# Post-Quantum Secure Identity-Based Encryption Scheme using Random Integer Lattices for IoT-enabled AI Applications 

Dharminder Dharminder, ${ }^{1}$ Ashok Kumar Das ${ }_{(0)}{ }^{2,3}$ Sourav Saha, ${ }^{2}$ Basudeb Bera, ${ }^{2}$ and Athanasios V. Vasilakos ${ }^{4}$<br>${ }^{1}$ Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Chennai, 601 103, India<br>${ }^{2}$ Center for Security, Theory and Algorithmic Research, International Institute of Information Technology, Hyderabad, 500 032, India<br>${ }^{3}$ Virginia Modeling, Analysis and Simulation Center, Old Dominion University, Suffolk, VA 23435, USA<br>${ }^{4}$ Center for AI Research (CAIR), University of Agder (UiA), Grimstad, Norway

Correspondence should be addressed to Ashok Kumar Das; iitkgp.akdas@gmail.com
Received 21 April 2022; Accepted 2 June 2022; Published 6 July 2022
Academic Editor: Wenbo Shi
Copyright © 2022 Dharminder Dharminder et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.


#### Abstract

Identity-based encryption is an important cryptographic system that is employed to ensure confidentiality of a message in communication. This article presents a provably secure identity based encryption based on post quantum security assumption. The security of the proposed encryption is based on the hard problem, namely Learning with Errors on integer lattices. This construction is anonymous and produces pseudo random ciphers. Both public-key size and ciphertext-size have been reduced in the proposed encryption as compared to those for other relevant schemes without compromising the security. Next, we incorporate the constructed identity based encryption (IBE) for Internet of Things (IoT) applications, where the IoT smart devices send securely the sensing data to their nearby gateway nodes(s) with the help of IBE and the gateway node(s) secure aggregate the data from the smart devices by decrypting the messages using the proposed IBE decryption. Later, the gateway nodes will securely send the aggregated data to the cloud server(s) and the Big data analytics is performed on the authenticated data using the Artificial Intelligence (AI)/Machine Learning (ML) algorithms for accurate and better predictions.


## 1. Introduction

According to [1], it is projected by 2027 the market of Internet of Things (IoT) industry will grow by $\$ 2$ trillion annually, which has already a market of $\$ 520$ billion in 2022. In the connected world, the IoT makes an environment where various smart devices are interconnected with each other. The advancement of information and communications technology (ICT) makes the IoT technologies and their solutions rich that have great impact to the society for improving the human life advanced and
easy. There are enormous applications of IoT, such as Industrial IoT (IIoT), smart cities, healthcare monitoring, smart home, and so on. In an IIoT, various IoT smart devices are connected in an industry to collect manufacturing data in order to predict the failure rates to increase productivity and efficiency [2]. In healthcare application, various smart devices like smartwatches and medical sensors are connected in the body of a patient to collect vital information and provide appropriate health condition of that person. Furthermore, in recent days, smart home application is in limelight where the smart
devices like smart locks and home appliances are connected with each other via the internet and they can be also controlled via the mobile devices. Though IoT has transformed the human life easier, there are various serious threats associated with IoT applications. For instance, it was found by HP that $70 \%$ of the devices connected IoT devices are vulnerable to various attacks [3].

In IoT applications, the smart devices exchange the sensitive data among each other and also with various other entities. In such a scenario, an unauthorized user or an attacker may take the advantage to compromise the data by eavesdropping, modifying, updating and deleting the information during the communication [4]. According to broadcom [5] in the year 2017, it was found that there was an approximately $600 \%$ hike in attacks against IoT devices in various applications. Therefore, there is a great need to design a secure IoT system to protect the data from the attackers [6].

Once the sensing information from the deployed smart devices in an IoT environment is aggregated by the nearby gateway node or access point, the gathered data needs to be also stored in semi-trusted cloud servers. Now, the stored data at the cloud is huge in volume and it needs data analytics. As a result, it is preferable to used some Big data analytics using traditional Artificial Intelligence (AI)/Machine Learning (ML) algorithms for accurate and better predictions [7, 8].

Ahanger et al. [9] provided various Machine Learning (ML) and Deep Learning (DL) based mechanisms for IoT paradigm. They also provided a taxonomy based on several IoT vulnerabilities, respective attackers and effects, as well as various threats. Iwendi et al. [10] pointed out the important of deep learning (DL) for detecting attacks in IoT paradigm. They suggested DL based mechanism to detect cyber-attacks on IoT using a long short term networks classifier.

Omolara et al. [11] gave an IoT concept and then provided the deep insights into possible solutions to the IoT security challenges due to the heterogeneous nature of IoT, and the respective emerging issues, opportunities, gaps as well as recommendations. Mukhopadhyay et al. [12] pointed out that IoT sensors need to be reliable, safety as well as privacy-aware for the users interacting with them. Thus, they discussed that IoT sensors having advanced AI capabilities will have the potential for identifying, detecting, and avoiding performance degradation as well as discovering new patterns.

Public-key cryptosystem works under a pair of keys (public key and private key), whereas the public key is made public that is accessible by everyone during communication, and the private key is kept secret and only known to the owner (sender/signer). The notion of the "Identity-Based Encryption (IBE)" due to Shamir [13], solves the certificate management problem. The existing Shor's algorithm [14] is a big threat to the existing number-theoretic identity-based encryptions. The main difference of IBE from certificate based public-key encryption schemes lies in the way how the public and secret
keys pair generated for a user. A private key generator, say $\mathscr{P} \mathscr{K} \mathscr{G}$ handles the process of secret key generation, but it executes user authentication process to confirm the validity of a legitimate user. In IBE process, a public key may be an information such as the user's email address or mobile number. The corresponding secret key is generated by embedding the user's identity with the $\mathscr{P} \mathscr{K} \mathscr{G}$ 's master secret. This process removes the need of certificate that is required for verification of a legitimate recipient's public key. The IBE process also solves the problems related to key generations and distributions in a multiuser settings. In case of limited resources, it can also offer the potential solution to make the process resource efficient.

In the literature, we have three important classes of identity-based encryptions (IBE) (see in Figure 1): 1) IBE based on bilinear pairings [15-18], 2) IBE based on quadratic residue [19, 20], and 3) IBE based on lattices [21]. To the best of our knowledge, most of the constructions proposed in the standard model relies on bilinear pairings.

Chamola et al. [22] reviewed that the disruption which the quantum computers have caused in the cryptographic field. They pointed out that the existing public key encryption schemes can be broken by the quantum computers, and as a result, there is a requirement for hunting the new cryptographic mechanisms that need to be secure in the post-quantum era. Hassija et al. [23] provided a review on several quantum computing applications that can be applied in different computer science areas, including "cryptography", "machine learning", "deep learning" as well as "quantum simulations". They also provided several real-life case studies on "risk analysis", "logistics", and "satellite communication". Hassija et al. [24] also discussed that with the help of online cloud services, the first generation of quantum computers can be programmed and accessed using the available software development kits. Moreover, they presented a growing trend in both the investments as well as patents in the quantum computing field. In recent years, the lattice-based cryptography has played a very important role in the post-quantum era for various real-life applications, such as "Vehicular Ad Hoc Network (VANET)" [25], " medical Cyber-Physical Systems (CPS)" [26] and "mobile communications" [27].
1.1. Research Contributions. There are two reasons to move towards post quantum secure lattice based cryptography: a) simple algebraic operations that are based on matrix multiplication and b) secure against existing quantum assisted algorithms. The main contributions of the work are listed below:

- This article presents a new identity-based encryption based on lattices without using the random oracles. The proposed encryption is anonymous in nature [28], which means that the cipher does not reveal the recipient's identity.
- Our proposed encryption is selective-ID secure [29], and can be converted to an adaptive-ID secure $[15,16,30]$ by


Figure 1: A communication model for IBE.
taking the bit-wise decomposition of the corresponding identities.

- The proposed encryption is inspired from the Water's [18] encryption and signature that use only non-zero positions of bits in the decomposition of the corresponding identity. The encryption is secure under "learning with errors" assumption without the random oracles.
- If $\kappa$ is an appropriate security parameter and $O\left(\kappa^{2}\right)$ is the size of a public key, we can relate the computation time in terms of security parameter complexity as $O\left(\kappa^{2}\right)$, and it can be compared with the size of classic public key (such as RSA [31] and ElGamal [32] cryptosystems) as $O(\kappa)$ and computation time in terms of security parameter as $O\left(\kappa^{3}\right)$, respectively [33].
- We then incorporate the constructed identity based encryption (IBE) for IoT applications, where the smart devices send securely the sensing data to their nearby gateway nodes(s) with the help of IBE and the gateway node(s) secure aggregate the data from the smart devices by decrypting the messages using the proposed IBE decryption. Later, the authenticated data stored at the cloud server(s) will be used for accurate and better predictions with the help of AI/ML algorithms.
1.2. Paper Outline. In Section 2, the security of an IdentityBased Encryption (IBE) is discussed. Section 3 provides a discussion of basic preliminaries that are needed to analyze the proposed scheme in Section 4. In Section 5, we incorporating our proposed IBE scheme for IoT-enabled AI applications. Next, the security analysis of the proposed scheme under standard models is discussed in Section 6. A comparative study among the proposed scheme and other relevant schemes is given in Section 7. Some concluding remarks are then provided in Section 8.


## 2. Security of an Identity-Based Encryption

An identity-based encryption (IBE) [15] comprises of four phases (algorithms): a) Set-up, b) Extraction, c) Encrypts, and d) Decrypts. The Set-up algorithm is run under the public parameters and a secret master key. The Extraction algorithm makes use of the master key to create a secret key respective to the given identity. The Encrypts algorithm
encrypts a message using the identity. Finally, the Decrypts algorithm decrypts a ciphertext with the help of the corresponding private key.
2.1. Both Selective and Adaptive Encryption. The security model of an IBE [15] defines the "indistinguishable adaptive chosen cipher and chosen identity (IND-ID-CCA2)" security. It allows a probabilistic polynomial time-adversary, say $\mathscr{A}$ to pick an identity on which it wants to target. A weaker version of an IBE security [34] restricts the adversary $\mathscr{A}$ to announce the target or identity at advance, that is known as the "indistinguishable adaptive chosen cipher and selective identity (IND-sID-CCA2)" security. We have described this system as a selective identity and chosen cipher secure identity-based encryption. In this version of encryption, we will not allow the adversary $\mathscr{A}$ to process decryption queries on the target identity, which implies a weaker notion of the "indistinguishable against adaptive chosen cipher and chosen identity (IND-ID-CCA2) and indistinguishable adaptive chosen cipher and selective identity (IND-sID-CCA2)", respectively. Another important notion is the "indistinguishable cipher against chosen plaintext attack (IND-CPA)", which is also called semantic security.
2.2. Security Model. We now define an IBE semantic security under the IND-sID-CCA2 with the help of a game that is played between a challenger, say $\mathscr{C}$ and an adversary $\mathscr{A}$. The description of the game is given below.

1. Target-phase: $\mathscr{A}$ declares the target identity ID* in advance.
2. Set-up-phase: $\mathscr{C}$ executes the Set-up-phase, generates the public parameters for $\mathscr{A}$, and keeps the master key as secret.

Phase-1. $\mathscr{A}$ submits queries $q_{1}, q_{2}, \ldots, q_{m}$ corresponding to the identities $\mathrm{ID}_{1}, \mathrm{ID}_{2}, \ldots, \mathrm{ID}_{m}$, respectively, where $\mathrm{ID}_{i} \neq \mathrm{ID}^{*}$ for $1 \leq i \leq m$. Now, $\mathscr{C}$ runs an algorithm, called Extraction $\left(\mathrm{Mk}, \mathrm{ID}_{i}\right)$ with the master key Mk and identity $\mathrm{ID}_{i}$ to obtain the private key $D_{i}$ corresponding to the identity $\mathrm{ID}_{i}, I D_{i}$, which is the public key. Then, it sends $D_{i}$ to $\mathscr{A}$, where all the queries are processed adaptively meaning that
$\mathscr{A}$ can make queries with the knowledge of the previous queries.
3. Challenge-phase: After completion of Phase-1, $\mathscr{A}$ submits two messages $m_{0}$ and $m_{1}$ from the message space on which it executes the challenge. The challenger $\mathscr{C}$ then picks $b \in\{0,1\}$ randomly, and outputs $c=$ Encrypts(paramts, $\mathrm{ID}^{*}, m_{b}$ ) and sends it to $\mathscr{A}$, where paramts, paramts are the parameters relevant to encryption.

Phase-2. $\mathscr{A}$ submits the adaptive extraction queries $q_{m+1}$, $q_{m+2}, \ldots, q_{n}$ corresponding to $\mathrm{ID}_{m+1}, \mathrm{ID}_{m+2}, \ldots, \mathrm{ID}_{n}$, where $\mathrm{ID}_{i} \neq \mathrm{ID}^{*}$, respectively. Next, $\mathscr{C}$ replies as in Phase-1.
4. Guess-phase: Finally, $\mathscr{A}$ requires to guess a bit $b^{\prime} \in\{0,1\}$. The game is won by $\mathscr{A}$ if $b^{\prime}=b$; otherwise, $\mathscr{A}$ looses the game.

We call such an adversary $\mathscr{A}$ as an IND-sID-CPA-adversary, and define the advantage of $\mathscr{A}$ attacking the identity-based encryption, say $\mathscr{P}$ as

$$
\begin{equation*}
A D V_{\mathscr{A}}(\mathscr{P})=\left|\operatorname{Pr}\left[b=b^{\prime}\right]-\frac{1}{2}\right| \tag{1}
\end{equation*}
$$

We can also describe an adaptive phase to the above notion by excluding the target phase, and permitting $\mathscr{A}$ to wait for the challenge phase to declare ID* as challenge identity. $\mathscr{A}$ can submit the arbitrary key extraction queries as in Phase-1, and then select an identity $\mathrm{ID}^{*}, I D^{*}$ as a target. But, the only condition imposed here is that $\mathscr{A}$ cannot submit extraction query on $\mathrm{ID}^{*}, I D^{*}$ in Phase-1, and the resulting notion is called as IND-ID-CPA security. In Ci-pher-Anonymity along with semantic security, we have another important notion of cipher anonymity under chosen plaintext attack.

## 3. Preliminaries

Let $\mathbb{R}$ be a set of real numbers and $x \in \mathbb{R}$ be a real number. We denote $\lfloor x\rfloor$ as the largest integer, but not greater than $x$, whereas $\left.\llcorner x\urcorner={ }_{\llcorner } x+1 / 2\right\urcorner$ denotes the integer closest to $x$, with ties broken upward. We apply a bold big letter $\mathbf{A}$ to denote a matrix and a bold small letter $\mathbf{x}$ to denote a column vector of the matrix $\mathbf{A}$, where $[\mathbf{A} \mid \mathbf{x}]$ denotes concatenation of the matrix $\mathbf{A}$ with a vector $\mathbf{x}$. Let $\mathbb{Z}$ denote the set of all integers and $\mathbb{Z}_{q}=\mathbb{Z} / q \mathbb{Z}$ be a quotient ring under integers modulo a prime $q$, that is, a collection of the (left or right) cosets $a+q \mathbb{Z}$ with addition and multiplication operations in the quotient ring $\mathbb{Z}_{q}$. It is worth noticing that $y=z(\bmod q)$ if and only if $y+q \mathbb{Z}=z+q \mathbb{Z}$, which is an obvious fact about the equality of cosets.
3.1. Lattice. A lattice $\Delta$ is defined with the following two properties: 1) it is an additive subgroup which implies $0 \in \Delta$, and $-x, x+y \in \Delta$ for all $x, y \in \Delta$, and 2) it is discrete that implies every $x \in \Delta$ possesses a neighborhood in $\mathbb{R}^{n}$ in which $x$ is the only lattice point in the neighborhood. More specifically, the $i^{\text {th }}$ successive minima $\lambda_{i}(\Delta)$ is the smallest


Figure 2: A two-dimensional lattice with bad basis.

Euclidean norm $\ell$ such that $\Delta$ possesses $i$ number of linearly independent vectors of norm less than or equal to $\ell$. Due to properties of a discrete group, one can observe that the quotient group $\mathbb{R}^{n} / \Delta$ of cosets $c+\Delta=\{c+v: v \in \Delta\}, c \in \mathbb{R}^{n}$, under the usual addition: $\left(c_{1}+\Delta\right)+\left(c_{2}+\Delta\right)=\left(c_{1}+c_{2}\right)+\Delta$ in the quotient group. A fundamental domain of $\Delta$ is a set $\mathbb{F} \subset \mathbb{R}^{n}$ that contains exactly one representative $\widehat{c} \in(c+\Delta) \cap \mathbb{F}$ of each coset $c+\Delta$.
3.2. Bases and Fundamental Parallelepiped. A lattice (see Fig. 2 ) is generated by a basis $B=\left\{b_{1}, b_{2}, \ldots, b_{m}\right\}$ and the integer linear combination of the linearly independent vectors $b_{1}, b_{2}$, $\ldots, b_{m}$ in the basis as $\Delta=\Delta(B)=\left\{\sum_{i=1}^{m} z_{i} b_{i}: z_{i} \in \mathbb{Z}\right\}$. The positive integer $m$ is the rank of the basis and $n$ represents the dimension of the space under consideration. We can consider $m=n$ to represent a full rank lattice. A lattice possesses infinitely many bases, because if $B$ is a basis then BU is also a basis for a unimodular matrix. If $B$ is a basis of the lattice $\Delta$, the fundamental domain is the parallelepiped $\mathbb{P}(B)=B[-1 / 2,1 / 2)$ centered at the origin. Note that parallelepiped is formed by "six parallelogram sides to result in a three-dimensional figure" or a "Prism", which contains a parallelogram base.

Definition 1. Let $b_{1}, b_{2}, \ldots, b_{m} \in \mathbb{R}^{n}$ be linearly independent tuples, a lattice $\Delta$ generated by a basis $B=\left\{b_{1}, b_{2}, \ldots, b_{m}\right\}$ is denoted $\Delta\left(b_{1}, b_{2}, \ldots, b_{m}\right)=\left\{\Sigma_{i} z_{i} b_{i}: z_{i} \in \mathbb{Z}\right\}$. The integers $m$ and $n$ denote the rank of the concerned matrix and the dimension of given lattice, respectively.

Definition 2. Let $b_{1}, b_{2}, \ldots, b_{m} \in \mathbb{R}^{n}$ be linearly independent tuples that generate $a$ lattice $\Delta\left(b_{1}, b_{2}, \ldots\right.$, $\left.b_{m}\right)=\left\{\Sigma_{i} z_{i} b_{i}: z_{i} \in \mathbb{Z}\right\}$, its dual lattice be $\Delta^{*}=\{z \in \mathbb{Z} \mid \forall y \in \Delta,\langle z, y\rangle \in \mathbb{Z}\}$, where $\Delta$ can be represented as


Figure 3: Gaussian distribution in multi dimensions.

$$
\Delta=\mathbf{A} \cdot \mathbf{z}=\left[\begin{array}{cccc}
\mid & \mid & . & \mid  \tag{2}\\
b_{1} & b_{2} & \ldots & b_{m} \\
\mid & \mid & . & \mid
\end{array}\right] \cdot\left[\begin{array}{c}
z_{1} \\
z_{2} \\
\vdots \\
z_{m}
\end{array}\right]
$$

3.3. q-ary Lattice. The $\mathbf{q}$-ary lattice satisfying $\mathbb{Z}_{q}^{m} \subseteq \Delta \subseteq \mathbb{Z}_{q}^{n}$, for some integer $q$, is called $q$-ary lattice because $q$ times vectors of lattice also belongs to it. Given a matrix modulo $q=$ poly $(n)$ (depends only dimension of lattice), denoted $A \in \mathbb{Z}_{q}^{n \times m}$, there are $n$-dimensional $q$-ary lattice $\Delta_{q}^{\perp}=\left\{z \in \mathbb{Z}^{n}: A z=0(\bmod q)\right\}$ and a coset of the lattice as $\Delta_{q}^{u}=\left\{z \in \mathbb{Z}^{n}: u=A z(\bmod q) \mid z \in \mathbb{Z}^{m}\right\}$, where $m, n$ and $q$ are integers and $m>n$. Here, $\mathbf{A} \cdot \mathbf{z}=0$ implies that:

$$
\left[\begin{array}{cccc}
\mid & \mid & \cdot & \mid  \tag{3}\\
b_{1} & b_{2} & \ldots & b_{m} \\
\mid & \mid & \cdot & \mid
\end{array}\right] \cdot\left[\begin{array}{c}
z_{1} \\
z_{2} \\
\vdots \\
z_{m}
\end{array}\right]=\left[\begin{array}{c}
0 \\
0 \\
\vdots \\
0
\end{array}\right]
$$

and $\mathbf{A} \cdot \mathbf{z}=u$ implies that:

$$
\left[\begin{array}{cccc}
\mid & \mid & \cdot & \mid  \tag{4}\\
b_{1} & b_{2} & \ldots & b_{m} \\
\mid & \mid & \cdot & \mid
\end{array}\right] \cdot\left[\begin{array}{c}
z_{1} \\
z_{2} \\
\vdots \\
z_{m}
\end{array}\right]=\left[\begin{array}{c}
u_{1} \\
u_{2} \\
\vdots \\
u_{m}
\end{array}\right] .
$$

These $q$-ary lattices are applied in the construction of cryptographic techniques. Now, if the matrix $\mathbf{A}$ is chosen randomly, solving the short vector problem on $\Delta^{\perp}(A)$ is equivalent to solve a hard problem in random lattice.
3.4. Gaussian Measures. Let $x, c \in \mathbb{R}^{n}$ and $\sigma>0$ be arbitrary. Then, $\rho_{\sigma, c}(x)=e^{-\pi\|(x-c)\|^{2} / \sigma^{2}}$ defines a Gaussian distribution function (see Fig. 3) with center $c$ and scaling $\sigma$, where the total measure corresponding to $\rho_{\sigma, c}$ is given by $\int_{x \in \mathbb{R}^{n}} \rho_{\sigma, c}(x) \mathrm{d} x=\sigma^{n}$. We can define the discrete Gaussian distribution as $D_{\Delta, \sigma, c}(z)=\rho_{\sigma, c}(z) / \rho_{\sigma, c}(\Delta)$, where $z \in \Delta$ is an
arbitrary lattice point. Note that $D_{\Delta, \sigma, c}(z)=D_{\sigma, c}(z) / D_{\sigma, c}(\Delta)=\rho_{\sigma, c}(z) / \rho_{\sigma, c}(\Delta)$.

We now introduce an advanced lattice parameter (called the smoothing parameter [35]) related to the Gaussian measures on random lattices as follows.

Definition 3. Let $\Delta$ be a lattice of dimension $n$ and $\varepsilon>0$ be an arbitrary small real number. The smoothing parameter is defined by $\xi_{\varepsilon}(\Delta)$ to be the smallest $\sigma>0$ such that $\sum_{z \neq 0 \in \Delta^{*}} \rho_{1 / \sigma, 0}(z) \leq \varepsilon$ holds.
3.5. Hard Assumptions Based on Learning with Errors. The "learning with errors" was introduced by Regev [36], which is secure against quantum computing. In the following, we state the assumption with respect to the Gaussian error distribution [35] and its parameterizations.

Definition 4 (see [36]). Let $n \in \mathbb{N}, q=q(n)>2$, and $s \in \mathbb{Z}_{q}^{n}$ be a secret. Then, $L W E_{s, \xi}$ is a distribution of $\left\langle b, b^{t} s+z\right\rangle$ over $\mathbb{Z}_{q}^{n} \times \mathbb{Z}_{q}$ with $b \in \mathbb{Z}_{q}^{n}$ is an arbitrary random and $z \in \mathbb{Z}_{q}$ is chosen from $\xi$, where $\xi$ is the Gaussian distribution.

Definition 5 (see [37]). The "Learning with Errors" decision problem is to distinguish betweenLWE $E_{s, \xi}$ which is the distribution of $\left\langle b, b^{t} s+z\right\rangle$ over $\mathbb{Z}_{q}^{n} \times \mathbb{Z}_{q}$ with randoms $\in \mathbb{Z}_{q}^{n}$ and the uniform random distribution over $\mathbb{Z}_{q}^{n} \times \mathbb{Z}_{q}$, given access to the random samples from the given distribution.

Regev [37] proved that the decision problem (learning with errors) under a suitable prime modulus $q$ and Gaussian distribution $\xi$ is as hard as solving the worst-case lattice problem, known as "short independent vector problem" and "decision short vector problem" in Euclidean norm, using quantum algorithms. Suppose $\mathbb{R} / \mathbb{Z}=[0,1)$ is a group with respect to modulo one operation. Let $\Phi_{\alpha}$ be the Gaussian distribution on $\mathbb{R} / \mathbb{Z}$ with mean 0 and standard deviation $\alpha / 2 \pi$, under modulo one, where $\alpha>0$ is a real number.

Theorem 1 (see [37]). Let $\alpha=\alpha(n) \in(0,1)$ be a real number, and $q=q(n)>0$ be a prime such that $\alpha q>2 n$ holds. If there exists a quantum algorithm that can solve $L W E_{q, \Phi_{\alpha}}$, there also exists a quantum algorithm to solve "short independent vectors problem" and approximate "decision short vector problem", in Euclidean norm, under the worst-case with in $O(n / \alpha)$ factors.
3.6. Regev's Dual Cryptosystem. If $\Delta$ is a lattice, its dual is the set $\Delta^{*}$ consisting of tuples $z \in L(\Delta)$, that is, a linear span of $\Delta$ such that inner product $\langle z, y\rangle$ is an integer for all $y \in \Delta$. Following the definition, one can easily observe that the dual of $\mathbb{Z}^{n}$ is $\mathbb{Z}^{n}$. The inner product between two $n$-tuples $x$ and $y$ is defined as $\langle x, y\rangle=x^{t} y=\sum_{i=1}^{n} x_{i} y_{i}$, where $x=\left(x_{1}, x_{2}, \ldots\right.$, $\left.x_{n}\right)$ and $y=\left(y_{1}, y_{2}, \ldots, y_{n}\right)$ are tuples with the real entries.

The dual space $\Delta^{*}$ has the same dimension as its primal space $\Delta$, and both are essentially isomorphic to each other. Therefore, a dual space $\Delta^{*}$ lies in the same space as the primal $\Delta$, and not necessarily be a sub-lattice of $\Delta$. The lattice $\Delta^{*}$ contains non-integers even $\Delta$ contains only integers
entries. The dual space is necessarily defined as follows in abstract vector space. If $V \subseteq \mathbb{R}^{n}$ is a vector space, a function $\Psi$ : $V \longrightarrow R$ is called a linear function if it satisfies the following conditions: 1) $\Psi\left(\vartheta_{1}+\vartheta_{2}\right)=\Psi\left(\vartheta_{1}\right)+\Psi\left(\vartheta_{2}\right)$, and 2) $\Psi\left(a \vartheta_{1}\right)=a \Psi\left(\vartheta_{1}\right)$, where $a, b \in \mathbb{R}$ and $\vartheta_{1}, \vartheta_{2} \in V$. The dual space of an abstract vector space $V$ is then the set of all linear functions, where a function $\Psi$ is represented as a tuple $\vartheta \in V$ such that $\Psi(x)=\langle\vartheta, x\rangle$, whereas the dual lattice is considered on the set of integers $\mathbb{Z}$ instead set of reals one $\mathbb{R}$. The dual of lattice $\Delta$ is the collection of linear functions of the forms: $\Psi: V \longrightarrow \mathbb{Z}$ represented as tuples in span $(\Delta)$. Each vector $\vartheta \in \Delta^{*}$ generates a linear function $\Psi_{\vartheta}(x)=\langle\vartheta, x\rangle$ satisfying $\Psi_{9}(\Delta) \subseteq \mathbb{Z}$ and partitions $\Delta$ into the layers as $\Delta=\cup_{i \in \mathbb{Z}}\left\{\varrho \in \mathbb{Z}: \quad \Psi_{9}(\varrho)=i\right\}$, where each layer $\Psi_{\vartheta}^{-1}(i)=\left\{\varrho \in \Delta: \Psi_{9}(\varrho)=i\right\}$ is necessarily a shifted copy of $\Delta \cap \vartheta^{\perp}=\{\varrho \in \Delta:\langle\vartheta, \varrho\rangle=0\}$, that is, a lower dimensional sub-lattice orthogonal to $\vartheta$ with distance between layers $1 /\|\vartheta\|$ implies that the sparser lattice has denser dual and vice-versa. Therefore, the dual of $c \Delta$ is $1 / c \Delta$, where $c>0$ is an arbitrary real.

Under the hard assumption "learning with errors", one can construct a public key cryptosystem under indistinguishable property of pseudo-random tuple $\left\langle b, b^{t} s+z\right\rangle$ from a random sample. The pseudo-random $b^{t} s+z \in \mathbb{Z}_{q}$ is used to mask a bit of the message in Regev's cryptosystem [37]. Furthermore, the dual Regev's cryptosystem consists of three phases: a) D-key-Gen, b) D-Encrypt, and c) D-Decrypt, which are discussed below.

1. D-key-Gen: Let $A \in \mathbb{Z}_{q}^{n \times m}$ be a random matrix, where $m \geq 2 n \log (q), f_{A}: \mathbb{Z}^{m} \longrightarrow \mathbb{Z}^{n}: e \mapsto A e(\bmod$ $q)$. Choose an error $e \leftarrow D_{\mathbb{Z}^{m}, \sigma}$. It then computes its syndrome as $u=f_{A}(e)$, where the secret vector $e \in \mathbb{Z}_{q}^{m}$ and the public key is $u \in \mathbb{Z}_{q}^{n}$.
2. D-Encrypt: Let $b \in\{0,1\}$ be a bit to be encrypted. Choose a random $s \in \mathbb{Z}_{q}^{n}$ with an error scalar $x \in \xi$ and an error vector $y \in \xi^{m}$. It then outputs $c=\left\langle c_{0}, c_{1}\right\rangle$, where $c_{0}=u^{t} s+x+b \cdot\lfloor q / 2\rfloor$ and $c_{1}=A^{t} s+y$.
3. D-De crypt: To perform the decryption on $c=\left\langle c_{0}, c_{1}\right\rangle$ using the secret $e$ under the matrix $A$, this phase computes $b=c_{0}-e^{t} c_{1} \in \mathbb{Z}_{q}$ and outputs 1 if $b$ is closer to $\lfloor q / 2\rfloor$; else, it is 0 .
3.7. Pre-image Samplable Family of Functions. Gentry et al. [21] defined a family of pre-image samplable functions that plays a very important role in the construction of the proposed encryption described in Section 4.

Definition 6. A family of pre-image samplable functions [21] consists of three phases: a) Trap-Gen, b) Sample-Dom, and c) Sample-Pre, which are given below.

- Trap-Gen $\left(1^{\kappa}\right)$ : Trap-Gen takes input as the parameterк, and outputs a pair $\langle A, T\rangle$, whereAis utilized in the functionf $A_{A}: D_{\kappa} \longrightarrow R_{\kappa}$ with recognizable domain $D_{\kappa}$ and range $R_{\kappa}$, andTis a trapdoor for the functionf $A_{A}$.
- Sample-Dom (A): Under function description A, it will sample $x \leftarrow \xi$ over the domain $D_{\kappa}$ in such a way the distribution of $f_{A}(x)$ is uniform over $R_{K}$, and outputs $x$ accordingly.
- Sample-Pre(T,y): Under a trapdoor $T$ and a value $y \in R_{\kappa}$, it will sample an element $x \in D_{\kappa}$ from the distribution $\xi$ under the criteria that $f_{A}(x)=y$, and it then outputs $x$.
3.7.1. Correctness. It is worth noticing that Sample-Dom samples $x \leftarrow \xi$ over the domain $D_{\kappa}$ such that $f_{A}(x)$ follows a uniform distribution over the range $R_{\kappa}$, and Sample-Pre samples $x \in D_{\kappa} \leftarrow \xi$ as in Sample-Dom under condition $f_{A}(x)=y$.
3.7.2. Security. The security of the pre-image samplable functions [21] is discussed below. The samplable functions [21] must satisfy the following properties:

1. One-way without trapdoor: If $\mathscr{A}$ is a probabilistic polynomial time adversary, the advantage $\mathscr{A}\left(1^{\kappa}, A, y\right) \leftarrow f_{A}^{-1}(y) \subset D_{\kappa}$ is negligible, where the advantage is considered over all the possible choices of $A$, the value $y \leftarrow R_{\kappa}$ is random, and $\mathscr{A}$ tosses the coin randomly.
2. Pre-image minimum entropy: If $y \leftarrow R_{\kappa}$, the conditioned minimum entropy of $x \leftarrow$ Sample $-\operatorname{Dom}(A)$ is least under the condition $f_{A}(x)=y$.
3. Collision-free without trapdoor: If $\mathscr{A}$ is a probabilistic polynomial time adversary, the advantage $\mathscr{A}\left(1^{\kappa}, A\right)$ results in the distinct $x, x^{\prime} \leftarrow D_{\kappa}$ with $f_{A}(x)=f_{A}\left(x^{\prime}\right)$ is negligible.

Theorem 2 (see [37]). If $q=p o l y(n)$ is an arbitrary large prime and $m \geq 5 n \log (q)$, there exists a probabilistic polynomial time algorithm [38] that takes input as $1^{n}$, and outputs a matrix $A \in \mathbb{Z}_{q}^{n \times m}$ and a full rank set $S \subset \Delta^{\perp}(A)$, where the distribution corresponding to $A$ is statistically close to a uniform distribution over $\mathbb{Z}_{q}^{n \times m}$ under the length $\|S\| \leq m^{2.5}$.

Another algorithm is known as the sampling Gaussian, denoted by Sample-Gauss, discussed by Gentry et al. [21], plays a very important role in cryptographic construction. The Sample-Gauss ( $B, \sigma, c$ ) uses a random basis $B$ in sampling from the Gaussian distribution centered at $c$ with the standard deviation $\sigma$ over the lattice $\Delta(B)$.

Theorem 3 (see [37]). The probabilistic polynomial time algorithm provided in [21] with inputs as a basisB, a lattice $\Delta(B)$, an appropriate parameter $\sigma \geq\left\|B^{*}\right\| \cdot \omega(\sqrt{\log (m)})$ and arbitraryc $\in \mathbb{R}^{m}$, results in a sample distribution that is statistically close to $D_{\Delta, \sigma, c}$.

The function defined in [21] is a sample pre-image consisting of three phases: a) Trap-Gen, b) Sample-Dom
and c) Sample-Pre. Let $\kappa$ be a security parameter, $n=\Theta(\kappa)$, $q=\operatorname{poly}(n)$ be a large prime, $m \geq 5 n \log (q), L=m^{2.5}$ and Gaussian parameter $\sigma \geq L \omega(\sqrt{\log (m)})$, respectively. Then,

- Trap-Gen $\left(1^{\kappa}, \sigma\right)$ : Under the algorithm in Theorem 3, choose a matrix $A \in \mathbb{Z}_{q}^{n \times m}$ and a trapdoor $T \in \Delta^{\perp}(A)$. Consider $D_{\kappa}=\left\{e \in A \mathbb{Z}_{q}^{m}:\|e\| \leq \sigma \sqrt{m}\right\}$ and $R_{\kappa}=\mathbb{Z}_{q}^{n}$ and $f_{A}: D_{\kappa} \longrightarrow R_{\kappa}$ such that $f_{A}(e)=A e(\bmod q)$. This phase then results $\langle A, T\rangle$.
- Sample-Dom $(A, \sigma)$ : Assuming $B^{\prime}$ as a standard basis for $\mathbb{Z}^{m}$, use Sample-Gauss ( $B^{\prime}, \sigma, 0$ ) to get sample from $D_{\mathbb{Z}^{m}, \sigma}$.
- Sample-Pre $(T, \sigma, y)$ : Let $k \in \mathbb{Z}^{m}$ be an arbitrary number under condition $\mathrm{Ak}=y(\bmod q)$. Then, use Sample-Gauss $(T, \sigma-k)$ [21] to sample $v$ from $D_{\Delta^{\perp}(A), \sigma,-k}$.

Theorem 4 (see [37]). Assume that the columns of $A \in \mathbb{Z}_{q}^{n \times m}$ span $\mathbb{Z}_{q}^{n}, \varepsilon \in(0,1 / 2)$, and $\sigma \geq \eta_{\varepsilon}\left(\Delta^{\perp}(A)\right)$. Then, the syndrome's distribution $u=A e(\bmod q)$ differs by a statistically distance equal to at most $2 \varepsilon$ from the uniform distribution over $\mathbb{Z}_{q}^{n}$.

To prove the correctness of the distribution $\xi=D_{\mathbb{Z}^{m}, \sigma}$, for a given $u \leftarrow \mathbb{Z}_{q}^{n}$ and $k \leftarrow \mathbb{Z}^{m}$ is a solution to $\mathrm{Ak}=u(\bmod q)$, the conditional probability distribution of $e \leftarrow D_{\mathbb{Z}^{m}, \sigma}$ under $A e=u(\bmod q)$ matches perfectly with $k+D_{\Delta^{\perp}(A), \sigma,-k}$. The correctness of the distribution is as follows. It can be observed that $f_{A}(e)=A e(\bmod q)$ is indistinguishable from the uniform distribution over $R_{\kappa}=\mathbb{Z}_{q}^{n}$, assuming the columns of $A \in \mathbb{Z}_{q}^{n \times m}$ spans $\mathbb{Z}_{q}^{n}$ [21] with the probability $1-q^{-n}$. Since $\sigma \geq L \omega(\sqrt{\log (m)})$, and $\|T\| \leq L$, the result in [21] implies $\sigma \geq \eta_{\varepsilon}\left(\Delta^{\perp}(A)\right)$. Thus, as a result, Sample-Pre $(v+k)$ is distributed under $D_{\mathbb{Z}^{m}, \sigma}$ under the condition $A(v+k)=y(\bmod q)$.

In the proof of security, we use the functions described in [21], which are one-way and collision resistant functions. A brief discussion of these two properties are given below.

- One-way without trapdoor: The process of inversion of $f_{A}$ under a uniform random $u \leftarrow R_{n}$ is equivalent to solving "in-homogeneous short integer solution" problem, say ISIS $_{q, m, \sigma \sqrt{m}}$ [21].
- Pre-image minimum entropy: Since all the pre-images follow the discrete Gaussian, it has minimum entropy [21].
- Collision-free without trapdoor: Let $z, z^{\prime} \leftarrow D_{\kappa}$. Then, a collision implies $A\left(z-z^{\prime}\right)=0(\bmod q)$, which actually solves the "short integer solution" problem, say SIS $_{q, m, 2 \sigma \sqrt{m}}$.


## 4. Proposed Identity-Based Encryption (IBE) Scheme in Standard Model

In this section, we propose a new provably secure identitybased encryption scheme. Note that such a scheme has a compact public key and also achieves adaptive security in the standard model [39].

Our proposed identity-based encryption scheme consists of four phases: a) Set-up, b) Extraction, c) Encrypts and d) Decrypts. We take an identity ID as an arbitrary $k$-bits string $\{0,1\}^{k}$, where $k=\Theta(\kappa)$ for a given security parameter $\kappa$. In the following, we now discuss the details of these four phases.
4.1. Set-up Phase. It includes the function Set-up ( $1^{\kappa}$ ). First, choose a suitable large prime $q$, a smoothing parameter $\sigma$ depending on the security parameter $\kappa$ and an arbitrary random matrix $A \in \mathbb{Z}_{q}^{n \times m}$, under a short basis for $\Delta^{\perp}(A)$, that is, $T_{A}$ with the help of Ajtai's construction [38]. Let $f_{A}$ : $\mathbb{Z}_{q}^{m} \longrightarrow \mathbb{Z}_{q}^{n}$ be a function defined as $f_{A}(e)=A e(\bmod q)$. Next, pick a tuple $u_{0} \in \mathbb{Z}_{q}^{n}$ and a random matrix $H_{i, b} \in \mathbb{Z}_{q}^{n \times \ell}$, where $\ell=m$ and $\hat{H}=\left\{\left\langle i, b, H_{i, b}\right\rangle: 1 \leq i \leq 2,0 \leq b \leq 1\right\}$ is the ordered set. The public parameters are $\left\{A, u_{0}, \hat{H}\right\}$, whereas $T_{A}$ is considered as the master secret.
4.2. Extraction Phase. This phase is accomplished by the function Extraction $\left(A, u_{0}, \widehat{H}, \mathrm{ID}, T_{A}\right)$. A decryption key is extracted related to the identity ID $\in\{0,1\}^{k}$ under the master secret $T_{A}$ as the trapdoor. The following steps need to be executed:

- Let $S=H$ (ID) and $U$ be the set of non-zero positions in the string $S$. After that, assemble an $n \times \ell$ matrix $H_{\text {ID }}=\left[H_{i_{1} \bmod 2, b i_{1} \bmod 2}\left|H_{i_{2} \bmod 2, b i_{2} \bmod 2}\right| \cdots \mid H_{i_{\ell} \bmod 2, b i_{\ell}}\right.$ $\bmod 2] \in \mathbb{Z}_{q}^{n \times \ell}, \quad$ where $\quad H_{i_{1} \bmod 2, b i_{1} \bmod 2}\left|H_{i_{2} \bmod 2, b i_{2} \bmod 2}\right|$ $\cdots \mid H_{i_{\ell} \bmod 2, b i_{\ell} \bmod 2} \in \widehat{H}$ as $H_{i_{1}, b 1}$ or $H_{i_{1}, b 0}$ is according to either $i_{1}(\bmod 2)=0$ or $i_{1}(\bmod 2)=1$, respectively.
- Now, sample $r_{i} \leftarrow \mathbb{Z}_{q}^{\ell}$ under Sample-Dom $\left(H_{i, b i \bmod 2}, \sigma\right)$, where $1 \leq i \leq k$, and consider $r \leftarrow \mathbb{Z}_{q}^{\ell}$ such that $r^{t}=\left[r_{1}^{t}\left|r_{2}^{t}\right| \cdots \mid-r_{\ell}^{t}\right]$.
- Let $u=u_{0}+H_{\mathrm{ID}} r \in \mathbb{Z}_{q}^{n}$. It can be observed as $u=u_{0}+\sum_{i=1}^{\ell^{\prime}} H_{i \bmod 2, b_{i(\bmod 2)}} r_{i}$, where $i$ is the non-zero position in the string $S$.
- Next, apply the Sample-Pre $\left(T_{A}, \sigma, u\right)$ under the trapdoor $T_{A}$ to find the pre-image $e$ of $u$ satisfying $u=A e(\bmod q)$, and outputs the private key $\langle e, r\rangle$.
4.3. Encrypts Phase. In this phase, we involve the function Encrypts ( $A, u_{0}, \hat{H}, \mathrm{ID}, b$ ). In order to process the encryption on a bit $b \in\{0,1\}$ under the identity ID $\in\{0,1\}^{k}$ using the master key $T_{A}$, the following steps are necessary:
- Let $H_{\mathrm{ID}}=\left[H_{i_{1} \bmod 2, b i_{1} \bmod 2}\left|H_{i_{2} \bmod 2, b i_{2} \bmod 2}\right| \cdots \mid H_{i_{\ell} \bmod 2, b i_{e}}\right.$ $\bmod 2] \in \mathbb{Z}_{q}^{n \times \ell}$, where $H_{i_{1} \bmod 2, b i_{1} \bmod 2}\left|H_{i_{2} \bmod 2, b i_{2} \bmod 2}\right|$ $\cdots \mid H_{i_{e} \bmod 2, b i_{e} \bmod 2} \in \hat{H}$ because $H_{i_{1} \bmod 2, b 1}$ or $H_{i_{1} \bmod 2, b 0}$ is based on either $i_{1}(\bmod 2)=0$ or $i_{1}(\bmod 2)=1$.
- Choose an arbitrary $s \in \mathbb{Z}_{q}^{n}$.
- Pick $x \in \mathbb{Z}_{q}, y=\left\langle y_{1}, y_{2}, \ldots, y_{m}\right\rangle \in \mathbb{Z}_{q}^{m}$ and $z=\left\langle z_{1}\right.$, $\left.z_{2}, \ldots, z^{\ell}\right\rangle \in \mathbb{Z}_{q}^{\ell}$ which are sampled from the distributions $\xi, \xi^{m}$, and $\xi^{\ell}$, respectively, based on the Regev's cryptosystem.


Figure 4: Network model for IoT-enabled AI applications.

- Now, calculate $c_{0}=u_{0}^{t} s+x+b\lfloor q / 2\rfloor \in \mathbb{Z}_{q}, c_{1}=A^{t} s+$ $y \in \mathbb{Z}_{q}^{m}$ and $c_{2}=H_{\mathrm{ID}}^{t} s+z \in \mathbb{Z}_{q}^{\ell}$.
- Finally, the initiator sends the output as the cipher $c=$ $\left\langle c_{0}, c_{1}, c_{2}\right\rangle$ to the responder.
4.4. Decrypts Phase. This phase is implemented by the function Decrypts $\left(A, u_{0}, \widehat{H}, \mathrm{ID}, k, c\right)$. After receiving the cipher $c=\left\langle c_{0}, c_{1}, c_{2}\right\rangle$, with the private key $\langle e, r\rangle \in \mathbb{Z}_{q}^{m+\ell}$, the responder executes the following steps:
- Compute $v=c_{0}-e^{t} c_{1}+r^{t} c_{2} \in \mathbb{Z}_{q}$, and then compare $v$ with $\lfloor q / 2\rfloor$ in $\mathbb{Z}$.
- If $|v-q / 2| \leq q / 4$, it results bit $b=1$; else, it outputs the bit $b=0$.


## 5. Incorporating Proposed IBE Scheme for IoTEnabled AI Applications

In this section, we first discuss the network model for IoTenabled AI applications, which is used for incorporating our proposed IBE scheme described in Section 4. Next, we describe the various phases where the proposed IBE scheme has been applied for IoT.
5.1. Network Model. The network model considered for IoTenabled AI applications using our proposed IBE scheme is presented in Figure 4. The model expresses various applications of IoT, such as traffic monitoring, smart home, and IIoT. In this model, different types of smart sensors, say $\left\{\mathrm{SS}_{i} \mid i=1,2,3, \ldots, n_{\mathrm{ss}}\right\}$ are connected with each other via the nearby gateway node(s) $\left\{\mathrm{GWN}_{j} \mid j=1,2,3, \ldots, n_{\mathrm{gwn}}\right\}$, where $n_{\text {ss }}$ and $n_{\text {gwn }}$ denote the number of smart sensors and gateway nodes to be deployed for each IoT application, respectively. Note that there might be multiple nodes that
are connected with a particular application and the gateways $\mathrm{GWN}_{j}$ are further connected with the cloud server(s), say $\left\{\mathrm{CLS}_{k} \mid k=1,2,3, \ldots, n_{\mathrm{cls}}\right\}$, where $n_{\mathrm{cls}}$ is the number of cloud servers. Before initiating any secure communications between $\mathrm{GWN}_{j}$ and CLS ${ }_{k}$, they need to complete their registration process which is performed by a fully-trusted registration authority (RA). Similarly, the RA also performs the registration of each smart sensor node to be deployed in various IoT applications. Next, a gateway node needs to perform the secure data aggregation where the data is collected through secure communication among the smart sensors and the gateway node. In this case, we apply the proposed IBE scheme for encryption/decryption of the data. After that the gateway nodes send the data securely to the cloud server(s) for secure data storage purpose. Finally, the cloud servers $\mathrm{CLS}_{k}$ can perform the Big data analytics using $\mathrm{AI} / \mathrm{ML}$ techniques with the data stored at $\mathrm{CLS}_{k}$.
5.2. Description of Various Phases. We have the following phases:

- In the pre-deployment of IoT devices phase, the trusted RA will perform the registration of each IoT smart device prior to their deployment in respective application. After deployment of the IoT devices, they need to communicate with their nearby gateway node(s). For avoiding various attacks by an adversary, we use the proposed IBE scheme for secure data transfer among the sensor nodes and their gateway node(s).
- In the registration of gateway nodes and cloud servers phase, the RA, RA also performs the registration of the deployed gateway nodes and cloud servers. For secure communication, we again use the proposed IBE scheme for secure data transfer among the gateway nodes and the cloud servers.


Figure 5: High-level overview of various phases.

- The secure data aggregation at gateway phase allows a gateway node to collect the data from its associated IoT smart devices securely using the proposed IBE scheme.
- The secure data storage at cloud servers phase permits storage of data at the cloud servers securely from the gateway nodes with the help of the proposed IBE scheme.
- Finally, the Big data analytics using AI phase is needed because the cloud servers store a huge volume of data from various IoT applications. Since the Big data analytics provides numerous advantages, such as better decision making and preventing fraudulent activities, it is preferable to do the Big data analytics on the data stored at the cloud servers.

A high-level description of various phases related to IoTenabled AI applications is given in Figure 5.
5.2.1. Pre-deployment of IoT Devices. Before deploying the IoT smart devices (smart sensors) $\mathrm{SS}_{i}$ in their respective application, the trusted RA, RA executes the Set-up phase described in Section 4.1 in order to select the system parameters. The steps are as follows:

- Step 1. The selected public parameters are $\left\{A, u_{0}, \widehat{H}\right\}$, whereas $T_{A}$ is as the master secret.
- Step 2. For each $\mathrm{SS}_{i}$, the RA, $R A$ assigns a unique identity $\mathrm{ID}_{\mathrm{SS}_{i}}$.
- Step 3. Next, for each $\mathrm{SS}_{i}$, the RA, RA executes the Extraction phase described in Section 4.2 to extract a decryption key related to $\mathrm{ID}_{\mathrm{SS}_{i}}$ under the trapdoor master secret $T_{A}$. The private key for $\mathrm{SS}_{i}$ is considered as $\left(e_{\mathrm{SS}_{i}}, r_{\mathrm{SS}_{i}}\right)$.
5.2.2. Registration of Gateway Nodes and Cloud Servers. The registration process for the deployed gateway nodes $\mathrm{GWN}_{j}$ and cloud servers $\mathrm{CLS}_{k}$ is also based on the execution of the Set-up phase, where the public parameters are
$\left\{A, u_{0}, \widehat{H}\right\}$, and $T_{A}$ is the trapdoor master secret. This phase involves the following steps:
- Step 1. For each $\mathrm{GWN}_{j}$, the RA, $R A$ assigns a unique identity $\mathrm{ID}_{\mathrm{GWN}_{j}}$. In a similar way, for each $\mathrm{CLS}_{k}$, the RA, $R A$ also assigns a unique identity $\mathrm{ID}_{\mathrm{CLS}_{k}}$.
- Step 2. For each $\mathrm{GWN}_{j}$ and $\mathrm{CLS}_{k}$, the RA, $R A$ executes the Extraction phase. After executing this process, the private keys for $\mathrm{GWN}_{j}$ and $\mathrm{CLS}_{k}$ are selected as $\left(e_{\mathrm{GWN}_{j}}, r_{\mathrm{GWN}_{j}}\right)$ and $\left(e_{\mathrm{CLS}_{k}}, r_{\mathrm{CLS}_{k}}\right)$, respectively.
5.2.3. Secure Data Aggregation at Gateway. In this phase, the following steps are involved:
- Step 1. Suppose the IoT smart sensors $\mathrm{SS}_{i}$ are deployed in their respective IoT applications as shown in Figure 4. The gateway nodes $\mathrm{GWN}_{j}$ and cloud servers $\mathrm{CLS}_{k}$ are also placed accordingly in the network. Let a smart sensor $\mathrm{SS}_{i}$ sense the information (data), say Data $_{\mathrm{Ss}_{i}}$ from its deployment area and want to communicate it securely with its gateway node $\mathrm{GWN}_{j}$, $G W N_{j}$. For this purpose, the $S S_{i}, S S_{i}$ generates a current timestamp, say $\mathrm{TS}_{\mathrm{SS}_{i}}$, prepares a message of the type $\mathrm{Msg}_{\mathrm{Ss}_{i}}=\left\{\mathrm{ID}_{\mathrm{SS}_{i}}, \mathrm{TS}_{\mathrm{SS}_{i},}\right.$, Data $\left._{\mathrm{SS}_{i}}\right\}$ and encrypts $\mathrm{Msg}_{\text {ss }_{j}}$ bit wise using the public parameters, identity of $\mathrm{GWN}_{j}, G W N_{j}$ and trapdoor master key $T_{A}$ to create the ciphertext $C_{\mathrm{SS}_{i}}=\left\{C_{0}, C_{1}, C_{2}\right\}$ as done in the Encrypts phase described in Section 4.3, where $C_{0}, C_{1}$ and $C_{2}$ are the encrypted bit strings corresponding to the bit strings of the $\mathrm{Msg}_{\mathrm{Ss}_{i}}$. Next, $\mathrm{SS}_{i}$ sends the encrypted message $\left\{C_{\mathrm{SS}_{i}}, \mathrm{TS}_{\mathrm{SS}_{i}}\right\}$ to its destination $\mathrm{GWN}_{j}, G W N_{j}$ via a public channel.
5.2.4. Secure Data Storage at Cloud Servers. In this phase, a cloud server $\mathrm{CLS}_{k}, C L S_{k}$ receives the encrypted data from the respective gateway nodes $\mathrm{GWN}_{j}$ residing in an IoT application, and stores the encrypted data in its database for further processing. In order to do this, the following steps are executed by the $\mathrm{CLS}_{k}, C L S_{k}$ :


Steps 1, 2, 3: Pre-deployment \& registration
Step 4: Secure data aggregation
Steps 5, 6: Secure data storage

Figure 6: Overall mechanism of the proposed IBE scheme for IoT-based AI applications.

- Step 1. Once the message $\left\{C_{\mathrm{GWN}_{j}}, \mathrm{TS}_{\mathrm{GWN}_{j}}\right\}$, is received at time $\mathrm{TS}_{\mathrm{GWN}_{j}}^{*}$, for checking replaying attacks, $\mathrm{CLS}_{k}, C L S_{k}$ checks the validity of the received timestamp by the condition: $\left|\mathrm{TS}_{\mathrm{GWN}_{j}}-\mathrm{TS}_{\mathrm{GWN}_{j}}^{*}\right|<\Delta T$. If the condition fails, the process is immediately terminated.
- If the timestamp validation is satisfied, the encrypted data $C_{\mathrm{GWN}_{j}}$ is then stored in the database of $\mathrm{CLS}_{k}$, $C L S_{k}$.
5.2.5. Big Data Analytics using AI. It is worth noticing that a cloud server $\mathrm{CLS}_{k}, C L S_{k}$ receives the encrypted data generated by the IoT smart sensors residing in various applications via the aggregator nodes (gateway nodes). $\mathrm{CLS}_{k}$, $C L S_{k}$ can then decrypt the stored data bit wise using its own private key ( $e_{\mathrm{CLS}_{k}}, r_{\mathrm{CLS}_{k}}$ ) and performs the Big Data analytics steps using AI/ML techniques, such as "data acquisition and filtering", "data extraction", "data aggregation and representation", "data analysis" as well as "data visualization". The results of this phase will provide some useful conclusions and predictions on the stored data.

The overall mechanism of the proposed IBE scheme for IoT-based AI applications is also illustrated in Figure 6. The pre-deployment and registration phases are performed through the steps 1,2 and 3 . Step 4 explains about the data aggregation phase. While the steps 5 and 6 are about secure data storage, Step 7 explains the Big data analytics using the AI techniques.

- Step 2. After receiving the message from $\mathrm{SS}_{i}, \mathrm{GWN}_{j}$, $G W N_{j}$ first checks the validity of the received timestamp by the condition: $\left|T S_{\mathrm{SS}_{i}}-T S_{\mathrm{SS}_{i}}^{*}\right|<\Delta T$, where $\mathrm{TS}_{\mathrm{SS}_{i}}^{*}$ and $\Delta T$ represent the time when the message was received and the maximum transmission delay, respectively. If the condition is satisfied, $G_{W N}{ }_{j}, G W N_{j}$ proceeds to decrypt $C_{\mathrm{SS}_{i}}$ bit wise using its private (secret) key ( $e_{\mathrm{GWN}_{j}}, r_{\mathrm{GWN}_{j}}$ ) with the help of the Decrypts phase described in Section 4.4 to obtain $\left\{\mathrm{ID}_{\mathrm{SS}_{i}}, T S_{\mathrm{SS}_{i}}{ }^{\prime}\right.$, Data $\left._{\mathrm{SS}_{i}}\right\}$. After that if the checking condition: $\mathrm{TS}_{\mathrm{SS}_{j}}^{\prime}=\mathrm{TS}_{\mathrm{SS}_{i}}$ is valid, $\mathrm{GWN}_{j}, G W N_{j}$ considers the data is fresh. Thus, no replay attack has been there
during this process with the timestamping mechanism. Of course, for this purpose, it is reasonable to assume that the network entities are synchronized with their clocks [8].
- Step 3. Now, $\mathrm{GWN}_{j}, G W N_{j}$ generates a current timestamp $\mathrm{TS}_{\mathrm{GWN}_{j}}$, encrypts the prepared message $\operatorname{Msg}_{\mathrm{GWN}_{j}}=\left\{\mathrm{Msg}_{\mathrm{SS}_{i}}, \mathrm{ID}_{\mathrm{GWN}_{j}}, \mathrm{TS}_{\mathrm{GWN}_{j}}\right\}=\left\{\left(\mathrm{ID}_{\mathrm{SS}_{i}}, \mathrm{TS}_{\mathrm{SS}_{i}}\right.\right.$, Data $\left.\left._{\text {SS }_{i}}\right), \operatorname{ID~GWN}{ }_{j}, \mathrm{TS}_{\mathrm{GWN}_{j}}\right\}$ bit wise using the public key of its corresponding cloud server $\mathrm{CLS}_{k}$ to obtain the ciphertext $C_{\mathrm{GWN}_{j}}=\left\{C_{0}^{\prime}, C_{1}^{\prime}, C_{2}^{\prime}\right\}$ as done in the Encrypts phase, and sends the encrypted message $\left\{C_{\mathrm{GWN}_{j}}, \mathrm{TS}_{\mathrm{GWN}_{j}}\right\}$ to its respective $\mathrm{CLS}_{k}, C L S_{k}$ via a public channel, where $C_{0}^{\prime}, C_{1}^{\prime}$ and $C_{2}^{\prime}$ are the encrypted bit strings corresponding to the bit strings of the $\mathrm{Msg}_{\mathrm{GWN}_{j}}$.


## 6. Security Analysis

In this section, we analyze the security of the proposed encryption scheme by using a sequence of games played between an adversary, say $\mathscr{A}$ and a challenger, say $\mathscr{B}$, namely the games $\mathscr{G}_{l}$, for $l=0,1,2,3,4$. The initial game $\mathscr{G}_{0}$ is considered as the real attack, whereas the final game $\mathscr{G}_{4}$ is the game that cannot be cracked by the adversary $\mathscr{A}$. Each transition from a game $\mathscr{G}_{i}$ to another game $\mathscr{G}_{i+1}$ is indistinguishable with a negligible advantage under some hard assumption. If there are polynomial time games, each of the transitions is also indistinguishable with the negligible advantage meaning that the advantage of $\mathscr{A}$ in real attack is negligible. We now define the games in order to ensure the indistinguishable transitions.
6.1. Games Descriptions. The following games are discussed below.

- $\operatorname{Game}\left(\mathscr{G}_{0}\right)$ : This game is played between the adversary $\mathscr{A}$ and the challenger $\mathscr{B}$ with both honest and indistinguishable properties under the IND-sIDCPA property. We have defined as earlier that, under "selective identity chosen plaintext attack

IND-sID-CPA" property, $\mathscr{A}$ needs to submit target identity at advance to the $\mathscr{B}$, before $\mathscr{B}$ runs the Set-up algorithm.

- $\operatorname{Game}\left(\mathscr{G}_{1}\right)$ : This game is same as $\mathscr{G}_{0}$ except in the Setup phase, $\mathscr{B}$ computes the matrices $H_{i, b}$, for $1 \leq i \leq k$ and $b \in\{0,1\}$ not directly, but as an arbitrary public key of random GPV trapdoors [21] corresponding to the trapdoor $T_{i, b}$.
- $\operatorname{Game}\left(\mathscr{G}_{2}\right)$ : This game is same as $\mathscr{G}_{1}$, except $\mathscr{B}$ neither uses the master secret $T_{A}$ nor the Extraction phase to answer the queries to private keys, but it uses another Trapdoor-Extraction phase and trapdoors $T_{i, b}$ for $1 \leq i \leq k$ and $b \in\{0,1\}$. The trapdoors are represented as $\widehat{T}=\left\{\left\langle i, b, T_{i, b}\right\rangle: 1 \leq i \leq k, 0 \leq b \leq 1\right\}$.
Trapdoor-Extraction $\left\langle A, u_{0}, \hat{H}, \mathrm{ID}, i^{*}, T_{i^{*}, b^{*}}\right\rangle$ : A key that corresponds to decryption is extracted for the identity ID, with the help of the trapdoor:

1. Let $b_{i}=\mathrm{bit}_{i(\bmod 2)}$ (ID) be the position of non zero bit for $1 \leq i \leq k$ and $b \in\{0,1\}$. Assemble an $n \times \ell$ matrix $H_{\mathrm{ID}}=\left[H_{i_{1}, b i_{1} \bmod 2}\left|H_{i_{2}, b i_{2} \bmod 2}\right| \cdots, \mid H_{i_{e}, b i_{\ell} \bmod 2}\right] \in \widehat{\mathbb{Z}_{q}^{n \times \ell}}$, where $H_{i_{1}, b i_{1} \bmod 2}\left|H_{i_{2}, b i_{2} \bmod 2}\right| \cdots, \mid H_{i_{e}, b i_{\ell} \bmod 2} \in \widehat{H}$ because $H_{i_{1}, b 1}$ or $H_{i_{1}, b 0}$ is according to either $i_{1}(\bmod 2)=$ 0 or $i_{1}(\bmod 2)=1$.
2. Sample $r_{i} \leftarrow \mathbb{Z}_{q}^{\ell}$ under Sample-Dom $\left(H_{i, b i \bmod 2}, \sigma\right)$, where $i \in\left\{1,2, \ldots,(k\}-\left\{i^{*}\right\}\right.$, that is, from the set $\mathbb{Z}_{q}^{\ell}$.
3. Let $\widehat{u}=u_{0}+H_{\mathrm{ID}} \widehat{r} \in \mathbb{Z}_{q}^{n}$. It can be then observed as $\widehat{u}=u_{0}+\sum_{i \in\{1,2, \ldots, k\}-\left\{i^{*}\right\}}^{\ell^{\prime}} H_{i, b_{i(\bmod 2)}} r_{i}$, where $i$ is the non zero position in the string $S=H$ (ID), and $\widehat{r}$ is the concatenation of all $r_{i}$ s, except 0 , which follows the distribution $r_{i^{*}}$.
4. Using the distribution $D_{\mathbb{Z}^{m}, \sigma}$, sample $e \leftarrow \mathbb{Z}_{q}^{m}$ under the Sample-Dom $(A, \sigma)$ algorithm.
5. Compute $u=A e \in \mathbb{Z}_{q}^{n}$ and $\vec{u}=u-\widehat{u}$, and then use the Sample-Pre $\left(T_{i^{*}, b^{*}}, \sigma, \widehat{u}\right)$ to sample $r_{i^{*}} \leftarrow \mathbb{Z}_{q}^{\ell}$ such that $\widehat{u}=H_{i^{*}, b^{*}} r_{i^{*}}$.
6. Let $r^{t}=\left[r_{1}^{t}\left|r_{2}^{t}\right| \cdots \mid r_{\ell}^{t}\right] \quad$ including $r_{i^{*}}$ such that $u=u_{0}+H_{\text {ID }} r$. Output a private key $K=\langle e, r\rangle$.

- $\operatorname{Game}\left(\mathscr{G}_{3}\right):$ This game is same as $\mathscr{G}_{2}$, except $\mathscr{B}$ computes $\widetilde{H}$ with the trapdoors $\widetilde{T}$. It knows only the trapdoor of $i^{\text {th }}$ index, but not corresponding to $i^{\text {th }}$, $i^{\text {th }}$-bit of the target ID*.

1. Let $\operatorname{bit}_{i(\bmod 2)}\left(\mathrm{ID}^{*}\right)$ for $i \in\{1,2, \ldots,(k\}$, be the modulo of non-zero $i^{\text {th }}, i^{\text {th }}$ position declared by $\mathscr{A}$ to $\mathscr{B}$ in the Set-up phase.
2. $\mathscr{B}$ generates $\tilde{H}$ by taking $b \in\{0,1\}, i \in\{1,2, \ldots,(k\}$ such that $b \neq \operatorname{bit}_{i(\bmod 2)}\left(\right.$ ID* $\left.^{*}\right)$, and executes GPV trapdoors [21] as in the $\mathscr{G}_{2}$ to obtain $H_{i, b}$ corresponding to $T_{i, b}$. Furthermore, it takes $b \in\{0,1\}$, $b \in\{0,1\}$ such that $b=\operatorname{bit}_{i(\bmod 2)}\left(\mathrm{ID}^{*}\right)$ for a random $i \in\left\{1, i \in\left\{1,2, \ldots,(k\},(k\}\right.\right.$, and takes $H_{i, b} \in \mathbb{Z}_{q}^{n \times \ell}$ with $T_{i, b}=\perp$.
3. To extract the private key for ID $\neq \mathrm{ID}^{*}, \mathscr{B}$ repeats the game $\mathscr{G}_{2}$, except $i^{*}$ is picked such that $\operatorname{bit}_{i(\bmod 2)}(\mathrm{ID}) \neq \operatorname{bit}_{i(\bmod 2)}\left(\mathrm{ID}^{*}\right)$ and $i^{*}$ corresponding to a legal query. If $b^{*}=\operatorname{bit}_{i^{*}(\bmod 2)}$ (ID) and $T_{i^{*}, b^{*}} \neq \perp$,
$\mathscr{B}$ executes Trapdoor-Extraction $\left\langle A, u_{0}, \hat{H}, \mathrm{ID}, i^{*}\right.$, $\left.T_{i^{*}, b^{*}}\right\rangle$ to generate the private key.

The challenge cipher then is generated by Encrypts ( $A$, $H_{0}, \vec{H}, \mathrm{ID}^{*}, b^{*}$ ) for an arbitrary $b^{*} \in\{0,1\}$, and outputs $c=\left\langle c_{0}, c_{1}, c_{2}\right\rangle$ as the challenge.

- $\operatorname{Game}\left(\mathscr{G}_{4}\right)$ : This game is also same as $\mathscr{G}_{3}$, except $\mathscr{B}$ gives a challenge to $\mathscr{A}$ that is not computed honestly, but it is a random cipher, that is, $c=\left\langle c_{0}, c_{1}, c_{2}\right\rangle$ is chosen randomly from $\mathbb{Z}_{q}^{1+m+\ell}$ distribution.
6.2. Games Transitions. In the following, we now show that each of the transitions between the successive games
$\left(\operatorname{Game}\left(\mathscr{G}_{0}\right), \quad \operatorname{Game}\left(\mathscr{G}_{1}\right), \quad \operatorname{Game}\left(\mathscr{G}_{2}\right), \quad \operatorname{Game}\left(\mathscr{G}_{3}\right)\right.$, $\left.\operatorname{Game}\left(\mathscr{G}_{4}\right)\right)$ is indistinguishable as follows.
- Transition: $\operatorname{Game}\left(\mathscr{G}_{0}\right),\left(\mathscr{G}_{0}\right) \longrightarrow \operatorname{Game}\left(\mathscr{G}_{1}\right),\left(\mathscr{G}_{1}\right)$ : Both games are identical with respect to $\mathscr{A}$, and $\mathscr{B}$ possesses the information regarding trapdoor $T_{i, b}$ corresponding to $H_{i, b}$ which is not known to $\mathscr{A}$.
- Transition: $\operatorname{Game}\left(\mathscr{G}_{1}\right),\left(\mathscr{G}_{1}\right) \longrightarrow \operatorname{Game}\left(\mathscr{G}_{2}\right),\left(\mathscr{G}_{2}\right)$ : Both games are identical with respect to $\mathscr{A}$, and $\mathscr{B}$ possesses a different algorithm for key extraction and it is invisible to $\mathscr{A}$.
- Transition: $\operatorname{Game}\left(\mathscr{G}_{2}\right),\left(\mathscr{G}_{2}\right) \longrightarrow \operatorname{Game}\left(\mathscr{G}_{3}\right),\left(\mathscr{G}_{3}\right)$ : Both games are identical with respect to $\mathscr{A}$, and $\mathscr{B}$ knows only half of all the hash-trapdoors and answers if the extraction queries are known, and these are invisible to $\mathscr{A}$.
- Transition: $\operatorname{Game}\left(\mathscr{G}_{3}\right),\left(\mathscr{G}_{3}\right) \longrightarrow \operatorname{Game}\left(\mathscr{G}_{4}\right),\left(\mathscr{G}_{4}\right)$ : The views are not identical with respect to $\mathscr{A}$, but are indistinguishable under "learning with errors" assumption.

1. In the beginning, $\mathscr{B}$ receives $1+m+\ell$ samples of "learning with errors" assumption $\left\langle a_{j}, b_{j}\right\rangle \in \mathbb{Z}_{q}^{n+1}$, for $1 \leq j \leq 1+m+\ell$, with random $a_{j} \in \mathbb{Z}_{q}^{n}$, and either $b_{j}$ for $1 \leq j \leq 1+m+\ell$ are random or $b_{j}=$ $a_{j}^{t} s+x_{j}$ for $1 \leq j \leq 1+m+\ell$ with a random $s \in \mathbb{Z}_{q}^{n}$ and Gaussian $x_{j} \leftarrow \xi$.
2. In the beginning, $\mathscr{B}$ also receives ID* from $\mathscr{A}$ to be challenged. By applying the Set-up phase, $\mathscr{B}$ computes $\vec{H}$. $\mathscr{B}$ picks $b \in\{0,1\}$ such that $b \neq \mathrm{bit}_{i(\mathrm{mod} 2)}$ (ID*) for $1 \leq i \leq k$, and executes GPV trapdoors [21] as in $\mathscr{G}_{2}$ to obtain random $H_{i, b}$ and its trapdoor $T_{i, b}$ as in another $\mathscr{G}_{3}$ and $\mathscr{G}_{4}$, respectively. Now, $\mathscr{B}$ picks $b \in\{0,1\}$ such that $b=\operatorname{bit}_{i(\bmod 2)}\left(\right.$ ID $\left.^{*}\right)$ for $1 \leq i \leq k$, random $H_{i, b}$ and its $j^{\text {th }}$-column "learning with errors" instance $a_{j}$, and then sets $T_{i, b}=\perp$.
3. $\mathscr{B}$ answers the private key queries as in the games $\mathscr{G}_{3}$ and $\mathscr{G}_{4}$ using the corresponding trapdoors. $\mathscr{B}$ picks random $b \in\{0,1\}$ and computes a challenge cipher $c_{0}^{*}=b_{1}+b\lfloor q / 2\rfloor, c_{1}^{*}=\left\langle b_{1+i}: 1 \leq i \leq m\right\rangle \in \mathbb{Z}_{q}^{m}$, $c_{2}^{*}=\left\langle b_{1+m+i}: 1 \leq i \leq \ell\right\rangle \in \mathbb{Z}_{q}^{\ell}$.
4. Finally, $\mathscr{A}$ guesses a bit $b^{*} \in\{0,1\}$, and $\mathscr{B}$ returns the correct $b^{*}=b$; else, returns a random bit as an answer to the "learning with errors" instances.

Table 1: A comparative study on recommended bit-size: Lattice versus classical discrete logarithm.

| Protocol | Primitive | Recommended bit- <br> size |
| :--- | :---: | :---: |
| DLP storage | $g \in \in \mathbb{Z}_{p}^{*}$ | $k^{\prime}=\log ^{2} p$ |
| Lattice storage | $A \in \mathbb{Z}_{q}^{m \times n}, \widehat{H}, u_{0} \in \mathbb{Z}_{q}^{n}$ | $16 \kappa^{2} \log ^{3} \kappa$ |
| DLP | $g^{r}, P g^{r x} \in \mathbb{Z}_{p}^{*}$ | $k^{\prime}=\log p$ |
| communication |  |  |
| Lattice |  |  |
| communication | $\left.u_{0}^{t} s+x+b \mid q / 2\right\rfloor \in \mathbb{Z}_{q}$, | $8 \log ^{2} \kappa$ |
|  | $A^{t} s+y \in \mathbb{Z}_{q}^{m}$, |  |
|  | $H_{\mathrm{ID}}^{t} s+z \in \mathbb{Z}_{q}^{e}$ |  |

It is thus worth noticing that $\mathscr{B}$ is indistinguishable in both the games $\mathscr{G}_{3}$ and $\mathscr{G}_{4}$ with respect to view of $\mathscr{A}$, excluding the challenge cipher. The "learning with errors" instance is random for the challenge cipher and components of $c^{*}$ has same distribution as in the game $\mathscr{G}_{3}$, and so they will be the components in $\mathscr{G}_{4}$.
6.3. Anonymous Cipher and Indistinguishablity. In this section, we discuss the notion of semantic security that is discussed in Section 3. It is observed that the proposed identity-based encryption scheme provides indistinguishable property of the ciphers from random strings of equal lengths, although an adversary can presume the identity of the receiver. The challenge cipher is then pseudo-random under the "learning with errors" assumption, which implies indistinguishability.

## 7. Performance Comparison

This section provides computation costs and recommended bit-size of the proposed identity-based encryption scheme and compares them with the other relevant approaches, such as discrete logarithm-based schemes, RSA public key cryptosystem [31] and ElGamal cryptosystem [32].
7.1. Comparison on Recommended Bit-size. Let $\kappa$ be an appropriate security parameter and $O\left(\kappa^{2}\right)$ be the size of public key. We can then relate the computation time in terms of security parameter complexity $O\left(\kappa^{2}\right), O\left(\kappa^{2}\right)$. It can be compared with the size of classic public key cryptosystems (RSA and ElGamal) which is $O(\kappa)$ and computation time in terms of security parameter as $O\left(\kappa^{3}\right)$ [33, 36].

We take $q=O\left(\kappa^{2}\right), \quad m=O(\kappa \log (q)) \quad$ and $n=O(\kappa \log (q))$ as the parameters, where $\kappa$ is the security parameter. Furthermore, we consider $q=\kappa^{2}, m=\kappa \log (q)$ and $n=\kappa \log (q)$ as the parameters to simplify the computation. The storage cost is $m n \log (q)=16 \kappa^{2} \log ^{3}(q)$ and the communication cost is $8 \kappa \log ^{2}(\kappa)$ in the proposed scheme. The cipher is computed as $c_{0}=u_{0}^{t} s+x+b\lfloor q / 2\rfloor \in \in \mathbb{Z}_{q}$, $c_{1}=A^{t} s+y \in \mathbb{Z}_{q}^{m}$ and $c_{3}=H_{\mathrm{ID}}^{t} s+z \in \mathbb{Z}_{q}^{\ell}$, that is, in the form of triplet $c=\left\langle c_{0}, c_{1}, c_{2}\right\rangle$. The size of public keys involves the security parameters $A \in \mathbb{Z}_{q}^{m \times n}, \hat{H}, u_{0} \in \mathbb{Z}_{q}^{n}$, which is roughly $16 \kappa^{2} \log ^{3}(\kappa)$, that is, $O\left(m^{2} \log (q)\right)=O\left(\kappa^{2} \log (\kappa)\right)$. In Table 1, a comparative study on recommended bit-size

Table 2: Key length and key generation time comparative study: RSA versus Lattice based cryptosystem.

| Approach | Key-length | Key generation time |
| :--- | :---: | :---: |
|  | (in bits) | (in milliseconds) |
|  | 512 | 360 |
| RSA | 1024 | 1280 |
|  | 2048 | 4195 |
|  | 1169 | 4 |
| Lattice-based | 1841 | 7.5 |
|  | 4024 | 17.5 |

Table 3: Encryption and decryption costs comparative study: RSA versus Lattice based cryptosystem.

| Approach | Key-length <br> (in bits) | Message <br> encryption | Message decryption |
| :--- | :---: | :---: | :---: |
|  |  | (blocks per <br> second) | (blocks per second) |
| RSA | 512 | 2440 | 120 |
| Lattice-based | 1024 | 930 | 20 |
|  | 2048 | 310 | 3 |
|  | 1169 | 5940 | 2820 |

with respect to Lattice and classical discrete logarithm due to the "discrete logarithm problem (DLP)" intractability.
7.2. Comparison on Computation Costs. In Table 2, the relationship between the length of keys in bits and the key generation time in milliseconds has been shown. Based on the results reported in [40], in RSA-based public cryptosystem, the key lengths of 512, 1024 and 2048 bits take 360, 1280 and 4195 milliseconds, respectively. On the other hand, in the proposed lattice-based scheme, the key lengths of 1170, 1841 and 4024 bits require the generation time having $4,7.5$ and 17.5 milliseconds, respectively [40]. This clearly shows that the lattice-based IBE scheme requires less computational time for key generation part as compared to other public key cryptosystems, such as RSA.

Table 3 shows a comparative analysis on the key length in bits with the encryption and decryption speed in terms of blocks per second based on the results reported in [40]. It is noticed that when the key size is smaller, the encryption and decryption processing time for the blocks per second are less. However, the lattice-based cryptosystem performs better than RSA-based public key cryptosystem even if the key size is large.
7.3. Comparison on Security. A comparative study on the key length size and the security aspect between the RSA-based public key cryptosystem and lattice-based cryptosystem has been presented in Table 4 based on the results reported in [40]. Million instructions per second (MIPS) is taken as an "approximate measure of a computer's raw processing power", which is considered in the comparative study. It is observed that in both the cases when the key size is large, the

Table 4: Key length and security comparative study: RSA versus Lattice based cryptosystem.

| Approach | Key-length <br> (in bits) | Security <br> (MIPS per year) |
| :--- | :---: | :---: |
|  | 512 | $4 \times 10^{5}$ |
| RSA | 1024 | $3 \times 10^{12}$ |
|  | 2048 | $3 \times 10^{21}$ |
| Lattice-based | 1169 | $2 \times 10^{6}$ |

security of the system increases. Moreover, even for a smaller key size the lattice based cryptosystem provides significantly better security as compared to that for an RSA-based cryptosystem.

In summary, the lattice-based cryptosystem has several advantages, such as: (a) "cryptographic resistance compared to RSA", (b) "faster key generation", and (c) "faster encryption and decryption of the messages". In addition, the prime advantage of the lattice-based cryptosystem is its resistance to quantum computer attacks.

## 8. Concluding Remarks

In this work, we attempted to design an advanced identitybased encryption that is a very important cryptographic tool to ensure confidentiality in the current quantum era. The proposed encryption is a provably post-quantum secure without random oracles. Since lattices depends on algebraic operations that are typically matrix addition and multiplication, they make the encryption much efficient as compared to other public key cryptosystems, such as RSA. In addition, the proposed scheme is also anonymous and it produces the pseudo-random ciphers. Finally, we incorporated the constructed identity based encryption (IBE) scheme for IoT applications and described how the Big data analytics using the AI/ML techniques will be helpful in such applications.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

## Acknowledgments

The authors would like to thank the anonymous reviewers and the Associate Editor for their valuable comments and suggestions which helped us to improve the presentation and quality of the paper.

## References

[1] B. Eshghi, "IoT Market Outlook for 2022 \& beyond," 2022, https://research.aimultiple.com/iot-future/.
[2] TE CONNECTIVITY, "Smart Factory Sensors and Industrial Internet of Things," 2022, https://www.te.com/usa-en/
industries/sensor-solutions/applications/iot-sensors/indust ry-4-0.html.
[3] Hp Internet of Things Security Study, "Smartwatches," 2017, https://www.ftc.gov/system/files/documents/public_commen ts/2015/10/00050-98093.pdf.
[4] A. K. Das, M. Wazid, A. R. Yannam, J. J. P. C. Rodrigues, Y. Park, and Y. Park, "Provably secure ECC-based device access control and key agreement protocol for IoT environment," IEEE Access, vol. 7, no. 1, pp. 55382-55397, 2019.
[5] ISTR, "Internet Security Threat Report (Istr)," 2018, https:// docs.broadcom.com/doc/istr-23-2018-en.
[6] M. Rana, Q. Mamun, and R. Islam, "Lightweight cryptography in IoT networks: a survey," Future Generation Computer Systems, vol. 129, pp. 77-89, 2022.
[7] G. S. Aujla, R. Chaudhary, N. Kumar, A. K. Das, J. J. P. C. Rodrigues, and P. C. Rodrigues, "SecSVA: secure storage, verification, and auditing of big data in the cloud environment," IEEE Communications Magazine, vol. 56, no. 1, pp. 78-85, 2018.
[8] J. Srinivas, A. K. Das, M. Wazid, and A. V. Vasilakos, "Designing secure user authentication protocol for big data collection in IoT-based intelligent transportation system," IEEE Internet of Things Journal, vol. 8, no. 9, pp. 7727-7744, 2021.
[9] T. Ahamed Ahanger, A. Aljumah, and M. Atiquzzaman, "State-of-the-art survey of artificial intelligent techniques for IoT security," Computer Networks, vol. 206, Article ID 108771, 2022.
[10] C. Iwendi, S. U. Rehman, A. R. Javed, S. Khan, and G. Srivastava, "Sustainable security for the internet of things using artificial intelligence architectures," ACM Transactions on Internet Technology, vol. 21, no. 3, 2021.
[11] A. E. Omolara, A. Alabdulatif, O. I. Abiodun et al., "The internet of things security: a survey encompassing unexplored areas and new insights," Computers \& Security, vol. 112, Article ID 102494, 2022.
[12] S. C. Mukhopadhyay, S. K. S. Tyagi, N. K. Suryadevara, V. Piuri, F. Scotti, and S. Zeadally, "Artificial intelligencebased sensors for Next generation IoT applications: a review," IEEE Sensors Journal, vol. 21, no. 22, pp. 24920-24932, 2021.
[13] A. Shamir, "Identity-based cryptosystems and signature schemes," in Workshop on the Theory and Application of Cryptographic Techniques, pp. 47-53, Springer, Berlin, Germany, 1984.
[14] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," SIAM Review, vol. 41, no. 2, pp. 303-332, 1999.
[15] D. Boneh and M. Franklin, "Identity-based encryption from the weil pairing," in Proceedings of the Annual international cryptology conference, pp. 213-229, Springer, Santa Barbara, CA, USA, August 2001.
[16] D. Boneh and X. Boyen, "Efficient selective-id secure identitybased encryption without random oracles," in Proceedings of the International conference on the theory and applications of cryptographic techniques, pp. 223-238, Springer, Interlaken, Switzerland, May 2004.
[17] G. Craig, "Practical identity-based encryption without random oracles," in Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques, pp. 445-464, Springer, Petersburg Russia, June 2006.
[18] B. Waters, "Efficient identity-based encryption without random oracles," Lecture Notes in Computer Science, Springer, in Proceedings of the Annual International Conference on the

Theory and Applications of Cryptographic Techniques, pp. 114-127, May 2005.
[19] D. Boneh, G. Craig, and M. Hamburg, "Space-efficient identity based encryptionwithout pairings," in Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07), pp. 647-657, IEEE, Providence, RI, USA, October 2007.
[20] C. Cocks, "An identity based encryption scheme based on quadratic residues," in Proceedings of the IMA international conference on cryptography and coding, pp. 360-363, Springer, Cirencester, UK, December 2001.
[21] G. Craig, C. Peikert, and V. Vaikuntanathan, "Trapdoors for hard lattices and new cryptographic constructions," in Proceedings of the fortieth annual ACM symposium on Theory of computing, pp. 197-206, Victoria British Columbia Canada, May 2008.
[22] V. Chamola, A. Jolfaei, V. Chanana, P. Parashari, and V. Hassija, "Information security in the post quantum era for 5G and beyond networks: threats to existing cryptography, and post-quantum cryptography," Computer Communications, vol. 176, pp. 99-118, 2021.
[23] V. Hassija, V. Chamola, A. Goyal, S. S. Kanhere, and N. Guizani, "Forthcoming applications of quantum computing: peeking into the future," IET Quantum Communication, vol. 1, no. 2, pp. 35-41, 2020.
[24] V. Hassija, V. Chamola, V. Saxena et al., "Present landscape of quantum computing," IET Quantum Communication, vol. 1, no. 2, pp. 42-48, 2020.
[25] Q. Li, D. He, Z. Yang, Q. Xie, and K.-K. R. Choo, "Latticebased conditional privacy-preserving authentication protocol for the vehicular Ad Hoc network," IEEE Transactions on Vehicular Technology, vol. 71, no. 4, pp. 4336-4347, 2022.
[26] Z. Xu, D. He, P. Vijayakumar, K.-K. R. Choo, and L. Li, "Efficient NTRU lattice-based certificateless signature scheme for medical cyber-physical systems," Journal of Medical Systems, vol. 44, no. 5, p. 92, 2020.
[27] F. Qi, D. He, S. Zeadally, N. Kumar, and K. Liang, "Ideal lattice-based anonymous authentication protocol for mobile devices," IEEE Systems Journal, vol. 13, no. 3, pp. 2775-2785, 2019.
[28] X. Boyen and B. Waters, "Anonymous hierarchical identitybased encryption (without random oracles)," in Proceedings of the Annual International Cryptology Conference, pp. 290-307, Springer, Santa Barbara, CA, USA, August 2006.
[29] C. Ran, S. Halevi, and J. Katz, "A forward-secure public-key encryption scheme," in Proceedings of the International Conference on the Theory and Applications of Cryptographic Techniques, pp. 255-271, Springer, Warsaw Poland, May 2003.
[30] D. Boneh and X. Boyen, "Secure identity based encryption without random oracles," in Proceedings of the Annual International Cryptology Conference, pp. 443-459, Springer, Santa Barbara, CA, USA, August 2004.
[31] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," Communications of the ACM, vol. 21, no. 2, pp. 120-126, 1978.
[32] T. ElGamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," IEEE Transactions on Information Theory, vol. 31, no. 4, pp. 469-472, 1985.
[33] O. Goldreich, S. Goldwasser, and S. Halevi, "Public-key cryptosystems from lattice reduction problems," in Proceedings of the Annual International Cryptology Conference, pp. 112-131, Springer, Santa Barbara, CA, USA, August 1997.
[34] C. Ran, S. Halevi, and J. Katz, "A forward-secure public-key encryption scheme," Journal of Cryptology, vol. 20, no. 3, pp. 265-294, 2007.
[35] D. Micciancio and O. Regev, "Worst-case to average-case reductions based on Gaussian measures," SIAM Journal on Computing, vol. 37, no. 1, pp. 267-302, 2007.
[36] O. Regev, "On lattices, learning with errors, random linear codes, and cryptography," Journal of the ACM, vol. 56, no. 6, p. 34, 2009.
[37] O. Regev, "Lattice-based cryptography," in Proceedings of the 26th Annual International Conference on Advances in Cryptology (CRYPTO'06), pp. 131-141, Santa Barbara, CA, USA, August 2006.
[38] M. Ajtai, "Generating hard instances of the short basis problem," in International Colloquium on Automata, Languages, and Programming (ICALP'99), Lecture Notes in Computer Sciencevol. 1644, pp. 1-9, Springer, 1999.
[39] D. Apon, X. Fan, and F.-H. Liu, "Compact Identity Based Encryption from LWE," 2016, https://ia.cr/2016/125.
[40] A. Gagnidze, M. Iavich, and I. Giorgi, "Analysis of post quantum cryptography use in practice," Bull. Georgian Natl. Acad. Sci, vol. 11, no. 2, pp. 29-36, 2017.

