# Side-Channel Auto-Encoder

논프로파일링 환경에서 오토인코더를 잡음 제거하는
전처리기로 사용한 기법

https://youtu.be/Z4bn3foLSjY

한성대학교 HANSUNG UNIVERSITY

CryptoCraft LAB
https://crypto.modoo.at

# 논프로파일링

- **프로파일링 장비없이 타겟 장비로부터만 전력 파형을 수집 가능한 환경에서의 부채널 분석 기법**

- 공격 대상에 고정된 비밀키와 무작위 평문에 대한 암호화를 여러번 수행시킨 뒤, 그때 발생하는 다수의 전력 파형을 수집하고 통계 분석을 통해 비밀키를 분석하는 기법

- 대표적으로 **차분 전력 분석**, **상관 전력 분석**이 있음

# 논프로파일링

- 차분 전력 분석 공격

  ✓ 단순전력분석 보다 더 강력한 공격 방법

  * 단순전력분석 : 암호 연산을 여러번 구동 시켜 수집한 전력 소비량의 시간에 따른 변화를 시각적으로 해석하는 공격 방법

  ✓ 암호 연산을 여러번 구동시킨 후, 얻을 수 있는 **전력 소비량 정보**를 다양한 신호 처리 방법을 사용하여 **분석해** 암호 기기 내부에 저장된 비밀 정보를 얻어냄.

- 상관 전력 분석 공격

  ✓ 공격자가 한꺼번에 선택한 암호문이 주어진다는 가정

  ✓ 평문을 선택하면 대응하는 암호문을 얻을 수 있는 상황에서 공격하는 것

# 논프로파일링

- 전력소비 모델 수식

$$P = \delta + HW(data) + Noise$$

P : 암호 연산을 수행하는 장비가 발생하는 소비 전력

data : 연산하는 중간값

δ : 고정된 상수 오프셋

HW : 해밍웨이트 함수

Noise : N(0 . δ^2)의 정규 분포를 따르는 랜덤 노이즈

- X와 Y의 상관계수 수식

$$\rho(X, Y) = \frac{Cov[X, Y]}{\sqrt{Var[X]\ Var[Y]}} = \frac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum(X_i - \overline{X})^2 \sum(Y_i - \overline{Y})^2}}$$

X : 장비가 비밀정보와 관련된 연산을 하고 있을 때, 소모하는 전력을 측정한 값

Y : 해당 연산의 결과인 중간값으로 추측한 값



권동근, 진성현, 김희석, 홍석희 (2019). 비프로파일링 기반 전력 분석의 성능 향상을 위한 오토인코더 기반 잡음 제거 기술. 정보보호학회논문지, 29(3), 491-501

# 딥러닝 기반 논프로파일링

- 2019년 Timon에 제안 (Differential deep learning Analysis_DDLA)
- 분석과정
  - 각 추측키를 다음과 같이 설정 후 신경망 학습
    - ✓ 입력: 파형으로 설정
    - ✓ 출력: 대상 연산 중간값에 대한 label 값으로 설정
  - 옳은 키로 계산한 label의 경우.
    - ✓ 파형과 관계된 값 -> 신경망 학습이 잘 됨
  - 틀린 키로 계산한 label의 경우,
    - ✓ 파형과 관계되지 않은 값을 준 것과 같은 효과->신경망 학습이 잘 되지 않음


한재승, 심보연, 임한섭, 김주환, 한동국 (2020). 효과적인 딥러닝 기반 비프로파일링 부채널 분석 모델 설계방안. 정보보호학회논문지, 30(6), 1291-1300.

# Auto-Encoder

- 딥러닝 알고리즘 중 대표적인 비지도 학습
- 뉴런 네트워크의 출력을 입력과 유사하도록 학습
- 데이터를 압축하거나 뉴런 네트워크의 사전 학습 등의 목적을 위해 사용


권동근, 진성현, 김희석, 홍석희 (2019). 비프로파일링 기반 전력 분석의 성능 향상을 위한 오토인코더 기반 잡음 제거 기술. 정보보호학회논문지, 29(3), 491-501

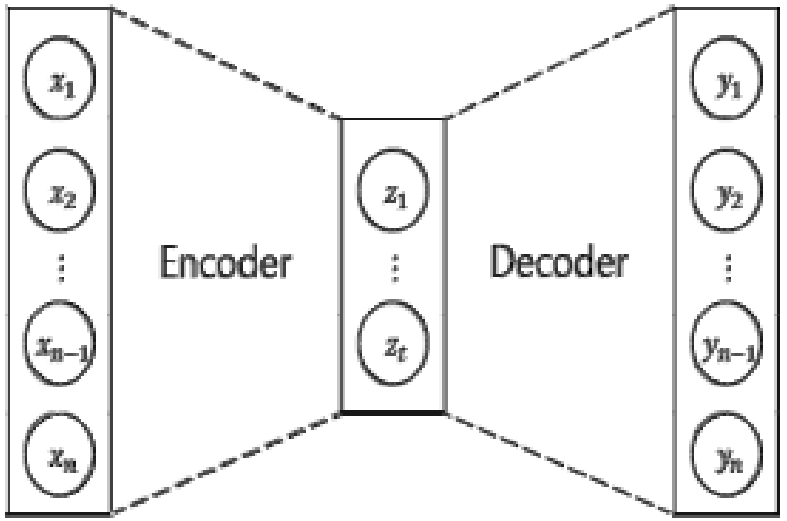
# Auto-Encoder

- 데이터의 차원을 압축하는 방법으로 사용 가능
- 차원 압축 목적으로 사용되는 오토인코더

  → 인코더 부분과 디코더 부분으로 구성

  ✓ 입력 데이터의 차원을 압축하는 인코더

  ✓ 인코더를 통해 압축된 데이터 원본의 입력 데이터로 다시 재구성하는 디코더





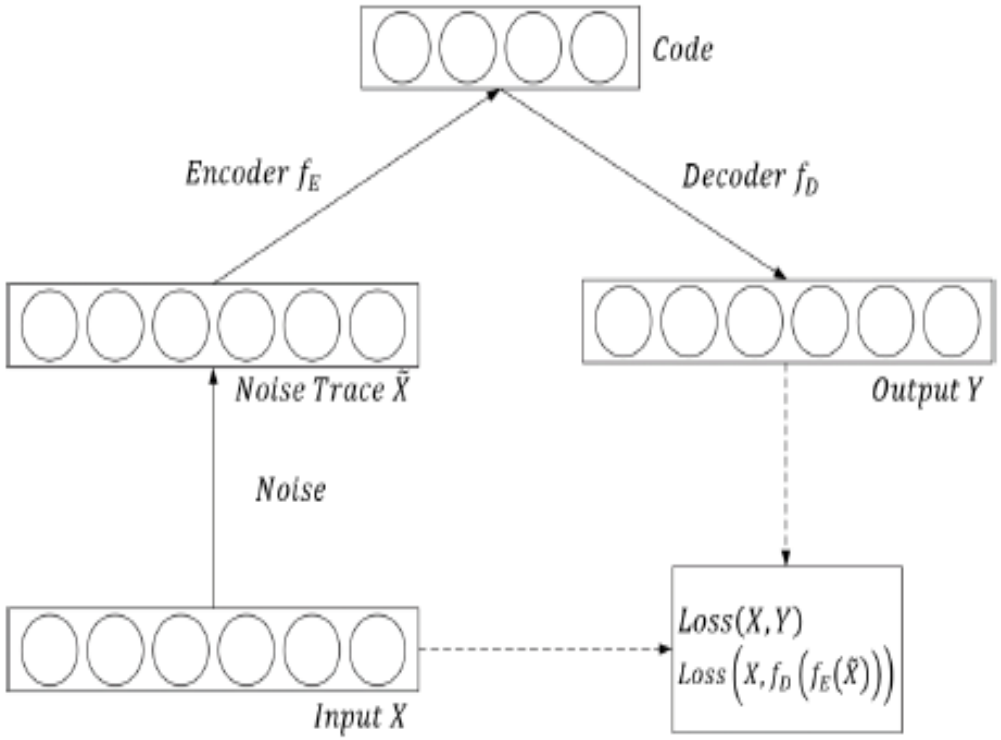

권동근, 진성현, 김희석, 홍석희 (2019). 비프로파일링 기반 전력 분석의 성능 향상을 위한 오토인코더 기반 잡음 제거 기술. 정보보호학회논문지, 29(3), 491-501

# Denoising Auto-Encoder

- 오토인코도와 네트워크의 구조 동일

  → 학습에 사용되는 데이터에서 차이 존재

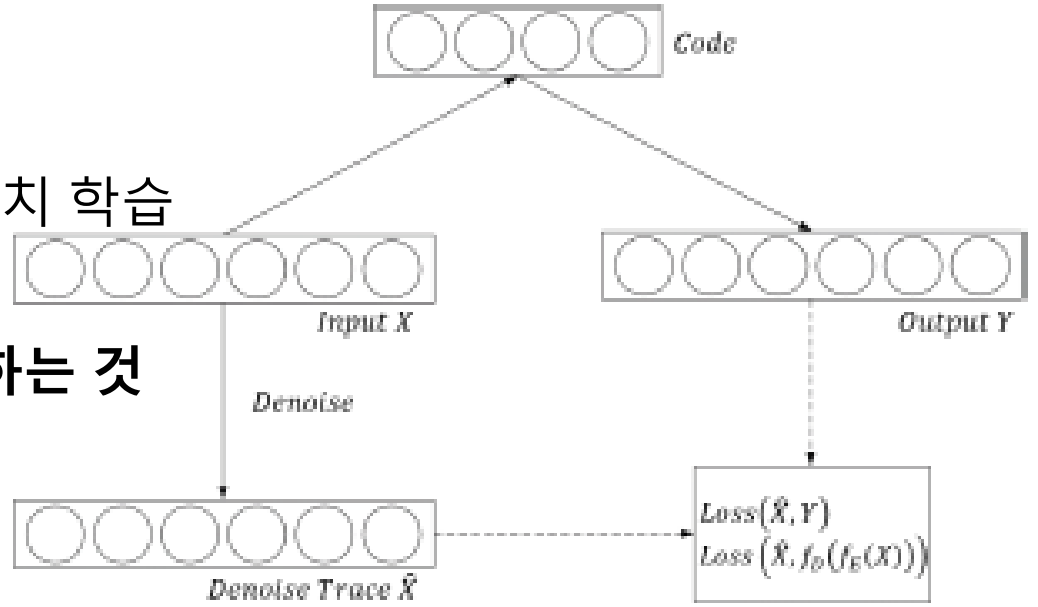
  ✓ 오토인코더
  원본데이터 X를 그대로 사용
  ✓ 노이즈 제거 오토인코더 (DAE)
  학습 데이터+노이즈가 들어간 데이터 X'
  를 뉴런 네트워크 입력으로 사용




- 공격자가 임의로 추가한 노이즈를 제거하여 원본 데이터를 복원할 수 있도록 뉴런 네트워크 학습

# Side-Channel Auto-Encoder

- 부채널 환경에서는 공격의 대상이 되는 연산은 어떤 특정 시점에서만 수행
  → 랜덤한 시점을 0으로 초기화하여 학습 데이터를 생성할 경우, 공격의 타겟이 되는 정보가 제외된 학습 데이터가 생성될 수 있어 부채널 분석 환경에 **적합하지 않음**
- DAE와 다르게 입력 데이터를 기본 오토인코더와  동일하게 학습데이터로 사용
- 전처리 기법을 이용하여 처리한 데이터를 라벨로 사용
- 출력데이터Y와 전처리데이터 X'와 손실을 계산하여 가중치 학습
- **수집한 전력파형에 존재하는 노이즈를 제거하도록 학습하는 것**
- 전처리 과정 이후, 논프로파일링에 해당하는
  기존 부채널분석 기법을 사용하여 비밀정보 복원 가능





권동근, 진성현, 김희석, 홍석희 (2019). 비프로파일링 기반 전력 분석의 성능 향상을 위한 오토인코더 기반 잡음 제거 기술. 정보보호학회논문지, 29(3), 491-501

# Q & A