Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
DOI: https://doi.org/10.17762/msea.v70i1.2513

Analysis of Computational Algebra for Cryptography and Coding Theory Applications

Ashok Singh Bhandari
Asst. Professor, Department of Mathematics, Graphic Era Hill University,
Dehradun Uttarakhand India

Article Info
Page Number: 598-606
Publication Issue:
Vol. 70 No. 1 (2021)

Article History
Article Received: *25 January 2021*
Revised: *24 February 2021*
Accepted: *15 March 2021*

Abstract

Secure communication is essential in many areas of today's interconnected society, from online commerce to international affairs. The foundation of data security is made up of encryption and decryption algorithms, and effective execution calls for an organised approach rather than a chance procedure. This study analyses several encryption techniques with an emphasis on those that apply algebraic coding theory to cryptography. These algebraic coding theory-based cryptographic systems are particularly significant for the future of data security because, in contrast to many existing systems, they provide resistance against attacks from quantum computers.This essay emphasises the necessity of an organised approach while outlining the essential concepts of encryption and decryption. It examines several encryption techniques, including symmetric and asymmetric encryption, pointing out their advantages and disadvantages.The study then explores how algebraic coding theory is used in cryptography. It explores the fundamental mathematical concepts that underlie algebraic coding methods, including finite fields and polynomial arithmetic. The McEliece cryptosystem and Niederreiter cryptosystem are two examples of unique algebraic coding theory-based cryptographic schemes that are the subject of this examination, which also highlights their benefits in terms of security and resilience to quantum assaults. These methods' computational complexity is also analysed, taking into account things like key sizes and encryption speed.With the help of the knowledge gleaned from this investigation, strong encryption systems that can survive upcoming developments in computing technology can be created, guaranteeing the secrecy and integrity of critical data in daily life.

Introduction

Applications like internet commerce, foreign diplomacy, and email communication require encryption. It is vital to use an organised strategy in order to assure effective encryption and decryption. The extensive use of the cryptosystem and system security are both improved by this structure.This study investigates several mathematical foundations for encryption and decryption techniques. Beginning with a description of the fundamentals of cryptography, it makes a distinction between private-key and public-key cryptosystems. Public-key cryptography assures the asymmetry of encryption and decryption to stop knowledge leakage, in contrast to private-key cryptography's symmetric encryption and decryption. With a focus on the rise of quantum computers, the importance of cryptography and the creation of alternative cryptosystems are examined [1].

The algebraic coding theory, [2] which is concerned with encoding and decoding messages to assure their dependability even in the presence of noise during transmission, is introduced in this section. The processes for adding redundancy and creating effective error-correcting codes are what are being concentrated on. In particular, Reed-Solomon codes are emphasised. These codes are created using polynomial evaluations at particular finite field elements, and polynomial interpolation is used for decoding.In the context of the McEliece cryptosystem, this section focuses on the integration of coding theory and cryptography. The [4]McEliece cryptosystem's encryption and decryption procedures are examined. Along with alternate codes like Reed-Solomon codes, the Niederreiter variant of the McEliece cryptosystem is also looked at.

To reduce key size in the McEliece cryptosystem, the implementation of Reed-Solomon codes is suggested. However, a flaw in this particular cryptosystem has been found. This flaw results from an attack that specifically takes advantage of the McEliece cryptosystem's Reed-Solomon code structure. This presents a problem because it restricts the system's ability to use Reed-Solomon codes directly.

I. Background for Cryptography

Encryption methods are used in cryptography to protect data and communications. Public-key (asymmetric) and private-key (symmetric) cryptosystems are the two main types of cryptographic systems [3] .A shared secret key is used in private-key cryptography, commonly referred to as symmetric cryptography, for both encryption and decryption. Both the sender and the recipient encrypt and decrypt the data using the same key. The communicating parties must securely share the key while maintaining its confidentiality. The Data Encryption Standard (DES), Advanced Encryption Standard (AES), and the Rivest Cypher (RC) family of algorithms are a few examples of private-key cryptosystems. Private-key cryptography is effective for encrypting vast amounts of data, but it presents difficulties for key distribution in massive communication networks [6].

A public key and a private key are used in public-key cryptography, commonly referred to as asymmetric cryptography [8]. While the private key is kept a secret by the owner, the public key is widely dispersed and can be freely shared. While the private key is required for decryption, the public key is utilised for encryption. Only the appropriate private key can be used to decrypt messages that have been encrypted with the public key. Due to its asymmetrical structure, it is possible to communicate securely without using a shared secret key. The RSA method, Diffie-Hellman key exchange, and elliptic curve cryptography (ECC) are examples of public-key cryptosystems. Public-key cryptography provides improved security and supports a number of cryptographic protocols, including key exchange and digital signatures [9].

Secure communication systems sometimes have a combination of private-key and public-key encryption as their foundation [10]. A more effective private-key technique can be used to encrypt subsequent communication once a safe key exchange has taken place using public-keycryptography. Utilising the benefits of both types of cryptosystems, this hybrid technique.Overall, private-key and public-key cryptosystems are crucial elements of contemporary cryptography. They both have unique advantages and cater to specific security needs in a variety of applications.

1. Private Key Cryptosystem

The same secret key is used for both data encryption and decryption in a private-key cryptosystem, also referred to as a symmetric cryptosystem. The key is kept secret and needs to be safely transferred between the persons involved in communication.In a private-key cryptosystem, the encryption process entails using the secret key and an encryption algorithm to transform plaintext (the original message) into ciphertext (the encrypted message). The ciphertext can then be sent via a channel that isn't secure. The ciphertext is decrypted on the recipient's end using the same secret key and a decryption algorithm [12][13].

Private-key cryptosystems [15] are appropriate for applications that require bulk data encryption, such as safe storage, secure file transfer, and symmetric key-based authentication protocols, since they efficiently perform encryption and decryption operations. The Data Encryption Standard (DES), Advanced Encryption Standard (AES), Triple Data Encryption Standard (3DES), and the Rivest Cypher (RC) family of algorithms are a few examples of private-key algorithms. Distributing the secret key to all parties involved in a private-key cryptosystem safely is one of the main difficulties. The key needs to be kept private and secured from unauthorised use. A secure and reliable channel must be utilised for key exchange in private-key cryptosystems because the same key is used for both encryption and decryption.

When compared [14] to public-key cryptosystems, private-key cryptosystems typically perform computations more quickly. They do not, however, offer the same amount of scaling and key management as public-key cryptosystems. As a result, hybrid cryptosystems that combine the benefits of both methods are frequently made using private-key and public-key cryptography.

2. Public Key Cryptosystem

Asymmetric cryptosystems are another name for public-key systems, which use two sets of mathematically related keys—a public key and a private key—for encryption and decryption. While the owner keeps the private key a secret, the public key is freely shared.A public-key cryptosystem encrypts data using the public key, whereas decrypting it requires the private key. Anyone can encrypt a message using the public key, but only the intended receiver and the associated private key can decrypt the message and obtain the original plaintext [13].

Public-key cryptography's core idea is based on the computational complexity of several mathematical puzzles, including prime factorization or discrete logarithm. The RSA algorithm, which depends on the difficulty of factoring big composite numbers, is the most popular public-key cryptosystem.Compared [17] to private-key cryptosystems, public-key cryptosystems have a number of benefits. First off, because the public key can be freely circulated, they do away with the necessity for secure key distribution mechanisms. As a result, public-key cryptography is more scalable and appropriate in situations where safe key exchange is difficult or in massive communication networks.

2.1 RSA Cryptosystem

A well-liked public-key cryptosystem is RSA, which depends on how difficult it is to factor huge integers. The sum of two big primes (n = pq) and the encoding exponent E make up its public keys. The secretive private keys are made up of the 300-digit-long prime factors p and q as well as the decoding exponent D. The Euler ϕ -function, which counts the amount of prime natural numbers that are fewer than n, is the foundation for the security of the RSA algorithm [18].

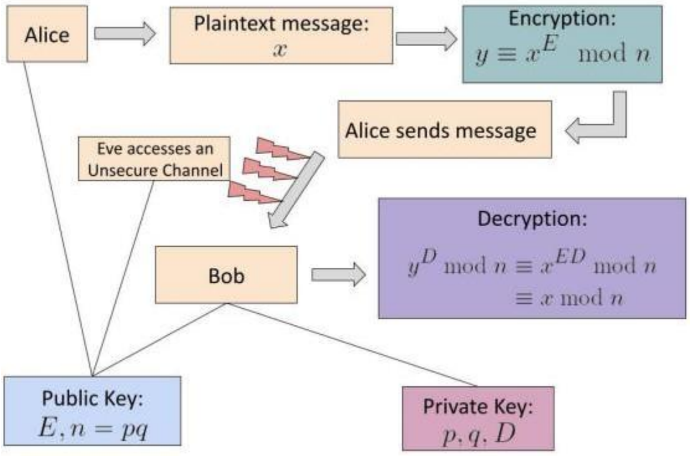
Alice
Plaintext message: x
Encryption: $y \equiv x^E \mod n$
Alice sends message
Eve accesses an Unsecure Channel
Bob
Decryption: $y^D \mod n \equiv x^{ED} \mod n \equiv x \mod n$
Public Key: $E, n = pq$
Private Key: p, q, D

Figure 1: Flow diagram of RSA working for Encryption and Decryption

Consider the example for RSA public Key cryptosystem,In this case, Alice wants to communicate with Bob using encryption. Bob has revealed his encryption settings to the world, including his public key, which contains the values n = 713 and E = 13. 420 is Alice's message. They can now continue working on the encryption and decryption calculations.

$Y= 420^{13}\text{MOD } 713$

$= 420^{8}* 420^{4} * 420 \text{ MOD } 713$

$= 18 * 100 * 420 \text{ MOD } 713$

$= 220$

and Alice sends 220 to Bob. Given that Bob is aware of n's factorization, he can calculate

$$\phi(\mathrm{n}) = (\mathrm{P} - 1)(\mathrm{Q} - 1)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

In order to ensure that ED 1 mod n, he utilises the Extended Euclidean Algorithm to obtain D such that ED + ϕ(n)k = 1 for some k ∈ Z. In particular, he may then calculate

$C = 220^{457}\text{MOD } 713$

$= 420$

The private key that Bob will use to decrypt is represented by the value of D. So he is aware of

$C= 360^{457} \text{ mod } 660$

and he is able to retrieve the initial message 420.

II. Coding Theory in Cryptosystem

The goal of coding theory is to make it possible for communications to be reliably transmitted even in the presence of noise that could cause some of the message to be distorted during transmission. To do this, the message is given structured redundancy, which enables mistake detection and correction. The [12] repetition code is a fundamental encoding system in which the message is repeated numerous times. Receiver can spot problems by comparing the repeated messages. There are, however, more effective techniques that use fewer superfluous bits while still ensuring that the recipient will understand the message. These cutting-edge coding methods establish a balance between effectiveness and redundancy, improving communication reliability [4].

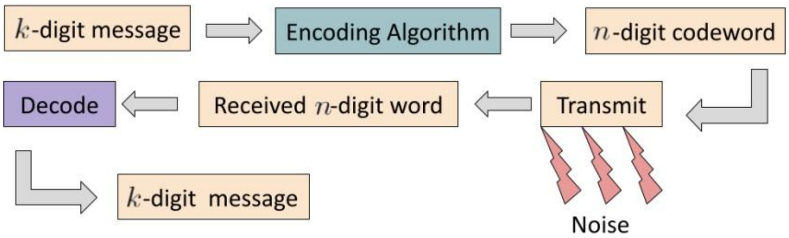

Figure 2: Standard coding theory model

According to coding theory, the codeword is communicated after being encoded. If a word is received without any errors, it is presumed that no errors occurred and the word is decoded exactly as it was. The received word will, however, deviate from all the codewords by at least one bit if mistakes resulting from transmission noise are present. The goal of coding theory is to create codes and decoding techniques that are capable of identifying and fixing a specific number of faults in a received word. Coding theory provides reliable error detection and repair during communication by creating codes with the right qualities.

The implementation of a parity check, which counts the number of 1s in a bit string, is another coding technique. To verify that the message contains an even amount of 1s, an extra bit is added. The receiver can tell if there is an error by looking at the parity of the received message and counting the 1s. An error is found if the parity is unusual. A parity check is not completely reliable for identifying and fixing errors, though. When more than one error occurs or when two distinct errors balance each other out to produce an even parity, it is ineffectual. A parity check also merely detects the error's presence rather than locating it precisely.

1. Linear Code and Block Codes

Block codes are an effective mathematical method that allows for the detection and correction of many faults in addition to simple error detection. Block codes include encoding a constant number of k binary digits within a larger sequence of n binary digits. An (n, k)-block code, then, consists of an encoding function that converts k digits into n digits and a decoding function that does the opposite. Block codes are more resilient and dependable for error control in communication systems because they can both detect and rectify faults.

For Encoding E : $Z_2^k \to Z_2^n$.. (2)

For Decoding D : $Z_2^n \to Z_2^k$.. (3)

A measure of dissimilarity between two n-tuples known as the Hamming distance is the number of bits that separate the two n-tuples. For instance, the 4-tuple 0100 and 0001 have a Hamming distance of 2 between them. The smallest Hamming distance between any two unique codewords within a code is referred to as the minimum distance of a code in coding theory. The ability of a code to discover and rectify errors is determined by the minimum distance, which is an important metric. Greater mistake detection and correction capability is implied by a bigger minimum distance. In order to assess the effectiveness and dependability of codes in terms of error detection and repair, the Hamming distance is crucial.

Example:
If a code has a minimum distance of 2M + 1, it possesses the capability to correct m or fewer errors and detect up to 2M errors.
Consider a code that has a minimum distance of 2M plus 1. Take into account a codeword x that is transmitted, and let z be the word that is received with a maximum of **M** mistakes.
Thus D(X, Z) ≤ M. Let y be another codeword such that x   y.
By the triangle inequality:$d(x, y) \leq d(x, z) + d(y, z)$
So it may represent as:
$M + 1 = 2M + 1 - M \leq D(X, Y) - D(X, Z) \leq D(Y, Z)$...(4)
This discovery reveals a crucial characteristic of codes with a minimum distance of at least m + 1. In such codes, the intended codeword x will be correctly decoded when the received word z differs from a codeword y by a Hamming distance of m or less. This is so that the received word will always be closer to the proper codeword than any other codeword, as long as the minimum distance is met. Additionally, the received word cannot be a valid codeword if the Hamming distance between it and the codeword x is 2m or less, and the existence of 2m mistakes will be recognised.This characteristic makes sure that codes with a minimum distance of at least m + 1 will be accurate in error repair and detection.The significant characteristic of linear codes is that they are the null space of a k by n matrix over Z2. The collection of all x Zn2 that satisfy the formula HxT = 0 is known as a linear code, and H Mkn(Z2) acts as the parity-check matrix for the linear code. According to a specific parity check for a codeword specified by each row of H, a subset of entries must contain an even number of 1s. Due to their effective encoding and decoding capabilities, linear codes are useful tools for applications such as data transfer and error control.

2. (RS) Reed-Solomon Codes

Algebraic structure is used by polynomial codes for both encoding and decoding. The traditional vector representation is expanded to include messages and codewords as polynomials. By specifying the coefficients of the polynomial, an m-tuple can be converted into a polynomial of degree m - 1. For instance, the polynomial 1 + x + 0x2 + 0x3 + x4 or just 1 + x + x4 can be used to represent the 5-tuple 11001. The encoding and decoding of polynomial codes can take advantage of algebraic characteristics thanks to this polynomial representation.
The original polynomial must be retrieved from the incoming data in order to decode the polynomial P(x) in polynomial codes. P(x) has a degree of k - 1, hence k points can be used to uniquely identify it. Thus, out of the n = pm accessible equations, decoding entails solving a system of k equations. This method makes it possible to recreate the original polynomial and makes it simpler to find and fix mistakes in polynomial codes.
P(0) = A0
P(A) = A0 + A1A + A2A 2 +   + AK−1A K−1
P(A 2) = A0 + A1A 2 + A2A 4 +   + AK−1A 2K−2
.
.
.
P(1) = A0 + A1 + A2 +⋯⋯+ AK−1
Reed-Solomon codes can also be represented as linear codes produced by a matrix with the following formula:

$$\begin{bmatrix} F\ 1\ 1\ 1 \ldots 1 \\ 0\ \alpha_1\ \alpha_2 \ldots \alpha_k \\ 0\ \alpha_1^2\ \alpha_2^2 \ldots \alpha_k^2 \\ \ldots \\ \ldots \\ 0\ \alpha_1^k\ \alpha_2^k \ldots \alpha_k^k \\ \ldots \\ \ldots \\ 0\ \alpha_1^m\ \alpha_2^m \ldots \alpha_k^m \end{bmatrix}$$

Where,
where the variables i and stand for the polynomial's coefficients and the field's constituents, respectively. The compact representation of the k equations needed to decode the polynomial in polynomial codes is provided by this matrix.

3. System of McEliece cryptography

Goppa codes, which are built similarly to Reed-Solomon codes, were the first type of coding used by the McEliece system. An algebraic curve's function field is evaluated at particular points along the curve to produce goppa codes.

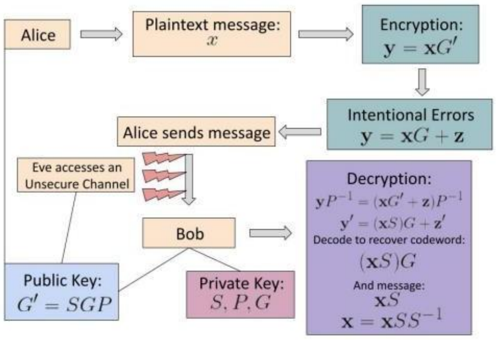
Alice
Plaintext message: x
Encryption: $\mathbf{y} = \mathbf{x}G'$
Intentional Errors $\mathbf{y} = \mathbf{x}G + \mathbf{z}$
Alice sends message
Eve accesses an Unsecure Channel
Bob
Decryption:
$\mathbf{y}P^{-1} = (\mathbf{x}G' + \mathbf{z})P^{-1}$
$\mathbf{y}' = (\mathbf{x}S)G + \mathbf{z}'$
Decode to recover codeword:
$(\mathbf{x}S)G$
And message:
$\mathbf{x}S$
$\mathbf{x} = \mathbf{x}SS^{-1}$
Public Key: $G' = SGP$
Private Key: S, P, G

Figure 3: System Flowchart for McEliece cryptography

These codes extend the idea of Reed-Solomon codes and are a subset of the larger class of algebraic geometry codes. Goppa codes are notable for having a length of N = 2M and a dimension of k n - sm, where s denotes the greatest number of faults that the code can fix.

After the reception of the vector y, the decryption algorithm is employed to recover the original codeword x. The initial step involves computing Y' = $PP^{\wedge(-1).}$ It is important to note that Y' corresponds to a distorted version of a codeword within the Goppa code. This can be observed as follows:

$K' = (XG' + Z)P^{-1}$
$= (XG')P{-}1 + ZP^{-1}$
$= XSG + ZP^{-1}$
$= (XS)G + Z'$

The McEliece cryptosystem is vulnerable from two directions. In the first attack, S, P, and G are extracted from the masked codeword G'. In the second attack, a foe tries to create the original vector x from the communicated vector y using a brute-force strategy. These assaults each have a specific goal, such as finding weaknesses in the system's components or doing a thorough search for the right decoding.

The parity-check matrix H can be recovered using this approach, making it possible for an eavesdropper to correctly decode communications. Reed-Solomon codes' innate structure, which enables more condensed representation and smaller key sizes, was the initial driving force behind their use. However, as the attack showed, this extra structure can be used by an eavesdropper to access the private key's parity-check matrix H and undermine the encryption. Because they introduce serious security flaws, Reed-Solomon codes cannot be directly used in either the Niederreiter variant or the original McEliece cryptosystem.

III. Conclusion

Public-key encryption is frequently performed using the RSA cryptosystem, which depends on the difficulty of factoring huge integers. Alternative cryptographic techniques are required because of the security danger posed by the emergence of quantum computers. One such option is the McEliece cryptosystem, which makes use of the computational difficulty of decoding random linear codes, an NP-complete problem from algebraic coding theory.Goppa codes, which are used in the original McEliece cryptosystem, are extremely safe but are impractical for everyday use because they call for enormous public keys. Due to the structured design of Reed-Solomon codes, which enable smaller public keys while retaining powerful error-correcting capabilities, they have been suggested as an alternative. Polynomial interpolation is used to decode Reed-Solomon codes, which are created by evaluating polynomials over particular finite fields. They are capable of correcting mistakes up to (d-1)/2, where (d) is the minimum distance, and (n,k) is the code length and dimension, respectively.However, a specific attack targeting the McEliece cryptosystem's Reed-Solomon code structure has exposed security flaws, making direct usage of these codes hazardous.This emphasises the necessity for additional study to examine the security of the McEliece cryptosystem using different codes or better Reed-Solomon code mascots. To assure the future development of reliable and secure cryptographic systems, this endeavour is essential.

References

[1] P.S. Menon and M. Ritwik. “A Comprehensive but not Complicated Survey on Quantum Computing”. In: IERI Procedia 10 (144-52) (2014).

[2] D’Anvers J P, Guo Q, Johansson T, et al. Decryption failure attacks on ind-cca secure lattice-based schemes. In: Proceedings of IACR International Workshop on Public Key Cryptography, 2019. 565–598

[3] D’Anvers J P, Rossi M, Virdia F. (One) failure is not an option: bootstrapping the search for failures in lattice-based encryption schemes. In: Proceedings of Annual International Conference on the Theory and Applications of Cryptographic Techniques, 2020. 3–33

[4] Guo Q, Johansson T, Yang J. A novel CCA attack using decryption errors against LAC. In: Proceedings of International Conference on the Theory and Application of Cryptology and Information Security, 2019. 82–111

[5] Fritzmann T, Pöppelmann T, Sepulveda J. Analysis of error-correcting codes for lattice-based key exchange. In: Proceedings of International Conference on Selected Areas in Cryptography, 2019. 369–390

[6] D'Anvers J P, Vercauteren F, Verbauwhede I. The impact of error dependencies on ring/mod-LWE/LWR based schemes. In: Proceedings of International Conference on Post-Quantum Cryptography, 2019. 103–115

[7] Gentry C, Peikert C, Vaikuntanathan V. Trapdoors for hard lattices and new cryptographic constructions. In: Proceedings of the 40th Annual ACM Symposium on Theory of Computing, 2008. 197–206

[8] Micciancio D, Peikert C. Trapdoors for lattices: simpler, tighter, faster, smaller. In: Proceedings of Annual International Conference on the Theory and Applications of Cryptographic Techniques, 2012. 700–718

[9] Wang Z, Ling C. On the geometric ergodicity of metropolis-hastings algorithms for lattice gaussian sampling. IEEE Trans Inform Theor, 2018, 64: 738–751

[10] Wang Z, Ling C. Lattice Gaussian sampling by Markov Chain Monte Carlo: bounded distance decoding and trapdoor sampling. IEEE Trans Inform Theor, 2019, 65: 3630–3645

[11] Lyubashevsky V. Fiat-Shamir with aborts: applications to lattice and factoring-based signatures. In: Proceedings of International Conference on the Theory and Application of Cryptology and Information Security, 2009. 598–616

[12] Lyubashevsky V. Lattice signatures without trapdoors. In: Proceedings of Annual International Conference on the Theory and Applications of Cryptographic Techniques,2012. 738–755

[13] Ducas L, Durmus A, Lepoint T, et al. Lattice signatures and bimodal gaussians. In: Proceedings of Annual Cryptology Conference, 2013. 40–56

[14] Peikert C. An efficient and parallel Gaussian sampler for lattices. In: Proceedings of Annual Cryptology Conference, 2010. 80–97

[15] Micciancio D, Walter M. Gaussian sampling over the integers: efficient, generic, constant-time. In: Proceedings of Annual International Cryptology Conference, 2017. 455–485

[16] Zhao R K, Steinfeld R, Sakzad A. FACCT: fast, compact, and constant-time discrete Gaussian sampler over integers. IEEE Trans Comput, 2020, 69: 126–137

[17] Misoczki R, Tillich J P, Sendrier N, et al. MDPC-mceliece: new mceliece variants from moderate density parity-check codes. In: Proceedings of IEEE International Symposium on Information Theory, 2013. 2069–2073