한국정보전자통신기술학회논문지(jkiiect)'22-8, Vol.15 No.4
논문 22-15-04-217 http://dx.doi.org/10.17661/jkiiect.2022.15.4.217

가변 ShiftRows를 이용한 하이브리드 기법에서 한글 메시지 은닉을 위한 이미지 스테가노그래피

지선수*

Image Steganography for Hiding Hangul Messages in Hybrid Technique using Variable ShiftRows

Seon-su Ji*

요 약 현대사회에서 정보는 중요한 역할을 한다. 대부분의 정보는 디지털 공간에서 처리되고, 이동된다. 사이버 공간에서 저항성과 보안성에 기반한 비밀 통신은 기본적인 사항이다. 네트워크를 통해 송신 및 수신되는 디지털 정보를 보호하는 것이 필수적이다. 그러나 권한이 없는 이용자에 의해 정보가 유출되고, 위변조 될 수 있다. 제3자에 의해 통신 내용을 파악하기 위한 혁신적인 기법이 적용됨에 따라 기존 보호 시스템의 효율성이 떨어진다. 스테가노그래피는 매개체의 특정 영역에 비밀정보를 삽입하는 기술이다. 스테가노그래피와 스테간 분석 기술은 상충관계에 있다. 고도화되어가는 스테간 분석에 대응하기 위해 새롭고, 정교한 구현 시스템이 필요하다. 단계별 확산 및 불규칙성을 강화하기 위해, 계층 암호화 및 가변 ShiftRows를 기반으로 하는 한글 메시지에 대한 이미지 스테가노그래피의 하이브리드 구현 기술을 제안한다. 제안된 스테가노그래피 효율성과 성능을 측정하기 위해 PSNR을 계산하였다. 기본 LSB 기법과 비교할 때 PSNR은 1.45% 감소하였으나 확산과 임의성을 증가시킬 수 있음을 보였다.

Abstract Information plays an important role in modern society. Most of the information is processed and moved in the digital space. In cyberspace, confidential communication based on resistance and security is fundamental. It is essential to protect the information sent and received over the network. However, information may be leaked and forged by unauthorized users. The effectiveness of the existing protection system decreases as an innovative technique is applied to identify the communication contents by a third party. Steganography is a technique for inserting secret information into a specific area of a medium. Stegganography and steganalysis techniques are at odds with each other. A new and sophisticatedly implemented system is needed to cope with the advanced steganalysis. To enhance step-by-step diffusion and irregularity, I propose a hybrid implementation technique of image steganography for Hangul messages based on layered encryption and variable ShiftRows. PSNR was calculated to measure the proposed steganography efficiency and performance. Compared to the basic LSB technique, it was shown that the diffusion and randomness can be increased even though the PSNR decreased by 1.45%.

Key Words : Embedding, Hybrid Techniques, Image Steganography, LSB, Variable ShiftRows

1. 서론

네트워크에서 통신은 전송 자료에 대한 기밀성을 유지하며, 안전하게 전송하는 것은 필수적이다. 또한 사이버 공간을 통해 공유되는 통신 규모가 급격하게 증가함에 따라 네트워크 보호는 점점 더 중

*Department of Computer Sciences&Engineering, Gangnung-Wonju National University
Received June 25, 2022 Revised July 10, 2022 Accepted July 18, 2022

요해지고 있다. 전송되는 정보를 보호하기 위해 암호화 및 정보 숨기기 등의 두 가지 유형의 기술이 상호 보완하여 사용된다. 정보 숨기기는 제3자에 의해 숨겨진 정보를 감지할 수 없도록 만들고, 숨겨진 상태를 유지할 수 있도록 조치해야 하는 것이 중요하다. 암호화는 혼돈(confusion)과 확산(diffusion) 기법을 사용하여 민감한 데이터의 기밀성을 유지하는 데 적용되는 방법이다. 스테가노그래피는 비밀정보를 커버 이미지의 특정 영역에 은닉한 후 수신자에게 안전한 방식으로 전송한다. 스테가노그래피는 정보 은닉의 흔적을 남기지 않기 위해 텍스트, 이미지, 오디오, 비디오, 네트워크 프로토콜과 같은 매개체에 비밀정보를 포함하는 의사소통의 기술이다. 비밀정보를 삽입하기 위해 다양한 방법을 사용한다. 스테가노그래피 성능은 페이로드 효율성, 이미지 품질과 저항성을 갖춘 보호 수준을 포함하는 3가지 평가 기준을 사용한다[1-2].

네트워크의 다양한 웹사이트 특정 영역을 정보 은닉 장소로 사용할 수 있으며, 디지털 공간에서 사용될 수 있는 시나리오와 매개체는 다양하다. 스테간 분석 및 스테가노그래피에 대한 연구는 지난 몇 년 동안 광범위하게 발전하였다. 합성곱 신경망(convolution neural network) 혹은 생성적 적대 신경망(generative adversarial networks) 등을 이용한 새로운 기법의 스테간 분석은 스테가노그래피에 취약점을 발생시킨다. 이와 같이 새롭고 정교한 기술의 출현으로 기존 데이터 보안 시스템은 자료 보호의 효율성이 떨어지고 있다[3]. 스테간 분석의 혁신적인 기법에 따라 스테가노그래피에 대한 새로운 기법과 기능의 병합(merge)이 요구된다. 예를 들어 다층 하이브리드 암호화 기법과 불규칙성이 포함된 시스템이 필요하다. 이 논문에서는 데이터 암호화 프로토콜, 열 단위와 행 단위에서 다층 시프팅 과정이 혼합된 가변적인 하이브리드 기법을 구현하고, 효율성과 성능을 계산하여 타당성을 확인한다.

논문의 2장에서 스테가노그래피와 스테간 부석에 대한 관련된 자료를 표현하였다. 제안하고자 하는 방법은 3장에서, 적용 방법과 효율성을 평가하기 위해 결과는 4장에서 제시하였다. 5장의 결론에서 마무리하였다.

2. 관련 연구

사이버 공간에서 정보의 이동은 중요한 역할을 한다. 전송되는 정보는 공격자로부터 보호되어야 하고, 무결성 상태를 유지해야 한다. A. Agrawal 등은 순환 이동(circular shift)과 LSB(least significant bit)를 이용하는 수정된 LSB 스테가노그래피 기술을 사용하여 이미지 파일 형식(bmp, png, jpeg)에 의료 정보를 포함하는 기법을 제안하여 이미지 품질 성능을 높였다. 특히 png 형식의 커버 이미지는 텍스트 정보를 마스킹하는데 효과적이며, 이미지 품질과 삽입용량에서 매우 적절한 매개체임을 확인하였다[1]. A. Dave 등은 AES-128 암호화 기법과 결합된 LSB 접근 방식에 기반한 해상도, 파일 형식 등의 이미지 속성을 사용하여 비밀정보를 보호하는 새로운 접근법의 스테가노그래피를 제안하였다. 송신 및 수신되는 메시지는 RGB 이미지의 세 평면 안에 숨겨지며, 이미지에 비밀 메시지를 포함할 때 RGB 평면을 가로지르는 순서를 변경하는 방법을 제안하였다. AES-128 암호화를 사용하며, LSB에 기반한 이미지 스테가노그래피를 탐색하면서 파일 유형 및 해상도와 같은 이미지 속성을 사용하여 알고리즘의 보안성을 향상시켰다. 평균 제곱 오차(RMSE, root mean square error) 및 최대 신호대잡음비(PSNR, peak signal to noise ratio)를 이용하여 향상된 성능의 측정치를 제시하였다[4]. V. Bhuvaneshwari 등은 RDH(reversible data hiding)와 LSB 기술을 혼합하여 암호화된 비밀정보를 삽입하는 과정을 보였다. 또한 이 과정에서 보안을 향상시키기 위해 데이터와 이미지에 대한 세 가지 키 즉, 세션키, 공개키, 개인키를 사용하는 기법을 제안하였다. AES 알고리즘을 사용하여 암호화된 원본 자료를 숨기는 기법을 제시하였다. 데이터 손실 없이 LSB 방식을 사용하여 우수한 결과

를 얻을 수 있음을 보였다[5]. R. Srivastava 등은 3개의 서로 다른 알고리즘을 사용하여 비밀정보를 3회씩 암호화하는 계층화된 암호화 구조를 사용하며 키와 정보를 보호하는 다층 하이브리드 기법을 제안하였다. 여기에서 AES와 SHA-1를 사용하였으며, Python 도구로 시스템을 구현하고 실험 결과가 효과적임을 제시하였다[6]. A. A. Ali 등은 미리 계산된 임의의 수학 방정식을 기반으로 산란(scattering) LSB와 함께 RSA를 사용하는 기법을 제시하였다. 스테가노그래피와 공개키 암호화를 2단계로 구현하는 보안 기술을 제안하였다. PSNR 값과 히스토그램의 향상된 결과를 제시하였다. RSA 암호 알고리즘을 통해 다단계 암호화 기법을 도입한 후 커버 매개체 비트의 위치와 값을 선택하기 위해 무작위 위치 지정을 적용하여 개선된 정보 은닉 방법을 제시하였다[2]. M. Grzelak 등은 스테간 분석을 위해 HUGO(highly undetectable stego) 등의 삽입 알고리즘에 의해 도입된 왜곡을 감지하도록 새로운 CNN 구조를 제시하였으며, 대규모 데이터 세트에서 학습된 구조가 총 오류율 측면에서 효과적인 성능을 보일 수 있었다[3]. Ji는 블록의 크기가 홀수와 짝수일 경우에 수식을 다르게 사용하는 것이 타당하며, 암호화와 선택적 셔플링을 함께 적용함으로 임의성(randomness)을 높이면서 보안성 및 저항성을 강화시킬 수 있음을 제시하였다[7]. 확산성을 높이기 위해 가변적인 ShiftRows와 LSB 기반의 이미지 스테가노그래피에서 단계별 하이브리드 방식을 적용하며, 한글 메시지를 은닉하는 방법을 제안한다.

3. 제안된 방법

스테가노그래피 구현에서 중요한 조건은 숨기려는 정보를 변형하지 않고 커버 매개체에 삽입하는 것이다. 암호화와 무작위화 과정을 이용하여 확산성을 높이고 저항성을 향상시키며, 제3자에게 숨겨진 비밀정보의 존재가 인지되지 않도록 조치하는 것이다. 한글 정보는 3가지 구성 요소인 초성, 중성, 종성자로 구성되어 있으며, 각각의 음절 요소는 사용 빈도에 따라[8] 표 1과 같이 표현될 수 있다. 괄호 안의 숫자는 사용 빈도이다. 분리된 각각의 음절 요소는 n비트로 표현되는 이진화 정보로 대응시킨다. 예를 들어 글자 '강'에서 각각의 음절 요소를 $n=4$비트로 표현된 정보에 대응시킬 경우, 결과는 '011110001011'이다.

표 1. 한글 음절 요소의 사용 빈도(%)
Table 1. Frequency of use of Hangul syllable elements(%)

Choseong	ㅇ(21.4), ㄱ(11.3), ㅅ(9.2), ㅈ(8.3), ㄷ(8.0), ㄴ(7.3), ㅎ(6.8), ㄹ(6.7), ㅁ(5.6), ㅂ(4.8), ㅊ(2.8), ㅌ(2.2), ㅋ(1.6), ㅍ(1.5), ㄲ(1.1), ㄸ(0.6), ㅉ(0.5), ㅆ/ㅃ(0.4)
Jungseong	ㅏ(22.5), ㅣ(16.8), ㅡ(10.8), ㅗ(10.6), ㅓ(9.6), ㅜ(6.6), ㅐ(5.3), ㅔ(4.8), ㅕ(4.0), ㅛ(1.6), ㅘ(1.4), ㅑ(1.2), ㅢ(1.0,), ㅝ(0.9), ㅠ(0.8), ㅚ(0.8), ㅟ(0.6), ㅒ(0.4), ㅙ(0.2), ㅞ(0.1), ㅖ(0.0)
Jongseong	null(56.5), ㄴ(14.1), ㄹ(7.7), ㅇ(7.0), ㄱ(4.1), ㅁ(3.7), ㅂ(1.8), ㅆ(1.6), ㅅ(1.4), ㅎ(0.3), ㅌ(0.3), ㄶ(0.3), ㅄ(0.3), ㅍ(0.2), ㅈ(0.2), ㄷ(0.2), ㅋ(0.1), ㅊ(0.1), ㄺ(0.1), ㄻ(0.1), ㄲ(0.1), ㄽ(0.0), ㅈ(0.0), ㄾ(0.0), ㅊ(0.0), ㄳ(0.0), ㄵ(0.0), ㄼ(0.0)

한글 음절 요소에 대응시킨 n비트 정보를 암호화한 후 열 단위와 행 단위에서 각각의 가변적인 이동 작업을 이용하는 하이브리드 혼합 과정을 적용한다.

3.1 가변 시프팅 과정

한글 정보로부터 분리한 초성자(s_1), 중성자(s_2), 종성자(s_3)에 각각의 n비트에 대응된 자료와 패딩 정보(s_4)의 자료, $n*4$는 암호화 단계를 적용한 후, 임의로 상태 행렬($S_{4\times 4}$)에 열 단위로 삽입한다.

[1단계]는 열 단위 섞음 과정이다. PRNG에 의해 생성된 순서에 따라 ⓪부터 ⑮까지 4개의 다른 영역으로 s_1, s_2, s_3, s_4 정보를 임의로 배치한다. 예

를 들어 2413을 적용할 경우 그림 1의 왼쪽과 같이 자료를 이동 배치(s_2, s_4, s_1, s_3)한다.

[2단계]는 행 단위 이동 과정이다. 추가 보안을 위해 행 단위 중심으로 ShiftRows를 이용하여 ⓪부터 ⓪까지 4개의 다른 영역으로 가변적 이동 작업을 적용한다. 이때 shifting 하는 순서 및 방향은 임의로 설정한다. 예를 들어 오른쪽에서 왼쪽으로 가변적 차등 이동, 2013을 적용할 경우 그림 1의 오른쪽과 같이 행 단위 자료가 이동 재배치된다.

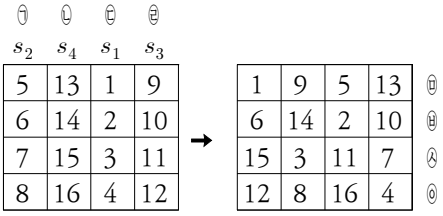

그림 1. 열 단위 뒤섞음과 ShiftRows를 이용한 행 단위 이동

Fig. 1. Column-wise shuffling and row-by-row shift using ShiftRows

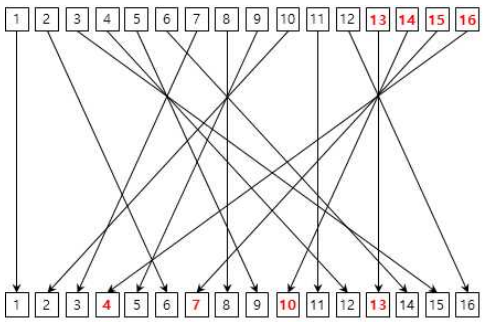

그림 2. ShiftRows를 사용한 이동 과정

Fig. 2. Shifting process using ShiftRows

비밀 자료를 암호화한다. 1단계를 적용한 후 2단계 가변적 이동에서 정보가 배치되는 과정은 그림 2에서 보여준다.

제안한 방법에서 행 단위 시프팅 연산은 이진화된 정보가 원하는 비트 수만큼 오른쪽에서 왼쪽 혹은 반대 방향으로 이동하면서 수행된다. 이때 수준 위치에 따라 이동 단위를 다르게 하는 가변적 방법을 적용한다. 가변적인 하이브리드 방법에 의해 원하는 양만큼 이동 관리와 임의성을 추가할 수 있기 때문에 제3자가 은닉된 비밀정보를 인지하는데 불가능한 수준으로 되어 저항성이 크게 향상될 수 있다.

3.2 삽입하는 과정

한글 정보를 음절 요소 단위로 분해한 후 대체되는 n*4비트 정보를 커버 매개체의 특정 위치에 다음의 단계별 과정을 적용하여 은닉한다.

1단계:커버 매개체인 이미지와 숨기려는 한글 메시지를 선택한다.

2단계:은닉시점, 세션키, n, 은닉 비트 크기(b) 등의 모수를 준비한다.

3단계:숨기려는 한글 정보를 음절 요소 단위로 분리한 후 [시프팅 과정]을 적용한다.

4단계:3단계 결과를 커버 이미지 RGB 각 채널의 b-LSB에 대체한다.

5단계:커버 이미지에 비밀정보의 끝부분까지 3단계와 4단계를 반복한다.

6단계:비밀정보를 숨기는 과정이 종료되는 다음 위치부터 종료 정보(e)를 k회 삽입한다.

7단계:완성된 스테고 이미지를 수신자에게 송신한다.

3.3 복원하는 과정

획득된 스테고 매개체로부터 숨겨진 비밀정보를 얻어내는 순서는 정보를 삽입하는 방법의 반대과정이다.

1단계:정보를 복원하기 위해 필요한 모수(b, n, 은닉시점, e, 세션키 등)를 준비한다.

2단계:수신된 스테고 이미지를 획득한다.

3단계:스테고 이미지에서 $b-$LSB의 비트로부터 정보를 추출하며, 종료 정보(e)를 확인한다.

4단계:수집된 비트 정보, n*4로부터 구성 요소를 재구성한다.

4.1 [시프팅 과정]과 반대되는 과정을 적용한다.

4.2 n*4씩 구분하여 대응되는 3개의 구성요소에 해당하는 음절을 재구성하여 해당 문자로 변환한다. 이때 패딩 자료를 제외시킨다.

5단계:커버 이미지에서 종료 정보를 찾을 때까지 3단계와 4단계를 반복한다.

6단계:n*3단위로 재구성된 정보를 가지고 비밀정보를 획득한다.

4. 적용 및 결과

커버 이미지에서 RGB 각 채널의 화소값을 참고한다. b-LSB에 정보를 숨기기 위해 한글 정보로부터 3가지 음절 요소를 분해한 후 $n=4$ 비트로 구성된 대응 정보에 대체하였다. 패딩 요소를 추가한다. n*3+4(padding)=16비트로 구성된 정보는 스트림 암호화를 적용한 후 상태 행렬($S_{4\times4}$)에 임의로 배치한다. 열 단위에서 '2431'을, ShiftRows를 이용한 행 단위의 가변적 이동을 위해 오른쪽에서 왼쪽, '1302'를 각각 설정하였다. e=1111($k=2$)을 사용하였다. 커버 이미지는 lena.gif(227 Kbyte)를 사용하였다. 정보를 숨기는 LSB 영역에서 $b=1,2,3,4,5$ 각각의 경우에 이미지 품질을 확인하는 PSNR을 계산하였다. 이를 위해 수식 (1)을 이용하였다.

$$PSNR = 10 \cdot \log_{10}\left(\frac{L^2}{MSE}\right)(dB) \qquad (1)$$

여기에서 최대 신호 수준인 $L=255$를 사용하였으며, MSE는 평균제곱오차를 의미한다.

커버와 스테고 매개체 사이의 유사성을 평가하기 위해 계산된 PSNR과 MSE는 표 2에서 제시하였으며, 이때 동일 크기의 다른 비밀 메시지 12셋을 적용하였다.

표 2. 제안된 방법의 성능

Table 2. Performance of the proposed method

Bit plane	LSB		Ji(2022)		Proposed	
	MSE	PSNR	MSE	PSNR	MSE	PSNR
1st	0.493	51.199	0.507	51.083	0.484	51.279
2nd	2.326	44.464	2.229	44.651	2.281	44.549
3rd	10.192	38.048	10.491	37.923	11.276	37.609
4th	21.839	34.738	31.260	33.181	33.659	32.860
5th	149.917	26.372	147.341	26.447	159.563	26.101

표 2에서와 같이 제안된 방법에서 PSNR 값은 기본 LSB, Ji의 결과와 비슷하며, MSE는 9.38%, 3.88% 각각 증가됨을 확인하였고, 상관계수가 0.9989로 유사성이 높음을 확인하였다. 정보은닉 전과 후의 RGB 값 변동에서 $b=1$일 때 50.7%, $b=3$일 경우 10.7%의 일치도를 보였다. PSNR은 기준값[9]보다 23.7% 높게 나타남을 확인하였다. 고도화 및 지능화되어지는 스테간 분석에 대응하기 위해 계층별 혼돈성이 반영되는 즉, 분리된 한글 정보와 패딩 정보를 임의 순서로 열 단위에 의한 섞는 과정, 상태 행렬에서 ShiftRows를 이용한 행 단위 이동 과정을 계층별 적용함으로 불규칙성을 강화시킬 수 있다.

5. 결론

숨기려는 한글 정보를 분리한 후 3가지 음절 요소에 대응된 자료와 패딩 정보를 암호화한 후 단계별 확산을 이용한 하이브리드 뒤섞음 과정을 적용함으로 확산성을 높이고 저항성과 보안성을 향상시킬 수 있는 대안임을 확인하였다. 즉, 기본 LSB와 비교할 때 PSNR 값은 0.9972의 유사도를 보였으며, 은닉 비트 크기가 작을 경우 0.17% 증가, 3이상일 경우 2.53% 감소되어 비슷함을 확인하였다. 제안된 방법이 기본 LSB와 유사한 이미지 품질을 유지하면서 확산과 임의성을 증가시키는 효과적인 방법임을 확인하였다.

REFERENCES

[1] O. C. Abikoye, R. O. Ogundokun, S. Misra, A. Agrawal, "Analytical Study on LSB - Based Image Steganography Approach", *Computational Intelligence in Machine Learning*, pp. 451-457, 2022.

[2] Y. M. Wazery, S. G. Haridy, A. A. Ali, "A Hybrid Technique based on RSA and Data Hiding for Securing Handwritten Signature", *International Journal of Advanced Computer Science and Applications*, Vol. 12, No. 4, pp. 726-735, 2021.

[3] K. Lichy, P. Lipinski, M. Grzelak, "Deep Convolutional Network for Steganalysis of HUGO, WOW and UNIWARD algorithms", *International Conference on Control, Automation, Robotics and Vision*, pp. 737-740, 2020.

[4] R. Dumre, A. Dave, "Exploring LSB Steganography Possibilities in RGB Images", *International Conference on Computing Communication and Networking Technologies*, 2021.

[5] V. Sivaranjani, V. Bhuvaneshwari, "Hybrid Approach for Data & Image Encryption using LSB, RDH and AES Algorithm", *International Journal of Engineering Research&Technology*, Vol. 3, Issue 11, pp. 740-744, 2014.

[6] T. S. Gupta, R. Srivastava, "Advanced -Data Encryption using Three- Layered Hybrid Cryptosystem and Secured Key Storing using Steganography", *International Research Journal of Engineering and Technology*, Vol. 8, Issue 8, pp. 3767-3773, 2021.

[7] S. S. Ji, "Selective Shuffling for Hiding Hangul Messages in Steganography", *The Korea Institute of Information Electronic Communication Technology,* Vol. 15, No. 3, pp. 8-14, 2022.

[8] Information Design, "Hangul Grapheme Frequency and Keyboard Heatmap", *https://story.pxd.co.kr/958*, Oct. 2014.

[9] C. K. Chan, L. M. Cheng, "Hiding Data in Images by Simple LSB Substitution", *The Journal of the Pattern the Recognition Society,* Vol. 37, pp. 469-474, 2004.

저자약력

지 선 수 (Seon-Su Ji) **[종신회원]**

- 충남대학교 계산통계학과(학사)
- 중앙대학교 응용통계학과(석사)
- 중앙대학교 응용통계학과(박사)
- 명지대학교 컴퓨터공학과(박사수료)
- (현)강릉원주대학교 컴퓨터공학과 교수

〈관심분야〉 정보보안(정보은닉), 스테가노그래피