JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 27 No. 1, pp. 43-50, January 2022
https://doi.org/10.9708/jksci.2022.27.01.043

Improved real-time power analysis attack using CPA and CNN

Ki-Hwan Kim*, HyunHo Kim*, Hoon Jae Lee*

*Professor, Dept. Computer Engineering, Dongseo University, Busan, Korea
*Professor, Dept. Computer Engineering, Dongseo University, Busan, Korea
*Professor, Dept. Computer Engineering, Dongseo University, Busan, Korea

[Abstract]

Correlation Power Analysis(CPA) is a sub-channel attack method that measures the detailed power consumption of attack target equipment equipped with cryptographic algorithms and guesses the secret key used in cryptographic algorithms with more than 90% probability. Since CPA performs analysis based on statistics, a large amount of data is necessarily required. Therefore, the CPA must measure power consumption for at least about 15 minutes for each attack. In this paper proposes a method of using a Convolutional Neural Network(CNN) capable of accumulating input data and predicting results to solve the data collection problem of CPA. By collecting and learning the power consumption of the target equipment in advance, entering any power consumption can immediately estimate the secret key, improving the computational speed and 96.7% of the secret key estimation accuracy.

▸ **Key words**: Side Channel Attack, AES, CPA, CNN, ATmega328, Oscilloscope

[요 약]

CPA(Correlation Power Analysis)는 암호 알고리즘이 탑재된 공격 대상 장비의 미세한 소비전력을 측정하여 90% 이상의 확률로 암호 알고리즘에 사용된 비밀키를 추측하는 부채널 공격 방법이다. CPA는 통계를 기반으로 분석을 수행하기 때문에 반드시 많은 양의 데이터가 요구된다. 따라서 CPA는 매회 공격을 위해 약 15분 이상 소비전력을 측정해야만 한다. 본 논문에서는 CPA의 데이터 수집 문제를 해결하기 위해 입력데이터를 축적하고 결과를 예측할 수 있는 CNN(Convolutional Neural Network)을 사용하는 방법을 제안한다. 사전에 공격 대상 장비의 소비전력을 수집 및 학습을 통해 임의의 소비전력을 입력시키면 즉각적으로 비밀키를 추정할 수 있어 연산속도를 향상하고 96.7%의 비밀키 추측 정확도를 나타냈다.

▸ **주제어**: 부채널 공격, 고급 암호화 표준, 차분 전력 공격, 합성곱 신경망, ATmega328, 오실로스코프

• First Author: Ki-Hwan Kim, Corresponding Author: Hoon Jae Lee
*Ki-Hwan Kim (ghksdl90@gdsu.dongseo.ac.kr), Dept. Computer Engineering, Dongseo University
*HyunHo Kim (feei_@naver.com), Dept. Computer Engineering, Dongseo University
*Hoon Jae Lee (hjlee@dongseo.ac.kr), Dept. Computer Engineering, Dongseo University
• Received: 2021. 10. 27, Revised: 2021. 12. 29, Accepted: 2021. 12. 29.

http://www.ksci.re.kr pISSN:1598-849X | eISSN:2383-9945

I. Introduction

현대인의 생활에 인터넷과 전기가 없는 삶은 상상할 수 없을 정도로 생활 곳곳에 다양한 전자제품이 자리 잡고 있다. TV, 냉장고, 세탁기 등은 사물인터넷(Internet of Things, IoT)을 접목하여 다양한 데이터를 서버에서 분석하여 사용자에게 편의 서비스로 제공하고 있다. 금융업무의 경우 IC 카드 및 보안토큰 등에 복잡한 암호 시스템을 사용하여 개인 인증 절차를 간소화하여 간편하게 서비스를 사용할 수 있다. 위와 같은 다양한 개인 인증 서비스와 데이터 암호화 기능은 수학적 안전성과 암호 학적 이론에 기반을 두어 보증되고 있다. 현존하는 암호화 기법에서 데이터 암호화하는 방법은 블록 기반의 암호화 방식의 AES(Advanced Encryption Standard)가 전 세계 수학자와 암호 연구자들에게 인정받은 국제 암호 표준이다. AES는 이론적으로 암호문에서 암호화에 사용된 원본 데이터와 비밀키는 알 수 없도록 설계되어 있어 현실적으로 공격할 수 없다.

1996년 Paul Kocher가 연산 시간을 이용한 타이밍 공격(Timing Attack, TA)을 시작으로 1999년 전력 분석(Power Analysis, PA)의 개념이 제안되면서 하드웨어에서 암호 알고리즘이 구현될 때, 전력 소모량, 수행 시간 등을 측정하여 비밀키를 유추하는 부채널 공격(Side Channel Analysis, SCA)이 등장하면서, AES의 안전성에 관한 이슈가 제기되었다[1-3]. 이 문제는 스마트 폰, IC 카드, 웨어러블 디바이스, IoT 센서 등 휴대성이 높거나 실외에 노출되어 누구든 물리적인 접근이 가능한 모든 전자장비를 대상으로 확산하였다. 부채널 공격은 주로 암호 알고리즘이 연산하는 순간 비밀키를 추측하기 위해 사용되는 공격수단이다[4-6]. 부채널 공격은 프로파일링과 논 프로파일링 기법으로 분류할 수 있다. 프로파일링은 공격 대상 장비와 동일한 환경에서 정보를 얻는 공격을 의미하고 논 프로파일링은 공격자가 공격 대상 장비로부터 정보를 얻는 환경에서 수행되는 공격을 의미한다.

부채널 공격에 해당하는 상관 전력 분석 공격(Correlation Power Analysis, CPA)은 통계적인 분석을 사용하므로 충분한 데이터가 있을 경우 약 90% 이상의 높은 공격 성공률을 보인다[7]. 그러나 데이터 수집에는 최소한 15분 이상의 많은 반복과정이 요구되며, 이는 현실적이지 못한 공격 기법이라고 볼 수 있다.

본 논문은 매회 공격을 시도할 경우 필수적으로 요구되는 측정을 생략하기 위해 머신러닝 기반의 합성곱 신경망(Convolutional Neural Network, CNN)을 결합한 공격 방법을 실험하고 실험 결과를 비교 분석하여 결론을 맺고자 한다. 실험에 사용할 하드웨어는 산업현장과 사물인터넷에 사용되는 ATmega 하드웨어이며, 공격을 시도할 암호 알고리즘은 국제적으로 널리 사용되는 AES이다. 실험의 성과비교를 위하여 전통적인 CPA공격과 제안하는 방법에 각각 임의의 비밀키와 소비전력을 데이터를 입력하고 공격 성공률을 비교분석하였다.

II. Related Work

2.1. Related research papers

김한빛 등[7]은 CHES(Cryptographic Hardware and Embedded Systems)에서 2018부터 2020년간 부채널 공격 논문 동향을 조사했다. 조사 결과에서 대칭키 암호는 AES가 다른 암호 알고리즘에 비하여 압도적으로 많은 논문이 투고되어 활발한 연구가 진행되고 있음을 알 수 있다. 기본적으로 CPA는 논 프로파일링에서 다량의 파형을 수집하여 통계적 분석을 통해 비밀 정보를 추론하는 방법이다. 통계적인 분석 방법은 변화량이 많은 데이터에서 분류 기준을 통해 정답을 도출할 수 있다. 단, 통계량을 계산하기 위해 반드시 소비전력의 시점이 일치시켜야만 한다.

Giovanni Camurati등[8]연구에서는 CPU와 무선 통신 센서가 통합된 시스템에서 CPU에서 누설되는 민감한 연산 정보를 무선 신호를 분석하여 수신하는 기법을 제안했으며, 최대 AES 128의 비밀키를 추론하기 위해 약 15분간 1,000~1,500개의 trace 데이터를 수집하고 공격에 성공했다. 또 다른 실험으로 이들의 연구에서는 거리에 따른 유출정보를 확인하여 키를 찾아내는 실험도 진행하였다. 실험 거리는 10m, 15m, 34m, 60m에서 각각 측정하여 진행하였으며, 10m와 45m 거리에서는 500~1500개의 trace, 15m 거리에서는 500~5000개의 trace, 60m에서는 10~500개의 trace에서 키를 찾는데 성공하였다.

Dennis R. E. Gnad등[9]은 마이크로컨트롤러 및 SoC 장치에 AES가 연산되는 상태를 대상으로 노이즈를 분석하기 위해 100,000번의 측정을 통해 연산 값을 추정했다. 이 외의 다른 동향을 포함한 연구 결과를 추론해보면 대부분의 부채널 공격은 통계적 추론하는 방법을 사용하는 상황이다. 이러한 상황은 현실에서는 끊임없는 연산 작업이 수행되기에 현실적이지 못한 문제점이 있다.

CPA는 공격 대상 장비의 미세한 소비전력까지 감지하여 특정 함수의 연산 과정에서 소비되는 전력량을 이용하여 연산에 사용된 값을 추측하는 공격 방법이다.

Owen Lo[11]가 AES와 유사한 블록암호 PRESENT를 대상으로 CPA 공격을 시도하였으나 AES의 반복 연산에 작업으로 정상적인 신호를 낮은 확률로 추론했다. Faisal Rahman[12]도 ATmega328p 하드웨어에 AES의 S-box를 공격하여 공격에 성공했다. 그러나 두 논문 모두 실패 확률이 높아 실용성과는 거리가 멀다. 다만 FRANÇOIS DURVAUX[13]는 ATmega328p와 ARM Cortex-M3를 사용하여 AES를 EM으로 공격하기 위하여 256bit 키를 1분 이내에 공격했다.

권홍필, 하재철[14]의 AES 1라운드 소비전력에서 AddRoundKey함수와 SubBytes 함수 부분을 공격한 논문은 RP 기법을 전처리하는 것으로 높은 정확도를 보였다. 이들의 공격 목표는 XMEGA128 보드로 부채널 공격의 정보 수집에 특화된 보드이다.

2.2 CPA overview and theory

CPA는 공격 대상 장비에 암호 알고리즘이 있다고 가정하고 공격 시점과 함수의 구조를 f(p, k)로 정의한다. 함수(f)는 암호 알고리즘에서 공격 시점에 연산하는 함수, 평문(p), 비밀키(k)를 의미한다. 위의 함수를 사용하면 선정된 시점에 실행된 연산 값을 분석하여 공격하게 된다[10].

식(1)은 임의의 암호 알고리즘(f)에 i 개의 서로 다른 평문과 중간키(m_k)를 공격 시점에 해당하는 함수에 입력하여 함수가 반환한 결과를 d로 표현한다. 식(2)는 i 개의 암호문이 연산하는 순간 측정된 각각의 소비전력 신호(T)로 아주 짧은 시간에 측정된 각각의 요소를 t로 정의하면 임의로 j개의 소비전력 요소로 정의할 수 있다.

$$d=\begin{bmatrix} f(p_1,k_1) \\ \vdots \\ f(p_i,k_i) \end{bmatrix}=\begin{bmatrix} c_1 \\ \vdots \\ c_i \end{bmatrix} \tag{1}$$

$$T=[t_1, t_2, \cdots, t_j] \tag{2}$$

이때 주의할 점은 통계적인 기법을 사용하기 때문에 측정된 소비전력 신호의 시작 지점이 같아야만 같은 속성의 데이터로 결합하여 분석에 사용할 수 있다. 그러나 미세한 소비전력 신호를 측정하는 오실로스코프만 사용하여 같은 시작 지점을 찾을 수 없다. 이 문제의 해결책으로 하드웨어에서 공격 목표인 함수를 호출할 때 5V 이외의 경우에는 0V를 출력하는 별도의 물리 회로를 추가하는 것이다.

암호 알고리즘에 임의의 평문과 비밀키를 입력하여 측정된 소비전력 신호는 불규칙한 길이(j)로 측정되며, 이를 수식으로 표현하면 식(3)과 같다.

$$T=\begin{bmatrix} t_{00}\cdots t_{0j} \\ \vdots \quad \ddots \quad \vdots \\ t_{ij} \ldots t_{ij} \end{bmatrix} \tag{3}$$

다음으로 통계분석을 위해 비교군 데이터를 연산한다. 추정된 비밀키(g)는 식(4)와 같이 AES의 AddRoundKey에서 입력하는 모든 파라미터를 0부터 255까지 8bit로 표현하는 것으로 총 256개의 경우에 수를 가진다.

$$g=(g_0, g_1 \cdots \cdots, g_{255}) \tag{4}$$

측정된 소비전력과 추정된 비밀키를 비교하기 위하여 식(5)처럼 1차원 배열인 평문(p)과 통계분석을 위해 비교할 추정된 비밀키(g)를 모두 곱하면 16진수의 2차원 추정된 배열(V)을 계산할 수 있다.

$$\begin{aligned} V&=f\left(\begin{bmatrix} p_1 \\ \vdots \\ p_i \end{bmatrix},[g_0,\cdots,g_{255}]\right) \\ &=\begin{bmatrix} f(p_1,g_0)\cdots f(p_1,g_{255}) \\ \vdots \quad \ddots \quad \vdots \\ f(p_i,g_0)\cdots f(p_i,g_{255}) \end{bmatrix} \\ &=\begin{bmatrix} v_{11}\cdots v_{1j} \\ \vdots \quad \ddots \quad \vdots \\ v_{i1}\cdots v_{ij} \end{bmatrix} \end{aligned} \tag{5}$$

식(6)은 앞서 측정한 소비전력과 추정된 배열을 활용하기 위해 논리적으로 표현한 것이다. 평문과 비밀키 암호문과의 관계가 특정 암호 함수와 소비전력이 관련이 있고 서로 치환할 수 있지만, 암호 함수의 결과와 소비전력은 같은 속성을 같지 않다는 것을 말한다.

$$\begin{aligned} &p \oplus k=c \\ &p \oplus k=f(p_i,k_i) \\ &c=T_i \\ &f(p_i,k_i)\neq T_i \end{aligned} \tag{6}$$

따라서 추정된 행렬을 해밍 거리, 해밍무게 이론을 사용하여 식(7)처럼 추정된 전력 신호 행렬(H)로 변환한다. 해밍무게 모델은 공격 시점에서 암호 알고리즘 연산 후의 값에서 임의의 열벡터를 피어슨 상관관계 연산에 사용한다.

$$H=\begin{bmatrix} h_{1,1}\,h_{1,2} & \ldots h_{1,j} & \ldots h_{1,K} \\ h_{2,1}\,h_{2,2} & \ldots h_{2,j} & \ldots h_{2,K} \\ \vdots & \vdots & \vdots \\ h_{D-1,1}\,h_{D-1,2}\ldots h_{D-1,j} & \ldots & h_{D-1,K} \\ h_{D,1} \quad h_{D,2} \quad \ldots \quad h_{D,j} \quad \ldots & & h_{D,K} \end{bmatrix} \tag{7}$$

피어슨 상관관계 함수는 식(8)이며, 추정된 전력 신호와 측정된 전력 신호 간의 상관계수를 계산한다. 여기서 $h_{d,i}$는 H행렬의 $i(1 \leq i \leq K)$ 번째 열의 요소이고, $\overline{h_i}$는 H행렬의 i번째 열의 평균값이다. $t_{d,j}$는 T행렬의

$j(1 \le j \le T)$번째 열의 요소이고, $\overline{t_i}$는 T행렬의 i번째 열의 평균값이다.

$$r_{i,j} = \frac{\sum_{d=1}^{D}(h_{d,i}-\overline{h_i})^2 \cdot (t_{d,j}-\overline{t_j})}{\sqrt{\sum_{d=1}^{D}(h_{d,i}-\overline{h_i})^2 \cdot \sum_{d=1}^{D}(t_{d,j}-\overline{t_j})^2}} \quad (8)$$

피어슨 상관계수 함수를 사용하여 측정된 소비전력 신호 행렬(T)의 j번째 열벡터 t_j와 추정된 전력 신호 행렬(H)의 j번째 열벡터 h_j와 계산하면 [그림 1]은 측정된 전력 신호와 추정된 전력 신호 간의 상관관계를 나타낸 것이다. $r_{i,j}$에서 i는 비밀키, j는 소비전력 길이를 나타내며, t_j와 h_j의 상관관계가 클수록 큰 값을 갖는다. 이때 매우 높은 상관관계가 발견된 비밀키가 찾고자 하는 비밀키다.

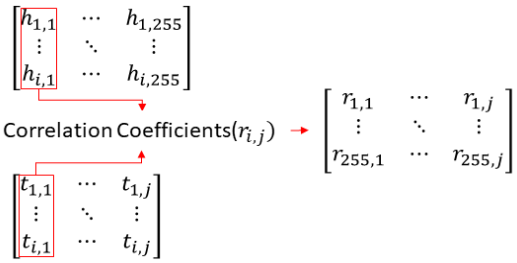

Fig. 1. The Correlation between measured Power Signal and the estimated Power Signal

즉, 전통적인 CPA는 공격 대상 하드웨어를 대상으로 많은 시간과 자원을 요구하며, 공격 성공확률을 높이기 위해 대량의 데이터를 요구하지만 한번 사용한 데이터를 재활용하는 방법을 고려하지 않는 문제점을 가지고 있다.

III. Proposal new CPA method

전통적인 CPA에서 최종적으로 비밀키를 예측하는 함수는 식 (8) 피어슨 상관계수 함수이며, 0부터 255까지 모든 비밀키를 대입하였을 경우 가장 높은 상관계수를 반환하는 비밀키를 최종 결과로 반환했다. 즉, 임의의 데이터를 입력받아 제한된 범주 가운데 한가지의 범주로 분류하는 방법과 같다.

본 논문에서는 인공지능에서 임의의 데이터를 입력받아 제한된 범주에서 하나의 범주를 반환하는 합성곱 신경망(Convolutional Neural Network, CNN)을 식(7)의 해밍거리, 해밍무게와 식(8)의 피어슨 상관계수 함수 대신 사용하여 임의의 소비전력신호를 통해 암호 알고리즘에 사용된 비밀키를 예측해보았다. CNN은 입력받은 데이터에서 정답과 관련성이 높은 데이터를 뽑아서 즉각적으로 결과를 예측한다. 대표적으로 사진에서 사람을 식별하고 위치를 감지할 수 있다. 이 원리를 이용하면 CPA의 많은 데이터를 누적해야 하는 문제를 회피할 수 있다. 실험을 위하여 본 논문에서 사용된 합성곱 신경망은 [Table 1]처럼 구성했다.

Table 1. Detailed settings of the training model

	Attribute	Value
Convolutional (5 layers)	Filter	32
	Kernel size	11
	Strides	2
Fully connection (3 layers)	Units	512
	Activation	ReLU
Model	Learning rate	0.0001
	Batch size	64
	Epoch	100
EarlyStopping	Patience	10

[그림 2](a)와 같이 공격 대상으로 마이크로컨트롤러(Microcontroller unit, MCU) 가운데 IoT 및 아두이노로 상용화된 ATmega328p를 선택하고 AES128의 AddRoundKey 함수의 소비전력을 오실로스코프(MSO2012B)로 측정했다. CPA의 공격이 수집된 데이터의 양에 얼마나 비례하는지 알아보기 위하여 [그림 2](b)처럼 x축은 시간, y축은 비밀키로 모든 비밀키별로 10개의 전력 파형을 수집하여 2,560개의 소비전력 표본을 수집하여 하나의 상자로 묶어 총 16개의 상자를 수집했다.

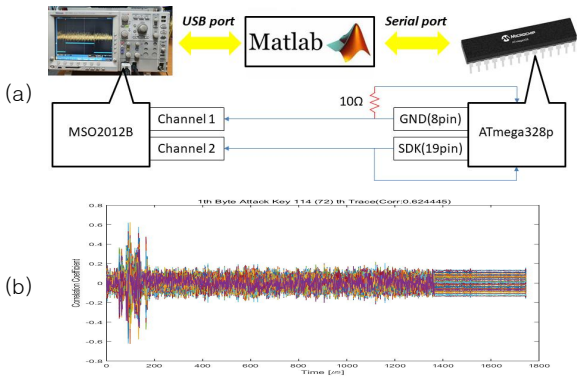

Fig. 2. CPA measurement. (a) Environment, (b) Power signal.

이 가운데 14개의 사례를 해당 모델에 학습에 사용하고 나머지는 검증용으로 사용하였다. [그림 3]은 x축이 AES AddRoundKey의 비밀키를 의미하며, y축이 정답률을 의미하며, CNN에 충분한 학습을 수행하면 소비전력량만 입

력하는 것으로 암호 알고리즘에 사용된 비밀키를 약 95% 확률로 맞출 수 있는 것을 볼 수 있다.

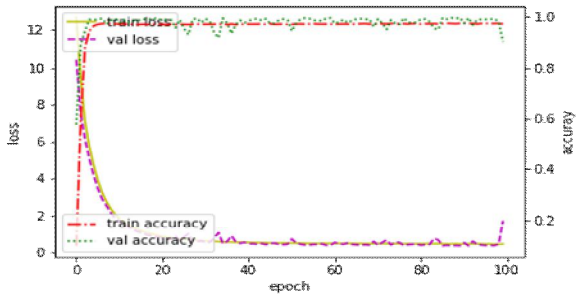

Fig. 3. Model training results with AI(T)=K, dropout, batch normalization and weight initialization

IV. Compared Traditional CPA and new CPA method

4.1 Traditional CPA analysis methods

본 논문의 방식의 차이점을 부각하기 위하여 전통적인 CPA방법을 사용하여 비밀키를 추측해보았다. 먼저 전통적인 CPA방식에서 20개의 소비전력 데이터만 사용하여 상관관계를 분석하면 [그림 4]처럼 뚜렷한 결과를 보여주지 못한다.

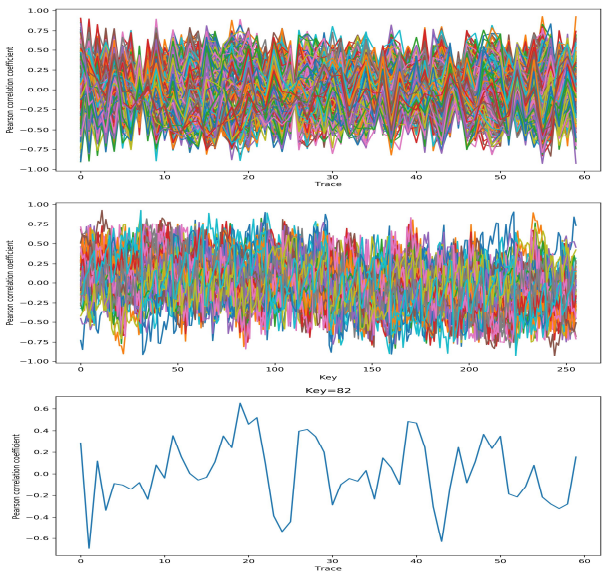

Fig. 4. In case of 20 power consumption data. Top: x-axis: time, y-axis: correlation, middle: x-axis: secret key, y-axis: correlation, bottom: x-axis: time, y-axis: correlation

[그림 4] 상단 그래프는 x축이 시간, y축이 상관계수로 어떤 순간에 비밀키가 연산 되었는지 관측할 수 있다.

중간 그래프는 x축이 비밀키, y축이 상관계수로 모든 비밀키에 대한 상관관계를 검증한 결과이다. 하단 그래프는 비밀키가 82인 경우 특정 시간에서 발견되는 상관관계를 나타낸 것으로 x축이 시간, y축이 상관관계로 표현되어 있다. 480개의 데이터만 사용하여 상관관계를 분석하면 [그림 5]처럼 상단의 시간 축과 상관관계에서 18번째와 27번째의 위치에서 높은 상관관계가 관측되는 것을 볼 수 있다. 또한 몇몇 비밀키가 상관성이 높게 평가되는 것을 중간 그래프로 볼 수 있다. [그림 4]와 비교하여 뚜렷한 차이를 보이지만 중간 그래프를 통해 가장 높은 상관관계의 비밀키가 82가 아닌 124로 잘못된 결과를 나타내고 있는 것을 볼 수 있다.

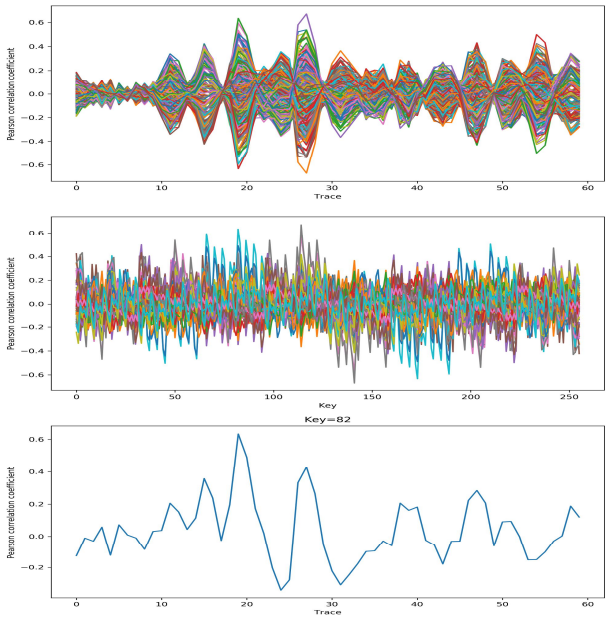

Fig. 5. In case of 480 power consumption data. Top: x-axis: time, y-axis: correlation, middle: x-axis: secret key, y-axis: correlation, bottom: x-axis: time, y-axis: correlation

마지막으로 1,000개의 데이터를 상관관계 분석하면 [그림 6]과 같은 결과를 보이며, [그림 5]와 큰 차이를 보이지 않는다. 다만 상단 그래프에서 가장 높은 상관관계를 나타내던 위치가 변경되고 중단 그래프에서 가장 높은 상관관계를 가지는 비밀키가 82로 변경되어 있음을 볼 수 있다.

다음으로 CPA가 모든 AES의 비밀키에 효과가 있는지 알아보았다. AddRoundKey에서 사용하는 비밀키는 8bit로 선택 가능한 모든 경우는 2^8=256이다. 사용가능한 비밀키 0부터 255까지 총 256가지의 경우를 각각 10번의 소비전력 측정을 수행하여 2,560개의 샘플 단위로 묶어 16개의 샘플을 수집하여 총 40,960개의 소비전력 샘플을 수집하고 1,000개의 서로 다른 소비전력 데이터로 모든 데이

터를 분석했다. 그 결과 모든 비밀키에 대하여 90%를 넘는 예측 정확도를 나타냈다. 그러나 각각의 소비전력 정보를 측정하는 것에 약 5초가 소요되어 많은 양의 데이터를 측정이 필수적인 전통적인 CPA는 효과적이지 못했다.

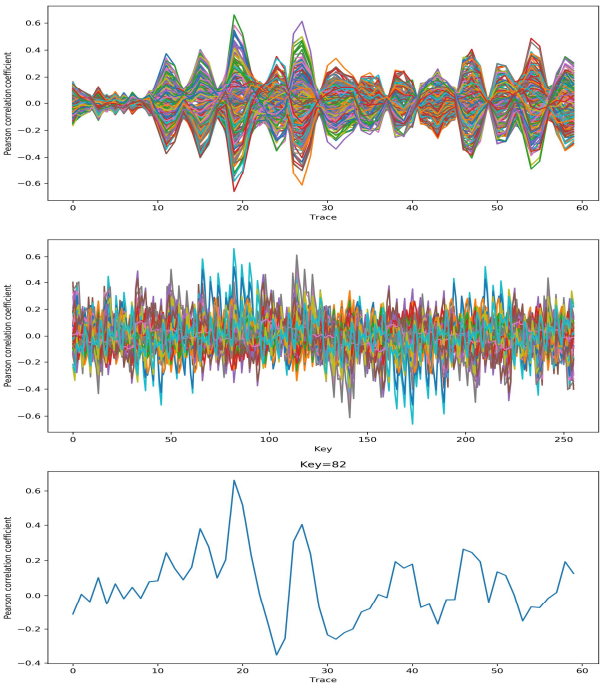

Fig. 6. In case of 1000 power consumption data. Top: x-axis: time, y-axis: correlation, middle: x-axis: secret key, y-axis: correlation, bottom: x-axis: time, y-axis: correlation

4.2 Proposed new CPA analysis method

임의의 소비전력 표본 데이터를 입력하여 예측한 중간키와 각 소비전력 표본별 평문을 연산하여 실제 중간키와 일치하는 횟수를 검증해보았다. [그림 7]은 CPA과 CNN을 복합 사용하는 경우 평균은 96.7%, 분산은 0.0236로 나타났으며, 중간 비밀키 0과 일부를 제외고 약 96%의 정답률을 보였다.

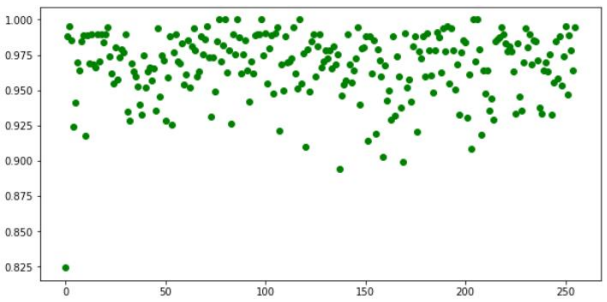

Fig. 7. Training model correct rate

비밀키가 0인 경우 연산이 존재하지 않고 측정부의 외부 잡음이 예측 결과에 많은 영향을 미치는 것으로 보인다. 다음 단계로 소비전력 데이터에서 비밀키의 위치를 알아내기 위해 합성곱 신경망 모델의 상위 8계층의 출력을 추출 및 시각화했다.

Softmax는 모든 출력의 합이 항상 1을 만족하므로 위의 8계층 정보를 Softmax와 결합하여 연산한 결과를 출력에서 값이 높을수록 노란색, 낮을수록 파란색으로 채색한 결과 [그림 8]의 상단 그래프처럼 나타났다. 가장 높은 값(노란색 영역)이 정답으로 예측된 결과와 높은 상관성이 있다는 시각적으로 표현할 수 있다. 하단 그래프는 8비트 비밀키에서 모든 경우에 대한 상관관계를 나타낸 것이다. 가장 높은 상관관계를 나타내는 비밀키가 58로 나타났다.

[그림 9]는 상단이 가중치, 하단이 소비전력을 나타낸다. 첫 번째 컨볼루션 층의 가중치를 내림차순으로 정렬하는 경우 x축이 855일 때, 170.625로 나타났다. 따라서 모든 컨볼루션 층의 가중치를 시각화하면 합성곱 신경망에서 예측 결과를 나타낼 때, 가장 많은 영향을 미친 부분을 추론할 수 있다. 실험결과 중간 영역이 예측 결과에 많은 부분을 차지하고 있음을 볼 수 있다.

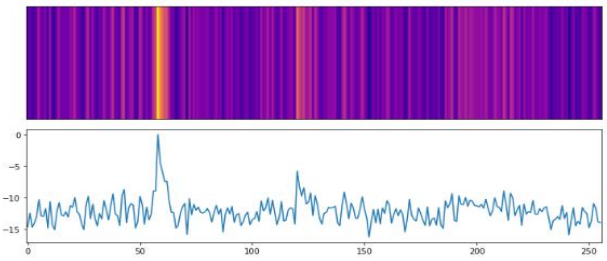

Fig. 8. Secret key verification for power consumption

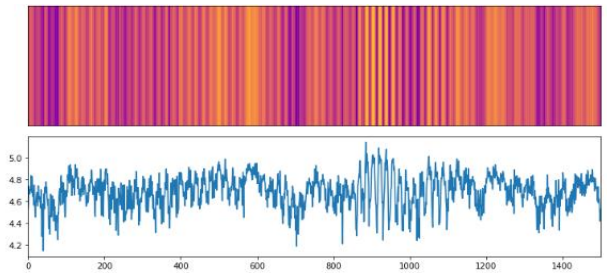

Fig. 9. Change in weight of the first Convolutional layer of the input data

입력된 소비전력 파형은 불규칙하고 급격하게 변하므로 Grad-Cam(Gradient-weighted Class Activation Mapping)을 사용하였다[15]. Grad-Cam은 CNN에서 정답과 관련된 입력 데이터 위치를 찾아내기에 적합한 방법

이다. Grad-Cam을 사용하여 활성함수(ReLU)의 신경망의 가중치와 소비전력신호를 겹치면 [그림 10]과 같다.

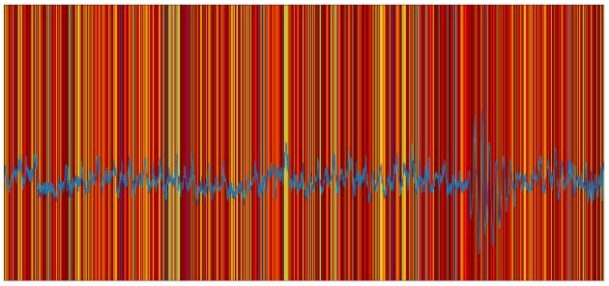

Fig. 10. Power consumption analysis based on the weight of the last hidden layer using Grad-Cam.

가중치가 가장 높은 489번째 구역으로 나타났지만 육안으로는 식별이 어려운 것을 볼 수 있다. CNN은 미세한 변화량을 감지하고 분석하여 결과를 도출 할 수 있다는 것이다.

V. Conclusions

전통적인 CPA는 마이크로컨트롤러에서 고정된 비밀키를 사용하는 암호 시스템에서 서로 다른 평문을 입력하여 다양한 데이터를 사용하여 고정된 비밀키를 예측하는 방법이다. 이 방법은 미세한 소비전력의 변화를 오랜 시간 계측해야 하므로 측정 환경에 민감하게 반응하여 공격을 위해 많은 준비와 시간이 요구되어 실용성이 떨어진다.

본 논문에서는 전통적인 CPA의 장점인 통계적 분석을 유지하면서 즉각적인 분석결과 반환이 가능하도록 CNN을 접목해 보았다. 실험결과 합성곱 신경망을 사용한 경우 모든 비밀키를 대상으로 평균 공격 성공률은 96.7%로 전통적인 CPA보다 좋은 성능을 나타났다. 또한 전통적인 CPA와 비교하여 즉각적인 추정결과 반환에서 극명한 차이점이 나타났다. 현재까지 부채널 공격을 방어하는 방법은 크게 두 가지로 측정을 불가능하게 물리적으로 봉쇄하거나 소프트웨어적으로 중요 함수의 연산 결과에 무작위 난수를 추가하는 것이다. 따라서 추후 연구방향으로 인공지능을 활용한 공격 기법을 응용하여 방어하는 함수를 설계하는 방법도 고려될 수 있다고 기대한다.

ACKNOWLEDGEMENT

This work was supported by Dongseo University, ``Dongseo Cluster Project'' Research Fund of 2021 (DSU-20210001).

REFERENCES

[1] S. Mangard, E. Oswald, T. Popp, “Power analysis attacks: Revealing the secrets of smart cards,” Springer Science & Business Media, 2008. ISBN:978-1-4419-4039-1

[2] P. C. Kocher, "Timing attacks on im- plementations of Diffie-Hellman, RSA, DSS, and other systems." Annual International Cryptology Conference. Springer, pp. 104-113, Berlin, Heidelberg, 1996. DOI: 10.1007/3-540-68697-5_9

[3] P. Kocher, J. Jaffe, and B. Jun, "Differential power analysis," Annual international cryptology conference. Springer, pp. 388-397, Berlin, Heidelberg, 1999. DOI: 10.1007/3-540-48405-1_25

[4] D. G. Kwon, al. “Study of Deep Learning Based Trace Preprocessing Method with Denoising Auto-Encoder,” Conference on Information Security and Cryptography-Winter, 2018.

[5] Loic Masure et al. “Gradient Visualization for General Characterization in Profiling Attacks,” International Workshop on Constructive Side-Channel Analysis and Secure Design, 2019. DOI: 10.1007/978-3-030-16350-1_9

[6] Pieter Robyns et al. “Improving CEMA using Correlation Optimization,” IACR Transactions on Cryptographic Hardware and Embedded Systems, pp. 1-24, 2019. DOI: 10.13154/tches.v2019.i1.1-24

[7] HanBit Kim, HeeSsok Kim, “The trend of research on sub-channel analysis security conference examined through CHES 2020,” Korea Institute of Information Security and Cryptology, Vol. 30, No. 6, pp. 67-81, 2020.

[8] Camurati, Giovanni, Aurélien Francillon, and François-Xavier Standaert, "Understanding screaming channels: From a detailed analysis to improved attacks," IACR Transactions on Cryptographic Hardware and Embedded Systems, pp. 358-401, 2020. DOI: 10.13154/tches.v2020.i3.358-401

[9] Gnad, D. R., Krautter, J., & Tahoori, M. B, “Leaky noise: New side-channel attack vectors in mixed-signal IoT devices,” IACR Transactions on Cryptographic Hardware and Embedded Systems, pp. 305-339, 2019. DOI: 10.13154/tches.v2019.i3.305-339

[10] E. Brier, C. Clavier, and F. Olivier, "Correlation Power Analysis with a Leakage Model", CHES'04, LNCS 3156, pp. 16-29, 2004. DOI: 10.1007/978-3-540-28632-5_2

[11] Lo, O. Buchanan, W. J, & Carson, D, “Correlation power analysis

on the PRESENT block cipher on an embedded device," In Proceedings of the 13th International Conference on Availability, Reliability and Security, pp. 1-6, 2018. DOI: 10.1145/3230833 .3232801

[12] Nuradha, F. R., Putra, S. D., Kurniawan, Y., & Rizqulloh, M. A., "Attack on AES Encryption Microcontroller Devices With Correlation Power Analysis," In 2019 International Symposium on Electronics and Smart Devices (ISESD), pp. 1-4. October 2019. DOI: 10.1109/ISESD.2019.8909447

[13] Durvaux, F., & Durvaux, M, "SCA-Pitaya: A Practical and Affordable Side-Channel Attack Setup for Power Leakage— Based Evaluations," Digital Threats: Research and Practice, 1(1), pp. 1-16, March 2020. DOI: 10.1145/3371393

[14] Hong-Pil Kwon, Jae-Cheol Ha, "Power Analysis Attack of Block Cipher AES Based on Convolutional Neural Network," Journal of the Korea Academia-Industrial cooperation Society, Vol. 21, No.5, 14-21. 2020. DOI: 10.5762/KAIS.2020.21.5.14

[15] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," arXiv, Preprint posted online 7, October 2016. DOI: 10.1007/s11263-019-01228-7

Authors

Ki-Hwan Kim received the B.S., M.S. Ph.D. degree in Computer Networking from Dongseo University, Republic of Korea in 2015. M.S., Ph..D. degree in Department of Ubiquitous IT from Dongseo University 2017, 2021. Dr. Kim received the B.S., M.S. Ph.D. degree in Computer Networking from Dongseo University, Republic of Korea in 2015. M.S., Ph..D. degree in Department of Ubiquitous IT from Dongseo University 2017, 2021. Since 2021 he is now working for the International College for the Department of Computer Engineering of Dongseo University as a visiting professor. His research interests are cryptography, Information security and Side-Channel Attack(SCA), Artificial Intelligent(AI).

HyunHo Kim received his B.S. degrees in computer science from Dongseo University in 2013. M.S., Ph.D. degree in Department of Ubiquitous IT from Dongseo University in 2015, 2020.

Dr. Kim received the B.S. degrees in computer science from Dongseo University in 2013. M.S., Ph.D. degree in Department of Ubiquitous IT from Dongseo University in 2015, 2020. Since 2020 he is now working for the Software and Convergence College for the Department of Computer Engineering of Dongseo University as a visiting professor. His research interests include Digital forensic, Information security, IoT security, and Network security.

Hoon Jae Lee received the B.S., M.S. and Ph.D. degree in Electrical Engineering from Kyungpook national university in 1985, 1987 and 1998, respectively. Dr. Lee had been engaged in the research on cryptography and network security at Agency for Defense Development from 1987 to 1998. Since 2002 he has been working for Department of Computer Engineering of Dongseo University as an associate professor, and now he is a full professor. His current research interests are in security communication system, side-channel attack, USN & RFID security. He is a member of the Korea institute of Information security and cryptology, IEEE Computer Society, IEEE Information Theory Society and etc.