한국정보전자통신기술학회논문지(jkiiect)'22-6, Vol.15 No.3
논문 22-15-03-211 http://dx.doi.org/10.17661/jkiiect.2022.15.3.211

스테가노그래피에서 한글 메시지 은닉을 위한 선택적 셔플링

지선수*

Selective Shuffling for Hiding Hangul Messages in Steganography

Seon-su Ji*

요 약 스테가노그래피 기술은 커버 매체의 특정 위치에 비밀 메시지를 대체시켜 숨겨진 정보의 존재를 추적할 수 없도록 보호 조치를 한다. 암호화와 스테가노그래피를 기반으로 다양한 복합적인 방법을 적용하여 보안성과 저항성을 강화한다. 특히 보안성을 향상시키기 위해 혼돈과 무작위성을 높이는 기법이 필요하다. 실제로 이산코사인변환(DCT)과 최하위 비트(LSB) 기반에서 셔플링 방식이 적용된 경우는 연구가 진행되어야 할 영역이다. 메시지 숨김의 복잡성을 추가할 수 있는 비트 정보 셔플링 방식을 통합하고, 공간 영역 기법을 스테가노그래피에 적용하여 한글 메시지의 비트 정보를 은닉하는 새로운 접근 방법을 제시한다. 메시지를 추출할 때 역셔플링을 적용한다. 이 논문에서, 삽입하려는 한글 메시지를 초성, 중성, 종성으로 분리한다. 대응된 정보에 기반한 선택적 셔플링 과정을 적용하여 보안성과 혼돈성을 향상시킨다. 제안된 방법의 성능을 확인하기 위해 상관계수와 PSNR을 이용하였다. 기준값과 비교했을 때 제안한 방법의 PSNR 값이 타당하다는 것을 확인하였다.

Abstract Steganography technology protects the existence of hidden information by embedding a secret message in a specific location on the cover medium. Security and resistance are strengthened by applying various hybrid methods based on encryption and steganography. In particular, techniques to increase chaos and randomness are needed to improve security. In fact, the case where the shuffling method is applied based on the discrete cosine transform(DCT) and the least significant bit(LSB) is an area that needs to be studied. I propose a new approach to hide the bit information of Hangul messages by integrating the selective shuffling method that can add the complexity of message hiding and applying the spatial domain technique to steganography. Inverse shuffling is applied when extracting messages. In this paper, the Hangul message to be inserted is decomposed into the choseong, jungseong and jongseong. It improves security and chaos by applying a selective shuffling process based on the corresponding information. The correlation coefficient and PSNR were used to confirm the performance of the proposed method. It was confirmed that the PSNR value of the proposed method was appropriate when compared with the reference value.

Key Words : Hangul Data Hiding, Image Steganography, LSB, Selective Shuffling, Pixel Decomposition

1. 서론

네트워크에서 다양한 기술의 발달로 우리가 사용하는 대부분의 디지털 데이터는 인터넷을 통해 송신 및 수신될 수 있다. 비밀 통신에서 전송되는 메시지는 내부 및 외부 취약점 등에 의해 제3자로부터 다양한 공격 문제를 겪을 수 있다. 스테가노그래피와 암호화를 기반으로 하여 다양한 하이브리드

*Department of Computer Sciences&Engineering, Gangnung-Wonju National University
Received May 02, 2022 Revised May 19, 2022 Accepted May 27, 2022

방법을 적용하여 보안성과 저항성을 강화한다. 즉 불법 접근으로부터 비밀 메시지를 보호하기 위해 암호화, 교차, 분해, 셔플링, 조합 등 다양한 혼합 방법으로 진화되고, 발전되어 왔다. 메시지의 은닉 기술인 스테가노그래피는 포함된 정보의 존재를 감지할 수 없도록 커버 매체에 비밀 정보를 숨긴다. 텍스트, 이미지, 오디오, 비디오, 매체들의 결합과 같은 다양한 디지털 미디어 플랫폼을 커버 매체로 사용할 수 있다. 실제적으로 시각 시스템에서 인간의 인지 능력 한계, 중복되는 정보를 다수 포함하여 충분한 삽입 용량을 확보할 수 있는 디지털 이미지가 커버 매체로 폭넓게 사용되고 있다[1].

일반적인 이미지 은닉 기술은 삽입하고자 하는 비트화된 메시지 정보에서 정교한 셔플링 방식을 고려하지 않고, 커버 이미지의 비트화된 특정 위치에 비트 정보를 삽입한다. 삽입 용량, 지각 불가능성, 견고성 및 저항성을 향상시킬 수 있는 스테가노그래피 구조를 설계하기 위해 새로운 요소가 필요하다. 이 논문에서는 메시지 숨김의 복잡성을 추가할 수 있는 변형된 셔플링 방식을 통합하여 공간 영역 기법을 스테가노그래피에 적용하고, 한글 메시지의 비트 정보를 은닉하는 새로운 접근 방법을 제시한다.

논문의 2장에서 셔플링 과정과 LSB(least significant bit)의 적용 기법 등과 관련된 자료를 제시하였다. 제안하고자 하는 방법은 3장에서 표현하였다. 제시된 방법의 타당성을 확인하기 위한 적용된 결과를 4장에서 설명하였다. 논문의 마무리는 5장에서 제시하였다.

2. 관련 연구

스테가노그래피 방법의 성능을 최적화하기 위해 다양한 방법을 적용할 수 있다. 이미지 스테가노그래피 기술은 LSB 대체와 LSB 매칭 등을 이용하는 공간 영역 기술, DCT(discrete cosine transform)와 DWT(discrete wavelet transform) 기반의 변환을 이용하는 주파수 영역 기술, 스펙트럼 확산 기술, 벡터 양자화에 기반한 압축 도메인 기술, 왜곡 기술과 통계 기술 등으로 분류되어 설명된다. DCT 및 양자화 단계 이후에 셔플링 방식을 도입하여 이미지 내부에 정보를 숨기는 방식으로 접근하는 방법이 효과적일 수 있다[2]. Das 등은 이산코사인 변환 작업의 전 단계에서 서로 다른 방식으로 셔플링 비트 문자열을 사용하여 본질적으로 메시지 숨김을 통합하는 것을 목표로 하는 방법을 제시하였다. 보안 수준을 향상시키기 위해 메시지 셔플링 방식을 제안하고, 적용된 이미지 스테가노그래피 방법이 기존의 표준화된 방법보다 성능이 우수할 수 있음을 보여주었다[3]. Aggarwal 등은 효율적이고 안전한 비밀 통신을 위해 AES 암호, 픽셀 관리, 서플링 방법을 적용하여 이미지 품질을 보장하는 방법을 제시하였다[4]. Jan 등은 공간 영역을 이용한 그레이스 케일 영상에 대해 PBSA(pattern based bits Shuffling algorithm)와 MLSB(matching LSB)를 기반으로 하는 감지할 수 없는 이미지 스테가노그래피 기법을 제안하였다. 비밀 정보는 스테고 키와 PBSA를 기반으로 암호화된 다음, MLSB 방식을 사용하여 커버 이미지에 삽입되어 전체 이미지 내부에 비밀 데이터를 흩어지게 하므로 공격자가 상대적으로 추출하기 어렵게 한다는 것을 제시하였다. 제안된 방법에서 계산된 최대 신호 대 잡음비(peak signal to noise ratio, PSNR)는 허용 가능한 수치보다 높게 나타난다는 것을 보였다[5]. Abood는 RC4 스트림 암호, RGB 픽셀 셔플링과 스테가노그래피를 이용하여 안전한 암호화와 복호화를 보장하기 위해 해시 함수를 사용하는 HLSB(hash LSB)를 가지고 커버 이미지의 RGB 픽셀의 LSB에 비밀 자료 비트를 삽입하는 방법을 제시하였다. 이미지의 품질에 영향을 미치지 않으면서 비밀 자료를 암호화하여 삽입할 수 있는 효과적인 방법임을 확인하였다[6]. Zhao 등은 커버 매체에서 허용 가능한 재구성 오류에 기반한 받아들일 수 있는 정교한 셔플링 체계를 바탕으로 효과적인 은닉 기법을 제안하였다. 데이터의 낮은 순위 구조를 활용하여 숨기려는 자료의 크기를 증가시킬 수 있음을 제시하였다[7]. 실제로 DCT와 LSB에 기반한 다양한

셔플링 방식이 적용되어 혼돈성을 강화시키는 다층 스테가노그래피는 많은 연구가 진행되어야 할 영역이다. 논문에서 LSB 기반의 이미지 스테가노그래피에서 셔플링/역셔플링 방식이 적용되어, 한글 메시지를 숨기는 새로운 방법을 제시한다.

3. 제안된 방법

스테가노그래피 기술은 커버 매체의 특정 위치에 비트화된 한글 메시지를 숨기며, 제3자에 의해 숨겨진 정보의 존재를 추적할 수 없도록 보호 조치를 한다. 보안성을 향상시키기 위해 한글 메시지의 음절 분리와 대체, 암호화, 셔플링 단계, l (= 1,2,3, ...)번째 가상 비트 평면을 이용하여 비밀 메시지를 은닉하는 이미지 스테가노그래피 방법을 제시한다. 비밀 메시지로 사용할 한글의 음절 요소는 표 1과 같이 표시할 수 있다. 분해된 음절 정보를 대체하기 위한 $b(=2,3,4,...)$ 비트를 적용하며, 비트 정보를 저장하는 형태에 따라 선택할 수 있다.

표 1. 사용 빈도에 기반한 비밀 메시지의 음절 요소
Table 1. Syllable elements of secret messages based on frequency of use

Choseong	ㅇ, ㄱ, ㅅ, ㅈ, ㄷ, ㄴ, ㅎ, ㄹ, ㅁ, ㅂ, ㅊ, ㅌ, ㅋ, ㅍ, ㄲ, ㄸ, ㅉ, ㅆ, ㅃ
Jungseong	ㅏ, ㅣ, ㅡ, ㅗ, ㅓ, ㅜ, ㅔ, ㅐ, ㅕ, ㅛ, ㅘ, ㅑ, ㅢ, ㅝ, ㅠ, ㅚ, ㅟ, ㅒ, ㅖ, ㅞ
Jongseong	‘ ‘, ㄴ, ㄹ, ㅇ, ㄱ, ㅁ, ㅂ, ㅆ, ㅅ, ㅎ, ㅌ, ㄶ, ㅄ, ㅍ, ㅈ, ㄷ, ㅋ, ㅊ, ㄺ, ㄾ, ㄲ, ㄻ, ㄵ, ㅐ, ㄿ, ㄳ, ㄽ, ㅀ

3.1 셔플링 과정

보안성을 증가시키기 위한 방법에서 혼돈과 무작위성을 높이는 기법이 필요하다. 이러한 이유로 암호화 및 스테가노그래피에서 셔플링 방식이 사용되어, 한글 메시지의 비트화된 정보를 뒤섞는 작업을 적용한다. 선형 매핑 $R_i : R \rightarrow R$ 연결로서 셔플링 과정을 표현할 수 있으며, 모델링의 불확실성을 높이기 위해 적용되는 순열을 고려한 다중 프로세서 구조에서 비트열 셔플링을 참조[3] 하였다. 이 논문에서 p블록의 q-셔플링 방식으로, 다음 수식으로 계산하는 것을 제시한다.

$$Su(i) = q \bullet i - 1 \bmod p, \quad \text{if } p \text{ is odd} \qquad (1)$$

$$q \bullet i \bmod p - 1, \quad \text{if } p \text{ is even} \qquad (2)$$

$$i = 1, 2, 3, ...$$

여기에서 p는 블록의 길이를 나타내며, q는 $p(=b*3)$에 따라 주어지는 상수이다.

표 2. 셔플링/역셔플링 적용 가능한 수 (가능횟수/시도횟수)
Table 2. Number of applicable Shuffling/inverse Shuffling (number of possible/number of attempts)

p	Das et. al.(2018)	Proposed-(1)
7	2/6	**6/6**
8	**6/7**	3/7
9	4/8	**6/8**
10	**6/9**	4/9
11	4/10	**10/10**
12	**10/11**	4/11

표 2에서와 같이 셔플링과 역셔플링 방법을 다양하게 선택할 수 있으며, p가 홀수와 짝수인 경우에 따라 수식 (1)과 (2)를 다르게 적용하는 것이 혼돈성 측면에서 효과적임을 확인할 수 있다.

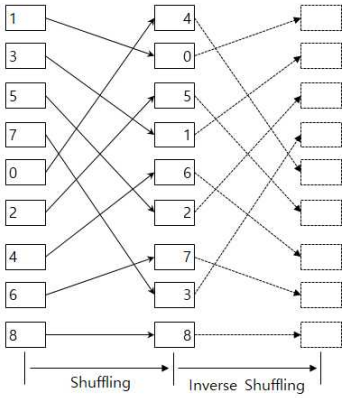

그림 1. p=9, q={2,5}가 주어진 경우 셔플링/역셔플링 과정
Fig. 1. Shuffling/Inverse Shuffling process given p=9, q={2,5}

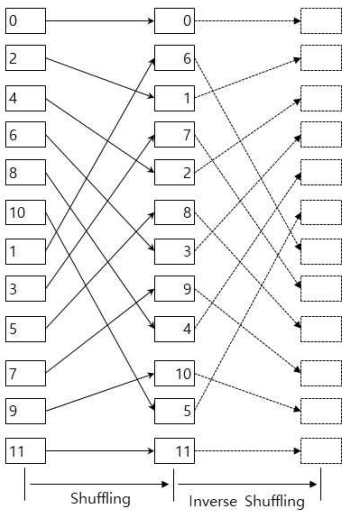

그림 2. p=12, q={2,6}가 주어진 경우 셔플링/역셔플링 과정
Fig. 2. Shuffling/Inverse Shuffling process given p=12, q={2,6}

$p=9$일 때 수식 (1)을 이용하여 계산된 결과, $q=\{1,1\}, \{2,5\}, \{4,7\}, \{8,8\}$을 각각 사용할 수 있다. 예를 들어 $q=\{2,5\}$의 경우 2-셔플링과 5-역셔플링을 적용할 수 있으며, 반대인 상황에서도 이용할 수 있다. 그림 1에서 셔플링과 역셔플링의 적용 과정을 보여주었다.

$p=12$일 때 수식 (2)를 적용한 계산 결과, $q=\{1,1\}, \{2,6\}, \{3,4\}, \{5,9\}, \{7,8\}, \{10,10\}$을 각각 사용할 수 있다. 예를 들어 $q=\{2,6\}$의 경우 2-셔플링과 6-역셔플링을 이용할 수 있으며, 반대인 상황에서도 적용할 수 있다. 그림 2에서 셔플링과 역셔플링의 적용 과정을 보여주었다.

한글 메시지 글자를 초성, 중성, 종성으로 분리한 후 대체된 홀수($b=1,3,...$) 비트열 정보를 사용하기 위해서는 수식 (1)을, 짝수($b=2,4,...$) 비트열 정보를 사용하기 위해서는 수식 (2)를 각각 사용하는 것이 타당하다는 것을 확인하였다.

3.2 숨기는 과정

선택된 커버 이미지의 가상 비트 평면의 특정 위치에 다음의 단계별 과정을 적용하여 비밀 메시지를 숨길 수 있다.

1단계:커버 이미지, 숨기려는 한글 메시지, 매개변수(k, b, p, q, l) 등을 준비한다.

2단계:숨기려는 글자를 음절 구조로 분리한다.

2.1 초성(h_1), 중성(h_2), 종성(h_3)으로 분리한 후 b비트에 대응되는 정보로 대체한다. 키(k)를 이용하여 암호화 과정을 적용($h_1' h_2' h_3'$) 한다.

2.2 $h_1' h_2' h_3'$의 $b*3$ 비트에 대한 선택적 셔플링 작업을 적용한다. 그림 1과 그림 2를 참조한다.

3단계:커버 이미지로 부터 RGB 화소 값에 각각의 비트화된 정보로 변환한다.

4단계:가상 비트 평면(l)에 2.2의 변환된 비트 정보를 대체한다.

5단계:숨기려는 정보가 커버 이미지에 모두 대체될 때까지 2단계부터 4단계 과정을 반복한다.

6단계:스테고 이미지를 완성한다.

3.3 추출하는 과정

스테고 매체로부터 숨겨진 비밀 정보를 추출하

는 것은 정보를 숨기는 방법의 역과정이다.

1단계:스테고 이미지, 매개변수 (k, b, p, q, l) 정보를 획득한다.

2단계:스테고 이미지의 은닉 시점에서 비트 정보를 선택한다.

2.1 스테고 이미지의 각각의 RGB 화소 값에서 l 번째 비트 정보를 획득한다.

2.2 b*3 비트 단위로 선택적 역서플링과 복호화 과정을 적용한다.

2.3 2.2에서 분리된 b비트에 대응되는 정보를 획득한다.

2.4 획득된 정보를 조합하여(b*3) 문자를 재구성한다.

3단계:은닉 종점까지 2단계 과정을 반복한다.

4단계:숨겨진 메시지를 완성한다.

4. 적용된 결과

숨기려는 한글 메시지의 음절을 3개의 요소로 분해한 후 $b=3$ 비트로 구성된 대응 정보에 대체하였다. 또한 스트림 암호화 방법을 적용한 후 b*3 비트로 구성된 정보를 가지고 셔플링 절차를 적용하였다. $\alpha=0.05(0.1)$을 이용한다. 한글 메시지 구조의 특수성 때문에 홀수인 경우 즉, $p=9$에서 $q=\{4,7\}$, 4-셔플링, 7-역셔플링을 적용하였다. $l-LSB$를 적용하기 위해 비트 평면 위치를 적용한다. 비밀 메시지는 '반도체설계기술독점인수제동우려혜택경쟁촉진'(42byte)을 이용하였다.

일반 LSB 등의 방법과 비교하기 위해 $l=1,2,3,4,5$의 각각의 경우에 이미지 품질을 비교하기 위해 PSNR과 상관계수를 확인하였다. 즉 제안된 방법의 성능을 측정하기 위해 PSNR을 사용하며, 수식 (4)을 이용하였다. 여기에서 H은 커버 매체의 높이이며, W는 커버 매체의 너비이며, L은 사용된 매체의 신호 수준을 의미하며, 최대값인 255를 사용하였다. C가 커버 매체일 때 스테고 매체는 $S=C+\alpha\sum_{i=1}R_i(M_i)$으로 표현할 수 있다. 이때 $R_i(M_i)$는 i번째 셔플링된 비밀 메시지를 의미하며, α는 보안과 효율성 간의 상충관계를 위한 강도를 관리하는 스칼라 값이다.

$$MSE=\frac{1}{HW}\sum_{x=1}^{H}\sum_{y=1}^{W}(S_{xy}-C_{xy})^2 \quad (3)$$

$$PSNR=10\cdot\log_{10}(\frac{L^2}{MSE})\,(dB) \quad (4)$$

커버 매체와 스테고 매체의 이미지 품질과 상관성 등의 성능을 확인하기 위해 계산된 PSNR과 상관계수는 표 3에서 결과를 제시하였다.

표 3. 제안된 방법의 결과 (PSNR)
Table 3. Results of the proposed method (PSNR)

Bit plane(l)	LSB		Das et.al. (2018)		Proposed	
	PSNR	Corr.	PSNR	Corr.	PSNR	Corr.
1st	50.403	0.9998	50.683	0.9998	50.921	0.9998
2nd	44.962	0.9993	45.284	0.9994	45.632	0.9994
3rd	38.700	0.9974	38.819	0.9975	39.132	0.9976
4th	33.448	0.9913	33.310	0.9914	32.921	0.9902
5th	26.659	0.9614	26.778	0.9633	27.090	0.9637

제안된 방법이 R, G, B 각각의 화소 정보에 1비트 정보 단위로 대체하기 때문에 PSNR이 높고, 상관계수가 1.0에 가깝게 나타나는 것을 확인하였다. 또한 한글은 3개의 구성 요소가 결합된 언어 구조이므로 블록의 크기(b*3)가 홀수일 경우 제안된 수식을 사용하는 것이 다양성 및 복잡성을 확보할 수 있음을 확인하였다. 제안된 방법에서 일반 LSB와 Das 등의 방법보다 PSNR이 0.8%와 0.4% 각각 높게 나타났으며, 상관계수는 비슷한 값을 보여주었다. PSNR은 기준치[8]보다 26.23% 높게 나타남을 확인하였다.

5. 결론

제안된 방법에서 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 확인하였다.

REFERENCES

[1] Y. Y. Wai, E. E. Myat, "Comparison of LSB, MSB and New Hybrid (NHB) of Steganography in Digital Image", *International Journal of Engineering Trends and Applications(IJETA),* Vol. 5, Issue 4, pp. 16-19, 2018.

[2] K. Peng, B. Feng, "A shuffling scheme with strict and strong security", *Proc of Fourth IEEE International Conference on Emerging Security Information Systems and Technologies (SECURWARE)*, pp. 201-206, 2010.

[3] P. Das, K. Chakraborty, S. Sinha, A. Das, "A New Image Steganography Method using Message Bits Shuffling", *Journal Mech. Cont. & Math. Sci.*, Vol. 13, No. 5, pp. 1-15, 2018.

[4] V. Sharma, U. Srivastava, S. Aggarwal, "Image Steganography Using Pixel Manipulation and Shuffling", *International Journal of Scientific & Engineering Research,* Vol. 10, Issue 4, pp. 956-958, 2019.

[5] K. Muhammad, J. Ahmad, H. Farman, Z. Jan, "A New Image Steganographic Technique using Pattern based Bits Shuffling and Magic LSB for Grayscale Images", *Sindh Univ. Res. Jour.*, Vol. 47, No. 4, pp. 723-728, 2015.

[6] M. H. Abood, "An Efficient Image Cryptography using Hash-LSB Steganography with RC4 and Pixel Shuffling Encryption Algorithms", *Annual Conference on New Trends in Information&Communications Technology Applications(NTICT 2017),* pp. 86-90, 2017.

[7] Z. Sun, C. Li, Q. Zhao, "Hide Chopin in the Music: Efficient Information Steganography via Random Shuffling", *ICASSP*, pp. 2370-2374, 2021.

[8] C. K. Chan, L. M. Cheng, "Hiding Data in Images by Simple LSB Substitution", *The Journal of the Pattern the Recognition Society 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