Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.30, NO.6, Dec. 2020 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2020.30.6.1291

효과적인 딥러닝 기반 비프로파일링 부채널 분석 모델 설계방안*

한 재 승,[1†] 심 보 연,[2] 임 한 섭,[1] 김 주 환,[3] 한 동 국[4‡]
[1,4]국민대학교 금융정보보안학과 (대학원생, 교수), [2,3]국민대학교 수학과 (연구교수, 학생)

Design of an Effective Deep Learning-Based Non-Profiling Side-Channel Analysis Model*

JaeSeung Han,[1†] Bo-Yeon Sim,[2] Han-Seop Lim,[1] Ju-Hwan Kim,[3] Dong-Guk Han[4‡]
[1,4]Department of Financial Information Security, Kookmin University
(Graduate student, Professor),
[2,3]Department of Mathematics, Kookmin University (Research professor, Undergraduate)

요 약

최근 딥러닝 기반 비프로파일링 부채널 분석이 제안됐다. 딥러닝 기반 비프로파일링 분석은 신경망 모델을 모든 추측키에 대해 학습시킨 뒤, 학습된 정도의 차이를 통해 올바른 비밀키를 찾아내는 기법이다. 이때, 신경망 학습모델 설계에 따라 비프로파일링 분석성능이 크게 달라지기 때문에 올바른 모델 설계의 기준이 필요하다. 본 논문은 학습모델 설계에 사용 가능한 2가지 loss 함수와 8가지 label 기법을 설명하고, 비프로파일링 분석과 소비전력모델 관점에서 각 label 기법의 분석성능을 예측했다. 해밍웨이트 소비전력모델을 가정했을 때의 비프로파일링 분석 특징을 고려해서 One-hot 인코딩을 적용하지 않은 HW(Hamming Weight) label과 CO(Correlation Optimization) loss를 적용한 학습모델이 가장 좋은 분석성능을 가질 것으로 예측했다. 그리고 AES-128 1라운드 Subbytes 연산 부분 데이터 집합 3가지에 대해 실제 분석을 수행했다. 제시한 각 label 기법과 loss 함수를 적용한 총 16가지 MLP(Multi-Layer Perceptron)기반 학습모델로 두 데이터 집합을 비프로파일링 분석하여 예측에 대해 검증했다.

ABSTRACT

Recently, a deep learning-based non-profiling side-channel analysis was proposed. The deep learning-based non-profiling analysis is a technique that trains a neural network model for all guessed keys and then finds the correct secret key through the difference in the training metrics. As the performance of non-profiling analysis varies greatly depending on the neural network training model design, a correct model design criterion is required. This paper describes the two types of loss functions and eight labeling methods used in the training model design. It predicts the analysis performance of each labeling method in terms of non-profiling analysis and power consumption model. Considering the characteristics of non-profiling analysis and the HW (Hamming Weight) power consumption model is assumed, we predict that the learning model applying the HW label without One-hot encoding and the Correlation Optimization (CO) loss will have the best analysis performance. And we performed actual analysis on three data sets that are Subbytes operation part of AES-128 1 round. We verified our prediction by non-profiling analyzing two data sets with a total 16 of MLP-based model, which we describe.

Keywords: Side-Channel Analysis, Deep Learning, Multi Layer Perceptron, AES

Received(10. 13. 2020), Modified(12. 03. 2020), Accepted(12. 03. 2020)

* 이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.2017-0-00520, SCR-Friendly 대칭키 암호 및 응용모드 개발)

† 주저자, jae1115@kookmin.ac.kr
‡ 교신저자, christa@kookmin.ac.kr(Corresponding author)

I. 서 론

부채널 분석은 알고리즘이 실제 장비에서 동작할 때 발생하는 부채널 정보(전기, 빛, 전자파 등)를 이용해 비밀키 등의 비밀 정보를 획득하는 기법이다. 부채널 분석 기법 중, 전력 분석 공격은 크게 프로파일링 분석, 비프로파일링 분석으로 구분된다. 프로파일링 분석은 사전에 공격 대상 장치와 동일한 환경의 장치로부터 전력 파형 프로파일을 생성하고, 공격 대상 장치로부터 획득한 전력 파형과 프로파일 간의 매칭을 통해 비밀키를 분석하는 기법이다.

대표적인 프로파일링 분석으로 TA(Template Attack)가 있다[1]. 비프로파일링 분석은 공격 대상 장치에 고정된 비밀키와 무작위 평문에 대한 암호화를 여러 번 수행시킨 뒤, 그때 발생하는 다수의 전력 파형을 수집하고 통계 분석을 통해 비밀키를 분석하는 기법이다. 대표적인 비프로파일링 분석으로 DPA(Differential Power Analysis), CPA(Correlation Power Analysis)가 있다[2].

딥러닝은 기계학습의 한 종류로, 어떤 집합 X, Y에 대한 함수 $F: X \rightarrow Y$가 있고, 실제 함수 F에 대해 알 수 없을 때, 훈련 데이터 집합을 이용한 깊은 신경망 학습을 통해 F와 유사한 함수 F'을 만드는 기법이다. 딥러닝은 이미지 인식, 음성 인식, 자연어 처리 등 넓은 분야에서 활용되고 있고 최근에는 부채널 분석 기반 암호분석에서도 딥러닝 기법을 적용하는 시도가 이루어지고 있다[3, 4].

부채널 분석에서도 딥러닝 기반 프로파일링, 비프로파일링 분석이 연구되고 있다. 주요 개념은 MLP(Multi-Layer Perceptron), CNN(Convolutional Neural Network) 등으로 설계한 신경망의 입력을 파형으로 주고 대상 연산 중간 결괏값과 관련된 값을 출력하도록 신경망을 학습시키고 학습된 신경망을 통해 공격대상의 비밀키를 판단하는 것이다[5, 6].

본 논문은 비프로파일링 분석과 소비전력모델의 특성을 고려한 신경망 학습의 label 기법을 세분화하여 제안하고, 분석성능에 대한 예측을 제시한다. 그리고 8개의 label 기법과 2개의 loss 함수를 적용한 16가지 학습모델로 대응기법이 적용되지 않은 3개의 데이터 집합에 대한 비프로파일링 분석을 수행하여 제시한 모델들의 분석성능을 보인다.

II. 관련 연구

2.1 MLP

MLP는 여러 개의 퍼셉트론을 여러 층으로 구성하여 입력에 대한 출력을 계산하는 신경망 구조이다. MLP는 Fig. 1.과 같이 크게 입력층, 은닉층, 출력층으로 이루어져 있고 현재 층의 각 퍼셉트론이 다음 층의 모든 퍼셉트론과 연결되어있는 완전연결구조이다. 충분한 양의 데이터와 각 데이터에 대한 정답인 label 값들이 주어진 경우, MLP 구조를 이 데이터 집합에 대해 반복 학습할 수 있고 이것을 지도학습이라 한다. 지도학습 과정은 신경망의 실제 출력값과 기대 출력값인 label 값으로 loss 함수를 통해 벌점을 계산하고, 벌점을 통해 신경망의 가중치를 갱신하는 방식으로 이루어진다. 잘 학습된 MLP는 학습에 쓰이지 않은 같은 속성의 데이터를 입력으로 받아 출력을 계산했을 때, 기대 출력값인 label 값이 계산된다. MLP를 이용한 딥러닝 기법을 통해 데이터의 분류, 회귀문제 등을 해결할 수 있다[7].

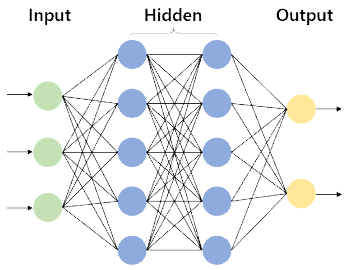


Fig. 1. Example of MLP

2.2 딥러닝 기반 비프로파일링 분석

비프로파일링 부채널 분석은 동일 비밀키를 사용한 암호 동작 시의 다수의 파형(전력, 전자파 등), 평문 정보를 이용해 비밀키를 분석하는 기법이다. 딥러닝 기반 비프로파일링 부채널 분석은 2019년도에 Benjamin Timon에 의해 제안됐다[8]. 딥러닝 기반 비프로파일링 분석 과정은 다음과 같다. 먼저, 각 추측키에 대해 입력을 파형, 출력을 대상 연산 중간값에 대한 label 값으로 설정하여 신경망을 학습시킨다. 그러면, 옳은키로 계산한 label은 파형과 관계된 값이므로 신경망이 잘 학습되고, 틀린키들은 파형과 관계없는 label 값을 준 것과 같은 효과가 생겨 학습이

잘되지 않는다. 그러므로 신경망을 가장 잘 학습시킨 추측키를 옳은키로 판단하여 비밀키를 분석한다. 딥러닝 기반 비프로파일링 분석기법 또한 기존 비프로파일링 분석과 같이 틀린키들의 중간값이 옳은키의 중간값과 상관도가 낮기를 기대하므로 비선형연산의 출력을 중간값으로 설정하는 것이 일반적이다.

Timon의 논문에서는 딥러닝 기반 비프로파일링 분석에서 중간값을 직접 label로 사용하면 분석이 불가능하다고 주장했고, 실험에서 label 값으로 MSB (Most Significant Bit)를 선택했을 때 가장 높은 분석성능을 가졌다고 언급했다.

III. 비프로파일링 분석을 위한 효과적인 모델 설계

본 논문은 각 추측키의 학습능력을 일반화 성능으로 측정하기 위해서 분석에 대한 검증 집합 손실(validation loss)을 학습능력의 지표로 사용했고 낮을수록 학습능력이 좋다고 판단했다[9]. 본 논문의 비프로파일링 분석 과정은 Fig. 2.와 같다.

Table 1.은 분석 모델 설계에서 일반적으로 채택할 수 있는 몇 가지 label 기법과 loss 함수에 대해 정리한 것이다. 본 장에서는 전력 소모가 해밍웨이트(Hamming Weight, HW) 소비전력모델을 따른다고 가정하고, 비프로파일링 분석관점을 고려할 때, label 기법 중 효과적인 기법이 무엇인지 설명한다. 또한, 학습모델 설계에 있어 선택 가능한 loss 함수 중 2가지를 설명한다.

Table 1.에서 label인 HW, MSB, LSB(Least Significant Bit), IV(Intermediate Value)는 각각 대상 중간값의 해밍웨이트, 최상위비트, 최하위

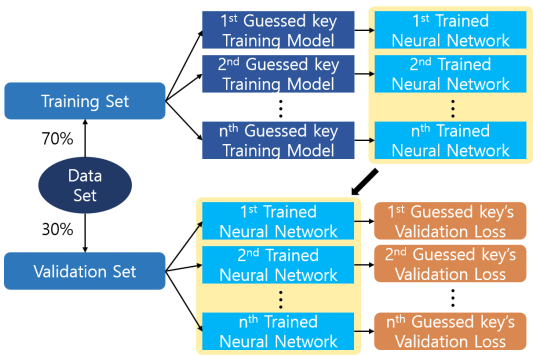


Fig. 2. Deep learning based non-profiling analysis process

Table 1. Methods of label, loss

	Method
Label	HW, LSB, MSB, IV, O-HW, O-LSB, O-MSB, O-IV
Loss	MSE, CO

비트, 마지막으로 중간값 자체를 label로 사용한다는 의미이다. O-XX는 XX label을 One-hot 인코딩하여 label로 사용한다는 의미이다. 예를 들어, One-hot 인코딩된 MSB label 이면 O-MSB로 표현하며, label은 (1,0), (0,1) 두 클래스로 나뉜다. HW, MSB, LSB, IV 모두 유한한 종류의 데이터로 구분될 수 있으므로, 대부분의 딥러닝 기반 부채널 분석에서는 label을 One-hot 인코딩하여 클래스 분류형태로 분석 모델을 학습시킨다[5, 6, 8]. loss 함수는 MSE(Mean Square Error), CO(Correlation Optimization) 두 가지를 다룬다.

3.1 Label

비프로파일링 분석 모델 설계는 다음을 만족해야 한다.

- [기준 1] : 분석 모델이 옳은키에 대해서 높은 학습능력을 가진다.
- [기준 2] : 분석 모델이 옳은키 외 다른 추측키에 대해서 낮은 학습능력을 가진다.

먼저, [기준 1] 관점에서 분석 대상의 소비전력모델이 해밍웨이트 모델을 따른다고 가정하면 소비전력모델과 정확히 일치하는 HW를 label 값으로 사용하는 것이 가장 타당하다. 그리고 HW와 일정 부분 선형관계를 갖는 MSB, LSB도 label 값으로 활용될 수 있을 것이라 보인다. IV 값 또한 HW 값과 완전한 선형관계를 갖는 것은 아니지만, IV 값이 높을수록 HW 값도 높게되는 부분적인 선형관계를 가지므로 IV label로도 [기준 1] 관점에서 학습능력을 가질 것으로 보인다.

[기준 2] 관점에서는 옳은키 중간값의 label 값들과 틀린키 중간값의 label 값들이 서로 상관도가 낮을수록 틀린키에 대한 학습능력이 낮아짐을 예측할 수 있다. 그리고 Subbytes 연산과 같은 비선형 연산의 출력을 중간값으로 사용할 때, 옳은키 중간값의 각 HW, MSB, LSB, IV label 값들이 틀린키의 각 label 값들과 선형적으로 상관도가 낮은 것은 자

명하다. 그러므로 4가지 label 모두 〔기준 2〕에 부합한다.

하지만 One-hot 인코딩을 적용하여 label 값을 설정하면 차이가 생긴다. One-hot 인코딩은 클래스들의 선형관계를 없애고 모든 클래스들을 서로 독립적인 관계로 만드는 특징이 있다. 즉, 클래스 간의 선형관계가 존재할 때 One-hot 인코딩을 적용하는 것은 적절하지 않다〔7〕. 2개의 클래스로 나뉘는 MSB, LSB의 경우엔 이런 특징에 의한 영향이 미미하나, 3개 이상의 클래스로 나뉘는 HW, IV는 영향이 크다.

O-HW의 경우, label 값 클래스들의 선형관계가 없어지므로 학습에 불리하다. 간단한 예시로, O-HW는 올바른 HW 값이 4일 때, 신경망이 HW 값을 3으로 잘못 추측한 것과 0으로 잘못 추측한 것에 신경망에 동등한 벌점을 준다. 하지만 해밍웨이트 소비전력모델 관점에서 HW 값은 3보다 0이 정답인 4와 더 멀기 때문에 0으로 잘못 추측했을 때 신경망에 더 큰 벌점을 부여해야 한다. 그러므로 HW의 경우 One-hot 인코딩을 하지 않는 것이 학습관점에서 타당하다.

또한, O-IV의 경우에는 비프로파일링 분석 불가능함이 기존 연구에 의해 밝혀져 있다〔8〕. IV로 각 학습파형의 label을 지정할 경우, 추측키와 상관없이 항상 동일한 평문에 대해서만 같은 클래스로 분류되게 된다. 그리고 One-hot 인코딩은 모든 클래스 간의 관계를 독립적으로 만들기 때문에 O-IV는 추측키와 관계없이 label 값을 부여하는 것이 된다. 즉, O-IV는 〔기준 2〕를 만족하지 못하므로 비프로파일링 분석이 되지 않는다.

즉, 소비전력모델과 비프로파일링 분석 모델 설계 관점에서 HW label이 가장 높은 분석성능을 가질 것으로 예측된다. 이는 2.2장에서 언급한 Timon의 논문과 상반되며 Timon의 논문에서는 One-hot 인코딩을 적용한 label 기법만을 고려하였다.

3.2 Loss

MSE는 흔히 사용되는 loss 함수이며 수식은 다음과 같다. x_i,y_i는 각각 예측치와 label 값을 의미한다.

$$MSE=\sum(x_i-y_i)^2 \tag{1}$$

결정 계수(Coefficient of Determination)는 MSE의 표준화된 버전으로 수식 (2)와 같다. D는 결정 계수, R은 피어슨 상관계수이다.

$$D=R^2=1-\frac{\sum(x_i-y_i)^2}{\sum(x_i-\overline{x})^2}=1-\frac{MSE}{Var(x)} \tag{2}$$

$\overline{x}$는 x의 평균, $Var(x)$는 x의 분산을 의미한다. MSE의 경우 종속 변수의 분산이 커지게 되면 모델의 적합도를 판단하기 어렵다. 그러나 수식 (2)와 같이, 결정 계수는 종속 변수의 분산에 의존하지 않기 때문에 0과 1 사이의 값으로 나타낼 수 있다. 즉, 수식 (3)과 같이 피어슨 상관계수를 기반으로 손실 값을 계산하는 CO는 MSE에 비해 비교적 분산에 영향 없이 loss를 계산하는 특징이 있다〔10〕.

$$CO=1-R \tag{3}$$

따라서, MSE보다 CO를 loss 함수로 사용하는 것이 모델 예측의 적합도 관점에서 좋을 것으로 예상된다.

IV. 실험 결과

3장에서 제시한 각 label 기법과 loss 함수의 분석성능을 확인하기 위해 각 기법에 따른 MLP기반 비프로파일링 분석을 대응기법이 적용되지 않은 3가지 데이터 집합에 대해 수행했다. 첫 번째 데이터 집합은 ChipWhisperer-Lite 보드〔11〕를 이용하여 파형을 수집하였고, 두 번째 데이터 집합으로는 ETRI에서 제공하는 공개 데이터 집합을 사용했다〔12〕. 세 번째는 〔8, 13〕에서 실험에 사용한 ASCAD 공개 데이터 집합이다〔14〕. ChipWhisperer-Lite와 ETRI 데이터 집합은 대응기법이 적용되어 있지 않은 집합이며, ASCAD는 첫 번째와 두 번째 바이트를 제외하고는 모두 1차 마스킹이 적용된 집합이다. 본 논문에서는 모두 첫 번째 바이트 분석 결과를 보인다.

Table 2. Features of data sets

	ChipWhisperer data set	ETRI data set	ASCAD data set
Target chip	XMEGA 128	ATMEGA 128	ATMEGA 8515
Target cipher	AES-128		
Target operation	Subbytes		

Table 3. Network structures for analysis

Layer	node (in, out)	kernel initializer
Input	(x,x)	
Batch Normalization	(x,x)	
Dense	$(x,32)$	he_uniform
Batch Normalization	$(32,32)$	
softplus	$(32,32)$	he_uniform
Dense	$(32,y)$	
sigmoid	(y,y)	

- Input Normalization: all values are within the range of -1 and 1
- Optimizer: Nadam (lr=0.0001, epsilon=1e-08)
- Batch size: 32

ETRI 파형이 나머지 두 데이터 집합과 비교하여 노이즈가 크다는 특징이 있다. 각 데이터 집합의 특징은 Table 2.와 같다. 본 논문에서 분석에 사용한 신경망 모델은 Table 3.과 같다.

Fig. 3.와 Fig. 4.는 loss함수를 CO으로 선택하고, 각 IV, O-IV를 label로 설정한 신경망의 ChipWhisperer 파형 2000개에 대한 학습결과이다. Epoch 수에 따른 validation loss를 256개 추측키에 대해 나타낸 것으로, 붉은선으로 그려진 옳은키의 loss가 다른 추측키들의 loss와 많은 차이가 날수록 분석이 잘된 것이라 할 수 있다. 3.1장에서 언급한 것과 같이 O-IV label로 학습한 신경망은 분석이 되지 않는 것을 볼 수 있다.

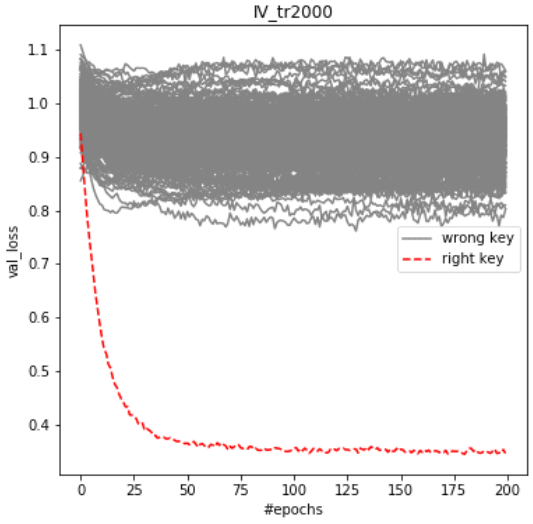


Fig. 3. IV labeling model analysis result on ChipWhisperer data set

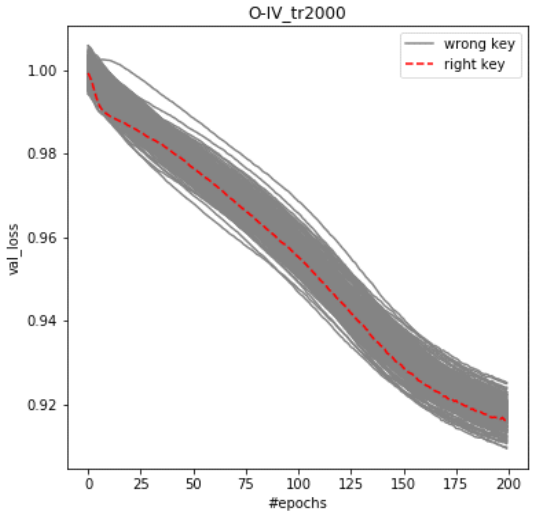


Fig. 4. O-IV labeling model analysis result on ChipWhisperer data set

Table 4.는 각 HW, MSB, LSB, IV, O-HW, O-MSB, O-LSB, O-IV label의 분석성능을 수치로 나타낸 것이다. ChipWhisperer 파형 2000개로 200 epoch 만큼 학습시켰다. Table 4.의 값들은 마지막 epoch를 기준으로 옳은키를 제외한 추측키들 중 가장 낮은 loss를 옳은키의 loss로 나눈 값이고 이를 ratio로 표기한다. 즉, ratio가 1.0 이상일 때

Table 4. Analysis performance for each labeling method (2000 traces, 200 epochs) on ChipWhisperer data set

CO	One-hot encoding X	One-hot encoding O
HW	7.902	1.211
MSB	1.858	1.926
LSB	1.139	1.085
IV	2.310	0.992
MSE	One-hot encoding X	One-hot encoding O
HW	5.597	1.074
MSB	1.458	1.442
LSB	1.033	0.967
IV	1.866	0.992

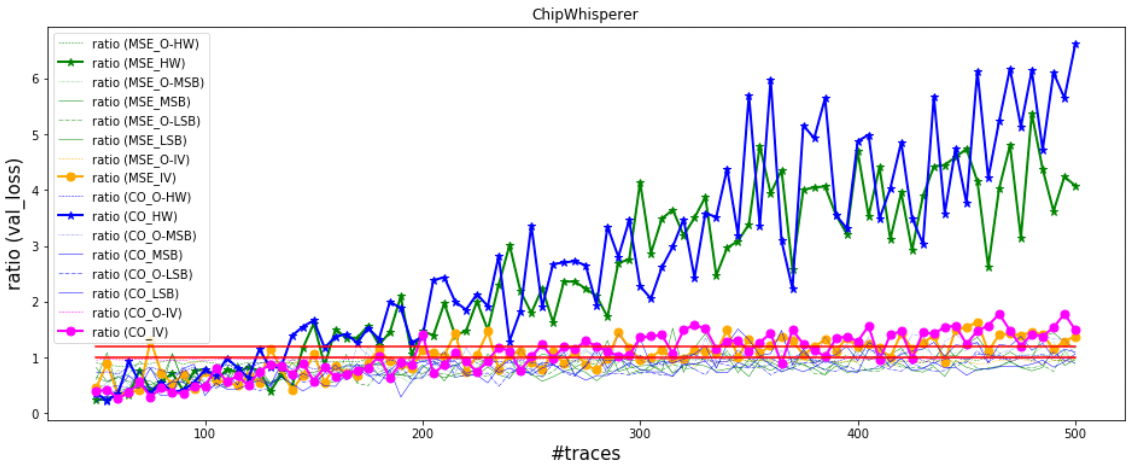


Fig. 5. Results of ChipWhisperer data set for each labeling, loss method

분석이 성공하며 높을수록 분석성능이 높다. 3.1장에서 예상한 것과 같이 MSB, LSB의 경우 One-hot 인코딩 여부에 큰 차이 없는 분석성능을 가지는 것을 볼 수 있고, HW는 One-hot 인코딩을 하지 않은 것이 더 좋은 분석성능을 가지는 걸 알 수 있다.

Fig. 5.는 ChipWhisperer 파형에 대한 각 label 기법, loss 함수를 적용했을 때, 분석에 사용한 파형 수에 따른 ratio를 나타낸 그래프다. 파형 수는 5개 단위로 수행했고 모두 200 epoch 만큼 학습했다. label 중에서는 One-hot 인코딩을 적용하지 않은 HW label이 가장 적은 파형 수로 분석됐고, MSE loss보다 CO loss가 더 좋은 분석성능을 가짐을 볼 수 있다. Table 5.는 Fig. 5.의 결과를 표로 나타낸 것으로, 500개 이하의 파형으로 분석됐다고 판단되는 4가지 모델들이 일정 ratio를 초과하는 최소 분석 파형 수를 나타냈다. Table 6.은 Table 4.와 같은 신경망 모델을 통해 ETRI 파형 2000개를 200 epoch 만큼 학습시킨 결과이다. ChipWhisperer 파형보다 노이즈가 많은 파형이므로 전체적으로 분석에 많은 파형 수가 필요한 것을 알 수 있다.

Table 7. Minimum number of traces that analyze successfully for each training model on ETRI data set

Model/Ratio	Ratio 〉 1.0	Ratio 〉 1.2
MSE_HW	495	700 ↑
CO_HW	370	530
MSE_IV	700 ↑	700 ↑
CO_IV	700 ↑	700 ↑

하지만 각 label 기법, loss 함수에 따른 상대적인 분석성능의 차이는 ChipWhisperer 파형의 결과와 유사하다. Fig. 6.은 ETRI 파형에 대한 각 label 기법, loss 함수를 적용했을 때, 분석에 사용한 파형 수에 따른 ratio를 나타낸 그래프다. Table 7.은 Table 5.와 같이 ETRI 데이터 집합에 대한 HW, I

Table 5. Minimum number of traces that analyze successfully for each training model on ChipWhisperer data set

Model/Ratio	Ratio 〉 1.0	Ratio 〉 1.2
MSE_HW	160	200
CO_HW	140	160
MSE_IV	445	495
CO_IV	430	430

Table 6. Analysis performance for each labeling method (2000 traces, 200 epochs) on ETRI data set

CO	One-hot encoding X	One-hot encoding O
HW	2.421	0.961
MSB	1.065	0.936
LSB	1.719	1.651
IV	1.126	0.993
MSE	One-hot encoding X	One-hot encoding O
HW	2.051	0.922
MSB	1.036	0.988
LSB	1.508	1.561
IV	1.027	0.799

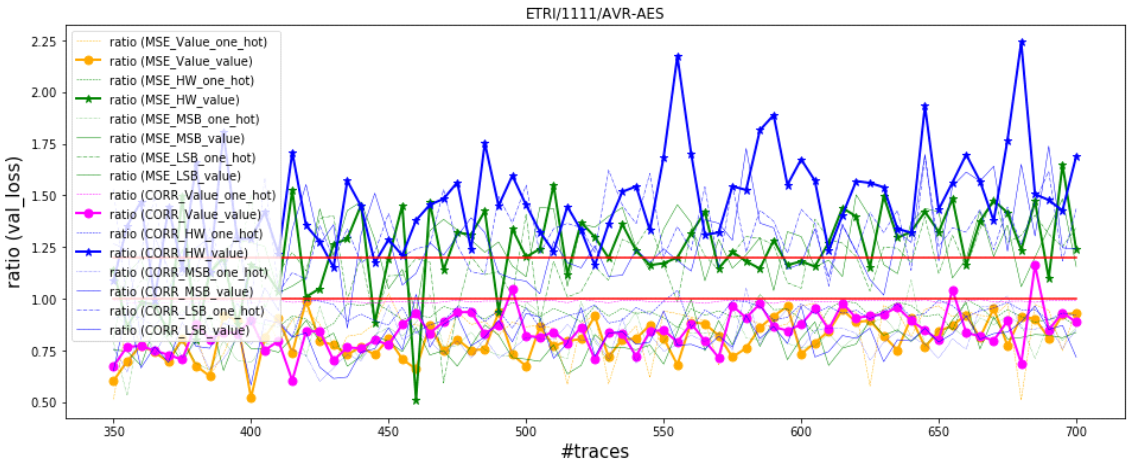


Fig. 6. Results of ETRI data set for each labeling method, loss function

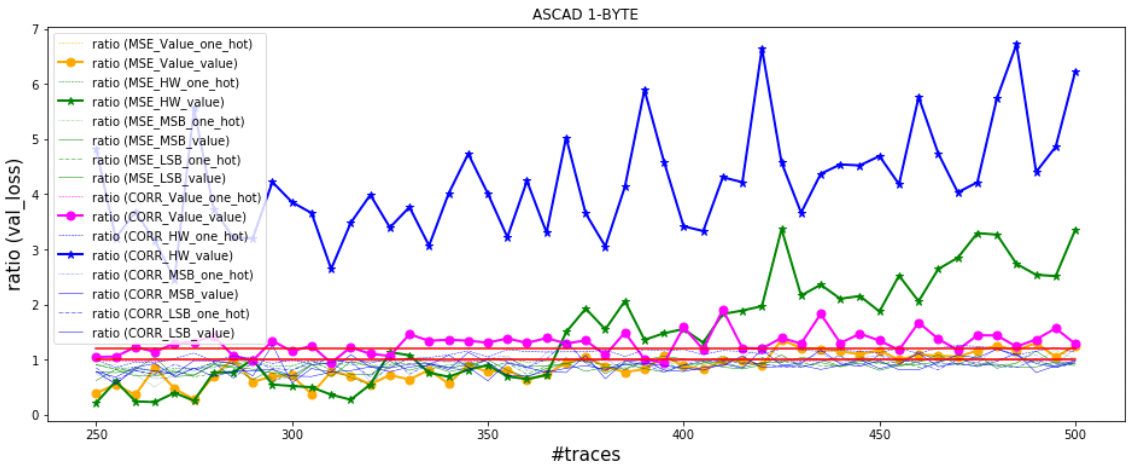


Fig. 7. Results of ASCAD data set for each labeling method, loss function

V의 최소 분석 파형 수를 정리한 것이다.

Fig. 7.은 ASCAD 파형에 대한 각 label 기법, loss 함수를 적용했을 때, 분석에 사용한 파형 수에 따른 ratio를 나타낸 그래프다. Table 8.도 Table 4, 6.과 같은 표로, ASCAD 데이터 집합에 대해 50 epoch만큼 학습한 결과이다. 앞선 데이터 집합의 분석 결과와 ratio 수치에 대한 차이는 존재하지만, 각 label 기법, loss 함수에 따른 상대적인 분석성능의 차이는 Table 4, 6.과 유사하다. Table 9.는 Table 5, 7.과 같이 ASCAD 데이터 집합에 대한 HW, IV의 최소 분석 파형 수를 정리한 것이다.

결론적으로 3.1장의 예측과 같이 3개의 데이터 집합 모두 One-hot 인코딩을 적용하지 않은 HW lab

Table 8. Analysis performance for each labeling method (5000 traces, 50 epochs) on ASCAD data set

CO	One-hot encoding X	One-hot encoding O
HW	10.670	1.280
MSB	1.325	1.359
LSB	1.338	1.288
IV	2.340	1.000
MSE	One-hot encoding X	One-hot encoding O
HW	6.636	1.098
MSB	1.135	1.176
LSB	1.096	1.155
IV	1.620	0.995

Table 9. Minimum number of traces that analyze successfully for each training model on ASCAD data set

Model/Ratio	Ratio 〉 1.0	Ratio 〉 1.2
MSE_HW	370	370
CO_HW	90	100
MSE_IV	460	500 ↑
CO_IV	400	475

el이 다른 label 기법과 비교하여 월등히 높은 분석 성능을 보였다. 또한, 같은 label 기법에 대해서 MS E loss보다 CO loss가 분석성능이 높았다.

V. 결 론

본 논문은 효과적인 딥러닝 기반 비프로파일링 분석 모델 설계방안을 제시한다. 학습모델을 설계할 때 선택 가능한 몇 label 기법, loss 함수를 제시하고 해밍웨이트 소비전력모델, 비프로파일링 분석관점을 고려하여 분석성능이 높은 label 기법을 예측했다. 그리고 각 학습모델 설계에 따른 실제 분석성능을 비교하기 위해서 MLP 기반의 학습모델로 대응기법이 적용되지 않은 3개의 데이터 집합에 대해 실험을 수행했다. 실험 결과, 예측과 같이 3개 데이터 집합 모두 One-hot 인코딩을 적용하지 않은 HW label로 학습한 모델이 가장 분석성능이 높은 것을 검증했고, loss 함수는 MSE loss보다 CO loss가 분석성능이 더 높은 것을 검증했다. 향후, 부울린 마스킹과 과 같은 대응기법이 적용된 데이터 집합에 대한 분석이 필요하다.

References

[1] S. Chari, J. Rao, and P. Rohatgi, "Template attacks," Cryptographic Hardware and Embedded Systems, CHES 2002, LNCS 2523, pp. 13-28, 2003.

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

[3] S. Jin, S. Kim, H. Kim, and S. Hong, "Recent advances in deep learning based side channel analysis," ETRI Journal, 42(2), pp. 292-304, Feb. 2020.

[4] B. Hettwer, S. Gehrer, and T. Güneysu, "Applications of machine learning techniques in side-channel attacks: a survey," Journal of Cryptographic Engineering, vol. 10, no. 2, pp. 135-162, Apr. 2019.

[5] Z. Martinasek, P. Dzurenda, and L. Malina, "Profiling power analysis attack based on mlp in DPA contest v4.2," Telecommunications and Signal Processing, pp. 223-226, Jun. 2016.

[6] E. Cagli, C. Dumas, and E. Prouff, "Convolutional neural networks with data augmentation against jitter-based countermeasures," Cryptographic Hardware and Embedded Systems, CHES 2017, pp. 45-68, Sep. 2017.

[7] I. Oh, Machine Learning, 1st Ed., Hanbit Academy, Dec. 2017.

[8] B. Timon, "Non-profiled deep learning-based side-channel attacks with sensitivity analysis," IACR Transactions on Cryptographic Hardware and Embedded Systems, TCHES, pp. 107-131, Feb. 2019.

[9] J. Park, D. Han, D. Jap, S. Bhasin, and Y. Won, "Non-profiled side channel attack based on deep learning using picture trace," ePrint, Oct. 2019.

[10] P. Robyns, P. Quax, and W. Lamotte, "Improving CEMA using correlation optimization," IACR Transactions on Cryptographic Hardware and Embedded Systems, TCHES, Nov. 2018.

[11] Inc,N.T.ChipWhisperer-Lite, https://wiki.newae.com/CW1173_ChipWhisperer-Lite.

[12] Inc, ETRI, https://trusthingz.org/index.php/scarf-data.

[13] R. Benadjila, E. Prouff, R. Strullu, E. Cagli, and C. Dumas, "Deep learning

for side-channel analysis and introduction to ASCAD database", Journal of Cryptographic Engineering, vol. 10, no. 2, pp. 163-188, Nov. 2019.

[14] ANSSI: Ascad database, https://github.com/ANSSI-FR/ASCAD.

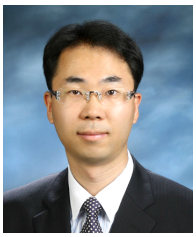
〈저자소개〉

한 재 승 (JaeSeung Han) 학생회원
2020년 2월: 국민대학교 정보보안암호수학과 학사
2020년 3월~현재: 국민대학교 금융정보보안학과 석사과정
〈관심분야〉 부채널 분석 및 대응법 설계, 딥러닝, 대칭키 암호, lattice 기반 암호

심 보 연 (Bo-Yeon Sim) 일반회원
2013년 2월: 국민대학교 수학과 학사
2015년 2월: 국민대학교 금융정보보안학과 석사
2020년 2월: 국민대학교 수학과 박사
2020년 3월~현재 : 국민대학교 산학협력단 연구교수
〈관심분야〉 공개키 암호 시스템, 부채널 분석 및 대응기법 설계, 경량 저전력 정보보호 기술

임 한 섭 (Han-Seop Lim) 학생회원
2019년 2월: 국민대학교 정보보안암호수학과 학사
2020년 3월~현재: 국민대학교 금융정보보안학과 석사과정
〈관심분야〉 부채널 분석 및 대응법 설계, 오류 주입 공격, 스마트 카드 보안

김 주 환 (Ju-Hwan Kim) 학생회원
2016년 3월~현재: 국민대학교 수학과 학사과정
〈관심분야〉 부채널 분석 및 대응법 설계, 딥러닝, 오류 주입 공격



한 동 국 (Dong-Guk Han) 종신회원
1999년 2월: 고려대학교 수학과 학사
2002년 2월: 고려대학교 수학과 이학석사
2005년 2월: 고려대학교 정보보호대학원 공학박사
2004년 4월~2005년 4월: 일본 Kyushu Univ., 방문연구원
2005년 4월~2006년 4월: 일본 Future Univ.-Hakodate, Post.Doc.
2006년 6월~2009년 2월: 한국전자통신연구원 정보보호연구단 선임연구원
2009년 3월~현재: 국민대학교 정보보안암호수학과 정교수
〈관심분야〉 공개키 암호시스템 안전성 분석 및 고속 구현, 부채널 분석 및 대응법 설계, IoT 정보보호 기술