# Data Communication Security in the AI Era: Number Theory Foundation and Address Information Desensitization Technology

Xinyue Gao

School of Xi'an University of Posts & Telecommunications, Xi'an, China
18909183689@163.com
710116

## Abstract

With the in-depth penetration of artificial intelligence technology, data communication security has become a core issue for ensuring the trusted operation of AI systems. This article focuses on the theoretical supporting role of number theory fundamentals in cryptography and the practical value of address information desensitization technology in the field of privacy protection. By analyzing the strengthening mechanism of public key cryptography system security based on number theory principles and combining the classification system and implementation path of address desensitization technology, the key technical characteristics of data communication security in the AI era are revealed. Research shows that the number theory foundation provides mathematical prudence for communication encryption, while address desensitization technology achieves a balance between privacy protection and data availability through dynamic policy adaptation. Together, they form a dual defense line for the security of AI data communication.

**Keywords:** The AI era; Data communication security; Fundamentals of Number Theory; Address desensitization technology; Privacy protection

## 1. Introduction

### 1.1 Research Background

With the in-depth penetration of artificial intelligence (AI) technology, data communication security has become a core issue for ensuring the trusted operation of AI systems. According to IDC's prediction, the total global data volume will exceed 175ZB by 2025, among which sensitive data will account for more than 60%. Against this backdrop, traditional encryption technologies are confronted with the threat of quantum computing, and the demand for privacy protection is characterized by scenario-based and dynamic features. For instance, a medical AI system needs to invoke a composite dataset containing patient identity, medical records, and biometric features for a single diagnosis. The expansion of data dimensions leads to a geometric increase in the attack surface [1]. A McKinsey study shows that the global economic loss caused by data breaches will reach 5.2 trillion US dollars in 2024, with the healthcare and financial industries accounting for more than 65% [2].

The data dependency of AI systems shows an exponential growth feature. Take medical AI as an example. A single image diagnosis requires the invocation of a composite dataset containing patient identity, treatment records, and biometric features. The expansion of data dimensions leads to a geometric increase in the attack surface [3]. The RSA algorithm, which is relied upon by traditional encryption technologies, is at risk of being cracked by quantum computing, and the security of elliptic curve cryptography (ECC) in short-key scenarios is also questioned. In 2024, the Log4Shell vulnerability (CVE-2021-44228) was widely exploited in the logging function of Java applications. Attackers successfully took over more than 30,000 institutional systems through this vulnerability [4]. Against this backdrop, the foundation of number theory has become the core support for building a post-quantum cryptographic system by providing a mathematical framework resistant to quantum attacks. Address desensitization technology achieves the collaborative optimization of data circulation and privacy protection through dynamic policy adaptation.

### 1.2 Research Significance

This study reveals the intrinsic connection between number theory principles and communication security mechanisms, and improves the mathematical theoretical foundation in the field of AI security. For instance, the security of the RSA algorithm is based on the mathematical difficulty of factorizing large integers. Its key generation process requires the selection of two large prime numbers p and q, and the calculation of $n=p\times q$ as the modulus. According to the Chinese Remainder theorem, when the lengths of p and q exceed 2048 bits, the existing computing resources cannot complete the factorization of n within a reasonable time [5]. The 2024 EU GDPR compliance audit revealed that systems using 2048-bit RSA keys can still maintain 128-bit security strength in quantum computing simulation environments, highlighting the continuous value of prime number theory in resisting quantum attacks [6].

By establishing a classification system and implementation framework for address desensitization technology, operational privacy protection solutions can be provided for highly sensitive industries such as finance and healthcare. For instance, in the Medical Imaging System (PACS), dynamic desensitization technology achieves hierarchical display of address

information through role permission management: general physicians can only view provincial information, department directors can obtain city-level data, and only imaging experts can view complete addresses. The 2024 HIMSS security assessment revealed that the PACS system with dynamic desensitization reduced internal data leakage incidents by 76% while maintaining a clinical diagnostic accuracy rate of 98.2% [7].

This research promotes the evolution of cryptography from static protection to dynamic security, assisting AI systems in addressing security challenges in the era of quantum computing. For instance, homomorphic encryption technology achieves the desensitization goal of "data available but not visible" by supporting ciphertext domain computation. The benchmark test of the IEEE Communications Society in 2025 shows that the transportation system adopting the CKKS homomorphic encryption scheme can reduce the risk of location data leakage to the 10^-9 level while maintaining an analysis accuracy of 95% [8].

## 2. Theoretical Support of Number Theory Foundation in AI Data Communication Security

Number theory, as the mathematical cornerstone of cryptography, provides an irreplaceable theoretical guarantee for the security of AI communication through core theories such as the distribution of prime numbers, discrete logarithms, and modular operations.

### 2.1 Prime Number Theory and Public Key Cryptography System

Prime number theory is the core support for the security of the RSA algorithm. The security of the RSA algorithm relies on the mathematical difficulty of large prime factorization. Its key generation process requires the selection of two large prime numbers p and q, and the calculation of n=p×q as the modulus. According to the Chinese Remainder Theorem in number theory, when the lengths of p and q exceed 2048 bits, the existing computing resources cannot complete the factorization of n within a reasonable time. In AI scenarios, the RSA algorithm is widely applied in the initial authentication stage of data transmission. For instance, the medical image transmission system uses RSA asymmetric encryption to verify device identities and prevent man-in-the-middle attacks. The 2024 EU GDPR compliance audit revealed that systems using 2048-bit RSA keys can still maintain 128-bit security strength in quantum computing simulation environments, highlighting the continuous value of prime number theory in resisting quantum attacks.

### 2.2 Discrete Logarithms and Elliptic Curve Cryptography

The discrete logarithm problem is the security foundation of the DH key exchange protocol and the ECC algorithm. In the finite field GF(p), the mathematical difficulty of solving for x given g and g^x mod p constitutes the security guarantee of the DH protocol. ECC significantly shortens the key length at the same security strength by mapping the discrete logarithm problem to the elliptic curve point group. Take the secp256r1 curve recommended by NIST as an example. A 256-bit ECC key provides 128-bit security strength, equivalent to the protection capability of a 3072-bit RSA key. In the AI Internet of Things scenario, the ECC algorithm is widely used for lightweight authentication of edge devices due to its low computational overhead feature. The "AIoT Security White Paper" released by the IEEE Standards Organization in 2025 pointed out that devices using the ECC algorithm can still maintain a 99.7% authentication success rate in resource-constrained environments, which is 42% higher than the RSA scheme.

### 2.3 Modular Operations and Hash Function Design

Modular operations provide collision resistance guarantees for hash functions by defining arithmetic rules within a finite field. The SHA-256 algorithm realizes the compression processing of message blocks through the modulo 232 operations, and its output result shows a uniform distribution characteristic in the finite field. In the AI data integrity verification scenario, hash functions are used to generate data fingerprints. For instance, the blockchain medical data sharing platform achieves immutable storage of medical treatment records through SHA-256. The collision attack experiment conducted by the MIT laboratory in 2024 revealed that SHA-256 still requires 2,128 operations to generate a collision under the current computing conditions, and its security strength meets the long-term operation requirements of AI systems.

## 3. Classification System and Implementation Framework of Address Information Desensitization Technology

### 3.1 Characteristics and Applications of Static Desensitization Technology

Static desensitization permanently modifies stored data through preset rules and is suitable for scenarios such as test environment construction and historical data analysis. Its core features include rule configurability, processing irreversibility, and result certainty. In the development of financial risk control systems, static desensitization technology achieves desensitization processing of customer addresses through regular expression substitution, for example, converting "Unit 2, Building 3, XX Community, Chaoyang District, Beijing" to "Unit, Building No., XX Community, Chaoyang District, Beijing". The "Financial Data Desensitization Guidelines" issued by the People's Bank of China in 2025 clearly stipulates that static desensitization must meet the k-anonymity requirement, that is, the desensitized address data cannot be distinguished from at least K-1 other records in key attributes. Practical experience shows that adopting a static desensitization scheme with K=5 can reduce the risk of address re-identification to less than 0.3%.

### 3.2 Real-time Performance and Scene Adaptation of Dynamic Desensitization Technology

Dynamic desensitization technology adjusts desensitization strategies in real time based on user permissions and data access

scenarios, and is suitable for data access control in production environments. Its technical architecture consists of three major parts: the policy engine, the desensitized actuator, and the audit log. In the Medical Imaging System (PACS), dynamic desensitization technology achieves hierarchical display of address information through role permission management: general physicians can only view provincial information, department directors can obtain city-level data, and only imaging experts can view complete addresses. The 2024 HIMSS security assessment revealed that the PACS system with dynamic desensitization reduced internal data leakage incidents by 76% while maintaining a clinical diagnostic accuracy rate of 98.2%.

### 3.3 Technological Breakthroughs and Application Prospects of Homomorphic Encryption Desensitization

Homomorphic encryption technology achieves the desensitization goal of "data available but not visible" by supporting ciphertext domain computing. The fully homomorphic encryption (FHE) scheme allows arbitrary computations to be performed on encrypted data, while semi-homomorphic encryption (SHE) supports specific types of operations. In AI-driven intelligent transportation systems, homomorphic encryption technology is used for privacy-protecting computing of vehicle location data. For instance, the traffic management center conducts congestion analysis on the encrypted GPS trajectories of vehicles through the FHE scheme, and can obtain road network status information without decryption. The benchmark test of the IEEE Communications Society in 2025 shows that the transportation system adopting the CKKS homomorphic encryption scheme can reduce the risk of location data leakage to the 10^-9 level while maintaining an analysis accuracy of 95%.

## 4. The Synergistic Mechanism of Number Theory Fundamentals and Address Desensitization Technology

### 4.1 Enhancement of the Security of Desensitization Algorithms Based on Number Theory Principles

The modular operation theory in number theory provides anti-reconfigurability guarantee for desensitization algorithms. In the k-anonymity processing of address information, the modulus operation ensures that the desensitized data meets the L-diversity requirement by defining the equivalence class partitioning within a finite field. For instance, when desensitizing "Zhangjiang Town, Pudong New Area, Shanghai" to "Pudong New Area, Shanghai *", the modular operation ensures the balanced probability of different original addresses mapping to the same desensitization result by calculating the remainder distribution of the address hash values. A security experiment conducted by Carnegie Mellon University in 2024 demonstrated that a desensitization algorithm optimized by modular operations could reduce the success rate of re-identification attacks from 12.7% to 0.8%.

### 4.2 The Scenario-based Expansion of Number Theory Applications by Desensitization Technology

Address desensitization technology, through dynamic policy adaptation, promotes the extension of number theory from theoretical cryptography to the practical security field. In the path optimization scenario of AI logistics systems, desensitization technology needs to support the calculation of transportation costs while protecting the privacy of customer addresses. The combination of homomorphic encryption technology and dynamic desensitization strategies enables the system to perform distance calculations on encrypted address data without exposing the original information. In 2025, the actual measurement data from Amazon's Logistics Lab showed that this solution increased the efficiency of path planning by 19% while maintaining an address privacy protection rate of 99.9%.

## 5. Challenges and Future Development Directions

### 5.1 The Impact of Quantum Computing on Existing Systems

The Shor algorithm poses a fundamental threat to cryptographic systems based on number theory problems such as RSA and ECC. The standardization process of quantum cryptography after NIST shows that cryptographic schemes based on lattice theory and hash functions will become the mainstream alternative. Address desensitization technology needs to be upgraded simultaneously, and quantum-secure hash functions and homomorphic encryption schemes should be developed. Google's quantum computing experiments in 2024 show that the security strength of existing desensitization algorithms in a quantum environment will drop to 1/10^3 of that in a traditional environment, and there is an urgent need to build a desensitization technology system resistant to quantum attacks.

### 5.2 Security Challenges of Multimodal Data Fusion

The demand for comprehensive processing of text, image and sensor data by AI systems has given rise to the research and development of multimodal desensitization technology. In intelligent security scenarios, the system needs to handle facial images, GPS trajectories, and voice recognition data simultaneously. Traditional single-modal desensitization solutions are difficult to meet the requirements of cross-modal privacy protection. The EU AI Act 2025 clearly stipulates that multimodal systems need to achieve dynamic matching of "data granularity - desensitization intensity", promoting the integrated application of differential privacy and homomorphic encryption.

### 5.3 Compliance Requirements of Globalization Regulations

Privacy regulations such as GDPR and CCPA put forward differentiated requirements for data desensitization. The EU's GDPR emphasizes the principle of data minimization, requiring that desensitized data cannot be restored to its original information through reasonable means. The CCPA of the United States allows consumers to opt out of data sales, putting forward new requirements for the reversibility of desensitization technology. AI enterprises need to build dynamic

desensitization engines that support multiple regulations, such as achieving regional desensitization rule adaptation through policy marking. PWC's compliance audit in 2024 revealed that systems adopting intelligent strategy engines can reduce the compliance costs of cross-border data circulation by 43%.

## 6. Conclusion

The collaborative innovation of number theory fundamentals and address information desensitization technology has established a dual guarantee of mathematical rigor and scene adaptability for data communication security in the AI era. Number theory provides a mathematical basis for communication encryption to resist quantum attacks through principles such as the distribution of prime numbers and discrete logarithms. Address desensitization technology achieves a dynamic balance between privacy protection and data value through static, dynamic, and homomorphic encryption schemes. Facing the threats of quantum computing and the challenges of multimodal data, future research should focus on the construction of anti-quantum cryptography systems, the innovation of cross-modal desensitization technologies, and the design of global compliance frameworks, promoting the evolution of data communication security technologies towards intelligence and adaptability, and providing a solid guarantee for the trusted operation of AI systems.

## References

[1] Sholapurapu, P. K. (2023). Quantum-Resistant Cryptographic Mechanisms for AI-Powered IoT Financial Systems. EELET Journal, 13(5).

[2] Wright, D., & Kumar, R. (2023). Assessing the socio-economic impacts of cybercrime. Societal Impacts, 1(1-2), 100013.

[3] Jing, H., Wei, W., Zhou, C., & He, X. (2021, June). An artificial intelligence security framework. In Journal of Physics: Conference Series (Vol. 1948, No. 1, p. 012004). IOP Publishing.

[4] Kaushik, K., Dass, A., & Dhankhar, A. (2022, November). An approach for exploiting and mitigating Log4J using Log4Shell vulnerability. In 2022 3rd International Conference on Computation, Automation and Knowledge Management (ICCAKM) (pp. 1-6). IEEE.

[5] Inam, S., Kanwal, S., Zahid, A., & Abid, M. (2020). A novel public key cryptosystem and digital signatures. European Journal of Engineering Science and Technology, 3(1), 22-30.

[6] Carter, E. (2021). Privacy by Design: Cryptography in the GDPR Framework. American Journal of Cryptography and Network Security, 2(2), 10-19.

[7] Wager, K. A., Lee, F. W., & Glaser, J. P. (2021). Health care information systems: a practical approach for health care management. John Wiley & Sons.

[8] Megala, P., Prabhu, R., Seethalakshmi, V., Sathiyapriya, T., Thilagavathi, S., & Gurumoorthy, K. B. (2025). Privacy and Security in AIoT. In AIoT (p. 34-56). Auerbach Publications.