bản (gồm 13728 ký tự ứng với 109824 bit) trong tập 10 tệp audio sử dụng kỹ thuật giấu đã đề xuất trình kết hợp giữa 7 bit MSB với 4 bit LSB bày trong mục 2.3 được bảng kết quả 3.5.

***Bảng 3.5**. Kết quả đánh giá PSNR chất lượng của audio sau khi giấu văn bản*

STT	**PSNR (Decibel – dB)**	
	Nhạc không lời	**Nhạc có lời**
1	85.6798 dB	91.4732 dB
2	89.6473 dB	90.0247 dB
3	83.4591 dB	92.0687 dB
4	84.673 dB	91.4382 dB
5	87.3452 dB	93.8745 dB

3.3. Nhận xét và đánh giá

Dựa vào bảng 3.2, 3.3, 3.4 và 3.5 chúng ta có thể đưa ra nhận xét như sau về các kỹ thuật giấu tin, ở đây các thông tin giấu có độ dài xấp xỉ nhau về số bit:

+ Kỹ thuật giấu tin trên miền biến đổi tần số Fourier có PSNR thấp nhất (bảng 3.3) (là trường hợp thay đổi nhiều giá trị ban đầu của tín hiệu audio gốc), so với các kỹ thuật giấu khác nó gẫy ra nhiễu tín hiệu gốc nhiều nhất nhưng chất lượng vẫn có thể chấp nhận được với các khán giả không quá khó tính. Tuy nhiên theo nhóm tác giả đề xuất kỹ thuật này cho rằng đây là phương pháp có thể giữ thủy vân đã giấu bền vững trước một số tần công về xử tín hiệu. Do nhóm nghiên cứu chưa có nhiều thời gian xây dựng một số phép biến đổi tín hiệu như lọc, giảm hoặc làm trơn nhiễu... nên chưa thể khẳng định được đây là các kỹ thuật thủy vân bền vững.

+ Kỹ thuật giấu văn bản trên LSB kết hợp với 7 bit MSB của nhóm tác giả đề xuất làm thay đổi rất ít giá trị của tín hiệu gốc nên giá tri của PSNR rất cao (bảng 3.5) (nghĩa là ít nhiễu so với giá trị ban đầu)

+ Kỹ thuật giấu ảnh cấp xám và ảnh nhị phân (bảng 3.2 và 3.4) có giá trị PSNR bằng nhau vì hai kỹ thuật này về cơ bản thay đổi số bit LSB của tín hiệu gốc như nhau.

Chưa nghiên cứu và thử nghiệm được trên âm thanh giọng nói (bản tin, thời sự,... kể truyện) để có thể đánh giá ảnh hưởng của thông tin giấu đến chất lượng âm thanh.

Các phương pháp đánh giá mới chỉ dựa trên tính giá trị nhiễu PSNR của tín hiệu, chưa đánh giá được dựa vào hệ thống thính giác của người nghe (thính giả) thực sự để có sự nhìn nhận về hệ thống giấu tin một cách trực quan và đa dạng.

KẾT LUẬN

Kỹ thuật giấu thông tin trong âm thanh số là một trong những hướng nghiên cứu chính của phương pháp giấu thông tin hiện nay và đã đạt được những kết quả khả quan. Đề tài nghiên cứu khoa học này đã tìm hiểu và trình bày lại một số kỹ thuật giấu tin điển hình theo hai hướng giấu tin chính đó là:

+ Giấu thông tin trên miền dữ liệu thời gian thực: chú trọng đến một số kỹ thuật giấu tin trên các bit LSB hoặc kết hợp giữa các bit MSB và LSB của tín hiệu âm thanh. Với kỹ thuật giấu tin này chúng ta có thể áp dụng để giấu văn bản, ảnh nhị phân, ảnh cấp xám hoặc một đoạn âm thanh vào tín hiệu âm thanh.

+ Giấu thông tin trên miền biến đổi: tìm hiểu phương pháp giấu tin trên miền biến đổi Fourier có thể áp dụng để giấu một đoạn âm thanh vào một đoạn âm thanh khác.

Từ các kỹ thuật tìm hiểu trên và trong quá trình cài đặt thử nghiệm tác giả của đề tài nghiên cứu này đã đề xuất một phương pháp giấu văn bản trên tín hiệu âm thanh bằng cách kết hợp giữa một số bit MSB và LSB với các chuỗi bit của ký tự văn bản. Với kỹ thuật giấu tin này chúng ta chỉ cần thay đổi một đến hai bit LSB của tín hiệu âm thanh và vẫn đảm bảo chất lượng tín hiệu audio ban đầu.

Do hạn chế về thời gian nên trong đề tài này chưa nghiên cứu được phương pháp cải tiến giấu tin trên miền biến đổi của dữ liệu audio. Đặc biệt chưa đánh giá được độ bền vững của các kỹ thuật giấu tin khi bị một số tấn công làm ảnh hưởng đến tín hiệu của audio đã mang thông tin. Các kỹ thuật giấu tin chủ yếu thử nghiệm trên tín hiệu audio định dạng tệp *.wav, chưa thực hiện trên các tệp âm thanh nén định dạng mp3, mp4, mpeg...

Tuy nhiên báo cáo của nghiên cứu có thể coi làm tài liệu cho sinh viên chuyên ngành CNTT, vì đã tổng hợp và trình bày những khái niệm cơ bản nhất về giấu tin trong tín hiệu audio làm nền tảng mới cho các nghiên cứu sau này về an toàn thông tin trong dữ liệu đa phương tiện.

Hướng nghiên cứu tiếp theo của đề tài sẽ nghiên cứu phương pháp giấu tin bền vững trên tín hiệu audio và trên dạng âm thanh nén: mp3, mp4,...

TÀI LIỆU THAM KHẢO

[1]. Aubrey de Seslincourt (1996), *Herodotus - The Histories*, Penguin Books, London.

[2]. Ingemar Cox, Jeffrey Bloom, Matthew Miller, Ton Kalker, Jessica Fridrich (2008), *Digital Watermarking and Steganography*, Second Edition, Morgan Kaufmann Press, USA.

[3]. Jessica Fridrich (2009), *Steganography in digital media: principles, algorithms, and applications*, Cambridge University Press

[4]. http://www.infotech.oulu.fi/Annual/2003/MTEAM.html

[5]. Nguyễn Xuân Huy, Trần Quốc Dũng, *Giáo trình giấu tin và thuỷ vân ảnh, Trung tâm thông tin tư liệu, TTKHTN - CN 2003.*

[6]. Min Wu, Multimedia Data Hiding, Princeton University, USA, 2001.

[7]. Michael Arnold, Dr. Christoph Busch, Watermarking of Audio, Music Scores and 3D Models, INI-GraphicsNet - Press & Media, 2003.

[8]. Chun-Shien Lu, Multimedia Security: Steganography and Digital Watermarking techniques for Protection of Intellectual Property, IDEA Group Publishing, 2005.

[9]. https://www.clear.rice.edu/elec301/Projects01/smokey_steg

[10]. Pradeep Kumar Singh, R.K.Aggrawal: Enhancement of LSB based Steganography for Hiding Image in Audio, (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 05, 2010, 1652-1658.