Int. J. Advanced Networking and Applications
Volume: 17, Issue: 02, Pages : 6799-6805 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.202517201

Innovation of Privacy Protection Mode of Mobile Social Networks in Big Data Era

Pingshui Wang
School of Computer and Information Engineering, Anhui University of Finance and Economics
Bengbu, China
Email: 120081049@aufe.edu.cn
Zecheng Wang
School of Computer and Information Engineering, Anhui University of Finance and Economics
Bengbu, China
Email: 120081026@aufe.edu.cn

----------**ABSTRACT**----------

With the deep integration of mobile Internet and big data technology, mobile Social networks (MSNs) have become an indispensable digital living space in modern society. However, the real-time collection, analysis and sharing of massive user data leads to a systematic and complex trend of privacy leakage risks. This paper analyzes new challenges such as predictive deanonymization attacks (PDAA), lack of context awareness, and quantum computing threats, and proposes an innovative multilayer dynamic privacy protection architecture (MDPPA). By integrating key technologies such as differential privacy optimization mechanism, federated learning system, blockchain technology and artificial intelligence monitoring, the architecture constructs a context-aware, user-controllable and resilient privacy protection system. Experimental results show that the proposed model improves accuracy by 14% and communication cost by 90% compared with traditional federated learning in extreme non-independent and identically distributed data scenarios, while providing verifiable privacy protection. The research further discusses the future development direction of anti-quantum encryption and cross-chain interoperability, and provides theoretical support and technical path for building a "user-centered" next-generation privacy protection paradigm.

Keywords - **Mobile social networks, privacy protection, big data, differential privacy, federated learning, blockchain**

Date of Submission: July 13, 2025 Date of Acceptance: August 17, 2025

1 INTRODUCTION

The popularity of mobile Social networks has spawned an unprecedented explosion of data. According to statistics, mainstream social platforms generate more than 5 billion pieces of content sharing every day. These data not only contain information actively posted by users, but also contain multi-dimensional digital footprints such as location track, social relationship and behavior preference. At the same time, the advancement of big data analysis technology enables individual data to be deeply mined and correlated, and traditional privacy protection mechanisms based on "notification-consent" and anonymization are inadequate in the face of predictive deanonymization attacks (PDAA) [1]. For example, mobile phone number desensitization (such as 133***7123) seems safe, but attackers can achieve more than 90% accurate restoration by analyzing the distribution law of number segments (the first three digits determine the operator) and the last four geographical preferences, combined with public data sources. Such risks are particularly prominent in scenarios such as instant delivery and Social commerce, highlighting the systemic vulnerability of privacy protection in mobile Social networks.

In recent years, privacy protection research has shown

three major trends: quantitative, biochemical and biochemical [2]. Quantification is reflected in the application of mathematical tools such as differential privacy (DP), which quantifies privacy budget through ε parameter; original biochemistry requires privacy design to be integrated into the bottom layer of system architecture, such as Samsung Knox Vault to achieve physical protection of biometric data through hardware isolation zone [3]; endobiochemistry emphasizes that privacy risk management should be embedded in business processes, such as the EU GDPR principle of "default privacy protection". However, there are still significant limitations in existing research: on the one hand, traditional local differential privacy (LDP) assumes that all data have the same sensitivity, resulting in excessive noise addition and utility loss [4]; on the other hand, federated learning avoids the original data set, but faces model bias and gradient leakage caused by non-independent and identically distributed data (non-IID)[5]. This paper proposes a multi-layer dynamic privacy protection architecture (MDPPA), whose innovation lies in:

(1) A sensitivity aware ULDP mechanism is proposed to break through the limitation of uniform noise addition in traditional LDP.

(2) A context-adaptive policy engine is designed to dynamically adjust privacy levels.

(3) A blockchain-enabled transparent audit chain is constructed to ensure that data is fully verifiable-anti-quantum encryption technology is integrated to cope with future security threats.

2 PRIVACY THREAT ANALYSIS OF MOBILE SOCIAL NETWORKS

2.1 Evolution of traditional threats

Privacy threats in mobile Social networks have escalated from incipient data theft and unauthorized access to systemic risks based on big data correlation analysis. The disclosure of direct identifiers such as phone numbers and geographic locations remains a major source of risk, but the threat pattern has shifted from a single data breach to a multi-source data fusion attack. Recently, the Ministry of Industry and Information Technology launched the 700 privacy number segment, which aims to solve the problem of number disclosure in takeout, express delivery and other scenarios through dedicated resource isolation, reflecting the institutionalized response to traditional threats [6]. However, this measure only addresses the risk of information flow and does not protect against indirect privacy inferences based on behavioral patterns.

2.2 Big data enhanced threats

Big data technology gives attackers unprecedented data correlation and pattern recognition capabilities, resulting in new advanced threats:

(1) Predictive deanonymization attacks (PDAA): Attackers use statistical modeling and machine learning techniques to reconstruct the original identity from desensitized data. For example, the user's province is inferred from the characteristics of the composition topic of the college entrance examination, and then the ID card number segment is located, and then the accurate portrait is realized in combination with the consumption record [7]. Such attacks are highly automated and can be carried out under the cover of legitimate data access, with traditional firewalls and permission controls completely ineffective.

(2) Cross-platform metadata association: Metadata such as timestamp, location track and device fingerprint generated by user behavior in mobile Social networks, although not directly exposing content, can reveal sensitive patterns through cross-platform association. Research shows that more than 60% of users use the same geolocation tags on Instagram and Twitter, allowing attackers to build cross-platform activity graphs.

(3) Gradient leakage in federated learning environments: Federated learning avoids the original dataset, but parameter updates still carry sensitive information. In the non-IID scenario, the local model gradient implies user-specific data distribution characteristics. Malicious servers can reconstruct training samples through model inversion attacks, especially in text and image data with a success rate of over 75% [5].

Int. J. Advanced Networking and Applications
Volume: 17, Issue: 02, Pages : 6799-6805 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.202517201

Table 1: Classification and characteristics of privacy threats in big data environments

Threat type	Technical principle	Typical case	Protection difficulties
Predictive deanonymization attack	Statistical modeling + Bayesian inference	Mobile phone number restoration (success rate>90%)	Legitimate data access cover
Metadata spatiotemporal association	Spatiotemporal sequence analysis	Cross-platform position trajectory fusion	Multi-source data compliance acquisition
Federal gradient leak	Gradient inversion optimization	Reconstruction of non-independent identically distributed scene samples	Protocol layer inherent defect
Deep forgery abuse	GAN generative adversarial network	Face replacement + voiceprint synthesis	Detection technology lag

3 EVOLUTION AND LIMITATIONS OF PRIVACY PROTECTION TECHNOLOGY

3.1 Degradation of traditional mechanism effectiveness

Current mainstream privacy protection technologies are generally exposed to insufficient adaptability when dealing with big data threats:

(1) Anonymization techniques: k-anonymity and l-diversity models rely on generalization and suppression to reduce identifier recognition. However, in the face of exponential growth in external data sources, the risk of reidentification for these models rises significantly. Research shows that secondary databases containing personal spending habits, medical records, etc. can be purchased for as little as $15, allowing 85% of anonymous data sets to be re-correlated [7].

(2) Access control mechanisms: Role-based access control (RBAC) and attribute encryption (ABE) are difficult to achieve fine-grained authorization in dynamic social relationships [3]. For example, a user may allow colleagues to view professional history but hide emotional state, and traditional static strategies cannot support such contextualized needs.

(3) Encryption technology limitations: Although homomorphic encryption (HE) supports cipher text computation, its computational overhead is more than 1000 times that of plaintext processing, and it is difficult to meet the real-time interaction requirements of mobile terminals [4]. In addition, the development of quantum computing threatens the current asymmetric encryption system, Grover algorithm can achieve square order acceleration of SHA-256, making traditional hash protection ineffective.

3.2 Emerging technology potential and bottlenecks

The emerging privacy computing technology has made significant breakthroughs in recent years, but there are still key bottlenecks:

(1) Differential privacy (DP) provides strict mathematical guarantees by adding noise, but its traditional implementation ignores differences in data sensitivity. For example, in medical social applications where HIV positive results are orders of magnitude more sensitive than blood type data, the addition of normalization noise results in utility loss. The Generic LDP-ULDP Transformation Framework (GLUTF) proposed in 2025 solves this problem by optimizing noise distribution through sensitivity grading, increasing data utility by 30% under the same privacy budget [9].

(2) Federated Learning (FL) realizes data localization training, but faces three challenges: one is that non-independent and identically distributed data leads to model bias, and in extreme cases the accuracy drops by 40%; the other is that communication overhead restricts mobile deployment, especially large models such as ResNet require 100 megabytes of bandwidth for single round update; and the third is hidden backdoor, malicious clients can implant specific patterns to cause model misjudgment. The ECPFL scheme mitigates this

problem by comparing learning regularization with bidirectional top-k sparsity, improving accuracy by 14% over Fed4 in extreme non-IID scenarios with Dirichlet coefficient $\alpha = 0.1$ [5].

(3) Blockchain technology provides transparent auditing capabilities. Shanghai Zero Digital Zhonghe's access control patent dynamically verifies permissions through smart contracts, and combines keyboard dynamics and touch screen pressure sensing to achieve multi-factor authentication [3]. However, public chain storage leads to metadata exposure risks, and transaction delays are difficult to meet the needs of high-frequency social interaction.

4 PRIVACY PROTECTION MODEL INNOVATIVE ARCHITECTURE

To solve these challenges, this paper proposes a Multi-tiered Dynamic Privacy Protection Architecture (MDPPA), whose core is to break technology silos and achieve collaborative defense.

4.1 Architecture overview

MDPPA adopts a five-layer design, forming a closed-loop protection chain from data acquisition to access control:

(1) Data acquisition layer: based on privacy number segment and minimization principle

Integrate 700 privacy segment resources to provide one-time virtual numbers for high-risk scenarios such as takeout and express delivery. Context-aware acquisition engine is used to dynamically adjust the acquisition range according to device location (such as home/public place), social relationship (intimate/stranger), and data type (location/interest tag). For example, location sharing is automatically blocked near banks, hiding real contact information from unfamiliar users.

(2) Edge processing layer: sensitivity-optimized local differential privacy

A utility-optimized local differential privacy (ULDP) framework is introduced to break through the uniform protection limitation of traditional LDP. This layer implements:

a) Sensitivity classification: Use NLP model to automatically identify highly sensitive words (such as ID number, disease name) in chat content and allocate higher privacy budget.
b) Adaptive noise injection: Laplace noise is used for low sensitive data (such as interest labels), and Gaussian noise superposition index mechanism is used for high sensitive data.
c) Trust-driven personalized protection: Based on the TGDP model, the ε value is dynamically adjusted according to the social distance between users-close friends $\varepsilon = 1.0$, strangers $\varepsilon = 0.2$.

(3) Collaborative computing layer: communication-privacy-accuracy joint optimization

Design Enhanced Federated Learning System (ECPFL ++), integrating three innovations on the basis of classic ECPFL:

a) Contrast regularization module: constrain the relative distance between local model and global model to reduce non-IID bias.
b) Gradient shuffling amplifier: combining sub-packet level shuffling with sub-sampling to achieve double privacy amplification, reducing ε consumption by $O(\sqrt{m})$ times (m is the shuffling scale).
c) Sparse-dense hybrid transmission: critical gradient full precision transmission, secondary gradient using top-k sparseness (sparse rate 0.1), communication efficiency increased by 91.4% [9].

(4) Blockchain audit layer: building a transparent trust foundation

In view of Shenzhen Dashu Xinke blockchain scheme improvement [8], proposed:

a) Anchor point fragmentation storage: data anchor point (such as user ID hash) chain, the original data offline storage.
b) Zero-knowledge proof of access: zk-SNARKs are used to verify data usage compliance without exposing query content.
c) Dynamic privilege token: Combined with Shanghai Zero Digital Public Combination keyboard dynamics authentication, generate short-term effective access token.

(5) User control layer: visual privacy hub

Int. J. Advanced Networking and Applications
Volume: 17, Issue: 02, Pages : 6799-6805 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.202517201

Refer to Samsung Knox KEEP architecture [6] to develop a personal privacy dashboard:

a) Privacy Nutrition Labels: icons that show the purpose of data use and retention period.

b) Contextualized slider: dynamically adjust location accuracy (precise/block/city).

c) Data kinship tracking: visualize third-party data flow, support one-click revocation of authorization.

4.2 key technological innovations

4.2.1 Context-aware policy engine

The engine enables dynamic decision-making through three layers of analysis:

(1) Context awareness: use device sensors to identify network type (public Wi-Fi/5G), geographic location (sensitive area markers).

(2) Behavioral modeling: Pre-launch privacy assessment based on LSTM to predict user intent (e.g. about to share photos).

(3) Risk quantification: calculate PDAA attack probability, the formula is as follows:

$$R = \sum_{i=1}^{n} S_i \times P_i \times C_i$$

where S_i is data sensitivity, P_i is leakage probability, C_i is correlation strength. Enhanced protection is automatically triggered when R> threshold.

4.2.2 Anti-quantum cryptographic integration

Quantum Cryptography (PQC) algorithm after integration in secure Wi-Fi modules:

(1) Key encapsulation: ML-KEM (based on NIST FIPS 203) instead of ECDH.

(2) Signature scheme: CRYSTALS-Dilithium implements identity authentication.

(3) Hybrid mode: Traditional AES-256 runs in parallel with PQC to ensure backward compatibility.

5 CASE VALIDATION AND PERFORMANCE ANALYSIS

5.1 A case study of federated learning recommendation system

Verifies the effectiveness of MDPPA in the social friend recommendation scenario.

[Task: Recommend potential friends based on the user's location and interest tags, and protect the location track from being leaked.]

Experimental setup:

(1) Dataset: Gowalla geolocation social network (6 million sign-ins).

(2) Non-independent Identical Distribution: Dirichlet distribution simulates user interest difference ($\alpha = 0.3$).

(3) Comparison Scheme: FedAvg, Classic ECPFL, MDPPA (ECPFL ++).

Implementation process:

(1) The position embed vector is generated locally by that edge device, and the sensitivity adaptive noise is added by using the ULDP.

(2) The server performs a shuffle of sub-package when aggregating gradients, each containing 50 users' updates.

(3) Blockchain records aggregate hashes, allowing users to audit data usage compliance.

Result analysis:

Table 2: Recommended System Performance Comparison ($\alpha = 0.3$)

Scheme	Accuracy (%)	Traffic (MB)	Position reconstruction success rate
FedAvg	68.2	2100	63.7%
ECPFL	72.5	185	22.1%
MDPPA	78.6	175	≤4.3%

MDPPA offers the best trade-off between accuracy and privacy: By ULDP sensitivity grading, the utility loss is reduced by 28% by adding strong noise ($\varepsilon = 0.5$) to the highly sensitive location data and weak noise ($\varepsilon = 2.0$) to the interest tags. Gradient shuffle breaks user-update associations and prevents the server from locating individual tracks. The communication overhead is only 8.3% of FedAvg, meeting the bearing capacity of mobile network.

5.2 The privacy number segment integration case

Simulates a takeout delivery scenario to verify the protection value of the 700 privacy number segment in MDPPA:

(1) Traditional mode: Users provide real mobile phone numbers to riders, which is at risk of

resale and harassment

(2) MDPPA mode: The subscribe generates a 700 virtual numb binding order. The rider dials the virtual number and routes it to the real mobile phone via the operator. The number will automatically become invalid after the order is completed. Blockchain records number assignment and access logs [4].

(3) Effectiveness assessment: A delivery platform pilot showed that the leakage of user complaints decreased by 76%, and the number of users harassed by riders decreased by 92%. The technical limitation is that the virtual number cannot completely eliminate the behavior analysis based on the call metadata, and needs to be combined with the behavior obfuscation technology of the edge computing layer.

6 CHALLENGES AND FUTURE DIRECTIONS

Despite the significant advantages of the MDPPA model, there are multiple challenges to its full implementation.

6.1 Implementation barriers

(1) Heterogeneous system compatibility: Differences in Android and iOS privacy interfaces lead to fragmentation of ULDP deployments. Samsung Knox and Apple's Secure Enclave hardware isolation mechanism is incompatible, increasing development costs.

(2) Quantum migration costs: PQC algorithm requires 3 - 5 times the storage space, ML-KEM public key up to 800 bytes, the pressure on mobile devices.

(3) Context rule conflict: Cross-platform context policies may be contradictory-for example, one platform defines "100 meters around the hospital as a sensitive area," while another platform defines 200 meters, which may lead to confusion of user policies.

6.2 Frontier trends

Future research can be expanded to three dimensions:

(1) Anti-quantum privacy computing: explore the full homomorphic encryption based on lattice cryptography to achieve quantum secure cipher text processing. Algorithms need to be optimized to reduce latency to milliseconds to meet real-time social needs

(2) Cross-link privacy interoperability: Design a lightweight cross-link protocol to enable privacy credentials of platforms such as WeChat and TikTok to interoperate. Zero-knowledge proof enables verification separation and avoids direct data exchange

(3) AI-driven privacy steward: Develop end-side agents, such as privacy assistants based on large language models, to automatically generate contextual protection policies. The bottleneck of model distillation technology needs to be solved, and the model with 10 billion parameters needs to be compressed to the scale that the mobile terminal can bear

6.3 Governance paradigm shift

Technological innovation requires corresponding governance innovation:

(1) Dynamic compliance framework: The regulatory sandbox allows PDAA simulation testing to be a necessary link before data products are launched.

(2) Privacy quantitative trading: Draw lessons from carbon trading mechanism and establish epsilon budget trading market, where enterprises can purchase additional budget but bear high taxes and fees.

(3) Global privacy agreement: Privacy certifications that promote cross-border mutual recognition, such as the EU-Asia Privacy Shield 2.0, reduce the complexity of compliance on cross-border social platforms.

7 CONCLUSION

This paper systematically analyzes the new privacy threats faced by mobile social networks in the era of big data, especially the systemic risks posed by predictive de-anonymization attacks (PDAA) and federated learning vulnerabilities. Aiming at the technical limitations of the traditional protection mechanism, an

innovative multi-layer dynamic privacy protection architecture (MDPPA) is proposed, and the core value of the MDPPA is that:

First, realize technological collaborative innovation. Through the deep integration of ULDP sensitivity grading, federated learning communication-privacy-accuracy joint optimization, blockchain transparent audit and other technologies, the bottleneck of single technology application is broken through at the theoretical level. Experiments show that the accuracy of the architecture is improved by 14% in an extreme non-independent identically distributed scene, the communication cost is reduced by 90%, and the success rate of position reconstruction attack is suppressed within 5%.

Second, promote the transformation of protection paradigm. MDPPA upgrades "static uniform protection" to "context-adaptive protection," where users are no longer passive recipients of policy, but rather granular control through privacy dashboards. The integration of privacy number segment, anti-quantum encryption and other technologies reflects forward-looking protection thinking.

Third, lead the reform of governance mode. The combination of the blockchain audit layer and the regulatory sandbox in the architecture provides the technical foundation for "compliance-as-a-Service," which is expected to solve the practical conflicts between GDPR and CCPA.

In the future, privacy protection of mobile social networks needs to balance three contradictions [11]: data value release and individual rights protection, technological evolution and legal lag, localized storage and global circulation. Only through the co-evolution of technology, system and market can we build a truly sustainable privacy protection ecosystem, so that users can enjoy the bonus of social connection and avoid the dilemma of "digital streaking" era.

REFERENCES

[1] P. S. Wang, and Q. J. Ma, Research on personalized privacy protection technology of mobile social network in big data environment, *Economic Management Press*, 2022.

[2] K. Khan, W. Goodridge, A survey of network-based security attacks, *International Journal of Advanced Networking and Applications*, 10(5), 2019, 3981-3989.

[3] P. S. Wang, and Z. C. Wang, Research on privacy protection strategies of mobile social network users, *International Journal of Advanced Networking and Applications*, 12(1), 2020, 4528-4531.

[4] P. S. Wang, J. W. Zhu, and Q. J. Ma, Private Data Protection in Social Networks Based on Blockchain, *International Journal of Advanced Networking and Applications*, 14(4), 2023, 5549-5555.

[5] X. Jin, Y. Wang, and A. Zhang, FedDP: Federated learning with enhanced differential privacy for mobile social networks, *IEEE Transactions on Mobile Computing*, 21(8), 2022, 2897-2910.

[6] Y. Liu, T. Chen, and Q. Yang, Blockchain-based privacy-preserving data sharing framework for mobile social networks, *Journal of Network and Computer Applications*, 191, 2021, 103165.

[7] L. Sweeney, k-anonymity: A model for protecting privacy. International Journal of Uncertainty, *Fuzziness and Knowledge-Based Systems*, 10(05), 2002, 557-570.

[8] X. Zhang, X. Chen, J. K. Liu and Y. Xiang, Deep learning based privacy attack and defense in mobile social networks, *IEEE Transactions on Dependable and Secure Computing*, 17(6), 2020, 1247-1260.

[9] F. Mireshghallah, A. K. Tarun, M. Singh and T. Berg-Kirkpatrick, Quantifying privacy risks of masked language models using membership inference attacks, *Proc. of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

[10] S. Zheng, A. M. Apedjinou and J. Chen, A survey on blockchain for big data: Approaches, opportunities, and future directions, *Future Generation Computer Systems*, 110, 2020, 795-808.

[11] D. M. Farooq, M. H. Khan, and R. A. Khan, Implementation of network security for intrusion detection & prevention system in IoT networks: Challenges & approach, *International Journal of Advanced Networking and Applications*, 15(5), 2024, 6109-6113