# CS771: Assignment 1

Anirudh [200128], Bhavya Gupta [22111017], Kunal Singh [200535]
Lakshmi Pravallika [200282], Sana Chaitanya [200599], Shaijal Tripathi [22111274]

August 17, 2023

## 1 To crack a simple XORRO PUF by a single linear model.

### 1.1 Time period of oscillation in a single XOR gate

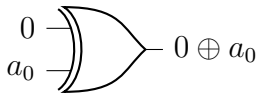


Figure 1: XOR gate circuit.

For a single XOR gate, let us define the following quantities:
$\delta_{00}$ : Time taken for signal transition from input to output when both the inputs are 0.
$\delta_{01}$ : Time taken for signal transition from input to output when inputs are 0 and 1.
$\delta_{10}$ : Time taken for signal transition from input to output when inputs are 1 and 0.
$\delta_{11}$ : Time taken for signal transition from input to output when both the inputs are 1.

Let $t_1$ be the time taken for the input to change from 0 to 1. Let $a_0$ be the configuration bit.Then,

$$t_1 = a_0\delta^0_{01} + (1 - a_0)\delta^0_{00}$$

Let $t_2$ be the time taken for the input to change from 1 to 0. Then,

$$t_2 = a_0\delta^0_{11} + (1 - a_0)\delta^0_{10}.$$

To calculate the time period $t$ for a single XOR gate:

$$t = t_1 + t_2 = a_0(\delta^0_{01} + \delta^0_{11}) + (1 - a_0)(\delta^0_{00} + \delta^0_{10}).$$

## 1.2 Time Period of oscillation in a chain of r-XOR gates

From the expression of $t$, we can determine the time period of oscillation of signal for a sequence of R number of XOR gates.

$$T = T_0 + T_1 + T_2 + ... + T_{R-1}$$

where $T_i$ represents the time period of the $i^{th}$ XOR gate.

$$T_i = a_i(\delta^i_{01} + \delta^i_{11}) + (1 - a_i)(\delta^i_{00} + \delta^i_{10}).$$

Let $x_i = (\delta^i_{01} + \delta^i_{11}) - (\delta^i_{00} + \delta^i_{10})$ and $y_i = (\delta^i_{00} + \delta^i_{10})$. Thus, we can say that

$$T_i = a_i x_i + y_i.$$

$$T = \sum_{i=0}^{R-1}(a_i x_i + y_i) = \sum_{i=0}^{R-1}(a_i x_i) + b$$

where $b = \sum_{i=0}^{R-1} y_i$. (The definitions of quantities $b$, $x_i$ are used in subsequent sections.)

If we let $x = [x_0, x_1, .., x_{R-1}]$ and $a = [a_0, a_1, .., a_{R-1}]$, then

$$T = x^T a + b.$$

## 1.3 Linear model for simple XORRO PUF

As the response of simple XORRO PUF depends on the frequency of the two XORROs we can introduce a new term $\triangle = F^u$- $F^l$ where, if $\triangle > 0$, the response is 1 otherwise 0.
In terms of time periods of XORROs, $\triangle = T^l - T^u$, where $T^l$ and $T^u$ represents the time period of the two XORROs used in a simple XORRO PUF.

$$\triangle = x_l^T a_l + b_l \text{ -}( x_u^T a_u + b_u)$$

As we have the same config bits, thus $a_l = a_u = a$. Also, for the two XORROs, the x vector (first defined in section 1.2) can be written as:

$$x^l = [x^l_0, x^l_1, .., x^l_{R-1}],\ x^u = [x^u_0, x^u_1, .., x^u_{R-1}]$$

Let $x_{diff} = [x_0^l - x_0^u, x_1^l - x_1^u, .., x_{R-1}^l - x_{R-1}^u]$ , b=$b_l - b_u$.

$$\triangle = x_{diff}^T.a + b$$

**is the linear model to crack the simple XORRO PUF.**

The response can be written as follows: $\frac{1+sign(\triangle)}{2}$
Comparing with the given equation for simple XORRO, is $\frac{1+sign(w^T\phi(c)+b)}{2}$ where c are challenge bits, **We get w=$x_{diff}$ and $\phi(c) = a = c$ ($a$ is the vector of the external inputs to the XORROs)**
$\phi : \{0,1\}^R \rightarrow \mathbb{R}^R$ implying that the number of dimensions is equal to the length of the vector $a$.

# 2 To crack the Advanced XORRO PUF

## 2.1 Introduction

Advanced XORRO PUF makes use of $2^S$ XORROs and the output of each of the oscillators goes to two multiplexers to finally select two of the XORROs out of $2^S$ XORROs.

The select bits ($S$ bits) of each of the multiplexer are defined as:

$$p = [p_0, p_1, .., p_{S-1}] \text{ and } q = [q_0, q_1, ..., q_{S-1}]$$

where, $p$ represents the select bits for MUX1 and $q$ represents the select bits for MUX2.
In general, there are $M = {}^nC_2 = 2^{S-1}(2^S - 1)$ different pairs of simple XORROs, where n=$2^S$.
Let $T_0, T_1, .., T_{2^s-1}$ denote the time periods of all the XORROs.

Each combination of $p$ bits selects one XORRO and each combination of $q$ bits selects one XORRO. These two together are selected for the final decision of the output bit of the advanced XORRO PUF.

As the response of Advanced XORRO PUF depends on the frequency of the two XORROs selected by two MUXs we can introduce a term defined in terms of time periods of the selected XORROs: $\triangle = T^l - T^u$, where $T^u$ is the time period of XORRO selected by MUX1 and $T^l$ is the time period of XORRO selected by MUX2.
In general, we can see $\triangle$ as one of the possible time period differences:

$$\triangle = \triangle_0 | \triangle_1 | ... | \triangle_M \text{ ,where}$$
$$\triangle_0 = T_0 - T_1,$$
$$\triangle_1 = T_1 - T_2 \text{ ,}$$
$$...$$
$$\triangle_M = T_{n-1} - T_{n-2}$$

## 2.2 Deriving the ensemble model

The brute force approach is to create an ensemble model of all the linear models created for a given pair of XORRO PUFs selected by the multiplexers. But, we mathematically investigate the equations that are formed when we try to create a general expression that defines the difference in the frequency values of any two selected XORROs.
So, the difference in frequency is represented as $\triangle = F^u - F^l$ ,

where $F^u$ and $F^l$ represents the oscillating frequencies of the two selected XORROs. Now, at any given point of time, the output of a multiplexer is output of one of the XORROs and we describe the time period of oscillation of the multiplexer with select bits as [$p_0$, $p_1$,..., $p_{s-1}$]:

$$T^u = (1-p_0)(1-p_1)(..)(1-p_{S-1})T_0 + (1-p_0)(1-p_1)..(p_{S-1})T_1 + (1-p0)(1-p1)..(p_{S-2})(1-p_{S-1})T_2 + ... + p_0p_1...p_{S-2}(1-p_{S-1})T_{n-2} + p_0p_1...p_{S-2}p_{S-1}T_{n-1}$$

Now, if we assume that the number of select bits is 4, then, different values of the select bits will evaluate the above expression of $T^u$ to one of the time periods of the XORROs of the advanced PUF. We demonstrate it as follows:
p=[0,0,..,0] implies $T^u = T_0$,
p=[0,0,..,1] implies $T^u = T_1$,
p=[0,0,..,1,0] implies $T^u = T_2$,
...
p=[1,1,..,1,0] implies $T^u = T_{n-2}$,
p=[1,1,..,1] implies $T^u = T_{n-1}$

Similarly , for another multiplexer with select bits as [$q_0$, $q_1$,..., $q_{s-1}$], the time period is

$$T^l = (1-q_0)(1-q_1)(..)(1-q_{S-1})T_0 + (1-q_0)(1-q_1)..(q_{S-1})T_1 + (1-q_0)(1-q_1)..(q_{S-2})(1-q_{S-1})T_2 + ... + q_0q_1...q_{S-2}(1-q_{S-1})T_{n-2} + q_0q_1...q_{S-2}q_{S-1}T_{n-1}$$

Now, we can transform the expression $\triangle = F^u - F^l$ to $\triangle = T^l - T^u$ (frequency is inversely proportional to time period).

$$\triangle = ((1-p_0)(1-p_1)(..)(1-p_{S-1}) - (1-q_0)(1-q_1)(..)(1-q_{S-1}))T_0 + ((1-p_0)(1-p_1)..(p_{S-1}) - (1-q_0)(1-q_1)..(q_{S-1}))T_1 + ((1-p0)(1-p1)..(p_{S-2})(1-p_{S-1}) - (1-q_0)(1-q_1)..(q_{S-2})(1-q_{S-1}))T_2 + ... + (p_0p_1...p_{S-2}(1-p_{S-1}) - q_0q_1...q_{S-2}(1-q_{S-1}))T_{n-2} + (p_0p_1...p_{S-2}p_{S-1} - q_0q_1...q_{S-2}q_{S-1})T_{n-1}$$

We can write it as

$$\triangle = k_0T_0 + k_1T_1 + .. + k_{n-1}T_{n-1}$$

where $k_i$ is a general expression to represent the difference in output values of the AND operation of the select bits of both multiplexers and $n = 2^S - 1$.
We saw before that $T_i$ can be written as $T = x^Ta + b$ (in section 1.2) and we transform the above expression into:

$$\triangle = \sum_{i=0}^{2^S-1} k_i(X_i^T a + b_i)$$

where,

$X_i = [x_0^i, x_1^i, ..x_{R-1}^i]$ (Refer to section 1.3)

On simplifying further,

$$\triangle = \sum_{i=0}^{2^S-1} U_i^T V_i$$

where $U_i = [a_0k_i, a_1k_i, .., a_{R-1}k_i, k_i]$, $V_i = [x_0^i, x_1^i, ..x_{R-1}^i, b_i]$

$\triangle$ represents the ensemble linear model to crack the advanced XORRO PUF and the response is given by the expression: $\frac{1+sign(\triangle)}{2}$

Total n=$2^S$ linear models are required to crack the Advanced XORRO PUF.

## 2.3 Finding features for Advanced XORRO

$\triangle = U_0^T V_0 + U_1^T V_1 + ... + U_{n-1}^T V_{n-1}$

$\triangle = U^T V$

where U=$[a_0k_0, a_1k_0, .., a_{R-1}k_0, k_0, a_0k_2, a_1k_2, .., a_{R-1}k_2, k_2, ....., a_0k_{n-1}, a_1k_{n-1}, .., a_{R-1}k_{n-1}, k_{n-1}]$,

V=$[x_0^0, x_1^0, ..x_{R-1}^0, b_0, x_0^1, x_1^1, ..x_{R-1}^1, b_1, ..., x_0^{n-1}, x_1^{n-1}, ..x_{R-1}^{n-1}, b_{n-1}]$

$\triangle = U^T V$ is linear model to crack Advanced XORRO where U are features and V are weights.

# 3 Reporting outcomes for various hyperparameters

The best accuracy that we obtain for our ensemble LinearSVC model is 0.9917 at hyperparamter values C = 10.0 and loss = 'squared hinge'. Below, we report the variations in accuracy and training time of the model on changing different hyperparameter values.

## 3.1 (a) Loss hyperparameter on LinearSVC

| Loss | Test Accuracy | Training time |
|---|---|---|
| loss = 'hinge' | 0.9875 | 1.61 sec |
| loss = 'squared hinge' | 0.9907 | 2.59 sec |

## 3.2 (b) C hyperparameter on LinearSVC and Logistic Regression

### 3.2.1 Effect of C on LinearSVC

| C Value | Test Accuracy | Training time |
|---|---|---|
| C = 1.0 (low) | 0.9906 | 0.89 sec |
| C = 10.0 (medium) | 0.9922 | 2.59 sec |
| C = 100.0 (high) | 0.9906 | 2.15 sec |

### 3.2.2 Effect of C on Logistic Regression

| C Value | Test Accuracy | Training time |
|---|---|---|
| C = 1.0 (low) | 0.9875 | 4.11 sec |
| C = 10.0 (medium) | 0.9915 | 11.11 sec |
| C = 100.0 (high) | 0.9926 | 11.35 sec |

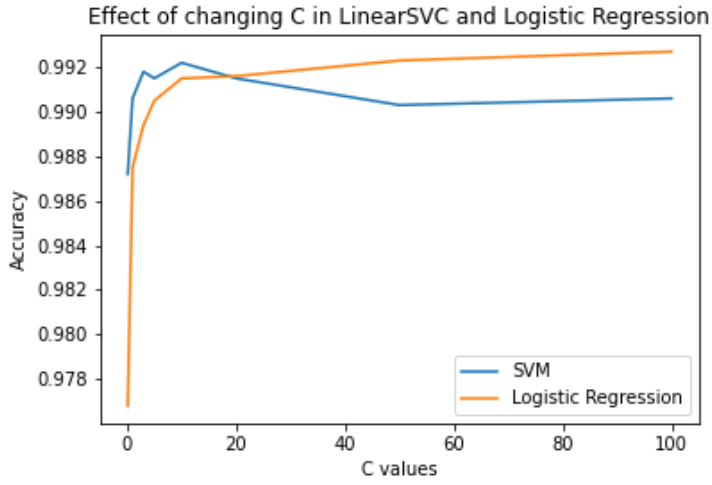


Figure 2: Effect of changing C in LinearSVC and Logistic Regression

As evident in Figure 2, the test accuracy for LinearSVC model peaks at around C = 10, and falls after that for higher values of C. For Logistic Regression, the model continues to perform better with increasing values for C.

## 3.3 (c) tol hyperparameter on LinearSVC and Logistic Regression

### 3.3.1 Effect of tol parameter on LinearSVC

| tol Value | Test Accuracy | Training time |
|---|---|---|
| tol = 0.000001 (low) | 0.9907 | 3.85 sec |
| tol = 0.0001 (medium) | 0.9907 | 2.51 sec |
| tol = 0.1 (high) | 0.9906 | 2.75 sec |

### 3.3.2 Effect of tol parameter on Logistic Regression

| tol Value | Test Accuracy | Training time |
|---|---|---|
| tol = 0.000001 (low) | 0.9875 | 11.07 sec |
| tol = 0.0001 (medium) | 0.9875 | 9.44 sec |
| tol = 0.1 (high) | 0.9875 | 13.31 sec |

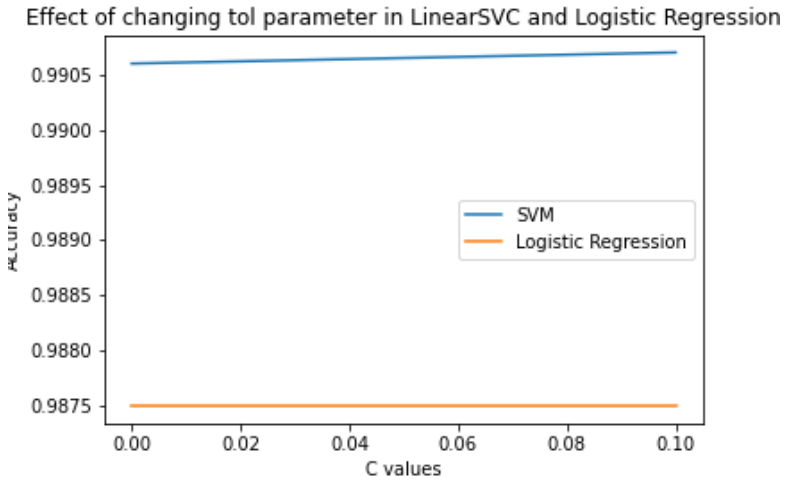


Figure 3: Effect of changing tol parameter on LinearSVC and Logistic Regression

As evident in Figure 3, the tolerance parameter has less effect on the performance of the model. The accuracy of the LinearSVC model is constant across all values of tol parameters. For logistic regression, the accuracy improves slightly as tol parameter is increased to a high value of 1.0