Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.31, NO.6, Dec. 2021 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2021.31.6.1127

연속 웨이블릿 변환을 사용한 비프로파일링 기반 전력 분석 공격*

배 대 현,[1†] 이 재 욱,[2] 하 재 철[3‡]
[1,2,3]호서대학교 (대학원생, 학생, 교수)

Non-Profiling Power Analysis Attacks Using Continuous Wavelet Transform Method*

Daehyeon Bae,[1†] Jaewook Lee,[2] Jaecheol Ha[3‡]
[1,2,3]Hoseo University (Graduate student, Student, Professor)

요 약

전력 분석 공격에서 소비 전력 파형의 잡음과 정렬 불량은 공격 성공 여부를 좌우하는 주요한 요인이다. 따라서 이를 완화하기 위한 여러 연구가 수행되고 있으며 웨이블릿 변환 기반의 신호처리 방법도 그중 하나이다. 대부분의 웨이블릿을 사용한 연구에서는 파형 압축할 수 있는 이산 웨이블릿 변환을 사용해 왔는데, 그 이유는 연속 웨이블릿 변환 기법이 선택된 스케일의 개수에 따라 데이터 크기 및 분석 시간이 증가할 뿐만 아니라 효율적인 스케일 선택 방법도 없기 때문이다. 본 논문에서는 전력 분석 공격에 최적화된 연속 웨이블릿 변환의 효율적인 스케일 선택 방법을 제안하며 이를 이용해 파형을 인코딩할 경우 분석 성능이 크게 향상될 수 있음을 보인다. 비프로파일링 공격인 CPA(Correlation Power Analysis) 및 DDLA(Differential Deep Learning Analysis) 공격 실험 결과, 제안하는 방법이 잡음 감쇄와 파형 정렬에 효과적임을 확인하였다.

ABSTRACT

In the field of power analysis attacks, electrical noise and misalignment of the power consumption trace are the major factors that determine the success of the attack. Therefore, several studies have been conducted to overcome this problem, and one of them is a signal processing method based on wavelet transform. Up to now, discrete wavelet transform, which can compress the trace, has been mostly used for power side-channel power analysis because continuous wavelet transform techniques increase data size and analysis time, and there is no efficient scale selection method. In this paper, we propose an efficient scale selection method optimized for power analysis attacks. Furthermore, we show that the analysis performance can be greatly improved when using the proposed method. As a result of the CPA(Correlation Power Analysis) and DDLA(Differential Deep Learning Analysis) experiments, which are non-profiling attacks, we confirmed that the proposed method is effective for noise reduction and trace alignment.

Keywords: Implementation Attack, Hardware Security, Artificial Intelligence, Deep Learning, Wavelet Transform

Received(09. 27. 2021), Modified(11. 05. 2021), Accepted(11. 08. 2021)

* 이 논문은 2020년도 호서대학교의 재원으로 학술연구비 지원을 받아 수행된 연구임.(20200848)

* 본 논문은 2021년도 한국정보보호학회 충청지부 학술대회에 발표한 우수논문을 개선 및 확장한 것임.

† 주저자, noeyheadb@gmail.com

‡ 교신저자, jcha@hoseo.edu(Corresponding author)

I. 서 론

부채널 분석은 하드웨어로부터 누출되는 부채널 정보를 분석해 암호 키 등의 비밀 정보를 추출하는 공격의 일종이다. 이때 부채널 정보란 합의되지 않은 모든 채널의 정보를 의미하며 대표적으로 소비 전력, 전자기파, 시간차, 소리, 발열량 등이 있다. 1996년 P. Kocher에 의해 시간차 정보를 이용한 타이밍 공격으로 처음 제안된 부채널 분석은 현재까지도 암호용 디바이스는 물론 상용 컴퓨터 프로세서 내의 비밀 정보까지 찾아낼 수 있는 위협적인 공격으로 평가받고 있다[1-2].

부채널 분석에서 가장 활발히 연구되는 전력 분석 공격은 스마트 카드나 마이크로 컨트롤러(MicroController Unit, MCU) 등과 같은 디바이스의 전력 소비량을 분석함으로써 공격이 수행되는데, 이를 위해서는 소비 전력 파형의 수집 단계가 선행되어야 한다. 파형 수집 단계에서는 공격 대상 암호 알고리즘이 동작하는 MCU의 접지 혹은 전원부에 저항을 직렬로 연결하고 해당 저항에 걸리는 전압을 오실로스코프로 측정한다. 이 과정에서 전력 파형에 전기적 잡음이 포함될 수 있으며 트리거 시점에 따라 신호가 정확히 정렬되지 않을 수도 있다. 이러한 잡음이나 파형의 정렬은 전력 분석 공격의 성공을 좌우하는 주요 요인이므로 이를 완화하기 위한 연구가 수행되고 있다[3-4].

이러한 연구의 일환으로 웨이블릿 변환 기반의 파형 전처리 연구도 수행되었는데, 대부분의 연구에서는 이산 웨이블릿 변환(Discrete Wavelet Transform, DWT)을 사용하고 있다.[5-7]. 이산 웨이블릿 변환과 연속 웨이블릿 변환(Continuous Wavelet Transform, CWT)의 가장 큰 차이점은, DWT는 2^n 스케일에 해당하는 주파수를 단계별로 필터링하여 파형을 압축시키는 반면 CWT는 특정 스케일을 선택하고 해당 스케일에 대응되는 주파수의 정보를 추출하며 파형을 분해한다는 것이다. 즉, CWT는 스케일의 개수만큼 데이터의 크기가 증가하게 되며 이에 따른 분석 시간 증가, 효율적인 스케일 선택 방법의 부재 등이 단점으로 작용해 부채널 정보 기반 역어셈블러(side-channel-based disassembler) 연구에서만 일부 사용하고 그 외에는 잘 사용되지 않는다[8-10].

본 논문에서는 이러한 연속 웨이블릿 변환을 이용해 파형을 인코딩함으로써 소비 전력 파형의 잡음 감쇄 및 정렬을 수행해 전력 분석 공격의 성능을 향상시킬 수 있는 효율적인 방안을 제시하고자 한다. 특히, MCU의 소비 전력 파형에 최적화된 모 웨이블릿(mother wavelet)의 스케일 선택 방법을 제안하며, 이렇게 선택된 최소한의 스케일만으로 파형을 인코딩했을 때의 성능 향상을 기존 전력 분석 공격의 성능과 비교해 볼 것이다. 여러 가지 데이터 셋을 이용하여 CPA(Correlation Power Analysis)[11] 및 DDLA(Differential Deep Learning Analysis)[12]와 같은 비프로파일링(non-profilig) 기반의 전력 분석 실험 결과, 제안하는 방법이 기존 DWT 기반 전처리 성능을 능가하며 전력 분석 공격의 성능을 크게 향상시킬 수 있음을 확인하였다.

II. 연속 웨이블릿 변환

웨이블릿 변환(wavelet transform)이란, 신호처리 분야에서 불확실성의 원리(uncertainty principle)로 인해 주파수 해상도와 시간 해상도간 상충관계가 존재하는 단시간 푸리에 변환(Short-Time Fourier Transform, STFT)의 한계점을 극복할 수 있는 다중 해상도의 시간-주파수 분석(time-frequency analysis) 기법이다. 웨이블릿 변환 함수 F는 시간 t에 대한 분석 대상 신호를 $x(t)$, 모 웨이블릿을 $\psi(t)$, 스케일 팩터를 s, 전이 팩터를 τ라 할 때 수식 (1)과 같이 정의된다.

$$F_{\psi}(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t)\, \Psi\left(\frac{t-\tau}{s}\right) dt \qquad (1)$$

이때 모 웨이블릿 Ψ는 다음 수식 (2)를 만족하며 0을 기준으로 증가와 감소하는 짧은 진동 형태의 함수이다. 이는 스케일 팩터 s에 따라 확장(stretched)되거나 압축(compressed)된다. 대표적인 모 웨이블릿으로는 Mexican hat, Morlet, Gaussian, Daubechies 등이 있다. 본 논문의 실험에는 Mexican hat 웨이블릿을 사용한다.

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \qquad (2)$$

웨이블릿 변환의 일종인 연속 웨이블릿 변환은 주어진 여러 스케일에 대하여 웨이블릿 변환을 수행해

1차원 시간 도메인의 신호로부터 2차원 시간-주파수 도메인의 데이터로 분해하며 이를 나타낸 이미지를 스켈로그램(scalogram)이라 한다. 이때 각각의 스케일 s로 추출할 수 있는 신호의 주파수 f_s는 다음 수식 (3)과 같다. 여기서 f_m는 모 웨이블릿의 중앙 주파수(center frequency)를, Δ는 각 샘플간 시간 간격(sampling period)을 의미하며 주파수는 Hz, 시간은 초 단위다.

$$f_s = \frac{f_m}{\Delta \times s} \tag{3}$$

CWT를 수행한 예시를 보이기 위해 상용 MCU의 클럭 사이클에 따른 소비 전력 파형을 분해하여 스켈로그램으로 나타낸 것이 Fig. 1.이다. Fig. 1.의 상단은 1차원 시간 도메인으로 이루어진 소비 전력 파형을 나타낸 것이며 하단은 이를 시간-주파수 도메인의 2차원 데이터로 분해한 것이다.

본 논문에서는 상대적인 개념인 고주파와 저주파를 MCU의 동작 주파수를 기준으로 구분한다. 즉, 동작 클럭보다 큰 주파수는 고주파 대역으로, 작은 주파수는 저주파 대역으로 통칭한다. 그 이유는 MCU의 소비 전력 정보에서 실질적으로 공격 성능에 영향을 미치는 상승 엣지와 하강 엣지 부분의 주기가 동작 클럭의 주기와 같으므로, 이보다 큰 주파수의 정보는 잡음으로 간주할 수 있기 때문이다.

전력 분석 공격에서는 전기적 잡음 감쇄와 파형의 정렬 효과 때문에 저주파 대역의 신호를 주로 사용한다. 전기적 잡음은 고주파 신호로 구성되는 점을 고려하여, CWT를 통해 저주파 신호를 추출한다면 고주파 정보를 자연스럽게 필터링하는 효과를 볼 수 있다. 그 이유는 모든 샘플에 존재하는 가우시안 분포의 전기적 잡음은 저주파 대역 신호를 계산할 때 누적되어 상쇄되기 때문이다.

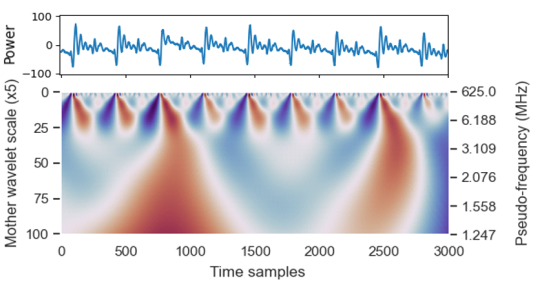


Fig. 1. Example of a decomposed power trace using CWT.

지터(jitter)로 인한 위상 잡음 역시 고주파 신호에서는 매우 크지만 저주파 신호일수록 영향이 적다. 그 이유는 저주파 대역의 신호를 추출할 때 여러 샘플이 하나의 CWT 계수(coefficient)로 조합되어 파형의 정렬에 둔감한 성질을 갖기 때문이다. 따라서 전력 분석 공격에 대응하기 위해 인위적인 지터나 잡음을 추가하는 하이딩(hiding) 기법을 구현하더라도 CWT를 사용한다면 이러한 대응책들을 일정 수준까지 무력화시킬 수 있다.

III. CWT 기반 전력 파형 인코딩

CWT를 이용해 파형을 인코딩하기에 앞서 모 웨이블릿의 스케일 선택 작업이 선행되어야 한다. 모 웨이블릿의 스케일은 추출되는 신호의 주파수를 결정하며 분석 당사자가 선택해야 하는 중요한 매개변수이다. 하지만 현재까지 대부분의 CWT를 적용한 SCBD 관련 연구에서는 명령어의 특징 추출 단계가 뒤따르기 때문에 Fig. 1.과 같이 폐구간 $[1,N]$에 속하는 자연수를 스케일로 사용하기도 한다. 그러나 선택된 스케일의 개수와 데이터의 크기가 비례하기 때문에 중요한 정보가 포함된 스케일만을 최소한으로 선택하는 것이 비용 측면에서 유리하다. 따라서 본 장에서는 전력 분석 공격의 특성을 반영한 모 웨이블릿의 스케일 선택 방법을 제안하고자 한다.

3.1 모 웨이블릿의 스케일 선택

일반적인 공격 시나리오에서는 공격자가 직접 오실로스코프와 같은 측정 장비를 사용해 대상 MCU의 전력 파형을 수집하기 때문에 표본화율과 동작 클럭 주파수를 알 수 있다고 가정한다. 상기한 수식 (3)에서 보는 바와 같이 스케일이 같을지라도 전력 파형이 측정될 때의 표본화율과 웨이블릿의 종류에 따라 추출할 수 있는 신호의 주파수가 달라진다. 따라서 앞선 수식 (3)의 관계를 이용해 추출하고자 하는 주파수로부터 스케일 s를 역으로 계산해 사용하는 것이 비용, 성능 등의 측면에서 유리할 것이다.

앞서 언급한 바와 같이 클럭 주파수보다 큰 고주파 신호는 잡음으로 간주할 수 있다. 또한, 전력 분석 공격에서 공격 대상이 되는 암호 알고리즘의 중간 연산 값은 오직 한 클럭이 아닌 여러 클럭에 거쳐 처리된다. 즉, 여러 클럭의 정보가 조합된 포괄적 정보인 저주파 신호를 사용하는 것이 가장 좋은 공격 성

능을 보일 것이다. 따라서 조합할 클럭의 개수를 c ($c \in \mathbb{R}^+$), MCU의 동작 클럭 주파수를 f_r이라 할 때 수식 (4)와 같은 스케일 결정 방법을 제안한다.

$$s = \frac{f_m}{\Delta \times \frac{f_r}{c}} \tag{4}$$

만약 표본화율이나 동작 주파수 등의 세부 정보를 알 수 없고 오직 파형에만 접근할 수 있다면 다음 수식 (5)에 의해 1클럭 사이클의 샘플 수를 의미하는 SPC(Samples Per Clock)를 사용해 s를 계산할 수 있다. SPC는 파형으로부터 반복되는 패턴을 찾아 추측할 수 있으며, 이러한 SPC의 근사치 $\widehat{SPC}$를 사용한다면 수식 (6)과 같이 스케일의 근사치 역시 계산할 수 있다.

$$SPC = \frac{1}{\Delta} \times \frac{1}{f_r} \approx \widehat{SPC} \tag{5}$$

$$s = \frac{f_m}{\Delta \times \frac{f_r}{c}} \approx \frac{f_m}{\frac{1}{\widehat{SPC}} \times \frac{1}{c}} \approx f_m \times \widehat{SPC} \times c \tag{6}$$

이때 조합할 클럭의 개수는 c개의 클럭 주기를 갖는 주파수 대역의 정보를 추출하기 위함이며 실제 웨이블릿과 컨볼루션 연산이 수행되는 샘플 수가 c개의 클럭 주기와 같음을 의미하지는 않는다. c는 공격 대상 연산의 종류, 구현 방법(어셈블리어의 구성)에 따라 달라질 수 있는데 이는 공격자가 모델링한 전력 값과 상관관계가 있는 암호 알고리즘의 중간 연산 값이 몇 개의 클럭에 거쳐 연산 되는지에 따라 선택되어야 한다. 또한, 지터의 정도가 클 경우에는 더욱 넓은 범위에 걸쳐 웨이블릿 변환을 수행해야 하므로, c는 지터의 정도와 비례하게 큰 값으로 사용되어야 한다. 최적의 c를 선택하는 결정적 알고리즘은 존재하지 않으며 공격자가 어셈블리어 분석을 통해 대략적인 c의 범위를 유추하여 후보군을 선택하고 이에 대해 전수조사를 수행해야 한다.

3.2 잡음 감쇄 및 파형 정렬 효과 검증

전력 분석 공격에서 CWT의 잡음 감쇄 및 파형 정렬에 대한 예시를 보이기 위해 32비트 STM32F3(Cortex-M4) MCU의 소비 전력 파형에 대한 신호대 잡음 비(Signal-to-Noise Ratio, SNR)를 분석하였다. 먼저 Fig. 2.의 상단은 아무런 전처리를 하지 않은 파형에 대한 SNR을 나타낸 것인데 1을 넘지 않음을 확인할 수 있다. 그러나 하단과 같이 제안한 스케일 결정 방법과 CWT를 통해 분해한다면 특정 주파수에서 ($c=8$, 0.92MHz) 5가 넘는 높은 SNR을 갖는 것을 확인할 수 있다.

따라서 이렇게 최적의 c를 사전에 분석하여 파형을 인코딩한다면 전력 분석 공격의 성능을 향상시킬 수 있다. 한편 기존 연구에서 많이 사용되던 DWT를 적용할 경우에는 Fig. 3.과 같이 모든 레벨에서 최대 SNR이 3을 넘지 못하는 것을 확인할 수 있다. 이를 통해 제안하는 CWT 방법이 기존 DWT 기반 전처리보다 우수함을 확인할 수 있다.

다음으로 파형 정렬 효과를 보이기 위해 앞서 사용된 파형에 균등분포 $U[0,300]$에서 임의로 선택한 정수만큼 샘플을 이동시킨 후 SNR을 계산한 결과는 Fig. 4.와 같다. 지터가 없는 파형에 비해 더 낮은 주파수($c=24$, 0.31MHz)에서 원본 파형보다도 높은 1.5정도의 SNR을 갖는 것을 확인할 수 있다.

또한, SNR이 최대인 지점에서 해밍 웨이트에 따른 전력 값의 확률 분포를 시각화하면 FIg. 5.와 같은데, 아무런 전처리를 하지 않은 상단의 파형은 분포가 많이 중첩되는 것을 확인할 수 있다. 반면 제안한 방법을 통해 인코딩한 하단 파형은 중첩의 정도가 작으며 더욱 잘 구분되는 것을 확인할 수 있다.

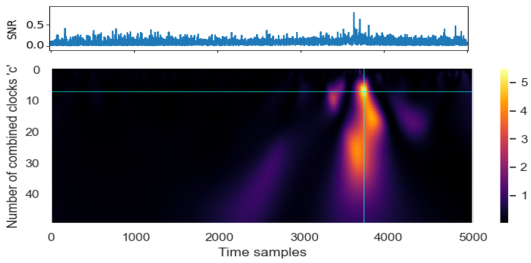


Fig. 2. Comparison of SNR values between raw and decomposed trace using CWT.

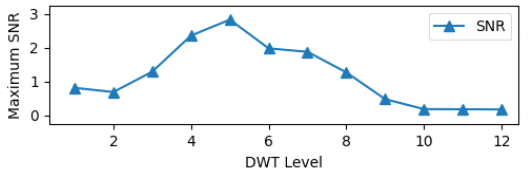


Fig. 3. Maximum SNR values according to DWT level.

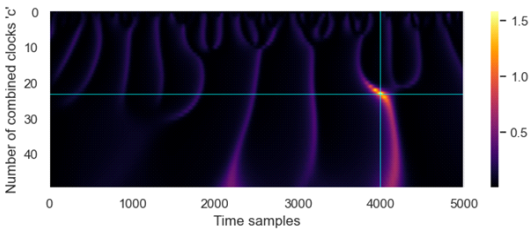


Fig. 4. The SNR values of the power consumption trace with artificial jitter.

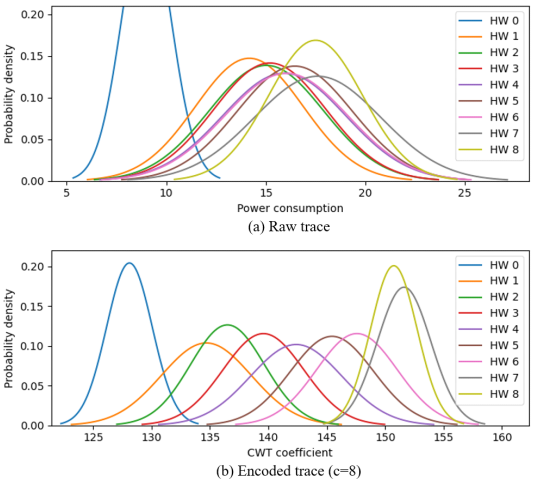


Fig. 5. Probability distribution of power consumptions according to Hamming weight.

IV. 비프로파일링 전력 분석 실험 및 결과 분석

본 논문에서 제안하는 스케일 선택 방법과 CWT 기반 파형 인코딩 방법이 지터와 잡음이 많은 파형 분석에서 좋은 성능을 보임을 검증하기 위해 3가지 데이터 셋을 이용하여 실험을 진행하였다. 모든 데이터 셋에서 실험 대상 암호 알고리즘은 AES-128이며 1라운드 SubBytes 연산만을 분석한다. 실험에 사용된 3가지 데이터 셋은 다음과 같다.

1) 직접 수집한 ChipWhisperer 보드 파형
2) 공개 데이터 셋 SCARF [13]
3) 공개 데이터 셋 ASCAD [14]

직접 수집한 데이터 셋은 Fig. 6.과 같은 환경에서 오실로스코프를 이용해 32비트 STM32F3 MCU가 탑재된 ChipWhisperer 타겟 보드의 소비 전력을 수집하였다. 이에 대한 구체적인 사양 및 정보는 Table 1.과 같다.

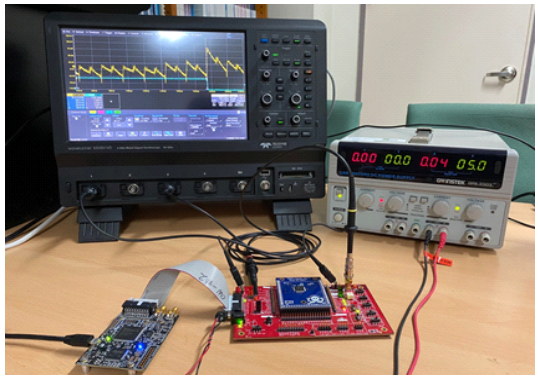


Fig. 6. Experimental setup for ChipWhisperer dataset.

Table 1. Experimental setup.

Category	Specification
Oscilloscope	Lecroy WaveRunner 8404M-MS (4GHz bandwidth, 8bit resolution)
Samp. rate	500MS/s
Target device	ChipWhisperer UFO board with Cortex-M4-based STM32F3 (32bit)
Clock freq.	7.37MHz

두 종류의 공개 데이터 셋은 실험의 객관성과 재현 가능성을 위해 사용하였다. 구체적으로 SCARF 데이터 셋은 32비트 STM32F4 MCU에서 부채널 분석 대응책이 적용되지 않은 AES-128가 고정된 키로 동작할 때 수집된 파형을 사용한다. 또한, ASCAD 데이터 셋 중에서는 ATMega8515 8비트 MCU에서 1차 마스킹 대응책이 적용된 AES-128이 고정된 키로 동작할 때 수집한 파형을 사용한다. 실험에 사용된 두 가지 공개 데이터 셋의 세부 사양은 Table 2.와 같다.

4.1 ChipWhisperer 데이터 셋에 대한 CPA 실험

가장 널리 사용되는 비프로파일링 부채널 분석 기법인 CPA 실험을 위해 인위적으로 지터 또는 잡음을 추가한 3가지 파형에 대해 실험을 진행한다. 첫째로 원본 파형에 대해, 둘째로 $U[0,300]$에서 샘플링한 지터를 추가한 파형에 대해, 마지막으로 지터가 추가된 파형에 가우시안 분포 $N(0,10)$에서 샘플링한 잡음을 추가한 파형에 대해 분석한다. 이때 공격에 필요한 최소 파형 수(Minimum Traces to

Table 2. Specification of open datasets for side-channel analysis.

Dataset	Specification	
SCARF	Target	Naive AES-128
	MCU	STM32F4 (32bit)
	Clock freq.	16MHz
	Samp. rate	250MS/s
ASCAD	Target	1^{st}-order Masked AES-128
	MCU	ATMega8515 (8bit)
	Clock freq.	4MHz
	Samp. rate	2GS/s

Disclosure, MTD)를 분석하기 위해 10개 ~ 1,000개의 파형에 대해 10개 단위로 개수를 증가시키며 공격을 수행한다.

원본 파형, 지터가 추가된 파형, 그리고 지터와 잡음이 추가된 파형에 대해 각각 전처리 여부에 따라 총 6가지의 파형을 분석하고 PGE(Partial Guessing Entropy)를 나타낸 것이 Fig. 7.과 같다. 이때 PGE란, 올바른 키가 예측된 후보 키의 몇 등(rank)에 위치했는지를 나타내는 지표이다.

실험 결과 인위적인 지터와 잡음이 추가되지 않은 경우에도 CWT를 전처리를 수행하면 MTD가 140에서 30까지 줄어들고, 지터를 추가했을 때는 1,000 이상에서 70까지, 잡음까지 추가되었을 때는 1,000 이상에서 100까지 줄어드는 것으로 보아 월등한 성능 향상이 있음을 알 수 있다. 이때 사용한 c는 앞서 언급한 방법과 같이 대략적인 c의 범위를 유추하고 후보군에 대한 전수조사를 통해 선택된 것이다. 이에 대한 세부 실험 결과는 Table 3.과 같다.

다음 Fig. 8.은 앞서 사용된 파형들을 각각 100개씩 겹쳐 그린 것이며 (a)는 원본 파형, (b)는 지터와 잡음이 추가된 파형, (c)는 (b)의 파형에 대해 $c=24$로 인코딩한 파형이다. (b)에서 보는 바와 같이 지터와 잡음이 추가된 파형은 형상을 구분하기 힘

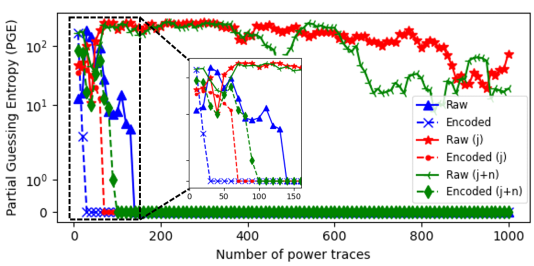


Fig. 7. The result of PGE analysis for ChipWhisperer dataset.

Table 3. Comparison of MTD according to hiding methods and encoding.

Hiding methods	Encoding	MTD
None	N/A	140
	$c=8$	30
Jitter ~ $U[0,300]$	N/A	1,000↑
	$c=24$	70
Jitter ~ $U[0,300]$ Noise ~ $N(0,10)$	N/A	1,000↑
	$c=24$	100

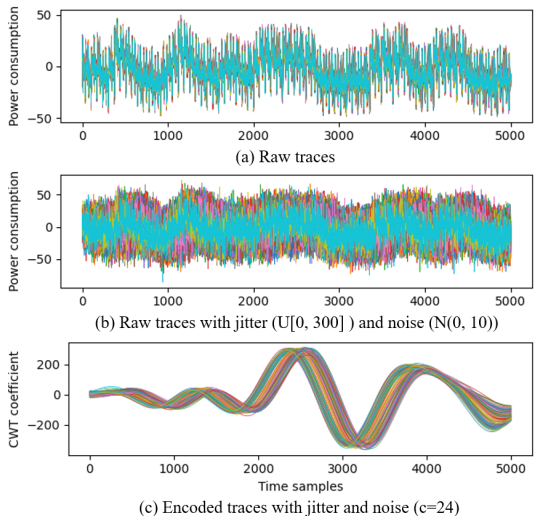


Fig. 8. Comparison of trace shapes according to hiding and encoding methods.

들지만, $c=24$로 인코딩한다면 상대적으로 정렬된 깨끗한 파형을 확인할 수 있다.

4.2 SCARF 데이터 셋에 대한 CPA 실험

다음으로 공개 데이터 셋인 SCARF 데이터 셋에 대한 CPA 분석을 수행하였다. 분석 대상은 AES의 1라운드 2번째 바이트 SubBytes 연산이며, ShiftRows 또는 MixColumns 부근의 파형 없이 오직 SubBytes 부근만 분석한다면 많은 잡음으로 인해서 제공된 2,000개의 파형 모두를 사용해도 키를 찾을 수 없다. 그러나 사전 분석을 통해 결정된 $c=6$을 이용해 파형을 인코딩하면 Fig. 9.와 같이 MTD가 2,000 이상에서 1,180까지 줄어드는 것을 확인할 수 있다. 이때 각 파형 수에 따른 상관계수의 절대값 추이는 Fig. 10.과 같다. 상관계수 역시 기존 0.13에서 0.23으로 증가한 것을 볼 수 있다.

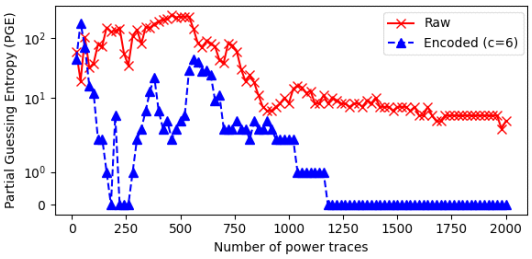


Fig. 9. The result of PGE analysis for SCARF dataset.

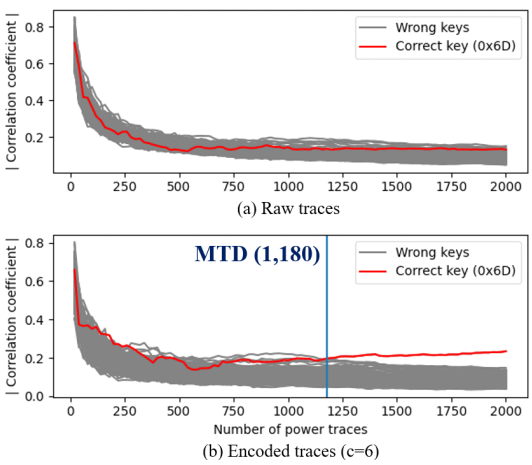


Fig. 10. Comparison of absolute correlation coefficients according to the number of traces and encoding.

4.3 ASCAD 데이터 셋에 대한 DDLA 실험

마지막으로 딥러닝 기반의 비프로파일링 부채널 분석 기법인 DDLA(Differential Deep Learning Analysis)[14] 실험을 수행하기 위해 ASCAD 데이터 셋에서 기본으로 제공하는 ASCAD.h5(desync0)와 ASCAD_desync50.h5 파형 셋을 사용하였다. ASCAD_desync50는 ASCAD 파형을 균등분포 $U[0,50]$에서 샘플링한 정수만큼 파형을 이동시킨 것이다.

구체적인 공격 대상은 1차 마스킹이 적용된 1라운드 3번째 바이트 SubBytes 연산이다. 문헌 [12]에 따르면 DDLA에서 마스킹이 적용된 알고리즘을 공격할 때, 공격 대상 레이블을 1차(first-order) 공격과 같게 설정한다면 딥러닝 모델이 마스크 값을 내부적으로 조합해 2차(second-order) 공격이 가능하다고 알려져 있다. 따라서 본 장의 실험에서도 이와 같은 방법으로 2차 DDLA 공격을 수행하였다. 공격에는 가장 단순한 형태의 딥러닝 모델인 다층 퍼셉트론(Multi-Layer Perceptron, MLP)을 사용하였으며, 모델 구성과 공격 관련 매개변수는 Table 4.와 같다.

실험에서는 desync0의 원본 파형, desync50의 원본 및 $c=1$로 인코딩된 파형에 대해 분석을 수행한다. 각각 파형 셋에서 1,000 ~ 10,000개의 파형에 대해 1,000개씩 증가시키며 DDLA 분석을 수행하고 가장 좋은 성능을 보이는 학습 평가 지표를 기록한다. 이때, 공격 성능의 평가 지표로는 문헌 [15]에서 제안한 NMM(Normalized Maximum Margin)을 사용한다. NMM은 올바른 키와 잘못된 키의 학습 평가 지표 값의 차이를 σ(표준편차) 단위로 나타낸 것이다. 만약 NMM이 0보다 크다면 키를 찾을 수 있음을, 0보다 작다면 키를 찾을 수 없음을 의미한다.

사용된 파형 수에 따른 NMM 분석 결과는 다음 Fig. 11.과 같다. 실험 결과, desync0 파형의 MTD는 약 3,000 그리고 인코딩된 desync50 파형의 MTD는 약 5,000으로 분석되며 원본 desync50 파형에 대해서는 10,000개를 사용해도 키를 찾을 수 없는 것을 알 수 있다. 보다 자세한 실험 결과를 위해 10,000개의 파형을 사용했을 때의 각각의 파형 셋에 대한 최적의 학습 평가 지표 값의

Table 4. Detailed parameters for DDLA.

Category	Specification	
Deep learning Model (MLP)	Input layer	700 nodes
	Hidden layer	30 nodes (ReLu)
	Output layer	2 nodes (Softmax)
Loss func.	Mean_Squared_Error (MSE)	
Optimizer	Adam (lr=0.001)	
Labeling method	LSB (Least Significant Bit)	
Epoch/ batch size	50 / $\frac{1}{10}\times$ # of training data	
Training/ validation ratio	0.75	
Scaling	Raw traces	Zero mean
	Encoded traces	Standard scaling

추이를 나타낸 것이 Fig. 12.이다. desync0 파형 분석 결과인 (a)와 인코딩된 desync50 파형 분석 결과인 (c)는 키가 구분되는 것을 볼 수 있는 반면 원본 desync50 파형의 분석 결과인 (b)는 키가 구분되지 않는 것을 볼 수 있다. 따라서 DDLA 실험 결과를 통해 제안한 방법이 기존의 통계 기반 비프로파일링 공격뿐만 아니라 딥러닝 기반의 공격에서도 좋은 성능을 발휘할 수 있음을 확인하였다.

V. 결 론

본 논문에서 제안한 스케일 선택 방법과 연속 웨이블릿 변환 기반 전력 파형 인코딩을 수행한다면 잡음 감쇄와 파형 정렬 효과로 인해 지터와 잡음이 많은 환경에서도 전력 분석 공격의 성능을 크게 향상됨을 실험을 통해 검증하였다. 또한, 전력 분석에 최적화된 스케일 선택 방법을 통해 CWT의 단점인 데이터 크기 및 분석 시간 증가 문제를 최소화하였으며 기존 비프로파일링 분석 기법인 CPA는 물론 딥러닝 기반 DDLA 분석에서도 뛰어난 성능을 보임을 확인했다. 따라서 제안한 전력 분석 공격 방법은 실험실 환경뿐만 아니라 잡음과 지터가 많은 실제 암호용 디바이스를 공격하는 환경에서도 사용될 수 있음을 주지할 필요가 있다.

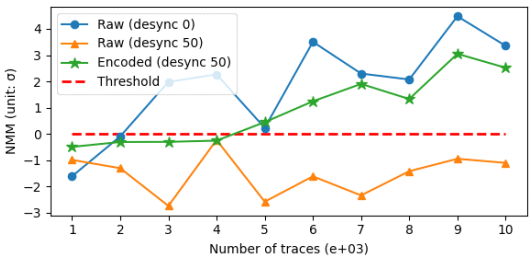


Fig. 11. The result of NMM analysis according to the number of traces.

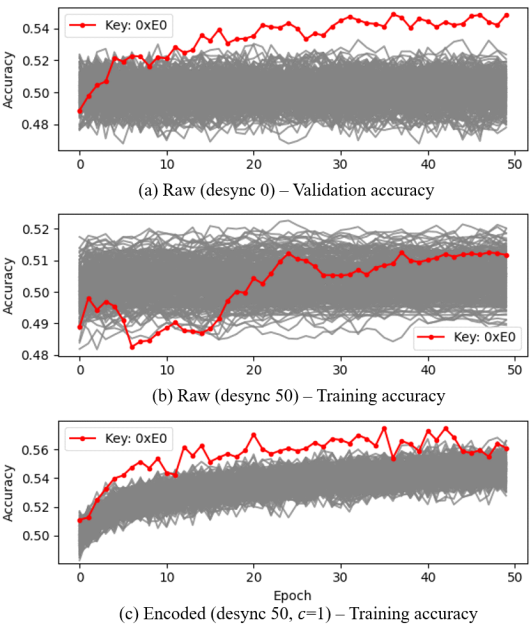


Fig. 12. Training metric values for each ASCAD trace set.

References

[1] P. Kocher, "Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems," CRYPTO'96, LNCS 1109, pp. 104-113, 1996.

[2] P. Kocher, J. Jaffe, and B. Jun, "Differential power analysis," CRYPTO'99, LNCS 1666, pp. 388-397, 1999.

[3] D. Kwon, S. Jin, H. Kim, and S. Hong, "Improving Non-Profiled Side-Channel Analysis Using Auto-Encoder Based Noise Reduction Preprocessing," Journal of the Korea Institute of Information Security & Cryptology, 29(3), pp. 491-501, May. 2019.

[4] T. Le, J. Clediere, C. Serviere, and J. Lacoume, "Noise Reduction in Side Channel Attack Using Fourth-Order Cumulant," IEEE Transactions on Information Forensics and Security, Vol. 2, No. 4, pp. 710-720, Nov. 2007.

[5] J. Ryoo, D. Han, S. Kim, H. Kim, T. Kim, and S. Lee, "A Proposal of Wavelet-based Differential Power Analysis Method," Journal of the Korea Institute of Information Security & Cryptology, 19(3), pp. 27-35, June. 2009.

[6] W. Kim, K. Song, Y. Lee, H.W. Kim,

and H.N. Kim, "Performance Improvement of Power Analysis Attacks based on Wavelet De-noising," The Journal of Korean Institute of Communications and Information Sciences, 35(9), pp. 1330-1342, Sep. 2010.

[7] N. Debande, Y. Souissi, M. Aabid, S. Guilley, and J. Danger, "Wavelet transform based pre-processing for side channel analysis," 45th Annual IEEE/ACM International Symposium on Microarchitecture Workshops, pp. 32-38, Dec. 2012.

[8] R. Gwinn, M. Matties, and A. Rubin, "Wavelet Selection and Employment for Side-Channel Disassembly," arXiv ePrint Archive, Available at https://arxiv.org/abs/2107.11870, 2021.

[9] D. Bae, and J. Ha, "Implementation of Instruction-Level Disassembler Based on Power Consumption Traces Using CNN," Journal of the Korea Institute of Information Security & Cryptology, 30(4), pp. 527-536, Aug. 2020.

[10] J. Park, X. Xu, Y. Jin and D. Forte, "Power-based side-channel instruction-level disassembler," Proceedings of the 55th Annual Design Automation Conference(DAC), pp. 1-6, June. 2018.

[11] E. Brier, C. Clavier, and F. Olivier, "Correlation power analysis with a leakage model," CHES'04, pp. 16-29, 2004.

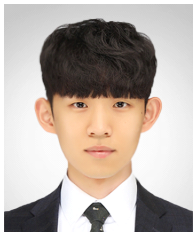
[12] B. Timon, "Non-profiled deep learning-based side-channel attacks with sensitivity analysis," IACR Transactions on Cryptographic Hardware and Embedded Systems, Vol. 2019, No. 2, pp. 107 - 131, Feb. 2019.

[13] ETRI, "SCARF Dataset," Available at https://trusthingz.org/index.php/scarf-data, 2019.

[14] E. Prouff, R. Strullu, R. Benadjila, E. Cagli, and C. Dumas, "Study of Deep Learning Techniques for Side-Channel Analysis and Introduction to ASCAD Database," IACR ePrint Archive, Available at https://eprint.iacr.org/2018/053, 2018.

[15] D. Bae, and J. Ha, "Performance Metric for Differential Deep Learning Analysis," Journal of Internet Services and Information Security (JISIS), Vol. 11, No. 2, pp. 22-33, May. 2021.

〈저 자 소 개〉



배 대 현 (Daehyeon Bae) 학생회원
2021년 2월: 호서대학교 컴퓨터정보공학부 학사
2021년 3월~현재: 호서대학교 정보보호학과 석사과정
〈관심분야〉 부채널 공격, 암호학, 정보보호

이 재 욱 (Jaewook Lee) 학생회원
2017년 3월~현재: 호서대학교 컴퓨터공학부 학부과정
〈관심분야〉 인공지능 보안, 자연어 처리, 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 컴퓨터공학부 교수
2013년 1월~현재: 한국정보보호학회 상임부회장
2009년 1월~현재: 한국산학기술학회 이사
〈관심분야〉 정보보호, 네트워크 보안, 부채널 공격