PeerJ Computer Science

An adaptive cryptosystem on a Finite Field

Awnon Bhowmik[1] and Unnikrishnan Menon[2]

[1] Department of Mathematics, The City College of New York, New York, NY, United States of America
[2] Department of Electrical and Electronics Engineering, Vellore Institute of Technology University, Vellore, Tamil Nadu, India

ABSTRACT

Owing to mathematical theory and computational power evolution, modern cryptosystems demand ingenious trapdoor functions as their foundation to extend the gap between an enthusiastic interceptor and sensitive information. This paper introduces an adaptive block encryption scheme. This system is based on product, exponent, and modulo operation on a finite field. At the heart of this algorithm lies an innovative and robust trapdoor function that operates in the Galois Field and is responsible for the superior speed and security offered by it. Prime number theorem plays a fundamental role in this system, to keep unwelcome adversaries at bay. This is a self-adjusting cryptosystem that autonomously optimizes the system parameters thereby reducing effort on the user's side while enhancing the level of security. This paper provides an extensive analysis of a few notable attributes of this cryptosystem such as its exponential rise in security with an increase in the length of plaintext while simultaneously ensuring that the operations are carried out in feasible runtime. Additionally, an experimental analysis is also performed to study the trends and relations between the cryptosystem parameters, including a few edge cases.

Subjects Computer Networks and Communications, Cryptography, Security and Privacy
Keywords Galois Field, Finite field, Product Exponent Modulo, Prime number theorem, Threshold cryptography, Autonomous encryption scheme

Submitted 9 April 2021
Accepted 20 June 2021
Published 16 August 2021

Corresponding author
Awnon Bhowmik,
abhowmik901@york.cuny.edu

Academic editor
Mamoun Alazab

Additional Information and Declarations can be found on page 16

DOI **10.7717/peerj-cs.637**

INTRODUCTION

Cryptography is the art of hiding messages to provide it with a certain level of security to maintain confidentiality and integrity. This new idea, whether it was to hide secret messages, or to transform the original message to make it look fancy, dignified, etc. continued through the medieval ages, the renaissance period saw the birth of the polyalphabetic substitution cipher, called the Vigenère Cipher (*Rubinstein-Salzedo, 2018*). An encryption device called the Enigma machine (*Singh, 1999*) was used by the Nazi Germans during World War II. Although history suggests that it has been in use for ages, systematic study of cryptology as a science (and perhaps an art) just started around one hundred years ago (*Sidhpurwala, 2013*).

But it was not until the 1970s, that studies in cryptography got serious. Data Encryption Standard (DES) was introduced by IBM in 1976 (*Tuchman, 1997*) followed by Diffie Hellman Key Exchange in the same year (*Kallam, 2015*). In 1977, RSA came along (*Calderbank, 2007*) and in 2002, AES was accepted as a standard security protocol

to be used in both hardware and software (*Dworkin et al., 2001*). And thus cryptography became popular.

The strength or foundation of a modern encryption protocol relies upon the inherent Trapdoor Function. As classical cryptography evolved, it has become clear that some key components are essential in making stronger trapdoor functions, also known as one-way functions. Studies have shown that prime numbers are an essential part of numerous cryptosystems, and with a bit of effort, numerous mathematical concepts can be used to generate stronger cryptosystems. Conventional, widely used algorithms such as RSA rely on integer products involving large primes. Breaking this system is essentially an attempt to solve the integer factorization problem, which can be readily attained using Shor's algorithm, Pollard's Rho algorithm, etc (*Aminudin & Cahyono, 2021*; *De Lima Marquezino, Portugal & Lavor, 2019*).

Cryptography algorithms rely on integer mathematics, in particular, number theory to perform invertible operations such as addition, multiplication, exponentiation, etc. over a finite set of integers. Finite Fields, also known as Galois Fields, are fundamental to any cryptographic understanding. A field can be defined as a set of numbers that we can add, subtract, multiply and divide together and only ever end up with a result that exists in our set of numbers. This is mainly advantageous in cryptography since we can only work with a small number of incredibly huge numbers (*Kohli, 2019*). When cryptography algorithms rely solely on converting raw string data in ASCII format, we are restricted to 256 different characters only. Doing such leaves us with only a handful amount of invertible operations in modulo 256 On the other hand, the Galois Field $GF(2^8)$ offers numerous such operations. In fact, Advanced Encryption Standard (AES) (*Daemen & Rijmen, 2001*) uses the multiplicative inverse in $GF(2^8)$. Using the Galois Field also shines forth the opportunity to use the concepts of irreducible polynomials (*Shoup, 1990*). In AES, addition and subtraction is a simple XOR operation. For multiplication, it uses the product modulo an irreducible polynomial. For example, the integer 283 refers to the irreducible polynomial $f(x) = x^8 + x^4 + x^3 + x + 1$ in $GF(2^8)$ whose coefficients are in $GF(2)$ (*Desoky & Ashikhmin, 2006*).

Threshold cryptography is a form of security lock where private keys are distributed among multiple clients or systems. They are even asked to provide digital signature authentication for verification purposes. Only when these keys are combined, can information be effectively decrypted. In practice, this lock is an electronic cryptosystem that protects confidential information, such as a bank account number or an authorization to transfer money from that account (*Henderson, 2020*). The encryption scheme described in this paper has traits that resemble threshold cryptography. Existing threshold cryptosystem protocols might benefit from the positive aspects of our system, making it a viable alternative contender soon. The suggested approach can also be integrated into intelligent systems that use master–slave communication topologies, such as swarm robots (*Chen & Ng, 2021*).

The technique suggested in this study uses an inventive trapdoor function based on the finite field to handle data encryption in cases with enormous string lengths in a reasonable amount of time, demonstrating that it is extremely light. This is a self-adjusting cryptosystem that optimizes the system parameters on its own, saving the user time and

effort while increasing security. The inherent lightness of this cryptosystem makes it an ideal contender for applications involving IoT devices with limited computational power. Confusion and diffusion, covered in 'Observed Security Features', are two aspects of a safe cipher's functioning in cryptography. Due to the system's demonstration of confusion and diffusion properties, it could potentially be used in scenarios such as encrypting bank transaction details, where a high degree of variance in the ciphertext is desirable upon altering few characters in the plaintext. Furthermore, data from our benchmarks in section 'Experimental Analysis' shows promising results when tested on large chunks of data proving that given sufficient computing power, this system could potentially be used for confidential military applications or as a layer of security for the compilation of large datasets in Big Data analytics.

The remainder of this paper is organized as follows. ''Literature Review'' gives a brief description of numerous sectors where the proposed system can be introduced. ''Trapdoor Function'' section explains in brief, the working of a traditional trapdoor function from a mathematical perspective. 'Prime number theorem', 'Galois Field in Cryptography', 'Generating Upper Bound for q' and 'Fermat's Factorization' describes the required preliminaries for a better understanding of the algorithm that follows in 'Proposed Algorithm'. Next section talks about two essential properties of the operation of a secure cipher, before moving onto 'Experimental Analysis'. Next, a few ways is covered in which an adversary might try to break into systems running this cryptosystem. 'From an interceptor's perspective' shows that it would be near impossible for them to achieve their goal. 'Remarks on edge cases' addresses an edge case of the system that revolves around the inbuilt block size optimization function. The paper concludes by briefly summarizing the study's overall accomplishments and providing important insights into future research directions.

LITERATURE REVIEW

Recently in the field of Internet of Things (IoT), research has been conducted on flexible privacy-preserving data publishing schemes in the sector of smart agriculture. Their study shows that over the years protection and privacy concerns for smart agriculture have grown in importance. In these IoT-enabled systems, the internet is used for communicating with participants. Since the cloud is often untrustworthy, higher privacy standards are needed (*Song et al., 2020*).

Security and privacy at the physical layer have become a serious challenge in recent years for numerous communication technologies. IoT networks are typically comprised of a network of interconnected sensors and information relaying units that communicate in real-time with one another. Individual nodes typically have specialized sensor units for detecting specific environmental attributes and have fewer computing resources available. For example, in a house, various technologies such as facial recognition, video monitoring, smart lighting, and so on will all function in tandem. Security and privacy are key impediments to the realistic deployment of smart home technologies (*Shen et al., 2018*).

The majority of the network's elements use sensitive user data and seamlessly exchange information with one another in real-time. To keep intruders out of such a network,

a dependable and stable solution based on edge computing is preferred. Another real-world application of secure edge computing lies in the domain of smart grids. Smart grids are recognized as the next-generation intelligent network that maximizes energy efficiency (*Wang et al., 2020*). Smart grid solutions help to monitor, measure and control power flow in real-time that can contribute to the identification of losses, and thereby appropriate technical and managerial actions can be taken to prevent the same. Smart grids generally rely on data recorded by energy meters from different houses. Since electricity usage data can be classified as confidential user metrics, there is a need for implementing a layer of security before transmitting this data to other parties for further analysis. Encryption of this data at the smart energy meter stage itself can be beneficial. However, this requires the development of a lightweight encryption protocol that can be easily integrated with microprocessors with minimal compute power.

The amount of data provided by users during numerous online activities has increased dramatically over the last decade. Celestine Iwendi et al. performed research that used a model-based data analysis technique for handling applications with Big Data Streaming to glean useful information from this massive amount of data. The method suggested in their research has been tested to add value to large text data processing (*Iwendi et al., 2019*). Our proposed schematic leverages an ingenious trapdoor function based on the finite field to handle data encryption in scenarios involving large string lengths within feasible runtime, proving it to be considerably lightweight.

TRAPDOOR FUNCTION

The essence of any cryptosystem relies on some special mathematical trapdoor function that makes it practically impossible for an unwelcome interceptor to gain access to secretive information. Simultaneously, these functions also ensure that the authorized parties (who know the secret key) can continue sharing data among themselves.

A trapdoor function is a mathematical transformation that is easy to compute in one direction, but extremely difficult (practically impossible) to compute in the opposite direction in feasible runtime unless some special information is known (private key). Analogously, this can be thought of like the lock and key in modern cryptography where until and unless someone has access to the exact key, they can't open the lock. In mathematical terms, if f is a trapdoor function, then $y = f(x)$ easy to calculate but $x = f^{-1}(y)$ is tremendously hard to compute without some special knowledge k (called key). In case k is known, it becomes easy to compute the inverse $x = f^{-1}(x, k)$.

The components of the proposed system in this paper that act as the trapdoor function is the modulo operation on a Galois field (*Benvenuto, 2012*).

PRIME NUMBER THEOREM

Positive integers that are divisible by 1 and itself, are known as prime numbers. The sequence begins like the following...

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, \ldots$$

Table 1 Prime density and approximation to logarithmic integral.

Search Size x	# of Primes	Density (%)	li(x)	$\frac{\text{li}(x)-\pi(x)}{\pi(x)}$ x 100
10	4	40	6.16	54.14
10^2	25	25	30.13	20.50
10^3	168	16.8	177.61	5.72
10^4	1229	12.3	1246.14	1.39
10^5	9592	9.6	9629.81	0.39
10^6	78498	7.8	78625	0.17
10^7	664579	6.6	664918	0.05
10^8	5761455	5.8	5.76×10^6	0.01

and has held untold fascination for mathematicians, both professionals and amateurs alike. A result that gives an idea about an asymptotic distribution of primes is known as the prime number theorem (*Goldstein, 1973*).

$\pi(x)$ is the prime-counting function that gives the number of primes less than or equal to x, for any real number x. This can be written as

$$\pi(x) = \sum_{p \leq x} 1 \tag{1}$$

It is seen *via* graphing, that $\frac{x}{\ln x}$ is a good approximation to $\pi(x)$, in the sense that the limit of the quotient of the two functions $\pi(x)$ and $\frac{x}{\ln x}$ as x increases without bound is 1.

$$\lim_{x \to \infty} \pi(x) \frac{\ln x}{x} = 1 \tag{2}$$

This result can be rewritten in asymptotic notation as

$$\pi(x) \sim \frac{x}{\ln x} \tag{3}$$

The logarithmic integral provides a good estimate to the prime density function.

$$\frac{\pi(x)}{x} \sim \text{li}(x) = \int_2^x \frac{1}{\ln t} \mathrm{d}t \tag{4}$$

To get an idea of the distribution of primes, it is important to count the number of primes in a given range and find the percentage of primes. Consider an infinitely tall tree, losing its leaves. The leaves represent prime numbers. Most leaves are found near the root, and the number of leaves reduces as we walk away from the center. But no matter how far we are from the center, we always find more leaves. These leaves are unpredictably scattered in an infinite area surrounding the tree. This is the situation with the distribution of primes as resembled by Table 1. Figure 1 shows the prime density and logarithmic integral on the left, and the asymptotic nature of the prime counting function on the right.

GALOIS FIELD IN CRYPTOGRAPHY

Galois Field, named after Evariste Galois, also known as finite field, refers to a field in which there exist finitely many elements. A computer only understands the binary data