Workingtitle
Postquantumcryptography

Jannis Priesnitz
University of Applied Sciences Darmstadt
Department of Computer Science
Schöfferstraße 3
64295 Darmstadt
Email: jannis.priesnitz@stud.h-da.de

***Abstract*—ABSTRACT**
***Index Terms*—Message Oriented Middleware, IoT**

I. Introduction

State of the art asymmetric cryptosystems compared with usage of sufficient long Keys are considered as ”safe” on current computer systems. This changes immediately when quantum computers enter the scene. But why is this the case? In this paper I briefly present the main differences between normal computer systems and quantum computers and why state of the art cryptosystems like RSA and ECC are not save in the quantum computation world. In addition to that I, present a extensive introduction to the McEliece Cryptosystem explain the strengths and downsides of the algorithm and give some details about the developments of codes which can be used for the system.

Since decades computer systems have been seen as digital circuits based on the rules of physics. A bit in this systems is seen as a value of voltage. If this value is above a certain, level the binary representation is one, otherwise it’s zero. On quantum computers, information is handled in binary, too. The main difference between the bit representation and the quantum bit representation (so called qubit) is that in addition to the two states zero and one, there are several more states which one qubit can reach due to a superposition. This means that a qubit is not either in state one or in state zero but in theory can have an arbitrary number of different states. All these states can appear with an certain probability. Each of these states is able to compute one possibility of a NP hard problem.

Due to this fact, quantum computers are able to solve problems which are NP hard in sense of complexity theory much faster than traditional computers.

One of these computational problems, which are traditionally hard to solve, is the prime factorization problem which is in serveral variations the core of many public key algorithms such as the RSA cryptosystem. The second main category, the elliptic curves cryptography is even more affected by this problem as the logarithm of a finite field of an elliptic curve can be computed efficiently and can be broken in less time compared to the RSA algorithm because of lower keysizes. In conclusion, there is no established asymmetric cryptoscheme which is post quantum resistant. Therefore completely different approaches are needed.

Outline

In section II, some background information regarding post quantum cryptography is given. With help of Shors algorithm the principle of quantum cryptography is described more precisely. section III takes a detailed look into the McEliece crypto system followed by the Niederreiter cryptosystem in .

make label

II. Post quantum cryptography

What makes cryptosystems strong?

In order to reach post quantum proof cryptographic algorithms, some fundamentally different then the established ones are required. Shors Algorithm in II-A reveals that cryptography based on integer factorization or the discrete logarithm problem is no longer an issue for quantum computers. Therefore other algorithm types are needed which are shortly described in II-B.

A. Shors Algorithm

Peter Shor presented in 1994 an algorithm which is able to factorize a composite number n into its prime factors. This algorithm is especially designed for quantum computers on which it only needs $logn$ qubits and has a runtime of $O((logn)^3)$ for finding a non trivial factor of n.

review section – more abstract

1) Procedure Shors algorithm divides into a classical part which can be executed on a conventional computer and a quantum part which has to be executed on a quantum computer in order tp perform the computation efficiently.

The basic idea is that the classical part reduces the problem while the quantum part finds the order of the group in which n is.

a) Classical part The classical part of the algorithm mainly contains computation of the greatest common devisor of a randomly selected number lower than n and n itself. Now we need to compute the order r of x which is where the quantum part emerges. The classical part is executed in a loop while the order r is odd or x to the power of $r/2$ is equivalent to $-1modn$ If this is not the case the gcd of $x^(r/2) - 1$ n is computed.

why?

make formula beau-

b) Quantum part At first there a q is determined which is a power of 2 and lies between $n2$ and $2n2$. A random a which is lower than n is selected and the input quantum register is initialized with all states of $amodq$. The output quantum register is initialized with all states of xa(mod n).

formel

A quantum Fourier transformation is computed on the input register.

The result values are gathered from the input register.

vllt. doch abstrakter

To sum up Shors algorithm with support of a quantum computer is able to find the period of a prime in polynomial time.

With this algorithm all cryptography based on prime factorization can be broken by a quantum computer in polynomial time.

2) Complexity of Shors Algorithm on normal pc and on quantum pc

Considering the complexity of Shors algorithm on quantum computers, it's easy to see that there is a complexity of $O(logn)$ which is in the class of BQP. This class is comparable to the class BPP on conventional computers. In this regard the an state of the art RSA key of 2048 bit length can be broken in xxxx. In the case of elliptic curves cryptosystem it's even worse. Due to the principal of the algorithm ECC is gathering more security per bit keylength on a conventional computer and only has keys up to 512 bit. This is not the case on quantum computer which makes them even easier to break compared to RSA keys.

kann solch eine Aussage treffen?

BQP / BPP mit reinnehmen??? Könnte interessant sein. Unter oder neben Shor?

B. Candidates for post quantum cryptography

In this section a short overview over promising state of the art post quantum cryptoschemes based on [1] is given.

Lattice-based cryptography One of the most studied types of algorithm is the lattice based cryptography which exists in several variants. Algorithm works on a lattice over a n-dimensional finite Euclidian field L with a strong periodicity property. A set of vectors provides the basis of L in the way that every element is uniquely represented. The cryptographic problem is to find the closest vector to a given lattice point e.g. by adding an error vector[1][2].

Multivariate cryptography Multivariate cryptography is based on a multivariate polynomials over a finite field F which are defined over both a ground and an extension field. In the case of solving systems they are NP-complete and due to this fact a candidate for post quantum cryptography. They are topic of studies for a long time and are promising especially for signature schemes[1][3].

Hash-based cryptography Hash-based algorithms such as Lamport-[4] and the Merkle[5] signature scheme are based on strong hash functions but have the drawback that only a limited count signatures can be created per key. The algorithm reduces the one time signature to an hash value using a hash function[1].

Code-based cryptography The forth group, the code based algorithms, are based on error-correcting codes. First investigations have been developed by Robert McEliece using random Goppa codes[6]. This paper deals with the properies of McEliece- and the related Niederreiter cryptosystem[1][7].

The raw overview of some state of the art algorithms shows different aproaches but in general it could be said that a much higher computational effort has to be taken to achieve a strong system compared to the traditional ones.

III. The McEliece cryptosystem

Back in 1978, Robert McEliece suggested an asymmetric quantum resistant cryptosystem based on the theory of algebraic codes. He selected binary Goppa codes with the property irreducibility as base a for the cryptosystem[8]. The chosen code C has a length of $n = 2^m$ and a dimension of $k >= n - tm$. These codes are able to correct any pattern of t or fewer errors. For each of this codes, there exists an irreducible polynom of degree t over $GF(2^m)$. The main reason for McEliece to select this setup is that there exists an fast algorithm to decode these codes[9].

Key generation

For the key-generation, a n and t with above mentioned properties are picked. Additionally, an irreducible polynomial of degree t over $GF(2^m)$ is selected randomly. The probability that this selection leads to an irreducible polynomial is $1/t$ and there is an efficient algorithm to prove this[10]. In the next step a generator matrix G which is of size $n \times k$ is produced. This can be transformed into canonical form.

Now the information of G has to be camouflaged. Therefore a random dense $k \times k$ matrix S which is nonsingular and a random $n \times n$ permutation matrix P is selected. Both of them are multiplied to $G' = SGP$. Due to the matrix multiplication properties, the linear code generated by G' has the same rate and distance like G. G' is the public generator matrix and is sent to the encrypting entity.

The following encryption algorithm is published so that the encrypting entity can use it.

Encryption

First of all, the message m which is to be encrypted has to be devided into k-bit blocks. The public key encryption is performed by $x = uG' + z$ with u beeing one of such a k-bit block. In this case, z is a randomly generated vector with length n and weight t [1].

x is the encrypted message which is transmitted to the private key owner who can decrypt the message block u as follows.

[1]The weight of an vector is defined as Hamming weight.

Decryption

The decryption of one block x starts with computing $x' = xP^{-1}$ with P^{-1} as inverse of the permutation matrix P.

With an error correcting algorithm for the code C, the codeword u' next to x', is calculated. To get a plaintext message block the computation $u = u'S^{-1}$ is performed[11][8]. As an efficient method for calculating the error corrections of x' to u', McEliece suggests the algorithm of Patterson[12]. Refer also to **??**.

Correctness

Assuming that P is a permutation matrix and randomvector z with length n and weight t, it is obvious that zP^{-1} has a weight of t or less. As discussed, the computation is $c' = cP - 1 = uG'P^{-1} + zP^{-1} = uSG + zP^{-1}$. The chosen Goppa code C is designed to correct up to t errors. On the other hand, mSG has a maximum distance from cP^{-1} of t. This leads to the fact that the correct code mS is determined by the algorithm. To obtain the message block u from uS, we can easily multiply the inverse $u = uSS^{-1}$[11].

Security properties

The security of the presented scheme refers on the one hand to the basics of learning with errors principle. More precise it is hypothesis of *Learning Parity with Noise*[13]. On the other hand it refers to the hypothesis that the generator matrix G is indistinguishable from any other $k \times n$-matrix. This leads to the property of a trapdoor function.

auswerten –¿ paper

IV. The Niederreiter cryptosystem in comparison to McEliece

The Niederreiter cryptosystem is highly comparable to the McEliece cryptosystem due to the fact that it is following the same basic idea.

Niederreiter designed his $(n, k, 2t+1)$ linear code C over a Galois field too. In contrast to McEliece the code size does not have to be a power of 2 instead of an arbitary integer $GF(q)$.

Another difference to McEliece is the usage of a $(n-k) \times (n)$ parity check matrix H instead of the generator matrix G. The nonsigular $(n-k) \times (n-k)$ matrix M is defined sightly different from McEliece $(k \times k)$.The permutation matrix P, an arbitrary $n \times n$ matrix, is exactly defined in the same way compared to McEliece. The private key is then defined as $M, HandP$ the public key consists of $H' = MHP$ and the hamming weight t which is quite the same compared to McEliece. The messages in the system of Niederreiter have to be n dimensional vectors over $GF(q)$ and they must have a hamming weight of t. This is an important fact issuing the signature creation in section V. Encryption with Niederreiter is performed with $z = yH'^T$ which again is comparable with the encryption operation $x = uG' + z$ in McEliece. The error vector z is not needed here because this is already represented in ...

why?

why?

why is it not needed to add the error vector here?

. The ciphertext has only n - k bit dimension compared to n bit in McEliece. The decryption is computed firstly with $(yP^T)H^T = z(M^T)^(- 1)$ Then H is eliminated by an error correction algorithm which leads to $(yP)^T$ which can easily computed to the plaintext y. [14][15][16]

V. Signing with Niederreiter

Besides en- and decryption signatures building and verification is an common requirement to an asymmetric cryptoscheme. In state of the art algorithms principal is quite simple: The message to be signed is *de*crypted with a given public key. The verifier *en*crypts the message with his private key an compares the result with the message.

In case of McEliece this isn't so easy because it is not possible to decrypt (= sign) an message before encrypting (= verify) it. More precise the process of signing produces a syndrom whose error pattern is bigger than the error correcting property t. In fact it is hard to create a ciphertext that fits to the error correcting properties of the encryption without using it.

dessen

Compared to ecnryption and decryption siging and verifying is much harder to realise with McEliece. Just in 2006 a digital signature scheme were presented by Courtois et. al. The problem with signing a given hash value n is that generally it is longer than the decoding capacity t of the used code. More general one can say that it is difficult to generate a random ciphertext without using the encyption algorithm. [17]

Complete decoding One possible solution would be to use complete decoding. Therefore not only the words within the radius of t can be decoded but all words laying in the code space. In other words with complete decoding we can find an erro pattern to any given syndrome as long as it is in the code space. This means that we have to add a δ with random collumns from the parity check matrixto t. The decoding works exactly when all of the δ-columns fit to an error position because then the syndrome will fit to an word of weight t. Else we have to add another δ to t and try again.

why pc-matrix

From this properties we now can construct a digital signature scheme: We have to select a δ which is small enough to get an usable key size but on the other hand has a good security.

For achieving a small δ the code has to be selected carefully in the way that it has to have a high density of decodable syndromes. This makes sure that the δ is kept small because the probability of finding a fitting one is high. For building up a signature the signer now takes a syndrome and hashes it together with the document. This is tried with an modified document (possibly with some kind of padding) as long as he gets a decodable syndrome. [17]

VI. Codes can be used for McEliece and Niederreiter

find parameter for this.

As described in the last sections the McEliece cryptosystem is based on the properties error correcting codes. In principle many codes with a good decoding algorithm could be used for the cryptosystem. However there are some properties which leads to higher security than others. In this section a

both the same?

short introduction to the basics of error correcting codes is presented followed by a deeper view on Goppa codes which are suggested by McEliece and are still the best choice.

what makes a code usable

A. Principles of error correcting codes

As the name states the basic purpose of error correcting code is to detect and correct errors which occur during the transition of messages. Therefore additional information is added to the message by the sender. Knowing the encoding algorithm receiver is able to decode the original word, detect errors and correct them with help of the additional information and the corresponding decode algorithm.

Binary Due to the fact that almost all transition is performed by computers and is in binary the explanations in following are focussed on the binary case too. Never the less all of them hold in the general case.

For generating these additional information a vector with binary indexes $V = [c1c2...cn]|ci \in 0, 1$ is created as set of $\mathbb{F}_2$. Further the addition is defined componend-wise modulo 2, which also can be seen as XOR operation. With this definition V covers all n-tuples in $\mathbb{F}_2$. Having this, a binary, linear $[n, k]$ code can be defined as subspace of $\mathbb{F}_2$ in which the n is the length and the k is the dimension. A codeword is represented as vector in the code.

Hamming distance Another important variable for error correcting codes is the hamming weight or more precise the minimum weight d of a vector. The weight is simply defined as count of non-zero positions in the vector. Furthermore the minimum weight is defined as minimum is the smallest weight of a codeword under the condition that it is not zero. A $[n, k, d]$ code can correct $t = \lfloor \frac{d-1}{2} \rfloor$ errors.

Goppa Codes

Goppa codes follow the principals described previously. In this section highlights the special basics (of Goppa codes) regarding the McEliece cryptosystem. Firstly some important basic properties and terminologies are presented. In the second part an example is discussed.

DO SO

Linearity Goppa Codes are linear codes. This is not a conincidence because non-linear codes do not provide an efficient decoding algorithm like Pattersons algorithm for linear codes does[12]. In [21] some remarkable results to achieve more efficiency are stated but also it is mentioned, that nonlinear codes are not able to achieve more comparable efficiency. However most of the popular codes are linear ones this condition does not rule out so many codes.

das könnte auch in ECC, da alle eccs durch polynomer erzeugt werden können und die irreduzibel sein müssen oder?

Irreducibility

To achieve a finite field with usable properties for a Goppa code the generator polynomial has to bew irreducible. This means that for a polynomial p over a finite field $GF(p^m)$ there exists no polynomial $p' \in GF(p^m)$ of lower degree which divides p.

Efficency(?) To keep the keyspace as efficient as possible the factor between the matrix dimensions n, k and the error correction range t (respectively the minimum hamming distance d) should be as high as possible. The most important point is that Goppa Codes are a suspect of research for years so the probability of is really low.

VII. Example of generating a Goppa Code and the McEliece cryptosytem

In III the threoretical the aspects of the McEliece cryptsystem were described and in VI-A the basics of generating a code were highlighted. In this section an example of of both is presented. Projecting these parameters onto the MCeliece cryptosystem ...

continue

ab hier eher ausblick

VIII. Variants of the McEliece cryptosystem

Due to the fact that the McEliece cryptosystem was published over 30 years ago and still is one of the most promissing post quantum security algorithms many variants came up. In this section two of the most important variants are presented.

A. Optimizing MECS

Wild MCE

In the previous sections we saw the importance of the McEliece crypto system and got an idea how the algorithm works. This sections focuses on the issues and their optimizations.

IX. Resistency against various attacker models

ausweite

a) Achieving chosen cipher text resistance Like the RSA cryptosystem in the original idea is not resistant against chosen cipher attacks.

This means if an attacker has access to an oracle which is able to decrypt an given ciphertext without knowing the key and the attacker is permitted to give all ciphertexts except the one he is asked to crack to the oracle he is not able to gather any reasonable information from the oracle. This goes over several iterations and is called CCA2-attack. In literature resistance against this attack is described as IND-CCA2.

Like the RSA cryptosystem uses padding standards such as PSS for signing and OAEP for encrypting to achieve IND-CCA2 the MECS has some padding schemes as well.

X. Conclusion

...

Quellen

[1] D. J. Bernstein, "Introduction to post-quantum cryptography," in *Post-quantum cryptography*, pp. 1–14, Springer, 2009.

[2] Wikipedia, "Lattice-based cryptography — wikipedia, the free encyclopedia," 2017. [Online; accessed 10-March-2017].

[3] Wikipedia, "Multivariate cryptography — wikipedia, the free encyclopedia," 2017. [Online; accessed 10-March-2017].

[4] Wikipedia, "Lamport signature — wikipedia, the free encyclopedia," 2017. [Online; accessed 10-March-2017].

[5] Wikipedia, "Merkle signature scheme — wikipedia, the free encyclopedia," 2017. [Online; accessed 10-March-2017].

[6] Wikipedia, "Binary goppa code — wikipedia, the free encyclopedia," 2015. [Online; accessed 10-March-2017].

[7] Wikipedia, "Niederreiter cryptosystem — wikipedia, the free encyclopedia," 2017. [Online; accessed 10-March-2017].

[8] R. J. McEliece, "A public-key cryptosystem based on algebraic," *Coding Thv*, vol. 4244, pp. 114–116, 1978.

[9] R. McEliece, *The theory of information and coding*. Cambridge University Press, 2002.

[10] E. R. Berlekamp, "Algebraic coding theory," 1968.

[11] Wikipedia, "Mceliece cryptosystem — wikipedia, the free encyclopedia," 2017. [Online; accessed 12-March-2017].

[12] N. Patterson, "The algebraic decoding of goppa codes," *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 203–207, 1975.

[13] K. Pietrzak, "Cryptography from learning parity with noise," in *International Conference on Current Trends in Theory and Practice of Computer Science*, pp. 99–114, Springer, 2012.

[14] N. Sendrier, "Niederreiter encryption scheme," in *Encyclopedia of cryptography and security*, pp. 842–843, Springer, 2011.

[15] Y. X. Li, R. H. Deng, and X. M. Wang, "On the equivalence of mceliece's and niederreiter's public-key cryptosystems," *IEEE Transactions on Information Theory*, vol. 40, no. 1, pp. 271–273, 1994.

[16] H. Niederreiter, "Knapsack-type cryptosystems and algebraic coding theory," *PROBLEMS OF CONTROL AND INFORMATION THEORY-PROBLEMY UPRAVLENIYA I TEORII INFORMATSII*, vol. 15, no. 2, pp. 159–166, 1986.

[17] N. T. Courtois, M. Finiasz, and N. Sendrier, "How to achieve a mceliece-based digital signature scheme," in *International Conference on the Theory and Application of Cryptology and Information Security*, pp. 157–174, Springer, 2001.

[18] P. Jiménez, F. Thomas, and C. Torras, "3d collision detection: a survey," *Computers & Graphics*, vol. 25, no. 2, pp. 269–285, 2001.

[19] Wikipedia, "Post-quantum cryptography — wikipedia, the free encyclopedia," 2017. [Online; accessed 10-March-2017].

[20] M. Baldi, "The mceliece and niederreiter cryptosystems," in *QC-LDPC Code-Based Cryptography*, pp. 65–89, Springer, 2014.

[21] F. Zeng, *Nonlinear codes: representation, constructions, minimum distance computation and decoding*. PhD thesis, Universitat Autònoma de Barcelona, 2014.