Side Channel Attack on 8051 microcontroller

Manda Yuktha, Pragati Patel, Vasantha M.H.
Department of Electonics & Communication Engineering
National Institute of Technology Goa, Goa, India
yukthamanda123@nitgoa.ac.in, pragati@nitgoa.ac.in, vasanthmh@nitgoa.ac.in

***Abstract*—Launching a Power Analysis Side-Channel Attack (SCA) involving predicting the contents of a register within an 8051 (AT89S52) microcontroller. This prediction is made possible by repeatedly executing the same instruction with various data inputs and monitoring the power consumption patterns using an Oscilloscope. Power traces are analyzed using machine learning techniques to forecast the data after noise reduction. The underlying concept of this method is that a microcontroller's power usage is influenced by two factors: the operations being performed, and the data being processed. By keeping the instruction constant, any variation in power consumption is attributed solely to the differences in the data being handled.**

***Index Terms*—Differential Power analysis, Side Channel Attack, microcontroller, 8051, machine learning**

I. Introduction

A. Cybersecurity in Semiconductors

Semiconductors are crucial components of electronic devices. Earlier, Side Channel attacks were only relevant to Cryptographic devices In the past few years, a variety of attacks targeting cryptographic devices have been disclosed. The primary objective of these attacks is to extract the confidential keys from such devices. Power analysis attacks take advantage of the reality that the instantaneous power usage of a cryptographic device is influenced by both the data it is handling and the specific operations it is executing. Importance of semiconductors is growing in the automotive industry due to the increase in vehicle automation, and the trend towards more connected vehicles. However, the cost of testing these integrated circuits very high. With the advancement of vehicle connectivity, there is also a heightened risk of cyber-attacks, which has significantly raised consumer awareness and demand for strong cybersecurity measures in connected vehicles. These attacks feasibility is rated based on the financial investment, duration, necessary equipment, and level of technical knowledge required. As a result, these attacks can be classified in multiple ways based on differing factors.

B. Power Analysis SCA

Power analysis exploits the instantaneous power which is effected by operation, data, noise and constant factor. Side channel attack is accessing the information that is exploitable. Simple Power Analysis (SPA)[1] is a type of SCA where the attacker examines the power consumption of a cryptographic device or a SoC to identify distinctive patterns. These patterns, observed through the inspection of power traces, can be linked to particular operations or instructions within the device. These attacks use a single trace. Differential Power Analysis (DPA) is an advanced side-channel attack that uses statistical techniques to analyze numerous power consumption readings. By correlating fluctuations in power usage with various hypothesized scenarios of the data being processed, DPA can reveal sensitive information, typically necessitating a substantial number of power traces for precise extraction. This technique used multiple traces and performs reverse engineering to extract the information from the power.

Performing power analysis attacks in the semiconductor industry can offer several advantages, particularly in the context of security testing and the development of robust cryptographic devices, and by addressing these vulnerabilities with improved countermeasures builds customer confidence.

C. 8051 microcontroller

8051 is considered target device in this paper due to its relatively simple architecture and modest bandwidth, which allows for a more straightforward analysis and understanding of side-channel leakage. Its low clock frequency is another advantageous feature, as it slows down the execution of instructions. This simplicity and the inherent potential vulnerability of the 8051 to SCAs make it an excellent starting point for developing methodologies that could later be applied to more complex and sophisticated (SoCs).

The duration required to perform an SCA on a microcontroller can vary greatly, depending on factors such as the complexity of the system, the level of noise in measurements, the proficiency of the attacker, the process can be time-consuming, potentially taking few years. However, the insights gained from such in-depth analysis are invaluable, as they contribute to the creation of more secure systems by informing the design of robust countermeasures against SCAs.

D. 8051 Architecture

In the 8051 microcontroller architecture[1], Consider executing an instruction is executed to move a 0 bit into the LSB of port 0. Initially, the CPU fetches the instruction from the Flash memory where the program is stored Fig. 1. This fetching process involves placing the address of the instruction onto the address bus and retrieving the instruction data.

Upon decoding the instruction, the CPU identifies the operation to write a value to an I/O port. It then proceeds to manipulate the corresponding Special Function Register

2024 IEEE International Symposium on Smart Electronic Systems (iSES) | 979-8-3315-3322-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISES63344.2024.00080

2832-3602/24/$31.00 ©2024 IEEE
DOI 10.1109/iSES63344.2024.00080

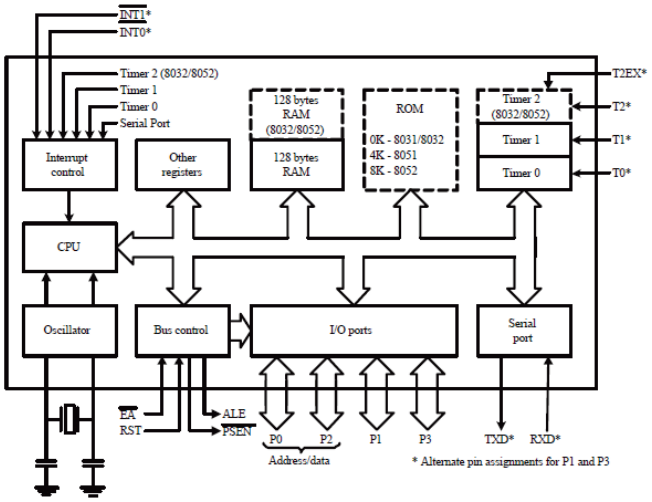

Fig. 1. Architecture diagram of 8051 Microcontroller
Courtesy:https://userdiagramjunker.z19.web.core.windows.net/explain-block-diagram-of-8051-microcontroller.html

(SFR) for port 0, specifically setting the LSB to 0. This action changes the electrical state of the LSB pin of port 0 to a logical 0. The power consumed by this process is being utilized to perform the Side Channel Attack.

II. Power Consumption

The Side channel attacks leverage the relationship between power consumption which is a characteristic of CMOS logic circuits and the data being processed[2]. Each data point in a power trace can be conceptualized as a composite of several components: an operation-dependent component P_{Op}, a data-dependent component P_{data}, electronic noise $P_{el.noise}$, and a constant baseline component P_{const}.

$$P_{Total} = P_{Op} + P_{data} + P_{el.noise} + P_{const} \quad (1)$$

III. Existing methods

A SPA using only single power trace has been performed on the 8051 microcontroller [1]. To examine the data dependency of power consumption in a microcontroller, either the LSB or the most significant bit (MSB) is toggled between 1 and 0. This alteration helps to study how power usage is related to the data being processed. Additionally, the study demonstrates a correlation between power consumption and the executed operations, such as AND, OR, JUMP, and NOOP, through SPA. This analysis confirms the susceptibility of the 8051 microcontroller to SPA attacks. This existing method uses the Limiting resistor in power or GND line and measures the Voltage accross it which is proportional to the Power consumption of the microcontroller.

A DPA attack was carried out on PIC18F2420 targeting its cryptographic algorithm [4].Secret key is being extracted using different bits of intermediate target values using statistical methods.

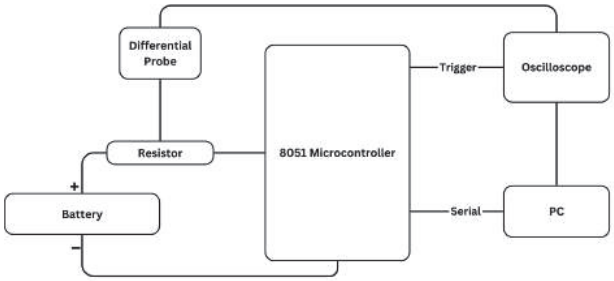

Fig. 2. Block diagram of Experimental setup overview

IV. Experimental setup

The development board has on board programmer allows easy connection with PC using USB type B cable for Programming. Resistor of 1 Ohm is to be connected in the power line of the Microcontroller across which the Voltage is being measured using differential probes as mentioned in [1]. Any noise originating from the device's power supply can introduce disturbances into the recorded power traces. Therefore, it is recommended to utilize a highly stable power source instead of powering the device through a PC's USB or PS/2 ports, which may introduce additional noise. Keil MicroVision (µVision) IDE is being used to code the 8051 microcontroller and debugging. PRG ISP is used for in-system programming of 8051 directly within the target system without removing the microcontroller from its application circuitry.

V. Methodology

A. Hardware circuit implementation

The block diagram represents the overview of setup in Fig 2.

B. Challenges

Selection of appropriate equipment is challenge due to the availability of different hardware cofigurtions of 8051 micocontroller. Each configuration may affect the signals differently, thus influencing the ease with which the attack can be executed and the quality of the data extracted. A breadboard setup with individual components such as capacitors, crystal oscillators, and the 8051 microcontroller allows customization but manually assembled circuits may have variations that affect repeatability of measurements. Microcontroller with Programmer Interface require additional work to integrate equipment. Development Board with Inbuilt Programmer is a better choice as it streamlines the setup process, offers more stable electrical connections, reducing noise. Development boards may provide easier access to power traces for capturing side-channel information. The uniformity of development boards means that results are more consistent and reproducible across different experiments or setups.

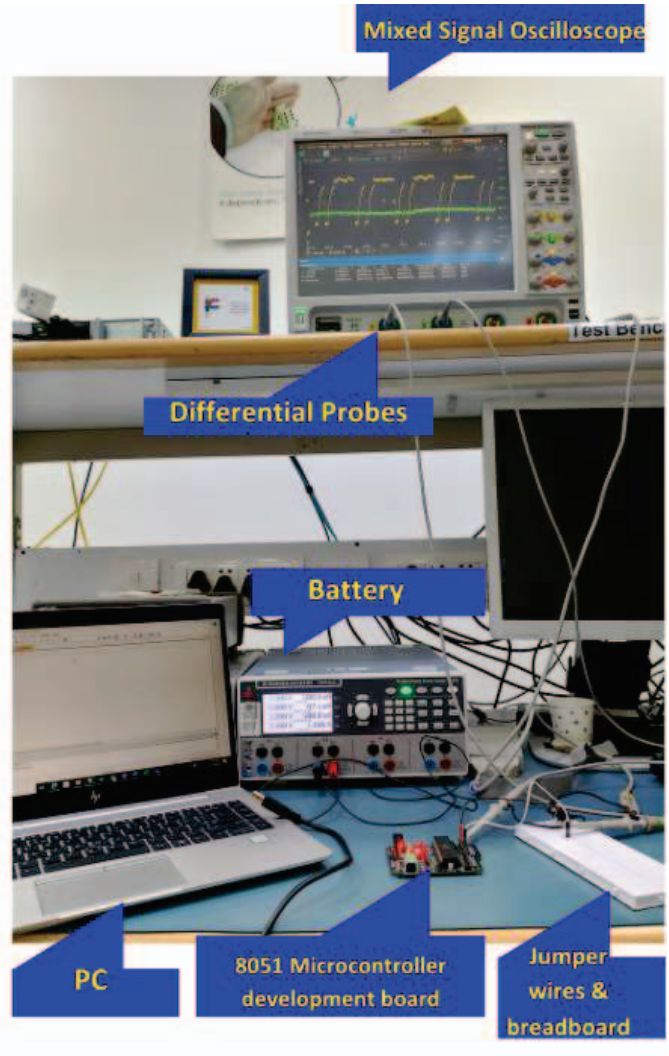

Fig. 3. Experimental setup for Power Analysis SCA

Another challenge is the difficulty of precisely measuring current consumption, as this requires current probes capable of detecting microamp-level changes. To circumvent this, a resistor is typically inserted into either the power or ground line of the microcontroller[1] as shown in Fig. 4. Using differential probes to measure the voltage drop (which is in the range of few milli-Volts) across this resistor in power line, one can infer the microcontroller's current usage. This voltage measurement is proportional to the current drawn, can be captured by an oscilloscope, thereby facilitating the execution of the side-channel attack.

C. Analysis

The Side channel attack has been performed to predict the register value stored in the LSB of port0. Assembly code is used to perform the task of toggling an output pin with a delay between state changes. The delay subroutine is crucial for creating a visible toggle effect on P0. Without it, the toggling would occur too rapidly to be observable. The actual duration of the delay will depend on the clock frequency of the microcontroller and may need to be calibrated for the desired toggle rate.

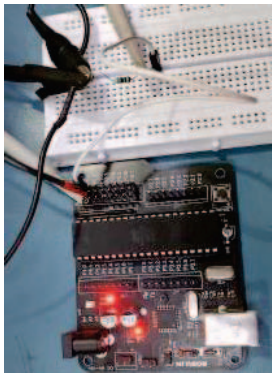

Fig. 4. Resistor connected in power line to measure the voltage accros

During the execution of the assembly code, power traces are captured. These traces are gathered using a mixed-signal oscilloscope, which facilitates the subsequent analysis in MATLAB. The collected traces are saved in the form of discrete signals within CSV (Comma-Separated Values) files. The analysis described involves processing the data through feature extraction and applying filters to prepare it for the classification stage.

D. Algorithm for Pre processing data

The MATLAB script is designed for data preprocessing in signal analysis, potentially for use in machine learning or pattern recognition applications. It sequentially processes several CSV files, retrieving features and labels, detecting signal peaks, and performing downsampling. The script further normalizes and applies a low-pass filter to the signal. Finally, it aggregates the refined data and corresponding labels for further analysis as shown in Fig. 5. The step-by-step algorithm describes the process that the MATLAB code is designed to perform. It involves reading data from multiple CSV files, extracting features, finding peaks, downsampling, normalizing, filtering, and finally concatenating the processed data.

E. Algorithm for Linear Discriminant Analysis

The preprocessed data is processed through Linear Discriminant Analysis (LDA) to classify the single bit hypothesized value as described in the algorithm in Fig. 6. The features and labels are extracted from the prepared .csv concatenated file. The data is split into 80% for training and 20% for testing and processed through LDA classifier with 5 fold cross validation.

1) Calculate the mean vectors for each class: For each class i, calculate the mean vector $\vec{m}_i$ of the samples in that class.
2) Compute the within-class scatter matrix S_W:

$$S_W = \sum_{i=1}^{c} S_i$$

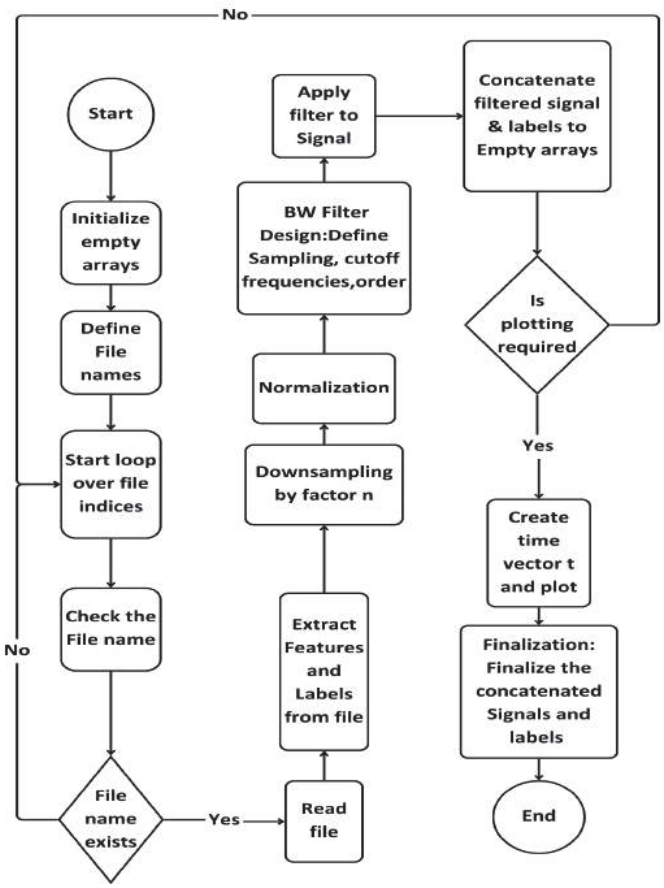

Fig. 5. Algorithm for Pre Processing all the traces

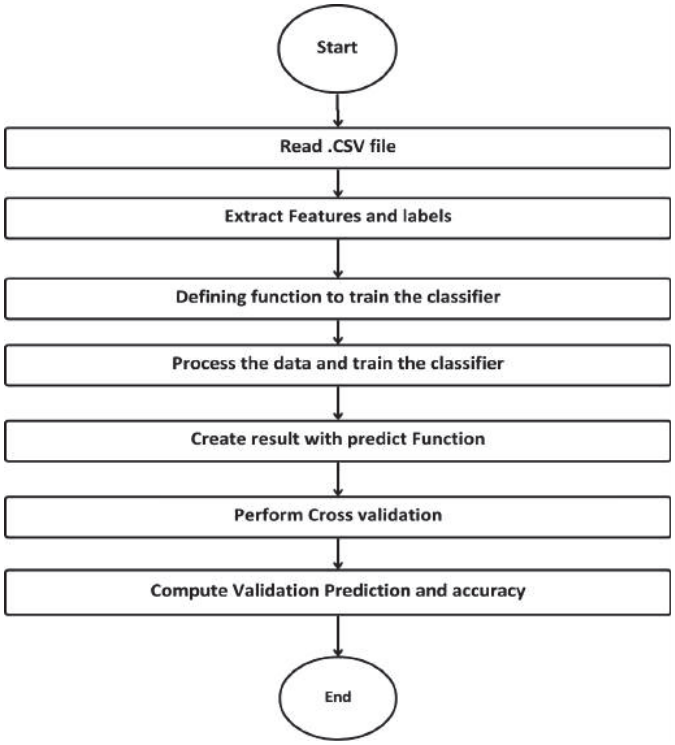

Fig. 6. Algorithm for Classification training and testing

where S_i is the scatter matrix for each class i:

$$S_i = \sum_{x \in D_i} (x - \vec{m}_i)(x - \vec{m}_i)^T$$

Here, D_i is the set of all samples belonging to class i.

3) Compute the between-class scatter matrix S_B:

$$S_B = \sum_{i=1}^{c} N_i (\vec{m}_i - \vec{m})(\vec{m}_i - \vec{m})^T$$

where $\vec{m}$ is the overall mean of the dataset, and N_i is the number of samples in class i.

4) Solve the eigenvalue problem for the matrix $S_W^{-1} S_B$: Find the eigenvalues and corresponding eigenvectors for $S_W^{-1} S_B$.
5) Select the linear discriminants for the new feature subspace: Choose the eigenvectors (linear discriminants) that correspond to the largest eigenvalues to form the matrix W. The number of eigenvectors chosen is less than or equal to $c - 1$, where c is the number of class labels. This is a mathematical step that involves finding directions in your data that best separate your classes.
6) Project the samples onto the new feature space: Use the matrix W to transform the samples onto the new subspace:

$$X_{\text{new}} = XW$$

Finally, the original data and the directions are selected to transform it into this new space. Now, the data in this new space, different classes are more distinct and easier to tell apart. This classifier learns to recognize patterns in the data that correspond to the different classes. It effectively learns where the boundaries between classes are in this new feature space.
The goal of all these steps is to find a new way to look at the data that makes it easier to see the differences between groups.

VI. Results

The accuracy of prediction is analysed using the Confusion matrix. The Confusion matrix gives the Positive Predictive Values (PPV) percentage and the False Discovery Values (FDV) for each predicted bit. In the Fig. 7 the bit 0 is predicted accurately as 0 by 95.5% but wrongly as 1 by 4.5% . The bit value 1 is correctly predicted as 1 by 79.8% and wrongly predicted as 0 by 20.2 %.

The Fig. 8 corresponds to the traces after Alignment and segmentation, original signal is the downsampled after extracting the peaks and the original signal filtered using the

Fig. 7. Confusion matrix for traces undergoing LDA Model

Fig. 9. Confusion Matrix for traces filtered using BW LPF and then classified using LDA

Low pass IIR Butterworth filter and is represented as Filtered signal(Low-Pass)

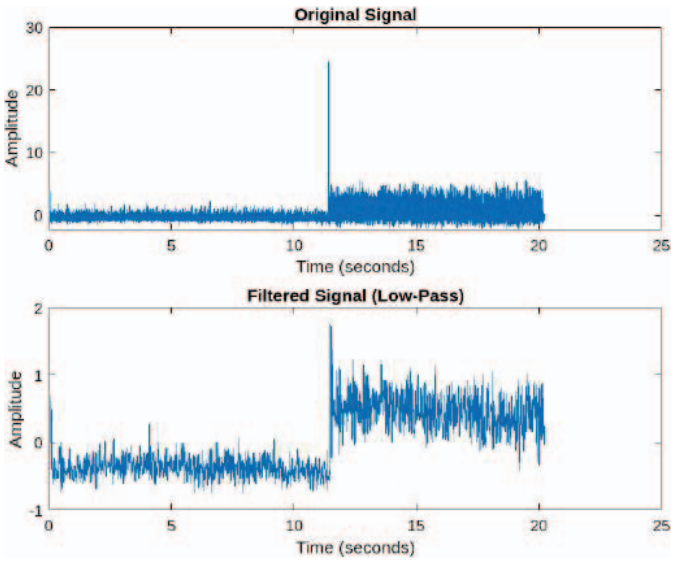

Fig. 8. Each trace is filtered using the IIR BW filter

The PPV has been improved to 98.8% for bit 0 prediction and the FDV is reduced down to 1.2% for the same. The PPV for predicting the bit 1 is improved to 91.9% and FDV 8.1% as shown in the model Fig.9.

VII. Comparision Results

Criteria	LDA with out filter	LDA with filter
Percentages	85.7%	93.8%

The comparision table represents the prediction accuracy percentage before and after using the filter.

VIII. Conclusion and Future Scope

This demonstrates that the 8051 is vulnerable to Power Analysis Side-Channel Attacks (SCA). By observing the assembly code execution timing, one can predict the bit value in the LSB of Port 0 when it toggles within specific time intervals. This process involves aligning and segmenting traces, preprocessing each trace using signal processing techniques, and then reducing noise with a low-pass IIR filter. The prediction is done with the help of machine learning techniques to classify the bit value. Through tailored preprocessing, higher-frequency bit value predictions can be achieved.

As SCA techniques continue to evolve, they can be applied to more sophisticated SoCs, which are integral to a vast array of modern devices. By extending the granularity of SCA to the instruction level, researchers and security professionals can gain a more nuanced understanding of the data processing within a microcontroller, leading to more accurate predictions of internal states. This increased accuracy is crucial for both the development of secure cryptographic implementations and for conducting effective vulnerability assessments.

To ensure the development of countermeasures to protect against SCAs to ensure integrity and confidentiality of data and to create more robust testing and validation protocols that ensure hardware components meet stringent security standards before deployment.

IX. Acknowledgement

Sincere thanks to STMicroelectronics for enabling to use the equipment and Priyank Sharma for mentoring the Project.

References

[1] P. Saravanan, N. Rajadurai and P. Kalpana, "Power analysis attack on 8051 microcontrollers," 2014 IEEE International Conference on Computational Intelligence and Computing Research, Coimbatore, India, 2014, pp. 1-4, doi: 10.1109/ICCIC.2014.7238441.

[2] M. Aigner, S. Mangard, R. Menicocci, M. Olivieri, G. Scotti and A. Trifiletti, "A novel CMOS logic style with data independent power consumption," 2005 IEEE International Symposium on Circuits and Systems, Kobe, Japan, 2005, pp. 1066-1069 Vol. 2, doi: 10.1109/ISCAS.2005.1464776.
[3] T. S. Messerges, E. A. Dabbish, and R. H. Sloan, "Examining smartcard security under the threat of power analysis attacks,"IEEE Trans. Comput., vol. 51, no. 5, pp. 541–552, 2002.
[4] K. Mpalane, H. D. Tsague, N. Gasela and B. M. Esiefarienrhe, "Bit-Level Differential Power Analysis Attack on Implementations of Advanced Encryption Standard Software Running Inside a PIC18F2420 Microcontroller," 2015 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2015, pp. 42-46, doi: 10.1109/CSCI.2015.115.
[5] H. Gupta et al., "Impact of Side Channel Attack in Information Security," 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, 2019, pp. 291-295, doi: 10.1109/ICCIKE47802.2019.9004435.
[6] N. Gattu, M. N. Imtiaz Khan, A. De and S. Ghosh, "Power Side Channel Attack Analysis and Detection," 2020 IEEE/ACM International Conference On Computer Aided Design (ICCAD), San Diego, CA, USA, 2020, pp. 1-7. keywords: Power grids;Resistance;Pins;Impedance;Cryptography;System-on-chip;Side-channel attacks;Power side channel attack detection;power grid;phase detection,
[7] Liyao, T. Ju and Z. Chunlian, "MLP-Based Power Analysis Attacks with Two-Point Joint Feature Selection," 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, 2020, pp. 250-254, doi: 10.1109/ICCWAMTIP51612.2020.9317303. keywords: Feature extraction;Side-channel attacks;Power demand;Cryptography;Correlation coefficient;Correlation;Analytical models;Side-channel attacks;Machine learning;Two-point joint;Correlation coefficient,