Analysis of code-based digital signature schemes

Rupali Khurana, Ekta Narwal
Department of Mathematics, Faculty of Physical Sciences, Maharshi Dayanand University, Rohtak, India

ABSTRACT
Digital signatures are in high demand because they allow authentication and non-repudiation. Existing digital signature systems, such as digital signature algorithm (DSA), elliptic curve digital signature algorithm (ECDSA), and others, are based on number theory problems such as discrete logarithmic problems and integer factorization problems. These recently used digital signatures are not secure with quantum computers. To protect against quantum computer attacks, many researchers propose digital signature schemes based on error-correcting codes such as linear, Goppa, polar, and so on. We studied 16 distinct papers based on various error-correcting codes and analyzed their various features such as signing and verification efficiency, signature size, public key size, and security against multiple attacks.

This is an open access article under the CC BY-SA license.

1. INTRODUCTION
In modern times, we use the internet to transfer data from one point to another. For secure transmission of data, we use cryptography [1]. Cryptography is a technique of protecting useful information by converting it into an unreadable format [2]. Many cryptographic algorithms are present that provide secure transmission via the insecure network [3]. But it has been forecasted that quantum computers [4] can break many classical cryptographic algorithms. Now, to protect these classical cryptographic algorithms from quantum computers, several algorithms are made that are also secure in the quantum era and are named post-quantum cryptography (PQC) [5]. Post-quantum cryptography mainly includes hash-based cryptography [5], lattice-based cryptography [6], and code-based cryptography [7].

In this paper, we will discuss only code-based cryptography. It includes all the cryptosystems in which security depends upon the hardness of encoding and decoding of the error-correcting codes like Goppa codes [8], Bose-Chaudhuri-Hocquenghem (BCH) codes [9], and Reed-Solomon codes [10]. In 1978, McEliece introduced the first code-based public-key cryptosystem [11]. This cryptosystem is based on a random binary Goppa code. In 1986, Harald Niederreiter proposed a variation of the McEliece cryptosystem [11], which uses a parity check matrix in place of the generator matrix used by McEliece. After that Stern identification scheme [12], Courtois-Finiasz-Sendrier (CFS) signature scheme [13], and many more came into existence.

There is a high chance of data being forged or stolen by the theft during transmission. Due to this fact, various digital signatures [14] are used during transmission to ensure that the data or message comes from the authorized entity and remains unchanged [15]. Digital signature [16] is a type of electronic signature in which the integrity and authentication of a message are verified by using mathematical algorithms. Recently used digital signatures like digital signature algorithm (DSA) are not secure against
quantum attacks. So, post-quantum digital signatures are very much in demand. There are many signature schemes present like hash-based signatures [17], lattice-based signatures [18], and code-based signatures [19]. Here, we focus only on code-based digital signatures. Code-based digital signature [19] is a type of digital signature that depends on the hardness of decoding the error-correcting codes like Goppa codes [8], linear codes [20], polar codes [21], and cyclic codes [22]. The major problem in code-based digital signatures is the large size of the public key and slow signature generation speed. In this paper, we will discuss some code-based signatures and identification schemes and find the scheme that will resolve this problem.

Here, we will discuss some basic definitions and some present threats against the signature schemes in section 2. Section 3 will discuss a few existing code-based digital signature schemes. We will analyze various properties of code-based digital signature schemes like signing and verification efficiency, signature size, public key size, and security against multiple attacks in section 4. Section 5 will explain the future plan. Lastly, we will conclude in section 6.

2. PRELIMINARIES

In coding theory, error-correcting codes are used to detect and correct errors. They are also used in post-quantum primitives in order to resist attacks against quantum computers. Here are some basic definitions and problems related to coding theory.

2.1. Linear codes

A \((n, k)\) linear code \(C\) over a field with \(q\) elements is a linear subspace of dimension \(k\) and length \(n\) of the linear space \(F_q^n\) [20]. A matrix \(G\) of order \(k \times n\) is a generator matrix for an \((n, k)\) linear code \(C\) if the rows of \(G\) form a basis of \(C\) over the field \(F_q\) [23]. A matrix \(H\) of order \((n - k) \times n\) is said to be the parity-check matrix [23] for an \((n, k)\) linear code \(C\) over the field having \(q\) elements if the rows of \(H\) form a basis for the orthogonal complement of \(C\), and it satisfies \(C' = \{y \in F_q^n ; Hy^T = 0\} \).

2.2. Difficult problems in coding theory

Here, we discuss syndrome decoding problem and Goppa code distinguishing problems. In syndrome decoding problem, we choose a matrix \(H\) of order \(m \times n\), and \(p\) be a target vector such that \(p \in F_q^m\) over a field \(F_q\), and an integer \(l \geq 0\). We find a vector \(y \in F_q^n\) having weight \(\leq l\) satisfying \(Hy^T = p^T\) [24]. In Goppa code distinguishing problem, we choose a matrix \(H\) of order \((n - k) \times n\) randomly over a field \(F_q\). Our goal is to decide whether \(H\) is a parity check matrix for a Goppa code \((n, k)\) or a \((n, k)\) random code parity check matrix [25].

2.3. Threats against code-based digital signatures

Digital signatures [26] are used very commonly in the digital world, and their main goal is to provide integrity and authentication [27]. But digital signatures are not secure and can be attacked by the attacker. Here are a few threats or attacks against digital signatures:

In a key recovery attack, the attacker attempts to recover the private/secret key with the help of the signer’s public key. After recovering the private/secret key, he can easily forge the digital signature [28]. A forgery attack is an attack in which the attacker attempts to create a valid signature for a document without the knowledge of the signer’s private key. If he creates a valid signature, he can forge the signature [28]. In chosen message attack, the attacker somehow makes the signer to sign one or more messages. Now, the attacker has some messages and signature pairs, and with the help of these pairs, the attacker analyzes the signature and tries to re-create it [29]. The attacker in a known-message attack has few messages and signature pairs. With the help of these signatures’ pairs, the attacker analyzes the messages and signature pairs. After that, he can easily recover the signer’s private key [30], and with the help of the private key, he can forge the signature. In a key substitution attack, the attacker has a public key and signature on a message ‘\(d\)’. Now, the attacker produces a different public key that validates the same signature on the same message ‘\(d\)’, which affects the authentication of a message [31].

Table 1 shows different attacks present on various code-based digital signature schemes. Key recovery attack against the rank quasi code-based signature. Forgery attack against CFS and identity-based signature. Chosen message attack against stern's identification and signature scheme. Known message attack against Kabastianskii-Krouk-Smeets (KKS) signature. Key substitution attack against CFS signature.

Analysis of code-based digital signature schemes (Rupali Khurana)
3. CODE-BASED DIGITAL SIGNATURE SCHEMES

Firstly, Xinmei [32] proposed a digital signature scheme, which was proven to be insecure [33]. After that, many digital signature schemes are designed which depend upon various error-correcting codes. Some of them are discussed here.

Initially, Stern [12] proposed an identification scheme that depends on the syndrome decoding problem for the error-correcting codes. Here, the author describes a basic-zero knowledge protocol that enables any prover $P$ to identify himself to any other verifier $V$. The probability of attacking the scheme is $(2/3)$. Using the Fiat-Shamir method, it is possible to convert Stern's identification scheme into a signature scheme [34].

Then, Kabatianskii et al. [35] proposed a digital signature scheme that depends on random linear error-correcting codes. The authors presented three different forms of the KKS scheme, and one modified form that helps to construct signatures from the codes containing low weight codewords. All the forms were based on different linear codes to enhance the scheme and resist some attacks. Also, the authors claimed that all the signature schemes were secure if the public parameters did not give any information. Cayrel et al. [36] proved that the attacker only needs a maximum of 20 signatures to break the KKS signature scheme. The authors gave new parameters to get a security of 40 signatures to resist this attack.

Later on, Courtois et al. [13] proposed the first practical code-based digital signature scheme depends on the McEliece cryptosystem. He made a signature using the Niederreiter cryptosystem. He used binary Goppa codes to develop the signature scheme. For any given integer $m$ and $t$, Goppa codes are of length $n = 2^m$ and of dimension $k = n – mt$. These codes can correct a maximum of $t$ errors. In this paper, the authors chose the parameters $m = 16$ and $t = 9$ and proposed a digital signature scheme of about 81 bits. This scheme depends upon a well-known syndrome decoding problem.

Zheng et al. [37] then presented the first code-based ring signature scheme by extending the CFS scheme. This practical ring signature was based on the syndrome decoding problem. Each signer uses $t$ error-correcting Goppa codes having length $n = 2^m$ and dimension $k = n – mt$, for some integers $m$ and $t$. In this paper, the authors chose the parameters $m = 16$ and $t = 9$, and made a ring signature of length about 144 + 126$l$, where $l$ is the number of ring members. The authors also showed that the probability of forging the signature was about $(1/2^n)$.

Following that, Melchor et al. [38] presented the first code-based threshold ring signature scheme. In this paper, the author generalizes stern's scheme into the $t$-out of $N$ threshold ring signature scheme with the help of the Fiat-Shamir paradigm. The proposed signature did not depend on the number of signers, i.e., $t$ and the signature size depends on the number of signers, i.e., $N$ in a ring. The signature length was $N$ times of stern's identification scheme, i.e., $20k_0 \times N$. Its security depends upon the syndrome decoding problem. The proposed signature has the cheating probability $(2/3)$.

Then, Dallot and Vergnaud [39] proposed the second code-based threshold ring signature scheme combining the technique of Bresson and the CFS scheme. The proposed scheme used an $(n, k, t)$ error correcting Goppa codes with $n = 2^m$ and dimension $k = n – mt$, for positive integer $m$ and $t$. Cayrel et al. [40] then proposed an improved identity-based identification scheme based on error-correcting codes. To develop this scheme, the author combines the modified Courtois-Finiasz-Sendrier (mCFS) signature scheme with the stern's identification scheme, in which security depends on the syndrome decoding problem. The authors used Goppa codes of length $n = 2^m$, and of dimension $k = n – mt$. They choose parameters $(16, 9)$ and produces a signature length of about $2^m \times (no. of rounds)$. For the first round, the signature length is approx. $1.1$ MB.

Alamelou et al. [41] presented the first code-based group signature scheme. The proposed group signature was obtained from the stern's identification protocol using the Fiat-Shamir paradigm. The idea behind the signature scheme was building a collision of two syndromes which was associated with two different matrices, one is the random matrix, and another is the trapdoor matrix. The security of the proposed scheme was based on the relaxation of the Bellare-Shi-Zhang (BSZ) model. The signature provides several properties such as anonymity, traceability [42], and non-frameability [41].
After that, Ren et al. [43] proposed an efficient code-based digital signature algorithm with the help of code-based hash function using the mCFS scheme. Here, the authors used the code-based hash function in place of the random hash function to better the signature’s efficiency. The signing process of the signature was improved by reducing the signing time by \( t! \) which increases the signing speed. The authors used Goppa codes of length \( n = 2^m \) and of dimension \( k = n - mt \).

Later on, Liu et al. [44] proposed a secure signature scheme using the improved version of a McElieic public key cryptosystem (PKC). The proposed scheme depends on the idea of the CFS scheme. The authors used binary \((n, k, d)\) Goppa codes whose length and dimension are \( n = 2^m \) and \( k = n - mt \), and distance \( d = 2t + 1 \). The authors gave the probability of signing a message was \((1/t!^2)\). Also, the authors claimed that a smaller \( t \) could be chosen in the presented signature, i.e., \( 1, 2, 3 \). With the help of this, the signer needs some attempts for signing a message which increases the speed of signing. The proposed scheme has properties like fast signing speed, high security, and strong practicability.

Then, Sahu and Tripathi [45] proposed a signature scheme with the help of modified quasi-cyclic low density parity check (QC-LDPC) codes in place of Goppa codes used in the CFS scheme. Here, the authors used a belief propagation (BP) decoding scheme that increases the decoding speed. The security and efficiency of the CFS scheme were also improved by using the BP decoding scheme. The proposed scheme is fast, secure, and has small public key and signature size with high security. He chooses the parameters \((16384, 12344)\) which reduces the key size to 6,140 bits.

Lee et al. [46] proposed a signature scheme that depends on modified reed-muller (RM) codes that reduce signing complexity. Here, the authors used \((U, U + V)\) codes with a high-dimensional hull in order to overcome the drawbacks of different code-based schemes. The presented signature scheme has a smaller key size and low signature time. This scheme also resists various known attacks like key substitution attacks. For classic security of 128 bits, the signature size is about 4096 bits, with a public key size less than 1 MB.

Forghani et al. [47] then proposed a digital signature scheme based on polar codes. This paper used polar codes [48] with the CFS scheme that reduces the public key size and signing time. The authors also proved that using polar codes in the CFS signature helps in improving security against forgery and key recovery attacks. The authors also showed that the proposed scheme is secure in random oracle model [49].

Hooshmand et al. [50] proposed two polar code-based identification schemes, Id-PC I, and Id-PC II, in which he replaces polar codes with random codes. The security of these schemes depends on the hardness of the syndrome decoding problem and the general decoding problem. As compared to the Stern and Veron scheme, the author reduces the size of public data by 90% and communication costs by 53%. The authors also showed in this paper that the proposed schemes Id-PC I and Id-PC II have a low cheating probability, i.e., \((2/3)^r\), where the protocol repeats ‘r’ times, and are secure against information set decoding attacks.

Then, Cho et al. [51] proposed enhanced pqsigRM, a signature scheme based on modified reed-muller codes. Here, the author replaces the Goppa codes in the CFS scheme with modified reed-muller codes. This scheme has several advantages, including a small signature size and fast verification speed. The author also showed that the proposed scheme is resistant to a variety of attacks based on the Reed-Muller code-based cryptosystem. For the security of 128 bits, the size of the proposed signature is 512 bytes. To improve its performance, the author modifies the public code used in pqsigRM.

4. SUMMARY

This section will summarize the different code-based digital signature schemes and analyze their properties. Table 2 shows the survey of the different code-based digital signature schemes based on different error-correcting codes, signature size, security proof. We have also considered their properties like key size, signing time, and security against multiple attacks. From this table, we observe that the CFS scheme has the smallest signature size among various studied signature schemes.

Table 3 gives the public key sizes used in various digital signature schemes concerning their parameters and security. Among various schemes, we see that PolarSig has the smallest size of public key. This makes PolarSig a more practical and efficient digital signature scheme.

Figure 1 shows the size of a public key in Kilobytes for various code-based signature schemes. Dallot’s signature scheme has the greatest public key size among all signature schemes, whereas PolarSig has the smallest public key size among all the signature schemes. This property makes the digital signature PolarSig more efficient.
Table 2. Different code-based digital signature schemes

<table>
<thead>
<tr>
<th>Signature scheme</th>
<th>Security against model</th>
<th>Signature size</th>
<th>Codes</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forghani et al. [47]</td>
<td>PolarSig</td>
<td>Random Oracle model</td>
<td>808 bits</td>
<td>Polar codes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Smaller key size, low signing time,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>secure against forgery and key recovery attacks</td>
</tr>
<tr>
<td>Liu et al. [44]</td>
<td>A secure and efficient signature scheme</td>
<td>Random Oracle model</td>
<td>81 bits</td>
<td>Goppa codes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fast signing speed, high security,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>strong practicability, and authentication</td>
</tr>
<tr>
<td>Ren et al. [43]</td>
<td>An efficient code-based signature scheme</td>
<td>Random Oracle model</td>
<td>144 bits</td>
<td>Goppa codes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fast signing speed, low signing time,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>and more practicability</td>
</tr>
<tr>
<td>Alaméou et al. [41]</td>
<td>Group signature</td>
<td>Random Oracle model</td>
<td>$\sqrt{N^2}$</td>
<td>Linear codes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Anonymous, traceable, non-frameable,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>and authentication</td>
</tr>
<tr>
<td>Cayrel et al. [40]</td>
<td>Improved identity-based identification and signature scheme</td>
<td>Random Oracle model</td>
<td>1.1 MB</td>
<td>Goppa codes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unforgeable and authentication</td>
</tr>
<tr>
<td>Melchot et al. [38]</td>
<td>Threshold ring signature scheme</td>
<td>Random Oracle model</td>
<td>$20k_0 \cdot N$</td>
<td>Linear codes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unforgeable, anonymous, and authentication</td>
</tr>
<tr>
<td>Zheng et al. [37]</td>
<td>Ring signature scheme</td>
<td>Random Oracle model</td>
<td>144+126l</td>
<td>Goppa codes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unforgeable, anonymous, practicability, and</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>authentication</td>
</tr>
<tr>
<td>Courtois et al. [13]</td>
<td>CFS signature scheme</td>
<td>Random Oracle model</td>
<td>81 bits</td>
<td>Goppa codes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Practicability, short signature, fast</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>verification, and authentication</td>
</tr>
<tr>
<td>Kabatianski et al. [35]</td>
<td>KKS signature scheme</td>
<td>Random Oracle model</td>
<td>16,976 bits</td>
<td>Linear codes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Authentication</td>
</tr>
<tr>
<td>Stern [12]</td>
<td>Stern identification and signature scheme</td>
<td>Random Oracle model</td>
<td>120 Kbits</td>
<td>Linear codes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unforgeable, simplicity</td>
</tr>
</tbody>
</table>

Table 3. Public key sizes of different code-based digital signature schemes

<table>
<thead>
<tr>
<th>Digital Signature Scheme</th>
<th>Public Key Size</th>
<th>Security (in bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KKS ($N, K, n, k, t_0, t_1$)</td>
<td>9,437,184 bits</td>
<td>83</td>
</tr>
<tr>
<td>Stern ($2n, n$) ($694, 347$)</td>
<td>347 bits</td>
<td>83</td>
</tr>
<tr>
<td>PolarSig ($N', K'$)</td>
<td>50.2 KB</td>
<td>128</td>
</tr>
<tr>
<td>Ring signature scheme ($m, t$)</td>
<td>0.7 MB</td>
<td>80</td>
</tr>
<tr>
<td>Modified pqsigRM ($r, m$)</td>
<td>0.773 MB</td>
<td>128</td>
</tr>
<tr>
<td>Parallel CFS ($m, t$)</td>
<td>1.2 MB</td>
<td>80</td>
</tr>
<tr>
<td>Enhanced pqsigRM ($r, m$)</td>
<td>1.2 MB</td>
<td>80</td>
</tr>
<tr>
<td>Id-PC identification scheme</td>
<td>13.7 MB</td>
<td>80</td>
</tr>
<tr>
<td>Id-PC I ($2k, k$)</td>
<td>2.41 MB</td>
<td>80</td>
</tr>
<tr>
<td>Id-PC II ($2k, k$)</td>
<td>70 MB</td>
<td>80</td>
</tr>
</tbody>
</table>

Figure 1. Comparison of public key sizes of various schemes
5. FUTURE PLAN

Polar codes have recently been used in 5G communications, making them a better choice among several error-correcting codes. Using polar codes in both the signature PolarSig and the identification scheme Id-PC gives them several advantages, such as reduced signature and key size, faster signing speed, and resistance to various attacks. These qualities make them the best of all schemes studied. In the future, we will create a digital signature scheme based on polar codes.

6. CONCLUSION

We presented various code-based digital signature schemes based on different error-correcting codes. We examined properties such as signature and public key size, signing and verification efficiency, and their resistance to various attacks. Digital signature PolarSig and the identification scheme Id-PC schemes based on polar codes are the best among the various studied digital signature and identification schemes because they have properties such as smaller key size and low signing time. These schemes are also resistant to a variety of attacks like key recovery attacks and forgery attacks.

REFERENCES


BIOGRAPHIES OF AUTHORS

Rupali Khurana did her M.Sc. (5-year Integrated M.Sc. Hons. Mathematics) from the Department of Mathematics, Maharshi Dayanand University, Rohtak. She is currently pursuing a Ph.D. Degree in Mathematics from the Department of Mathematics, Maharshi Dayanand University, Rohtak. Her area of research is coding theory and cryptography. She can be contacted at email: rupali.rs.maths@mdurohtak.ac.in.

Ekta Narwal did her B.Sc. (Computer Science, Mathematics, Physics), M.C.A., Ph.D. in Computer Science and Applications. She is working as an Assistant Professor in the Department of Mathematics, at Maharshi Dayanand University, Rohtak (Haryana) India since 2012. She worked as Guest Lecturer, in Computer Science in the Department of Mathematics, M. D. University Rohtak from January 2011 to March 1, 2012. Her major research areas are coding theory, cryptography, network security, and artificial neural networks. She has research experience of 8 years and teaching experience of nearly 11 years. She has published 11 research papers. She can be contacted at email: ektnarwal.math@mdurohtak.ac.in.