통계 분석에 강인한 심층 암호

*유정재, *이광수, *이상진, **박일환
*고려대학교 정보 보호 대학원, **국가 보안 연구소
e-mail : *{shakehds, kslee}@cist.korea.ac.kr, *sangjin@korea.ac.kr,
**ilhpark@etri.re.kr

A Secure Steganographic Scheme against Statistical analyses

*Jeong-jae Yu, *Kwang-su Lee, *Sangjin Lee, **IL-hwan Park
*CIST, Korea University, **NSRI, Daejeon, Korea

Abstract

Westfeld[1] analyzed a sequential LSB embedding steganography effectively through the χ^2-statistical test which measures the frequencies of PoVs(pairs of values). Fridrich[2] also proposed another statistical analysis, so-called RS steganalysis by which the embedding message rate can be estimated. In this paper, we propose a new steganographic scheme which preserves the above two statistics. The proposed scheme embeds the secret message in the innocent image by randomly adding one to real pixel value or subtracting one from it, then adjusts the statistical measures to equal those of the original image.

I. 서 론

초창기 심층 암호의 대부분은 원본 영상의 최하위 비트를 비밀 메시지 비트로 치환하는 방식이었기 때문에 인간의 감각으로는 메시지 삽입 여부를 구별해낼 수 없었지만 통계적 분석에 의하여 원본과 은닉물의 구별은 물론, 비밀 메시지의 삽입량까지도 거의 추정해낼 수 있을 만큼 취약점을 내포하고 있었다. 우리는 Westfeld[1]와 Fridrich[2]가 판단의 기준으로 정한 통계량을 각각 분석하였고, 이에 근거하여 원본의 통계량을 유지하면서도 대용량의 메시지를 삽입할 수 있는 방법을 제안하고자 한다. 제안하는 방식은 단순히 원본 영상의 최하위 비트를 변화시켜 메시지를 삽입하는 방식이 아닌 원본의 실제 화소값이 랜덤하게 증가하거나 감소하는 방식으로 메시지를 삽입하게 된다.

실험 결과, 위에서 제시한 통계적 분석법들은 물론 Provos 등이 제안한 변형된 χ^2-통계 분석법에도 탐지되지 않았다.

논문의 구성은 다음의 순서로 이루어진다. 2절에서는 지금까지 제안된 심층 암호의 통계 분석 방법들에 대하여 살펴보고 3절에서 우리가 SES(a Steganography Evading Statistical analyses)라고 부르는 심층 암호의 구체적인 알고리즘을 설명할 것이다. 4절에서는 실험 결과를 분석하고 5절에서 결론 및 앞으로의 과제에 대하여 논의하겠다.

II. 심층 암호의 통계 분석

2.1 χ^2-통계 분석

심층 암호 통신을 실생활에 적용하려면 삽입 메시지의 기밀성을 추구하기 위하여 암호화 과정을 거치게 되므로 일반적으로 원본 영상에 암호문이 삽입된다.

단순한 최하위 비트 치환 심층 암호라면 이 과정에서 시각적으로 드러나진 않지만 원본과 은닉물 사이에 심각한 통계적 차이를 발생시킬 수 있다. 즉, 메시지를 삽입하지 않은 원본의 인접한 두 화소, 혹은 색인값들(PoVs, pair of values)이 발생한 빈도를 살펴보면 변형을 가하지 않을 경우 서로 다르게 나타나지만 은닉물에서는 인접한 두 화소값들의 발생 빈도가 암호문의 통계적 분포 때문에 거의 비슷하게 되는 것이다.

이 차이를 Westfeld[1]등이 암호문의 랜덤성 테스트[3]와 비교하여 메시지 삽입 확률을 정량화하였고 이러한 심층 암호 분석 기법을 χ^2-통계 분석이라고 명명하였다. 그리고 Provos[4]는 탐지 구간을 세분화하거

나 PoVs의 샘플을 화소 x와 $(x+1)$에서 x와 $(x-1)$로 변형하여 χ^2-통계 분석 방법을 다시 적용한 변형된 χ^2-통계 분석을 제안하면서, 메시지 삽입 후에도 원본의 PoVs 통계량을 유지하는 Outguess를 발표하였다. Outguess는 랜덤하게 원본의 최하위 비트를 비밀 메시지 비트로 변화시킨 후 삽입에 이용되지 않은 나머지 원본의 최하위 비트들로 원본의 PoVs 통계량과 일치하도록 변화시키는 심층 암호 알고리즘이다.

2.2 RS 통계 분석

Fridrich[2] 등은 실험에 기반하여 심층 암호를 개발하던 중 원본 영상의 고유한 특성을 발견하고 RS 통계 분석법을 제안하였다. RS 통계 분석은 최하위 비트 변환 기반의 심층 암호 탐지와 삽입 메시지량을 추정하는 분석법이다. 그 핵심 내용은 다음과 같다.

먼저 원본 C를 서로 공통 부분이 없으며 각각 n개의 원소를 갖는 집합들로 분할한다. 이러한 집합을 $G=(x_1,\ldots,x_n)$라고 할 때 판별 함수 f를 다음과 같이 정의한다.

$$f(x_1,x_2,\ldots,x_n)=\sum_{i=1}^{n-1}|x_{i+1}-x_i| \qquad (1)$$

그리고 주기 2를 가지는 가역함수 F_1과 F_{-1}는 다음의 성질을 만족해야 한다.

$F_i(F_i(x))=F_0(x)=x,\ i\in\{-1,1\}$

$F_1: 0\leftrightarrow1, 2\leftrightarrow3, \ldots, 254\leftrightarrow255$

$F_{-1}: -1\leftrightarrow0, 1\leftrightarrow2, \ldots, 253\leftrightarrow254, 255\leftrightarrow256$

그러면 플립(flipping) 가역함수 F_1과 F_{-1}사이에는 다음 식이 성립하게 된다.

$$F_{-1}(x)=F_1(x+1)-1 \quad \text{for all } x \qquad (2)$$

이 때 집합 G의 특성을 다음과 같이 결정한다.

Regular group : $G\in R\Leftrightarrow f(F_i(G))>f(G)$

Singular group : $G\in S\Leftrightarrow f(F_i(G))<f(G)$

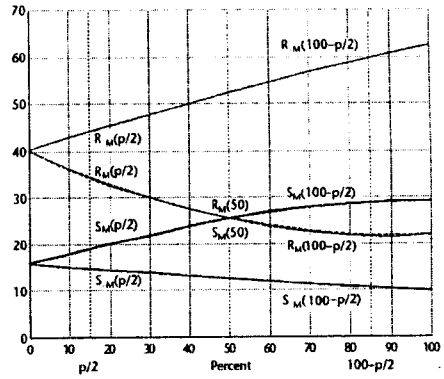
Unusable group : $G\in U\Leftrightarrow f(F_i(G))=f(G)$

또한 임의의 마스크 M을 집합 G에 적용한 것을 $F_M(G)=(F_{M(1)}(x_1),\ldots,F_{M(n)}(x_n))$이라고 나타낼 때, 이러한 마스크 M에 대하여 원본 C를 regular, singular, unusable 그룹으로 분류할 수 있다. 많은 실험을 거쳐 Fridrich 등은 최하위 비트를 조작하지 않은 원본의 경우 대부분 다음과 같은 통계적 특성을 만족함을 발견하였다.

$$R_M+S_M\le100\ ,\quad R_{-M}+S_{-M}\le100$$

$$R_M\cong R_{-M}\quad\text{and}\quad S_M\cong S_{-M} \qquad (3)$$

여기에서 R_M과 S_M은 전체 영상에 대한 regular 그룹과 singular 그룹의 백분율을 각각 나타낸다. 그리하여 위의 통계적 가설을 바탕으로 은닉물의 최하위 비트 메시지 삽입량을 추정할 수 있다. Fridrich는 원본 영상 전체 화소의 최하위 비트에 난수를 삽입하였을 때 R_M과 S_M이 유사함을 실험적 가설로 추가하여 [그림 1] 과 같은 RS 통계 그래프를 유추하였다.



[그림 1] RS 통계 그래프

$$R_M(50)\cong S_M(50) \qquad (4)$$

R_M과 S_M은 2차 보간으로, R_{-M}과 S_{-M}은 1차 보간법을 사용하였을 때 식 (5)와 같은 방정식을 유도할 수 있으며, 메시지 삽입 확률 p는 식 (6)으로 주어진다.

$$2(d_1+d_0)x^2+(d_{-0}-d_{-1}-d_1-3d_0)x+d_0-d_{-0}=0 \qquad (5)$$

$$d_0=R_M(p/2)-S_M(p/2)$$

$$d_1=R_M(1-p/2)-S_M(1-p/2)$$

$$d_{-0}=R_{-M}(p/2)-S_{-M}(p/2)$$

$$d_{-1}=R_{-M}(1-p/2)-S_{-M}(1-p/2)$$

$$p=x/(x-1/2) \qquad (6)$$

메시지 삽입 추정 확률은 식 (5)에서 절대값이 작은 근을 취하는데 우리는 이 삽입 확률이 0이 되도록 메시지 삽입 후 R_M과 R_{-M}의 비율을 조정할 것이다. 플립핑 함수는 주기가 2인 특성을 가지므로, 다음 식 $F_M(R)=S,\ F_M(S)=R$은 항상 성립하게 된다.

Fridrich 등이 실험 결과로 제시한 것처럼 RS 통계 분석은 최소 2%내외의 삽입 메시지 추정 오차가 발생한다. 제안하는 SES 심층 암호는 랜덤 플립핑 함수를 이용하여 메시지를 삽입하므로 은닉물의 RS 통계량이 원본 RS 통계량과 큰 차이가 발생하지 않고, 메시지 삽입 후 안전성을 위하여 메시지 삽입에 이용되지 않는 원본 영상 화소를 조정하여 은닉물의 RS 통계량을 오차 범위 내로 일치시킨다.

III. SES 심층 암호

SES(a Steganography Evading Statistical analyses) 심층 암호는 이름의 유래에서 알 수 있는 것처럼 원본 영상의 알려진 통계량을 만족하도록 고안된 심층 암호 기법이다. SES 알고리즘은 랜덤 플립핑과 간단한 실수 계산만으로 이루어졌기 때문에 구현 또한 용이하다. 구체적인 과정은 다음과 같다.

1. 원본 영상의 RS 통계량을 계산하여 메모리에 저장한다.

2. 비밀 메시지를 암호화하고 메시지의 길이를 앞부분에 연접시킨다.

$$S = l \| c_1 \| \ldots \| c_n = s_1 \| \ldots \| s_{n+1}, \quad l = n$$

3. 삽입될 메시지(메시지 크기+암호문)의 비트와 원본 영상 화소의 최하위 비트를 순차적으로 비교한다. 일치하면 플립 함수 F_0를 적용하고 일치하지 않으면 플립 함수 F_1과 F_{-1}을 랜덤하게 적용하여 은닉 영상 화소의 최하위 비트를 삽입 메시지 비트와 일치시킨다.

$$x'_i = \begin{cases} x_i & \text{if} \quad s_i = LSB(x_i) \\ F_j(x_i) & otherwise \end{cases}$$

여기에서 j는 암호학적으로 안전한 난수 발생기로부터 랜덤하게 선택한다.

$$j = r \in \{-1, 1\}$$

4. 메시지 삽입이 모두 끝난 후 은닉물 전체의 RS 통계량을 계산한다. 원본의 RS 통계량과 비교하여 2%이상 오차 발생 시 삽입하고 남은 원본 영상의 화소를 조정하여 은닉물의 RS 통계량을 원본과 2% 이내로 일치시킨다. $d_0 - d_{-0} = 0$이 되면 식 (5)와 (6)에서 알 수 있듯이 삽입 메시지 추정 확률 p도 0이 된다. 다음 식 $F_M(R) = S$, $F_M(S) = R$을 이용하여 R_M과 R_{-M}의 비율을 일치시키도록 한다.

Westfeld 등이 제안한 PoVs 통계량은 은닉 영상 화소값들이 두 개씩 서로 쌍을 이루어 비슷하게 분포하게 되는 특성에 근거한 것이므로 위의 과정 3처럼 똑같은 값을 가지는 화소 x 에 대하여 랜덤하게 더하거나 뺀다면 x 가 $(x+1)$ 이나 $(x-1)$ 과 서로 상관관계를 갖지 않게 된다. 또한 Provos가 제안한 변형된 χ^2-통계 분석법처럼 샘플링 구간을 달리하거나 샘플을 x 와 $(x+1)$ 에서 x 와 $(x-1)$ 로 변형하여 분석해도 SES 심층 암호로 삽입한 은닉 영상에서 별다른 특징을 발견할 수 없었다.

[그림 2] 40KB 삽입 은닉 영상(1bit/pixel)

보다 안전한 심층 암호 통신을 해야 할 경우 원본 영상 1/2 이상의 화소를 RS 통계량 조정에 사용할 수 있다면 거의 모든 실험 영상에 대하여 RS 통계 분석을 통한 탐지를 회피할 수 있었다 - [표 1] 참조.

IV. 실험 결과

χ^2-통계 분석과 RS 통계 분석을 통해서 삽입 확률이 모두 0인 160개의 BMP 파일 영상을 대상으로 하여 각각 10kbytes, 40kbytes, 70kbytes의 암호문을 삽입한 후 통계 분석을 하였다. 실험 영상은 디지털 카메라 촬영 후 BMP로 변환한 영상과 컴퓨터로 생성한 영상(fractal image), 그리고 24 비트 만화 등 여러 가지 특성의 영상들을 선정하였다. 또한 삽입 용량의 비교가 편리하도록 원본의 크기는 모두 $512 \times 379 \times 24 \cong 570$ kbytes으로 고정하였다. 암호문과 난수의 생성 방식은 비밀키 암호인 AES[5]를 사용하였다.

[그림 2]는 한 화소당 1비트 삽입 방식으로 40KB의 메시지를 삽입한 은닉 영상이다.

[표 1]은 160개의 은닉 영상에 대하여 각각의 통계 분석법에 의해 탐지된 영상의 개수를 나타낸다.

RS 통계 분석의 경우에는 탐지의 임계치 p를 각각 2%와 4%로 달리 했을 때의 결과를 보여준다.

[표 1]에서 시사하는 바와 같이 암호문의 삽입량에 상관없이 χ^2-통계 분석으로는 SES 심층 암호에 의한 삽입 영상을 구별할 수 없었으며, 암호문의 삽입량이 적을수록 즉, RS 통계량을 조절할 수 있는 영역이 많을수록 RS 통계 분석에도 강인함을 볼 수 있다. $570/8 = 71.25$이므로 한 화소당 1비트의 메시지를 삽입할 경우 최대 삽입량의 약 50% 가량을 RS 통계량 조정에 사용(40KB)했을 때 거의 탐지되지 않음을 확인할 수 있다.

[표 1] 160개 은닉 영상의 분석 결과(1bit/pixel)

	10KB	40KB	70KB
χ^2-통계 분석	0/160	0/160	0/160
RS 통계 분석(4%)	0/160	0/160	4/160
RS 통계 분석(2%)	0/160	3/160	9/160

[표 2]는 [그림 2]와 원본 영상과의 RS 통계량을 보여준다. 비록 원본 영상의 RS 통계량과는 다소 차이가 있지만, 은닉물의 R_M과 R_{-M}, 그리고 S_M과 S_{-M}의 개수가 거의 유사하여 식 (3)의 통계량을 만족하고 있다.

[표 2] 원본과 은닉 영상의 RS 통계량

	원본영상	은닉영상(40KB)
R_M의 개수	59160	58100
R_{-M}의 개수	58952	58304
S_M의 개수	26878	27919
S_{-M}의 개수	26943	27850

V. 결 론

본 논문에서 제시하는 SES 심층 암호는 잘 알려진 통계적 심층 암호의 분석법을 고려하여 이러한 통계량을 유지시킨 심층 암호이기 때문에 χ^2-통계 분석이나 RS 통계 분석으로는 탐지되지 않는다. 또한 Outguess[4]나 F5[6] 심층 암호처럼 어느 한 가지만의 통계 특성을 만족시킨다거나 제한된 메시지 삽입량을 가지지도 않는다. 지금까지 알려진 통계 분석에 안전하게 메시지를 삽입하면서도 최하위 k 비트로의 확장이 가능하므로 단순 최하위 비트 치환 심층 암호보다 많은 메시지 삽입이 가능하다.

향후 메시지 삽입 시 비밀키에 의해 랜덤하게 삽입할 수 있도록 일대일 함수(permutation)기법과 은닉물의 RS 통계량이 보다 원본의 통계량에 근접하도록 하는 연구가 보완되어질 예정이다. 또한 공간 영역에서 제안된 심층 암호 기법을 주파수 영역으로 확장하여야 보다 실용적인 심층 암호 통신을 가능하게 할 것이다.

참고문헌

[1] A. Westfeld and A. Pfitzmann: "Attacks on Steganographic Systems," Information Hiding, LNCS vol. 1768, Springer-Verlag, 1999, pp.61~76

[2] J. Fridrich, M. Goljan, and R. Du, "Detecting LSB steganography in color and gray-scale image," Magazine of IEEE Multimedia, 2001, pp. 22-28.

[3] U. Maurer, "A universal statistical test for random bit generators", in Advances in Cryptology - CRYPTO'90, A. J. Menezes and S. A. Vanstone, eds., vol. 537 of Lecture Notes in Computer Science, Springer- Verlag, 1991, pp.409~426.

[4] N. Provos, "Defending Against Statistical Steganalysis," in Proceedings of the 10th USENIX Security Symposium, 2001, pp.323-335.

[5] http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf.

[6] A. Westfeld, "F5--A Steganographic Algorithm," Information Hiding, 4th International Workshop, LNCS 2137, Springer-Verlag 2001, pp. 289-302