Confidence in Post-Quantum Algorithms

cloud security alliance®

The permanent and official location for Cloud Security Alliance Quantum Safe Security research is https://cloudsecurityalliance.org/working-groups/quantum-safe-security/

Acknowledgments

Quantum-Safe Security Working Group Co-chairs:

Bruno Huttner
Ludovic Perret

Quantum-Safe Security Working Group:

Roberta Faux, *Lead Author*
Ludovic Perret, *Reviewer*

CSA Staff:

Hillary Baron, *Research Analyst*
AnnMarie Ulskey, *Graphic Design*

Table of Contents

Summary

NIST recently announced its Round 3 candidates for future post-quantum cryptography or quantum-resistant standards. As the world prepares to transition to quantum-resistant cryptography, it is essential to understand how much analysis has been done on the security of individual quantum-resistant algorithms and classes of algorithms.

This document focuses on quantifying the cryptanalytic and mathematical research that builds meaningful confidence in the algorithm's security as evidenced in publications. This does not include analysis of implementation, performance nor application to protocols.

Public Key Cryptography Today

It is widely known that large-scale quantum computers will render public-key cryptography vulnerable. Today's public-key cryptosystems are based on the hard mathematical problems of integer factorization and computing discrete logarithms. Both of these problems can be solved efficiently by the quantum Shor's algorithm on a sufficiently large quantum computer. The public key systems such as Diffie-Hellman, RSA, DSA, and ECDSA form the security backbone of ubiquitously deployed encryption, digital signing mechanisms, authentication frameworks, and more.

Large Scale Quantum Computer will break modern public-key cryptography used in these domains:

- Public Key Infrastructure (PKI)
- Key Management Systems
- Authenticated Web Communication (TLS)
- Secure Point-to-Point (SSH)
- Transport Security (Osc)
- Key Agreement
- Identification and Authentication
- Password-Authenticated Key Exchange (PAKE)
- PGP/GPG
- Secure / Multipurpose Internet Mail Extensions (S/MIME)
- Over-the-Air Rekeying (OTAR)
- Domain Name System Security Extensions (DNSSEC)
- Encrypted File System • Internet Key Exchange (IKE)
- ZRTP (Secure VoIP Protoc1D1)

NIST and other standards bodies such as ETSI and ISO are leading efforts to vet quantum-resistant (or post-quantum) cryptography. New quantum-resistant standards are expected in the 2022-2025 timeframe. It is anticipated that the new post-quantum standards will include a number of different cryptographic algorithms in order to accommodate the various applications of cryptography. Many of these new post-quantum algorithms have complexities that traditional public-key algorithms do not have, such as decryption failures, very large ciphertext, or requirements to maintain the state of a process between applications of the algorithm.

There are advantages and disadvantages to the emerging quantum-resistant algorithms. Nearly all quantum-safe algorithms have much larger key sizes than today's public key algorithms. While some have higher computational requirements, others are competitive with or even faster than today's public key cryptography. Some of the proposed algorithms have ciphertext sizes that are even smaller. More important than key size and computational time for any algorithm, is how much confidence there is in the difficulty of the underlying mathematical problem on which the security of the algorithms rests. Confidence level is mainly built from amassed time dedicated to solving a particular mathematical problem. As the amount of cryptanalysis time increases, confidence in the security of a system improves.

Integer factorization has remained one of number theory's greatest problems. Integer factorization dates to ~300 BCE when Euclid defined primes and the concept of unique factorization. Trial division was the only method for factorization until 1643 when Fermat demonstrated that an integer could be written as the difference between two square numbers – an idea that has influenced some of the fastest integer factorization methods today. The discrete logarithm problem is comparatively a much more recent problem. Consistent efforts to solve the discrete logarithm problem date to from the early 1970s. For some specialized cases, such as discrete log over small characteristic fields progress has been made on quasi-polynomial time algorithms. However, for sufficiently-sized parameters, both these problems are considered to be computationally intractable for classical computers. In 1994, Peter Shor presented a polynomial time integer factorization and discrete logarithm algorithm that could run on a quantum computer. While there have been some sub-exponential approaches to solve certain specific cases of the discrete logarithm, until the advent of large-scale quantum computing there remains much confidence in today›s security parameters based on large integer factorization and discrete logarithms.

Quantum-Resistant Cryptographic Publications

Quantum-Resistant cryptographic algorithms fall into a few classes: code-based, hash-based, multivariate, lattice-based, multivariate, hash-based, code-based, and supersingular isogeny along with a few others.[1] The security of each class is based on a particular hard problem. The table below shows the type, along with the associated hard problem, first scheme, and year of publication.

[1] There are a few other emerging post-quantum schemes based on the security of zero-knowledge proof systems, symmetric key primitive, and others. Such schemes were not considered in this report.

Type	Based on Hard Problem	Hallmark Schemes	First Proposed PK Cryptosystem
RSA	Integer factorization	RSA	1977
Code	Hardness of decoding a linear code NP-hard	McEliece	1978
Hash	Statistical properties of key and function interaction	Lamport Signature	1979
Multivariate	Solving systems of multivariate polynomials NP-Hard	C* scheme HFE	1986
Lattice	Shortest vector problem	Ajtai-Dwork Encryption NTRU Regev	 1997 1998 2005
Isogeny	Problem of finding the isogeny mapping between two supersingular elliptic curves	SIDH	2011[2]

The International Association for Cryptologic Research maintains the Cryptology ePrint Archive. The ePrint Archive was started in 1997 and contains an electronic database of cryptographic related papers. One way to begin to quantify the amount of cryptanalytic research that has gone into each category is simply by counting the number of relevant papers published in the archive. The determination of relevance based on the title alone, that is, without reading all the papers, is subjective to a degree. It is also noted that this method will not capture some research. With the goal of quantifying the amount of cryptanalytic research, a review was undertaken to see if some insights from the cumulative literature would be illuminating.

Specifically, the process first collects papers from the ePrint archive based on certain keywords in the title. Duplicates are removed. Titles with keywords that are irrelevant to quantum-resistant cryptography are also removed. Next the titles are classified as relevant to cryptanalysis or not. This process introduces some subjectivity, and admittedly the approach will miss some relevant research not captured by keyword searching. Still, the results hopefully add some insight to confidence levels in quantum-resistant cryptography.

As of October 1, 2020, there were a little over 15,300 papers in the archive. The archive was filtered by title according to keywords. Keywords, included below, were the name of specific algorithms submitted to NIST[3] as well as the name of the general category. This first pass revealed 1,893 papers.

[2] In 2006, Alexander Rostovtsev and Anton Stolbunov proposed a public-key cryptosystem based on isogenies of ordinary elliptic curves. https://eprint.iacr.org/2006/145.pdf

[3] NIST Round 1 Submission can be found here: https://csrc.nist.gov/projects/post-quantum-cryptography/round-1-submissions Note: The following algorithm returned no results: Big Quake, CFPKM, Compact LWE, Ding Key, DME, DualMode, Edon-K, EMBLEM, GeMSS, Guess Again, Gui, KCL, KINDI, Lake, LEDAKEM, Lepton, LIMA, LOCKER, LOTUS, McNie, MQDSS, Ouroboros-R, Ramstake, RLCE-KEM, RQC, RVC, RVB, Titanium, SRTPI.

Many of these would turn out to not be relevant. Even if the keywords lattice, code, and hash were entirely eliminated, there still are roughly 700 papers potentially related to quantum-resistant cryptography.

Type	Keyword	Number of Papers
Code	*bike*	3
Code	*code*	299
Code	*dags*	5
Code	*hqc*	2
Code	*ledapkc*	1
Code	*mceliece*	34
Code	*nts-kem*	1
Code	*pqsigrm*	1
Code	*racoss*	1
Code	*ranksign*	2
Code	*rollo*	2
Hash	*hash*	501
Hash	*sphincs*	8
Multivariate	*giophantus*	1
Multivariate	*himq*	1
Multivariate	*luov*	3
Multivariate	*multivariate*	112
Multivariate	*rainbow*	17
Lattice	*dilithium*	5
Lattice	*falcon*	5
Lattice	*frodo*	4
Lattice	*hila5*	2
Lattice	*kyber*	7
Lattice	*lac*	125
Lattice	*lattice*	396
Lattice	*lizard*	5

Lattice	*newhope*	10
Lattice	*ntru*	89
Lattice	*qtesla*	3
Lattice	*round2*	1
Lattice	*round5*	5
Lattice	*saber*	7
Isogeny	*isogeny*	71
Isogeny	*sidh*	39
Isogeny	*sike*	12
To Be Determined	*post-quantum*	112
To Be Determined	*post-quantum*	1
	Total	**1893**

The papers were categorized according to type based on the keyword, i.e., titles with the keywords *frodo* or *kyber* were labeled type "LATTICE,"; titles with the keywords of *bike* or *mceliece* were labeled type "CODE." Each of the types (*code, hash, multivariate, lattice, and isogeny*) were compiled. Duplicates were then removed as some papers were captured repeatedly because they had multiple keywords in their titles. The remaining papers in the "post-quantum" keyword category were quickly labeled by type manually. Papers that dealt with multiple post-quantum schemes or overviews of the post-quantum landscape were marked a "general" type.

Only 3 of 125 papers from the keyword "lac" were relevant to the LAC Key Exchange. Titles including "blackbox" or "black box" or "replacement" were removed.

Next, the elimination of non-relevant papers was done based on the keywords that would not be inconsistent with quantum-resistant key exchange mechanisms and digital signatures. These included the following: *symmetric, FHE, homomorphic, zero-knowledge, functional encryption, MPC, blockchain, RSA, IBE, ABE, Oblivious*. Papers containing those words were considered not applicable to quantum-resistant cryptography and hence removed. Specifically, the following keywords were used to remove paper titles from each category.

Type	Keyword Removal Filter
Code	PUF, voting, hash, multivariate, Tardos, authentication code, turbo, lattice, LWE, Machine Code, game, detection codes, wire*, biometric, matroids, codex, Secure Erasure, batch, private codes, passwords, wire, source code, frameproof Codes, spammed, prolific, retrievability, MAC
Multivariate	zero, homomorphic, hash, FHE, block, RSA

Lattice	FHE, untrusted, storage, privacy, IBE, RSA, DSA, factoring, Attestation, SEAL, secret sharing, ZNIZK, wallet, pseudorandom, ledger, secure comput*, intrusion, stream cipher
Isogeny	hash, oblivious, pairing

Due to the large volume of work on general cryptographic hash functions, for the hash category, papers were included only if they had both the word "hash" and one of following: "signature" or "LMS" or "XMSS" or "SPHINCS."

This left the following number of papers for each category.

Type	Number of Papers
Code	167
Hash	43
Multivariate	118
Lattice	350
Isogeny	114
TOTAL	792

With this narrowed list, it was desired to classify each remaining entry simply as *cryptanalysis-related* or *non-cryptanalysis*.

The cryptanalysis classification was broadly defined to include any paper that would potentially be influential in developing a fundamental attack against the underlying problem of a given type of cryptosystem. It should be noted the cryptanalysis classification deliberately excluded timing attacks, fault injection, and differential power analysis, as these attacks, while extremely important, are typically implementation-dependent and do not affect the fundamental analysis of hard problems behind the cryptosystem. Paper titles that included the keywords "cryptanalysis" or "attack" were put into an initial "cryptanalysis" category.

The *non-cryptanalysis* classification was defined to include new construction, protocols, implementation, and timing related attacks. Hence, titles were marked as *non-cryptanalysis* if any of the following appeared in the title:

- *timing attacks, fault injection, and differential power analysis, efficien*, implementations, FPGA, fast, speed, processor, ARM, hardware, software, scheme, fault, channel, timing, practical, performance, differential power, compression*

A few examples are given:

- *Efficient BIKE Hardware Design with Constant-Time Decoder*
- *Side-Channel Attacks on the McEliece and Niederreiter Public-Key Cryptosystems*
- *Key Agreement Protocols Based on Multivariate Polynomials over Fq*
- *NTRU-KE: A Lattice-based Public Key Exchange Protocol*

At this point, a cursory manual inspection of remaining unclassified titles was done and obvious adjustments were made in the cryptanalysis. With the goal of looking at cryptanalysis specifically, papers were sorted by cryptanalysis and non-cryptanalysis in order to examine any key observations.

Cryptanalytic Publications by Type

Code

Code-based cryptography grew out of the field of error correction codes which dates back to 1940's. The early emergence of a public key cryptosystem based on coding theory significantly predates the ePrint archive; this makes it difficult to appropriately capture research potentially relevant to cryptanalysis of quantum-resistant code-based cryptography. Notably, code-based McEliece algorithm from 1978 has remained secure, modulo parameter sizes, although different families of codes have been broken. It is a well-studied area with many papers on cryptanalysis and implementation improvements. Code-based cryptography has withstood scrutiny for over 40 years.

Figure 1

- 167 papers relevant to Code-based cryptography between 1999-2020
- 39 papers relevant to cryptanalysis of Code cryptography = 23%

It should be noted that there are certainly other papers prior to 2001 which also address the security of code-based cryptography. The graph above shows very consistent study over the last 15 years.

Hash

Most hash functions are designed around ad-hoc constructions in which the bits of the message are mixed to produce an output, not based on mathematical problems. Hash functions are considered to be "one-way" functions, impossible to reverse. While such algorithms are accepted as hard to break, there is no formal proof.

In 1979, Leslie Lamport invented one-time hash-based signatures. Later, one-time signatures were combined with Merkle tree structures. More recently variants of hash-based signatures include the eXtended Merkle Signature Scheme (XMSS), Leighton-Micali Signatures (LMS), and SPHINCS. The use of hash functions relies on statistical properties of key and function interaction – with a vanishingly small probability.

Hash algorithms in general have received an enormous amount of study with over 500 papers related to hashing over the past 20 years. While the cryptanalysis of using hashes for signatures is only a handful of papers, since security of hash-based signatures directly depends on security of the hash, there is high degree of confidence in hash-based cryptography to date.

Figure 2

- 501 paper related to hashing in general
- 42 papers relevant to Hash-based cryptography from 1999-2020
- 12 papers relevant to cryptanalysis of Hash cryptography = 28%

Multivariate

In 1988, Tsutomu Matsumoto and Hideki Imai presented the first multivariate cryptosystem known as C*. While this was broken by Jacques Patarin in 1995, it launched numerous follow-up proposals in the area, including Patarin's Hidden Monomial Cryptosystems, which has inspired some of the multivariate schemes in use today.

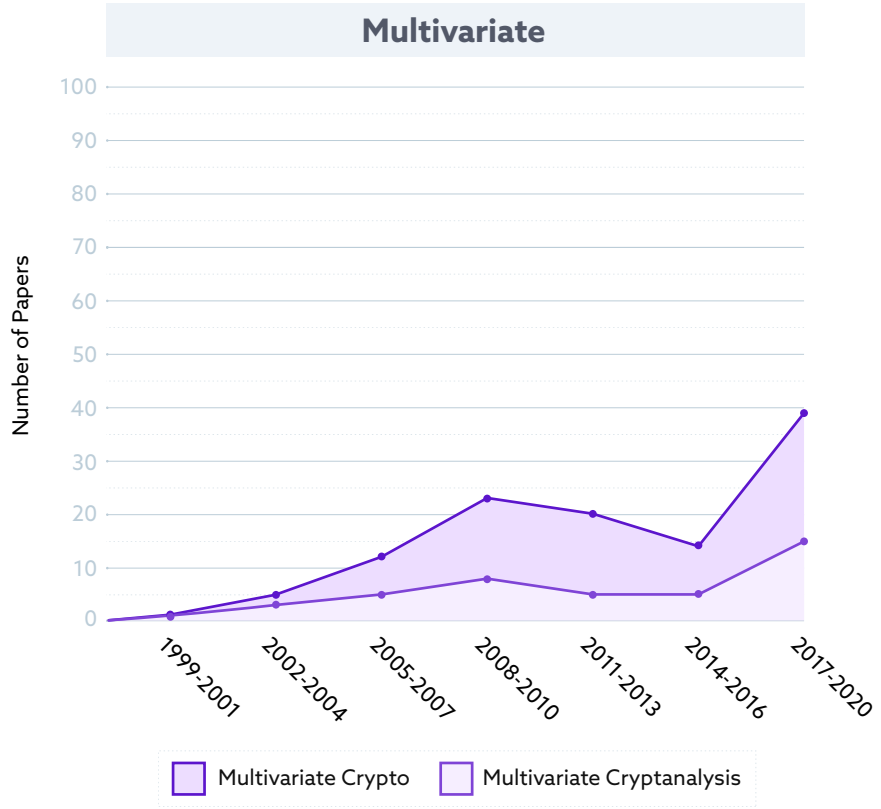

Figure 3

- 118 papers relevant to Multivariate cryptography from 2001-2020
- 42 papers relevant to cryptanalysis of Multivariate cryptography = 36%

Lattice

There has been a tremendous amount of work focusing on lattice cryptography over the past 15 years, and this area has clearly received the most study. The first lattice-based cryptosystem was introduced in 1996 by Miklos Ajtai. In 1998, the lattice-based NTRU scheme was published by Hoffstein, Pipher, and Silverman. In 2005, Oded Regev introduced the first lattice-based public-key encryption scheme whose security was proven under worst-case hardness assumptions.

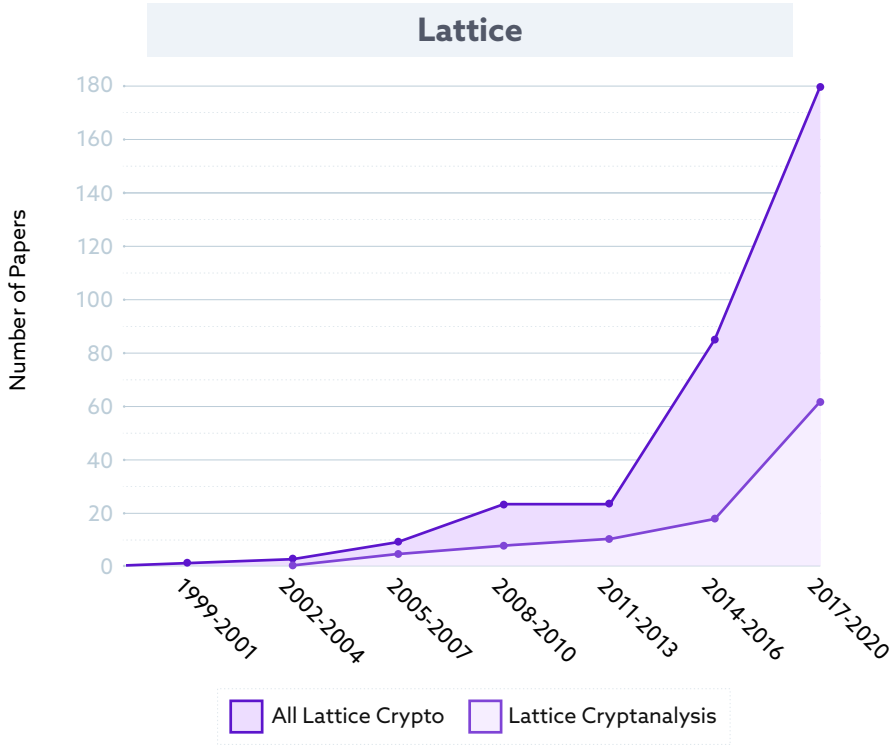

Figure 4

- 350 papers relevant to lattice cryptography from 2001-2020
- 121 papers relevant to cryptanalysis of lattice cryptography = 35%

Isogeny

Isogeny based cryptography is the most recent of proposed quantum-resistant schemes. It is based on the hard problem of finding an isogeny or map between two supersingular elliptic curves with the same number of points. There are 115 papers in the ePrint archive on isogeny cryptography that have been published between 2011-2020 as well as a 2005 paper on Efficient Scalar Multiplication by Isogeny Decompositions.

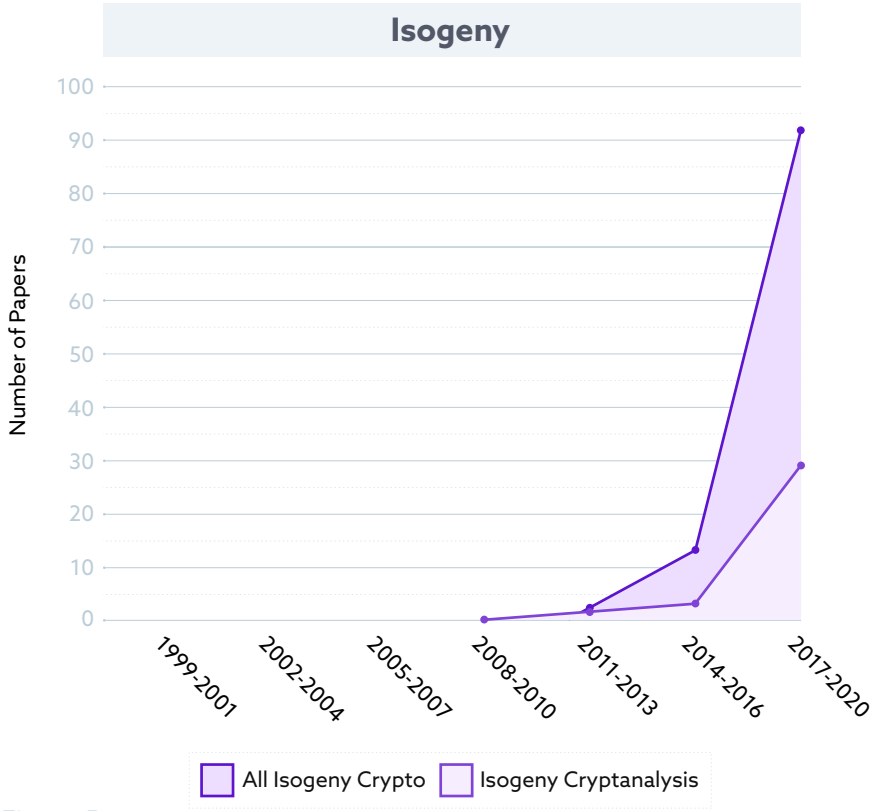

Figure 5

- 115 papers relevant to Isogeny cryptography
- 37 papers relevant to cryptanalysis of Isogeny cryptography = 32%

Classical Public Key Problems: Factorization and Discrete Logarithms

Besides the long history of factorization, there has been much study on RSA and large integer factorization, as well as the similarly hard problem of solving discrete logarithms. Similar to the process above, papers were collected based on keywords in titles. The IACR archive has 727 unique papers with the keywords RSA, Factoring, Diffie, Discrete Log. Duplicates were removed. Keywords removal filters included the following:

- voting, homomorphic, oblivious, hash, multiparty, multi-party, LWE, threshold, bitcoin, blockchain, SIDH, Isogeny, zero-knowledge, adversaries, adversarial, universal, versatile, NIZK, password, identity, identification, ID-based, multivariate, post-quantum.

The non-cryptanalysis classification was indicated if any of the following appeared in the title: timing, fault, differential, efficien*, implementation, FPGA, fast, speed, processor, ARM, hardware, scheme, channel, practical, performance.

This left 382 papers relevant to RSA, Diffie Hellman, Factoring, Discrete Logs with 159 (or about 43%) identified as being relevant to cryptanalysis. Some of these techniques may only be applicable to a factor or discrete log variants such as small exponents, unbalanced moduli, weak prime factor, etc.

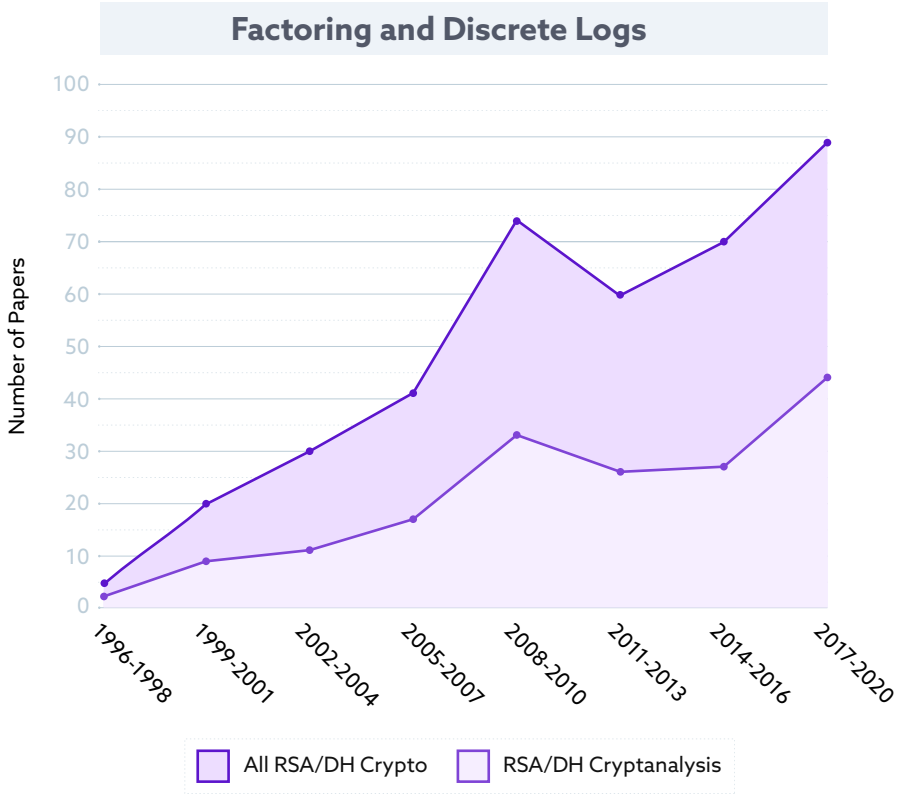

Figure 6

- 382 papers relevant to RSA and Diffie-Hellman
- 159 papers relevant to cryptanalysis of RSA and Diffie-Hellman = 43%

Conclusion

It is important to note that this is only a cursory look at cryptanalytic efforts. There was no metric for quality. It is recognized that a simple keyword search undoubtedly missed papers that may have been relevant but did not include one of the keywords. There is some margin of error in the classification process. Despite these shortcomings, it is hoped that this data offers a glimpse of the work to date.

It is not easy to achieve high security along with efficiency, but it is critically important. Future broken cryptographic algorithms are a constant looming threat to real-world security. Post-Quantum Cryptography is a relatively recent research area that has witnessed enormous global participation in finding practical quantum-resistant cryptographic solutions. The challenge in post-quantum cryptography remains to balance demands for usability and flexibility without sacrificing trust.

The world is moving closer to the standardization of new cryptographic algorithms to resistant attacks from future large-scale quantum computers. It is hoped that this paper will spark additional reflection about the analysis of quantum-resistant cryptography.