DEEP LEARNING-BASED POWER ANALYSIS ATTACK FOR EXTRACTING AES KEYS

Authors : **Ismail Negabi, Smail Ait El Asri, Samir El Adib, Naoufal Raissouni**

Publisher : **Arabian Journal for Science and Engineering** ***(19 Sept 2023)***

R. KIRTHIKA ***(23MCS001)***

CRYPTOGRAPHY AND NETWORK SECURITY (CS-741)

Department of Computer Science and Engineering

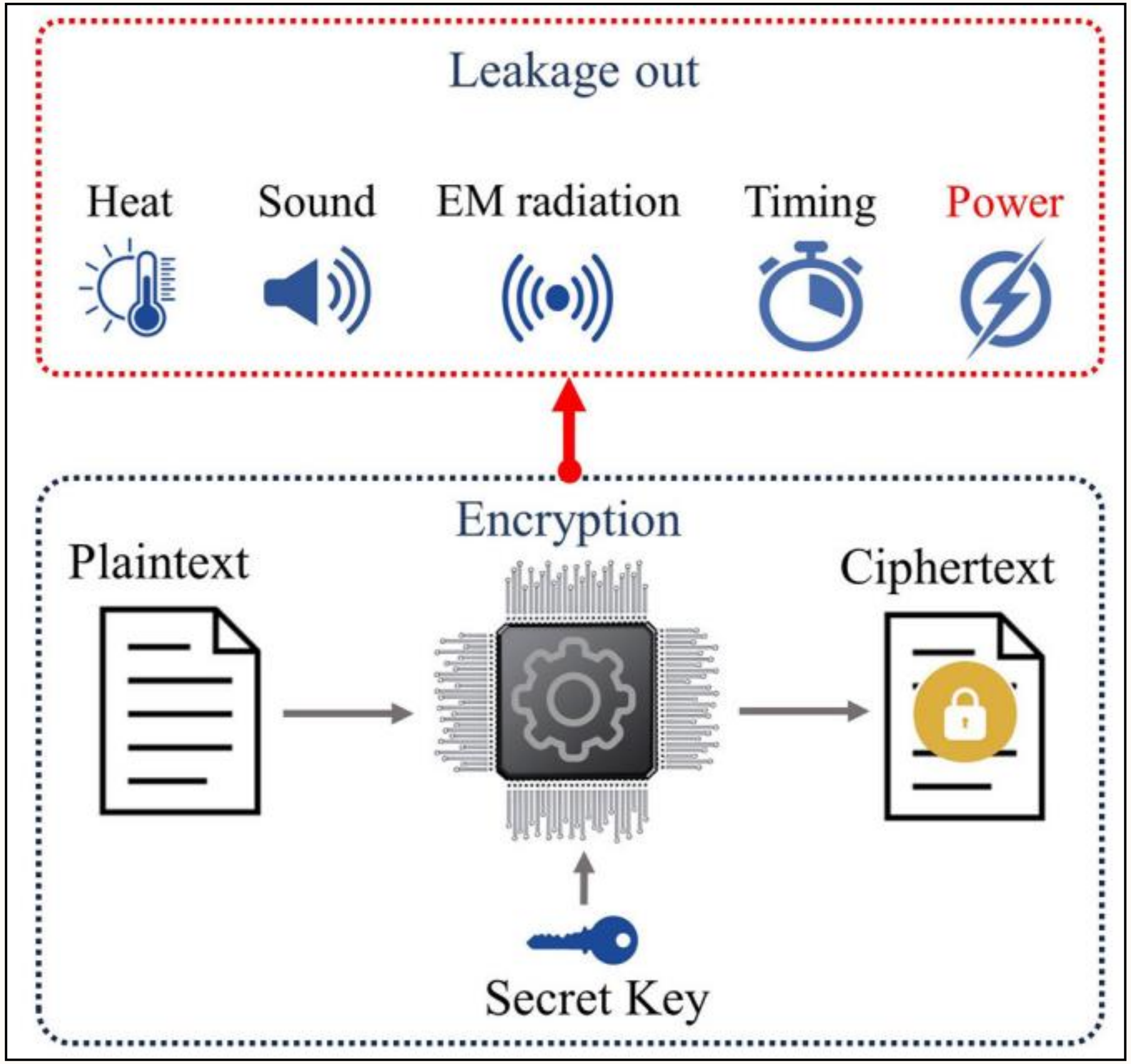
National Institute of Technology, Hamirpur

March 2024

WE WILL DISCUSS :

- Side-Channel Attacks
    - ➢ Power Analysis Attack
- Role of DL (CNN)
- Research Paper Methodology (Algo + CNN)
- Research Paper Results & Extensions
- References

INTRODUCTION



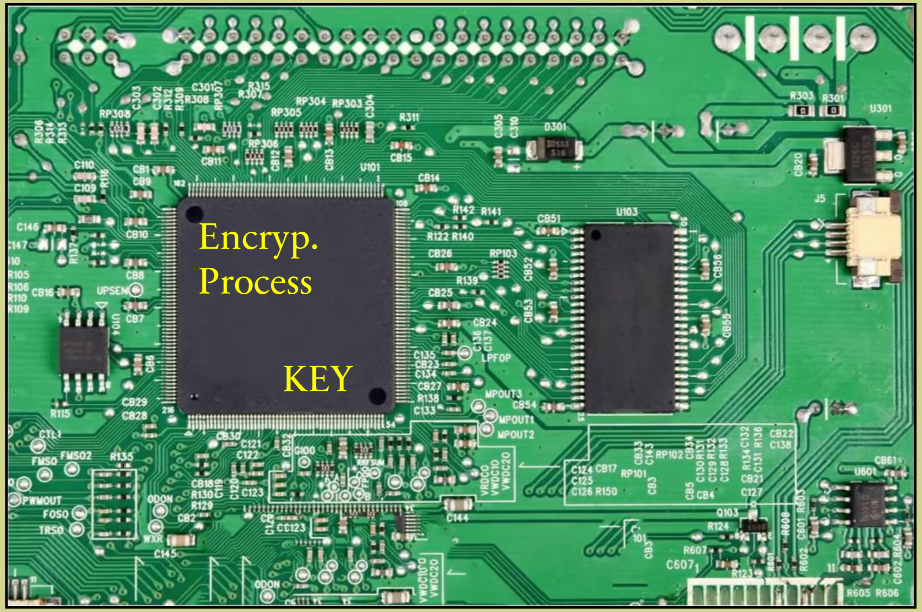
Side Channel Attack (SCA) exploits unintended leaks of information during the implementation of cryptographic algorithms.

SCA Types :
Timing Attack *(Kocher,1996)*
Power Analysis Attack
Electromagnetic Attack
Fault Analysis
Acoustic Analysis
Heat Analysis

Power analysis attack (**PAA**) is a method of SCA that aims to infer sensitive information, such as passwords or encryption keys, by analyzing the energy consumption fluctuations of a device.

SIDE-CHANNEL ATTACKS

POWER ANALSIS ATTACK



Output of Device(Target)

Power Consumption

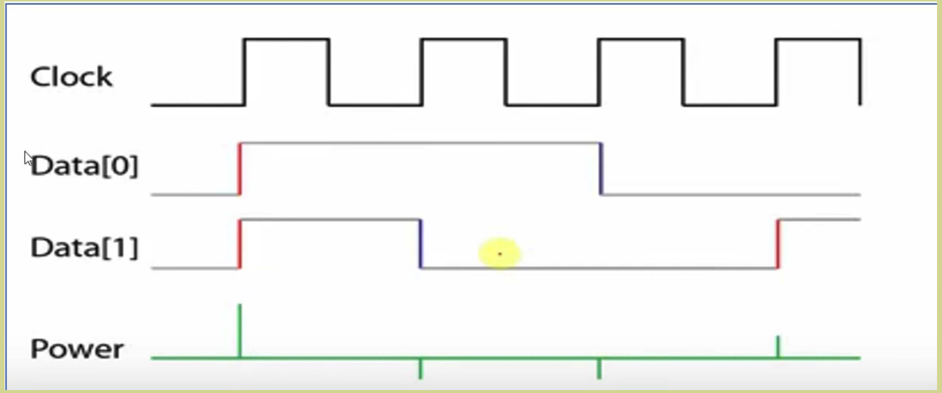
- Power is consumed when there is transition state 0→1 or 1→0.
- We measure this power using an Oscilloscope in order to determine the output state.

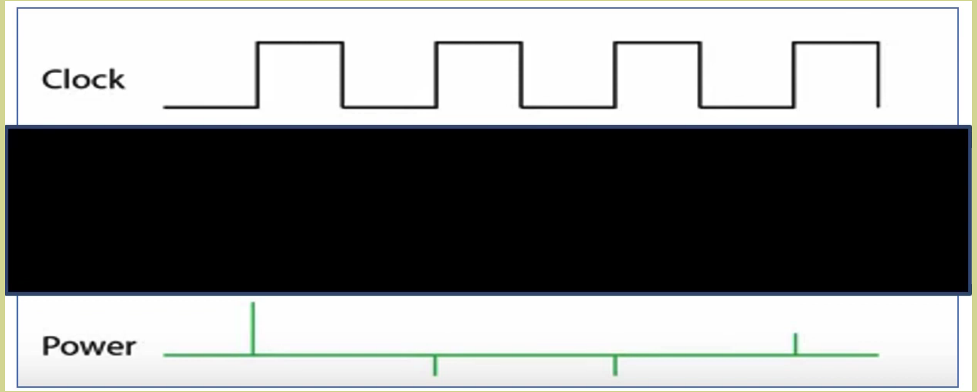
Assumptions :

- Attacker has the access to the target device that he want to attack.
- Attacker is able to monitor and regulate(control) the various inputs passed to the target device.
- Attacker is able to tap the power line data during any encryption process

POWER ANALSIS ATTACK (Contd.)

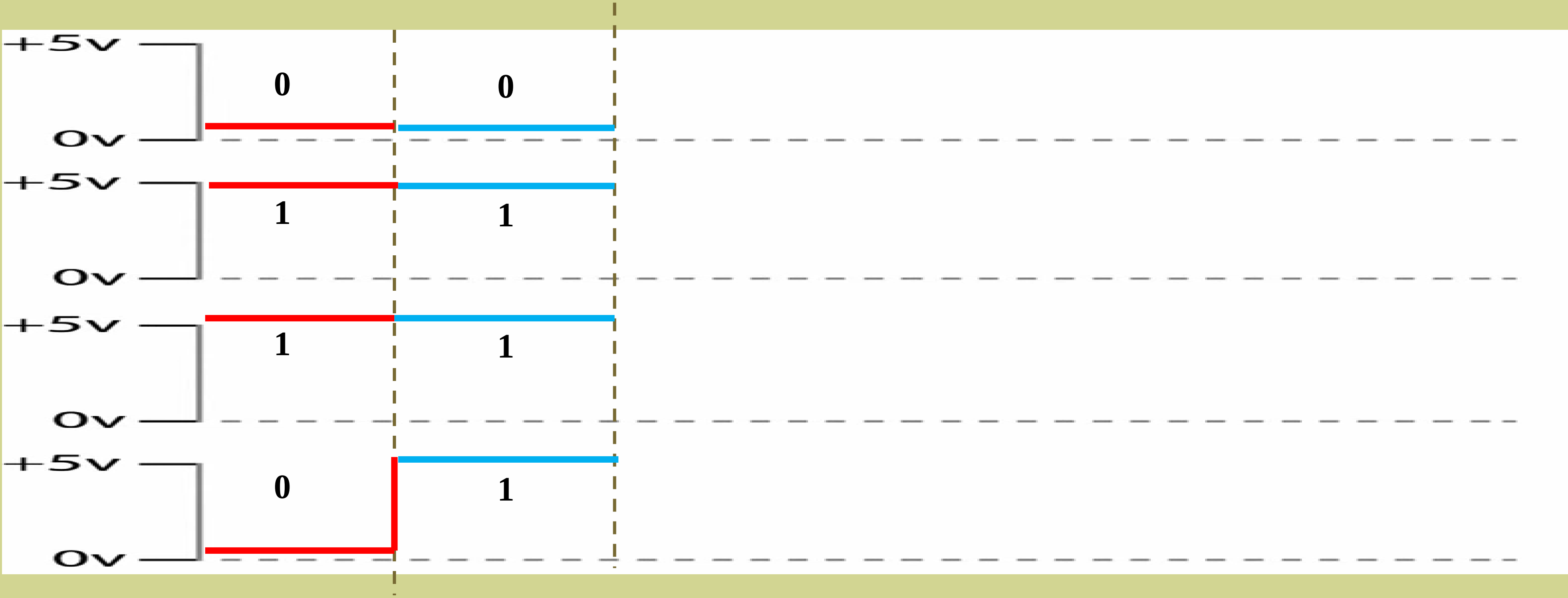
In a Nutshell…



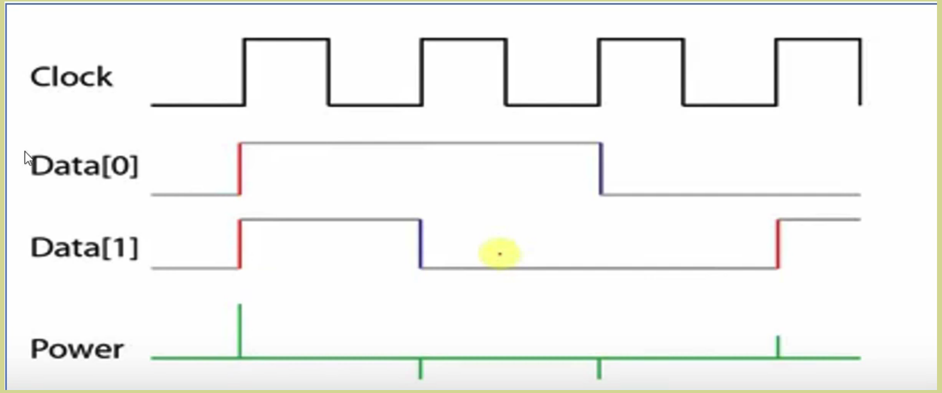


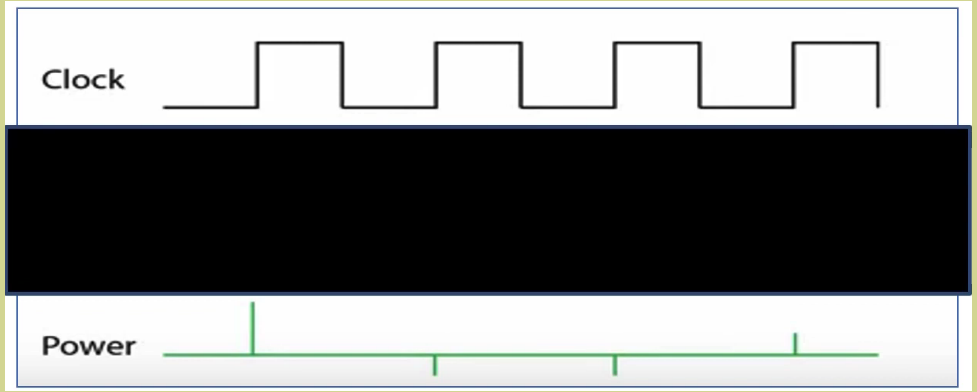
SIDE-CHANNEL ATTACKS

POWER ANALSIS ATTACK (Contd.)



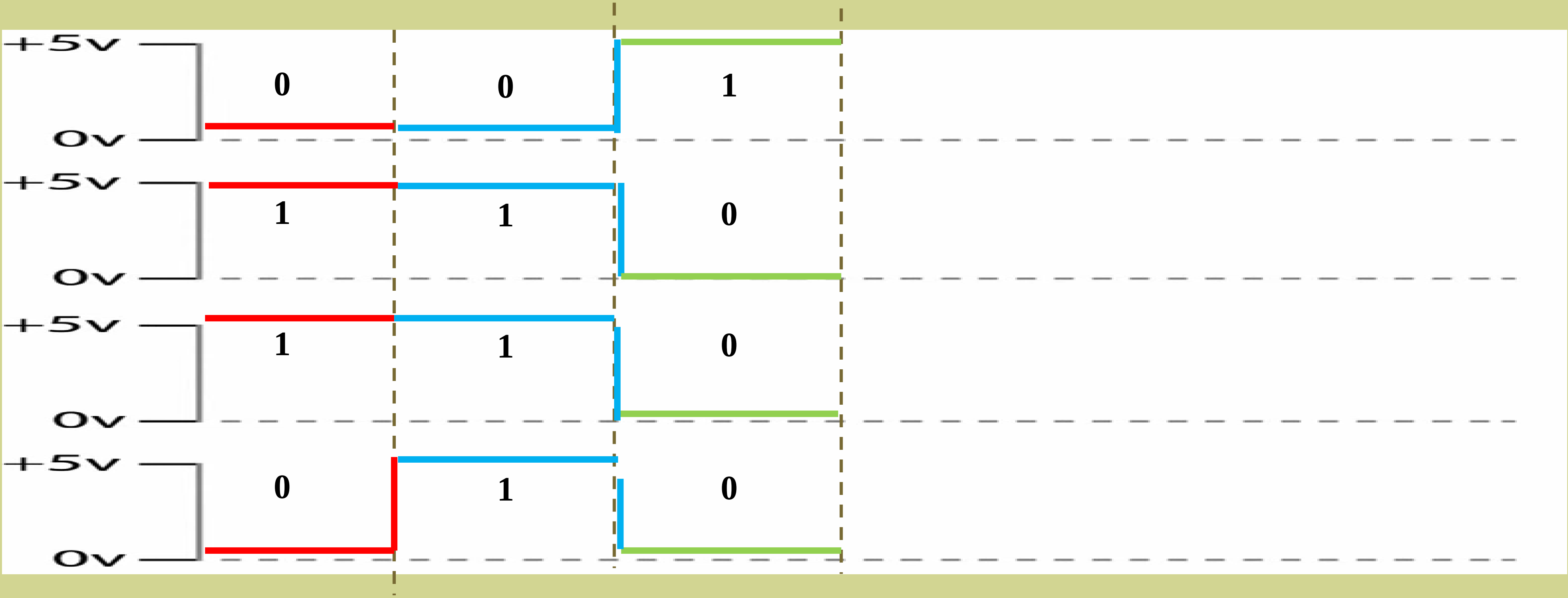
In a Nutshell…



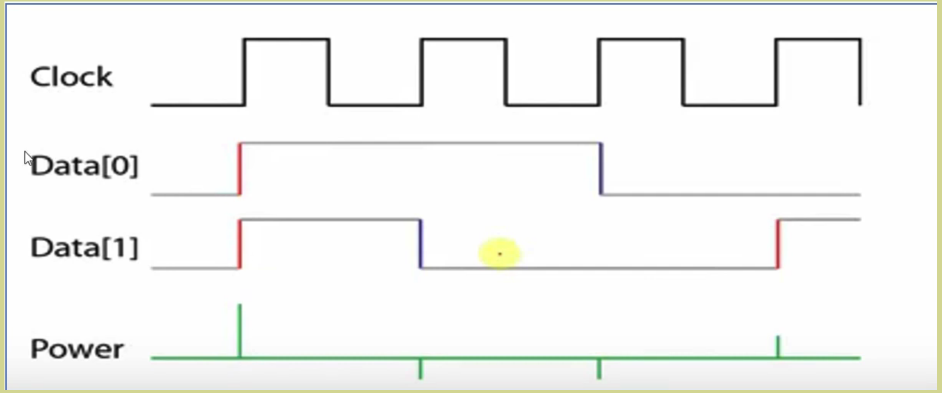


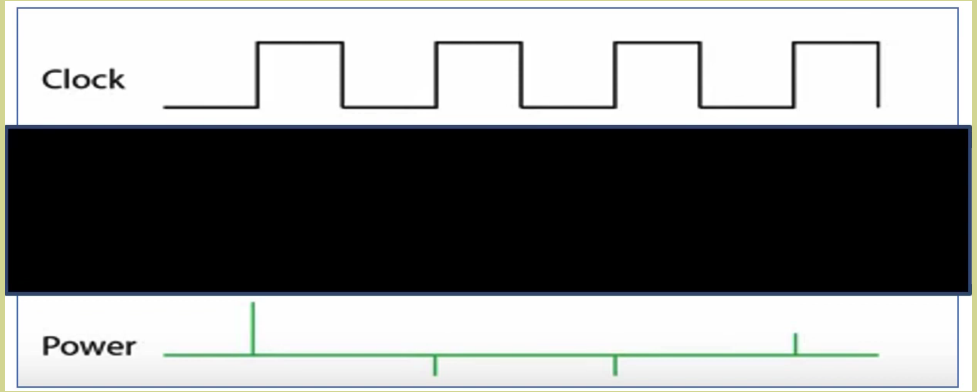
SIDE-CHANNEL ATTACKS

POWER ANALSIS ATTACK (Contd.)



In a Nutshell…

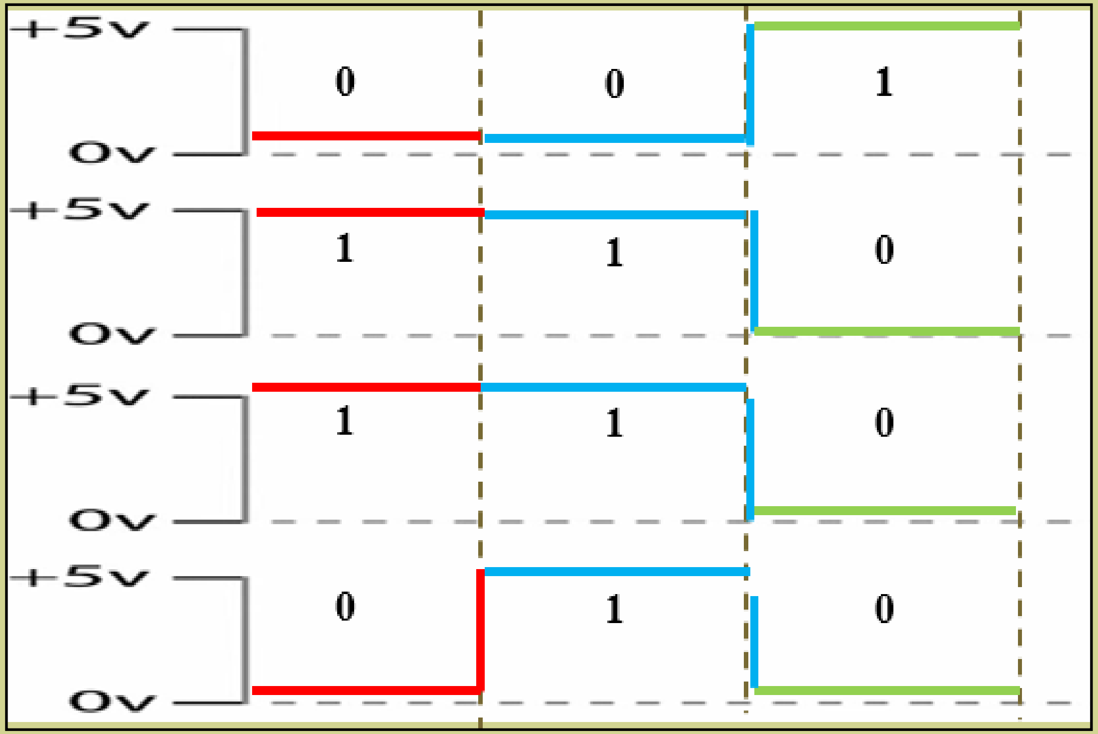


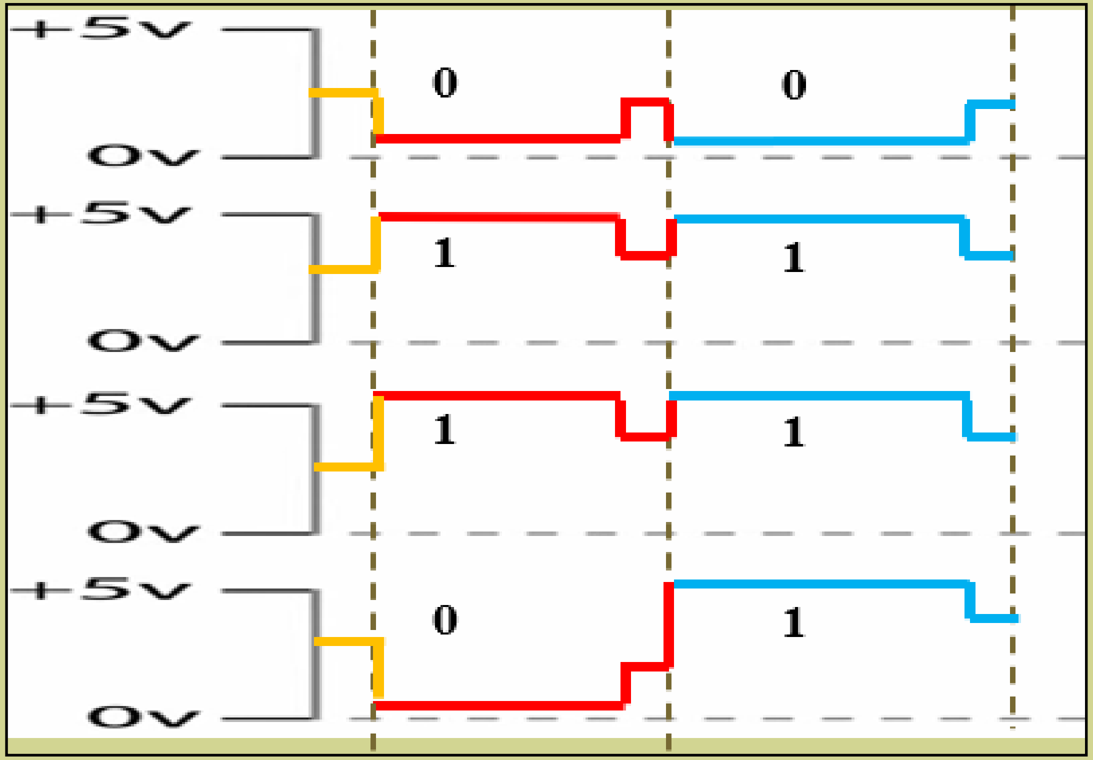


SIDE-CHANNEL ATTACKS

POWER ANALSIS ATTACK (Contd.)

Measuring the power (2 Methods) -





Hamming Distance – Power equivalent to number of bits changing on data bus.

(E.g.) (1011) → (1101) → (1001) → (0010) → (0011)
2 **1** **3** **1**

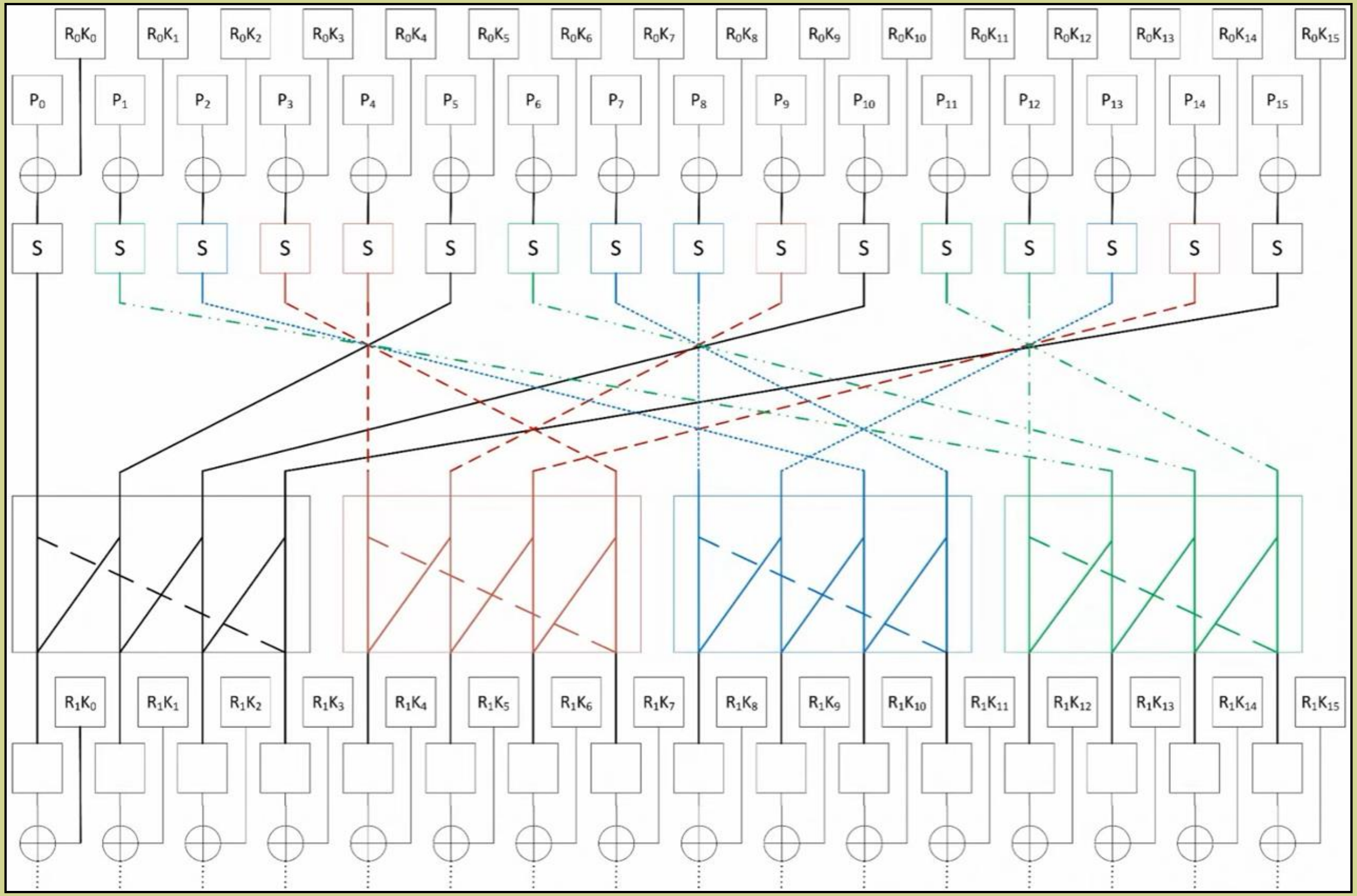
Hamming Weight – Power equivalent to number of bits set to 1 on data bus.

(E.g.) (1011) → (1101) → (1001) → (0010) → (0011)
3 **2** **1** **3**

SIDE-CHANNEL ATTACKS

POWER ANALSIS ATTACK (Contd.)

Applying to AES



AES KEY SPACE :

Brute Force :

2 ^128 = 3.4028 x 10^38

~1078289752 Trillion Yr

SCA :

Checking Key for each section

= 2^8 = 256

Overall Round (10 or 16)

= 256 x 10 (or 16)

= 2560 (or 4096)

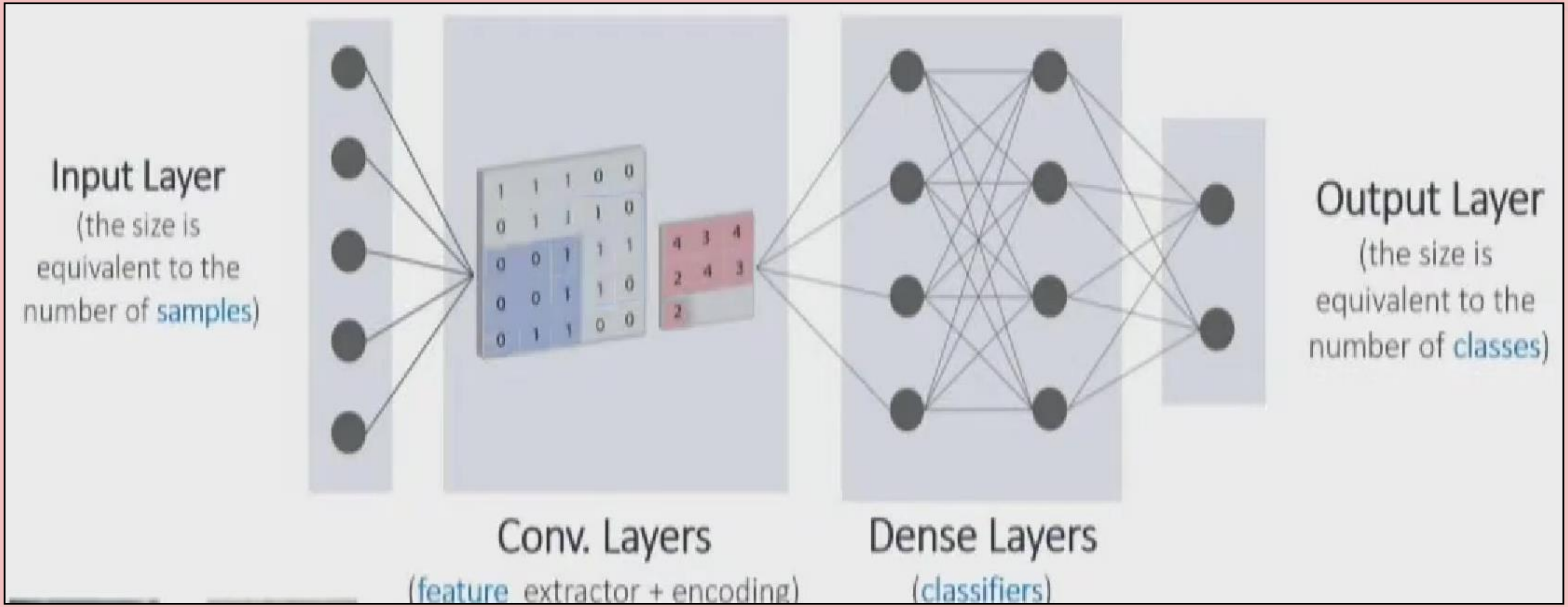
SIDE-CHANNEL ATTACKS

ROLE OF DEEP-LEARNING

Why Deep Learning Model (CNN) approach ??

- Used to improve efficiency and accuracy of the attacks
- Used to identify hidden patterns & correlations.

Convolution Neural Network (CNN) Basics -



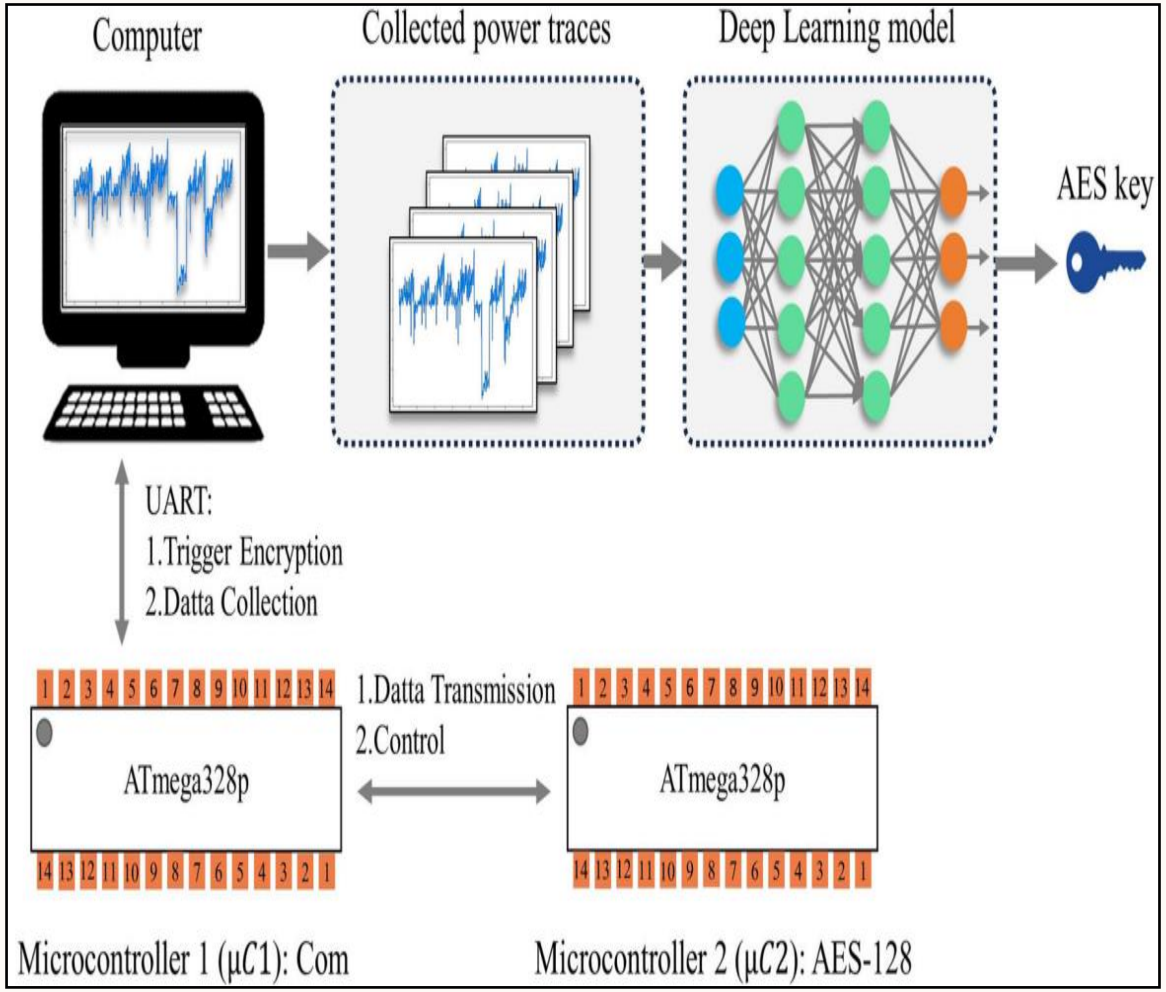
HARDWARE AND SOFTWARE CONFIGURATION

- ATmega328p microcontroller (µC2) – *(Target Device)*
- Oscilloscope
- ATmega328p microcontroller (µC1) – *(Interface)*
- Dell computer(i5-6300U, CPU@2.5 GHz, 8 GB of RAM and a 256 GB SSD.) – *(Attacker's Device)*
- Arduino IDE
- Python

DATASET (Power Consumption Traces)

Training Dataset : 100k traces (out of which 10%(10k traces for Validation)
Testing Dataset : 2k traces

DATA CAPTURE WORKFLOW

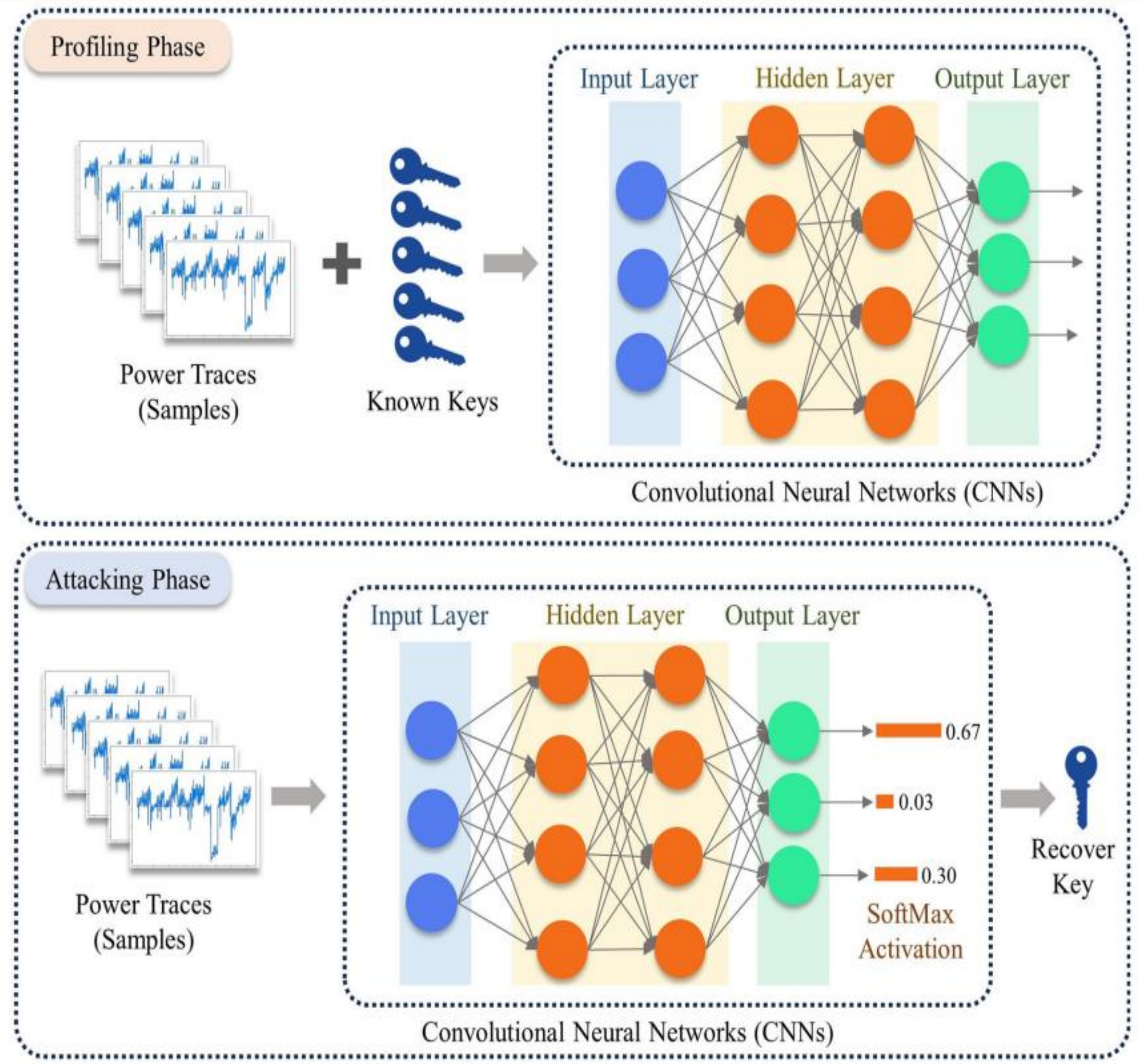


Algorithm 1 Data Collection for Side Channel Attacks

```
function DATA_COLLECTION(trainingPhase)
    if trainingPhase then
        key ← generate_random_key()
    else
        key ← fixedKey
    end if
    for i = 1 to 10 do
        text ← generate_random_text()
        send_to_μC1_via_UART(text, key)
         μC2_data ← receive_from_μC1_via_I2C()
         μC2.prepare_data(μC2_data)
         μC2.signal_start()
         trace ← μC1.collect_data()
         send_trace_to_python(trace)
     end for
     return average_results()
end function
```

RESEARCH METHODOLOGY

PROPOSED CNN ARCHITECTURE



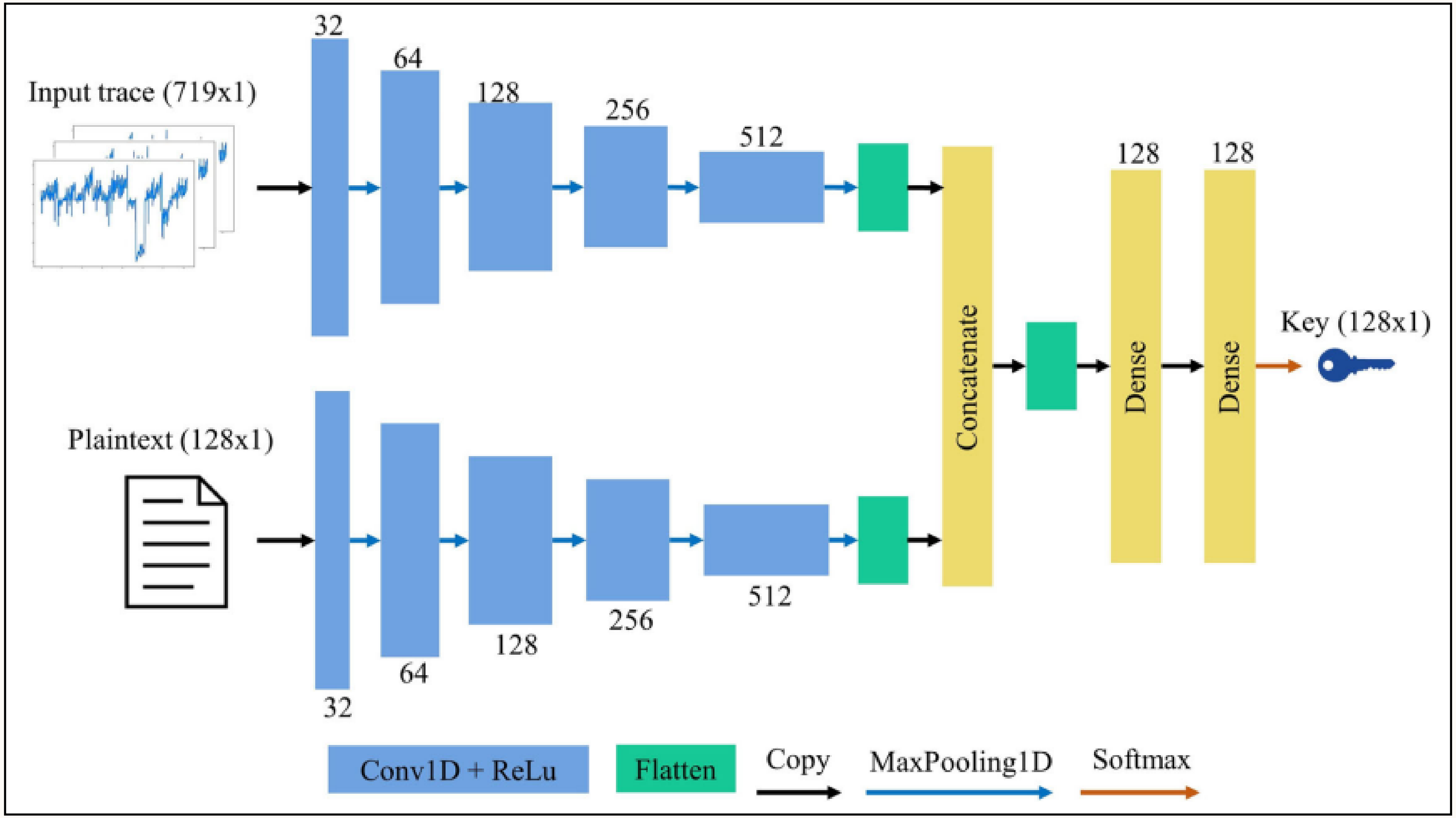
Profiling Step:

- collecting and analyzing power consumption data in order to build a model of the target device.
- Generate a profile of energy consumption by analyzing a considerable number of power trace
- The DL model is trained to map the power traces to the data being processed.
- The output of the DL model is a score vector which represents the probability that the data being processed at the attack point is a specific value

Attacking Step:

- The attacker uses the trained DL model to classify the power traces captured from the target device and obtain a score vector
- Then identifies the subkey Ki with the highest probability in the score vector and compares it to the true subkey.
- If the two values match, the subkey has been successfully recovered

PROPOSED CNN ARCHITECTURE



RESEARCH METHODOLOGY

RESEARCH RESULTS

RESULT

The author successfully shows that using the discussed approach he was able to recover the AES Key with only **1200 Traces** on avg.

LIMITATIONS

Issue of **Overfitting** during the validation phase

- Can be addressed by decreasing the model complexity as necessary

EXTENSION

- Extend our study to other popular microcontrollers.
- Explore other commonly used encryption algorithms and see if the said approach can also be used to extract keys from these algorithms.
- Possible countermeasures to protect microcontrollers against this type of attack

ANY QUESTIONS ?

THANK YOU

R.Kirthika

23MCS001

23MCS001@nith.ac.in