Cryptographic Algorithms and Protocols for Financial Data Analysis

Marco Pedicini[1] and Carla Mascia[2]

[1] Roma Tre University, Italy
[2] DataKrypto, Italy

The digital transformation of financial services has intensified the demand for robust cryptographic techniques that ensure data privacy, security, and regulatory compliance. This tutorial provides a comprehensive overview of modern cryptographic tools designed to address the unique challenges of financial data analysis. Its purpose is not to create cryptographers but to equip practitioners with a practical toolbox of techniques, enabling them to understand which tools are available and how to apply them effectively.

Noise

At the core of cryptographic security lies the concept of noise as an encryption key. Drawing from Shannon's theory, security is measured in terms of entropy: higher entropy translates to greater uncertainty and, consequently, stronger security [6]. In cryptography, noise represents a transformation that preserves the entropy of the original signal rather than a loss of information. The ciphertext, when properly encrypted, appears as noise and carries no discernible information about the original message without the decryption key. This principle can be illustrated through signal-plus-noise diagrams, where the addition of noise obscures the original signal, rendering the ciphertext indistinguishable from random data. The effectiveness of this approach depends on how the noise is added, as different operations—such as bitwise logical ANDor XOR—produce distinct outcomes. Properly managed noise ensures that the ciphertext remains secure; however, excessive noise can hinder decryption, highlighting the need for precision in cryptographic design.

Malleability

Another fundamental concept is malleability in cryptography, where the ability to modify ciphertexts without decryption is considered a feature rather than a vulnerability. This property forms the basis of homomorphic encryption, which allows computations to be performed directly on encrypted data.

De Cifris Koine – FCiR25 ACTA https://doi.org/10.69091/koine/vol-8-P04

This raises a crucial question:

Can this property be extended to every function?

If so, encrypted data would never need to be decrypted for processing, revolutionizing data security.

To this end, current cryptographic techniques such as fully homomorphic encryption (FHE), zero-knowledge proofs (ZKP), secure multi-party computation (SMPC) offer practical solutions for secure data analysis.

Fully Homomorphic Encryption (FHE), enables computations on encrypted data without decryption, preserving privacy throughout the process. However, many FHE implementations tend to be computationally intensive and often unnecessary for many applications. Conversely, Somewhat Homomorphic Encryption (SHE) supports a limited set of arithmetic operations and is sufficient for numerous use cases, emphasizing the importance of selecting the right tool for the problem at hand.
Homomorphic Encryption ensures privacy, correctness of computation, and client-side decryption, even if it may also introduce challenges such as computation overhead and noise accumulation. The bootstrap technique in cryptography addresses the issue of noise growth in ciphertexts, ensuring that encrypted data remains usable after multiple operations [2].

Secure Multi-Party Computation (SMPC), extends the principles of secure computation to scenarios where multiple parties wish to jointly compute a function without revealing their raw data. This technique is particularly relevant in financial contexts, such as banks collaboratively estimating systemic risk without disclosing sensitive information.
SMPC guarantees privacy, correctness, independence of inputs, and reaches thus fairness, that is, it ensures that no party gains an unfair advantage [3].
The comparison between FHE and SMPC reveals that they are complementary tools rather than competing approaches, each offering unique strengths depending on the context. FHE excels in scenarios with encrypted computation on a single server, while SMPC is ideal for distributed computations among multiple parties.

Zero-Knowledge Proofs (ZKP), provide a mechanism for proving the validity of a statement without revealing any additional information. In financial applications, ZKPs enable institutions to demonstrate compliance or solvency without exposing sensitive data, such as detailed balance sheets or account information [4]. This capability is invaluable in regulatory contexts where transparency and privacy must coexist.

Financial Analysis

Financial analysis is increasingly intertwined with machine learning techniques. Consequently, it has become essential to protect the confidentiality of the under-

lying data and also to ensure the security and integrity of the machine learning models themselves, while preserving their effectiveness.

The training phase represents the initial stage of this process, where large datasets—often containing sensitive information such as customer transactions, credit histories, insurance claims, and trading records—are utilized. To enable model training while safeguarding individual privacy, *Differential Privacy* is commonly utilized. This technique introduces carefully calibrated noise that hides personal information yet preserves the statistical patterns essential for effective learning [1]. In scenarios where multiple institutions — such as banks or regulatory bodies — seek to collaborate without sharing raw datasets, *Federated Learning* provides an effective solution. It enables the joint training of a shared model while data remain securely on local servers, with only the necessary model updates exchanged among participants [5].

Once trained, a machine learning model is defined by its weights—the parameters that encapsulate all the knowledge it has acquired. In many sectors, such models represent valuable intellectual property, and their theft can result in severe economic and reputational damage. Equally concerning is the risk of malicious tampering, such as the injection of harmful code or the manipulation of model parameters. To address these threats, Homomorphic Encryption offers a robust solution. By encrypting the model's weights, it ensures that only authorized entities possessing the appropriate cryptographic keys can perform computations on the model. Crucially, the model remains encrypted at all times—never exposed in plaintext—thereby preventing both unauthorized access and manipulation.

Finally, during the inference stage—when prompts are submitted and responses generated—sensitive information may still be exposed. Consider, for instance, financial institutions increasingly deploying AI assistants for customer service, compliance, and investment research, where both queries and outputs may contain confidential data, such as details from investigations into suspicious transactions. *End-to-End Encryption* safeguards this communication channel, ensuring that only authorized parties can access the exchanged information.

While each of these techniques appears to address a specific phase of the AI pipeline—training, collaboration, model protection, or inference—it is crucial to recognize that true security and privacy can only be achieved through their coordinated integration. Advanced privacy-preserving solutions should, therefore, be designed holistically, combining Differential Privacy, Federated Learning, Homomorphic Encryption, and End-to-End Encryption into a unified framework. Such a combined approach ensures continuous protection of both data and models throughout their entire lifecycle, from training to deployment and real-time interaction.

Conclusion

In conclusion, cryptography should not be seen as an obstacle but rather as a powerful enabler—an evolving toolbox that makes secure, collaborative, and privacy-preserving financial analysis possible. The guiding principle that "a ciphertext is forever" highlights the enduring need to design systems whose security withstands technological progress and future computational advances.
The techniques discussed — Homomorphic Encryption, Secure Multi-Party Computation, Zero-Knowledge Proof, Differential Privacy, Federated Learning, and End-to-End Encryption — together form the foundation for next-generation financial infrastructures that are both innovative and resilient. By integrating these methods into a coherent framework, we can ensure that data confidentiality and model integrity remain protected throughout the entire lifecycle of financial AI systems, fostering trust and enabling innovation in an increasingly digital world.

References

1. Dwork, C.: Differential Privacy (Invited Paper). In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (eds.) ICALP 2006, Part II. LNCS, pp. 1–12. Springer, Berlin, Heidelberg (2006). https://doi.org/10.1007/11787006_1
2. Gentry, C.: Fully homomorphic encryption using ideal lattices. In: Mitzenmacher, M. (ed.) 41st ACM STOC, pp. 169–178. ACM Press (2009). https://doi.org/10.1145/1536414.1536440
3. Goldreich, O., Micali, S., Wigderson, A.: How to Play any Mental Game or A Completeness Theorem for Protocols with Honest Majority. In: Aho, A. (ed.) 19th ACM STOC, pp. 218–229. ACM Press (1987). https://doi.org/10.1145/28395.28420
4. Goldwasser, S., Micali, S., Rackoff, C.: The Knowledge Complexity of Interactive Proof-Systems (Extended Abstract). In: 17th ACM STOC, pp. 291–304. ACM Press (1985). https://doi.org/10.1145/22145.22178
5. Li, T., Sahu, A.K., Talwalkar, A., Smith, V.: Federated Learning: Challenges, Methods, and Future Directions. IEEE Signal Processing Magazine **37**(3), 50–60 (2020). https://doi.org/10.1109/msp.2020.2975749. http://dx.doi.org/10.1109/MSP.2020.2975749
6. Shannon, C.E.: Communication theory of secrecy systems. Bell Systems Technical Journal **28**(4), 656–715 (1949)