SHIELD: Scalable Homomorphic Implementation of Encrypted Data-Classifiers

Alhassan Khedr, *Member, IEEE,* Glenn Gulak, *Senior Member, IEEE,* and Vinod Vaikuntanathan

Abstract—Homomorphic encryption (HE) systems enable computations on encrypted data, without decrypting and without knowledge of the secret key. In this work, we describe an optimized Ring Learning With Errors (RLWE) based implementation of a variant of the HE system recently proposed by Gentry, Sahai and Waters (GSW). Although this system was widely believed to be less efficient than its contemporaries, we demonstrate quite the opposite behavior for a large class of applications.

We first highlight and carefully exploit the algebraic features of the system to achieve significant speedup over the state-of-the-art HE implementation, namely the IBM homomorphic encryption library (HElib). We introduce several optimizations on top of our HE implementation, and use the resulting scheme to construct a homomorphic Bayesian spam filter, secure multiple keyword search, and a homomorphic evaluator for binary decision trees.

Our results show a factor of $10\times$ improvement in performance (under the same security settings and CPU platforms) compared to IBM HElib for these applications. Our system is built to be easily portable to GPUs (unlike IBM HElib) which results in an additional speedup of up to a factor of $103.5\times$ to offer an overall speedup of $1035\times$.

Index Terms—Homomorphic Encryption, FHE, Ring LWE, Bayesian Filter, Secure Search, Decision Trees, Implementation, GPU.

---◆---

1 INTRODUCTION

A fully homomorphic encryption scheme (FHE) is an encryption scheme that allows evaluation of arbitrary functions on encrypted data. Starting with Gentry's mathematical breakthrough constructing the first plausible FHE scheme [1], we have seen rapid development in the theory and implementation of homomorphic encryption (HE) schemes. HE schemes can now be based on a variety of cryptographic assumptions – approximate greatest common divisors [2], [3], learning with errors (LWE) [4], [5], [6], [7], and Ring-LWE (RLWE) [8], [9], [10].

Due to the growing use of cloud computing, privacy concerns have begun to escalate. Secure data classifiers constructed from FHE schemes can present a very useful tool to provide an answer for these concerns. It is important that these secure data classifiers protect the privacy of the input data and also the classifier model itself.

Consider the secure email spam filter as a representative example for data-classifier class of applications. In the training phase, the spam filter creates its model from plaintext emails. In the testing phase however, the secure email spam filter is required to classify the input encrypted email as a spam or not without any knowledge about the actual contents of the email. The same concept can be applied to the secure search application, which is basically a variant of the spam filter. In the secure search problem, it is also required to search for multiple encrypted keywords inside files and return the number of matching keywords without the knowledge of the file contents.

The main contribution of this paper is the construction of a secure email spam filter and a secure multi-keyword search. We have also prototyped a secure binary decision trees as a proof of concept and compare our work to other similar implementations [11]. Detailed algorithms for building those applications, using the GSW HE scheme, are presented in this work along with their GPU performance results. Our data-classifier takes the idea of Private Information Retrieval (PIR) one step forward, and homomorphically computes on the encrypted data retrieved by the PIR to obtain useful pieces of information. We show that a careful consideration of the mathematics underlying the recently proposed GSW scheme and Brakerski and Vaikuntanathan (BV) scheme [7], together with the features of the specific applications, results in significant speedups. The performance of the data-classifier as well as the secure search engine depends on the size of the database/file and can be as low as a few seconds with present off-the-shelf technology.

The applications presented in this paper take advantage of the observation about a key feature in the GSW encryption scheme, namely, when multiplying a sequence of encrypted numbers using the GSW scheme, if the final result happens to be a zero, then the error level is reduced to the error level of a fresh cyphertext. Of course, because of the security of the scheme, there is no way for the homomorphic evaluator to tell if and when this happens, but if such an event is guaranteed to happen often during homomorphic evaluation, we are guaranteed to have small

- *Alhassan Khedr is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada.*
  E-mail: alhassan@ece.utoronto.ca
- *Glenn Gulak is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada.*
  E-mail: gulak@ece.utoronto.ca
- *Vinod Vaikuntanathan is with the Department of Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA, USA.*
  E-mail: vinodv@mit.edu

error growth. As an illustrative example of this feature, consider evaluating an expression of the form

$$F(x_1, \ldots, x_v) = \sum_{(y_1, \ldots, y_v) \in S} \left(\prod_{i=1}^{v} \overline{(x_i \oplus y_i)} \right) \quad (1)$$

where $x_1, \ldots, x_v$ are v-tuples of input encrypted bits, $y_1, \ldots, y_v$ are v-tuples of bits in some set S, operation $\overline{(x_i \oplus y_i)}$ represents binary XNOR between bits x_i and y_i. Since the form of the expression in (1) guarantees us that exactly one of the terms may survive ($F = 1$ when $x_1, \ldots, x_v \in S$, otherwise $F = 0$), the homomorphic evaluator is guaranteed to have a small total error growth (even though the evaluator does not know precisely which term will survive). It can be noticed that (1) is identical to the server's computation in a PIR scheme. Indeed, the same format appears in several applications, including the ones we discuss in Section 5.

In this work, algorithmic optimizations were also introduced to the work in [6], [7] to reduce their computational complexity as will be discussed in Sections 3.1 and 3.1.1. We also introduce the notion of *decrypting a flag submerged in noise* described in Section 5.1 which can be used to argue that a decryption error in fact gives us a meaningful bit of information! This is unlike all other lattice-based HE schemes that we are aware of where one gives up hope the moment the error exceeds a certain threshold.

1.0.0.1 Implementation Results: We carefully exploit the parallelism in the encryption system by implementing it on a GPU platform. The number theoretic transform (NTT) engine was carefully designed to address the limitations present in current GPUs. Our GPU implementation of the HE scheme scores a ciphertext (Ctxt) multiplication run time of 3.477 milliseconds with a $1035\times$ speedup over IBM HElib for circuits depth more than 5 and 80-bit security. Also for the same security level, our ciphertext size is smaller than HElib by a factor of $1.5\times$. Our improvement is realized through a reduction in parameters for the same security level and homomorphic capacity, stemming from our observations about noise growth. This ultimately leads to faster implementation and smaller ciphertext sizes. Our data-classifiers are completely scalable and their running time can be reduced proportional to the number of GPU cores utilized.

The rest of the paper is organized as follows. Section 2 presents related work. In Section 3 we introduce the improved encryption scheme. Next in Section 4, the NTT architecture and Solinas primes used in our system are introduced. The encrypted data-classifier design, secure multiple keyword search engine and encrypted binary decision trees are introduced in Section 5. Performance results are introduced in Section 6. Finally we conclude in Section 7.

2 RELATED WORK

Previous constructions of RLWE-based FHE schemes include [4], [5], [10]. One of the drawbacks of these schemes is the need to maintain a so-called "modulus chain" which increases the size of the prime number and consequently increases the ring dimension for the same security level. They also need to perform expensive modulus and key switching operations. Bootstrapping was introduced to the GSW scheme in [12], [13], with the bootstrapping in [13] completing in under a second.

Based on [4], Halevi and Shoup designed a homomorphic encryption library "IBM HElib" [10], [14], but due to the need of some additional large data structures and functions, the performance of their library has large execution time. A performance comparison between our library and the IBM HElib library is presented in Section 6. In [9] they implemented a variant of the RLWE FHE scheme. Our results also show considerable speedups over their implementation. Another homomorphic library was developed by Rohloff, Cousins, and Peikert [15]. In their paper they implement primary building blocks in hardware to accelerate their system. There are no performance results available yet publicly to compare our library with theirs. In [16] significant speedups were gained from the use of GPU computing, but their implementation suffered from having very large memory requirements, which eventually becomes the bottleneck of their implementation. Other implementation attempts were made but they were either incomplete implementations of an HE scheme capable only of performing one multiplication operation [17], or based on other cryptographic assumptions [18], [19], [20].

Applications analyzed in this paper were primarily inspired from [9], [21]. We extend their ideas and developed full algorithms. The work on CryptDB [22] used a combination of very simple HE schemes to implement a subset of encrypted SQL queries, and the work on "ML Confidential" [23] implemented simple classification tasks on encrypted data. Secure computations on binary decision trees were introduced in [11], yet they have long execution time due to the use of the IBM HElib library. Searching an encrypted database was previeously addressed by [24], [25], [26], [27]. One drawback in [24], [25] was the need for a special key to aid the server in performing the search request. They achieve a weaker security notion, namely one where partial information about the data access pattern is leaked. In particular, in the work of [26], the same server requests would generate the same tags. For a simple and general overview of homomorphic encryption concepts the reader is encouraged to read [28], [29], [30].

3 THE ENCRYPTION SYSTEM

3.0.0.1 Notation: For an odd prime number q we identify the ring $\mathbb{Z}/q\mathbb{Z}$ (or $\mathbb{Z}_q$) with the interval $(-q/2, q/2) \cap \mathbb{Z}$. The notation $[x]_q$ denotes reducing x modulo q. Our implementation uses polynomial rings defined by the cyclotomic polynomials $R = \mathbb{Z}[X]/\Phi_m(X)$, where $\Phi_m(X) = x^n + 1$ is the irreducible mth cyclotomic polynomial, in which n is a power of 2 and $m = 2n$. We let $R_q = R/qR$. Any type of multiplication including matrix and polynomial multiplication is denoted by the multiplication operator '·'. Rounding up to the nearest

integer is denoted by $\lceil a \rceil$. Matrices of rings are defined as $A_{M \times N}$, where $A_{ij} \in R_q$ and M, N are the matrix dimensions. $I_{N \times N}$ represents the identity matrix of rings. Row vectors are represented as $[a \; b]$, where a and b are the vector elements. Column vectors on the other hand are represented as $[a \; ; \; b]$.

3.0.0.2 *Ring Learning With Errors:* The ring learning with errors problem (RLWE) was introduced in [31]. It is the mapping of the LWE problem from the vectors over $\mathbb{Z}_q$ to polynomial rings over R_q. The RLWE problem is to distinguish between the following two distributions. The first distribution is to draw (a, b) uniformly from R_q^2. The second is to first draw t uniformly from R_q. Then sample (a, b) as follows. Draw a uniformly from R_q, sample e from a discrete Gaussian error distribution $e \leftarrow D_{R_q,\sigma}$, and set $b = a \cdot t + e$.

3.1 The Encryption Scheme

The parameters of the system are n, the degree of the number field; q, the modulus; σ_k and σ_c, the standard deviation of the discrete Gaussian error distribution in the keyspace and ciphertext space, respectively; $\ell \triangleq \lceil \log q \rceil$; and $N = 2\ell$ that governs the number of ring elements in a ciphertext. The setting of these parameters depends on the security level λ (e.g., $\lambda = 80$ or 128 bits) as well as the complexity of functions we expect to evaluate on ciphertexts.

3.1.0.1 *Bit Decompose Function "BD()":* The bit decompose function $BD(integer)$ takes an ℓ-bit input integer, then outputs a row vector with size ℓ containing the bit decomposition of this integer. Similarly, $BD(polynomial)$ takes an input polynomial of size n, where each coefficient is an ℓ-bit integer, then outputs an ℓ-sized row vector of polynomials (each of size n) containing the bit decomposition of each coefficient of the input polynomial, yielding a matrix of size $l \times n$. Finally, $BD(Matrix\ of\ polynomials)$ takes an input matrix of polynomials of size $x \times y$ (each polynomial is of size n with integer coefficients), then outputs a matrix of polynomials expanded by a factor ℓ in the column dimension, yielding a matrix of size $x \times y\ell$, where each consecutive ℓ elements along the row contain the bit representation of each coefficient of each of the input polynomials. For example, the bit decompose of the input polynomial matrix $B_{x \times y \times n}$ is $BD(B_{x \times y \times n}) = \beta_{x \times y\ell \times n}$. The reader should note that despite the fact that the polynomial coefficients of matrix $\beta_{x \times y\ell \times n}$ are single bit values, the storage requirement of matrix β in CPU or GPU memory is not equal to $x \times y\ell \times n$ bits. This is due to the fact that the smallest addressable unit of memory is a byte (i.e., Byte Addressable). Hence, β requires $x \times y\ell \times n$ *bytes* of storage. This results in the further observation that the storage requirement of $\beta_{x \times y\ell \times n}$ is $8\times$ the storage requirement of $B_{x \times y \times n}$.

3.1.0.2 *Bit Decompose Inverse Function "BDI()":* As the name reveals, the BDI() function is the inverse of BD(). The BDI() function groups consecutive ℓ coefficients along a row (the coefficients don't need to be binary),

and outputs the integer corresponding to those ℓ bits. Mathematically, the BDI() function can be defined as multiplying the *expanded* matrix of polynomials $\beta_{x \times y\ell}$ from the right by the matrix $\alpha_{y\ell \times y}$ defined in (2) (polynomial dimension n will be omitted from this point forward for clarity). Hence $B_{x \times y} = BDI(\beta_{x \times y\ell}) = \beta_{x \times y\ell} \cdot \alpha_{y\ell \times y}$.

$$
\alpha_{y\ell \times y} = \begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
2 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
2^\ell & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
0 & 2 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 2^\ell & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 \\
0 & 0 & 0 & \cdots & 2 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 2^\ell
\end{pmatrix}
\tag{2}
$$

In this work, we introduced some algorithmic optimizations to the encryption system in [6], [7] in order to reduce computational complexity and to speedup our operations, as will be detailed below. Our encryption system works as follows.

- Keygen(1^λ): Choose polynomial $t \leftarrow D_{R_q,\sigma_k}$. The secret key $sk = s_{2\times 1} \leftarrow [1; -t] \in R_q^2$. Uniformly sample $a \leftarrow R_q$, $e \leftarrow D_{R_q,\sigma_k}$, set $b = a \cdot t + e$, The public key $pk = A_{1\times 2} = [b \; a]$. Note that

$$A_{1\times 2} \cdot s_{2\times 1} = b - a \cdot t = e \tag{3}$$

where "$\cdot$" is the inner product over the ring R_q.

- Enc(pk, μ): The message space of our encryption scheme is R_q. $r_{N \times 1}$ is a matrix that consists of N polynomials with random coefficients in the range $\{0, 1\}$, $E_{N \times 2} \leftarrow D_{R_q^{N \times 2}, \sigma_c}$, encrypt the plain text polynomial $\mu \in R_q$ by calculating

$$C_{N \times 2} = \mu \cdot BDI(I_{N \times N}) + r_{N \times 1} \cdot A_{1 \times 2} + E_{N \times 2} \tag{4}$$

(as opposed to $C_{N \times N}$ in [6], [7], we have a smaller ciphertext by a factor of $8\times$ as discussed previously in the BD() definition.)

- Dec(sk, C): Given the ciphertext C, the plaintext $\mu \in R_q$ is restored by multiplying C by the secret-key s as follows :

$$
\begin{aligned}
C_{N \times 2} \cdot s_{2 \times 1} &= (\mu \cdot BDI(I_{N \times N}) \\
&\quad + r_{N \times 1} \cdot A_{1 \times 2} + E_{N \times 2}) \cdot s_{2 \times 1} \\
&= \mu \cdot BDI(I_{N \times N}) \cdot s_{2 \times 1} \\
&\quad + r_{N \times 1} \cdot A_{1 \times 2} \cdot s_{2 \times 1} + E_{N \times 2} \cdot s_{2 \times 1} \\
&= \mu \cdot BDI(I_{N \times N}) \cdot s_{2 \times 1} \\
&\quad + r_{N \times 1} \cdot e + E_{N \times 2} \cdot s_{2 \times 1} \\
&= \mu \cdot BDI(I_{N \times N}) \cdot s_{2 \times 1} + error
\end{aligned}
\tag{5}
$$

(as opposed to $\mathsf{Dec}(\mathsf{sk}, C) = C_{N \times N} \cdot v_{N \times 1}$ in [6], [7], we have fewer operations in Dec by a factor of ℓ times)

Observe that the first ℓ coefficients in the first term of the last equation in (5) are in the form $\mu, 2\mu, \cdots, 2^{\ell-1}\mu$. This means that the element at location $i \in [0, \ell - 1]$ is in the form $\mu \cdot 2^i + error$. That is, the most significant bit of each entry carries a single bit from the number μ assuming that $error < q/2$ and there is no wrap-around mod q as was described in [6].

3.1.1 Homomorphic Operations

For input ciphertexts $C_{N \times 2}$ and $D_{N \times 2} \in R_q^{N \times 2}$ encrypting μ_1 and μ_2 respectively, homomorphic operations are defined as follows.

- ADD(C, D): To add the two ciphertexts $C_{N \times 2}$ and $D_{N \times 2}$, simply output $C_{N \times 2} + D_{N \times 2}$, which is an entry-wise addition.

- MULT(C, D): To multiply the two ciphertexts $C_{N \times 2}$ and $D_{N \times 2}$, output $\mathsf{BD}(C_{N \times 2}) \cdot D_{N \times 2}$.

   (as opposed to $\mathsf{MULT}(C, D) = \mathsf{FLATTEN}(C_{N \times N} \cdot D_{N \times N})$ in [6], [7], where $\mathsf{FLATTEN}(A)$ is defined as $\mathsf{BD}(\mathsf{BDI}(A))$. We have fewer operations in MULT by a factor of at least ℓ times)

Correctness of homomorphic addition is immediate, however it is not that obvious for the homomorphic multiplication. It is clear that the multiplication algorithm is asymmetric in the input ciphertexts C and D. That is, we treat the components of D as a whole, whereas the components of C are broken up into their "bit-wise decompositions". This is a "feature" that is inherited from the work of BV [7]. It is shown below that this multiplication method is correct and gives a slow noise-growth rate.

The correctness of the multiplication operation can be noticed from the decryption operation in (6). Matrix dimensions are removed for clarity.

$$
\begin{aligned}
\mathsf{BD}(C) \cdot D \cdot s &= \mathsf{BD}(C) \cdot (\mu_2 \cdot \mathsf{BDI}(I) + r_2 \cdot A + E_2) \cdot s \\
&= \mathsf{BD}(C) \cdot (\mu_2 \cdot \mathsf{BDI}(I) \cdot s + r_2 \cdot e + E_2 \cdot s) \\
&= \mu_2 \cdot C \cdot s + \mathsf{BD}(C) \cdot (r_2 \cdot e + E_2 \cdot s) \\
&= \mu_2 \cdot (\mu_1 \cdot \mathsf{BDI}(I) \cdot s + r_1 \cdot e + E_1 \cdot s) \\
&\quad + \mathsf{BD}(C) \cdot (r_2 \cdot e + E_2 \cdot s) \\
&= \mu_2 \cdot \mu_1 \cdot \mathsf{BDI}(I) \cdot s + \mu_2 \cdot (r_1 \cdot e + E_1 \cdot s) \\
&\quad + \mathsf{BD}(C) \cdot (r_2 \cdot e + E_2 \cdot s) \\
&= \mu_2 \cdot \mu_1 \cdot \mathsf{BDI}(I) \cdot s + \mu_2 \cdot error_1 \\
&\quad + \mathsf{BD}(C) \cdot error_2 \\
&= \mu_2 \cdot \mu_1 \cdot \mathsf{BDI}(I) \cdot s + error
\end{aligned}
$$
(6)

It can be noticed from the last line in (6) that it is the encryption of $\mu = \mu_2 \cdot \mu_1$. Note that $\mathsf{BD}(C_{N \times 2}) \cdot \mathsf{BDI}(I_{N \times N}) = I_{N \times N} \cdot C_{N \times 2} = C$.

Function 1: Multiply "v" Ciphertexts Function

Input: "v", Ctxts: $C_1, C_2, \cdots, C_v$
Output: C_{accum}
     The multiplication result of "v" input Ctxts.

$C_{accum} = C_1$
For i from 2 to v {
     $C_{accum} = C_{accum} \times C_i$
}
Return C_{accum}

3.1.1.1 Noise Analysis: Correct decryption depends crucially on the ciphertext noise being bounded. Thus, it is important to understand how homomorphic operations increase ciphertext noise. Let C be a fresh ciphertext. We make the following observations, after [7].

- Homomorphic addition of v ciphertexts increases the noise by a factor of v in the worst case. In practice, since the coefficients of the error polynomials follow a Gaussian distribution, the factor is closer to $O(\sqrt{v})$.

- Homomorphic multiplication is significantly more interesting. Multiplication of two ciphertexts $C = \mathsf{Enc}(\mu_1)$ and $D = \mathsf{Enc}(\mu_2)$ with error magnitudes B_1 and B_2, respectively, increases the error to $O(B_1 \cdot \|\mu_2\|_1 + B_2 \cdot n \log q)$ in the worst case, and $O(B_1 \cdot \|\mu_2\|_1 + B_2 \cdot \sqrt{n \log q})$ in practice. Here, $\|\mu\|_1$ denotes the ℓ_1 norm of the message polynomial μ. *The key fact to note here is that the error dependence on the two ciphertexts is asymmetric.*

3.1.1.2 Better Error in Homomorphic Multiplication: To multiply v ciphertexts it is crucial to pay attention to the order of multiplication. In the applications presented in this work, input μ will typically be 0 or 1, meaning that the growth is simply additive with respect to B_1. Thus, the best way to multiply v ciphertexts with (the same) error level B is through an accumulator-like algorithm as in Function 1, rather than using a binary tree of multiplications (which grows the error at superpolynomial rates). The resulting error growth is $O(B \cdot vn \log q)$ in the worst case, and $O(B \cdot \sqrt{vn \log q})$ in practice.

For example consider (1), the noise grows to $O(B \cdot vn \log q \cdot |S|)$ in the worst case, or $O(B \cdot \sqrt{vn \log q |S|})$ in the typical case. This is in contrast to $O(B \cdot \sqrt{(n \log q)^{\log v} |S|})$ when using the Brakerski-Gentry-Vaikuntanathan [4] encryption scheme, implemented in IBM HElib. Indeed, such expressions, as in (1), are far from atypical – they occur quite naturally in evaluating decision trees and PIR-like functions as will be discussed in Section 5.

3.1.1.3 Zero Plaintext, Zero Error: Yet another source of improvement is evident when we inspect the error term $B_1 \cdot \|\mu_2\|_1 + B_2 \cdot n \log q$. When we multiply using an accumulator as in Function 1, B_2 represents the smaller error in the fresh ciphertexts C_i, and B_1 represents the larger error in the accumulated ciphertext C_{accum}. We see that if C_i encrypts $\mu_2 = 0$, then the larger error term B_1 vanishes in the error expression!

Table 1
Parameter Selection and Keys/Ctxt Sizes.

Parameter	RLWE (This work)
λ	80
n	1024
ℓ	31
N	$2 \cdot \ell = 62$
σ_k, σ_c	10
SK size	$2 \times n \times \ell = 7.936\ KBytes$
PK size	$2 \times n \times \ell = 7.936\ KBytes$
Ctxt size	$N \times 2 \times n \times \ell = 492.032\ KBytes$

This phenomenon manifests itself in evaluating the expression in (1) as well. When evaluating each of the products in (1), the error grows proportional not to v, the total number of multiplications, but rather with k, the longest continuous chain of 1's starting from the end. This is because the last time a zero is encountered in the multiplication chain, the error vanishes, by the observation above. Assuming that S is a "typical set", the expected length of a continuous chain of trailing 1's is $\sum_{i=1}^{v} i \cdot 2^{-i} < 2$. In other words, the multiplicative factor of v vanishes from the error expression as well, and we get error growth close to $O(B \cdot \sqrt{n \log q} |S|)$. This is the same effect as if one were merely adding $|S|$ ciphertexts.

3.1.1.4   How to Set Parameters: Let f be the function that we are evaluating, for example the expression in (1). Let $\text{error}_f(B, n, q)$ denote how much the error grows when evaluating a function f on ciphertexts in R_q with an initial error of magnitude B. For correct decryption, we need

$$\text{error}_f(B, n, q) < q/2 \qquad (7)$$

Since errors grow slower in our scheme, q can be set to be correspondingly smaller for the same security level. Following the analysis of Lindner and Peikert [32], for a security level of λ bits, we need

$$n > \log(q/\sigma)(\lambda + 110)/7.2 \qquad (8)$$

Since our $\log q$ is smaller, we can set our n to be smaller, for the same security level λ. In turn, since we now have a smaller n, our new $\text{error}_f(B, n, q)$ is smaller, leading to an even smaller q, and so on. The optimal parameters are obtained by solving both the above inequalities together. Table 1 summarizes our final parameter selection.

4   POLYNOMIAL MULTIPLICATION

4.1   Number Theoretic Transform

The Number Theoretic Transform (NTT), analogous to the well known FFT, is used to speedup the polynomial convolution operation to $O(n \log(n))$ for the finite field modular polynomial multiplications as was described in Schonhage-Strassen multiplication algorithm [33]. Our target hardware platform is a GPU, however, random memory access on GPUs may hurt the performance. Having this in mind, the NTT engine was carefully chosen to exploit serial memory accesses. In [34], several designs for FFT engines were introduced, one of which was explicitly designed for sequential data accesses and storage, which
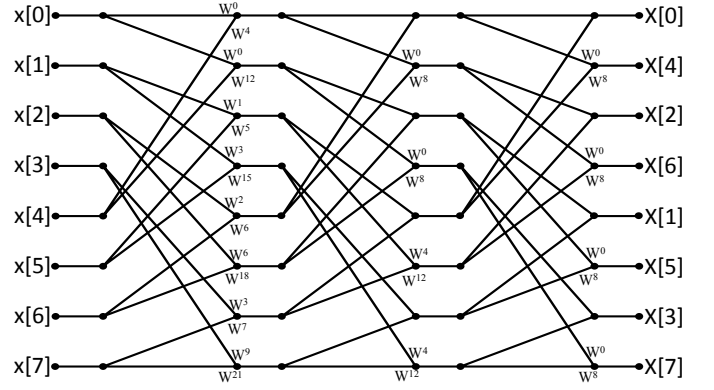

Fig. 1. Forward 8-point NTT engine optimized for sequential data accesses and storage.

is more suited for a GPU global memory architecture. This design was also attractive due to its hardware regularity, which is also better suited for the GPU kernels.

To convert the polynomial to its NTT representation we evaluate the polynomial at the roots of unity of Φ_m. The roots of unity of $\Phi_m(X) = x^n + 1$ are in the form of odd powers of ζ (i.e. roots=ζ^{2k+1} for $0 \le k < n$), where ζ is the nth root of unity. For ζ to be a valid n^{th} root of unity, it must satisfy both these conditions: a) $\zeta^{2n} = 1 \bmod q$ and b) $\zeta^i \neq 1 \bmod q$ for $i < 2n$. The equation for the N-point forward NTT transform is

$$X(k) = \sum_{n=0}^{N-1} x(n) W^{n(2k+1)} \qquad , 0 \le k < N \qquad (9)$$

where $W = \zeta$. W is also called the twiddle factor. The analysis made in [34] was followed to port the NTT equation to the same $n \cdot log(n)$ FFT architecture.

The final NTT architecture for an 8-point NTT is shown in Fig. 1. This architecture has the same structure for each level and supports sequential memory accesses, which is well suited for our GPU implementation. The twiddle factors of the modified NTT engine in Fig. 1 are re-formatted to reduce the number of modulus operations needed. The N-point inverse NTT equation is shown in (10). The final inverse NTT engine can be seen if Fig. 1 is viewed from the right side.

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(K) W_N^{-n(2k+1)} \qquad , 0 \le n < N \qquad (10)$$

4.2   Solinas Prime Numbers

Finite field NTT is performed modulo a specific modulus q as was described in Section 4.1. Modulus reduction is simply a division operation but with the downside that it is an expensive operation in hardware. Alternatively, modulus reduction can be done through successive addition and subtraction operations modulo the same prime q. Solinas primes are special prime numbers introduced in [35] which support high efficiency modulo reduction. We chose one of the Solinas primes $q = 0 \times 7FFE001$ to fit our FHE prime

number bit width $\ell = 31$ bits. For example, if an input number a is in the form of

$$a = a_1 \cdot 2^\ell + a_0 \tag{11}$$

the modulus operation modulo $q = 0 \times 7FFE001$ will simply be

$$a \bmod q = a_0 - a_1 + (a_1 << 17) \tag{12}$$

where "$<<$" is a shift left operation.

5 CANDIDATE APPLICATIONS

5.1 Homomorphic Spam Filter

We implement a homomorphic version of Bayesian spam filters [36]. The main idea behind a Bayesian classifier is that words have certain probabilities of occurrence in *authentic* emails (sometimes called *ham* emails) and in *spam* emails. Since the filter doesn't know these probabilities in advance, email training sets are used to estimate these probabilities. The training phase is assumed to take place on unencrypted training sets, and results in a database of words together with probabilities associated to each word arising in spam e-mails. Once this database is created, the word probabilities are used to classify new emails.

Let p_w denote the probability that a word w occurs in spam e-mails. Given an e-mail with key words $(w_1, \ldots, w_K)$, there are many techniques to combine the probabilities of each word to compute a final estimate of whether the e-mail should be classified as spam [37]. The simplest perhaps is to use Bayes rule. This results in the following expression for p, the probability that the e-mail will be classified as spam.

$$p = \frac{p_{w_1} p_{w_2} \cdots p_{w_K}}{p_{w_1} p_{w_2} \cdots p_{w_K} + (1 - p_{w_1})(1 - p_{w_2}) \cdots (1 - p_{w_K})} \tag{13}$$

At a high level, the email server will receive encrypted words w_i, and map them, homomorphically, into the numbers p_w. Once we obtain these numbers p_w, we wish to compute the expression in (13) to obtain p.

The first downside of (13), when it comes to homomorphic computations, is that integer divisions are extremely expensive to carry out using current homomorphic encryption schemes. In order to overcome this, we make a number of reformulations of the equation above, as follows.

$$\eta \overset{\Delta}{=} \log(1 - p) - \log p = \sum_{i=1}^{K} (\log(1 - p_{w_i}) - \log p_{w_i}) \tag{14}$$

and we will let

$$\eta_{w_i} \overset{\Delta}{=} \log(1 - p_{w_i}) - \log p_{w_i} \tag{15}$$

In other words, the email training phase will store the numbers η_w for each word w in the dictionary (rather than the numbers p_w). The numbers η_w are represented as binary fixed-point numbers, whose bits are encoded into the coefficients of polynomial π_w. For example, $\eta_w = 101_b$ is represented as the polynomial $\pi_w = x^0 + x^2$. The addition of two binary polynomials will not generate a carry

between adjacent polynomial elements, rather polynomial elements will grow individually and will be appropriately reconstructed after decryption (e.g. $101_b + 111_b = 212$, which will be constructed back after decryption to 1100_b). The encrypted spam filter computation will take as input an encrypted word w, map it first into an encrypted η_w as will be described in Function 5, and then simply perform a homomorphic addition of the η_w to get an encrypted η. This is then sent back to the client who decrypts, recovers η using her secret key, and computes $p = 1/(2^\eta + 1)$ in the clear.

The only remaining question is how to map encrypted words w into encrypted η_w. When input e-mails and words are not encrypted, matching a certain word is an easy task. Each email word can be searched across the database. If the word is found, the corresponding number η_w is fetched. On the other hand, when input emails are encrypted, matching words become much harder. This "lookup problem" is the same as the problem of private information retrieval (PIR) [38], [39]. Our data-classifier takes the idea of PIR one step forward, it homomorphically computes on the encrypted data retrieved by the PIR to obtain useful pieces of information. Other PIR constructions [5], [38], [39], [40], [41], [42] cannot implement data-classifiers the way we do because they either: (a) cannot compute with the PIR responses, or, (b) their plaintext field is only mod 2 (or modulo a small prime, for efficiency purposes) and thus they cannot do integer addition as required by (14). Our HE on the other hand has the advantage of being able to use the full modulo-q domain for plaintext additions. As should be clear from the description above, spam filtering is just an example of a class of "lookup-and-compute" type of applications for which we can use our HE scheme.

Function 2 shows how to encrypt individual words in a given list (email). Function 3 shows how we match an input encrypted word versus another unencrypted word from the database. The matching function in Function 3 can be used to construct our encrypted-email spam-filter by simply multiplying the database word probabilities by the "match" output as in Function 5. Only the words that find a match in the database will contribute towards the final probability. It is also possible to keep the database encrypted to protect it from outside attackers. To do this, the matching function should be replaced by `EncWordMatch` presented in Function 4, which performs bit matching for two encrypted inputs, but at an extra cost of two extra Ctxt multiplications to implement the XNOR operation.

In order to increase the performance and efficiency of the spam filter, several optimizations are introduced:

5.1.0.1 Optimization 1: By storing probability numbers in a single polynomial entry (ex. $\eta = 5, \pi = 5x^0$), the other polynomial entries will be unused. This will also lead to the rapid growth of the final result. We decided to store a probability number as binary bits in adjacent polynomial entries (e.g. $\eta = 5 = 101_b, \pi = x^0 + x^2$). By doing so, we will benefit from the unused slots and also when the adjacent slots are added without a carry

---

Function 2: Word List Encryption

Input: Set of words in a list (email)
Output: Encrypted words using HE

```
For each word in the list {
    a = Hash( word ).
    For each bit i in a {
        C_i = Encrypt( a_i ).
        Store C_i to the output list.
    }
}
```

Function 3: Word Matching `WordMatch`

$Input_1$: Encrypted bits of $Word_1$ (C_i)
$Input_2$: Plaintext $Word_2$
Output: Encrypted binary bit $"match"$ = 1 if words match, 0 otherwise

```
match = 1
a = Hash( Word_2 ).
For each bit i in a {
    If ( a_i = 1 )
        B_i = C_i
    Else
        B_i = 1 − C_i
    match = match × B_i
}
Return match
```

Function 4: Enc. Word Matching `EncWordMatch`

$Input_1$: Encrypted bits of $Word_1$ (C_i)
$Input_2$: Encrypted bits of $Word_2$ (D_i)
Output: Encrypted binary bit $"match"$ = 1 if words match, 0 otherwise

```
match = 1
For each bit i {
    B_i = \overline{C_i ⊕ D_i}
    match = match × B_i
}
Return match
```

Function 5: Homomorphic Email Spam Filter

$Input_1$: Encrypted email
$Input_2$: Spam database (DB)
Output: Encrypted email spam/ham probability

```
prob = 0
For each encrypted word "i" in the email {
    For each word "j" in the database {
        match = WordMatch(EmailWord_i , DBWord_j)
        prob += match × WordProbability_j
    }
}
Return prob
```

---

propagate, values in individual slots will grow much slower than before (grows logarithmically). By having individual polynomial slot values grow logarithmically, we will also have a logarithmic growth in the Ctxt noise as was discussed in Section 3.1.1.

5.1.0.2 Optimization 2: The matching Function 3 was stated naively for simplicity. This is done by matching bits one-by-one to get the matching flag. Another more clever way to do the same task when database words are not encrypted, we can rearrange database entries in ascending order. By doing so, we can infer consecutive matching bits in adjacent plaintext entries in the database to skip redundant computations. As a simple example, assume the following two 4-bit database entries: 1001 and 1011, both those entries share the left-most two bits "10". Instead of doing 6 multiplication operations to match an input encrypted word with those two entries as in Function 3, we can store partial matching results of the left-most two bits "10" and reduce these multiplication operations to 4 operations. Experimental results for a database of size 10^5 and hash numbers of size 32-bits show that the number of multiplications needed for matching one word across the entire database decrease from $32 \cdot 10^5$ to $16 \cdot 10^5$ which is a factor of 2 reduction in the number of multiplications.

5.1.0.3 Optimization 3: The interesting property of *zero plaintext, zero error*, described in Section 3.1.1.3, can be used for applications where we can correctly decrypt a *binary flag* even when it is totally submerged in noise! For example, if the application in hand needs many multiplication operations to be done to match one entry as in (1), this may lead to the rapid growth of the noise in the Ctxt to the limit that it may not be decrypted correctly. On the other hand, as discussed in Section 3.1.1.3, in the case of non-matching items, the result will have much less noise. This means that when the resulting flag is "0", it will most probably be decrypted correctly. Otherwise, if we get an error in decryption, this most probably means that the resulting flag was a "1".

5.2 Secure Multiple Keyword Search

Another interesting problem is the problem of searching for a set of input encrypted keywords in encrypted files [21]. Consider an application at an airport where an agent can encrypt passenger names and search for them in an encrypted watchlist present in the cloud. This would be crucial to preserve the security of the watchlist without compromising the privacy of the passengers. Another useful security application would be in monitoring encrypted emails for keywords without compromising the privacy of users. This problem is somewhat parallel to the problem of the data classifier discussed in Section 5.1. The only difference is that Function 5 will be replaced by Function 6 which will compute the number of matched keywords in a given file. The computational complexity of this search problem can be decreased if the input keywords are not encrypted (plaintext keywords) but still the files being searched are still encrypted (henceforth referred to as Partially Secure Keyword Search). In this case, `EncWordMatch` can be replaced by

Function 6: Secure Multiple Keyword Search In Encrypted Files

---

Input$_1$: Set of encrypted keywords
Input$_2$: Encrypted file
Output: Keywords Found "$Result = 1$", otherwise "0"

$result = 1$
For each encrypted keyword "i" {
    For each encrypted word "j" in the file {
        $match = \texttt{EncWordMatch}(FileWord_j, Keyword_i)$
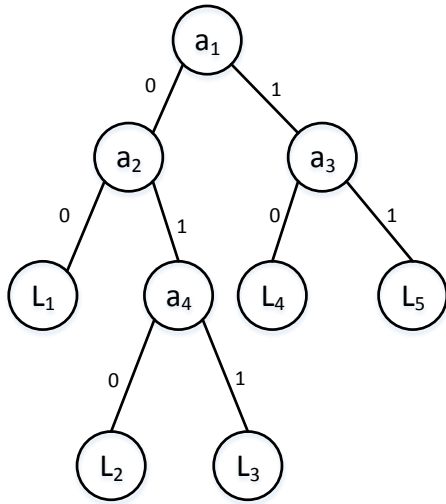 $result\ += match$
    }
}
Return $result$

Fig. 2. Binary Decision Tree with nodes a_i and leafs L_i.

`WordMatch` defined in Function 3, which is computationally less expensive.

5.3 Binary Decision Trees

Binary decision trees are classifiers consisting of interior nodes and leaf nodes. Interior nodes are decision nodes which decide which direction the tree should follow. Leaf nodes are the final tree decision. Binary decision trees are considered as a simplified version of the spam filter described previously, which is considered as a complete decision tree. Fig. 2 shows an example of a binary decision tree with 4 nodes and 5 leafs.

The decision tree in Fig. 2 can be expressed as polynomial equation as in (16). Such a polynomial equation can be efficiently implemented using our HE scheme.

$$
\begin{aligned}
T(a_1, a_2, a_3, a_4) = {} & a_1(a_3 \cdot L_5 + (1 - a_3) \cdot L4) \\
& + (1 - a_1)(a_2(a_4 \cdot L_3 + (1 - a_4) \cdot L_2) \\
& + (1 - a_2) \cdot L_1)
\end{aligned}
$$

$$(16)$$

Table 2
Design Environment.

Item	Specification
CPU	Intel Core-i7 5930K
# of CPU Cores	4
# of Threads	8
CPU Frequency	3.5 GHz
Cache Size	15 MB
System Memory	32 GB DDR4
Operating System	Windows 8.1 Ultimate 64-bits
Programming IDE	Visual Studio 2012 Ultimate edition
GPU	NVIDIA GeForce GTX980
Maxwell Version	GM204
# of CUDA Cores	2048
GPU Core Frequency	1126 MHz
GPU Memory	4 GB
GPU L2 Cache	2 MB

5.4 Security against attacks

Our homomorphic encryption scheme and algorithms, and indeed all known FHE schemes, are proven secure in the IND-CPA sense (i.e., under a chosen plaintext attack). This is the standard notion of security for FHE schemes as in [1], [5], [43]. The algorithms in Section 5 are secure against external attackers. They are also secure against an honest but curious server that wants to learn the underlying encrypted data without trying to actively change it. It is trivial that any homomorphic encryption scheme can be broken by CCA2 (i.e., if the adversary can make decryption queries after the challenge). It can also be broken by CCA1 attacks [44] (i.e., if the adversary can make decryption queries, but only before the challenge). The correctness of any FHE algorithm relies on the honesty of the server that it will execute the exact algorithm.

6 PERFORMANCE RESULTS

6.0.0.1 *Design Environment:* A summary of the specifications of the system used to implement our work for the purpose of benchmarking is found in Table 2.

6.0.0.2 *Test Environment:* Visual studio instrumentation profiling and CUDA NSIGHT performance analyzer were used in measuring the performance of our code. Performance results presented in this section were measured using two experiments

- Ctxt Multiplication time vs. Circuit Depth: The circuit depth d of a circuit is the number of multiplication levels needed to implement a certain function. For example, multiplying 32 numbers requires a circuit depth $d = 5$. To test the running time of Ctxt multiplication, correctness must be tested as well. To test the running time of the executable with specific parameters for circuit level d, 2^d Ctxts are independently created each encrypting a "1" polynomial. All 2^d Ctxts are multiplied into one accumulator which is then decrypted and the resulting plain text is checked for correctness. The running time is averaged over 100 iterations. If the parameters chosen for a given
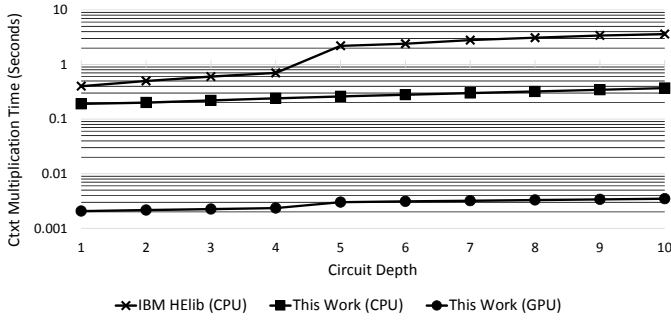
Fig. 3. Ctxt multiplication time of our work when a) running on a single CPU core b) running on a GPU, compared to the IBM HElib when running on a single CPU core across different circuit depths. The running time is plotted on a log scale.
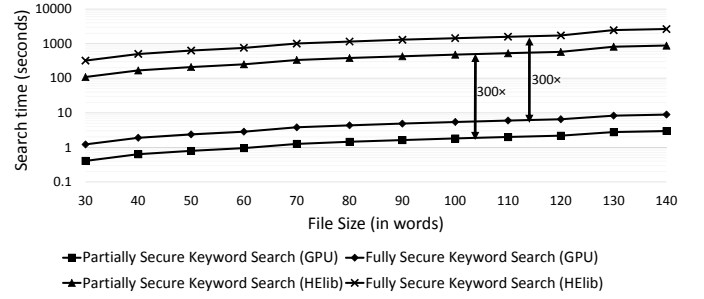

Fig. 4. Fully Secure and Partially Secure keyword search running time, for single keyword, versus different file sizes compared to IBM HElib (Fully Secure search means both the input keywords and the files being searched are encrypted, whereas Partially Secure search means input keywords are in plaintext but the files being searched are still encrypted).

implementation were not sufficient to maintain its correctness, then we increase these parameters and repeat the tests until correct results are acquired.

- Algorithm Implementation: The running time of our algorithms are tested by averaging 10 full iterations of each algorithm using the appropriate parameters.

6.1 Ctxt Multiplication

Ctxt multiplication is considered the main bottleneck for most of the homomorphic applications. Thus, we next report the performance of the Ctxt multiplication operation on different platforms (CPU and GPU). To exploit the parallelism in our work, the GPU implementation partitioned the polynomial operations across GPU cores. The downside of the IBM HElib is that it is not parallelizable.

The performance of the Ctxt multiplication in our library a) running on a single CPU core and b) running on a GPU, compared to the IBM HElib library running on a single CPU core, across different circuit depths, is shown in Fig. 3. It can be noticed from Fig. 3 that our CPU and GPU implementations for our library scores speedups up to $10\times$ and $1035\times$ respectively, compared to IBM HElib, across circuit depths larger than 5. Note also that there is a jump in the running time of the IBM library between circuit depth 4 and 5; this is due to a large increase in the polynomial degree used in the IBM HElib to maintain the same security level and correctness of the scheme. This is an implementation dependent choice by the HElib developers in the selection of parameters in the IBM HElib library.

Table 3 summarizes the performance results of the complete homomorphic operations for our library compared to [9], [14] at a circuit depth equal 10. It can be seen from this table that we have a $10\times$ speedup for the multiplication operation of our CPU implementation compared to the IBM HElib library. By additionally exploring the parallelizable properties that our HE library has, we get another $103.5\times$ speedup by distributing the HE computations on the GPU cores. This resulted in an overall $1035\times$ speedup for the multiplication operations compared to the IBM HElib library and a $36310\times$ compared to [9]. The comparison between the Ctxt size of this work and

the Ctxt in the IBM HElib library is shown in Table 4. It can be noticed from Table 4 that the Ctxt size in the IBM HElib is about $1.5\times$ larger than the Ctxt size in this work. This is due to the fact that in the IBM HElib, the need for a chain of moduli required for the noise management leads to a modulus width approximately $10\times$ larger. Since the polynomial degree is directly related to the modulus width by equation (8), this resulted in another $10\times$ growth in the polynomial degree in the IBM HElib. This increase in the modulus size and the polynomial degree resulted in a Ctxt in the IBM HElib $1.5\times$ larger than this work despite the $log(q)$ difference present in the equation of the Ctxt size in Table 4.

6.2 Secure Multiple Keyword Search Performance

The performance of the Fully Secure and Partially Secure keyword search engines, described in Section 5.2, compared to IBM HElib versus different file sizes is shown in Fig. 4. We observe a $300\times$ speedup for the Fully Secure keyword search on a GPU compared to IBM HElib (and $1035\times$ speedup for circuit depth ≥ 5 as indicated in Fig. 3). It is worth mentioning that our implementation is totally scalable and parallelizable. Increasing the number of GPUs inside the server by a factor G, will automatically scale down the running time of the our search engine by the same ratio.

6.3 Secure Binary Decision Tree Performance

The performance of the decision tree depends on the tree structure and the number of nodes and leafs, which will affect our parameter selection and Ctxt operation running times. The decision tree running time depends mainly on the number of multiplications needed. For example, the polynomial equation in (16) that describes the tree in Section 5.3 has 8 multiplication operations, each multiplication operation takes 3.477 milliseconds, which results in a total running time of 27 milliseconds compared to multiple seconds in [11].

7 CONCLUSION

We described, optimized, and implemented an RLWE-based variant of the HE scheme of [6], [7] which achieves much

Table 3
Performance comparison between this work and the IBM HElib. Running time is in seconds.

| | This Work | | IBM | GPU Speedup | Work in [9] | GPU Speedup |
	CPU	GPU	HElib	over IBM HElib		over [9]
Startup	0.27	0.27	85.3	$316\times$	5	$18.5\times$
Encrypt	0.383	0.023	0.59	$25.6\times$	4.8	$208.7\times$
Decrypt	0.3	0.005	0.39	$78\times$	2.27	$454\times$
Add	0.006	200×10^{-6}	0.002	$10\times$	0.013	$65\times$
Multiply	0.372	3.477×10^{-3}	3.6	$1035\times$	126.25	$36310\times$

Table 4
Ctxt Size Comparison.

	This Work	IBM HElib
$SecurityLevel(\lambda)$	80	80
$Depth(L)$	10	10
Modulus Width (bits)	$log(q_1) = 31$	$log(q_2) = 301$
Poly. Degree [32]	$n > log(q)(\lambda + 110)/7.2$	
n	$n_1 = 1024$	$n_2 = 13981$
$Ctxt$ size (bits)	$4 \cdot n_1 \cdot log^2(q_1)$ ≈ 3.9 Million	$2 \cdot n_2 \cdot log(q_2)$ ≈ 6 Million

slower growth of noise, and thus much better parameters than previous HE schemes for the same security level. We implement three representative applications, namely encrypted spam filters, secure multiple keyword search, and secure binary decision trees, using this HE scheme. Compared to the IBM HElib library, our GPU implementation scores a speedup of $1035\times$ in Ctxt multiplication, which represents the bottleneck for most HE schemes. Our secure search engine application runs in a few seconds on small to moderate file sizes, and our decision tree computations run in a milliseconds for moderate size decision trees.

ACKNOWLEDGMENTS

REFERENCES

[1] C. Gentry, "A fully homomorphic encryption scheme," Ph.D. dissertation, Stanford University, 2009, crypto.stanford.edu/craig.

[2] J.-S. Coron, A. Mandal, D. Naccache, and M. Tibouchi, "Fully Homomorphic Encryption over the Integers with Shorter Public Keys," in *Advances in Cryptology – CRYPTO 2011*, ser. Lecture Notes in Computer Science, P. Rogaway, Ed. Springer Berlin Heidelberg, 2011, vol. 6841, pp. 487–504. [Online]. Available: dx.doi.org/10.1007/978-3-642-22792-9_28

[3] M. Dijk, C. Gentry, S. Halevi, and V. Vaikuntanathan, "Fully Homomorphic Encryption over the Integers," in *Advances in Cryptology – EUROCRYPT 2010*, ser. Lecture Notes in Computer Science, H. Gilbert, Ed. Springer Berlin Heidelberg, 2010, vol. 6110, pp. 24–43. [Online]. Available: dx.doi.org/10.1007/978-3-642-13190-5_2

[4] Z. Brakerski, C. Gentry, and V. Vaikuntanathan, "(Leveled) Fully Homomorphic Encryption Without Bootstrapping," in *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, ser. ITCS '12, New York, NY, USA, 2012, pp. 309–325. [Online]. Available: doi.acm.org/10.1145/2090236.2090262

[5] Z. Brakerski and V. Vaikuntanathan, "Efficient Fully Homomorphic Encryption from (Standard) LWE," in *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, 2011, pp. 97–106.

[6] C. Gentry, A. Sahai, and B. Waters, "Homomorphic Encryption from Learning with Errors: Conceptually-Simpler, Asymptotically-Faster, Attribute-Based," in *Advances in Cryptology – CRYPTO 2013*, ser. Lecture Notes in Computer Science, R. Canetti and J. Garay, Eds. Springer Berlin Heidelberg, 2013, vol. 8042, pp. 75–92. [Online]. Available: dx.doi.org/10.1007/978-3-642-40041-4_5

[7] Z. Brakerski and V. Vaikuntanathan, "Lattice-based FHE As Secure As PKE," in *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science*, ser. ITCS '14, New York, NY, USA, 2014, pp. 1–12. [Online]. Available: doi.org/10.1145/2554797.2554799

[8] ——, "Fully Homomorphic Encryption from Ring-LWE and Security for Key Dependent Messages," in *Advances in Cryptology – CRYPTO 2011*, ser. Lecture Notes in Computer Science, P. Rogaway, Ed. Springer Berlin Heidelberg, 2011, vol. 6841, pp. 505–524. [Online]. Available: dx.doi.org/10.1007/978-3-642-22792-9_29

[9] M. Naehrig, K. Lauter, and V. Vaikuntanathan, "Can Homomorphic Encryption Be Practical?," in *Proceedings of the 3rd ACM Workshop on Cloud Computing Security Workshop*, ser. CCSW '11, New York, NY, USA, 2011, pp. 113–124. [Online]. Available: doi.acm.org/10.1145/2046660.2046682

[10] C. Gentry, S. Halevi, and N. Smart, "Homomorphic Evaluation of the AES Circuit," in *Advances in Cryptology – CRYPTO 2012*, ser. Lecture Notes in Computer Science, R. Safavi-Naini and R. Canetti, Eds. Springer Berlin Heidelberg, 2012, vol. 7417, pp. 850–867. [Online]. Available: dx.doi.org/10.1007/978-3-642-32009-5_49

[11] R. Bost, R. A. Popa, S. Tu, and S. Goldwasser, "Machine Learning Classification over Encrypted Data," Cryptology ePrint Archive, Report 2014/331, 2014, http://eprint.iacr.org/.

[12] J. Alperin-Sheriff and C. Peikert, "Faster Bootstrapping with Polynomial Error," in *Advances in Cryptology – CRYPTO 2014*, ser. Lecture Notes in Computer Science, J. Garay and R. Gennaro, Eds. Springer Berlin Heidelberg, 2014, vol. 8616, pp. 297–314. [Online]. Available: http://dx.doi.org/10.1007/978-3-662-44371-2_17

[13] L. Ducas and D. Micciancio, "FHE Bootstrapping in less than a second," Cryptology ePrint Archive, Report 2014/816, 2014, http://eprint.iacr.org/.

[14] S. Halevi and V. Shoup. (2013) Design and Implementation of a Homomorphic-Encryption Library. researcher.ibm.com/researcher/files/us-shaih/he-library.pdf.

[15] D. Cousins, K. Rohloff, C. Peikert, and R. Schantz, "An update on SIPHER (Scalable Implementation of Primitives for Homomorphic EncRyption) ; FPGA implementation using Simulink," in *High Performance Extreme Computing (HPEC), 2012 IEEE Conference on*, 2012, pp. 1–5.

[16] W. Wang, Y. Hu, L. Chen, X. Huang, and B. Sunar, "Exploring the Feasibility of Fully Homomorphic Encryption." Los Alamitos, CA, USA: IEEE Computer Society, 2013, vol. 99, pp. 1–10.

[17] M. Yasuda, T. Shimoyama, J. Kogure, K. Yokoyama, and T. Koshiba, "Secure pattern matching using somewhat homomorphic encryption," in *Proceedings of the 2013 ACM Workshop on Cloud Computing Security Workshop*, ser. CCSW '13, New York, NY, USA, 2013, pp. 65–76. [Online]. Available: http://doi.acm.org/10.1145/2517488.2517497

[18] W. Wang, Y. Hu, L. Chen, X. Huang, and B. Sunar, "Accelerating fully homomorphic encryption using GPU," in *High Performance Extreme Computing (HPEC), 2012 IEEE Conference on*, 2012, pp. 1–5.

[19] M. Brenner, H. Perl, and M. Smith. (2011) Scarab library . hcrypt.com/scarab-library/.

[20] H. Perl, M. Brenner, and M. Smith, "Poster: An Implementation

of the Fully Homomorphic Smart-vercauteren Crypto-system," in *Proceedings of the 18th ACM Conference on Computer and Communications Security*, ser. CCS '11, New York, NY, USA, 2011, pp. 837–840. [Online]. Available: doi.acm.org/10.1145/2046707.2093506

[21] E. M. RUSLI. (2014) A Darpa Director on Fully Homomorphic Encryption (or One Way the U.S. Could Collect Data) . http://blogs.wsj.com/digits/2014/03/09/a-darpa-director-on-fully-homomorphic-encryption/-or-one-way-the-u-s-could-collect-data/.

[22] R. A. Popa, C. M. S. Redfield, N. Zeldovich, and H. Balakrishnan, "CryptDB: Protecting Confidentiality with Encrypted Query Processing," in *Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles*, ser. SOSP '11, New York, NY, USA, 2011, pp. 85–100. [Online]. Available: doi.acm.org/10.1145/2043556.2043566

[23] T. Graepel, K. Lauter, and M. Naehrig, "ML Confidential: Machine Learning on Encrypted Data," in *Information Security and Cryptology – ICISC 2012*, ser. Lecture Notes in Computer Science, T. Kwon, M.-K. Lee, and D. Kwon, Eds. Springer Berlin Heidelberg, 2013, vol. 7839, pp. 1–21. [Online]. Available: dx.doi.org/10.1007/978-3-642-37682-5_1

[24] D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public Key Encryption with Keyword Search," 2004, pp. 506–522.

[25] J. Baek, R. Safavi-Naini, and W. Susilo, "Public Key Encryption with Keyword Search Revisited," in *Computational Science and Its Applications – ICCSA 2008*, ser. Lecture Notes in Computer Science, O. Gervasi, B. Murgante, A. Laganà, D. Taniar, Y. Mun, and M. Gavrilova, Eds. Springer Berlin Heidelberg, 2008, vol. 5072, pp. 1249–1259. [Online]. Available: dx.doi.org/10.1007/978-3-540-69839-5_96

[26] D. Boneh, C. Gentry, S. Halevi, F. Wang, and D. Wu, "Private Database Queries Using Somewhat Homomorphic Encryption," in *Applied Cryptography and Network Security*, ser. Lecture Notes in Computer Science, M. Jacobson, M. Locasto, P. Mohassel, and R. Safavi-Naini, Eds. Springer Berlin Heidelberg, 2013, vol. 7954, pp. 102–118. [Online]. Available: dx.doi.org/10.1007/978-3-642-38980-1_7

[27] D. Cash, S. Jarecki, C. Jutla, H. Krawczyk, M.-C. Roşu, and M. Steiner, "Highly-Scalable Searchable Symmetric Encryption with Support for Boolean Queries," in *Advances in Cryptology – CRYPTO 2013*, ser. Lecture Notes in Computer Science, R. Canetti and J. Garay, Eds. Springer Berlin Heidelberg, 2013, vol. 8042, pp. 353–373. [Online]. Available: dx.doi.org/10.1007/978-3-642-40041-4_20

[28] C. Aguilar-Melchor, S. Fau, C. Fontaine, G. Gogniat, and R. Sirdey, "Recent Advances in Homomorphic Encryption: A Possible Future for Signal Processing in the Encrypted Domain," 2013, vol. 30, no. 2, pp. 108–117.

[29] B. Hayes, "Alice and Bob in Cipherspace," 2012, vol. 100, no. 5, pp. 362–367.

[30] J. Benaloh and V. Vaikuntanathan. (2013) Homomorphic Encryption: What, How, and Why . people.csail.mit.edu/vinodv/Homomorphic-TECHFEST-Final.ppsx.

[31] V. Lyubashevsky, C. Peikert, and O. Regev, "On Ideal Lattices and Learning with Errors over Rings," in *Advances in Cryptology – EUROCRYPT 2010*, ser. Lecture Notes in Computer Science, H. Gilbert, Ed. Springer Berlin Heidelberg, 2010, vol. 6110, pp. 1–23. [Online]. Available: dx.doi.org/10.1007/978-3-642-13190-5_1

[32] R. Lindner and C. Peikert, "Better Key Sizes (and Attacks) for LWE-based Encryption," in *Proceedings of the 11th International Conference on Topics in Cryptology: CT-RSA 2011*, ser. CT-RSA'11. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 319–339. [Online]. Available: http://dl.acm.org/citation.cfm?id=1964621.1964651

[33] A. Schönhage and V. Strassen, "Schnelle multiplikation großer zahlen," *Computing*, vol. 7, no. 3-4, pp. 281–292, 1971. [Online]. Available: http://dx.doi.org/10.1007/BF02242355

[34] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Pearson Education, 2006. [Online]. Available: books.google.ca/books?id=geTn5W47KEsC

[35] J. A. Solinas, "Generalized Mersenne Numbers," University of Waterloo, Tech. Rep. CORR 99-39, 1999.

[36] M. Sahami, S. Dumais, D. Heckerman, and E. Horvitz, "A Bayesian Approach to Filtering Junk E-Mail," in *Learning for Text Categorization: Papers from the 1998 Workshop*. Madison, Wisconsin: AAAI Technical Report WS-98-05, 1998. [Online]. Available: citeseer.ist.psu.edu/sahami98bayesian.html

[37] Wikipedia. (2004) Naive Bayes spam filtering . http://en.wikipedia.org/wiki/Naive_Bayes_spam_filtering.

[38] Y. Doroz, B. Sunar, and G. Hammouri, "Bandwidth Efficient PIR from NTRU," in *Cryptology ePrint Archive*, 2014, pp. 1–12.

[39] X. Yi, M. G. Kaosar, R. Paulet, and E. Bertino, "Single-Database Private Information Retrieval from Fully Homomorphic Encryption," May 2013, vol. 25, no. 5, pp. 1125–1134.

[40] C. Aguilar-melchor and P. Gaborite, "A Lattice-Based Computationally-Efficient Private Information Retrieval Protocol," in *WEWoRC 2007*, 2007, pp. 1–13.

[41] N. Fotiou, D. Trossen, G. Marias, A. Kostopoulos, and G. Polyzos, "Enhancing information lookup privacy through homomorphic encryption," *Security and Communication Networks*, vol. 7, no. 12, pp. 2804–2814, 2014. [Online]. Available: http://dx.doi.org/10.1002/sec.910

[42] C. A. Melchor, "High-Speed Single-Database PIR Implementation," in *The 8th Privacy Enhancing Technologies Symposium (PETS 2008)*, 2008, pp. 1–12.

[43] C. Gentry, "Fully Homomorphic Encryption Using Ideal Lattices," in *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, ser. STOC '09, New York, NY, USA, 2009, pp. 169–178. [Online]. Available: doi.acm.org/10.1145/1536414.1536440

[44] M. Chenal and Q. Tang, "On Key Recovery Attacks against Existing Somewhat Homomorphic Encryption Schemes," in *The third International Conference on Cryptology and Information Security in Latin America, Latincrypt 2014*, 2014, pp. 1–28. [Online]. Available: https://orbilu.uni.lu/handle/10993/18106

Alhassan Khedr received his M.Sc and B.Sc degrees from Electronics and Electrical Communications Engineering Department, Faculty of Engineering, Cairo University, Cairo, Egypt in 2008 and 2011 respectively. After graduation, he was appointed as a Teaching Assistant in Cairo University and American University of Cairo for 3 years. He received numerous awards for his excellence as a student and as a teaching assistant. He was among the team responsible for developing and fabricating CUSPARC the first fully developed Egyptian embedded processor. Alhassan joined Electronics and Computer Engineering Department at University of Toronto to pursue his PhD degree in 2011. Alhassan main research interests include algorithm optimization and VLSI implementation of high performance algorithms. He is also interested in parallel and multi/many core processor architecture design and computer arithmetic.

Dr. Glenn Gulak is a Professor in the Department of Electrical and Computer Engineering at the University of Toronto. He is a Senior Member of the IEEE and a registered Professional Engineer in the Province of Ontario. His present research interests are in the areas of algorithms, circuits, and CMOS implementations of high-performance baseband digital communication systems and, additionally, in the area of CMOS biosensors. He has authored or co-authored more than 150 publications in refereed journals and refereed conference proceedings. In addition, he has received numerous teaching awards for undergraduate courses taught in both the Department of Computer Science and the Department of Electrical and Computer Engineering at the University of Toronto. From Jan. 1985 to Jan. 1988 he was a Research Associate in the Information Systems Laboratory and the Computer Systems Laboratory at Stanford University. He has served on the ISSCC Signal Processing Technical Subcommittee from 1990 to 1999, ISSCC Technical Vice-Chair in 2000 and served as the Technical Program Chair for ISSCC 2001. He served on the Technology Directions Subcommittee for ISSCC from 2005 to 2008. He currently serves as the Vice-President of the Publications Committee for the IEEE Solid-State Circuits Society and a member of the IEEE PSPB.

Vinod Vaikuntanathan is a Steven and Renee Finn Career Development Assistant Professor of Computer Science at MIT. His main research interest is in the theory and practice of cryptography. He works on lattice-based cryptography, building advanced cryptographic primitives using integer lattices; leakage-resilient cryptography, defining and developing algorithms resilient against adversarial information leakage; and more recently, the theory and practice of computing on encrypted data, constructing powerful cryptographic objects such as fully homomorphic encryption and functional encryption. Vinod obtained his Ph.D. from MIT where he received a 2009 George M. Sprowls Award for the best MIT Ph.D. thesis in Computer Science. He is also a recipient of the 2008 IBM Josef Raviv Postdoctoral Fellowship, the 2013 Alfred P. Sloan Research Fellowship, the 2014 Microsoft Faculty Fellowship, and a 2014 NSF CAREER award.