Non-interactive inner products in privacy-preserving biometric identification using secure multiparty computation

# Introduction

Inner products are fundamental operation in privacy-preserving applications. In addition to this report along the code implementation project, which focus on LWE implementation based on (Couteau & Zarezadeh, Non-Interactive Secure Computation of Inner-Product from LPN and LWE, 2023) my submission for the "Secure and Private Computing" 3 ECTS course at HSG also includes a presentation about biometric identification, a presentation about the status of the code and various reviews on other papers that were presented by peers.

## Non-Interactive Key Exchange

When two parties decide to communicate among each other, they first must compute shared private key K. Let’s make an example how two parties (Alice and Bob) can get a shared private key K:

The great thing about this interaction pattern is that it avoids n2 pairwise key exchange among all the parties. Small n refers to the number of parties being involved. So, we have just n public keys. Please refer to Figure 1 for clarification.

## Non-interactive secure computing

It would be nice if there exist a pattern to securely exchange encodings that works between n parties. Parties are broadcasting an encoding of their input:

Two parties can than compute and from their own state plus the other party’s encoding.

This interaction pattern avoids as above n2 pairwise key exchange among all the parties. Please refer again to Figure 1 for clarification.

## Non-interactive multi party computation (MPC)

The goal of the implementation is to do non-interactive MPC for shares of inner products. So instead of encoding some xi or xj, shares of are calculated.

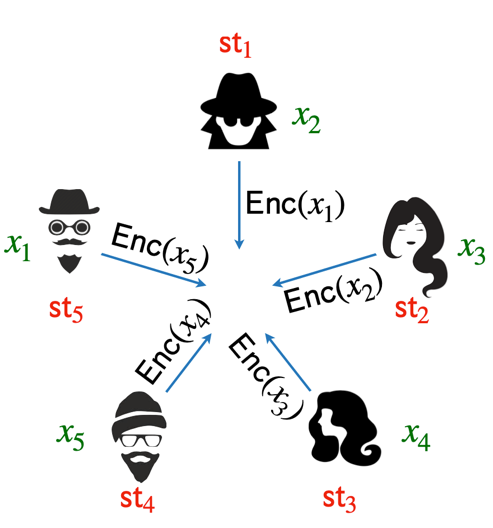
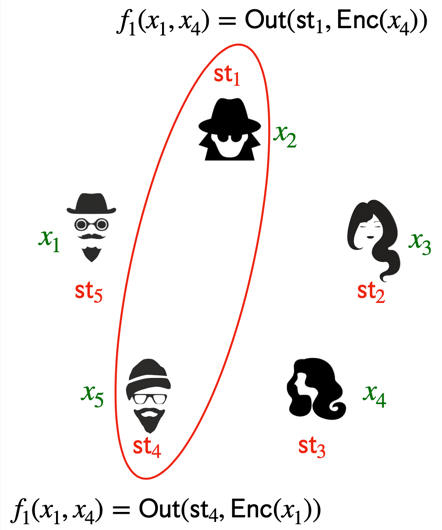
# Protocol Description

Here some questions and answers to understand the protocol which will be implemented. Please note that all pictures in this chapter are from slides from (Couteau, Geoffroycouteau, 2022) and the explanations are founded on (Couteau, Youtube, 2022).

1. What does the protocol do?

The protocol is used to exchange shares of vectors among different parties. Assuming there are two parties, Bob and Alice, each holding a vector, x1 and x4. Then this protocol can be used to securely exchange shares of the locally computed vector encoding.

Figure 1: Non-Interactive-Key Exchange and non-interactive secure computing



1. How does the protocol work?

The protocol's exact mechanism is quite intricate. It is based on Alekhnovich’s key exchange (AKE), which inherently involves an inner product calculation. Can this framework embed an inner product computation between x1 and x4 within AKE? Absolutely! The protocol operates by performing an inner product computation along with some noise term e. Essentially, it enables a secure inner product computation layered on top of AKE, which itself relies on inner product computation. Please refer to Figure 2.

Additionally, it supports the homomorphic property. This means that encrypting the product of two elements is feasible when one element is in plaintext, and the other is encrypted. For instance, consider the following equation where x is in plaintext and n is an integer value.

A diagram of a complex equation

Description automatically generated with medium confidence

Figure 2: Alekhnovich's key exchange (AKE)

The correctness of the protocol can be demonstrated by examining AKE. Specifically, when t2/n << 1 (where t denotes the Hamming weight and n the dimension of a vector), it holds true. In MPC, correctness errors can lead to information leakage when a "detectable" error occurs, though the system remains secure albeit with some loss. However, using LWE makes the error negligible, which is why LWE is preferred. Please refer to Figure 3. It's important to note that **security** is guaranteed only with appropriate parameters. Due to the challenges and complexities in determining the correct parameters for security, the focus was placed on **correctness**. Selecting LWE parameters is a nontrivial task, as there isn't a straightforward formula for it. A script is needed to estimate runtime against state-of-the-art cryptoanalysis.

A screenshot of a computer

Description automatically generated

Figure 3: Learning with Error (LWE)

1. Where can such a protocol be used?

It can for example be used to exchange securely fingerprint, iris scans, any other biometric data or in general secured data. Another example would be authentication from a user to a server as can be depicted from Figure 4.

A black and white symbols

Description automatically generated with medium confidence

Figure 4: Usage of protocol

1. How can it be used for biometric identification?

By combining LWE with MPC, biometric identification can be performed in a way that respects the privacy and security of the individual's sensitive biometric data, while still allowing for accurate and secure comparisons. The protocol involves computing the inner product between the stored encrypted biometric data and the newly provided biometric data (also encrypted). The LWE scheme's homomorphic properties facilitate this computation without revealing the actual data. By relying on cosine similarity for example as stated in (Ernst & Mitrokotsa, 2023) on page 18, biometric matching can easily be computed as an inner product.

1. What advantage does the protocol provide (over other protocols) for biometric identification?

It provides post quantum security. It is very efficient. The error tolerance makes the protocol more robust. Ahead of time, each party can publish the shares of inner product. That makes the protocol more efficient.

# Implementation Overview

Table 1: Parameter description for implementation

|  |  |
| --- | --- |
| Security parameter |  |
| Upper case letters denote matrices |  |
| Bold lower-case letters denote row vectors |  |
| Transpose on row vectors denote column vectors |  |
| Horizontal concatenation of horizontal vectors |  |
| Vertical concatenation of vertical vectors |  |
| x is uniformly sampled from the set X |  |
| Finite set S uniform distribution is denoted by |  |
| Bernoulli distribution with parameter |  |

In Table 1 the parameters are described. Please note: means that the random variable evaluates to 1 with probability and to 0 with probability . More generally, denotes the distribution that outputs a uniformly random element of with probability , and 0 otherwise.

According to (Couteau & Zarezadeh, Non-Interactive Secure Computation of Inner-Product from LPN and LWE, 2023) in section 3, there exists a “general non-interactive protocol for securely computing the inner product between two vectors over , with correctness error (independent of the value of the inputs).”The great thing is that the protocol can either be instantiated under LPN (learning parity with noise, the error will be noticeable but arbitrary small) or under the LWE (learning with errors, the error can be made negligible) assumption. As described above and depicted in Figure 3 focus will be on LWE.

But what data is been used? The tutors provided two arrays (.npy) with shape 13’232x128 (data) and 13’232x1 (labels). After having quickly looked at the arrays and with some research done, it turns out that the data must be LFW (labeled faces in the wild). The LFW dataset contains 13,233 images of faces collected from the web. This dataset consists of the 5749 identities with 1680 people with two or more images.[[1]](#footnote-1)

## Implementation in Detail

Implementation is done according to (Couteau & Zarezadeh, Non-Interactive Secure Computation of Inner-Product from LPN and LWE, 2023) in section 3.3 (non-interactive inner product from LWE) as described in Figure 5. A white background with black text

Description automatically generated

Figure 5: LWE-based non-interactive inner-product

So the first step is to implement the setup method. The output will be .

Table 2: Parameter values

|  |  |  |
| --- | --- | --- |
| Description | Parameter | Value |
| Generator matrix of linear code |  | 512x384 |
| Dimension, eg. row dimension of u |  | 128 |
| Dimension |  | 4n |
| Dimension |  | 2n |
| Dimension |  | 2^15 |
| Dimension |  | 2^14 |

Setup is normally done by an trusted party but here it is done in directly in the code. Dimension of is . Actually it is an matrix with the dimension 512x384 as listed in Table 2

In two further steps, encode and decode are implemented. Plese refer directly to the code. Please note that some matrices had to be padded with zeros or random values according to the describtion of the algorithms in Figure 5.

### Validation 1

So the next step will be to apply the encode and decode algorithms accordingly.

A math equations and formulas

Description automatically generated with medium confidence

Figure 6: Application of Encode and Decode

Alright lets say we have two parties, X and Y. Now calucaltion of pkx, skx, pky and sky is performed with Encode function.

We want to validate if . If so, we are done! Please refer to Figure 6

### Validation 2

Another possible way to validate according to (Couteau & Zarezadeh, Non-Interactive Secure Computation of Inner-Product from LPN and LWE, 2023) from page 9 as depicted in Figure 7.

A black text with black letters

Description automatically generated with medium confidence

Figure 7: Validation as in LPN (Learning Parity with Noise)

## Further Reading

Please note that the distances are either Euclidean or Hamming distance calculations. According to (Couteau & Zarezadeh, Non-Interactive Secure Computation of Inner-Product from LPN and LWE, 2023) computing the Hamming distance between two vectors x and y is where HW denotes the Hamming weight. Please also note that there is the possibility to use cosine similarity for biometric matching as described in (Ernst & Mitrokotsa, 2023) and already mentioned above.

1. https://paperswithcode.com/dataset/lfw [↑](#footnote-ref-1)