Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.30, NO.1, Feb. 2020 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2020.30.1.39

블록 암호 SM4에 대한 부채널 공격 및 마스킹 기반 대응기법 분석*

배 대 현,[1†] 남 승 현,[2] 하 재 철[3‡]
[1,3]호서대학교(학생, 교수), [2](주) 라닉스(연구소장)

Side Channel Attack on Block Cipher SM4 and Analysis of Masking-Based Countermeasure*

Daehyeon Bae,[1†] Seunghyun Nam,[2] Jaecheol Ha[3‡]
[1,3]Hoseo University(Student, Professor), [2]RANIX Co., Ltd(Chief Technical Officer)

요 약

본 논문에서는 중국 표준 블록 암호 알고리즘인 SM4가 부채널 공격에 취약함을 보이고 그에 대한 대응책을 제안하고자 한다. 먼저, SM4는 차분 전력 분석(DPA)과 상관 전력 분석(CPA)에 기반한 공격에 의해 쉽게 비밀 키가 노출됨을 확인하였다. 논문에서는 공격 취약 요소를 분석하고 데이터 마스킹에 기반한 전력 분석 공격 대응 기법을 설계하였다. 제안한 SM4에 대한 1차 마스킹 기법은 딥 러닝 기반의 다층 퍼셉트론(MLP) 모델을 이용한 공격 프로파일링(profiling) 기반 공격에는 여전히 취약하지만, 차분 전력 분석이나 상관 전력 분석과 같은 비프로파일링(non-profiling) 공격에는 충분히 대응할 수 있음을 확인하였다.

ABSTRACT

In this paper, we show that the Chinese standard block cipher SM4 is vulnerable to the side channel attacks and present a countermeasure to resist them. We firstly validate that the secret key of SM4 can be recovered by differential power analysis(DPA) and correlation power analysis(CPA) attacks. Therefore we analyze the vulnerable element caused by power attack and propose a first order masking-based countermeasure to defeat DPA and CPA attacks. Although the proposed countermeasure unfortunately is still vulnerable to the profiling power attacks such as deep learning-based multi layer perceptron(MLP), it can sufficiently overcome the non-profiling attacks such as DPA and CPA.

Keywords: Side-Channel Analysis, Power Analysis Attack, DPA, CPA, Data Masking, MLP

I. 서 론

두 통신자간 주고 받는 정보에 대한 기밀성을 제공하기 위한 방법으로 암호화 기법을 사용하고 있으며 지금까지 수많은 블록 암호 알고리즘들이 제안되었다. 블록 암호화를 수행하기 위해서는 암호용 디바이스 내부에 비밀 키를 저장하고 이를 사용하여 메시지를 혼돈시키거나 확산시킨다.

대표적인 블록 암호 알고리즘은 국제 표준 암호 알고리즘인 AES[1], SEED[2], LEA[3] 등이 있다. 현재는 블록 암호를 사물 인터넷 장치 등에 쉽게 활용하기 위한 경량화, 고속화 연구가 지속적으로 진행되고 있다. 중국에서는 SM4[4, 5]라고 명명된 독자적인 블록 암호 알고리즘을 개발하여 중국 국가 표준(GB, Guojia Biaozhun)으로 지정하였으며, 이

Received(02. 03. 2020), Accepted(02. 14. 2020)
* 본 논문은 2019년도 한국정보보호학회 충청지부 학술대회에 발표한 우수논문을 개선 및 확장한 것임
† 주저자, noeyheadb@gmail.com
‡ 교신저자, jcha@hoseo.edu(Corresponding author)

알고리즘은 2017년 국제 표준화 기구(ISO)에서 ISO-IEC 18033-3 AMD2로 표준화가 진행된 상태이다.

한편, 암호용 디바이스에 암호 알고리즘을 구현하는 과정에서 발생하는 취약점으로 인해 비밀 키와 관련한 정보가 누설되기도 한다. 부채널 분석(side channel analysis) 공격이란 암호용 디바이스가 암호 연산을 수행할 때 얻을 수 있는 소비 전력, 전자기파, 시간 정보 등을 기반으로 비밀 키를 분석하는 공격 방법이다[6]. 부채널 분석 공격 중에는 디바이스 구동 시 발생하는 소비 전력을 오실로스코프와 같은 장비로 수집하여 이를 바탕으로 비밀 정보를 탐색하는 전력 분석 공격이 대표적이다[7].

전력 분석 공격에는 단순히 소비 파형만 분석하는 단순 전력 분석(Simple Power Analysis, SPA)[8]을 비롯하여 차분 전력 분석(Differential Power Analysis, DPA)[9], 상관 전력 분석(Correlation Power Analysis, CPA) 공격[10]으로 나누어진다. 이러한 공격은 공격 목표가 되는 디바이스와 그 디바이스로부터 측정한 전력 파형만 이용하므로 비프로파일링(non-profiling) 공격이라고 부른다. 최근에는 공격 대상이 되는 디바이스와 동일하거나 비슷한 사양을 갖는 다른 디바이스를 통해 공격용 프로파일을 생성하고, 실제 공격 대상 디바이스로부터 얻은 파형과 프로파일 간의 일치성을 비교 비교함으로써 비밀 키를 알아내는 프로파일링(profiling) 공격 방법이 많은 주목을 받고 있다.

본 논문에서는 중국 표준 블록 암호 알고리즘인 SM4를 개발용 보드에 그대로 구현하면 전력 분석 공격에 의해 비밀 키가 노출될 수 있음을 보이고자 한다. 또한, 전력 분석 공격에 대응할 수 있도록 SM4 알고리즘 특성과 추가 연산량을 고려한 마스킹 기법을 설계하여 제안하였다. 제안한 1차 마스킹 기법은 딥 러닝(deep learning) 기반의 다층 퍼셉트론(Multi Layer Perceptron, MLP) 모델[11-14] 등을 이용한 프로파일링 공격에는 여전히 취약하지만, DPA나 CPA와 같은 비프로파일링 공격에는 충분히 대응할 수 있음을 확인하였다.

II. 부채널 공격에 대한 배경 지식

2.1 전력 분석 공격

전력 분석 공격은 암호 알고리즘이 구축되어 있는 디바이스가 구동될 때 소비되는 전력을 바탕으로 비밀 키를 추출하는 공격 방법이다. 특히, 각 암호 라운드 연산이 수행되는 동안 비밀 키 값이 사용될 때의 전력 파형을 찾아내고 해당 지점의 파형과 중간 데이터간의 통계적 분석을 통해 비밀 정보를 추측하게 된다.

전력 분석 공격은 공격 방법에 따라 크게 비프로파일링 공격과 프로파일링 공격으로 구분할 수 있다. 먼저, 비프로파일링 공격은 암호용 디바이스에 랜덤한 평문을 입력하여 암호 연산을 수행하고, 이 과정에서 소비 전력을 측정하여 수집한 수십~수백 개의 파형들을 분석하여 비밀 키를 추출하는 공격으로서 오직 목표가 되는 디바이스만 공격에 사용된다. 비프로파일링 공격 중에서 DPA와 CPA 공격은 측정한 파형들을 기반으로 추측된 비밀 키에 의한 차분 값을 계산하거나 비밀 키와 파형간의 상관도를 구하여 비밀 키를 추출하는 매우 위협적인 공격 방법이다.

프로파일링 공격에서 공격자는 먼저 학습용 프로파일을 생성하기 위해 공격 대상 디바이스와 동일하거나 비슷한 사양을 갖는 다른 디바이스를 가지고 있어야 한다. 이 공격은 프로파일용 디바이스를 통해 얻은 정보와 실제 공격 대상 디바이스로부터 얻은 파형과의 매칭 확률을 비교함으로써 비밀 키를 알아내는 공격 방법이다. 기존 프로파일링 형태의 공격에는 템플릿 공격(Template Attack, TA)[15]이 있으며 최근 딥 러닝 모델의 일종인 MLP나 CNN(Convolutional Neural Network) 기반 전력 분석 공격[16, 17] 등이 있다. 최근의 연구 결과에 의하면 전력 분석 공격에 대응하기 위해 1차 마스킹 기법을 적용한 AES 암호 알고리즘도 MLP나 CNN과 같은 프로파일링 공격에는 비밀 키가 노출될 수 있음이 밝혀졌다[18, 19].

2.2 딥 러닝 기반 MLP

인간의 뇌 구조 즉, 생물학적 뉴런(neuron)의 구조를 바탕으로 컴퓨터에서 대량의 데이터를 처리할 수 있도록 설계한 딥 러닝 알고리즘을 인공 신경망(Artificial Neuron Network, ANN)이라고 하며 대표적으로 퍼셉트론(perceptron) 구조가 있다.

간단한 퍼셉트론은 입력 계층(input layer)과 출력 계층(output layer)을 가지며, 입력 계층에서 데이터를 입력받아 이를 가중치(weight)와 곱하고 바이어스(bias)를 더한 후, 이 값을 활성화 함수

(activation function)인 Step 함수에 입력하여 최종적인 값을 출력하게 된다. 그러나 퍼셉트론은 기본적으로 선형 이진 분류기 형태이기 때문에 XOR과 같은 비선형적 데이터에 대해서는 분류가 불가능하다는 한계가 존재한다. 따라서 단층 퍼셉트론의 한계를 극복하기 위해 입력 계층과 출력 계층 사이에 은닉 계층(hidden layer)을 두어 비선형적으로 분류되는 데이터에 대해서도 학습할 수 있도록 고안하였는데 이것이 다층 퍼셉트론, 즉 MLP이다. 다음 Fig. 1은 MLP 모델을 도식화한 것이다. 그림에서 보는 바와 같이 MLP는 입력 계층, 은닉 계층, 출력 계층으로 이루어져 있는데 은닉 계층에서는 데이터의 입·출력 과정에서 직접적으로 보이지 않는 숨겨진 특징을 학습하는 역할을 한다.

MLP는 여러 개의 단층 퍼셉트론으로 이루어져 있기 때문에 각 노드간 여러 가중치 값들이 존재하고 노드의 값들이 Sigmoid, Tanh, ReLU와 같은 비선형 활성화 함수에 입력됨으로써 여러 가지로 분류될 수 있는 데이터에 대한 학습을 수행할 수 있게 된다. MLP는 딥 러닝 기반의 여러 학습 모델 중에서 구조가 간단하며 연산 처리가 빨라 주로 1차원 데이터를 처리하는데 사용되고 있다. 특히, CNN과 같은 딥 러닝 기법은 이미지와 같은 2차원 데이터 처리에 적합한 구조이므로 본 논문에서는 소비 전력 파형을 학습하는 구조임을 고려하여 MLP 기법을 사용한다.

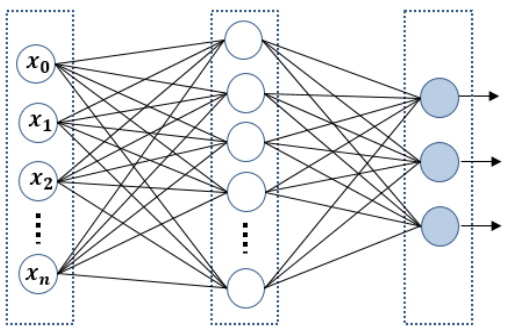


Fig. 1. Structure of Multi Layer Perceptron

III. 블록 암호 알고리즘 SM4

중국 표준 블록 암호 알고리즘인 SM4는 S. Lu에 의해 2003년에 무선 네트워크를 보호할 목적으로 SMS4라는 이름으로 제안되었다가 2012년에 SM4로 변경되었다. 현재 사용 중인 SM4는 중국 국가표준으로 지정되었으며, 2017년 국제 표준화 기구에서 ISO-IEC 18033-3 표준 알고리즘으로 공표되었다. SM4는 하나의 규격으로 설계되었는데 비밀 키와 블록의 크기는 동일하게 32비트 4개 워드로 구성된 128비트이다. 그리고 암호화 및 복호화 그리고 키 확장 과정은 모두 32라운드로 비교적 간단하게 구성되어 있다.

3.1 암호화 알고리즘

SM4 블록 암호는 불균형 페이스텔(unbalanced Feistel) 구조 암호이며 블록 크기와 키 크기는 모두 128비트, 라운드 함수는 32번을 수행하게 된다. 입력 메시지 즉, 평문 블록은 4개의 워드(32비트)로 나누어 처리되는데 Fig. 2는 SM4의 한 라운드 암호화 과정을 나타낸 것이다. 입력 $X_1 \sim X_4$가 각각 하나의 워드일 때, 각 라운드는 3개의 워드($X_2 \sim X_4$)를 라운드 함수 F에 입력하여 나온 출력 값을 이용해 X_1을 암호화한다. 이와 같은 라운드 연산을 총 32번 수행한 이후에 나온 4개의 워드는 순서를 역순으로 교체하여 최종 암호문으로 출력된다. 최종적인 암호문 출력 순서를 조정하는 이유는 페이스텔 구조인 SM4의 암호화 과정과 복호화 과정이 같아지도록 하기 위함이다.

SM4의 F 함수는 Fig. 3과 같이 입력으로 라운드 키와 3개의 평문 워드 값을 사용하며 비선형 치환(nonlinear-substitution) 연산인 τ(tau) 연산과 선형 치환(linear-substitution) 연산인 L연산을 거치게 된다.

F 함수에서 사용되는 τ 연산은 4개의 입력을 모두 XOR한 후 바이트별로 나누어 S-Box를 통해 치

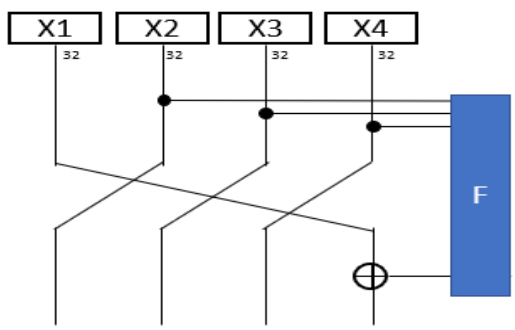


Fig. 2. One round operation of SM4

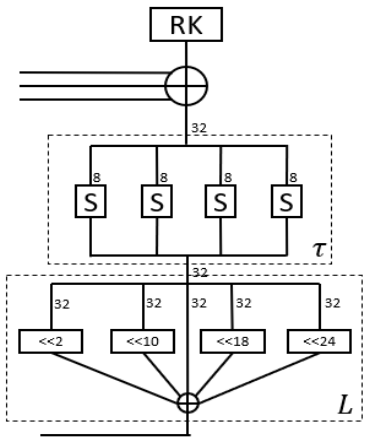


Fig. 3. The F function of SM4

환하는 연산이다. 선형 치환 연산인 L 연산은 S-Box를 거친 후 다시 병합된 32비트 워드를 각각 2, 10, 18, 24번 좌측으로 순환(rotation)시킨 후 4개의 워드를 모두 XOR 하는 연산이다. 비선형 치환 연산인 τ연산에 사용되는 S-Box는 SM4의 공식 문서에서 256개의 원소를 가진 룩업 테이블(lookup table)을 사용하도록 명시하고 있다. 해당 테이블을 만드는 실제 S-Box 연산에서는 기약 다항식 $f(x)$를 사용하는데 입력 x에 대해 $GF(2^8)$상에서의 다음 연산을 수행한다. 여기서 C_1은 $(11001011)_2$이며 C_2는 $(11010011)_2$이다.

$$S(x) = A_2(A_1 \cdot x + C_1)^{-1} + C_2$$

$$A_1 = \begin{bmatrix} 1\,0\,1\,0\,0\,0\,1\,1\,1 \\ 0\,1\,0\,0\,1\,1\,1\,1 \\ 1\,0\,0\,1\,1\,1\,1\,1\,0 \\ 0\,0\,1\,1\,1\,1\,0\,1 \\ 0\,1\,1\,1\,1\,0\,1\,0 \\ 1\,1\,1\,1\,0\,1\,0\,0 \\ 1\,1\,1\,0\,1\,0\,0\,1 \\ 1\,1\,0\,1\,0\,0\,1\,1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 1\,1\,0\,0\,1\,0\,1\,1 \\ 1\,0\,0\,1\,0\,1\,1\,1 \\ 0\,0\,1\,0\,1\,1\,1\,1 \\ 0\,1\,0\,1\,1\,1\,1\,0 \\ 1\,0\,1\,1\,1\,1\,0\,0 \\ 0\,1\,1\,1\,1\,0\,0\,1 \\ 1\,1\,1\,1\,0\,0\,1\,1 \end{bmatrix}$$

3.2 키 확장 알고리즘

SM4에서 사용하는 라운드 키는 128비트의 마스터 비밀 키로부터 32라운드 연산 과정을 거쳐 확장된다. 다음 Fig. 4는 라운드 키 생성 알고리즘을 나타낸 것으로 $i = 0$에서 $i = 31$ 까지 총 32번 반복하여 $k_4 \sim k_{35}$를 생성한다.

키 확장 과정에서 $k_0 \sim k_3$은 마스터 키를 워드 단위로 구분한 것이며, 이 때 생성된 $k_4 \sim k_{35}$가 차례로 라운드 키 $rk_1 \sim rk_{32}$로 사용된다. 키 스케줄 과정에서도 암·복호화 과정에서 사용되는 동일한 S-Box 테이블이 사용되며 처음과 중간 부분에 두 종류의 상수 FK와 CK가 XOR 연산에 사용된다.

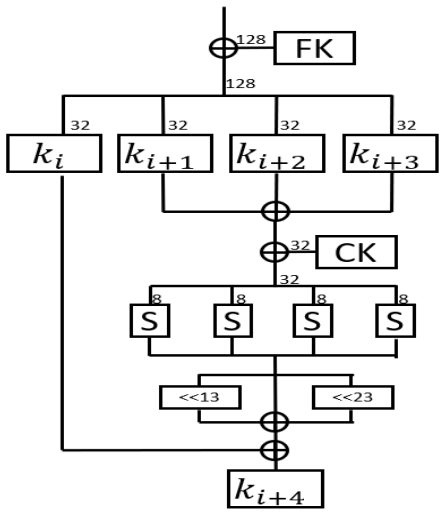


Fig. 4. One round operation for key expansion

IV. 전력 분석 공격 및 마스킹 기반 대응책

4.1 SM4에 대한 DPA/CPA 공격

블록 암호에 대한 전력 분석 공격을 효과적으로 수행하기 위해서는 분석할 전력 파형을 수집할 POI(Point Of Interest) 구간을 설정하여야 한다. AES 암호 알고리즘에서는 비선형 함수인 S-Box 연산이 수행된 이후 지점을 선택하였다[20, 21]. 본 논문에서 공격하고자 하는 SM4에서의 일차적인 전력 분석 공격 지점은 Fig. 5와 같이 F 함수 내의 비선형 치환 연산인 τ연산이 수행되는 S-Box 결과가 출력되는 구간으로 설정하였다.

τ 연산은 3개의 입력 메시지와 라운드 키가 모두 XOR 된 결과를 바이트 단위로 나누어 S-Box를 통과하는 구조이다. 이때 입력되는 3개 메시지 값은 공격자가 알 수 있고 한 바이트인 라운드 키는 알 수가 없다. 따라서 한 바이트의 라운드 키를 가정하고 이 가정된 키를 이용하여 S-Box 결과 값을 계산한 후 이 값이 실제 전력 파형과 어떤 연관성을 갖는지를 확인하는 방법으로 DPA와 CPA 공격을 시도할

수 있다.

이후 논문의 실험을 통해서 알 수 있듯이 한 바이트에 대한 키 설정은 256가지 경우의 추측 키를 가정한 것이어서 한 바이트의 라운드 키를 찾는 것은 어렵지 않다. 이 과정을 4번 수행하면 한 워드의 라운드 키를 찾을 수 있고 동일한 방법을 반복함으로써 128비트 마스터 비밀 키를 모두 추출할 수 있다.

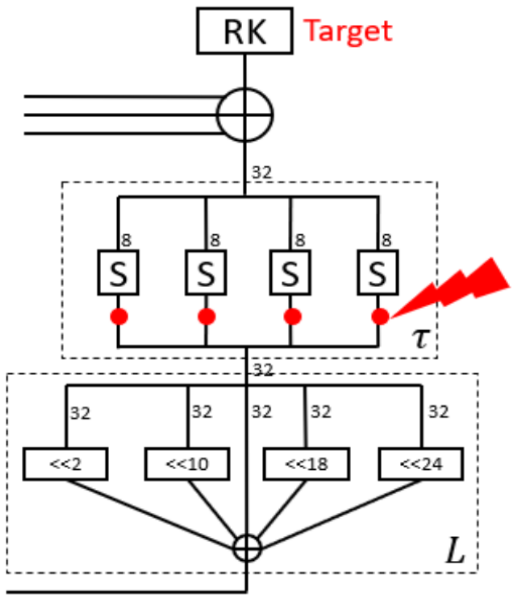


Fig. 5. Point of interest for power analysis on SM4

4.2 SM4에 대한 MLP 딥 러닝 공격

딥 러닝 기반의 전력 분석 공격에서는 라벨링 작업이 이루어진 N개의 학습용 파형을 학습시킨 후, 새로운 M개의 테스트용 파형으로 모델의 성능을 검증한다. MLP를 이용한 SM4의 공격 지점 역시 F 함수 내의 S-Box 연산 지점으로 설정하였다. 즉, 공격에서의 라벨링 값을 S-Box 중간 결과 값으로 지정해 줄 수 있으며 모델의 입력으로는 소비 전력 신호가 들어간다. MLP 전력 분석 공격에서 사용하는 입력 계층의 노드의 수는 파형의 입력 범위, 즉 샘플 수에 따라 달라지는데 일반적으로 POI 구간의 파형 샘플 수로 한다.

본 논문에서의 은닉 계층은 두 개의 층으로 구성하였으며 각각 300개의 노드로 구성하였다. 또한, 출력으로는 MLP 모델이 추측하는 중간 값으로 SM4 라운드 함수 내의 S-Box 결과 값이다. 즉, 한 바이트씩 S-Box 결과 값을 분류하는 모델이므로 256개의 노드가 필요하다. 이를 통해 S-Box 결과 값을 확정할 수 있다면 이것은 라운드 키를 한 바이트 찾은 것과 동일하다. 본 논문의 MLP에서 사용한 손실 함수는 교차-엔트로피(cross-entropy) 함수이며 최적화기(optimizer)로는 Adam 함수를 사용하였다. 실험 결과에 의하면, 딥 러닝 기반의 MLP를 이용한 프로파일링 공격에서도 높은 정확도(accuracy)로 라운드 키를 찾을 수 있어 암호 알고리즘 SM4는 전력 분석 공격에 매우 취약한 특성이 있음을 알 수 있었다. 다만 MLP 기반의 공격에 사용되는 하이퍼 파라미터(hyper parameter)는 모델의 성능과 수행 시간에 중요한 영향을 미친다. 본 논문에서는 입력 파형의 샘플 수와 라벨 수 등을 고려한 시뮬레이션을 통해 은닉 계층의 층수와 노드 수를 결정하였으며 다중 분류 특성을 고려하여 일반적으로 많이 사용하는 손실 함수와 최적화기를 선정하였다.

4.3 마스킹 기법을 이용한 전력 분석 대응

전력 분석 공격에 대응하기 위한 방법 중 하나는 메시지에 대한 마스킹을 적용하는 방법이다. 그러나 마스킹 기법은 암호 연산에 사용하는 연산자에 따라 많은 부가 연산 시간 및 메모리가 필요한 단점이 있다. 일부 연구 결과에서는 마스킹을 적용하는 것이 원래 암호 알고리즘을 수행하는 시간보다 4~10배 이상까지 증가하기도 하였다[22, 23]. 본 논문에서는 SM4에 마스킹 기법을 적용하되 추가되는 연산 시간이나 메모리가 최소화되도록 설계하고자 한다. 마스킹이 적용된 SM4 블록 암호 과정을 쉽게 설명하기 위해 Fig. 6에 나타내었다.

SM4의 암호 연산은 S-Box 연산을 제외하면 모두 부울(Boolean) 연산으로만 이루어져 부울 마스크를 우선적으로 적용하는 것이 효과적이다. 논문에서 제안하는 마스킹 기법에서는 초기 마스크 값으로 4개의 워드와 2개의 바이트를 사용하는데 이 값은 하나의 메시지 암호를 수행할 때마다 매번 랜덤한 값으로 생성되어야 한다. 생성된 4개의 워드 및 2개의 바이트를 $M_1, M_2, M_3, M_4, M_i, M_o$라 표기하며 M_1~M_4는 평문 블록에 적용되는 마스크 값이며 M_i, M_o는 S-Box에 적용되는 마스크 값이다. 일반적으로

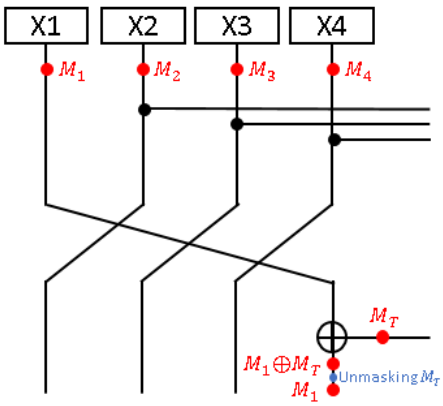


Fig. 6. Masking value for one round on SM4

Masked S-Box인 MS는 S-Box 테이블을 이용하여 다음과 같이 사전에 생성한다[24, 25].

$$Output = MS[Input \oplus M_i] = S[Input] \oplus M_o$$

사전에 마스킹 테이블이 생성된 후에는 입력 메시지에 랜덤하게 생성된 M_1, M_2, M_3, M_4를 사용하여 마스킹 연산을 수행한다. 그리고 F 함수의 결과에 씌워진 M_T 마스킹을 다음 라운드가 진행되기 전에 이를 제거하여 매 라운드마다 4개의 마스크 값이 그대로 유지되게 하는 것이 중요하다.

이와 같이 한 라운드 동안의 마스크 값이 모두 결정되면 F 함수에서는 Fig. 7과 같이 메인 가지 중 3개가 입력되어 라운드 키와 함께 XOR 된다. 따라서 마스킹된 입력들이 XOR된 이후 마스크 값은 1라운드에서 $M_s = M_2 \oplus M_3 \oplus M_4$이며 2라운드에서는 $M_s = M_1 \oplus M_3 \oplus M_4$가 되며 이는 고정되지 않고 4번의 주기로 순환하게 된다. 이를 구현에 적합하도록 표현하면, 진행되는 라운드 수가 $r(1 \le r \le 32)$일 때 마스크 값은 $M_s = M_1 \oplus M_2 \oplus M_3 \oplus M_4 \oplus M_{(r-1) \bmod 4+1}$이 된다.

다음으로 입력된 값을 마스킹된 S-Box 연산을 수행하기 위해 먼저 $M_i^4 (= M_i \| M_i \| M_i \| M_i)$값을 먼저 마스킹 처리를 하고 이후에 M_s 값을 제거하게 된다. 그러면 S-Box 입력 전 마스크 값은 M_i가 되며 출력 후 마스크 값은 M_0가 된다. 이후 32비트의 마스크 값 $M_o^4 (= M_o \| M_o \| M_o \| M_o)$도 L 연산을 거치면서 마스크 값이 변하게 된다. 결국, L 연산까지 진행되

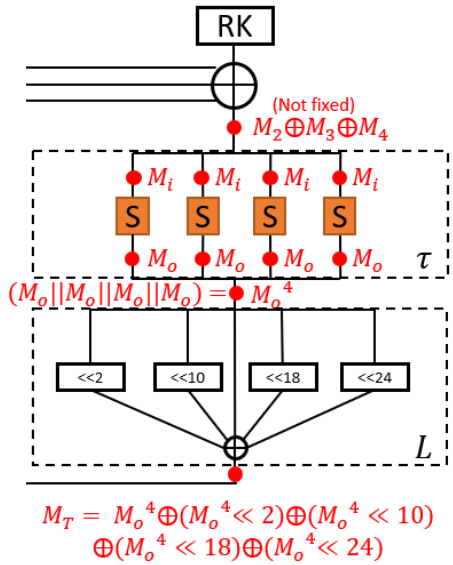


Fig. 7. Masking value for F function

었을 때 F 함수의 최종 마스크 값은 M_T이다. 그런데 이 마스크 값은 M_o에 따라 결정되는 암호화 과정에서는 고정된 값이다.

V. 전력 분석 공격 실험 및 안전성 검증

5.1 전력 분석 공격 실험 환경

상기한 바와 같이 SM4는 비선형 연산인 S-Box의 출력 값을 대상으로 전력 분석 공격이 가능하다. 해당 지점의 중간 결과 값을 직접 예측하여 키를 찾는 DPA나 CPA 비프로파일링 공격 기법을 사용할 수도 있으며, 중간 결과 값을 라벨로 두어 딥 러닝 모델을 학습시킨 후 공격하는 프로파일링 공격 기법들을 모두 사용할 수 있다. 한 바이트의 라운드 키를 찾을 수 있으면 동일한 과정을 거쳐 라운드 키 전체를 찾을 수 있으며 나아가 마스터 비밀 키 추출이 가능하다.

본 논문에서는 블록 암호 SM4에 대한 전력 분석 공격 및 마스킹 기법의 안전성을 검증하기 위해 ChipWhisperer-Lite[26]를 사용할 수 있는 개발 보드에 SM4 알고리즘을 구현하여 실험하였다. 이 개발 보드는 8비트 MCU인 XMEGA128가 탑재되어 있으며 동작 클럭은 7.37MHz이다.

5.2 마스킹이 적용되지 않은 SM4에 대한 공격

5.2.1 DPA 및 CPA 공격

전력 분석 공격 대응책이 없는 순수한 SM4에 대한 DPA 및 CPA 공격을 위해서 각 220개의 샘플로 이루어진 1,000개의 파형을 사용하였다. 실험에서 첫 번째 라운드의 첫 번째 바이트 키에 대해 DPA를 진행한 결과 Fig. 8과 같이 쉽게 라운드 키의 일부인 0xF1를 찾는 것을 확인할 수 있다. 동일한 전력 파형으로 DPA와 유사하게 CPA 공격 기법으로도 실험한 결과, Fig. 9와 같이 쉽게 라운드 키 중 한 바이트인 0xF1을 찾는 것을 볼 수 있다. Fig. 8에서의 세로축은 각 추측 키에 따른 분류된 파형의 차분 값을 의미하며 Fig. 9에서의 세로축은 중간 값과 전력 파형간의 상관도를 의미한다.

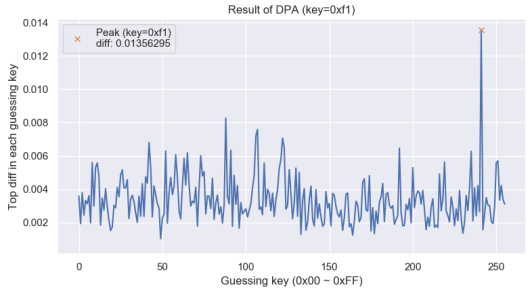


Fig. 8. The result of DPA on unmasked SM4

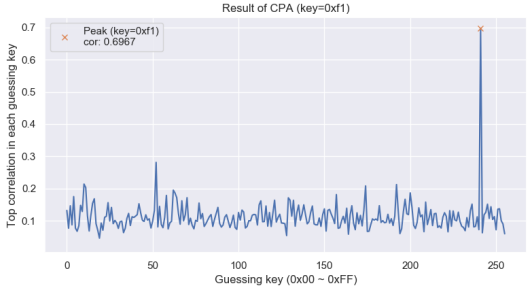


Fig. 9. The result of CPA on unmasked SM4

5.2.2 MLP 공격

MLP 모델을 이용한 실험에서는 공격 대상인 S-Box 부근의 1,300개의 샘플로 이루어진 8,000개의 파형을 사용하였다. 이 중에서 모델 학습에 7,000개를 검증에 1,000개를 사용하였다. 따라서 각 파형의 데이터 셋 라벨은 S-Box의 결과 값으로 사용하였다.

MLP 모델은 파형의 구간에서 각 파형에 해당하는 라벨과의 연관성을 자동으로 학습한다. 7,000개의 파형으로 학습시킨 후 학습에 이용되지 않은 1,000개의 파형으로 검증한 결과 약 99.9%의 정확도로 주어진 파형에 대한 S-Box 출력 결과를 찾아냄을 확인할 수 있었다. Fig. 10은 모델의 정확도와 손실 값의 추이를 나타낸 것이다. 에포크(epoch)가 40인 지점에서 정확도가 90%를 나타내었으며 에포크 값을 증가시켜 더 많은 학습을 진행하면 정확도는 더욱 상승하였다.

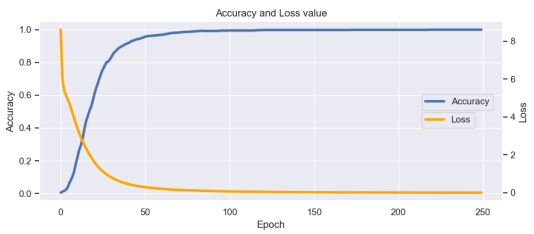


Fig. 10. The accuracy and loss value of the MLP attack model

5.3 마스킹 기법이 적용된 SM4의 성능

제안한 마스킹 기법이 적용된 SM4에 대한 성능을 비교하기 위해 Atmel Studio의 시뮬레이터를 이용하여 수행 속도를 측정하였다. SM4의 마스킹 기법 적용 유무에 따른 클록 사이클을 측정한 표는 Table 1과 같다. 결과적으로 제안한 마스킹 기법이 적용된 SM4가 대응책이 없는 SM4보다 암호화 과정에서는 클록 사이클이 약 5.5%만 증가한 것을 확인할 수 있었다.

원래 SM4에서의 사전 계산 과정은 키 확장 연산을 의미하며 이 과정은 35,403 사이클 정도의 시간이 필요하다. 마스킹 기법이 적용될 때에는 마스킹 테이블을 만드는 과정과 마스크 값에 의해 고정된 값을 계산하는 과정에서 21,583 사이클이 필요하게 되어 전체적으로 약 61% 증가하게 된다. 하지만, 이는 블록 암호 운용 모드를 이용해 여러 블록의 메시지를 암·복호화할 때 처음 한 번만 수행되므로 메시지 암호의 큰 영향 요소는 아니다. 따라서 긴 메시지를 암·복호화할수록 추가되는 사전 계산량은 전체 연산량에 비해 큰 비중을 차지하지 않게 된다. 다만,

Table 1. Comparison of clock cycles on original SM4 and Masked-SM4

	Original SM4	Masked SM4	Increase Rate
Pre-computation	35,403	56,986	61%
Encryption	42,477	44,845	5.5%

마스킹 기법을 적용할 경우에는 Masked S-box 값을 저장하기 위한 256바이트 정도의 임시 메모리가 추가로 필요하다.

5.4 마스킹 기법이 적용된 SM4 공격

5.4.1 DPA 및 CPA 공격

제안한 마스킹 기법이 적용된 SM4를 대상으로 이전 DPA와 동일한 실험을 시도하였다. Fig. 11에서 보는 바와 같이 정확한 비밀 키를 나타내는 피크가 생성되지 않음을 알 수 있다. 또한, 동일한 방법으로 마스크가 적용된 SM4에 대해서 CPA를 적용해도 비밀 키 구간에서 상관도가 낮게 생성됨을 Fig. 12에서도 볼 수 있다. 이를 통해 제안하는 마스킹 기법은 DPA 및 CPA 공격에 충분히 대응할 수 있음을 알 수 있다.

Fig. 11. The result of DPA on masked SM4

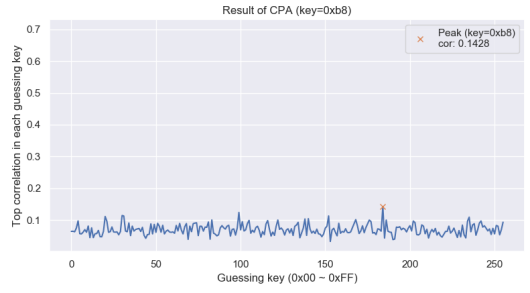


Fig. 12. The result of CPA on masked SM4

5.4.2 MLP 공격

프로파일링 공격은 공격자가 공격 목표로 하는 디바이스와 동일하거나 유사한 디바이스를 마음대로 조작할 수 있어야 하며 학습용 데이터에 대해 평문과 암호문은 물론 내부적으로 사용된 마스크 값 또한 알 수 있다는 보다 강력한 가정을 전제로 한다. 물론 공격 목표 디바이스에서는 비밀 키와 마스크 값은 알 수 없지만 입력되는 평문과 출력 암호문 그리고 구동 시 소비되는 전력 파형은 측정할 수 있는 환경임을 가정한다.

마스킹이 적용된 SM4에 대한 MLP 기반 프로파일링 공격 모델은 Fig. 13과 같다. Masked S-Box의 결과와 마스크 값 M_o를 각각 학습한 두 모델을 이용해 공격 대상 파형으로부터 $MSbox[x \oplus M_i]$와 M_o를 찾아낸다. 그런데 $MSbox[x \oplus M_i]$의 값은 $SBox[x] \oplus M_o$와 같으므로 두 모델을 통해 추론한 각각의 값을 XOR 해주면 마스크 값이 씌워지지 않았을 때의 $SBox$의 입력 값을 구할 수 있게 된다. 따라서 이를 평문 워드인 X_2, X_3, X_4와 함께 XOR 연산을 하면 rk_1을 찾을 수 있게 된다. 즉, MLP를 이용하여 마스킹이 적용된 SM4를 공격하기 위해서는 독립된 두 개의 모델 학습과 테스트를 거쳐 $MSbox$의 출력과 마스크 값 M_o를 찾아내고 이 공격 모델에 따라 비밀 키 추출이 가능하다.

마스크 값 M_o를 찾기 위해서는 암호화 연산이 이루어지기 전 $MSBox$를 생성할 때 M_o를 256번 연

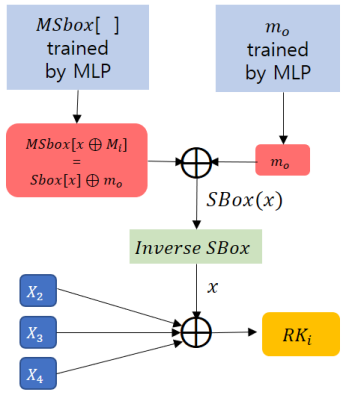


Fig. 13. MLP attack model for masked SM4

산에 직접 사용한다는 점을 주목할 필요가 있다. 따라서 이 시점에서 256개의 일정 패턴을 보이는 구간을 찾아 M_o 값에 대한 학습을 진행할 수 있다. 또한, $MSBox[x \oplus M_i]$ 결과는 마스킹이 적용되지 않았을 때의 공격 지점과 동일하게 $MSBox$ 출력 지점을 학습시키면 된다.

MLP 실험에서는 M_o 학습과 $MSBox[x \oplus M_i]$ 학습을 위해 각각 500개, 1,300개의 샘플로 이루어진 8,000개씩의 파형을 사용하였다. 전자는 $MSBox$를 만드는 256번의 반복 연산 중 두 번째까지의 파형을 후자는 $MSBox$ 연산 인근 파형 구간을 사용하였다. 두 모델 모두 7,000개의 학습용 데이터로 학습시킨 후 1,000개의 테스트 데이터로 검증하는 형태로 진행하였다.

마스킹이 적용된 SM4 알고리즘에 대해 MLP 모델에서 라운드 키를 찾는 실험을 한 결과, 올바른 마스크 값과 중간 값을 각각 98.2%, 99.9%의 정확도로 찾을 수 있었다. 다음 Fig. 14는 M_o를 학습한 모델의 입력 계층의 파형(위)과 은닉 계층의 가중치(아래)를 함께 도시한 그래프이다. 파형에 나타난 가중치의 절대 값이 큰 부분이 MLP 모델이 집중적으로 학습한 구간을 의미한다.

또한, Fig. 15는 $MSBox[x \oplus M_i]$를 학습한 모델의 입력 계층의 파형(위)과 은닉 계층의 가중치(아래)를 도시한 그래프이다. 결론적으로 마스킹이 적용된 SM4 암호 알고리즘은 MLP를 이용한 프로파일링 공격에는 약 98.1%의 정확도로 비밀 키가 노출되는 취약성이 있다고 할 수 있다.

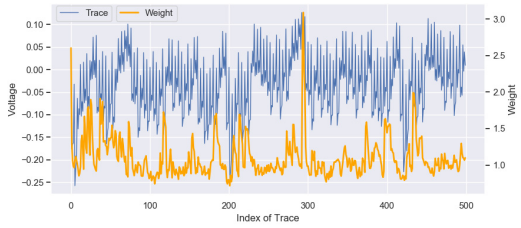


Fig. 14. The input trace and weight value of the model when the M_o is learned

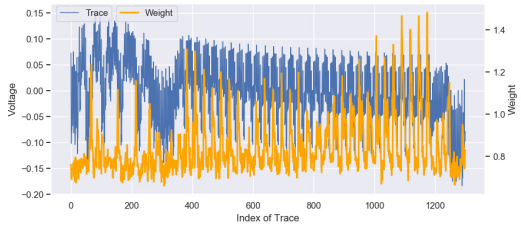


Fig. 15. The input trace and weight value when the $MSBox[x \oplus M_i]$ is learned

VI. 결 론

본 논문에서는 중국 표준 블록 암호 알고리즘인 SM4를 ChipWhisperer-Lite를 사용할 수 있는 개발 보드에 구현하여 실험한 결과 전력 분석 공격에 비밀 키가 노출될 수 있음을 검증하였다. 전력 분석 공격에는 DPA, CPA 등과 같은 비프로파일링 기법뿐만 아니라 공격 가정이 더욱 강력한 MLP와 같은 프로파일링 공격도 적용할 수 있음도 확인하였다.

본 논문에서는 전력 분석 공격에 대응하고자 마스킹 기반의 대응 기법을 제안하였다. 제안한 마스킹 기법은 순수하게 SM4를 구현하는 것보다 암호화 과정에서 5.5% 정도의 연산만 추가되어 매우 효율적으로 설계되었으며 DPA, CPA 등과 같은 비프로파일링 공격에 대응할 수 있음을 실험적으로 확인하였다. 그럼에도 1차 마스킹 기법은 모듈의 안전성 검증 기준에 따른 테스트를 향후에 수행할 필요가 있다. 다만, 논문에서는 1차 마스킹 기법만으로는 MLP나 CNN과 같은 딥 러닝 기반의 프로파일링 공격을 막기에는 여전히 한계가 있음을 확인하였다. 따라서 SM4 암호 알고리즘을 사용할 경우에는 고차 마스킹 기법과 같은 고수준의 전력 분석 공격 대응책과 더불어 하드웨어적인 대응 기법을 병행하여 설계하여야 안전한 암호 시스템을 구현할 수 있다.

References

[1] National Institute of Standards and Technology, "Advanced Encryption Standards," NIST FIPS PUB 197, 2001.

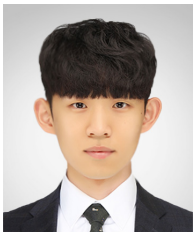
[2] Korea Information Security Agency, "A Design and Analysis of 128-bit Symmetric Block Cipher (SEED)," Seoul, Korea, 1999.

[3] D. Hong, J. Lee, D. Kim, D. Kwon, K. Ryu, and D. Lee, "LEA, A 128-bit block cipher for fast encryption on common processors," WISA'13, LNCS

8267, pp. 3-27, 2014.
〔4〕 Office of State Commercial Cryptography Administration, P.R.China, Functionality and Interface Specification of Cryptographic Support Platform for Trusted Computing (in Chinese), http://www.oscca.gov.cn, 2012.
〔5〕 W. Diffie and G. Ledin. "SMS4 Encryption Algorithm for Wireless Networks," IACR Cryptology e-print Archive, 2008.
〔6〕 P. Kocher, "Timing Attacks on Implementation of Diffie-Hellman, RSA, DSS, and Other Systems," CRYPTO'96, LNCS 1109, pp. 104-113, 1996.
〔7〕 T. Messerges, "Securing the AES finalists against power analysis attacks," FSE'00, LNCS 1978, pp. 150-164, 2001
〔8〕 P. Kocher, J. Jaffe, and B. Jun, "Differential power analysis," CRYPTO'99, LNCS 1666, pp. 388-397, 1999
〔9〕 J. Coron, "Resistance against differential power analysis for elliptic curve cryptosystems", CHES'99, LNCS 1717, pp. 292-302, Springer-Verlag, 1999.
〔10〕 E. Brier, C. Clavier, and F. Olivier, "Correlation Power Analysis with a Leakage Model," CHES'04, LNCS 3156. pp. 16 - 29, 2004
〔11〕 F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," Psychological Review, Vol. 65, No. 6, 1958.
〔12〕 R. Collobert and S. Benjio, "Links between perceptrons, MLPs and SVMs," Proceedings of the twenty-first international conference on Machine learning, ICML'04, p. 23, 2004.
〔13〕 Z. Martinasek, and V. Zeman, "Innovative method of the power analysis," Radioengineering, Vol. 22, No. 2, pp. 589-594, 2013.
〔14〕 Z. Martinasek, J. Hajny, and L. Malina, "Optimization of power analysis using neural network," CARDIS'13, LNCS 8419, pp. 94-107, 2014.
〔15〕 S. Chari, J. R. Rao, and P. Rohatgi, "Template Attacks," CHES'02, LNCS 2523, pp. 13-28, 2002.
〔16〕 S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a Convolutional Neural Network," ICET '17, pp. 1-6, 2017.
〔17〕 J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," Neural Networks, Vol. 61, pp. 85-117, 2015.
〔18〕 A. Golder, D. Das, J. Danial, S. Ghosh, S. Sen, and A. Raychowdhury, "Practical Approaches Towards Deep-Learning Based Cross-Device Power Side Channel Attack," IEEE Trans. on VLSI systems, Vol. 27, No. 12, pp. 2720-2733, 2019.
〔19〕 L. Wei, B. Luo, Y. Li, Y. Liu, and Q. Xu, "I Know What You See: Power Side-Channel Attack on Convolutional Neural Network Accelerators," ACSAC'18, pp. 393-406, 2018.
〔20〕 W. Shan, L. Wang, Q. Li, L. Guo, S. Liu, and Z. Zhang, "A chosen-plaintext method of CPA on SM4 block cipher," 10th International Conference on Computational Intelligence and Security, pp. 363-366, 2014.
〔21〕 G. Bai, H. Fu, W. Li, and X. Wu, "Differential power attack on SM4 block cipher," IEEE International Conference On Trust, Security And Privacy In Computing And Communications, pp. 1494-1497, 2018.
〔22〕 C. Kim, J. Park, D, Han, and D. Lee,

"Investigation of masking based side channel countermeasures for LEA", Journal of The Korea Institute of Information Security & Cryptology(JKIISC), Vol. 26 No. 6, pp. 1431-1441, 2016.
[23] E. Park, S. Oh, and J. Ha, "Masking-based block cipher LEA resistant to side channel attacks," Journal of The Korea Institute of Information Security & Cryptology(JKIISC), Vol. 27 No. 5, pp. 1023-1032, 2014.
[24] C. Herbst, E. Oswald, and S. Mangard, "An AES smart card implementation resistant to power analysis Aatacks," ACNS'06, LNCS 3989, pp. 239-252, 2006.
[25] E. Oswald and K. Schramm. "An efficient masking scheme for AES software implementations," WISA'05, LNCS 3786, pp. 292-305, 2006.
[26] ChipWhisperer - NewAE Technology Inc., "chipwhisperer," Available at http://newae.com/tools/chipwhisperer, 2017.

〈저자소개〉



배 대 현 (Daehyeon Bae) 학생회원
2017년 3월: 호서대학교 컴퓨터정보공학부 입학
2017년 3월~현재: 호서대학교 컴퓨터정보공학부 학부과정
〈관심분야〉 암호학, 부채널 공격, 인공지능 보안

남 승 현 (Seunghyun Nam) 정회원
1986년 2월: 연세대학교 전자공학과 학사
1988년 9월: 연세대학교 전자공학과 석사
1995년 3월: 연세대학교 전자공학과 박사
2017년 11월~현재: ㈜ 라닉스 핵심기술연구소장
〈관심분야〉 V2X 통신 칩, 자동차 네트워크 보안, 보안 칩 설계, 부채널 공격

하 재 철 (Jaecheol Ha) 종신회원
1989년 2월: 경북대학교 전자공학과 학사
1993년 8월: 경북대학교 전자공학과 석사
1998년 2월: 경북대학교 전자공학과 박사
1998년 3월~2007년 2월: 나사렛대학교 정보통신학과 교수
2007년 3월~현재: 호서대학교 컴퓨터정보공학부 교수
2013년 1월~현재: 한국정보보호학회 상임부회장
2009년 1월~현재: 한국산학기술학회 이사
〈관심분야〉 정보보호, 네트워크 보안, 부채널 공격