

Introducing a New Post-quantum Cryptography Scheme: Intersection of NLP and Cryptography

Zahra Dorostkar

Abstract. In this work, we introduce a quantum-resistant public-key cryptosystem that leverages structured linguistic transformations in place of traditional number-theoretic problems. By embedding encryption and decryption operations within the combinatorial complexity and semantic variability of natural language, the scheme defines novel hardness assumptions that defeat known quantum-algorithmic attacks. Comparative evaluation against leading post-quantum families—lattice, code, multivariate, isogeny, and group-based—demonstrates that our linguistic-based approach provides a complementary trade-off space between performance and security. This work inaugurates a new class of post-quantum cryptographic primitives rooted in natural-language complexity.

Introduction and Problem Statement

There are wide-ranging applications of public-key cryptography algorithms in securing communications, protecting data, authenticating users, and ensuring the integrity of digital transactions across various industries and technologies.

The emergence of quantum computing poses a significant threat to the security of traditional public key digital signature schemes, as quantum algorithms such as Shor's algorithm have the potential to efficiently break cryptographic protocols that underpin digital signatures. This vulnerability raises concerns about the long-term integrity and confidentiality of digital signatures in the context of quantum computational capabilities.

In light of this challenge, there is a critical need to explore innovative cryptographic methodologies that can withstand quantum attacks and ensure the resilience of

digital signature schemes in quantum computing environments. The existing cryptographic landscape lacks robust solutions that are specifically designed to address the vulnerabilities exposed by quantum algorithms, necessitating the development of novel approaches that can fortify digital signature schemes against quantum computational threats.

To address this imperative, the proposed work aims to integrate linguistic features into a public key cryptography scheme to develop a quantum-resistant digital signature solution. By leveraging the complexities of natural language structures, the work seeks to enhance the security properties of the digital signature scheme and fortify it against potential quantum attacks, thereby addressing the pressing security concerns posed by quantum computing.

In order to explain why we think the proposed approach may be resistant to quantum attacks, it's important to highlight the following points:

- **Linguistic Complexity:** The integration of linguistic structures introduces a level of complexity that may pose significant challenges for quantum algorithms to efficiently break. This complexity is derived from the intricate nature of linguistic structures, which may inherently hinder the ability of quantum algorithms to exploit patterns and vulnerabilities.
- **Quantum-Resistant Design Principles:** The design principles underlying the integration of linguistic elements into cryptographic systems aim to leverage the complexities of natural language structures to create cryptographic functions that are inherently difficult for quantum algorithms to compromise. This approach is founded on the premise that linguistic-based cryptography may introduce novel challenges for quantum algorithms, thereby enhancing the scheme's resistance to quantum attacks.

By emphasizing these technical aspects, the project aims to establish a foundation for a quantum-resistant digital signature scheme that leverages linguistic complexities and randomness to fortify cryptographic operations against potential quantum threats.

Related Works

In fact, there are cryptography algorithms that use Natural Language Processing (NLP) techniques. One interesting application is in the field of steganography, which is the practice of concealing messages or information within other non-secret data. In the context of NLP, steganography techniques can be used to hide secret messages within natural language text.

Short Paper Title

One approach is to embed secret information in the structure or semantics of the text itself. For example, certain words or phrases in a document could be chosen to represent specific letters or symbols in a secret message. By carefully selecting these indicators, the hidden message can be decoded using NLP algorithms.

Another application is in text-based encryption, where NLP techniques are used to transform plain text messages into a form that is not easily readable without the appropriate decryption key. This could involve techniques such as text tokenization, substitution ciphers based on language patterns, or even using language models to generate cipher text.

While NLP-based cryptography may not be as common or standardized as traditional cryptographic algorithms, there is ongoing research in this area exploring the potential of combining NLP and cryptography for secure communication and data hiding.

To the best of our knowledge there are not any public-key cryptography algorithms that are specifically based on NLP techniques. Public-key cryptography algorithms, such as RSA or Elliptic Curve Cryptography, are typically based on mathematical principles and operations rather than NLP concepts.

Public-key cryptography algorithms rely on mathematical properties of prime numbers, prime fields, and discrete logarithms to provide secure key exchange and encryption/decryption operations. These algorithms are widely used and well-studied in the field of cryptography and are not typically combined with NLP techniques.

However, it is possible to incorporate NLP techniques within the context of public-key cryptography. For example, NLP techniques can be used to:

- Analyze the security properties of cryptographic algorithms
- Design secure cryptographic protocols
- Implement cryptographic algorithms using natural language processing libraries or tools

But these are separate applications of NLP within the broader field of cryptography rather than specific NLP-based public-key cryptography algorithms.

Regarding that, there are no widely recognized post-quantum cryptography algorithms that are specifically based on NLP (Natural Language Processing) techniques. Post-quantum cryptography focuses on developing cryptographic algorithms that are secure against attacks by quantum computers.

It's important to note that the field of cryptography is vast and constantly evolving, so there may be emerging research that explores novel approaches combining

NLP and post-quantum cryptography. However, at present, the mainstream post-quantum cryptographic algorithms do not have a direct basis in NLP techniques.

Based on our current understanding, the cryptographic domain lacks linguistic-based algorithms. In contrast, the existing quantum-safe algorithms fall into categories such as:

- Lattice-based cryptography
- Code-based cryptography
- Multivariate cryptography
- Isogeny-based cryptography
- Group-based cryptography

The absence of linguistic-based cryptography underscores a unique opportunity for innovative exploration in developing novel cryptographic solutions resilient to quantum threats [1, 2, 3].

References

- [1] E. Ortega, J. *Enhancing Networking Cipher Algorithms with Natural Language..* 43–54. <https://doi.org/10.5121/csit.2022.121013>
- [2] Mansour, A. M. A., Fouad, M. A. M. *Cryptography Protocol: A Novel Multilingual Adaptive Encryption Technique with Phonetic Based Ciphering.* International Journal of Security and Its Applications, 11(9), 13–22. <https://doi.org/10.14257/ijisia.2017.11.9.02>
- [3] Jing, X., Hao, Y., Fei, H., Li, Z. *ext encryption algorithm based on natural language processing.* Proceedings - 2012 4th International Conference on Multimedia and Security, MINES 2012, 670–672. <https://doi.org/10.1109/MINES.2012.216>

Zahra Dorostkar
Skolkovo Institute of Science and Technology
Moscow, Russia
e-mail: zahradorostkar68@gmail.com