해밍코드의 신드롬을 이용한 데이터 은폐

Data hiding Scheme based on Syndrome of Hamming(7,4) Code

김 천 식*, 김 형 중**
Cheonshik Kim, Hyoung Joong Kim

Abstract

According to researches[1], good image quality and amount of hiding information is a main point of steganography. In this point of view, [3] is a very good scheme to hide in an image. However, it cannot hide a lot of information. In order to solve this problem, we propose new method to hide more information than that of [3]. In addition, it can be reduced distortion of an image than that of [4]. Thus, our method is a very efficient and novel scheme.

Keywords : title, abstract, keywords, introduction, heading

I. 서 론

스테가노그라피(Steganograpy)와 데이터 은폐(Data Hiding)는 멀티미디어 매체에 데이터를 은폐한다는 점에서 유사하지만, 데이터 은폐는 외부 공격에 스테가노그라피 보다는 취약한 면이 있다. 따라서, 데이터 은폐를 목적으로 할 때 데이터를 얼마나 많이 은폐 하는가는 이 분야에서 중요한 요소이다. 데이터를 은폐하는 목적은 멀티미디어의 저작권을 위해서 혹은 비밀 통신을 위한 채널 등 다양한 목적으로 사용 가능하다[1-2]. 이미지에 데이터를 은폐한 경우, 이미지의 질은 매우 중요하다. 왜냐하면, 이미지의 질이 좋지 않은 경우 데이터 은폐 사실을 간접적으로 보여주게 되는 것이다. 따라서, 이미지의 질의 높게 유지하면서 많은 양의 데이터를 저장하는 것은 알고리즘을 개발하는 전문가에게 매우 흥미 있는 과제가 된다. 데이터를 은폐하는 방법으로 가장 많이 사용하는 방법이 공간 도메인 환경에서 데이터를 은폐하는 것으로서, 주소 LSB를 flip함으로써 데이터를 은폐하고 있다. 본 논문에서는 데이터를 Hamming code[2] 기법을 사용해서 데이터를 은폐하고자 한다. 기존에 [4-5]의 기법이 데이터 은폐에 적용되었다. [4]의 방법에서는 syndrome을 이용하여 데이터를 은폐했기 때문에 1비트를 바꾸고, 3비트를 저장하는 방법을 제안하였다. [5]의 연구에서는 데이터의 은폐 성능을 높이기 위해서 7비트를 바꾸고 7비트를 저장하는 방법이 제안되었다. 본 논문에서는 syndrome을 이용하여 4비트를 바꾸고, 6비트를 저장하는 방법을 제안하고자 한다.

접수일자 : 2009년 8월 10일
최종완료 : 2009년 8월 14일
*안양대학교 교양학부
**고려대학교 정보경영공학부
교신저자, E-mail : database.lab@gmail.com

II. 본 문

1. *Hamming Code*

패리티 비트(Parity Bit)에 의한 오류 검출은 단지 오류 검출만 되지만 해밍코드 (Hamming Code)는 오류 검출 후 오류 정정까지 가능하다. 해밍 코드는 R. W. Hamming에 의해 고안되었다. 해밍(7,4)은 4비트 데이터를 전송할 때 3비트의 패리티 비트를 추가해서 채널에 데이터를 전달할 때, 오류를 찾고 정정이 가능하다. 두 해밍 행렬은 코드 생성자 (G)와 패리티 체크 행렬 (H) 로 정의 된다.

$$c = \mathrm{mod}((G \times d^T)^T, 2) \quad \cdots\cdots (1)$$

우리가 $d = (1101)$를 전송하기 위해서 인코등한다면 $c(codeword)$는 (1101001)이 된다.

$$s = (\mathrm{mod}(H \times C^T)^T, 2) \quad \cdots\cdots (2)$$

수신측에서는 공식(2)를 이용해서 데이터가 패리티 코드를 제외한 값이 정상적인 값인가를 그렇지 않은 값인가를 판단한다. 이때, s는 $syndrome$으로서 데이터가 무결성인 경우 $syndome$은 (0 0 0)이 된다.

2. 해밍코드를 이용한 데이터 은폐

코드를 보내는 측은 커버 이미지 X를 전송한다. X는 n개의 요소 (x_i^n)로 구성된다. 이 경우, $x_i \in J$ 와 같다. 이때, 우리는 공간 도메인을 사용하기 때문에 $J = \{0, 1, \ldots, 255\}$ 이 된다.

X로부터 9개씩을 분리하여 사용한다. 이중 7개의 x를 수식 (3)과 같이 추출하여 코드워드 c로 사용한다.

$$x = (LSB(x_1), LSB(x_2), \ldots, LSB(x_n)) \quad \cdots\cdots (3)$$

추출한 코드워드 c를 공식 (2)에 적용하여 신드롬을 구한다. 이때, 구한 첫 번째, 신드롬을 $syndrome_1$으로 한다. 메시지 M은 n개의 요소 (m_i^n)로 구성된다. 데이터를 저장하기 위해서 $syndrome_1$ 3비트와 메시지 3비트에 대해서 xor 연산을 수식 (4)와 같이 실행한다.

$$syndrome_2 = exor(syndrome_1, message_1^3) \ldots\ldots (4)$$

$syndrome_2$를 10진수로 바꾼 값이 해밍 시퀀스에서 오류가 있는 위치를 의미한다. 해당 위치의 값이 0이면 1로 1이면 0으로 대치한다.

$syndrome_2$와 새로운 메시지 3비트를 가져와서 수식 (5)와 같이 연산을 수행한다.

$$syndrome_3 = exor(syndrome_2, message_1^3) \ldots\ldots (5)$$

$syndome_3$의 값을 이용해서 시퀀스의 해당 위치를 바꾸면, $syndrome_2$를 복원할 수 없기 때문에, $syndome_3$의 값을 $LSB(x_8)$과 $LSB(x_9)$에 나누어 저장한다. 총 3비트를 저장해야 하므로 x_8에는 1비트를 x_9에는 2비트를 저장한다.

수신측에서는 공식 (3)에서 추출한 값에 코드워드 c 에 $LSB(x_8)$과 $LSB(x_9)$에 표시한 값을 10진수로 변환 한 후, c에서 해당 위치의 값이 1이면 0으로 0이면 1로 변환한 후 (2)를 적용하면 은익 했던 두 번째 메시지를 구할 수 있다. 그다음, 첫 번 째 메시지를 구하기 위해서 두 번 째 메시지를 구하기 위해서 코드워드 c에 처리 되었던 부분을 복원한다. 그런 다음, 공식 (2)를 수행하면 첫 번 째 메시지를 구할 수 있다.

III. 결 론

Hamming+1 방법[3]은 3×3에 2비트를 flip하고 4비트를 저장하는 방법이고, [4]의 방법은 7비트를 flip하고 7비트를 저장하는 방법이다. 하지만, 본 논문에서 제안하는 방법은 4비트를 flip하고 6비트를 저장하는 방법으로서 이미지 질과 데이터 은폐의 량을 모두 고려한 방법으로서 위의 방법보다 효율적이라고 할 수 있다. 향후, 본 논문에서 제안한 방법을 구현하여 실효성을 증명하고자 한다.

[참 고 문 헌]

[1] Provos, N.and Honeyman, P., *Hide and seek: an introduction to steganography*, IEEE Security & Privacy, vol.1, no. 3, pp:32- 44, June 2003.

[2] R. W. Hamming, *Error detecting and error correcting codes*, Bell System Tech. J. 29, pp.147-160, 1950.

[3] W.Zhang, S. Wang, and X. Zhang, *Improve embedding efficiency of covering codes for applications in steganography*, IEEE Communications Letters, 11(8):680-682, August 2007.

[4] The Duc Kieu and Yung-Chen Chou, *A High Payload Steganographic Scheme Based on (7, 4) Hamming Code for Digital Images*, ISECS : International Symposium on Electronic Commerce and Security 2008, pp. 16-21, 2008.

김 천 식

1997년 한국외국어대학교 컴퓨터및정보통신공학과 (공학석사)
2003년 한국외국어대학교 컴퓨터및정보통신공학과 (공학박사)
2000년~2003년 경동대학교 교수
2004년~현재 안양대학교 교수
2007년~현재 대한전자공학회 컴퓨터소사이어티 분과위원장
2006년~현재 인터넷 정보학회 학회편집위원
2006년~현재 대한교통학회 정회원
2005년~현재 한국데이터베이스학회 정회원
2008년~인터넷방송통신tv학회 상임이사
2008년~ICHIT 2008 committee member
2009년~ICHIT 2009 committee member
<관심분야> 데이터베이스, 데이터마이닝, 이미지처리, e-Learning, Agent system
<e-mail> database.lab@gmail.com

김 형 중

1978년 서울대학교 전기공학과 (공학사)
1986년 서울대학교 제어계측공학과(공학석사)
1989년 서울대학교 제어계측공학과(공학박사)
1992년~1993년 USC 방문교수
1989년~2006년 강원대학교 교수
2006년~현재 고려대학교 정보경영공학부 교수
<관심분야> 멀티미디어보안, 분산처리, 콘텐츠공학 등
<e-mail>khj-@korea.ac.kr