Article



# Secure Cloud Computing By A dual-Layer Encryption Mechanism

Sanaa Ali Jabber * , Soukaena Hashem , Shtha Jafer

Posted Date: 8 December 2023

doi: 10.20944/preprints202312.0615.v1

Keywords: cloud; storage; secure; SOA; AES; RSA; LSTM

*Article*

# Secure Cloud Computing by a Dual-Layer Encryption Mechanism

**Sanaa Ali Jabber [1,*], Soukaena H hashem [2] and Shatha H Jafer [3]**

[1] AL-Muthanna University; sanaa.ali@mu.edu.iq
[2,3] University of Technology; Soukaena.h.hashem@uotechnology.edu.iq and Shtha.h.jafer@uotechnology.edu.iq
* sanaa.ali@mu.edu.iq; 07703143080.

**Abstract:** In today's rapidly expanding era of the Internet of Things (IoT) and the Industrial Internet of Things (IIoT), the emphasis on robust data security and insightful data interpreta- tion is more pronounced than ever. This research introduces a comprehensive approach to both data protection and predictive analytics, leveraging the diverse dataset TON IoT.csv, sourced from a myriad of IoT and IIoT environments. For data security, a dual-encryption technique incorporating both AES and RSA algorithms is established. Its efficacy is evidenced by a perfect match between the original and decrypted datasets, under- scoring the integrity of our encryption process. Concurrently, the study ventures into predictive modeling using a modified Snake Optimization Algorithm (SOA) to streamline hyperparam- eter selection. This subsequently aids in the development and fine-tuning of an LSTM network, which exhibits remarkable predictive accuracy. Additionally, the paper provides an in- depth examination of various encryption methodologies like Elliptic Curve Cryptography (ECC), Lightweight Cryptography for Cloud Computing, and Homomorphic Encryption, while also emphasizing the nuances of encryption in cloud setups, particularly contrasting server-side with client-side encryption and efficient key management. The insights presented serve as a cornerstone for ensuing research, promising a bright future for advancements in IoT and IIoT data protection and analysis.

**Keywords:** cloud; storage; secure; SOA; AES; RSA; LSTM

## 1. Introduction

In the proliferative era of cloud computing, wherein a substantial quantum of data is ceaselessly transferred, processed, and stored in a virtual environment, the eminence of encryption cannot be overstated [1]. The quintessence of encryption lies in its capability to transmogrify data into an indecipherable format, mitigating the peril of unauthorized access and ensuring data integrity, confidentiality, and availability in the multifaceted cloud ecosystem [2]. The latter becomes especially poignant, given the dispersed architecture of cloud computing, which inherently encompasses numerous potential vulnerabilities due to the broadened attack surface, encompassing interactions among multiple platforms, devices, and interfaces. Scientifically, encryption mechanisms—whether symmetric, where the same key is utilized to encrypt and decrypt data, or asymmetric, involving a pair of public and private keys operate employing mathematical algorithms to convert plaintext into ciphertext, thereby erecting a virtually impervious barrier against nefarious actors aiming to exploit sensitive data [3]. In a cloud context, this technological stratagem becomes pivotal not merely for safeguarding data at rest, in transit, or during processing but also as a fundamental instrument in assuring regulatory compliance, fortifying user trust, and subsequently enhancing the overarching security posture of cloud service models (IaaS, PaaS, SaaS) [4]. It's worthy of note that encryption also plays a vital role in facilitating secure multi-tenancy within cloud environments, ensuring that the data of different users (tenants) remains isolated and inaccessible to one another despite residing on shared infrastructure [5]. Thus, in synthesizing the panoramic view of encrypted data management

 doi:10.20944/preprints202312.0615.v1



in cloud computing, one navigates through an intricate web of technological, legal, and ethical dimensions, each demanding scrupulous exploration and perpetual advancements to adeptly secure digital assets in an ever-evolving cyber landscape [6].

Secure data storage, particularly within the realms of cloud environments, has burgeoned into a paramount concern amidst an escalating landscape of cyber threats and the voluminous surge of data generation [7]. The significance of ensuring secure data in cloud storage is multi-faceted, transcending technical and operational aspects and anchoring itself deeply into organizational viability, legal compliance, and safeguarding user trust [8]. The act of storing data on the cloud inherently exposes it to a myriad of potential vulnerabilities, from unauthorized access to malicious attacks, given that data is often stored in a shared and virtually accessible environment [9]. Consequently, businesses, governments, and individuals necessitate meticulous encryption, access control, and regular auditing to shield their digital assets from potential breaches, data leaks, or unauthorized manipulations, thereby protecting not only the integrity and confidentiality of the data but also the privacy of the entities and individuals involved.

Moreover, from an economic and reputation standpoint, secure data storage is instrumental in preventing the potentially catastrophic repercussions of data breaches, which can include financial losses, reputational damage, and legal consequences [10]. Particularly in an era where data is regarded as the ”new oil,” protecting this invaluable asset is imperative for sustaining competitive advantage and ensuring continuity in digital transformation endeavors [11]. Legal and ethical dimensions also significantly weigh into the importance of secure cloud data storage, as organizations are increasingly mandated by regulatory frameworks (like GDPR, HIPAA, etc.) to ardently protect consumer and stakeholder data, necessitating them to employ stringent security measures and demonstrate due diligence in managing digital information securely [12]. Therefore, secure data in cloud storage is not merely a technical necessity but a comprehensive, multifaceted imperative that amalgamates aspects of cybersecurity, legal compliance, ethical consideration, and organizational survival and prosperity in the digitally intertwined global ecosystem.

This research paper seeks to delineate a comprehensive comparative analysis of various encryption methods meticulously tailored for securing data in cloud environments, aiming to ascertain their respective strengths, vulnerabilities, and aptness across diverse use-case scenarios in cloud computing. Anchoring on the pivotal role that encryption plays in safeguarding data integrity, confidentiality, and availability in cloud storage and transactions, this inquiry strives to illuminate the technical intricacies, performance implications, operational feasibilities, and security robustness of each scrutinized encryption methodology. Through a systematic exploration of symmetric, asymmetric, homomorphic, and attribute-based encryption techniques, this investigation shall delve into a myriad of dimensions including algorithmic complexity, computational overhead, scalability, key management, and resilience against cyber-attacks, with an overarching aim to carve out a structured framework that assists stakeholders in adeptly selecting and implementing encryption strategies that harmonize with their specific cloud data security requirements and operational contexts. Further, the paper envisages contributing to the scholarly dialogue on cloud data security by identifying potential gaps and future research trajectories in the domain of cloud data encryption, thereby paving the way for innovative approaches that adeptly navigate the evolving challenges in securing cloud environments.

## 2. Background and Related Work

The history of encryption is a long and evolving narrative stretching back to ancient times. The inception of written cryptography is documented around 1900 B.C. in ancient Egypt [13], where a scribe employed non-standard hieroglyphs in an inscription. The term ”encryption” has its roots in the Greek word ”krypto,” signifying something hidden or secret [14]. Throughout history, diverse cultures and civilizations have developed and employed encryption methods for various purposes, including military and political strategy. The Greeks, Romans, and later historical periods saw the use of simple substitution ciphers, which evolved over centuries into more complex systems like the Caesar cipher and eventually modern public-key systems [15]. The evolution continued through the

 doi:10.20944/preprints202312.0615.v1



World Wars, spurring the creation of more sophisticated encryption methods, culminating in the development of digital encryption with the advent of computers. The modern era has witnessed the establishment of robust encryption standards like the Advanced Encryption Standard (AES) and is now venturing into the realm of post-quantum cryptography as the next frontier [16].

In the exploration conducted in [17], the spotlight was placed on enhancing the security protocols inherent to cloud storage, particularly through the lens of searchable encryption models, acknowledging the challenges and limitations posed by single cloud service provider systems. The authors introduced a novel searchable encryption scheme tailored for a multi-cloud environment, underpinned by blockchain technology. A system model was defined across multiple clouds, employing a consortium chain to assimilate multiple cloud service providers for data storage. The proposed scheme ensures encrypted documents and indexes are securely housed within The Interplanetary File System (IPFS), while the hash value and IPFS address of documents are archived in the blockchain. This innovative approach facilitates outsourced encrypted data retrieval based on multiple keywords and also incorporates verification systems to confirm the integrity of retrieved files. The security and elevated performance of the scheme were substantiated through theoretical analyses and practical experiments employing real-world data.

In [18], the researchers delve into the pivotal role of updatable encryption within the context of cloud storage, attributing significance to its capabilities in offering update functionalities for ciphertext data and fortifying defenses against key compromise attacks. Contrasting most prevalent updatable encryption approaches, which predominantly leverage partitioning strategies such as leakage sets or firewalls - limiting the adversary's ability to query the key at challenge-equal epochs and barely securing forward and backward security - the authors augment the existing security model. They introduce an enhanced model that incorporates a corruption oracle, permitting any secret key query and surpassing the reliance on leakage sets. The paper unveils the first updatable public-key encryption scheme, formulated on the foundational ElGamal encryption scheme, that permits no- directional key updates and unidirectional ciphertext updates. This proposed scheme adeptly minimizes information leakage of update tokens, employing the indistinguishable obfuscation technique from the punctured program, and under the developed security model, demonstrates its IND-CPA security. It additionally eschews leakage sets and firewalls, thereby capturing both forward and backward security, presenting an advantageous alternative to extant updatable encryption schemes.

The work presented in [19] addresses several challenges inherent to conventional speech encryption schemes in cloud storage, which include security vulnerabilities, undue communication consumption, insufficient robustness against various attack vectors, and the suboptimal efficiency of the speech homomorphic encryption scheme. The researchers proposed an adaptive speech homomorphic encryption scheme, which is especially oriented toward energy in cloud storage. Initially, an adaptive classifier is crafted employing the improved Adaboost algorithm, which is done by contrasting the threshold of speech energy and subsequently partitioning the speech data into sound and silent segments based on the energy threshold. The BGV homomorphic encryption algorithm is applied to encrypt the sound portion of the data, while the silent segment is encrypted utilizing the Paillier homomorphic encryption algorithm. Ultimately, the two ciphertext parts are integrated to enable ciphertext domain operation and adaptive decryption. Experimental analysis corroborates that the proposed scheme offers commendable encryption and decryption efficiency, minimal ciphertext expansion, and robustness against a spectrum of attacks, including statistical, entropy, and chosen-plaintext attacks.

In the research presented in [20], a salient issue surrounding multiuser security sharing, as well as the privacy protection of speech data stored in the cloud, is addressed. With the objective of enabling efficient encrypted speech retrieval, the authors proposed a scheme rooted in multiuser searchable encryption. Initially, the paper intertwines the strengths of ciphertext-policy attribute-based encryption (CP-ABE) with searchable encryption (SE). This combined approach facilitates a multiuser searchable speech encryption scheme, ensuring not only the encryption but also the nuanced access control of speech data. As a secondary step, the Mel frequency cepstral coefficient

 doi:10.20944/preprints202312.0615.v1



(MFCC) feature, intrinsic to the original speech, is harvested and subsequently fed into the long- and short-term memory network (LSTM) for deep semantic feature extraction, serving as the speech keywords. These keywords undergo encryption, resulting in a secure index, which, when tethered to the encrypted speech, is stored within the cloud. During retrieval, user queries extract speech keywords via the pre-trained LSTM, producing a search trapdoor. This trapdoor, once uploaded to the cloud server, utilizes Euclidean distance for matching against the secure index. Notably, a proxy server is interjected into the framework, undertaking partial ciphertext decryption operations. This addition is strategically designed to mitigate computational overhead and economize on storage space. Both theoretical and experimental evaluations validate the scheme's superior security, precise retrieval capabilities, and its aptitude for securely storing voluminous speech data while supporting multiuser data sharing.

In [21], the authors navigate through the nuanced challenges pertaining to the privacy and security of sensitive speech data in cloud storage, focusing specifically on the realization of efficient, privacy-preserving retrieval for encrypted speech data within cloud storage. A searchable encryption over encrypted speech retrieval scheme is introduced in the cloud storage domain, initiating with the data owner encrypting the original speech data using the symmetric encryption algorithm Lorenz chaotic mapping. Subsequently, this encrypted speech data is uploaded to a cloud server for storage. Simultaneously, the Mel frequency cepstrum coefficient (MFCC) features of the original speech are extracted and used as input for the convolutional neural network (CNN) to perform deep semantic feature extraction, eventually serving as speech keywords. These extracted keywords are then encrypted and stored in the cloud. During retrieval, an authorized user dispatches a retrieval request, employing the trained CNN to extract keywords from the speech intended for retrieval, subsequently generating a search trapdoor which is transmitted to the cloud server. Euclidean distance is utilized to match encrypted keywords with the search trapdoor during retrieval. Both theoretical and experimental evaluations validate the scheme's heightened security and retrieval accuracy, endorsing its aptitude for encrypted storage of speech data and efficient, secure retrieval.

In the work encapsulated in [22], the authors delve into the pivotal realm of cloud computing, a technology synonymous with the evolution of information technology that goes beyond merely furnishing users with high-performance computing to addressing the exigencies of large-scale data storage. However, the paper underscores a prevalent user distrust stemming from the opaque nature of the storage service provided by cloud computing, where users remain uninformed about the security status of their data within the cloud computing environment, thereby hindering its development. The paper demystifies the fundamental knowledge and system architecture of cloud storage and elucidates the current developmental status of cloud storage. The focal point of the paper is its exploration into data encryption algorithms with a proposition of a cross-encryption scheme crafted for data security storage within a cloud computing environment, aiming to fortify the storage security of user data therein. A comparative analysis with traditional hybrid encryption methods reveals that the proposed scheme boasts commendable encryption and decryption outcomes, swift execution speed, and heightened security, positing it as an exemplary scheme for data security storage in cloud computing environments.

In [23], the focus resides on securing expansive volumes of medical reports, including electronic patient records and medical images, which are imperative to be stored securely for prospective reference. Despite cloud storage services stepping into to meet demands with their scalability and availability, there lies a crucial issue in that numerous cloud service providers store client data in an unencrypted text format. Thus, the onus falls onto cloud users to devise strategies to safeguard their medical data. While existing image encryption solutions prevail, they often demonstrate vulnerabilities to chosen-plaintext attacks due to continually advancing computer power and hacker ingenuity. The paper brings forth an encryption technique influenced by Hopfield neural network (HNN) aimed at bolstering resilience against an array of attacks, optimizing and refining the system via persistent learning and updating. Remarkably, the methods encompass a dynamic security feature that self-adapts to real-world fluctuations and advancements. This scheme employs the back propagation neural network to concoct image-specific keys that bolster resilience against potential

 doi:10.20944/preprints202312.0615.v1



hacker intrusions. These generated keys serve as an initial seed for the generation of confusion and diffusion sequences through HNN, thereby enhancing security parameters.

In the study presented in [24], the authors spotlight the growing significance of cloud storage services, a salient facet of cloud computing. With the ever-increasing migration of data owners depositing their data on remote cloud platforms, there's an escalating concern regarding data security and the safeguarding of privacy, given the considerable distance between users and these platforms. One promising avenue for fortifying security and privacy within cloud storage is through Identity-based broadcast encryption (IBBE). However, the susceptibility of cryptographic systems to side-channel attacks, which may unveil pivotal key information, is a profound challenge to system security. Addressing this challenge, the paper introduces an identity-based broadcast encryption with leakage resilience through state partition, termed as LR-SP-IBBE. This novel scheme compensates for entropy loss in the symmetric key caused by side-channel onslaughts by leveraging a binary extractor and further randomizes the encapsulated symmetric key. The innovative approach of partitioning the private key into bifurcated segments and executing decryption in dual stages further amplifies its security stance. The employment of the double- system encryption method aids in substantiating its security and leakage resilience within a composite order group model.

In a related work, an innovative approach towards tackling the issues of inefficient encryption and decryption, as well as the notable expansion of speech ciphertext in current speech homomorphic encryption schemes for cloud storage, was explored [25]. The study acknowledged the DGHV fully homomorphic encryption scheme but highlighted its limitation to a singular-bit encryption form. In the proposed method, the researchers introduced a speech fully homomorphic encryption strategy for DGHV, pivoting on multithreading technology. Initially, the scheme involved the conversion of floating-point speech data to integer data through a preprocessing stage. Subsequent steps involved disassembling the preprocessed data into binary strings bit-wise, which were then converted into a matrix to facilitate cyclic encryption. Leveraging multithreading technology and a many-to-one homomorphic speech encryption scheme, parallel encryption was employed to secure the final ciphertext speech data. Performance and experimental outcomes emphasized the scheme's adeptness in ensuring high security, efficient encryption, and decryption with minimal ciphertext expansion while maintaining resilience against various conventional attacks. The Summary of related work are presented in Table 1 depending on the Problem Addressed**,** Method and Contribution.

**Table 1.** Summary of Related Work.

| Ref | Problem Addressed | Method/Contribution |
|---|---|---|
| [17] | Enhancing security in cloud storage | Introduced a blockchain-based searchable encryption for multi-cloud environments. |
| [18] | Updatable encryption within cloud storage | Enhanced security model and presented an updatable public-key encryption scheme. |
| [19] | Issues in conventional speech encryption | Proposed an adaptive speech homomorphic encryption scheme optimized for cloud storage. |
| [20] | Multiuser secure sharing of speech data | Proposed a scheme utilizing CP-ABE and SE for encrypted speech retrieval. |
| [21] | Updatable encryption functionality | Developed an updatable encryption scheme, moving beyond leakage sets and firewalls. |
| [22] | Privacy of speech data in cloud storage | Introduced a searchable encryption scheme for encrypted speech retrieval in cloud. |
| [23] | Securing medical data in cloud storage | Proposed a dynamic encryption technique utilizing Hopfield neural network (HNN). |
| [24] | Handling side-channel attacks in cloud storage | Presented an LR-SP-IBBE scheme with bifurcated private key decryption. |
| [25] | Inefficiencies in speech homomorphic encryption | Proposed a multithreaded DGHV-based speech homomorphic encryption scheme. |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2023 doi:10.20944/preprints202312.0615.v1



## 3. Encryption Methods Under Study

### *3.1. Advanced Encryption Standard (AES)*

The AES holds a widely acknowledged position in the realm of symmetric encryption methods, establishing itself as a quintessential tool in safeguarding digital information [26]. AES, recognized for its resilient security attributes and inherent capability to guard against various cyber-attacks, has permeated numerous domains of digital security, offering a fortified shield against unauthorized data access and potential breaches [27].

As the Figure 1 illustrates, the AES operates by utilizing different key lengths: 128-bit, 192-bit, and 256-bit. Each of these key lengths denotes the size of the key used in the encryption process, with a longer key generally offering a stronger level of encryption [28].

In the provided diagram, the process begins with the input of both the 'Secret Key' and the 'Plain Text'. Depending on the security requirements and the application's needs, one can choose from the three key lengths mentioned. The selected key, combined with the plain text, undergoes an encryption process represented by the "Cipher" block in the figure.

The outcome of this encryption process is the 'Cipher Text', which is essentially the encrypted version of the original plain text. The 'Cipher Text' can then be safely transmitted or stored without exposing the actual information contained within the plain text. Only individuals with the correct 'Secret Key' can decrypt the 'Cipher Text' back to its original form.

This diverse range in key lengths ensures flexibility. For applications that require a high level of security, the 256-bit key length can be employed, while for scenarios with less stringent security needs, the 128-bit or 192-bit key lengths might be more appropriate. The adaptability of AES in accommodating various key lengths makes it versatile for a wide range of applications across different platforms and systems.

In the expansive and ever-evolving digital landscape of cloud computing, AES has demonstrated profound utility, notably attributed to its exemplary balance between security and computational overhead. The versatility and computational efficiency of AES make it a preferred choice in cloud environments, where safeguarding sensitive data is paramount while concurrently maintaining optimal system performance. This encryption standard seamlessly integrates with various cloud applications, offering a secure conduit for data transmission and storage, thereby ensuring that the digital assets residing in the cloud are enveloped in a secure, impenetrable cipher, mitigating risks and enhancing the overall cloud security posture. Consequently, AES has become an indispensable component in sculpting a secure and reliable digital infrastructure in cloud environments across the globe.

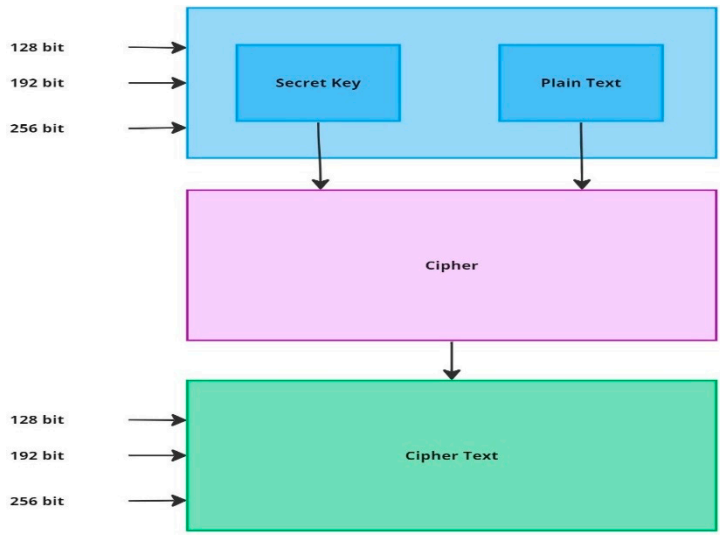


**Figure 1.** AES Architecture [28].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2023 doi:10.20944/preprints202312.0615.v1



*3.2. Elliptic Curve Cryptography (ECC)*

ECC stands as a pivotal advancement in the domain of asymmetric encryption. What differentiates ECC from other cryptographic methods is its exceptional efficiency and suitability for lightweight devices [29]. This efficiency stems from ECC's ability to provide strong encryption with shorter key lengths, which translates to faster encryption and decryption processes. As a consequence of this inherent property, ECC demands a lower computational cost compared to traditional cryptographic methods. As illustrated in this Figure 2, the architecture of ECC can be understood in terms of its role in digital signature generation and verification. This specific diagram exemplifies the broader process of digital signing and its subsequent verification using ECC.

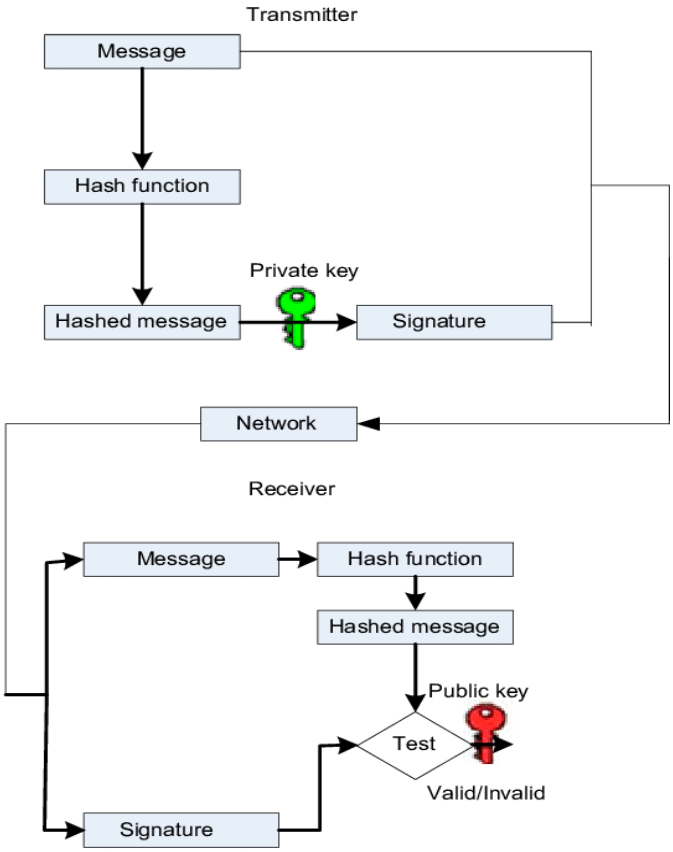


**Figure 2.** Elliptic Curve Cryptography [30].

In the "Transmitter" section, the process begins with a "Message" that needs to be sent securely. This message undergoes a "Hash function" which computes a unique hash representation, termed the "Hashed message." The hashed message is then combined with a "Private key" to generate a unique "Signature." This signature serves as a mathematical proof of the authenticity and integrity of the message.

The message and its associated signature are then transmit- ted over the "Network" to the intended recipient.

Upon reaching the "Receiver" section, the received message is subjected once again to the same "Hash function" to derive a "Hashed message." Subsequently, using the sender's "Public key," the receiver conducts a "Test" on the hashed message and the received signature. If the test proves successful, it validates the authenticity and integrity of the message, indicating that the message has remained unchanged during transmission and that it has indeed been sent by the claimed sender. If there's any discrepancy, the test would result in an "Invalid" outcome, implying potential tampering or issues with the message's authenticity.

Through this ECC-based mechanism, parties can securely exchange information with the assurance that the data is genuine and untampered, bolstering trust and reliability in digital

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2023 doi:10.20944/preprints202312.0615.v1



communications. This computational advantage is especially pronounced in cloud environments. As cloud platforms often handle vast amounts of data and require rapid encryption and decryption processes, the efficiency of ECC becomes highly valuable. Its ability to deliver robust encryption without imposing a significant computational burden makes ECC an ideal choice for cloud infrastructures, ensuring both security and optimal performance. In essence, Elliptic Curve Cryptography offers a harmonious blend of strength and speed, making it a preferred choice for modern cryptographic applications, especially in the realm of cloud computing.

### *3.3. RSA (Rivest–Shamir–Adleman)*

RSA, which stands for Rivest–Shamir–Adleman, is a widely recognized asymmetric encryption method. Its prominence in the realm of cryptography is largely due to its robust security features that ensure the confidentiality and integrity of data. While RSA is renowned for its robustness, it is important to note that this strength comes with a trade-off [31]. The RSA algorithm inherently requires a higher computational overhead compared to some other cryptographic methods. This means that encrypting and decrypting using RSA might be more resource-intensive and potentially slower. However, many con- sider this a worthwhile compromise given the high level of security that RSA affords. In essence, RSA embodies a blend of computational complexity and unparalleled security, solid- ifying its place as one of the pillars in modern cryptographic practices. As illustrated in this Figure 3, the architecture of the RSA encryption method can be understood in a sequence of steps that demonstrate how data is securely transmitted between two parties.

The process initiates with the ”Sender,” who has some ”Plaintext data” they wish to send securely. To ensure the data remains confidential during transmission, it is encrypted using a ”Public Key.” Once the plaintext data undergoes encryption with the public key, it transforms into ”Ciphered Data.” This encrypted form ensures that any eavesdropper or malicious actor cannot easily decipher the original message even if they intercept the transmission.

Upon reaching the intended ”Recipient,” the ciphered data is then decrypted. For this decryption process, a ”Private Key” is utilized. It’s essential to understand that the private key is unique and only known to the recipient. After successful decryption, the ciphered data reverts to its original ”Decrypted Plaintext data” form, enabling the recipient to access the original message sent by the sender [32] .

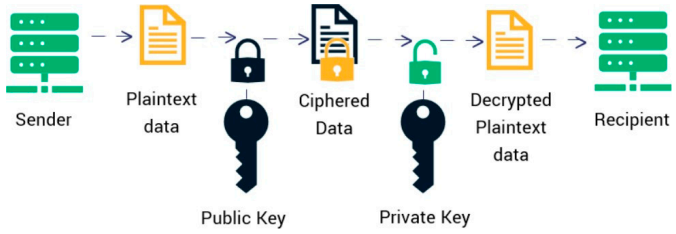


**Figure 3.** Rivest–Shamir–Adleman [31].

### *3.4. Lightweight Cryptography for Cloud Computing*

In the realm of cloud computing, there’s a burgeoning interest in what’s known as lightweight cryptography. This area of study zeroes in on cryptographic methods that are tailor-made for devices with constraints, such as limited pro- cessing power, memory, or energy resources. Unlike traditional cryptographic systems that can be computationally intense, lightweight cryptography is finely tuned to work efficiently on smaller or embedded systems [33]. The crux of lightweight cryptography lies in striking a balance. While it’s essential to maintain robust security, there’s an equal emphasis on ensuring computational efficiency. It acknowledges the fact that while high-end security is paramount, it shouldn’t come at the cost of overwhelming constrained devices or slowing down their operations [34].

 doi:10.20944/preprints202312.0615.v1



### *3.5. Homomorphic Encryption*

Homomorphic encryption stands out as a revolutionary cryptographic technique, especially in the context of cloud environments [35]. Its unique selling proposition is its ability to enable computations directly on encrypted data. In conven- tional encryption methods, if one needed to perform operations on the data, they would first have to decrypt it, process it, and then possibly re-encrypt it. Homomorphic encryption bypasses this need. It allows data to remain encrypted while still being processed or computed. This feature holds immense significance in cloud settings. When data is stored in the cloud, concerns about its confidentiality and potential access by unauthorized parties arise. With homomorphic encryption, cloud service providers can perform computations on the client's encrypted data without ever needing to decrypt it [36]. This ensures that the data remains confidential, boosting trust and opening up new possibilities for secure cloud-based applications and services.

## 4. Cloud Encryption

### *4.1. Importance of Encryption in Cloud*

As the world becomes more connected, the reliance on cloud storage has seen a significant increase. This shift towards cloud storage is driven by its convenience, allowing users to access their data from anywhere, and scalability, adapting to the ever- growing data needs of both individuals and organizations. This has led to an enormous amount of sensitive and critical data being stored in the cloud, making its security paramount.

However, as more data is hosted online, it becomes an enticing target for cybercriminals. Data breaches, where unauthorized parties gain access to secured information, have become a concerning reality of the digital age. Such breaches can result in a plethora of negative consequences, ranging from financial losses due to fraud or ransom demands, damage to a company's reputation, and potential legal ramifications. In light of these threats, the significance of robust encryption becomes clear. Encryption ensures that even if data is accessed without authorization, it remains unintelligible without the correct decryption keys, thus safeguarding the information's confidentiality and integrity.

### *4.2. Server-side vs. Client-side Encryption*

The process of encryption can occur at various points, but two primary methods are server-side and client-side encryption [37]. The distinction between the two is essential for under- standing cloud data security. In server-side encryption, data is encrypted once it reaches the cloud provider's infrastructure. The main advantage of this method is its ease of use for the end-user; they simply upload their data, and the cloud provider handles the encryption process. However, since the encryption takes place on the provider's side, they have control over the encryption keys. This can be a potential security concern if the provider's systems are compromised.

Contrastingly, client-side encryption refers to data being encrypted on the user's end before it is transmitted to the cloud. This ensures that the data is always encrypted during transit and remains encrypted on the server. One significant benefit of this approach is that the encryption keys are typically managed by the user or their organization, ensuring the cloud provider cannot decrypt the data even if they wanted to. However, this added security comes with the responsibility of key management on the user's part.

### *4.3. Key Management in Cloud*

In cloud environments, the proper management of encryp- tion keys is essential for data security. Key management encompasses the processes of creating, storing, distributing, and retiring encryption keys. The initial challenge lies in generating strong keys that are unpredictable, using trusted cryptographic algorithms. Once created, these keys need se- cure storage, typically in

 doi:10.20944/preprints202312.0615.v1



specialized solutions like hardware security modules (HSMs), ensuring they're isolated from potential threats.

But it's not just about keeping keys safe. Over time, for reasons ranging from computational advances to potential vulnerabilities in algorithms, keys might need to be rotated or replaced. This necessitates re-encrypting data with new keys. Access controls are also crucial, ensuring only authorized personnel can handle the keys. Furthermore, to prevent irreversible data loss, backups of these keys are a must, but they too require secure storage.

Lastly, as with many elements of cybersecurity, monitoring and auditing are vital. Organizations need to keep a vigilant eye on how encryption keys are accessed and used, ensuring there's a record of all key-related activities. This not only helps in maintaining security but also ensures compliance with various data protection regulations.

## 5. Methodology

Embarking on a journey through machine learning with the IoT dataset, our methodology meticulously intertwines data security, preprocessing, and model development to forge a robust predictive model. Initially, we safeguard data through a multi-layered encryption and decryption process, ensuring secure storage and transmission. The data is first subjected to a thorough preprocessing phase, ensuring it is of the highest quality and in the appropriate format. This data is then strategically divided into training and testing sets to ensure a robust modeling process. For optimal model parameter selec- tion, we utilize a modified version of the Snake Optimization Algorithm (SOA). Following this, a deep learning model, specifically an LSTM network, is built. During the training process, measures like early stopping and dropout layers are employed, not just to improve the model's performance but also to counteract overfitting. The ultimate objective is to ensure our model can consistently and accurately predict outcomes using unseen data. The reason for the in-depth expla- nation of each step in the subsequent subsections is to provide readers with a transparent, comprehensive understanding of our methodology and to set a clear path for researchers aiming for reproducibility. Further insights can be drawn from Figure 4.

### *5.1. Dataset Description: Insights into TON IoT.csv*

We utilize the TON IoT dataset, a rich and diverse col- lection of data, amalgamating various sources to facilitate a comprehensive analysis in the realm of IoT and IIoT. Originating from the Cyber Range and IoT Labs at the School of Engineering and Information Technology (SEIT), UNSW Canberra @ the Australian Defence Force Academy (ADFA), this dataset is a product of a meticulously designed, large-scale network, specifically crafted for Industry 4.0, encompassing IoT and IIoT networks. The creation of the testbed involved deploying numerous virtual machines and hosts, utilizing various operating systems like Windows, Linux, and Kali, to manage the intricate interconnections between IoT, Cloud, and Edge/Fog systems.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2023 doi:10.20944/preprints202312.0615.v1



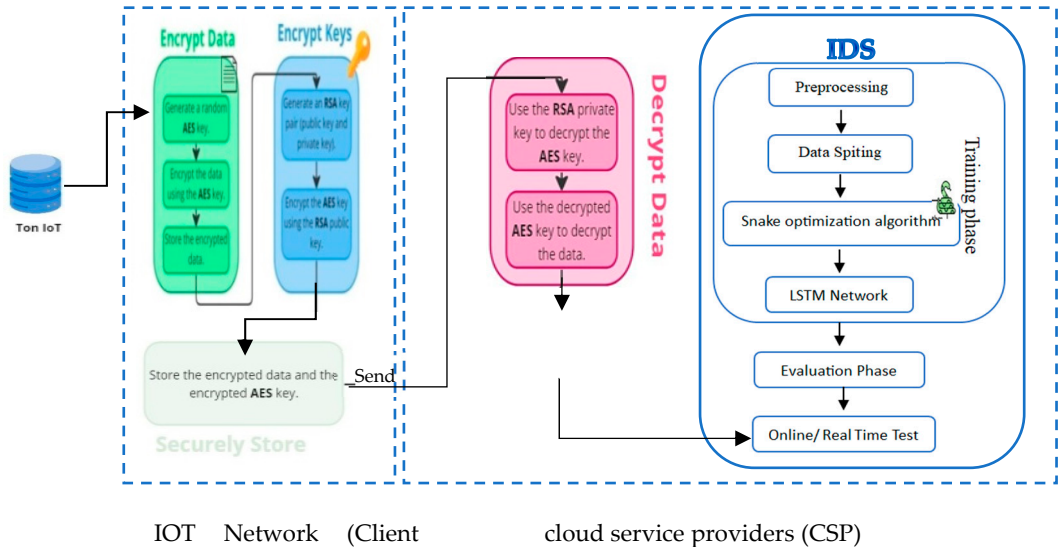


**Figure 4.** The general structure of the proposed system.

The TON IoT dataset is a confluence of data from telemetry datasets of IoT and IIoT sensors, operating system datasets from Windows 7 and 10, as well as Ubuntu 14 and 18 TLS, and network traffic datasets. It was curated in a parallel processing environment, capturing both normal and cyber-attack events from network traffic, Windows audit traces, Linux audit traces, and telemetry data of IoT services, providing a robust foun- dation for exploring various attacking techniques, including DoS, DDoS, and ransomware, against web applications, IoT gateways, and computer systems across the IoT/IIoT network.

In terms of directory structure, the TON IoT datasets are organized as follows:

- IoT/IIoT Datasets: Logged in both log and CSV files, capturing telemetry data from over 10 IoT and IIoT sensors, including weather and Modbus sensors.
- Network Datasets: Collected in packet capture (pcap) formats, log files, and CSV files, utilizing the ZEEK (Bro) tool.
- Linux Datasets: Obtained by employing a tracing tool, particularly atop, on Ubuntu 14 and 18 systems, logging desk, process, processor, memory, and network activities in TXT and CSV formats.
- Windows Datasets: Captured by deploying dataset col- lectors of the Performance Monitor Tool on Windows 7 and 10 systems, collecting activities of desk, process, processor, memory, and network activities in a CSV format, with raw datasets being collected in a blg format.

This dataset, with its multifaceted nature, serves as a pivotal element in our methodology, enabling us to navigate through various aspects of machine learning in the IoT domain.

### *5.2. Experimental Setup*

In the pursuit of conducting our experiments with the TON IoT dataset, we leveraged a blend of cloud and local computational resources to ensure optimal performance and efficiency. Specifically, we utilized Google Colab, a cloud- based platform that offers a conducive environment for ma- chine learning research, providing access to robust computational resources and facilitating collaborative research efforts. Google Colab was instrumental in enabling us to perform extensive computations without being constrained by local hardware limitations.

In addition to cloud resources, our local experimental setup was anchored by a Dell machine, fortified with an Intel(R) Core(TM) i5-10500H CPU, clocking at 2.50GHz and boasting 6 cores and 12 logical processors, ensuring a potent computational capability. The machine was further equipped with 16.0 GB of installed physical memory (RAM), providing ample space to handle large datasets and perform memory-intensive operations.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2023 doi:10.20944/preprints202312.0615.v1



This hybrid setup, combining the flexibility and scalability of Colab with the reliable and consistent performance of our local Dell machine, allowed us to navigate through the various stages of our methodology, from data preprocessing to model training and evaluation, in a seamless and efficient manner. This strategic amalgamation of hardware and software tools ensured that our experimental workflow was both robust and adaptable, accommodating the diverse and dynamic needs of our machine learning experiments.

### *5.3. Evaluation Metrics: Scrutinizing Security Aspects*

In the domain of security, our methodology is carefully designed to ensure the highest levels of protection and integrity for the TON IoT dataset. The focus areas include encryption and decryption times, computational overhead, memory usage, and pertinent security metrics. The process begins with data encryption, leveraging the AES algorithm with a 256-bit key (32 bytes in length). A random AES key is generated for this purpose and is then utilized to encrypt the dataset. Once encrypted, the data is stored with utmost security, ensuring it is shielded from unauthorized access. .

Subsequently, an RSA key pair is generated to encrypt the AES key, providing an additional layer of security. The encrypted data and AES key are then either transmitted or stored securely, ensuring that both are safeguarded during any data transfer processes or while at rest. The decryption process, which is crucial for data analysis, involves using the RSA private key to decrypt the AES key, which is then used to decrypt the data.

A pivotal aspect of our methodology is the data integrity check, which validates that the decrypted data precisely matches the original data, ensuring both its integrity and authenticity. This is achieved by comparing the original and decrypted datasets, ensuring that no discrepancies have arisen as a result of the encryption and decryption processes. This meticulous approach to security ensures that the data remains both secure and reliable throughout the experimental process, safeguarding the validity of our subsequent analyses.

## 6. Experimental Results

### *Security*

Ensuring the security of the TON IoT dataset was paramount, given the sensitive nature of IoT and IIoT data. The encryption process, which utilized a combination of AES and RSA algorithms, was executed to safeguard the data against unauthorized access and potential cyber threats. The AES algorithm, known for its robustness and widespread use in securing data, was employed to encrypt the actual dataset. A random AES key was generated and used to transform the original data into an encrypted format. This key itself was then encrypted using an RSA public key, providing a dual-layer security mechanism.

Upon decryption, a crucial step was the validation of data integrity and authenticity. This was achieved by comparing the original and decrypted datasets, ensuring that the encryption and decryption processes had not inadvertently altered the data. The results were impeccable, with a 100% match across all 461,043 rows, affirming that the decrypted data was an exact replica of the original dataset, thereby ensuring that subsequent analyses could be conducted with confidence in the data's integrity.

### *Snake Optimization Algorithm (SOA)*

The SOA was employed to navigate the hyperparameter space, seeking to identify the optimal configuration for the LSTM model. The algorithm, which utilizes a blend of exploitation and exploration strategies, achieved an impressive accuracy of 0.9942. This indicates that the SOA was able to efficiently traverse the parameter space, identifying a configuration that yielded near-optimal model performance.

 doi:10.20944/preprints202312.0615.v1



*LSTM Model*

The LSTM model was constructed and trained using the optimal parameters identified by the SOA. The model demonstrated robust predictive performance, achieving an overall accuracy of 0.98. The precision, recall, and F1-score for both classes (0 and 1) were also commendable, indicating that the model was capable of accurately classifying the IoT data with minimal error.

**Table 2.** Summary of Experimental Results.

| Metric/Algorithm | Result |
|---|---|
| **Security** | |
| Matching Rows | 461,043 / 461,043 |
| Percentage Match | 100% |
| **Snake Optimization Algorithm (SOA)** | |
| Accuracy | 0.9942 |
| **LSTM Model** | |
| Overall Accuracy | 0.98 |
| Precision (Class 0) | 0.98 |
| Precision (Class 1) | 0.98 |
| Recall (Class 0) | 0.99 |
| Recall (Class 1) | 0.97 |
| F1-Score (Class 0) | 0.99 |
| F1-Score (Class 1) | 0.98 |

## 7. DISCUSSION

The experimental journey through the security and predictive modeling of the IoT dataset has unfolded a plethora of insights and findings. The security aspect, which was meticulously handled through a dual-layer encryption mechanism involving AES and RSA algorithms, demonstrated impeccable integrity in safeguarding the data. The 100% match between the original and decrypted datasets not only affirmed the reliability of the encryption-decryption process but also underscored the importance of implementing such security measures, especially considering the sensitive and critical nature of IoT and IIoT data. In scenarios where data is stored in cloud environments, this dual-layer encryption method proves to be pivotal, providing a robust shield against potential unauthorized access and cyber threats, thereby ensuring that the data remains unscathed and authentic, even in a remotely accessed and managed environment.

The Snake Optimization Algorithm (SOA) and LSTM model, on the other hand, showcased the potential of employing advanced algorithms and models in extracting meaningful insights from IoT data. The SOA, with its strategic blend of exploitation and exploration in navigating the hyperparameter space, and the LSTM model, with its adept handling of sequential data, together forged a pathway towards accurate and reliable predictive modeling. The achieved accuracy and other metric scores indicate that such a combination of optimization algorithm and predictive model can be instrumental in har- nessing the true potential of IoT data, enabling organizations to predict, and thereby mitigate, potential issues before they escalate into critical problems.

In light of the experimental findings, it is recommended that a dual-layer encryption method, similar to the one employed in this study, be adopted widely across various IoT scenarios, especially those involving cloud storage and transmission of data across networks. The additional layer of security provided by encrypting the AES key with an RSA public key adds an extra shield, making it considerably more challenging for unauthorized entities to gain access to the original data. Furthermore, considering the potential of the SOA and LSTM in navigating the hyperparameter space and handling sequential data respectively, it is recommended that such methodologies be explored further in various IoT scenarios, potentially un- locking new possibilities and insights in the realm of predictive modeling and data analysis in IoT and IIoT environments.

## References


1. S. Ali Jabber, S. H. Hashem and S. H. Jafer, Task Scheduling and Resource Allocation in Cloud Computing : A Review and Analysis, *3rd International Conference on Emerging Smart Technologies and Applications (eSmarTA), Taiz, Yemen* **2023**, pp. 01-08, doi: 10.1109/eSmarTA59349.2023.10293517.
2. Xin'an Zhou, Jiale Guan, Luyi Xing, and Zhiyun Qian. Perils and mitigation of security risks of cooperation in mobile-as-a-gateway iot. In Proceedings of the *ACM SIGSAC Conference on Computer and Communications Security* **2022,** pages 3285–3299,.
3. Aryan Parekh, Mayav Antani, Kartik Suvarna, Ramchandra Mangrulkar, and Meera Narvekar. Multilayer symmetric and asymmetric technique for audiovisual cryptography. *Multimedia Tools and Applications* **2023,** pages 1–39.
4. Sarah Ahmed and Muhammad Khan. Securing the internet of things (iot): A comprehensive study on the intersection of cybersecurity, privacy, and connectivity in the iot ecosystem. *AI, IoT and the Fourth Industrial Revolution Review* **2023,** 13(9):1–17.
5. Tommaso Crepax and Siddharth Prakash Rao. Blockchain in the cloud: a primer on data security for blockchain as a service (baas). Available at *SSRN* **2020,** 3766900.
6. Hanane Allioui and Youssef Mourdi. Exploring the full potentials of iot for better financial growth and stability: A comprehensive survey. *Sensors*, **2023,** 23(19):8015,.
7. Pesqueira, A., Sousa, M. J., & Costa, J. Exploring the Role of Big Data Analytics and Dynamic Capabilities in ESG Programs within Pharmaceuticals. **2023**.
8. Ponnam Lalitha and Rohita Yamaganti. Investigation into security challenges and approaches in cloud computing. *Journal of Engineering Sciences* **2023,** 14(08).
9. Mahmoud Abbasi, Javier Prieto, Amin Shahraki, and Juan M Corchado. Industrial data monetization: A blockchain-based industrial iot data trading system. *Internet of Things* **2023,** page 100959.
10. Srinath Perera, Xiaohua Jin, Alana Maurushat, and De-Graft Joe Opoku. Factors affecting reputational damage to organisations due to cyberat- tacks. *In Informatics*, MDPI **2022,** volume 9, page 28.
11. Alshammari, S. A., & Seno, S. H. A Cooperation of Fog Computing and Smart Gateways in a Secure and Efficient Architecture for IoT-Based Smart Homes. *Engineering and Technology Journal* **2019**, 37(7 Part A).
12. Sarath Sabu, H.M. Ramalingam, M Vishaka, H.R. Swapna, Swaraj Hegde. Implementation of a secure and privacy-aware E-Health record and IoT data sharing using blockchain. *Global Transitions Proceedings* **2021,** Volume 2, pp. 429-433.
13. Joseph J Boutros. The role of cryptography in our information-based society. **2020**.
14. Rashed J Al-Hamadin. A New Approach for Data Symmetric Key Cryptography Using Fast Neural Networks with Single Step of Back- propagation and Finite Fields. PhD thesis, Princess Sumaya University for Technology (Jordan), 2021.
15. Abdalbasit Mohammed Qadir and Nurhayat Varol. A review paper on cryptography. In 2019 7th international symposium on digital forensics and security (ISDFS), pages 1–6. IEEE, 2019.
16. Randall Seymour. Designing Improved Minimum Resource Recommen- dations for Virtual Environments with Layered Encryption Mechanisms. PhD thesis, Colorado Technical University, 2022.
17. Shaojing Fu, Chao Zhang, and Weijun Ao. Searchable encryption scheme for multiple cloud storage using double-layer blockchain. *Con- currency and Computation: Practice and Experience* **2022**, 34(16):e5860.
18. Zhenhua Liu, Jingwan Gong, Yuanju Ma, Yaxin Niu, and Baocang Wang. Updatable elgamal encryption scheme with forward and back- ward security for cloud storage. In International Conference on Frontiers in Cyber Security, pages 324–345. Springer, 2022.
19. Qiu-Yu Zhang and Yu-Jiao Ba. An adaptive speech homomorphic encryption scheme based on energy in cloud storage. *International Journal of Network Security* **2022**, 24(4):628–641.
20. Qiuyu Zhang, Minrui Fu, Yibo Huang, Zhenyu Zhao, et al. Encrypted speech retrieval scheme based on multiuser searchable encryption in cloud storage. *Security and Communication Networks*, **2022**, 2022.
21. Qiuyu Zhang, Minrui Fu, Zhenyu Zhao, and Yibo Huang. Searchable encryption over encrypted speech retrieval scheme in cloud storage. *Journal of Information Security and Applications* **2023**, 76:103542.
22. Haiyan Kang and Jie Deng. A cross encryption scheme for data security storage in cloud computing environment. *International Journal of Internet Protocol Technology* **2023**, 16(1):1–10.
23. C Lakshmi, Karuppusamy Thenmozhi, John Bosco Balaguru Rayap- pan, Sundararaman Rajagopalan, Rengarajan Amirtharajan, and Nithya Chidambaram. Neural-assisted image-dependent encryption scheme for medical image cloud storage. *Neural Computing and Applications* **2021**, 33:6671–6684.

24. Qihong Yu, Jiguo Li, and Sai Ji. Identity-based and leakage-resilient broadcast encryption scheme for cloud storage service. *Applied Sciences* **2022**, 12(22):11495.
25. QY Zhang and YG Jia. A speech fully homomorphic encryption scheme for dghv based on multithreading in cloud storage. *Int J Netw Secur* **2022**, 24(6):1042–55.
26. Nicolas Moura, Joaquim Lucena, Eduardo Pereira, Ney Calazans, Lu- ciano Ost, Fernando Moraes, and Rafael Garibotti. Assessment of lightweight cryptography algorithms on arm cortex-m processors. In 2023 36th SBC/SBMicro/IEEE/ACM Symposium on Integrated Circuits and Systems Design (SBCCI), pages 1–6. IEEE, 2023.
27. Radhi, Samara Mohammed; OGLA, Raheem. In-Depth Assessment of Cryptographic Algorithms Namely DES, 3DES, AES, RSA, and Blowfish. *Iraqi Journal of Computers, Communications, Control and Systems Engineering*, **2023**, 23.3: 125-138.
28. Dwi Kuswanto. Performances combination schemes aes-turbo code based-on keys length. In IOP Conference Series: Materials Science and Engineering, volume 1125, page 012047. IOP Publishing, 2021.
29. Shamsher Ullah, Jiangbin Zheng, Nizamud Din, Muhammad Tanveer Hussain, Farhan Ullah, and Mahwish Yousaf. Elliptic curve cryptog- raphy; applications, challenges, recent advances, and future trends: A comprehensive survey. *Computer Science Review* **2023**, 47:100530.
30. Nabil Ghanmy, Naziha Khlif, Lamia Fourati, and Karim Lotfi. Hardware implementation of elliptic curve digital signature algorithm (ecdsa) on koblitz curves. pages 1–6, 07 2012.
31. Mohammed Abujoodeh, Liana Tamimi, and Radwan Tahboub. Toward lightweight cryptography: A survey. In Computational Semantics. IntechOpen, 2023.
32. matthias, d.; osakwe, b. p.; anireh, V. I. E. A Secure Model on Cloud using a Modified Rivest, Shamir and Adleman Algorithm along with Gray Codes. *International Journal of Computers & Technology*, **2021,** page 207-214, 8(1),.
33. Zane Mechalke Sullivan, Maj Bobby Birrer, Sameul Dick, and Jordon Cochran. Analysis of practical application of lightweight cryptographic algorithm ascon jeffrey avery, phd, bryson fraelich, william duran, andrew lee, agustin.
34. Abd Zaid, Mustafa M., and Soukaena Hassan. Proposal Framework to Light Weight Cryptography Primitives. *Engineering and Technology Journal* **2022**, 40(04):516-526.
35. Stefania Loredana Nita and Marius Iulian Mihailescu. Advances to Homomorphic and Searchable Encryption. Springer Nature, 2023.
36. Kundan Munjal and Rekha Bhatia. A systematic review of homomorphic encryption and its contributions in healthcare industry. *Complex & Intelligent Systems* **2023**, 9(4):3759–3786.
37. Yufei Xing and Shuguo Li. An efficient implementation of the newhope key exchange on fpgas. IEEE Transactions on Circuits and Systems I: Regular Papers, 67(3):866–878, 2019