

Diving into Decimals: Navigating Privacy Challenges in Floating-Point Data with Order-Preserving Encryption



By

Areesha Nazish

Registration No: 00000360640

Supervisor

Assoc Prof Shibli Nisar

A thesis submitted to the faculty of the Information Security Department,
Military College of Signals, National University of Sciences and Technology,
Rawalpindi in partial fulfillment of the requirements for the degree of MS in
Information Security

July 2024

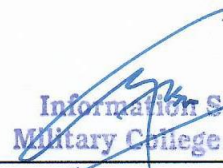
THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS Thesis written by Miss Areesha Nazish, Registration No. 00000360640, of Military College of Signals has been vetted by undersigned, found complete in all respects as per NUST Statutes/Regulations/MS Policy, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS degree. It is further certified that necessary amendments as pointed out by GEC members and local evaluators of the scholar have also been incorporated in the said thesis.


Signature: _____ 

Name of Supervisor: Assoc Prof Shibli Nisar

Date: 29 August, 2024

Signature (HOD): _____ 
Information Security
Military College of Sigs

Date: 29 August, 2024

Signature (Dean/Principal) _____ 

Date: 13/09/2024

NATIONAL UNIVERSITY OF SCIENCES & TECHNOLOGY
MASTER THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by **NS Areesha Nazish, MSIS-20 Course** Regn No **00000360640** Titled: "**Diving into Decimals: Navigating Privacy Challenges in Floating-Point Data with Order-Preserving Encryption**" be accepted in partial fulfillment of the requirements for the award of **MS Information Security** degree.

Examination Committee Members

1. Name: **Asst Prof Shahzaib Tahir**

Signature: 

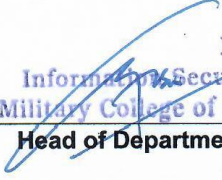
2. Name: **Engr Bilal Ahmed**

Signature: 

Supervisor's Name: **Assoc Prof Shibli Nisar**

Signature: 

Date: 29 August, 2024


 HoD
 Information Security
 Military College of Sigs
 Head of Department

29 August, 2024
 Date

COUNTERSIGNED

Date: 13/09/2024




 Dean

CERTIFICATE OF APPROVAL

This is to certify that the research work presented in this thesis, entitled "Diving into Decimals: Navigating Privacy Challenges in Floating-Point Data with Order-Preserving Encryption." was conducted by Aresha Nazish under the supervision of Assoc Prof Shibli Nisar. No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the Military College of Signals, National University of Science & Technology Information Security Department in partial fulfillment of the requirements for the degree of Master of Science in Field of Information Security Department of Information Security National University of Sciences and Technology, Islamabad.

Student Name: Aresha Nazish

Signature: 

Examination Committee:

a) External Examiner 1: Asst Prof Shahzaib Tahir. (MCS)

Signature: 

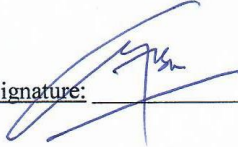
b) External Examiner 2: Name Major Bilal Ahmed. (MCS)

Signature: 

Name of Supervisor: Assoc Prof Shibli Nisar

Signature: 

Name of Dean/HOD. Assoc Prof Dr. Muhammad Faisal Amjad

Signature: 

AUTHOR'S DECLARATION

I Areesha Nazish hereby state that my MS thesis Diving into Decimals: Navigating Privacy Challenges in Floating-Point Data with Order-Preserving Encryption is my own work and has not been submitted previously by me for taking any degree from National University of Sciences and Technology, Islamabad or anywhere else in the country/ world. At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Student Signature: 

Name: Areesha Nazish

Date: 16 August, 2024

PLAGIARISM UNDERTAKING

I solemnly declare that research work presented in the thesis titled **Diving into Decimals: Navigating Privacy Challenges in Floating-Point Data with Order-Preserving Encryption** is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero-tolerance policy of the HEC and National University of Sciences and Technology (NUST), Islamabad towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS degree, the University reserves the rights to withdraw/ revoke my MS degree and that HEC and NUST, Islamabad has the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized thesis.

Student Signature: _____ *Areesha Nazish*

Name: Areesha Nazish

Date: 16 - August, 2024

Dedication

This Thesis is dedicated to my beloved Parents, Siblings, Teachers, and Friends who all have been my endless source of love, encouragement, and strength. Your unwavering belief in my abilities, countless sacrifices, and relentless support have been the foundation upon which I built my academic pursuits. Without their love and support this research work would not have been made possible.

Acknowledgments

In the name of Allah (S.W.A), the Creator and Sustainer of the Universe, to whom belongs all glory and power. He alone has the authority to elevate and humble individuals as He pleases. Truly, nothing can be accomplished without His will. From the moment I stepped foot into NUST until the day of my departure, it was by His divine blessings and guidance that I was able to navigate the path of success. His unwavering support and the opportunities He bestowed upon me were instrumental in completing my research journey.

I humbly acknowledge that no words or actions can fully express my gratitude for the countless blessings He has showered upon me throughout this research period. I am indebted to His boundless bounties and am forever grateful for His divine intervention in my academic pursuits. To Allah (S.W.A), I dedicate this thesis as a humble tribute, recognizing His infinite wisdom and benevolence. It is through His mercy that I have reached this milestone, and I pray that my work may be of benefit to others and serve as a means of pleasing Him.

I would also like to express my heartfelt appreciation to my thesis supervisor, **Assoc Prof Shibli Nisar**, and **Asst Prof Fawad Khan**, for his unwavering support and guidance throughout my thesis. His knowledge, expertise, and dedication to his field have been a source of inspiration to me, and I am grateful for the time and effort he invested in my success. Whenever I encountered any difficulties, he was always available to offer his assistance and provide me with insightful feedback.

In addition, I extend my gratitude to my GEC members, **Asst Prof Shahzaib Tahir** and **Major Bilal**, for their continuous availability for assistance and support throughout my degree, both in coursework and thesis. His expertise and knowledge have been invaluable to me, and I am grateful for his unwavering support and guidance.

Lastly, all praises and thanks be to Allah (S.W.A), the Most Merciful and the Most Gracious.

Areesh Nazish

Table of Contents

Chapter 1	1
Introduction	1
1.1 Overview.....	1
1.2 Order Preserving Encryption.....	2
1.3 Motivation.....	6
1.4 Advantages.....	8
1.5 Applications.....	10
1.6 Problem Statement.....	11
1.7 Research Objectives.....	11
1.8 Proposed Work.....	12
1.9 Thesis Organization.....	12
1.10 Summary.....	16
Chapter 2	18
Literature Review	18
2.1 Overview.....	18
2.2 Literature Review	18
2.3 Critical Review.....	28
2.3.1 Probabilistic	28
2.3.2 Deterministic.....	29
2.3.3 Homomorphic	29
2.3.4 IND Standards.....	30
2.3.5 Computational Hardness.....	31
2.4.6 Complexity	32
2.4 Summary	33
Chapter 3	35
Proposed Methodology	35
3.1 Overview.....	35
3.2 System Model.....	36
3.3 Threat Model.....	37
3.4 Research Methodology.....	38
3.5 Problem Analysis.....	39
3.6 Proposed Methodology.....	40
3.6.1 Design of the Hybrid OPE Scheme using Approximate Common Divisors.....	40
3.6.2 Algorithm Development.....	41
3.6.2.1 Key Generation	42
3.6.2.2 Encryption with Normalization.....	43
3.6.2.3 Decryption with De-normalization.....	44
3.7 Dataset Description.....	45

3.8 Security Analysis.....	46
3.8.1 Security of the Scheme.....	46
3.8.1.1 Security models	47
3.8.1.2 One-wayness.....	48
3.8.2 Enhancement of Security with Floating-Point Data	48
3.8.3 Security Benefits of the Hybrid GACD-Based OPE Scheme Against Various Attack Types.....	49
3.8 Summary.....	50
Chapter4.....	52
Analysis and Results.....	52
4.1 Overview.....	52
4.2 Experimental Results	52
4.2.1 Comparison of Encryption Time	55
4.2.2 Comparison Decryption Time.....	56
4.2.3 Comparison of Range Query Time.....	57
4.3 Ciphertext Expansion Factor.....	58
4.4 Evaluation and Analysis.....	59
4.5 Comparative Analysis.....	60
4.6 Summary.....	62
Chapter 5.....	64
Conclusion and Future Work.....	64
5.1 Overview.....	64
5.2 Conclusion of Research.....	64
5.3 Future Work.....	65
5.4 Summary.....	67
References.....	68

List of Figures

Figure 1.1 Working of Order-Preserving Encryption (OPE).....	5
Figure 1.2 Thesis Organization	15
Figure 3.1 System Model Architecture for Order-Preserving Encryption (OPE) in Cloud Environments.....	36
Figure 3.2 Architecture of the Proposed Hybrid OPE Scheme.....	41
Figure 4.1 Average encryption times.	55
Figure 4.2 Average decryption times.	56
Figure 4.3 Average range query times.	57
Figure 4.4 Average total execution times.....	60

List of Tables

Table 1.1 Comparative Analysis of Encryption Methods and Order-Preserving Encryption (OPE).....	4
Table 2.1 Comparison table to analyze different order-preserving schemes concerning security and performance.....	33
Table 4.1 Performance analysis of the proposed Hybrid GACD-based OPE and existing GACD-based OPE scheme.....	54
Table 4.2 Ciphertext expansion factor for given bit lengths.....	58
Table 4.2 Comparison table to analyze existing and proposed OPE schemes concerning security, performance, and functionality.....	61

List of Abbreviation

OPE	Order Preserving Encryption
AES	Advanced Encryption Standard
HPC	High-Performance Computing
SGX	Software Guard Extensions
MPC	Secure Multiparty Computation
FHE	Fully Homomorphic Encryption
RNS	Residue Number System
TFHE	Fast Fully Homomorphic Encryption
HE	Homomorphic Encryption
MPC	Multi-Party Computation
IND	Indistinguishability
CPA	Chosen Plaintext Attack
CCA	Chosen Ciphertext Attack
OCPA	Ordered Chosen-Plaintext Attack
OCCA	Ordered Chosen-Ciphertext Attack
DBMS	Database Management System
CT	Cipher Text
PT	Plain Text
BYOE	Bring Your Own Encryption
LWE	Learning with Errors
SVP	Shortest Vector Problem

LAN	Local Area Network
KPA	Known-Plaintext Attacks
COA	Ciphertext-Only Attacks
WAN	Wide Area Network
HPC	High-Performance Computing
RSSE	Ranked Searchable Symmetric Encryption
GACD	General Approximate Common Divisor
SOPE	Semi-Order Preserving Encryption
MDOPE	Multi-Dimensional Order-Preserving Encryption scheme

Abstract

Order preservation is useful for encrypted databases because it provides a means for range queries on encrypted data while preserving the order of the plaintext in the ciphertext. This characteristic is very effective for applications that ask for data to be searched, and sorted and where other operations can be carried out within the encrypted data without the data being decrypted first. OPE is used in several encrypted database systems to enable efficient and secure search on the encrypted data keeping the original order of the data intact. This thesis proposes a novel hybrid OPE scheme based on the General Approximate Common Divisor (GACD) problem. Our approach enhances the earlier utilized OPE methods, which were restricted to solving problems of integer data nature in the past. Just like in previous work, our scheme also supports only integer data; in this work, we expand it to floating-point data, hence improving its applicability. Further, compared with other older OPE methods based on a less reliable security model, our scheme utilizes the computational hardness of the GACD problem. This approach not only improves the protection against some potential risks but also raises the standard of data protection. Our GACDP scheme shows high efficiency as it takes an average time for the encryption and decryption operation with minimal time complexity equals $O(1)$. Results demonstrate that our scheme enhances the functioning of both integer and floating-point data types and proves to be superior to previous approaches.

Introduction

1.1 Overview

The first chapter deals with a general introduction to Order-Preserving Encryption (OPE), the importance of which is predicated on its ability to support operations like sorting and range queries on the encrypted data maintaining the order of plaintexts. This capability is quite useful in different solutions, specifically those involving secured databases as well as the cloud where data confidentiality along with processing speed is crucial. The chapter also spells out the major goals of the research as follows: the formulation of the new uniform OPE method that can process the integer and float data; to improve security; and to ensure that efficient queries are performed. This research is motivated because new ways of data processing using security measures should not worsen usability and this issue is topical for environments that process a great amount of information that is considered sensitive. Then the benefits of the proposed work are described to demonstrate how the new OPE method which is suggested in the work is better than existing techniques by providing security and functionality for incorporation of keys. This section also emphasizes the concerns where the floating-point data is continuously adding to the vulnerability of the cipher space. The problem statement is then given that seeks to define the main issue at hand, and this is the ability to devise an encryption scheme that will preserve the order of data while on the same note, ensuring that the data is secure, especially

when the idea is implemented on both discrete and continuous data. After this, a brief account of the planned work is given explaining the strategy adopted in the present work that aimed at the design and implementation of the improved OPE method. This also involves explanations of the methods that were employed to realize the laid down research objectives as well as the anticipated results. Last of all, the current chapter is rounded off by the presentation of the thesis organization followed by a summary of the following chapters and their implication to the general research. This section enables the reader to have a checklist of the thesis and how each segment contributes to the next one to solve the research problem systematically.

1.2 Order Preserving Encryption

As far as outsourcing is concerned, computation has become more vital in businesses, governments, and academia. Most of these computations are carried out in distributed computing environments including clouds, grids, or HPC clusters. However, dealing with sensitive data in such environments induces some difficulties. Security of computation on the above platforms is very paramount. It must be noted that traditional cryptography can effectively protect data whether it is in storage or during transmission. However, many encryption techniques such as the AES (Advance Encryption Standard) do not enable arithmetic operations conducted on the ciphertext [1]. As a result, if data is encrypted through AES, it must be decrypted to enable computation and expose the data. One of these solutions is the secure computation hardware, for instance, Intel's Software Guard Extensions (SGX) [2]. SGX enables the data to be decrypted and processed within the safe and confidential memory region or processor. Nevertheless, as SGX depends on certain hardware, it may not be as efficient for any environment with a different set of stocks [3].

Another method is in protocols called secure multiparty computation (MPC). MPC is a concept that enables two or more parties to solve a certain function while none of the parties is aware

of one another's inputs. Also, there has been a massive development in the realization of MPC in practice [4]. Second, while much progress has been made in individually scalable MPC, this subfield's growth has been less substantial, largely because of issues with communication and computation [5]. Practical optimizable MPC protocols tend to rely on some form of hardware [6].

Another method is homomorphic encryption which will allow computations to be made directly on the encrypted data [7]. The outcome does not change data is encrypted during the analysis and, when decoded, the analyzed data is accurate and identical to the original. Unlike the above method, this does not require any specific hardware and applies to various distributed systems; moreover, the computing party deals only with ciphertexts [8].

As for the solutions, fully homomorphic encryption (FHE) has been suggested as the most globally applicable one. However, FHE is currently inapplicable since it is based on computation over arithmetic circuits that are unfortunately space-consuming [9]. In the case of each gate in the circuit, the 'data' it is working on is encrypted bits, and therefore each of the ciphertexts is larger than their corresponding plaintexts and the computational time for the ciphertexts is considerably lesser than it is for the plaintexts [10, 11, 12]. Hence, somewhat HE, solely homomorphic for some specific input functions, is more feasible to an extent in the present world [13]. For uses that require kind of tasks such as sorting and comparisons of data, there is a need to employ a cryptosystem that supports the homomorphic operations on the ciphertexts. Order-preserving encryption (OPE) is relatively new to this need [14,15]. Further, it is pointed out that property-preserving encrypted databases contain certain risks waiting to be solved after recent studies [16,17].

Table 1.1 Comparative Analysis of Encryption Methods and Order-Preserving Encryption (OPE)

Category	Description	Advantages	Limitations
Traditional Cryptography [1]	Standard encryption for messages such as through AES preserves data but does not allow for computations to be done on the encrypted messages.	Security of data stored in databases as well as transport of data between databases.	Need decryption for any computation and, thus, makes the data itself exposed and may cause an issue with privacy.
Hardware-Based Solutions [2,3]	Deploys specific equipment for instance Intel’s SGX to perform efficient computations in secure zones.	When used, it creates a secure environment for computation while keeping the data in an encrypted format that cannot be read.	Relies on a particular piece of hardware, which cannot be used in all computing contexts, especially not in heterogeneous ones.
Advanced Encryption Techniques [4,5,6,7,8,9,10]	Includes methods such as Homomorphic Encryption (HE) and Secure Multiparty Computation (MPC) that enable computation on encrypted data.	Enables one to perform computations on encrypted data without revealing the plain text. Provides strong privacy guarantees.	Most of the time it is expensive from both communication and computational perspectives. Not suitable for ranged queries, sorting, or if value calculations are needed. FHE is less suitable for many operations, especially those specified here.
Order-Preserving Encryption (OPE) [14,15,16,17]	There is also a technique called Order-Preserving Encryption (OPE) which permits data to be encrypted but the order of the plaintext remains the same in the ciphertext.	Facilitates querying and range queries and performs operations such as find least/greatest, comparison value operations.	OPE is somewhere in the middle of the secure and usable continuum. On one side, it gives a high level of efficient access and operations on encrypted data but on the other side, it represents less security than other stronger encryption techniques.

Crypt DB is a valuable solution in this sense because it provides confidentiality when the queries are processed in an encrypted manner [18]. Frameworks for education, contemporary cryptography, enciphering books, and other complicated techniques for secure computation purporting to solve issues of privacy and anonymity buttress the establishment of effective encryption procedures [19, 20, 21, 22, 23].

With the growth of digital systems, it becomes critical to try and secure numeric data since it is sensitive in most industries. Techniques like encrypting data help to protect such data through the translation of plain text to a form that is not understandable by anyone. However, these traditional encryption methods hide semantic meaning from the encrypted data so that the encrypted data is useless for queries and analysis. Ordering in the encrypted domain addresses this limitation through order-preserving encryption (OPE), which preserves the ordinal structure of the numeric data while encrypting it, enabling efficient computation over encrypted data [24].

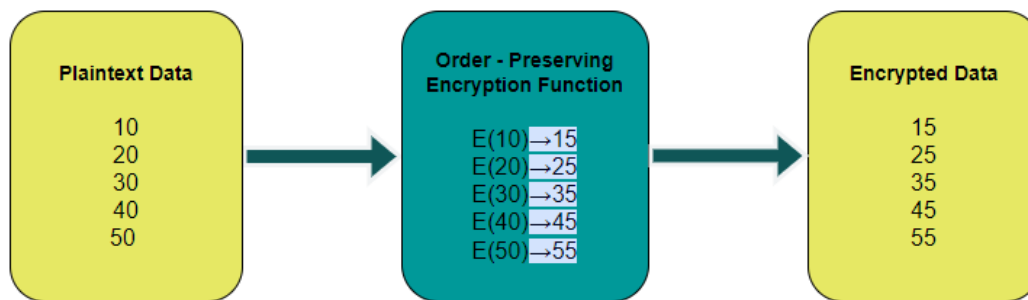


Figure 1.1 Working of Order-Preserving Encryption (OPE)

Figure 1.1 explains the scheme of Order-Preserving Encryption known as OPE, which maintains the order of plaintext values in the ciphertext. It is most useful for floating-point data, where precision can be an issue. The floating-type numbers provide real numbers with variable accuracy with an emphasis on scientific and financial operations using continuous measurements [25]. Earlier OPE schemes primarily dealt with integer operand values, which proved ineffective when it came to dealing with floating-point characteristics such as inadequate precision and having infinite characteristics [26]. To extend the potential problem domain of OPE and address the domain's security concerns simultaneously, it is possible to devise special adaptations of OPE for floating-point attributes.

The most critical trade-off of floating-point OPE is when to prioritize utilitarian and privacy concerns [27]. While the ability to maintain order facilitates logical operations on the encrypted data, this de-anonymization technique can potentially reveal statistical first-order dependencies. Visible data distribution poses a significant threat since attackers can easily find the plaintext value by analyzing the data. Schemes must be optimized in a way that they will be useful on their own but do not divulge sensitive information on the other hand.

Further, certain floating-point details, such as variable precision, carries, or edge-case peculiarities can interfere with order encoding [28]. Arithmetic instability leads to confusion of order in mapping. Pertinent, strong schemes should employ complementary measures against such vices when managing floating-point edges. Finally, performance degradation occurs in areas where large floating-point datasets are required, as in analytics. Effective encryption processes, indexing, and targeted architecture can enhance response time [29].

Modern works discuss improving order encodings regarding security by employing non-linear transforms, adding noise, and employing cryptographic techniques like homomorphic encryption [30]. Other potential benefits are also related to the improvement of precision control, possibly through the utilization of granular schemes. The operations themselves become efficient with the help of aggregate functions, while optimized data structures and hardware acceleration are used with analytics. Subsequent models may incorporate learning from various fields such as differential privacy and secure multiparty computation also [31].

1.3 Motivation

The selection of the topic " Diving into Decimals: Navigating Privacy Challenges in Floating-Point Data with Order-Preserving Encryption " is justified for several reasons, one of which is that it will serve as the central theme of this research:

- The incorporation of OPE for both integers as well as floating point numbers ensures that the same encryption model can be adopted for all the numerical data formats in each system. This makes management, configurations, and access controls easier since there is only one scheme: the complex integers instead of having one method for integers and another for floating-point data. It also makes analysis and computations to operate on data in a mixed-type environment without any ambiguity.
- Integers are often unable to continuously provide highly accurate orders across extensive areas because of the fixed precision. On the other hand, floating-point OPE allows for variations in precision and scale for encrypting integer data and the order thereof, where necessary, across a wider domain. Variable precision serves to reduce problems that may occur when encrypting integers.
- Floating point numbers can range over orders of magnitude and can be located seamlessly on a numeric scale. Thus, Floating point OPE can smoothly proceed with the encryption of any real number value at this full spectrum. This flexibility can handle different ranges and scales of data most common in the analysis datasets.
- Fixed-point representations introduce rounding errors into the data which can hide underlying plaintext values. This supplementary ‘noise’ increases integer OPE’s levels of privacy to protect against attempts at analyzing frequencies to obtain secret information.
- Floating point OPE gives fine grain and resolution options when the accuracy of the order of precision is desired, but lets the programmer choose the amount. This additional flexibility assists in enhancing the security and encoding of integer data to allow for better defenses.
- Floating point OPE enlarges the horizon on encrypted data size from very small fractions to very huge numbers. This broad span covers all the ranges of engineering and scientific data that may be needed in engineering and scientific data.

- Floating point OPE provides a possibility for meaningful mathematical operations on float data encrypted within OPE. This allows calculations such as the aggregation statistics without using intermediate decryption just as it used floating point-based analysis.

1.4 Advantages

Our newly designed OPE scheme works on two data types, integers as well as floating numbers yields various advantages:

- **Enhanced Security:** The integration of float data in Order-Preserving Encryption (OPE) works to the advantage of security since the field complexity and variability of the ciphertext increases. The continuous flotation of float values widens the scale of the cipher space and makes it more difficult to predict the initial values and choose the corresponding attack, enhancing the general protection of the information, in turn.
- **Security with Utility:** The security of integer and floating-point data during processing is maintained by the proposed OPE scheme with modest aspects of operation efficiency. The proposed OPE scheme works as a secure integer and floating-point data processing while maintaining efficient operational security. This way, the information that is needed to be kept secure is never exposed but at the same time, the essential operations can be conducted with ease. As the scheme is probabilistic and uses one-wayness to increase security, it is well suited for applications that require both security and fast processing such as financial and scientific applications that include transactions and numerical computing.
- **Greater Flexibility:** The traditional OPE scheme is replaced with a floating point OPE scheme as proposed here. One major benefit of flexibility is that it enables it to support integers and floating-point numbers. The dual capability makes the scheme well suited to these mixed data type scenarios, where the array might be used for financial transactions, scientific computation or any other of the many possible applications where a wider variety of data types

are used. Maintaining parity regarding kinds of data, the scheme guarantees reliable safety and efficiency step by step, which makes the tools more multifunctional and effective for protecting and working with data.

- **Preservation of Data Utility:** In our proposed hybrid OPE scheme, data utility is maintained. Unlike previous approaches that addressed arithmetic operations while ignoring security, the scheme also keeps the data encrypted for integer or floating-point numbers, as required, to preserve the security of the data while still permitting the data to be used for operations like range queries, k-queries (smallest, largest value) for instance. This rift between the strength of the encryption on one hand and the usability of the data on the other, it is most sensitive in cases where data manipulation must be precise and fast such as in financial analysis and scientific processes.
- **Scalability:** The presented scheme is therefore designed to be scalable in that it can be scaled up and down depending on the amount of data that is being processed to achieve good performance. In this scheme, both integer and floating-point data types are used so that in future as the size of datasets increases, the effectiveness of the encryption and decryption is not compromised. This scalability ensures that the scheme can be used in large-scale systems like clouds and big data analytical systems and without really compromising greatly on time and security.
- **Improved Performance in Cloud Environments:** The proposed scheme is optimized for cloud environments, enhancing the performance of executing queries directly on encrypted data. By efficiently managing both integer and floating-point data types, the scheme enables faster and more accurate query processing without compromising data security. This improved performance is particularly beneficial in cloud-based applications where large volumes of sensitive data need to be processed swiftly and securely, ensuring that the advantages of cloud computing are fully realized while maintaining robust encryption standards.

1.5 Applications

Order-Preserving Encryption abbreviated as OPE for integer and floating-point data type is useful in several ways in different fields such as:

- **Financial services:** This way, OPE can assist in protecting assets such as the balance in checking or savings accounts, value of stocks and bonds, insurance compensation, and so on, but will enable cluster analysis at the same time.
- **Healthcare:** OPE allows the problem of assessing patient's test results such as laboratory values and sensor readings while keeping patients' personal information secure.
- **Recommender systems:** OPE does acquire user ratings & preferences so that it retains control of them while at the same time permitting collaborative filtering as well as personalization.
- **Supply chain:** OPE can perform analysis of any shipment location, inventory status, etc. in a supply chain network and simultaneously protect the relevant logistics information.
- **Smart grid:** To maintain privacy in customer electricity usage data while permitting data analysis in general for grid management, OPE is employed.
- **Scientific research:** It is useful for time series analysis, and statistical hypothesis testing on encrypted sensor/SIM data derived from fields such as meteorology, physics, epidemiology, etc.
- **Government sector:** OPE preserves numeric security information such as census, economic statistics, defense spending, etc., while maintaining the analytical usefulness of the data.

Thus, in general, OPE allows us to perform a wide range of numeric analytics in many verticals while keeping personal information safe; therefore, it is incredibly useful for security-critical

tasks with float or integer data. Most particularly it proves very helpful in k queries (i.e. Smallest, largest, middle value), range queries, comparisons, and so on.

1.6 Problem statement

Creating a unified encryption method that solves the problem of encryption of both integer and decimal numbers is one of the main hurdles that have been encountered. The objective is the ability to manipulate the data in such a way that its order remains conserved and thus enables meaningful operations and analysis while at the same time maintaining data confidentiality. Also, there is a question of whether this method can preserve data privacy and be functional when it is applied to continuous data. This thesis intends to meet these challenges by proposing and analyzing an encryption method that would concurrently maintain the order of the records as well as ensure the highest levels of confidentiality and security of information.

1.7 Research Objectives

It enables the generation of a new OPE scheme that will be capable of encrypting integers and floating-point numbers along with the capability of maintaining the actual order of the encrypted numbers. The objective targets the existing deficiency in the encryption algorithms that are tailored to perform the computation on integers and floating-point numbers.

- Developing a unified OPE framework that specifically caters to both integer and floating-point data types, aiming to provide a seamless encryption mechanism that maintains order preservation while addressing the nuances of both numeric domains.
- Explore how the inherent imprecision of floating-point representation can be strategically leveraged to strengthen privacy and deter inference attacks, addressing challenges that are unique to floating-point data.

- Investigating how using floating-point-based Order-Preserving Encryption (OPE) enhances privacy protection in scenarios involving continuous data, where preserving the original order of values is crucial for meaningful analysis.

1.8 Proposed Work

The research will design a hybrid general approximate common divisor-based (GACD) OPE scheme that has the features of simplicity in the concept level and constant time complexity $O(1)$, which will be equipped for supporting both integer and floating points. The proposed hybrid GACD-based OPE scheme explored above provides more functionality and security than the existing ones. From the functional perspective, it offers a single unified approach that can be used both for integer and floating-point data, and the order of encrypted data must remain the same as of the original data ensuring that encrypted datasets remain sortable and query-able with a specified range. Thus, this flexibility is beneficial in the particularities of application calls that need to process and output consistent data regardless of the data type. Concerning its security, the scheme builds proactively on the fact that floating-point data is continuously exchanged and therefore the cipher space becomes larger and harder for an opponent to look for patterns to decipher the original numbers which makes data even more concealed and less susceptible to threats. Such improvements are vital to guarantee that not only the scheme will be able to meet the requirements of securing the data processing but also will offer rather a high level of protection against different types of cryptographic attacks while still maintaining the commercial value of the information that is to be protected.

1.9 Thesis Organization

This research paper's subsequent sections provide a thorough summation of supporting evidence for the main premise. The sections outlined below are depicted in Figure 1.2. This is organized into five chapters building up to a coherent continuous work that addresses the

extension of OPE to handle both integer and floating-point data, its efficiency analysis, and implications.

Chapter 1: Introduction

This first chapter provides an overview of the field of Order-Preserving Encryption (OPE), which has initially been defined for integer data only, and the need and obstacles related to generalizing this approach to floating-point data. This gives some information about OPE, beginning with the problem statement, which describes the realities and importance of incorporating floating-point numbers into OPE solutions. The objectives are outlined very well and focus on the need to establish a single encryption mechanism that will be able to support multiple data types. The motivation for the research is described, stating the increasing demand for powerful encryption techniques as data types require more spectrum in different applications. In the final section of the chapter, a brief introduction to the proposed method is provided before it is introduced in more detail in the following chapters.

Chapter 2: Literature Review

Chapter two focuses on a critical analysis of literature in the context of OPE. It provides the historical context of OPE and reviews the very early work which is concerned with the optimization of integer data, and the more recent and current work on how to extend OPE to handle floating point numbers. Considering this, it offers a critical literature review to justify the proposed improvements, noting that current methods are grossly inefficient when applied to floating point data.

Chapter 3: Proposed Methodology

This chapter concludes with a detailed explanation of the method used in creating the improved OPE scheme. It starts with the design of the architectural system and the algorithmic structure of the new encryption scheme, which is to handle both integers and floating-point information.

Additionally, an explanation of the security mechanism incorporated in the scheme is provided to highlight the added layer of security of data for the users. Some of the aspects captured in its implementation or design include the software tools and programming languages that have been used and this helps the reader to understand the feasibility and realistic nature of the proposed scheme.

Chapter 4: Results and Analysis

The fourth chapter overviews the conclusion derived from the study on the performance evaluation and security assessment of the developed OPE scheme. It describes how the tests were performed and in what environment, including the parameters used to measure the degree of success of the scheme with integer and floating-point numbers. The chapter also presents an overview of the encryption scheme quality, based on its performance, evaluating its security against various cryptographic attacks. A comparison between the computational complexity and efficiency of the proposed floating-point OPE and the conventional integer OPE has also been performed to exhibit the progress and improvements with the new scheme. lastly, some comparisons between the functionality and security are made in this chapter, and we show that the proposed scheme achieves enhancement of the prior methods.

Chapter 5: Conclusion and Future Work

Provides a brief conclusion of the thesis and highlights the main contributions of the study towards the advancement of cryptography and data security. It explains the constraints of the study and the extent of the current research, ways that provide a realistic perspective on the work done. It also lays out the potential of extending future work.

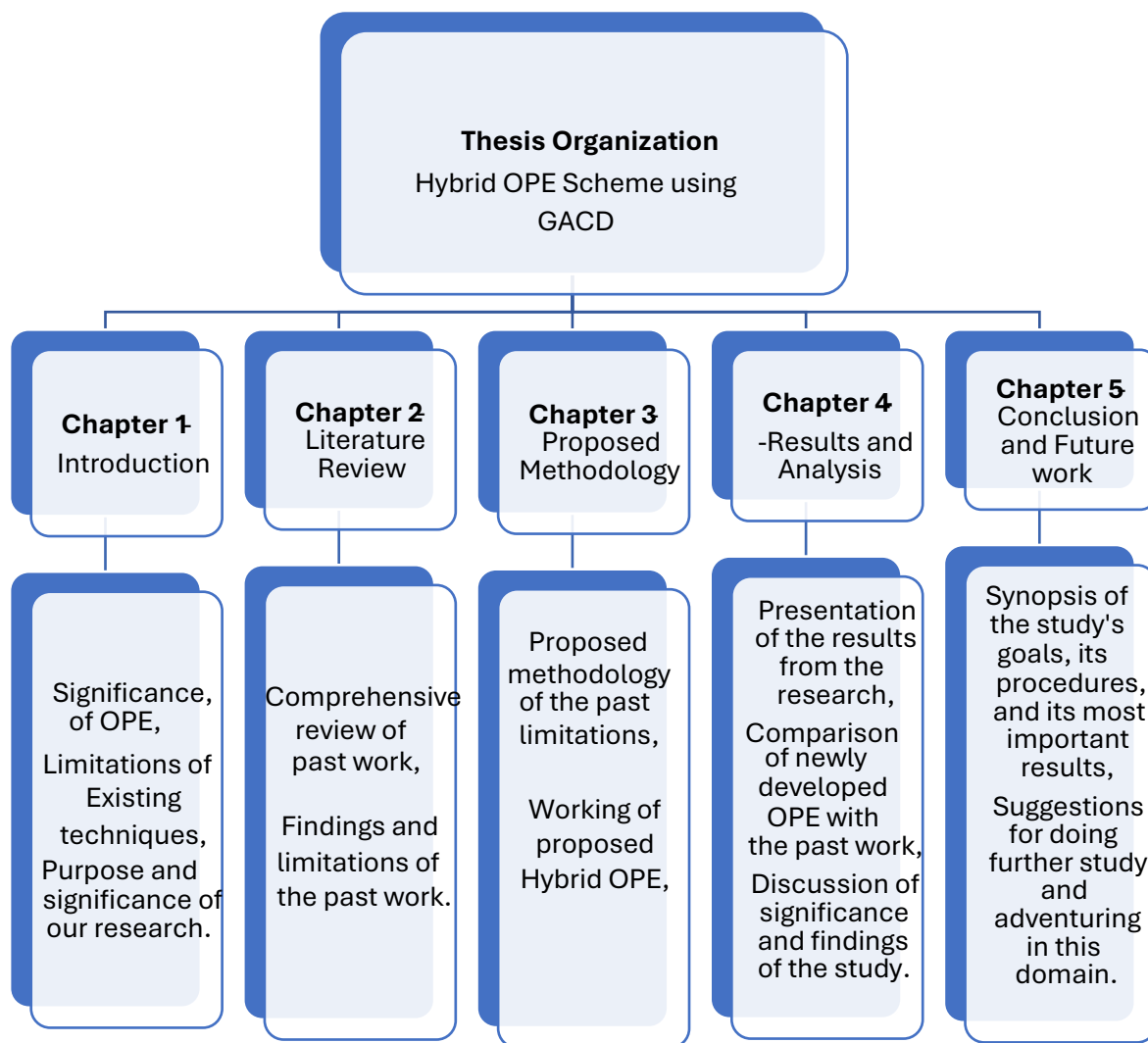


Figure 1.2 Thesis Organization

1.10 Summary

In the thesis, Chapter 1 lays the groundwork for this thesis by presenting Order-Preserving Encryption (OPE) and its generalized form, which includes both integers and floating-point numbers. The chapter starts with an extensive overview of the traditional OPE systems that have been developed and implemented mainly for integer data, as they process ordinal information most naturally. It emphasizes the need to include floating-point data, which is habitually used in scientific computation, financial applications, and other settings that involve precise decimal computations while maintaining privacy and security. The problem statement derived in this chapter concerns the difficulties in building an OPE system capable of preserving the inherent order of float numbers while considering difficulties and threats concerning security and privacy. It highlights the challenges in maintaining the order of data and at the same time making amendments to the security features to counter hypothetical and real cryptographic assaults. The objectives laid out in this chapter are twofold: firstly, to provide a single method for encryption of a variety of data types within the same framework, and secondly to consider the key management schemes which are dynamic to the level of encryption needed for security. These objectives stem from the existing need that is filled by most encryption technologies that are good in either security, performance, or handling of multiple forms of data. Interest in this research comes from the elevation of the necessity for strong encryption mechanisms that can guarantee maximum security while retaining the functionality of information. The current chapter describes the real-world implications of such an advanced OPE scheme, including the potential of such a scheme to transform data security in sectors where data sensitivity is critical. This gives an anticipation of the subsequent details of the technical development and application of the proposed OPE enhancements. In conclusion, Chapter 1 provides appropriate groundwork for the thesis by presenting insights on impediments, aims, and reasons for the study. It also predetermines the direction to discuss

more sophisticated cryptographic techniques for addressing the emerging requirements in the context of data protection in the conditions of the digital environment.

Literature Review

2.1 Overview

The literature review critically evaluates past OPE schemes, focusing on attributes such as determinism vs. probabilism, homomorphic properties, IND standards, computational hardness, and complexity. Each attribute provides a lens through which the security and functionality of OPE schemes can be assessed. The basic property of OPE schemes is that it is possible to compare and sort the data that is encrypted with the help of the applied scheme without decrypting it. From the literature, this paper critically evaluates previous OPE schemes by evaluating them concerning several key features; determinism or probabilism, homomorphism, indistinguishability standards, computational hardness, and complexity. The literature review chapter outlines various methods of developing the OPE and presents a comparative analysis of the known methods with the context of their efficiency and security.

2.2 Literature Review

Traditionally, for such scenarios, indexed encryption techniques have been derived to facilitate querying the encrypted data and at the same time ensure data privacy. These methods, as a rule, make it possible to perform equal searches, that is, the document which contains the keys to a query is retrieved. However, searchable encryption schemes have limitations where equality searches are the usual option, and searches such as comparisons cannot be performed. This limitation prompted the development of Order-Preserving Encryption (OPE) by Agrawal et al. in their study [24] in the year 2004 where the ciphertext maintains the order of the plaintext,

thereby allowing comparison functions on the encrypted numeric data. Thus, while deriving several operational benefits, Agrawal et al. failed to define a formal security that governs their OPE scheme.

To fill this gap, in 2009, Boldyreva et al. [25] suggested the first clear security notion for OPE and presented that it is possible to provide security proof under the framework of IND-OCPA (Indistinguishability of Order Preserving Encryption under Chosen Plaintext Attack). However, their analysis revealed a critical limitation: Originally the operations that are performed in OPE schemes to obscure the data that is to be transmitted leak information particularly half of the plaintext's bits, thus the schemes are prone to certain types of attacks. In addition, Boldyreva et al. even stated that no OPE scheme can be IND-CPA secure and that no deterministic encryption can be Secure under IND-OCPA, which led them to the understanding that no ideal security is possible for OPE.

The paper [26] by Wang et al. presents solutions for secure and efficient rank of keyword search on encrypted cloud data. There are large quantities of searchable encryption schemes in the literature that only allow for Boolean search, which can lead to ineffective retrieval and excess network load. To this end, Wang et al. elaborate an order-preserving solution for this problem; it improves comprehensibility by ordering the resulting answers in terms of relevance. This method deals with the concealing of keywords and asserts to safeguard the keyword against IND-OCPA attacks. This is expanded by presenting a new definition for ranked searchable symmetric encryption (RSSE) and a construction based on order-preserving symmetric encryption (OPSE). This advancement offers very strong security promises and proof and does substantial experiments thus enhancing the realism of privacy-preserving distribution of data to the cloud. While striving to improve the usability of the encrypted databases, Liu et al. [27] proposed an efficient scheme for the formation of order-preserving indexes that allows for effective organizing of the range of searches in encrypted data. The kind of method they have

proposed revolves around the generation of indexes that have random noise added to them to come up with linear expressions that would consequently ease how range queries are conducted. This approach employs the AES algorithm combined with the index-preserving method; however, it cannot be considered as an encryption system. The first strength of the work presented by Liu et al. is to facilitate a range of queries on encrypted data while not putting the original content into the spotlight. The random noise that is introduced during the indexing is such that, although the indexes remain usable for querying the data, it is not possible to generate the plaintext from the indexes. This method gives a research way of responding to the issues of asking for information from the encrypted database to enhance the security of data and privatization of the latter in the cloud environments. The authors claim that their scheme is information-theoretically secure, which means that this type of security is based on the theories that lie at the foundation of information theory rather than on the specific assumptions that have to do with the capabilities of computing. Thus, randomizing the indexes and enabling immediate programming of basic indexing expressions, the scheme led to hiding the distribution of the initial data, which contributes to privacy.

In this paper, Liu et al. [28] extend their previous work on the order-preserving indexing scheme through the introduction of a nonlinear technique that can improve range queries on encrypted data. This updated method has the nonlinear transforming function to perform homomorphic operations on an encrypted database. The scheme increases security because range queries are effectively addressed even if there are many duplications of plaintexts and plaintext distribution is also hidden from indexes. Most importantly, it doesn't assume the data distribution, the range of observations, or the number of records in the database as does ordinary statistical analysis, so it is ideal for databases that change over time. This advancement provides a more competent means of Privacy of Data in Cloud Computing.

Popa et al. [29] put forward a stateful interaction-based scheme for designing binary indexes of encrypted data that hides all the information and only allows the order of plaintext to be exposed. They ensure that the obtained security is optimal under the IND-OCPA model according to which ciphertexts should reveal no information about plaintexts except their order. Although the protocol provided the high-security measures as brought out here, it cannot be directly classified as the OPE scheme. As for the methods described in the paper of Popa et al., let us notice that the authors propose an order-preserving encryption scheme that aims at ideal security. This scheme leverages homomorphic ciphertexts which are time-varying for a small domain of plaintext values, the requirement for achieving the ideal security. This interactive protocol which is more complicated than other OPE schemes offers a better performance within the range of 1-2 times than the OPE schemes currently in use. Originally tested and further analyzed in benchmarks and an encrypted MySQL database, the proposed scheme proves that the ideal security can be obtained and maintained along with efficiency, thereby contributing to the development of the state of the art in secure order-preserving encryption.

Kerschbaum et al. [30] describe OPE strategies that use binary search trees that would increase the efficiency of the method while preserving security. It is an approach whose objective is to attain maximum security within a realm of Theoretical Framework known as the indistinguishability standard of IND-OCPA. It is shown that this scheme can defeat the standard security game, and it constructs CryptDB for efficiently operating database tasks. In their paper, Kerschbaum et al. respond to the high insertion costs which are present in the current ideal-secure OPE schemes like Popa et al' high encryption costs. From the works of the authors on the OPE scheme, the readers will learn how to develop one that will meet the following two factors: ideal security and efficiency. Based on Reed's work on the average height of random binary search trees, it is smoothly incorporated with adaptive encryption used in Crypt DB. Their approach appears to have a performance advantage of up to 81% over the

LAN environment and 95% of the WAN environment; hence, their work is more efficient compared to previous works. Thus, this advancement leads to the significance of developing more applied research in order-preserving encryption since the amount of computational overhead within insertions has been minimized.

In their work, Krendelev et al. [31] discuss two different OPE schemes for their investigation target. The first scheme focuses on arithmetic coding while the second one implies the usage of the sequence of the matrices for encryption. Both schemes are intended to solve the problem of cryptographic hardness concerning the requirements for protecting messages from ciphertext-only as well as chosen plaintext attacks. Still, they don't quite touch on privacy issues.

For context, their paper begins with a discussion on the state of existing work on OPE which the authors describe as recent advancements. The first type preserves the order of the encrypted data using arithmetic coding and the second one applies matrix operations for encryption. While the schemes make various nodes secure theoretically, they have advantages and disadvantages, the major one being the inability to protect privacy.

Chenette et al. [32] examine the use of MOPE with two algorithms aimed at SQL database queries, which primary weaknesses prior techniques have. They improve basic OPE by adding a secret modular offset but note that not very smart implementation can reveal this offset, which shrinks the security advantage. Their work includes two new query execution algorithms: One model is designed for balanced distribution while the other is for skewed distribution; however, the latter introduces the least significant bits of information leakage. This paper presents new security modes for MOPE, formally discusses the security of their proposals, and showcases the efficiency of their schemes through an implemented system that integrates with real databases, tunes SQL query optimization against various datasets, and offers higher security.

Jayashri et al. [33] present a MOPE that seeks to improve the basic techniques of OPE. Their mechanism is to add a unique offset to every message before encryption, which remains the same across multiple messages, and uses the Multivariate Hypergeometric Distribution (MHGD) to enhance the system's security level. The scheme intends to obviate leakage of plaintext locality thus improving the security over and above the basic OPE. Jayashri et al. state that based on experience, the proposed methodology is pseudorandom and POPE-secure. Its research focuses on the problem of the efficient searching of encrypted data within cloud infrastructures and aims at enhancing the security of existing methods of matrix operations on private elements, or MOPE. The approach is proved theoretically, and it aims to eliminate any additional information leakage about the position of the plaintext in the cloud and thus provides a higher degree of security in searchable encryption systems in the cloud.

Yang et al. [34] present a new method of semi-order preserving encryption (SOPE) to solve the problem of searching the encrypted text because they pointed out that the traditional OPE has some problems in this aspect. To this end, the novel type of functional encryption called SOPE is proposed in this paper, which besides offering higher security in comparison with OPE has lower storage costs but lower precision. The outcomes of this research focus on the effects of implementing different degrees of semi-order-preserving maps degree on precision, security, and ciphertext expansion where enhancing the degree of degree leads to increased security, decreased precision, and increased expansion of the ciphertext. Their approach is also flexible and can therefore provide a good balance of these factors by adjusting the value of the degree to suit the needs of the real-world application. The authors propose the theoretical aspects of such a method called SOPE and its real implementation that proves that the suggested pseudo-range-querying method can work with encrypted data, preserving a sufficient precise security ratio.

Khoury et al. [35] propose their new OPE scheme and compare it with other Approaches for OPE. A comparison is made with the Dyer et al. scheme on how well their new OPE scheme performs in terms of rank preservation on the input plaintexts as well as efficient encrypted data search operations. This leads to lower power consumption and prompts the results of new applications in the wireless networks by using the proposed scheme: to permit the protected data aggregation operations such as MAX & MIN without decryption. However, the study also informs the following error that has brought inefficiency to the use of the scheme, and it is because of the symmetric key that changes its key at a certain interval of time. The authors describe the state of the art of OPE schemes, evaluate their performance and security, and present the developed scheme as more efficient and less complex for application in WSN.

Quan et al. [36] have proposed a mutable Top Order-Preserving Encryption (TOPE) to achieve a better query processing time and to reduce the leakage of data. Their scheme focuses on solving the problem of performing query-by-example sorted top-k queries on encrypted datasets essential in developing applications such as the PageRank ranking system, healthcare data analysis, and decision-making in cloud infrastructures. For the TOPE scheme, there are main two phases of implementation. First, it allows top 1 queries like finding min or max while preventing the disclosure of the other entries. Second, it extends to support top-k queries, meaning that ciphertexts corresponding to the top-k values are kept in the relative order of top-k in the encrypted domain, and other ciphertexts for the non-top-k values are permuted in some meaningless order. This approach keeps the information leakage in check as compared to the other methods of OPE. The authors clearly and formally specify the security of TOPE under mixed indistinguishability against top-ordered chosen-plaintext attacks and analyze the work well. The experiments that they conducted in search of top-k using both synthetic and actual data are Quality, specificity, and relevance of the TOPE In their experimental results, the

authors that used synthetic as well as actual datasets assert that TOPE offers search performance for top-k queries that is almost on the par with the time it takes to query plaintexts. To reduce the cost and augment security, Reddy et al. [37] proposed a new scheme named Secure and Cost-Efficient Order-Preserving Encryption, known as SCOPE. This scheme is meant to be directly more secure than Modular Order-Preserving Encryption (MOPE) since it has solved all the problems of MOPE. In particular, the SCOPE guarantees confidentiality according to a security definition that relates to ordered chosen repeated plaintext distribution attacks (IND-OCR-PDA). This new security model is designed to enhance the security of encrypted data against different types of attacks and at the same time contain costs. In the authors' view, the enhancements given by SCOPE are enough to own up to a better security stance than the previous OPE schemes and act as a reliable method of securing cloud data.

Dyer et al. [38] introduced two order-preserving encryption (OPE) schemes based on different computational hardness problems: there is made the General Approximate Common Divisor Problem (GACDP) and the Decisional Polynomial Approximate Common Divisor Problem (DPolyACDP). The first scheme of theirs based on the GACDP is seen to be significantly different from other OPE schemes in so far as it does not use the concept of security games but computational hardness. The strategy guarantees that it ensures safety against IND-KPA (Key-Prefix Attacks) or CPA (Chosen Plaintext Attacks), as well as a type of CCA (Chosen Ciphertext Attacks) since the datasets are static in the actual context. It must be remarked that the GACDP-based scheme offers more efficiency and uses $O(1)$ operations for both encryption and decryption. As for efficiency, in addition to near-optimal information leakage, this is substantiated by computational experiments, including MapReduce computation over encrypted data, for which better execution time is observed compared to other related OPE schemes.

Ahmed et al. [39] add more protection to the encrypted databases using a novel semi-order-preserving encryption (SOPE) technique which is based on the general order-preserving encryption strategy. Their approach can be seen to solve the main problem of OPE schemes where normally the data's order contains sensitive information that does not get protected. The performance goal of the SOPE in combating ciphertext-only attacks (COA) and known-plaintext attacks (KPA) needs to be achieved. Adding up extra layers of protection on the ordinary OPE enhances security in data and the sequence of encrypted data. This advancement facilitates managing the comparison operations sufficiently without decreasing the performance despite those arguing that the order of data is sensitive. The weakness of various OPE schemes, especially in the aspect of the IND-OCPA, has been pointed out and to address this problem Chen et al. [40] propose the BOPE scheme. BOPE consists of two components: Two types of algorithms included in this patent are a searching algorithm and an updating algorithm. This search algorithm also uses another data structure known as the boundary tree to minimize the area and the search operations performed. The updating algorithm deals with stale encodings by deciding when to renew the lookup table, so as not to frequently update unused entries. Compared to the previous mutable OPE schemes, these flows provide a substantial improvement in granularity and response time with a greater than 10% improvement in realistic tests.

In the paper by Shen et al. [41], the authors establish a stateless non-deterministic order-preserving encryption (OPE) that they claim to be more efficient yet highly secure. Unlike stateful schemes that depend on state information, and which are close to mutations, their stateless solution only depends on encryption keys, which provide more performance. This scheme protects Indistinguishability for adaptively chosen plaintext attacks under commitment, which means that the attackers are allowed to choose the challenge pairs before the keys for encryption are created. Hence, using the prior knowledge of the data the scheme provides good

security for the static datasets, and the leakage of the access pattern is significantly minimized at the cost of not being able to provide IND-CCPA security for the dynamic datasets. Different types of attacks are briefly described and methods to avoid them and to eliminate the leakage of access patterns are suggested, thus improving the real-life security of the encryption algorithm.

Yang et al. [42] propose a Frequency-Hiding Order-Preserving Encryption (FH-OPE) to improve the security aspect due to frequency-based threats. This scheme, which is aimed at security against Indistinguishability under Frequency and Ordered Chosen Plaintext Attack (IND FA-OCPA), provides an increased level of security in comparison with basic IND-OCPA schemes. As earlier models such as Kerschbaum's FH-OPE were accused to be imprecise and costly in terms of overhead, Yang et al. 's propose a new security model that certainly outperforms the earlier schemes. It enhances the efficiency of encryption procedures for the datasets containing the repeating plaintexts as they included the complex update algorithm in their method. Nevertheless, all privacy problems are not wholly eliminated as the scheme needs the order information of all the encrypted plaintexts, which is a drawback concerning the security of the scheme.

Zhan et al. [43] have proposed a more advanced scheme known as the Multi-Dimensional Order-Preserving Encryption scheme (MDOPE) to address the queries, which are more complicated and involved with multidimensional data. Depending on the context, traditional OPE schemes have been mainly used to deal with primarily single-dimensional information; however, this new direction enables OPE to be much more effective in executing multi-dimensional range queries. The MDOPE scheme uses a network-based data structure to build the query index for every dimension in the variable space. This enhancement enables us to have a selective query while at the same time having the benefit of order-preserving encryption. More to it, in MDOPE the external MDS-attacker, says cloud servers cannot retrieve extra

information from the encrypted data other than the order of the data during the query. As earlier stated, MDOPE protects the security of the data under their made security definition management by preventing the leakage function that has been described by the authors, which is a critical aspect of multi-dimensional data handling since someone's identity, behavior medical status, or even political leaning, etc., are enough to bruise that person reputation or even subject him/her to severe punishment, stigma or discrimination.

2.3 Critical Review

This section presents a critical review of existing OPE Schemes and provides a detailed comparison of the performance and security of existing OPE schemes concerning the attributes of Deterministic, Probabilistic, Homomorphic properties, IND Standards, Computational hardness, and Complexity.

2.3.1 Probabilistic

The probabilistic feature of a cryptographic algorithm relates to the facet of the process whereby, with the encryption of the certain plaintext a numerous variety of ciphertexts is produced every time the key is used. This randomness is usually realized through the introduction of a key, constant, or component that is random during the actual encoding [44]. The first main advantage of probabilistic encryption schemes is that of semantic security while an opponent can decrypt the text encrypted with the aid of probabilistic encryption, he or she is unable to gain any information about the plaintext, even if he or she manages to compare more text encryptions executed from the same plaintext [45]. This unpredictability is important especially for security against cryptanalysis like a frequency analysis or a chosen plaintext attack where one tries to compare the plaintext to the encrypted text. Further, by producing different ciphertexts for the same plaintexts, probabilistic encryption schemes successfully randomize any possible recognizable patterns and make it much more difficult for the

adversaries to perform good analyses or draw conclusions [46]. The recipient of such a property increases the overall security of the encryption and thus probabilistic algorithms can be used to securely encrypt data in cases where the security of a data value must be utmost. The probabilistic column in Table 2.1 shows whether the scheme is probabilistic or not.

2.3.2 Deterministic

This property is most deterministic, meaning that every time the same plaintexts are encrypted using the same key, each input plaintext will generate identical encrypted data. This feature proves to be critical in its use such as database encryption since it helps in creating order in mapping hence increasing the efficiency in retrieval of data [47]. But at the same time, it generates security issues, as patterns in the ciphertext can produce various attacks, for instance, frequent analysis [48]. However, due to their deterministic flow, these algorithms can be quite useful for specific search operations and at the same time need a well-designed approach to minimize security flaws. The deterministic column in Table 2.1 shows whether the scheme is deterministic or not.

2.3.3 Homomorphic

In homomorphic encryption, the encrypted texts are characterized in a particular manner that allows performing addition or multiplication of texts as well as other operations between them and getting results, which, when decrypted, will be equivalent to results of analogous operations between the original plaintexts. This capability will allow the encrypted data to be directly processed without the ability to decrypt them, making the data secure all through the computation process. For instance, if two numbers are encrypted individually and then added mathematically in the encrypted form, it will be equal to the method when two numbers are added before the actual process of encryption. It is especially useful in secure data processing environments, namely cloud computing where while data processing, the data must remain

encrypted [49,50]. The column homomorphic in Table 2.1 shows that the OPE scheme provides the functionality of homomorphic operations on encrypted data.

2.3.4 IND Standards

IND in the context of OPE is slightly different because of the nature of OPE schemes which is why OPE-IND is even more challenging. The main objective of OPE is the ability to maintain the order of plaintexts in their encrypted counterparts such that one can freely operate such as a range query, on the encrypted data. However, this order preservation means that ciphertexts leak some information – the relative order of the plaintexts and it is in principle impossible to reach the traditional security models, such as IND-CPA (Indistinguishability under Chosen-Plaintext Attack) or IND-CCA (Indistinguishability under Chosen-Ciphertext Attack), where no relatively sound information about the plaintexts should be obtained at all from the ciphertexts. Consequently, OPE schemes do not satisfy the mentioned stringent security requirements strictly. Instead, the security of OPE is analyzed under looser definitions or conceptions that are chosen to consider the inherent order leakage. The best possible level of IND safety that aims at OPE is IND-OCOA (Indistinguishability under Ordered Chosen-Ciphertext Attack). This notion is of the form that the order of ciphertexts can be known, but no more information should be leaked even when an attacker has access to a decryption oracle. In the same fashion, there is another security standard that is slightly weaker than IND-CCA2, known as IND-OCOA (Indistinguishability under Ordered Chosen-Plaintext Attack) under which the attacker can choose plaintexts and get their respective ciphertexts but have not much idea of the content except the order of the content [48][52][38].

The column IND standard in Table 2.1 shows whether the mentioned schemes have been tested against IND standards or not.

2.3.5 Computational Hardness

In a post-quantum setup, most of the conventional cryptographic issues are insecure because a quantum computer can solve problems such as integer factorization and discrete logarithms easily. Nonetheless, some elements of cryptographic problems are still quantum attack proof, especially those that are based on the lattices. For instance, although the GACD, SVP, and LWE, are some of the problems believed to be secure against both classical and quantum computers. The GACD problem which seeks to determine a common divisor for a set of approximate integers is a tough nut that has presented a challenge to most computational models including Quantum that do not offer an efficient solution systematically. Similarly, lattice-based cryptographic schemes, for instance, those that rely on the Learning with Errors (LWE) problem or the Shortest Vector Problem (SVP) are thought to have very sound foundations since even with the form of quantum computing, the problems are hard to solve [51][53][54][38].

Also, as a consequence of the above discussion on IND standards and the computational hardness of OPE, three crucial information can be gained: Thus, when constructing an OPE scheme we have to base it on computationally hard problems along with satisfying IND security standards, the reason being that using computationally hard problems makes the attacks on such schemes substantially more difficult because the methods that are used in solving the underlining cryptographic problems are hard to implement. This approach enhances the security of the scheme under the consideration of order preservation so that the attacker cannot take advantage of potential vulnerabilities. Thus, although OPE inherently reveals order information, these drawbacks can be again partly offset by enhancing computational hardness assumptions and providing overall security. That is why to judge whether the schemes rely on computationally hard problems for their security or are merely relying on the notions of IND standards, one must consider the security over both parameters: the computational hardness of

the used problems. The column computational hardness shows whether the mentioned scheme is based on any computational hard problem or not.

2.4.6 Complexity

The amount of time that is necessary for an algorithm to finish its execution about the quantity of the input is referred to as the temporal complexity of the algorithm. To get an understanding of how efficient an algorithm is, it measures the amount of time it takes for the algorithm to carry out each operation and execute each line of code. It is common practice to represent time complexity using the Big O notation, which categorizes algorithms according to the amount of time they take to execute in the worst-case scenario or the average-case scenario, as a function of the size of the input. This enables a comparison of the effectiveness of various algorithms and contributes to a better understanding of how these algorithms will scale up as their inputs are increased.

When analyzing the performance of algorithms, time complexity is an essential component to take into consideration, particularly when dealing with large-scale data processing or computing activities [55][56][57]. The complexity column in Table 2.1 presents the time complexities of the scheme.

Table 2.1 shows a full comparison of different order-preserving encryption (OPE) methods, focusing on both their speed and security features. It checks how well each plan protects against important security issues like being vulnerable to various types of threats and rates how fast and how much it costs to use. This study helps us understand the costs and benefits of various OPE methods when it comes to security levels and operating performance.

Table 2.1 Comparison table to analyze different order-preserving schemes concerning security and performance

Seq.	Author	Probabilistic	Deterministic	Homomorphic	IND Standards			Computational Hardness	Complexity
					Tested	OCPA	OCCA		
1	Agrawal et al. 2004 [24]	×	✓	×	×	N/A	N/A	×	O(n)
2	Boldyreva et al. 2009[25]	×	✓	×	✓	✓	×	×	O(n)
3	Kerschbaum et al.2014 [30]	×	✓	×	✓	✓	×	×	O (n ² /log n)
4	Chenette et al. 2015 [32]	×	✓	×	✓	✓	×	×	O (n)
5	Wang et al.2010 [26]	×	✓	×	✓	✓	×	×	O (log n)
6	Popa et al.2013 [29]	×	✓	×	✓	✓	×	×	O (n log n)
7	Liu et al.2012 [27]	×	✓	✓	×	N/A	N/A	×	O(n)
8	Liu et al.2013 [28]	×	✓	✓	×	N/A	N/A	×	O(n ²)
9	Jayashri et al.2015[33]	×	✓	×	✓	✓	×	×	O (log n)
10	Krendelev et al.2014[31]	×	✓	×	×	N/A	N/A	×	O(n ³)
11	Khoury et al. 2018 [35]	×	✓	×	×	N/A	N/A	×	O (log n)
12	Yang et al. 2017[34]	×	✓	×	×	N/A	N/A	×	O (n)
13	Chen et al.2021 [40]	×	✓	×	×	N/A	N/A	×	O (log n)
14	Ahmed et al.2019 [39]	×	✓	×	×	N/A	N/A	×	O (n log n)
15	Quan et al.2018 [36]	×	✓	×	×	N/A	N/A	×	O (log n)
16	Shen et al.2021 [41]	✓	×	×	✓	✓	✓	×	O (n log n)
17	Yang et al.2021 [42]	×	✓	×	✓	✓	×	×	O (log n)
18	Dyer et al.2019 [38]	✓	×	×	✓	✓	✓	✓	O (1)
19	Reddy et al.2019 [37]	×	✓	×	✓	✓	×	×	O (n)
20	Zhan et al.2022 [43]	×	✓	×	×	N/A	N/A	×	O (1)

2.4 Summary

However, it is worthy of note that user anonymity will be a feature in future OPE schemes in order not to expose the user’s identity to other people. Cryptographic algorithms should ensure they use probabilistic encryption so that the actual values of some numbers are not disclosed while computing the sum. Homomorphic Support. They should have homomorphic support which allows computations

on encrypted data to be performed whilst it remains secure. Cryptographic hardness with the current and emerging forms of attacks, the schemes should be cryptographically hard to compromise. Optimized Complexity schemes should be efficient in both security and privacy and the complexity that is required to deploy in different environments. Concisely, the design recommendation includes security and flexibility where the user is anonymous, using the higher level of the encryption method, homomorphic, and cryptographic resistance, and optimizing the complexity.

Proposed Methodology

3.1 Overview

This chapter brings out the approach that was followed when trying to solve the challenges that arise with Order-Prescribing Encryption (OPE). The initial step of the research methodology entails conducting a critical analysis of the current OPE schemes with preference drawn to their shortcomings, most of which do not support floating point data. To overcome these limitations, the proposed methodology used the extension of Dyer et al's General Approximate Common Divisor (GACD) based scheme [38]. This scheme was chosen because of the robust security genesis that performs single operations of $O(1)$ for encryption and decryption. The system model is aimed at including integer and floating-point data which increases the effectiveness of encryption and decryption processes and respects the order of the data. The threat model reveals the risks of security drawbacks as well as the protection measures to be applied, especially when both forms of data are incorporated. The problem analysis section elaborates on the impact of the identified issues on achieving the goal of OPE and offers a rationale for the suggested approach to handling multiple data types. Indeed, this chapter is devoted to the formulation of the hybrid OPE scheme based on Dyer et al 's GACD-based approach and some other methods to address floating-point data. This paper analyzes the performance and security of the proposed scheme in detail and analyzes it in the literature section. This methodology will improve the adaptability of using OPE to the problems of different fields, and through meeting

the development requirements of the first data, serve the purpose of improving secure and efficient encrypted data processing systems.

3.2 System Model

Order-preserving encryption (OPE) is especially beneficial in cloud scenarios where data must be encrypted for security but querying and sorting must be done efficiently. In the cloud environment, it allows users to perform a range of queries and ordered searches on the encrypted data without decrypting it, making it secure and fast. OPE has been made to meet the challenge of ordering the data in the encrypted form which in turn enhances data operations on encrypted databases. Figure 3.1 shows the flow of our scenario.

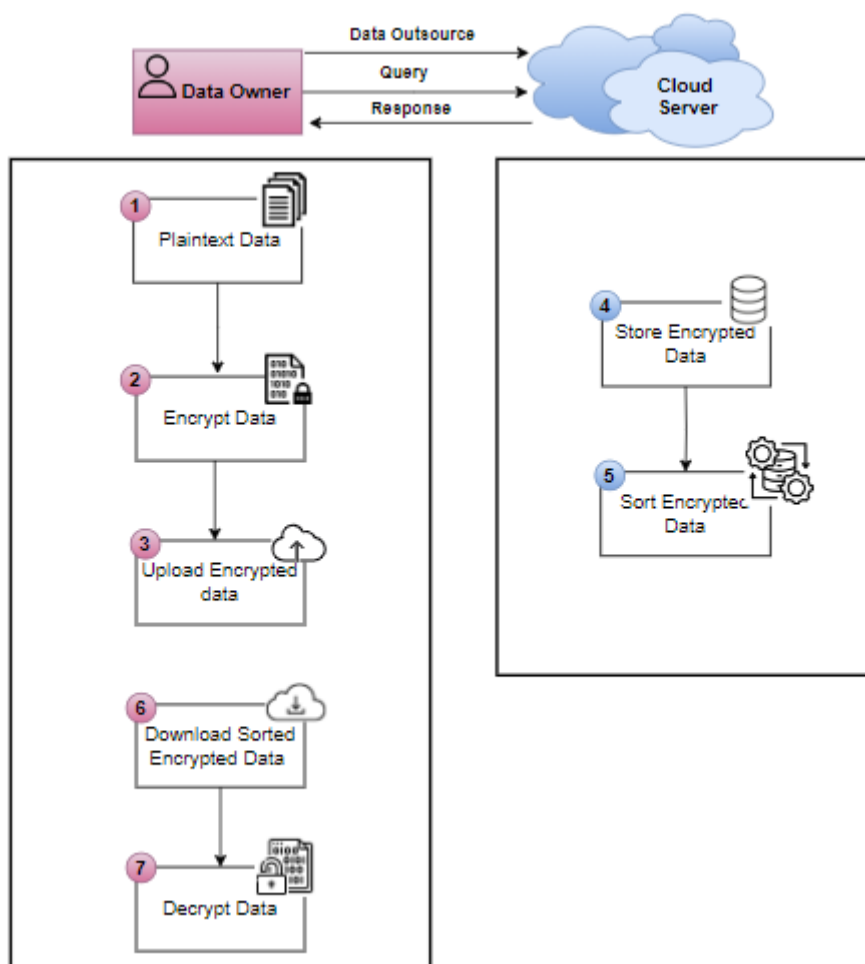


Figure 3.1 System Model Architecture for Order-Preserving Encryption (OPE) in Cloud Environments

However, OPE can also be utilized in an on-premises manner for the same purposes, security, and functionality of on-premises storage and computing. Within the scope of this undertaking, the system model is made up of the following entities: a data owner and a cloud server.

It is the responsibility of the data owner to build the Order-Preserving Encryption (OPE) scheme and encipher the data set before sending the plaintext data into the cloud environment. After encrypting a plaintext message, the data homeowner uploads the encrypted data through the cloud network and stores it in the cloud server. When the data homeowner has to 'search' or do a range 'search' for a particular data, he builds an encrypted query using the OPE scheme and transmits it to the cloud server. For the specific query on the cloud server side, the operations are executed over the encrypted data to make maximum use of the order-preserving characteristics of the encryption to sort and carry out any number of necessary operations without decryption. The query response generated by the cloud server is encrypted and it is returned to the client side in such an encrypted format. On receiving the above-mentioned encrypted response, the data owner or the client-side application deciphers the response to get the original plain data. This approach guarantees that the data will not only be kept secure on the cloud but passed through the needed transformations securely and at the same time be easily searchable. Since OPE is optimal for cloud computing to provide safe and effective means of data handling, for the same results it can be applied to the on-premises infrastructure [75].

3.3 Threat Model

Assumption: In our threat model, the cloud is semi-trusted meaning that cloud service providers will be expected to adopt best security practices and enforce strict access controls however, total security of the cloud environment cannot be fully relied on by the service provider. Insider threats are ruled out in this model since we can trust that the cloud provider is not predisposed to hacking into his infrastructure and follows defined norms of security.

- **Adversary Capabilities:** The threat model describes the potential capabilities of the adversary and the kind of attacks the encryption scheme must be capable of resisting. The adversary is thought to know all the data encrypted and can work with the ciphertexts to gain information about the plaintext quantities. They know that the ordering of plaintexts is reflected by the encryption scheme, and so if one ciphertext is less than another, the corresponding plaintexts are also less by the transformation. The adversary might have some information about plaintext and ciphertexts (called known-plaintext attack) or even might be able to choose plain texts and see their encrypted versions (called chosen-plaintext attack).
- **Attack Vector:** Specific attacks include the frequency analysis attack which the adversary tries to map the frequency distribution of ciphertexts to possible plaintext values, and the order inference attack which takes advantage of the order-preserving nature of the encryption to estimate plaintext values through the ordering of ciphertexts. The adversary may also use range queries to make unauthorized searches and maybe leak relations between plaintexts. Moreover, an attack on the implementation of the specific encryption scheme can be made to reveal vulnerabilities.
- **Security Goal:** The security objectives of OPE schemes are to present confidentiality of data, that is, it is quite hard to deduce something big about plaintexts from ciphertexts other than what is given by the order-preserving nature. It should also be safe against both the known-plaintext attacks as well as the chosen-plaintext attacks to minimize information leakage.

3.4 Research Methodology

In this work, an attempt is made to generalize the existing OPE scheme to encrypt both the integer as well as the floating-point data. The following are the steps that need to be done for the given data to be encrypted for security while preserving the order. Hence, after converting the data type, they are used concurrently with the enhanced OPE scheme that will be discussed

later because the current OPE schemes were mostly designed to work with integer data and they have some critical drawbacks. Therefore, this work aims to design an encryption system to enable secure additional data processing accompanied by their sorting. However, this form of the system should, due to the pre-processing of the data and the integration of the improved OPE scheme into the query, be expected to refine the query processing and the encrypted data's accuracy. This research entails a case study and literature review analysis to deduce the deficits and voids of the current methodologies. Such reviews are informative to the enhancement of the models for encryption and the processing systems, to optimize the encryption initiatives, and to increase the effectiveness of the encryption procedures. The next paragraphs contain a general description of all the stages in the procedure and a detailed description of the analysis.

3.5 Problem Analysis

The current OPE schemes specifically speaking are bespoke for integer data completely neglecting the requirement of floating-point data. This does not go well in areas that require very sensitive arithmetic operations such as finance and healthcare where even the decimal point is sometimes important such as when calculating money or measuring a person's temperature. The absence of endorsement for both data types hinders OPE schemes' flexibility and efficiency in achieving tasks and hampers their prospects across different domains. To overcome the above-mentioned problem, this research work suggests enhancing the Dyer et al. [38] GACD-based scheme to design a new OPE technique named Hybrid OPE where two types of dataset integer and floating-point are effectively implemented. The Dyer et al. scheme has been adopted because it has a secure background, and its general concept is rather simple comprising only $O(1)$ operations for encrypting and decrypting the data which also speaks of the algorithm's high performance. Some other OPE schemes can be compared in detail with their advantages and limitations described in detail in the literature section. That is why this extension is to improve the usability of OPE schemes that can become unchanged for

accommodating the various requirements of distinct sectors, in which the stability of data and their organization is important.

3.6 Proposed Methodology

The proposed methodology introduces a novel hybrid Order-Preserving Encryption (OPE) scheme planned to address the limitations of existing order-preserving encryption schemes that predominately handle integer data. This Hybrid approach is important because it expands OPE's range of applications and opens it up for float data which has obvious benefits, especially in such fields as finance and healthcare, where both kinds of data are usually encountered. The methodology provided in this paper presents an extensive approach for ordering and maintaining the order of both integer and float data using sophisticated enshrinement techniques. Thus, improving upon the poor outcomes of prior OPE schemes that offered limited data security and query/ processing functionality, this approach empowers global data utilization for multiple applications. With this advancement, it is projected that the application of OPE will be enhanced in real-life situations where reliability, accuracy, and security of data are desired.

3.6.1 Design of the Hybrid OPE Scheme using Approximate Common Divisors

Our Order-Preserving Encryption (OPE) scheme is a symmetric encryption system composed of five components: Key Generation (KGen), Encryption (Enc), Decryption (Dec), normalization (Norm), and de-normalization (De-norm). Figure 3.2 presents the detailed architecture and flow of our proposed hybrid general approximate common divisor problem-based OPE scheme.

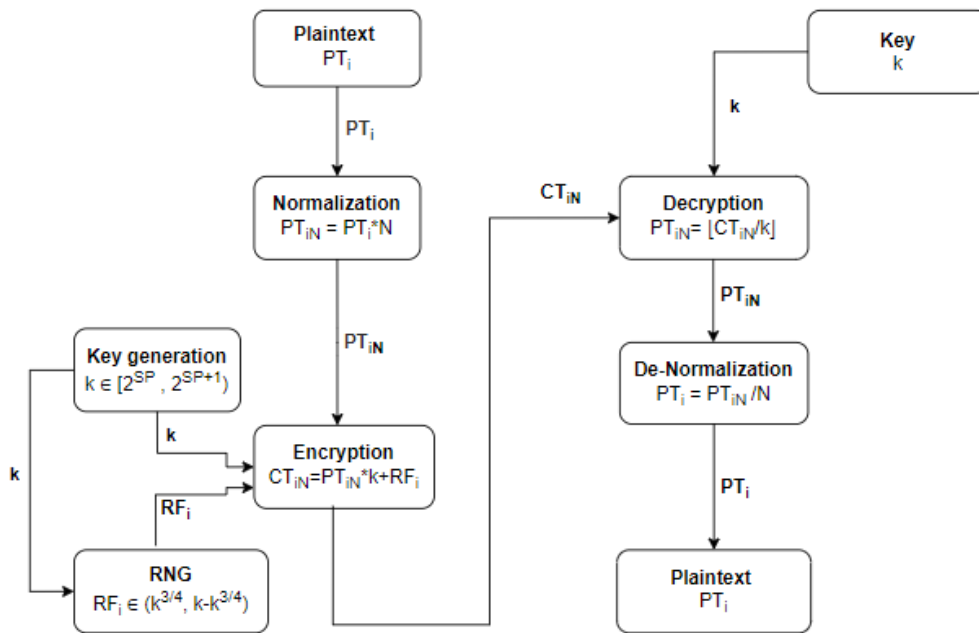


Figure 3.2 Architecture of the Proposed Hybrid OPE Scheme

In the past, the OPE has generally consisted of three primary kinds of operations: These are important processes involved in encryption, the generation of the key, and decryption. But for entering Integers as well as Floating point numbers initially we have included the phase of normalization and denormalization.

3.6.2 Algorithm Development

Our proposed scheme is simplest at the conceptual level and probabilistic. The plaintext space PT_space is $[0, PT_space]$, and the ciphertext space CT_space is $[0, CT_space]$, where $CT_space > PT_space$. We assume a set of plaintexts $PT_i \in PT_space, i \in [1, n]$, such that $0 < PT_1 \leq PT_2 \leq \dots \leq PT_n \leq PT_space$.

The proposed algorithm is simplest at the conceptual level and probabilistic. To encrypt, we first normalize the plaintext PT_i by multiplying it by a large integer $N \in 10^n$. Next, we scale the plaintext PT_{iN} by a large integer k . We add a random factor RF_i , specific to the individual ciphertext, where $RF_i < k$. To decrypt, we must divide the Ciphertext by forgetting the remainder and then de-normalizing the quotient. To check that the ciphertext is maintaining the order, suppose we have two plaintext messages PT_1 and PT_2 such that $PT_1 < PT_2$, PT_1 is

encrypted using $CT_{1N} = k*PT_{1N} + RF_1$, and PT_2 is encrypted using $CT_{2N} = k*PT_{2N} + RF_2$; To maintain the ordering, we need $CT_{2N} - CT_{1N} > 0$, This follows because, This follows because, in the left side of the inequality, it is at least k while in the right, it is at most $k-1$ (it is $k-k_{3/4}$ which is less than k). Recall that $PT_1=PT_2$, i.e. We are encrypting a plaintext twice, as for the order of the two encryptions it is random because $\Pr (RF_2>RF_1) \approx 1/2 - 1/k \approx 1/2$ as, $k \gg 1$. Every component of the proposed framework is explained below sequentially.

3.6.2.1 Key Generation

Both the security parameter space SP_space and the key space (key_space) belong to the set of positive integers. Given a security parameter $SP \in SP_space$, with $SP > 8/3 \log PT_space$. The choice of an integer in Algorithm 1 is randomly made, $k \in [2^{SP}, 2^{SP+1})$ as a secret key. So, k is a $(SP+1)$ -bit integer such that $k > PT_space^{8/3}$. Also, k is not required to be axiomatically prime.

Algorithm 1: Key Generation Algorithm KGen.

- a) Input:
 - A security parameter SP such that $SP \in SP_space$ and $SP > 8/3 \lg (PT_space)$.
 - A parameter PT_space .
- b) Initialization:
 - Compute (2^{SP}) , store it in k_{min} , minimum allowed key value.
 - Compute $2^{(SP+1)}$, store it in k_{max} , maximum allowed value of the key.
- c) Generate key:
 - Randomly generate a key k such that $k_{min} \leq k \leq k_{max}$ and $k > PT_space^{(8/3)}$;
- d) Output:
 - An key k , for given security parameters.

3.6.2.2 Encryption with Normalization

To manage float point data, we introduce the normalization step. Algorithm 2 firstly normalizes the message PT_i by multiplying it by a positive integer $N \in 10^n$. So that each message changes into normalized msg PT_{iN} which cannot be the float value. A normalized plaintext message PT_{iN} , is encrypted using Algorithm 3 and ciphertext CT_{iN} is generated for each normalized message which store in array CTN.

Algorithm 2: Algorithm for Norm and Enc.

a) Input:

- An array of plaintext PT.
- A pre-generated key k where $k \in \text{key_space}$.
- A normalization factor N , where $N \in 10^n$.

b) Initialization:

- Compute $RF_{\min} \leftarrow k^{(3/4)}$, the minimum allowed value for random factor RF.
- Compute $RF_{\max} \leftarrow k - k^{(3/4)}$, the maximum allowed value for random factor RF.
- Initialize an array CTN for generating ciphertext.
- Compute length of array PT, store in m .

c) Generate ciphertext:

- For $1 \leq i \leq m$:
 - Randomly generate RF_i where $RF_{\min} \leq RF_i \leq RF_{\max}$.
 - Compute the normalized value PT_{iN} such that $PT_{iN} = PT_i * N$;
 - Compute CT_{iN} such that $CT_{iN} \leftarrow (PT_{iN} * k) + RF_i$;
 - Store CT_{iN} in array CTN.

d) Output:

- An array of CTN of generated ciphertext.

3.6.2.3 Decryption with De-normalization

A decrypted ciphertext DT_{iN} is decrypted by Algorithm 3. Decrypted data is de-normalized back to the original using Algorithm 3 after decryption has been completed and the original decrypted text array DT is recovered. That is the final stage of our proposed scheme.

Algorithm 3: Algorithm for Dec and De-norm.

a) Input:

- An array CTN of Cipher text.
- A key k for given ciphertext array.
- A de-normalization factor N , where $N \in 10^n$.

b) Initialization:

- Initiate an array DT for storing decrypted ciphertext.
- Compute length of array CTN, store it in d .

c) Generate Decrypted text:

- For $1 \leq i \leq d$:
 - Compute the decrypted value DT_{iN} for this ciphertext, such that
$$DT_{iN} = \text{int}(CT_{iN} / k);$$
 - Compute the de-normalized value DT_i for decrypted value such that DT_{iN} / N ;
 - Store DT_i in array DT.

d) Output:

- An array of DT of decrypted text.

3.7 Dataset Description

To compare our presented hybrid, OPE scheme with the previous scheme, two different datasets of the same bit length were applied. First, to match the previous work by Dyer et al., we created a dataset with 100 random integer values for each bit length: It is 7, 15, 31, 63, and 127. These values were then used to examine the existing OPE scheme as was done in the main study. This enables a comparison of the effectiveness of the two schemes where they are both subjected to similar conditions.

Secondly, we generated another dataset for the proposed hybrid OPE scheme. This set also contains 100 random values for each bit length (7, 15, 31, 63, and 127), but this variant consists of integer and floating-point numbers more exactly, float numbers. The float values in the dataset are up to four decimal places. This precision of 4 decimal points for the input values is based on the normalization factor that will be later used in the encryption. For the current parameters of the proposed scheme, this precision was four. This arrangement helps in comparing the two schemes fairly by applying the bit lengths of the two sets of data. Also, the source dataset is assumed to be static, just like in the previous research. This prevents the attacker from submitting plaintexts to the data owner. It is noteworthy that such datasets are to replicate real-life conditions. For instance, while using finance, most of the transactions require a specific decimal value to be involved, while in using health, most of the measurements involved in the handling of patients may require floating-point numbers. This approach also addresses national requirements on how data should be secured to prevent leakage and improve the level of security. The hybrid OPE scheme that we have developed is quite general and can find application in many fields such as telecommunication, scientific research, and so on.

3.8 Security Analysis

Our contribution is in the proposal of a new Hybrid OPE scheme, for which, contrary to prior approaches that use security games, its security relies on computational hardness rather than security games, specifically the General Approximate Common Divisor (GACD) problem challenge [34].

3.8.1 Security of the Scheme

The security of our proposed encryption scheme is based on the General Approximate Common Divisor Problem (GACDP) which is acknowledged as a computation complexity problem. In the case of GACDP, the question deals with finding an unknown integer. The fact that solving GACDP is pivotal to the security of a countless number of cryptosystems, it has been shown that GACDP reduction is as computationally intensive as solving an LWE, which forms the basis of several post-quantum cryptosystems such as LIMA and Lizard [58,59]. Regarding the framework of our scheme, the critical aspect that must be addressed is the common divisor. The primary constituent that is difficult to find is essential. Howgrave-Graham [60] described two attacks directed at GACDP, which relied on the search of divisors under certain conditions. However, relevant potential vulnerabilities are eliminated in our system since we carefully choose the parameters, and thus the common divisor k remains elusive. This is accomplished by setting the offsets and guarantees that ' k ' is sized appropriately which makes the system immune to advanced attacks.

Also, Cohn and Heninger [61] generalized Howgrave-Graham's technique to multiple integers; however, the worst case of the described algorithm has exponential time consumption. Chen and Nguyen [62] noted that this algorithm may take longer than brute force methods sometimes, although the time complexity of the algorithm under consideration is low. In our scheme, we use large offsets, and, thus, such kinds of attacks are less important, and the security of the scheme is enhanced even more. In conclusion, the proposed scheme provides sufficient

security based on the hardness of the GACDP for both the classical and post-quantum cryptanalysis, in terms of the state-of-the-art assessment of Galbraith et al.'s [63].

3.8.1.1 Security models

The security of our scheme falls under the traditional security models of encryption concerning the security results obtained within the bounds of OPE schemes that are limiting inherent ways. As usual, OPE schemes cannot be IND-CPA secure because of the ordering of the ciphertexts that provide some relational information between the plaintexts. This characteristic makes the IND-CPA irrelevant to the use of OPE in practice as pointed out by other authors [64,65,66,67] such as Boldyreva et al., who have developed different indistinguishability assumptions that put impractical constraints on an adversary. In this regard, we should consider Indistinguishability under ordered CPA (IND-OCPA) by Boldyreva et al. which enhances the OPE security model's agenda by identifying the way the ordered nature of cipher texts affects security goals namely confidentiality and integrity. Thus, our hybrid OPE scheme is secure when IND-OCPA is used as the security model. It is good enough in preventing the leakage of information other than the order of the plaintexts and at the same time concedes that given the nature of OPE, it is impossible to eliminate disclosure of information. This model is realistic and realistic enough to be implemented in OPE schemes & Order preservation is necessary for quite a lot of applications and provides reasonable security.

Thus, we assess the security of our scheme under the one-wayness of windows that is provided by Boldyreva et al. In our opinion, this is a much more appropriate and realistic model for our purposes. It can simply be mentioned that achieving some indistinguishability, for example, IND-OCPA, does not necessarily mean that the cryptosystem to be constructed is secure, and vice versa, its failure to satisfy some of the requirements mentioned above does not, in any way, mean that the system under construction is defeatable. We believe our scheme is secure within the parameters indicated by the window one-wayness model while recognizing that OPE

schemes can't be completely indistinguishable from real random numbers and that this is not essential in most installations given that truly random numbers are not generally available, even where they are employed for key generation the random will normally be pseudorandom.

3.8.1.2 One-wayness

The security of the proposed scheme is based on the one-way property of the function to be used which is defined as $CT(PT)=kPT+RF$ and is related to the General Approximate Common Divisor (GACD) problem. A one-way function is judged by how easy it is to compute in one direction and how difficult it is to reverse precluding knowledge in the specific case of recovering PT from $CT(PT)$ is practically impossible to solve without the secret key k . Furthermore, it is crucial to note that the hardness of the GACD problem which forms the basis of this scheme likewise adds to this one-wayness. The problem used in the GACD is known for its resistance to polynomial time attacks; therefore, it is quite reliable for constructing secure encryption. As such, our scheme's security model continues to adhere to a one-way function, such that while encryption is always feasible, decryption using anything less than the right key is impossible, thus guarding the encrypted data from the predictors [64]. For survey and evaluation see Dyer et al. [38].

3.8.2 Enhancement of Security with Floating-Point Data

Adding floating-point data into Order-Preserving Encryption (OPE) can be highly beneficial when it comes to enhancing security because of the shortcomings in integer-only schemes. In addition to solving the problem of generalization of OPE in practical conditions, this advancement enhances the application's protection against different types of attacks.

- **Precision Preservation and Reduced Leakage:** Precision Preservation and Reduced Leakage: Floating-point data gives a finer level of detail as compared to integer data as it minutely preserves precision and does not leak information. Specifically, native OPE schemes,

which are originally developed for integers, have the problem of precision loss, as floating-point numbers get rounded or truncated. The ability to use floating-point data reduces the risk of losing confidential information and increases the effectiveness of the encryption/decryption processes. Some recent works have stressed the importance of pushing the encryption schemes to be able to handle floating-point data better to counter precision-based attacks since the encoded data must look as close to the original data as possible [68][69].

- **Obfuscation of Data Patterns:** The conversion to floating-point representations inevitably comes with variability and complexity which in turn hides patterns in the cipher data. This added feature decreases the probability of any entity identifying statistical patterns or regularly occurring patterns. For example, possible patterns of incorporating floating-point data can lower the impact of attacks based on averaging, counting, and distribution of encrypted quantities. These schemes incorporate a layer of obfuscation using floating point numbers; thus, these schemes are even more secure against such attacks [70][71].

- **Versatility in Mixed Data Environments:** Versatility in Mixed Data Environments: Most of the practical scenarios require data that are in integer and floating-point format as records from the financial sector and scientific investigations. Thus, the expansion of opportunities for OPE schemes for both data types enhances the flexibility of its use and protection. This approach of treating mixed data types uniformly results in better security of the encryption as a whole and dependably maintains the encryption's security across various datasets. This adaptation also avoids weaknesses that are associated with the use of specific schemes that can perform efficiently in only one type of data [72][73].

3.8.3 Security Benefits of the Hybrid GACD-Based OPE Scheme Against Various Attack Types

- **Frequency and Statistical Attacks:** With the introduction of floating-point data, there is a new level of chaos and hence the attackers will not be able to use frequencies or statistical

methods. As encrypted data is not random and does not follow a recognizable pattern, these schemes are very effective in protecting against attacks that take advantage of the text's statistical characteristics [75].

- **Order Preservation Vulnerabilities:** Actually, using floating-point data can help minimize order preservation vulnerabilities. Due to the limitations of traditional integer-based OPE schemes, the processes may reveal the order of the plaintexts therefore being more vulnerable to attacks. However, considering the floating-point data would complicate the order-related information flows, which would make it difficult for the attackers to derive or hack their way into the detailed information [76].

3.9 Summary

The chapter gives a clear and concise background on the research approach, system architecture, threat assessment, issue identification, the solution proposed, the dataset used, and the concerns of security in the present investigation. The Research Methodology states the study procedures that will be followed in the research process: a critical analysis of the current OPE schemes that have been employed; and the design of a new hybrid OPE scheme, which can be used to process integer and floating-point data. The System Model specifies the structure and parts of the encryption system on which the evaluation will be conducted. The Threat Model points out the risks and weaknesses of a system. The Problem Analysis outlines current OPE strategies' shortcomings, stressing their inability to handle floating point data, and the need for cooperation to extend their versatility. The Proposed Scheme builds upon Dyer et al. The GACD-based scheme in which it inherits the strong computational base of hardness while it offers optimum security. The Dataset Description details the use of two datasets: one for the existing scheme consisting of integer values and the other one having a mixed set of integers and floating points for the proposed scheme. Last, the Security Aspects explain how some of the challenges do not apply to the proposed scheme since rather than a security game, it builds

upon a computational hard problem, while it enhances the solutions presented in the literature. Thus, this chapter which provides detailed information and analysis, seeks to further the cause of applying OPE in settings that entail issues of data security and order.

Analysis and Results

4.1 Overview

This chapter provides a comprehensive comparison of the experimental results of the proposed hybrid OPE against the prior GACD scheme by Dyer et al, the section first states the datasets used in this study are integer only for the existing scheme while the proposed scheme used integers and fixed-point numbers. The chapter then focuses on providing the results of encryption, decryption, and range queries concerning different bits. To compare the performance, plots are given for our scheme and existing GACD, proving that the proposed scheme has constant output in terms of efficiency and higher bit lengths while the GACD has fluctuating results. Chapter six provides a summary of the findings based on the experimented results and underlines our schemes' strengths for both different formats of data and performance stability.

4.2 Experimental Results

In the case of our experimental evaluation, we developed two unique datasets to ascertain the competence of Dyer et al. 's GACD scheme as well as our hybrid GACD scheme. The first dataset was generated for this study to test the previous GACD scheme which is restricted to operate on only positive integers. This dataset is composed only of integer values, which helps estimate the scheme within the scope of its expected use only. On the other hand, the second set of data was specifically created for the improved GACD

scheme in this paper which compliments the training of integers and floating-point numbers. Another significance of this dataset is that, unlike the first one, this one comprises both integer and floating-point type numbers; this helps to establish the fact that our scheme is even more flexible than previously thought. Specifications of these datasets, the bit lengths used and the way that normalization for a floating-point set is made have been presented in proposed methodology, the proposed methodology. It will therefore be possible to achieve the purpose of showcasing the applicability and efficiency of the proposed hybrid scheme as compared to the conventional GACD strategy with the help of these datasets.

To assess our schemes in depth, the current average time for selected cryptographic computations was compared. We collected the average time to encrypt a single message as well as decrypt a single message and the time to encrypt and decrypt a single message collectively. Also, determine the meantime taken in seconds to perform a single range query on the datasets.

This analysis was carried out over five distinct sets of bit values: 7, 15, 31, 63, and 127, corresponding to the bit lengths used in the previous works that follow this one. Therefore, the bit lengths were chosen to be specific to the designated parameters so that the comparisons made in this work are consistent with other works, which would enable us to quantify the enhancements made to the speed and real-world applicability of the proposed hybrid GACD scheme. Table 4.1 represents the findings of the experimental results and also indicates the timing of each experimental configuration carried out where $n=100$ is the number of messages used for each bit length. The Variable ρ denotes the number of bits of the input data and the value of ρ is (7, 15, 31, 63, 127).

Table 4.1 Performance analysis of the proposed Hybrid GACD-based OPE and existing GACD-based OPE scheme

Scheme	ρ	Encryption(μ s)	Decryption(μ s)	Total exec. time(μ s)	Range query(μ s)
Hybrid GACD-based OPE	7	2.4920	0.5091	3.0011	45.400
	15	2.9799	0.5229	3.5028	47.339
	31	3.4529	0.5659	4.0188	47.984
	63	3.9330	0.6090	4.5432	48.904
	127	4.7548	0.6777	5.4325	55.364
Dyer's GACD-based OPE [38]	7	2.3477	0.4156	2.7633	42.382
	15	2.4252	0.4595	2.8847	42.946
	31	2.6099	0.5081	3.1180	43.983
	63	2.9123	0.5307	3.4430	47.241
	127	3.9123	0.5560	4.4683	51.961

Encryption is the mean time taken to encrypt a single message while decryption is the mean time taken to decrypt a message. To derive the “Total Execution Time” it is the sum of the mean time taken for encryption and decryption of each message. Further, “Range Query Time” gives the mean time required for a single range query over the encrypted data. As seen in the table, all time measurements are in microseconds giving a much deeper insight into the performance of each of the operations. The mean encryption time for a single message in our proposed scheme covers the generation of the key once, the generation of RF for the reach message as well as the encryption and normalization. Thus, the reported mean time for encryption incorporates time for the additional (key generation, random factor generation, normalization) processes as well. The same is correct for the decryption function in our scheme as the de-normalizing step is also included in the mean time for the decryption operation. For the existing scheme, the meantime to encrypt one message is the time includes key generation phase. The mean time to query includes the time taken to encrypt a query and for searching.

Additional detailed graphical expiation of the results is described in the next sections which will give a better understanding of the performance between the existing and proposed OPE schemes. These diagrams emphasize the differences of time spent on encryption and decryption

of messages, as well as on the other operations, depending on the bit length of the messages which helps to compare and decide which scheme works best and how it is effective under the given circumstances. It will also ease the comparison and contrasting process between the two methods in terms of the efficiency and performance of the intervention.

4.2.1 Comparison of Encryption Time

This part figured out the actual comparison of the hybrid GACD scheme (our scheme) and the integral GACD scheme (Dyer et al. 's scheme). The average time taken for encryption of a single message comparison is done and presented in a line graph as shown below in Figure 4.

1. Every point that flies on the graph is the time taken to encrypt the message and the length of the message bits used.

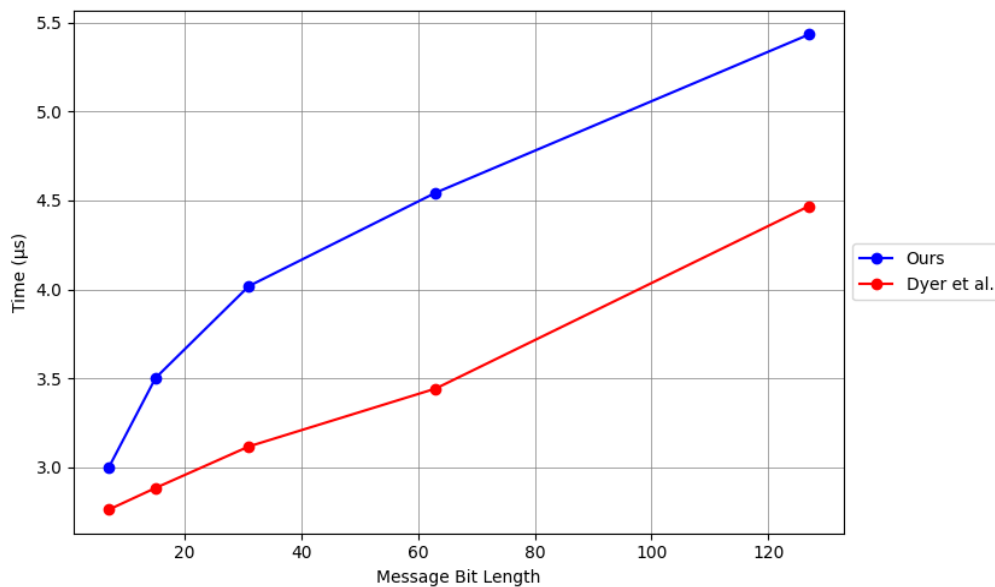


Figure 4.1 Average encryption times.

As noted earlier, the [38] GACD-based scheme and our scheme use the same primary encryption process. However, the present scheme adds normalization and denormalization steps in addition to the original key steps. These are additional operations to facilitate the processing of the data type integer as well as data type float to augment the message space while at the same time extending the security parameter space and key size. Therefore, the

present scheme consumes much time in calculating the encryption as compared to the existing scheme which has the facility of operating only one data type. This extra complication originates from the inclusion of a wider array of data varieties and the increased parameter area, which strengthens the scheme's flexibility as well as the security it offers; nonetheless, at the expense of more time used in computations.

4.2.2 Comparison of Decryption Time

In this section, we have explained the comparison of the hybrid GACD-based proposed scheme and the existing GACD-based OPE scheme.

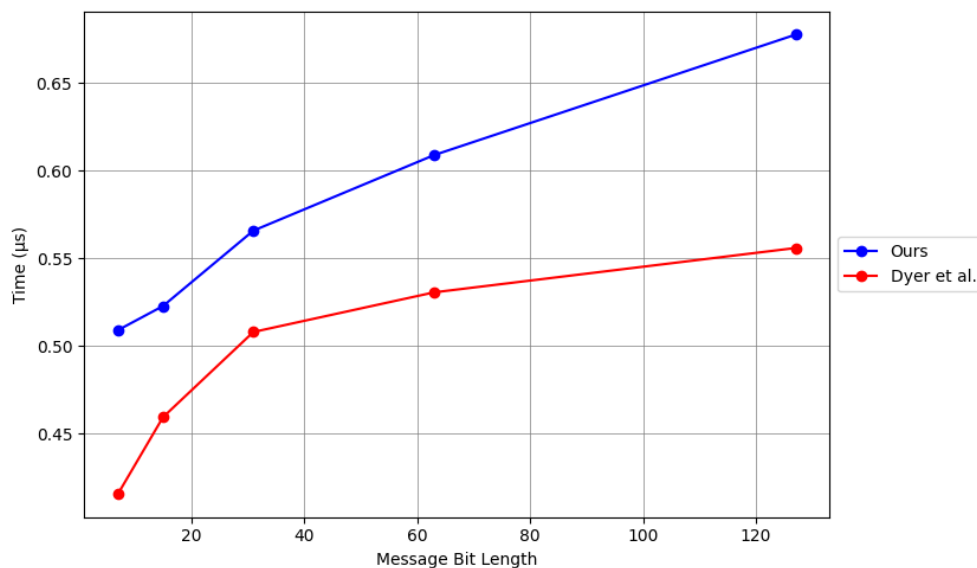


Figure 4.2 Average decryption times.

This time is calculated and depicted in the line graph below referred to as Figure 4.2, where the average time taken for the decryption of a single message is determined. Every dot on the graph represents the time it took to decipher a message relative to the message's bit length.

For decryption, our scheme requires two operations: first, to decode the encoded messages and, second, to de-normalize the decoded results. Unlike the existing OPE algorithm, this research only performs decryption. Thus, the Executes of our scheme consume more time to accomplish a decryption because of the denormalization step. However, time also rises slowly in our

scheme as well as in the existing algorithm with the growth of the bit length that describes the calculations required for the larger bit length.

4.2.3 Comparison of Range Query Time

An important functional requirement of order-preserving encryption is range queries. Here in this section, Figure 4.4 depicts the average time it took to perform a single range query of the messages of different bit lengths in the case of both schemes.

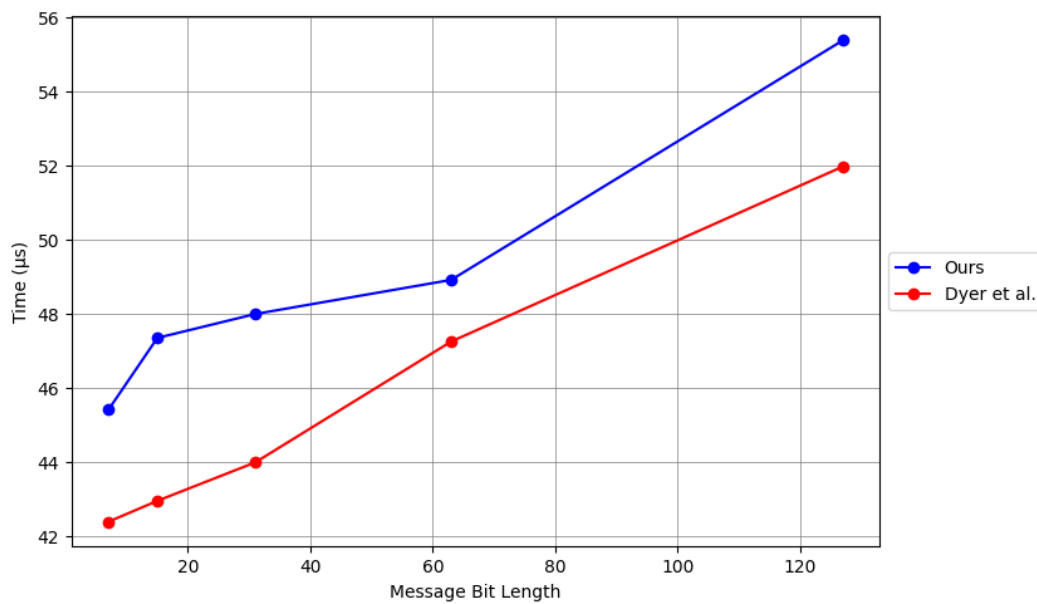


Figure 4.3 Average range query times.

The graph shows the time each of the schemes takes to execute the range query in microseconds when faced with messages of various sizes. The graph varies the total time of range query on average for a single operation depending on the bit length. In both algorithms, the actual search is performed not on the original space but on the space of ciphertexts. For smaller bit lengths (7, 15, 31), the time taken to encrypt and search too rises steadily. Still, the time needing to be spent increases more as the bit length rises from 63 to 127. This is because the bit length difference is large, so there is much more searching on the larger bit lengths. Also, floating-point data employed in our scheme makes it more computationally intensive when compared with integer-only data. Hence, as before, for floating-point data ciphertexts are constructed

complex, a priori significantly increasing in time depending on the length in bits. The current algorithm also has reduced response time with increased bit length, which indicates this trend.

4.3 Ciphertext Expansion Factor

The Expansion Factor is defined as the ratio of the bit length of ciphertext, to the bit length of ciphertext. It measures the increase in size of encrypted data, as compared to the original message data. It is a critical evaluation metric used for measuring the efficiency of an encryption algorithm. Table 4.2 shows the ciphertext expansion factor for our proposed hybrid GACD-based OPE algorithm for given bit lengths [76].

Mathematical Formula:

$$\text{CT Expansion Factor} = \frac{\text{ciphertext bit length}}{\text{original message bit length}}$$

Table 4.2 Ciphertext expansion factor for given bit lengths

Message Bit length	Expansion Factor	Ciphertext Bit length
7	1.4	11
15	1.53	23
31	2	62
63	3.6	226
127	5.2	661

Table 4.2 shows the trend of expansion factor in the proposed scheme for bit lengths of 7,15,31,63 and 127. The expansion factor increases from 1.4 to 6.67 as the bit length increases from 7 to 127. This shows that the encryption scheme shows optimum performance till the message bit size of 31. But as message bit size increases further, the expansion factor increases drastically, and hence, the ciphertext bit length increases exponentially. The factors that contribute to the increase in expansion factor at higher bit lengths for the proposed scheme are described below.

- **Key Generation:** As the bit length of the message list increases, the parameter lambda also increases. This parameter is further used in key generation, increasing key size, hence affecting the encryption process.
- **Encryption logic:** The encryption logic affects the ciphertext size. The proposed logic multiplies the message with a key and further adds a nonce value. The key size as explained above, increases with bit length, thereby increasing the nonce value which depends on the key. Since both the variables used in encryption i.e. key size and nonce, increase with bit length, this contributes to a significant increase in ciphertext length.

4.4 Evaluation and Analysis

In our work, we have developed our Order-Preserving Encryption (OPE) that supports both integer and floating-point data types, making it possible to protect databases with high variability in data formats. To deal with issues of floating-point data our scheme involves normalization just before the encryption process and denormalization of the data after the decryption process. Nevertheless, these steps add further computations and thus, the overall execution time is higher, particularly with a rise in the bit length of parameters. Further, there is a grave ciphertext expansiveness of our scheme where the size of the ciphertext increases with the larger bit length which is due to the complexity of float type data. Nonetheless, the following are the problems that we have encountered in our work: Still, the security of our scheme is based on the computational hardness of some mathematical problem that increases the cipher space and makes it more secure from different types of attacks. This makes our scheme suitable for environments where security is of paramount importance and when multiple types of data are to be stored and encrypted despite the performance overhead and the ciphertext size.

4.5 Comparative Analysis

Our scheme is inspired by [38] (Dyer et al.) which is rooted in the General Approximate Common Divisor (GACD) in the context of Order-Preserving Encryption (OPE). The primary Dyer scheme works directly on integers and hence, the proposed scheme has been further generalized to handle integer as well as floating point data types. Security parameters, key generation, encryption, and decryption procedures are the same as the basic scheme given by Dyer. However, to accommodate the mixed data types, we have introduced two additional steps: additional procedures such as normalization of the message before encryption and denormalization of the message after decryption. These steps are crucial to maintain an accuracy of floating-point forms of data so much more so since they work by involving computations far more complicated than mere integers.

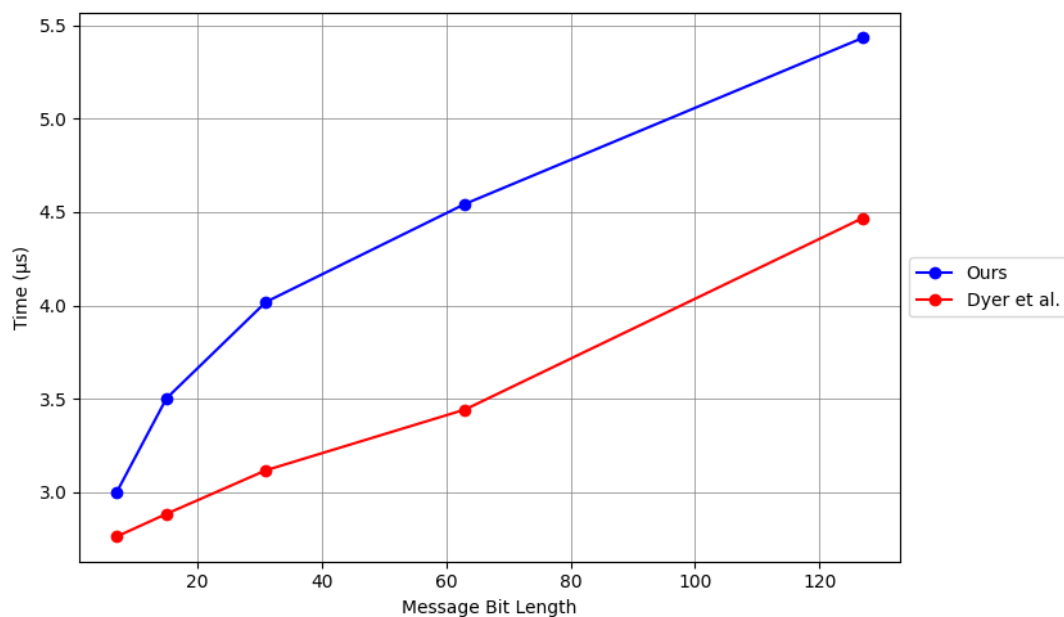


Figure 4.4 Average total execution times

As with the base paper, for performance comparison, we have experimented with the same bit lengths 7, 15, 31, 63, and 127. We compared the performance of our scheme with the original

Dyer scheme using two datasets: one with all the integer terms (for Dyer’s scheme) and the other one with the mixed integer terms and the floating-point terms for our scheme. The evoked results, which are shown in the following graphs, suggest that our scheme is slightly less efficient than the original Dyer algorithm in terms of time consumption. This increase is however owed to the following factors. First, the enlargement of the message space by including floating-point data brings in the same consequences in the security parameters, the key size, and other computational costs. Also, normalization as well as denormalization phases that are required for the management of mixed kinds of data impose a certain impact on the execution time. Although the time and space complexities of our scheme are higher than that of the two previous schemes, our scheme has an added layer with the capacity to support the handling of integer as well as raw decimal data. This increased functionality improves the practical utility of OPE in many contexts and our scheme is more general than previous OPE schemes, which are designed to work only for one data type. In addition to the computational time results, Table 4.2 shows a detailed performance study comparing both the existing and proposed schemes along with their security and flexibility is given below.

Table 4.2 Comparison table to analyze existing and proposed OPE schemes concerning security, performance, and functionality

Scheme	Probabilistic	Deterministic	Homomorphic	Computational Hardness	IND Standards	Data Type (Int, Float, Both)	Complexity
Hybrid GACD-based OPE	✓	×	×	✓	OCCA	Both	$O(1)$
Dyer’s GACD-based OPE [38]	✓	×	×	✓	OCCA	Integer	$O(1)$

The analysis done here clearly shows that both schemes are probabilistic and have IND-OCCA security, which can be considered more secure under the indistinguishability standards. Adding to that, both are based on the computationally hard General Approximate Common Divisor

(GACD) problem which enhances their security of data. It is worth mentioning that concerning the time complexity, both encryption Schemes say, $O(1)$ make the encryption process efficient. Still, as for the general applicability of the methods suggested, it should be noted that in our case we can handle the floating point and integer data whereas the existing scheme can handle only integer data. This enhancement in turn improves the flexibility of our scheme, to cover various types of data and potential applications. As it is evidenced by both proposed schemes, both data security and efficiency characteristics look rather impressive. On the other hand, our scheme adds more computational steps making the programs take more time to execute than the conventional one; but with more additional functionality for accommodating both float and integer data types. This added capability increases flexibility and flexibility and makes our scheme more usable when applied to different scenarios. Thus, the decision for which of the schemes should be implemented depends on the characteristics of the application and the needs that must be met with a minimal impact on execution time.

4.6 Summary

Under the chapter, 'Proposed Methodology,' we first outline the system and threat model for this work to establish the nature of the work and the areas of security that the proposed and improved OPE needs to safeguard against. Based on the GACD problem, the present work generalizes the Dyer et al. scheme to include integer or floating-point data types. To overcome the problem of float data we incorporate normalization before encryption and denormalization after decryption so that float data type can also be passed effectively through this scheme though we have designed this scheme for accommodating mixed data types. In the chapter, the author describes the processes of encryption and decryption with a focus on extra computational steps which, though lead to longer time of the scheme execution, add to the functionality of the scheme. A performance evaluation examines our scheme with other

methods making a distinction between a higher runtime and a heightened adaptability of the process. In sum, the chapter presents the whole picture in which our scheme improves OPE on different data types, maintains a well-balanced between performance and functionality, and enhances flexibility for different applications.

Conclusion and Future Work

5.1 Overview

The findings of our research were summarized in the final chapter, presenting a discourse on new possible developments of the study. It provides the primary consequence of the designed hybrid OPE scheme concerning increased functions and security improvements. This chapter also raises some areas of future work such as the addition of homomorphic operations, the management of key change matters, and the implementation of the scheme into practical uses. This overview agrees with the intent of the chapter that was conducted to present the research findings and areas of future research studies.

5.2 Conclusion of Research

In conclusion, Order Preserving Encryption (OPE) is crucial in the optimization of query and sort operations regarding ordered and encrypted data without compromising the order of plaintext. However, traditional OPE schemes face serious difficulties, especially in terms of security and efficiency of their functioning. As this thesis has pointed out in the analysis of existing OPE schemes, most of them are based on relatively insecure security models like IND-OCPA. As for integer data, the above schemes work well; however, there are issues and threats arising from the enhanced security models that remain unsolved. In particular, the approaches of traditional OPE do not conform to the needful requirements concerning the floating-point

data and do not meet all the demands for the security of contemporary software. To overcome these constraints, we have presented a new hybrid OPE scheme which is based on a newly proposed extension to general approximate common divisor (GACD). Our proposed scheme supports integer and floating-point data type data security under the scheme is proved under IND-OCCA which is a stronger security model than IND-OCPA. Also, our scheme adopts the use of the given GACD problem to add to our security in such a way that it will not be easily penetrable by well-orchestrated attacks. Here we incorporate normalization and denormalization to enforce some variation in mixed-type data to enhance function but preserve probabilistic security. However, all these additional steps add more time to the scheme, but the time complexity is kept $O(1)$ in this scheme. Hence, the results of the performance evaluation prove that our scheme is slightly slower as has been evidenced by computations required for the same, but it provides better functionality and security.

Thus, this research shows that, when progressing the OPE methodologies to integrate the integer and floating-point data types, we secure the scheme by making use of the IND-OCCA model based on the GACD problem, which improves the study significantly. Our hybrid scheme achieves higher functionality while being highly secure and operationally efficient to overcome critical flaws in state-of-the-art OPE solutions.

5.3 Future Work

The future work will be devoted to the solution of one of the major shortcomings of the present GACD-based OPE scheme, where the message list uses one key only. This makes it rather insecure in the sense that if there is a logical key, then there is the danger of a single-point failure especially when dealing with massive data sets. If this key were lost/disclosed to a third party, the entire dataset would be potentially at risk. To improve versatility and address this issue, there is a plan to work on a procedure to split the data set into several subgroups where

each can be protected using a different key. Here, the goal is to decentralize the encryption responsibility, and consequently, lessen the threat of compromising any individual key. Since there will be different keys used for different groups of messages, extended security will have better protection against possible attacks concerning key management systems and will provide higher survivability. This multi-key strategy is aimed at providing a more efficient and versatile concept that allows achieving various degrees of security and providing more practical improvements in the protection of information.

The next study will be related to the extension of the described hybrid OPE scheme by the ability to perform homomorphic operations. This enhancement will enable the scheme to compute over encrypted data to realize their computations without decrypting the data and therefore ensures raw data privacy concerns during processing. Therefore, with homomorphic capabilities integrated, there should be enhancements in the application of the scheme to cover the aspect of secure data computation and analysis in such a way that it will still enable the extra features and at the same time be highly secure and efficient in supporting their implementation as well.

As for future work, we want to learn more about how to use our mixed OPE method in real life because it seems to be working well so far. This needs to be investigated in a few key areas. First, we plan to test the suggested scheme in several real-world settings, including healthcare, banking services, and cloud computing, to see how well it works and how useful it is in these different settings. To find out how well and efficiently the plan works, we will do a lot of testing to see how it handles big datasets and complicated questions. Integration testing is important to see how well the plan works with current security and data management systems, which will make sure everything runs smoothly. We will also be checking that the method is safe against new cryptographic risks, such as those that come from quantum computing. Getting feedback

from people who are already using the plan and experts in the field will help improve it and answer worries about how it can be put into action. Lastly, adding our mixed OPE plan to industry standards and best practices will make it easier for more people to use it and incorporate it into safe data management systems. This all-around method is meant to make our encryption plan more useful and effective, meeting the changing needs of current data protection.

5.4 Summary

The last chapter summarizes the study recommendations by noting the advancement under the context of the advanced hybrid OPE scheme proposed in this study. It can also solve several problems of the existing OPE schemes and improve the interaction of the schemes with floating point data and less secure techniques. Nonetheless, this new scheme that we propose is based on the French General Approximate Common Divisor (GACD) problem; this new scheme therefore provides additional functionalities while the security level is increased; the time complexity remains approximately $O(1)$. The future work includes the extension of the scheme which includes homomorphic operations, implementation of all the changes concerned with the key update on the message groups for eradicating single-point failure, and the relative applicability in realistic scenarios.

References

- [1] FIPS, P. (2001). 197: Advanced encryption standard (AES). National Institute of Standards and Technology, 5(5), 1-51.
- [2] Costan, V., & Devadas, S. (2016). Intel SGX explained. IACR Cryptol. ePrint Arch., 2016, 86.
- [3] Arnautov, S., Trach, B., Gregor, F., Knauth, T., Martin, A., Priebe, C., ... & Fetzer, C. (2016). SCONE: Secure computing on nodes with minimal trust. In 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16) (pp. 689-703).
- [4] Evans, D., Kolesnikov, V., & Rosulek, M. (2018). A pragmatic introduction to secure multi-party computation. Foundations and Trends® in Privacy and Security, 2(2-3), 70-246.
- [5] Huang, Y., Evans, D., Katz, J., & Malka, L. (2011). Faster secure two-party computation using garbled circuits. In USENIX Security Symposium (pp. 539-554).
- [6] Wang, X., Ranellucci, S., & Katz, J. (2017). Global-scale secure multiparty computation. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (pp. 39-56).
- [7] Gentry, C. (2009). A fully homomorphic encryption scheme (Doctoral dissertation, Stanford University).
- [8] Van Dijk, M., Gentry, C., Halevi, S., & Vaikuntanathan, V. (2010). Fully homomorphic encryption over the integers. In Annual International Conference on the Theory and Applications of Cryptographic Techniques (pp. 24-43). Springer.

- [9] Bajard, J. C., Eynard, J., Hasan, M. A., & Zucca, V. (2016). A full RNS variant of FV like somewhat homomorphic encryption schemes. In *International Conference on Selected Areas in Cryptography* (pp. 423-442). Springer.
- [10] Chillotti, I., Gama, N., Georgieva, M., & Izabachène, M. (2020). TFHE: fast fully homomorphic encryption library. *Journal of Cryptographic Engineering*, 10(1), 3-37.
- [11] Liu, J., Asokan, N., Pinkas, B., & Schneider, T. (2017). Secure intersection of ordered sets. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1691-1706).
- [12] Acar, A., Aksu, H., Uluagac, A. S., & Conti, M. (2018). A survey on homomorphic encryption schemes: Theory and implementation. *ACM Computing Surveys (CSUR)*, 51(4), 1-35.
- [13] Brakerski, Z., & Vaikuntanathan, V. (2014). Efficient fully homomorphic encryption from (standard) LWE. *SIAM Journal on Computing*, 43(2), 831-871.
- [14] Boldyreva, A., Chenette, N., Lee, Y., & O'Neill, A. (2009). Order-preserving symmetric encryption. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques* (pp. 224-241). Springer.
- [15] Naveed, M., Kamara, S., & Wright, C. V. (2015). Inference attacks on property-preserving encrypted databases. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security* (pp. 644-655).
- [16] Kerschbaum, F., & Schröpfer, A. (2014). Optimal average-complexity ideal-security order-preserving encryption. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security* (pp. 275-286).

- [17] Popa, R. A., Redfield, C., Zeldovich, N., & Balakrishnan, H. (2011). CryptDB: Protecting confidentiality with encrypted query processing. In Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles (pp. 85-100).
- [18] Boneh, D., Goh, E. J., & Nissim, K. (2005). Evaluating 2-DNF formulas on ciphertexts. In Theory of Cryptography Conference (pp. 325-341). Springer.
- [19] Boelter, T., Davidson, A., Bindschaedler, V., & Shmatikov, V. (2016). Riffle: An efficient communication system with strong anonymity. Proceedings on Privacy Enhancing Technologies, 2016(2), 115-134.
- [20] Chatterjee, S., & Sarkar, P. (2011). HCCA: A homomorphic encryption scheme based on the CCA2 secure scheme of Sahai. In Progress in Cryptology-INDOCRYPT 2011 (pp. 146-162). Springer.
- [21] Landecker, W., Scott, S. L., Wasser, L., & Grammel, J. (2012). A framework for teaching cryptography with JCE. Journal of Computing Sciences in Colleges, 27(6), 117-125.
- [22] Katz, J., & Lindell, Y. (2008). Introduction to modern cryptography. Chapman and Hall/CRC.
- [23] Paillier, P. (1999). Public-key cryptosystems based on composite degree residuosity classes. In International Conference on the Theory and Applications of Cryptographic Techniques (pp. 223-238). Springer.
- [24] Agrawal, R., Kiernan, J., Srikant, R., & Xu, Y., "Order-Preserving Encryption for Numeric Data", In Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data, 2004, pp. 563-574.

- [25] Boldyreva, A., Chenette, N., Lee, Y., & O’neill, A, “Order-Preserving Symmetric Encryption”, In Annual International Conference on the Theory and Applications of Cryptographic Techniques, April 2009, pp. 224-241.
- [26] Wang, C., Cao, N., Li, J., Ren, K., & Lou, W., “Secure Ranked Keyword Search Over Encrypted Cloud Data”, In 2010 IEEE 30th International Conference on Distributed Computing Systems, June 2010, pp. 253-262.
- [27] Liu, D., & Wang, S., “Programmable Order-Preserving Secure Index for Encrypted Database Query”, In 2012 IEEE Fifth International Conference on Cloud Computing, June 2013, pp. 502-509.
- [28] D. Liu and S. Wang, “Nonlinear Order-Preserving Index for the Encrypted Database Query in Service Cloud Environments”, *Concurrency and Computation: Practice and Experience*, 25(13), 2013.
- [29] Popa, R. A., Li, F. H., & Zeldovich, N., “An Ideal-Security Protocol for Order-Preserving Encoding”, In 2013 IEEE Symposium on Security and Privacy, May 2013, pp. 463-477.
- [30] Kerschbaum, F., & Schröpfer, A., “Optimal Average-Complexity Ideal Security Order-Preserving Encryption”, In Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, November 2014, pp. 275-286.
- [31] Krendelev, S. F., Yakovlev, M., & Usoltseva, M, “Order-Preserving Encryption Schemes Based on Arithmetic Coding and Matrices”, In 2014 Federated Conference on Computer Science and Information Systems, September 2014, pp. 891-899.
- [32] Mavroforakis, C., Chenette, N., O’Neill, A., Kollios, G., & Canetti, R., “Modular Order-Preserving Encryption Revisited”, In Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, May 2015, pp. 763-777.

- [33] Jayashri, N., & Chakravarthy, T, “Effective Modular Order-Preserving Encryption on Cloud Using MHGD”, IOSR Journal of Computer Engineering (IOSR-JCE), August 2015, e-ISSN: 2278-0661, p-ISSN: 2278-8727, volume 17, Issue 4, Ver. II, pp. 16-24.
- [34] Yang, C., Zhang, W., & Yu, N., “Semi-Order-Preserving Encryption”, Information Sciences, 2017, volume 387, pp. 266-279.
- [35] Khoury, E., Medlej, M., Abou Jaoude, C., & Guyeux, C, “Novel OrderPreserving Encryption Scheme for Wireless Sensor Networks”, In 2018 IEEE Middle East and North Africa Communications Conference (MENACOMM), April 2018, pp. 1-6.
- [36] Quan, H., Wang, B., Zhang, Y., & Wu, G., “Efficient and Secure Top-K Queries with Top Order-Preserving Encryption”, IEEE Access, 2018, volume 6, pp. 31525-31540.
- [37] Reddy, K. S., & Ramachandram, S., “A Secure, Fast Insert and Efficient Search Order-Preserving Encryption Scheme for Outsourced Databases”, International Journal of Advanced Intelligence Paradigms, 2019, volume 13(1-2), pp. 155-177.
- [38] Dyer, J., Dyer, M., & Djemame, K., “Order-Preserving Encryption Using Approximate Common Divisors”, Journal of Information Security and Applications, 2019, volume 49, 102391.
- [39] Ahmed, S., Zaman, A., Zhang, Z., Alam, K. M. R., & Morimoto, Y., “Semi-Order-Preserving Encryption Technique for the Numeric Database”, International Journal of Networking and Computing, 2019, 9(1), pp. 111-129.
- [40] Chen, S., Li, L., Zhang, W., Chang, X., & Han, Z, “BOPE: Boundary Order-Preserving Encryption Scheme in the Relational Database System”, IEEE Access, 2021, volume 9, pp. 30124-30134.

- [41] Shen, N., Yeh, J. H., Sun, H. M., & Chen, C. M., “A Practical and Secure Stateless Order-Preserving Encryption for Outsourced Databases”, In 2021 IEEE 26th Pacific Rim International Symposium on Dependable Computing (PRDC), December 2021, pp. 133-142.
- [42] Yang, J., & Kim, K. S., “Practical Frequency-Hiding Order-Preserving Encryption with Improved Update”, Security and Communication Networks, 2021.
- [43] Zhan, Y., Shen, D., Duan, P., Zhang, B., Hong, Z., & Wang, B., “MDOPE: Efficient Multi-Dimensional Data Order-Preserving Encryption Scheme”, Information Sciences, 2022, volume 595, pp. 334- 343.
- [44] Bellare, M., Desai, A., Jokipii, E., & Rogaway, P. (1997). A concrete security treatment of symmetric encryption. In Proceedings of 38th Annual Symposium on Foundations of Computer Science (pp. 394-403). IEEE.
- [45] Goldwasser, S., & Micali, S. (1984). Probabilistic encryption & how to play mental poker keeping secret all partial information. In *Proceedings of the fourteenth annual ACM symposium on Theory of computing* (pp. 365-377).
- [46] Rivest, R. L. (2002). Chosen Ciphertext Attacks and Security Proofs. *Journal of Cryptology*, 15(4), 213-222.
- [47] Boneh, D., & Waters, B. (2007). Conjunctive, subset, and range queries on encrypted data. In *Theory of Cryptography Conference* (pp. 535-554). Springer.
- [48] Boldyreva, A., Chenette, N., & O’Neill, A. (2012). Order-preserving encryption revisited: Improved security analysis and alternative solutions. In *Advances in Cryptology – CRYPTO 2012* (pp. 578-595). Springer.

- [49] Yang, C., Zhang, W., & Yu, N., “Semi-Order-Preserving Encryption”, *Information Sciences*, 2017, volume 387, pp. 266-279.
- [50] Gentry, C. (2009). A fully homomorphic encryption scheme. *Stanford University*.
- [51] Dyer, K. P., Coull, S. E., Ristenpart, T., & Shrimpton, T. (2010). Peek-a-boo, I still see you: Why efficient traffic analysis countermeasures fail. *IEEE Symposium on Security and Privacy*, 332-346.
- [52] Naor, M., & Reingold, O. (1997). Number-theoretic constructions of efficient pseudorandom functions. *Proceedings of the 29th Annual ACM Symposium on Theory of Computing (STOC)*, 203-213.
- [53] Peikert, C. (2016). A decade of lattice cryptography. *Foundations and Trends® in Theoretical Computer Science*, 10(4), 283-424.
- [54] Regev, O. (2005). On lattices, learning with errors, random linear codes, and cryptography. *Proceedings of the 36th Annual ACM Symposium on Theory of Computing (STOC)*, 84-93.
- [55] Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to algorithms* (3rd ed.). MIT Press.
- [56] Knuth, D. E. (1998). *The art of computer programming: Volume 1, Fundamental algorithms* (3rd ed.). Addison-Wesley.
- [57] Sedgewick, R., & Wayne, K. (2011). *Algorithms* (4th ed.). Addison-Wesley.
- [58] Smart NP, Albrecht MR, Lindell Y, Orsini E, Osheter V, Paterson KG, et al. LIMA: A PQC encryption scheme Tech. Rep.. University of Bristol; 2017.

- [59] Cheon JH, Kim D, Lee J, Song Y. Lizard: cut off the tail! Practical post-quantum public-key encryption from LWE and LWR. Cryptology ePrint Archive, Report 2016/1126; 2016b.
- [60] Kellaris G, Kollios G, Nissim K, O’Neill A. Generic attacks on secure outsourced databases. In: Proceedings of the 2016 ACM SIGSAC conference on computer and communications security (CCS ’16). ACM; 2016. p. 1329–40.
- [61] Cohn H, Heninger N. Approximate common divisors via lattices. In: Proceedings of the 10th algorithmic number theory symposium (ANTS-X), 1. Mathematical Sciences Publishers; 2012. p. 271–93. doi:10.2140/obs.2013.1.271
- [62] Chen Y, Nguyen PQ. Faster algorithms for approximate common divisors: breaking fully homomorphic encryption challenges over the integers. Cryptology ePrint Archive, Report 2011/436; 2011.
- [63] Galbraith SD, Gebregiyorgis SW, Murphy S. Algorithms for the approximate common divisor problem. LMS J Comput Math 2016;19(A):58–72.
- [64] Boldyreva A, Chenette N, Lee Y, O’Neill A. Order-preserving symmetric encryption. In: Proceedings of the 28th annual international conference on the theory and applications of cryptographic techniques (EUROCRYPT 2009). Springer Verlag; 2009. p. 224–41.
- [65] Boldyreva A, Chenette N, O’Neill A. Order-preserving encryption revisited: improved security analysis and alternative solutions. In: Proceedings of the 31st annual cryptology conference (CRYPTO 2011). Springer-Verlag; 2011. p. 578–95.
- [66] Teranishi I, Yung M, Malkin T. Order-preserving encryption secure beyond one-wayness. In: Proceedings of the 20th international conference on the theory and application of

- cryptology and information security (ASIACRYPT 2014). Springer-Verlag; 2014. p. 42–61.
- [67] Xiao L, Yen I-L. Security analysis for order-preserving encryption schemes. In: Proceedings of the 46th annual conference on information sciences and systems (CISS 2012). IEEE; 2012. p. 1–6.
- [68] Gentry, C., & Peikert, C. (2012). Fully Homomorphic Encryption with Relatively Small Key and Ciphertext Sizes. In *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing (STOC)*, 169-178.
- [69] Kerschbaum, F. (2015). Frequency-Hiding Order-Preserving Encryption. *Journal of Cryptology*, 28(1), 22-42.
- [70] Agrawal, D., & Kiernan, J. (2004). Order-Preserving Encryption for Numeric Data. In Proceedings of the ACM SIGMOD International Conference on Management of Data, 563-574.
- [71] Fuchs, E., & Jarecki, S. (2017). Practical Techniques for Order-Preserving Encryption with Support for Range Queries. *IEEE Transactions on Knowledge and Data Engineering*, 29(8), 1811-1825.
- [72] Liu, X., & Wang, S. (2020). Enhancing Order-Preserving Encryption for Secure Multi-Dimensional Query Processing. *ACM Transactions on Privacy and Security*, 23(4), 1-27.
- [73] Naor, M., & Nissim, K. (2009). Integrity-Sensitive Order-Preserving Encryption. In *Proceedings of the 30th Annual International Cryptology Conference (CRYPTO)*, 108-125.

- [74] Boldyreva, A., Chenette, N., Lee, Y., & O'Neill, A. (2009). Order-preserving symmetric encryption. *Advances in Cryptology–EUROCRYPT 2009*. Springer, Berlin, Heidelberg.
- [75] P. Wang, J. Yang, and C. Hsu, "Order-Preserving Encryption Scheme for Single-Party Cloud Computing," *Journal of Cloud Computing: Advances, Systems and Applications*, vol. 6, no. 1, pp. 15-30, 2017.
- [76] Katz, J., & Lindell, Y. (2007). *Introduction to Modern Cryptography: Principles and Protocols*. Chapman & Hall/CRC.