Creativity and Innovation
2025, VOL. 9, NO. 6, 161-166
DOI: 10.47297/wspciWSP2516-252736.20250906

Anti- "Human Flesh Search" Scheme Based on Blockchain and Zero-Knowledge Proofs

Zihang Luo

Chongqing Normal University, Chongqing, 401331, China

ABSTRACT
The public ledger characteristic of blockchain grants data immutability but simultaneously introduces privacy leakage risks, making association analysis between on-chain behaviors and real-world identities possible. Existing privacy protection schemes struggle to balance the anonymity of the querier with the traceability of malicious behaviors. On one hand, legitimate inquiry behaviors are easily reverse-tracked by third parties through on-chain records (i.e., "human flesh search" targeting the querier); on the other hand, a completely anonymous environment may lead to data abuse without the possibility of accountability. To address this issue, this paper proposes an anti-"human flesh search" privacy protection system based on blockchain and zero-knowledge proofs. Addressing the aforementioned contradictions, this paper presents a blockchain data sharing scheme that balances privacy and regulation. The scheme utilizes IPFS to implement graded encrypted storage for large files. The core innovation lies in combining the Schnorr protocol and Chameleon Hash to construct a Blockchain Designated Verifier Proof (BDVP). While verifying user query permissions through blockchain smart contracts, the system utilizes the trapdoor property of the Chameleon Hash to achieve the non-transferability of proofs, preventing third parties from reverse-tracking the querier's identity by analyzing on-chain records[2]. Furthermore, the system introduces a threshold private key held by regulatory agencies to ensure that, in the event of data abuse, malicious users can be de-anonymized and held accountable according to the law.

KEYWORDS
Blockchain; Zero-knowledge proof; Human flesh search; Chameleon hash; Privacy protection

1 Introduction

1.1 Research Background

The open and transparent ledger characteristic of blockchain introduces a series of privacy security risks. In typical social network scenarios, merely associating an account with a relevant phone number and IP address allows malicious attackers to obtain identity information. Related research points out that traditional blockchain systems can only provide "Pseudonymity" rather than true "Unlinkability"[1]. Attackers can link on-chain pseudonym addresses to real-world identities by analyzing transaction graphs and network traffic, thereby triggering severe "human flesh search" (doxing) and privacy leakage events.

1.2 Problem Statement

In existing data sharing models, privacy protection and regulatory traceability are often an irreconcilable contradiction. Completely anonymous systems (such as Zcash [12]) protect user privacy but also provide a breeding ground for illegal transactions and malicious behaviors [4]; meanwhile, real-name systems face the risk of single-point leakage caused by centralized data storage. How to build a system that protects queriers from reverse tracking while enabling effective regulation of malicious behaviors is an urgent problem that needs to be solved.

1.3 Contributions

This paper proposes a privacy protection scheme based on Blockchain Designated Verifier Proof (BDVP). It protects Data Owners through encryption (preventing unauthorized data viewing) and protects Legitimate Queriers through non-transferability (preventing query behaviors from being subjected to association analysis that exposes identity). The main contributions are as follows:

(1) Designed a graded encrypted storage architecture based on IPFS to address the issue of data privacy leakage.

(2) Proposed a non-interactive zero-knowledge proof protocol combining the Schnorr protocol and Chameleon Hash. This achieves the non-transferability of query behaviors, preventing reverse "human flesh search" targeting the queriers.

(3) Introduced a regulatory mechanism based on threshold cryptography, realizing a balanced governance mode of "anonymous by default, traceable when necessary".

2 Related Work

2.1 Blockchain Privacy Protection Technology

Zhang et al. summarized blockchain privacy requirements into three dimensions: identity anonymity, transaction confidentiality, and unlinkability [1]. Early solutions mainly relied on Mixing technology, but they suffered from low efficiency and centralized risks. Subsequently, related research reviewed the application of zero-knowledge proofs in blockchain, highlighting the successful implementation of zk-SNARKs in projects like Zcash, while also emphasizing their performance bottlenecks in large-scale concurrent scenarios [4].

2.2 Zero-Knowledge Proofs and Identity Authentication

Zero-Knowledge Proof (ZKP), first proposed by Goldwasser et al., allows a prover to demonstrate knowledge of a secret without revealing the secret itself [11]. A non-interactive zero-knowledge proof scheme based on the Schnorr protocol was designed specifically for authentication in resource-constrained IoT devices, demonstrating its feasibility in lightweight environments [3]. However, traditional ZKP schemes on the blockchain face the issue of "permanent retention of proof records." Third parties may infer user behavioral patterns through long-term accumulation of proof records, thereby leading to privacy leakage risks.

2.3 Designated Verifier Proof (DVP)

Although traditional DVP schemes restrict the identity of the verifier, they often fail in scenarios of public verification on the blockchain. Chi et al. utilized Chameleon Hash functions to construct a special trapdoor mechanism, enabling the verifier (or regulator) to forge proofs, thereby endowing the proof records with "non-transferability" [2]. This paper builds upon this concept to construct a query verification protocol resistant to "human flesh search" (doxing).

3 Method

3.1 System Architecture

This chapter proposes an anti- "human flesh search" privacy protection system based on blockchain and zero-knowledge proofs. The system aims to resolve the contradiction between the vulnerability of data privacy in traditional centralized architectures and the difficulty of tracing and defending against malicious behaviors in decentralized architectures. By combining "Off-chain Storage (IPFS)"[10] and "On-chain Auditing (Blockchain)[8]," the system guarantees user data sovereignty while introducing regulatory agencies to effectively hold malicious "human flesh search" behaviors accountable.

3.1.1 System Entities

The system primarily consists of the following four types of participating entities:

(1) User (User / Data Owner & Querier): General participants in the system. Users switch between two roles depending on the interaction scenario: Data Owner (responsible for uploading data and setting access policies) and Querier (the subject initiating data access requests). When involving sensitive data queries, the Querier must act as a Prover to generate a Zero-Knowledge Proof, verifying their authority without exposing their identity. Users must undergo real-name authentication during registration and perform graded encryption on sensitive information when uploading. When accessing others' data, users generate zero-knowledge proofs to verify legitimate permissions while maintaining identity anonymity.

(2) Regulator (Regulator / Trusted Authority): The regulatory body is a federated trust network composed of multiple independent nodes (e.g., judicial, public security, audit departments). Its responsibilities include:

① Generating master key parameters during the system initialization phase;

② Assisting users in completing real-name registration;

③ In the event of malicious leakage, reconstructing permissions through multi-party collaboration to decrypt on-chain records and hold users accountable [1].

(3) Blockchain Network: As an immutable public ledger, the blockchain primarily assumes the function of "Auditing and Verification". It does not store original large data files but only stores data hash indexes, access control policies, and user query records (ZKP proofs). Smart contracts [9] are deployed on the chain to automatically execute permission verification logic [6].

(4) IPFS: Responsible for storing encrypted original large files and returning hash indexes for on-chain anchoring [4].

Given the high storage costs and low efficiency of blockchain, this system adopts IPFS as the off-chain storage layer. User privacy data (such as ID numbers, photos) is encrypted and stored in IPFS nodes, and the returned Content Identifier (CID) is anchored on the blockchain to ensure data has not been tampered with.

3.2 Data Classification and Encryption Storage Strategy

To balance the security of privacy protection with the storage efficiency of the blockchain system, this study designs a hybrid architecture of "off-chain storage and on-chain anchoring," combined with a data sensitivity classification mechanism, ensuring that only authorized entities (the user themselves or the regulatory agency) can decrypt sensitive data [1].

3.2.1 Data Classification

According to blockchain privacy requirements, mere pseudonymity does not equal unlinkability. To reduce the risk of association analysis after data leakage, this system divides user data into three security levels:

(1) Public Data (D_{pub}): Such as nicknames or tags voluntarily disclosed by the user. This data is stored in plaintext and used to build a basic social index.

(2) Semi-Private Data(D_{semi}) : Such as blurred geographical locations or institutional affiliations. This data adopts lightweight hashing or symmetric encryption storage and is visible only to specific groups.

(3) Strong Privacy Data (D_{priv}): Such as real names, ID numbers, mobile phone numbers, and other key identity information. As pointed out by Chi et al., in blockchain applications, laws and regulations require strictly protecting private information. Therefore, such data must adopt a high-strength hybrid encryption scheme and can only be decrypted when regulatory conditions are met[2].

3.2.2 Hybrid Encryption Scheme

For strong privacy data(D_{priv}) , directly using asymmetric encryption to process large files would result in excessive computational overhead. Dwivedi et al. pointed out that designing efficient data encryption algorithms is crucial for privacy protection in constrained environments. Therefore, this system designs the following hybrid encryption process [3]:

(1) Data Encryption Layer: The system generates a random symmetric key K_{sym}(e.g., AES-256) for each data upload and uses this key to encrypt privacy data, generating the data ciphertext:

$$C_{data} = Enc_{sym}(D_{priv}, K_{sym})$$

This ensures the efficiency of the data encryption and decryption process after authentication is passed [3].

(2) Key Encapsulation Layer (Regulatory Backdoor): To realize the regulatory requirement of "confidential in normal times, traceable when necessary," we need to securely transmit K_{sym} to the regulatory agency. The system uses the Master Public Key PK_{Reg}, negotiated by the regulatory consortium through the Distributed Key Generation (DKG) protocol during the initialization phase, to encrypt K_{sym}:

$$C_{key} = Enc_{asym}(K_{sym}, PK_{Reg})$$

This design ensures that no single entity can view the key. Only by collecting private key shards satisfying the threshold quantity can C_{key}

be unlocked through collusive reconstruction.

3.2.3 On-chain and Off-chain Collaborative Storage Based on IPFS

Since on-chain storage costs for blockchain are extremely high and unsuitable for storing large amounts of unstructured data, scaling solutions must be adopted. We use IPFS to solve the storage bottleneck. This system adopts the following storage strategy [4]:

(1) Off-chain Storage: The encrypted data package（containing C_{data} and C_{key}）is uploaded to the IPFS network. IPFS returns a unique hash fingerprint (CID) as the data index [4].

(2) On-chain Anchoring: The system packages the CID returned by IPFS and the data metadata (Metadata) and uploads them to the blockchain smart contract:

$$Tx_{store} = \{Hash(Owner) || CID || Timestamp\}$$

This "hash on-chain" method utilizes the decentralized nature of IPFS to ensure data is difficult to tamper with at a single point, while also ensuring the authenticity of the index through the immutability of the blockchain, greatly reducing on-chain storage pressure.

3.3 Core Protocol Workflow

The entire protocol consists of three key phases: system initialization and registration, privacy-preserving query verification, and regulatory tracing.

3.3.1 Phase 1: Initialization & Registration

(1) Setup: The regulatory agency generates the elliptic curve parameters G and the generator g, and publishes its public key PK_{Reg} for subsequent identity encryption. Meanwhile, a smart contract is deployed on the blockchain to verify ZK proofs4.

(2) User Registration:

① Key Generation:User Ueferencing Schnorr's key generation algorithm [5], enerates a private key $x \in Z_q^*$ and a public key $Y = g^x$ locally.

② Identity Anchoring：The user must encrypt their real identity information ID_{real}(e.g., ID number)using the regulator's public key to generate the identity ciphertext $C_{ID} = Enc(ID_{real}, PK_{Reg})$.

On-chain:The user sends $\{Y, C_{ID}\}$ to the smart contract.After the contract verifies it as correct, it records the user's public key Yin the Merkle Tree, regarding them as a legally registered user.

3.3.2 Phase 2: Privacy-Preserving Query

This is the core of the "anti-human flesh search" mechanism. When a querier wants to access sensitive data, they must prove they are a registered user without exposing their specific identity. To prevent third parties from reverse-deducing the querier's identity by analyzing permanent on-chain records, this study introduces the BDVP (Blockchain Designated Verifier Proof) mechanism [2].

(1) Step 1：Commitment Generation: The querier selects a random number k and calculates the commitment $R = g^k$. To enhance privacy, the Chameleon Hash function $CH(\cdot)$[6], is introduced here to package the commitment as $H_{cham} = CH(R, PK_{Reg})$.Since the regulatory agency holds the key, it has the ability to forge hash collisions. This means that for a third party, the proof on the chain could be generated by a user or simulated by the regulator. This "non-transferability" cuts off the tracking path for external attackers against query behaviors [2].

(2) Step 2：Non-Interactive Proof Generation: To adapt to the non-interactive environment of the blockchain, the querier utilizes the Fiat-Shamir transformation [7] to generate the proof:

Calculate Challenge：$c = Hash(H_{cham}, Y, Message)$

Calculate Response:$s = k + c \cdot x (mod q)$

The final generated zero-knowledge proof is $\pi = \{H_{cham}, c, s\}$.

(3) Step 3：On-chain Verification: The querier sends the query request and the proof π to the smart contract. The contract performs the following verification:：

$$g^s \cdot Y^{-c} \stackrel{?}{=} R'$$

And verifies whether $CH(R', PK_{Reg})$ equals H_{cham}. If the verification passes, the contract releases the data access permissions.

3.3.3 Phase3: Regulation and Tracing

When a malicious "human flesh search" or data abuse event occurs, the regulatory agency intervenes to perform de-anonymization operations.

(1) Locate Transaction：Based on clues provided by the victim, the regulatory agency locks the suspicious query transaction $Tx_{suspicious}$ on the blockchain.

(2) Decryption and Identity Reconstruction: The regulatory agency cannot decrypt it alone and must initiate a (t,n) threshold decryption protocol.

1) Proposal Initiation：An authorized node in the regulatory network (e. g., the court) initiates an accountability proposal and publicizes the hash value of the malicious transaction.

2) Shard Aggregation：When more than the threshold t egulatory nodes (e. g., public security, audit, etc.) sign and confirm the proposal, they each submit their private key shards sk_i to a Secure Multi-Party Computation (MPC) environment or smart contract.

3) Identity Restoration：The system utilizes Lagrange Interpolation to synthesize the decryption factor without revealing the complete private key SK_{Reg} and decrypts the C_{ID} to finally restore the malicious user's real identity ID_{real}.

4 Security Analysis and Evaluation

This chapter evaluates the security of the proposed anti- "human flesh search" system from four dimensions: data privacy, identity unlinkability, attack resistance, and forward security.

4.1 Data Confidentiality and Integrity

Although traditional blockchain systems possess characteristics of openness and transparency, they struggle to directly satisfy the confidentiality requirements of private data[1]. This system resolves this issue through a hybrid encryption mechanism.

4.1.1 Confidentiality

We adopt a combination of symmetric encryption (AES) and asymmetric encryption (RSA) to ensure that sensitive data is visible only to the regulatory agency or data owners who possess the specific private keys.

4.1.2 Integrity

The combination of off-chain storage (IPFS) and on-chain hash anchoring can effectively prevent data tampering [4]. Any malicious modification to off-chain data will result in a change in its hash value (CID), thereby preventing it from passing the verification of the on-chain smart contract; this utilizes the blockchain's inherent tamper-proof characteristic [1].

4.2 Anonymity and Unlinkability

Since blockchain systems provide "Pseudonymity," attackers can still deduce user identities through transaction graph analysis, making it impossible to achieve true "Unlinkability"[1]. This system compensates for this defect through the following mechanisms:

4.2.1 Application of Zero-Knowledge Proofs

Utilizing the non-interactive proof of the Schnorr protocol allows queriers to complete permission verification without transmitting any private keys or plaintext identity information over the network, thereby cutting off the direct link between identity and behavior [3].

4.2.2 Non-transferability of Proofs

This is the key to the system's defense [2].By using the BDVP scheme and utilizing the trapdoor property, the verification records retained on the chain possess "forgeability" (from the perspective of a third party). This means that a third party cannot confirm whether the record actually occurred or was simulated, thus thoroughly preventing tracking attacks targeting query history.

4.3 System Robustness

4.3.1 Anti-DDoS Attacks

Benefiting from the decentralized network topology of the blockchain, the system naturally inherits the ability to withstand Distributed Denial of Service (DDoS) attacks. Single points of failure or attacks targeting specific nodes will not cause the paralysis of the entire verification service, ensuring the high availability of the system [1].

4.3.2 Anti-Replay Attacks

In the ZKP generation process, the protocol introduces a random Challenge and a timestamp mechanism[3]. This "challenge-response" mode ensures the uniqueness of each query request, preventing malicious attackers from intercepting old proofs for illegal replay [3].

5 Conclusion and Future Work

5.1 Conclusion

Addressing the increasingly severe risk of "human flesh search" in social networks, this paper designs a privacy protection system based on blockchain and zero-knowledge proofs. It introduces IPFS off-chain storage and AES-RSA hybrid encryption mechanisms. The core protocol layer adopts the improved Schnorr protocol and Chameleon Hash technology to construct an identity verification process with "non-transferability," blocking the reverse tracking path of third parties against the querier's identity. At the same time, the design of the regulatory tracing mechanism fills the gap of lacking accountability means in decentralized networks.

5.2 Future Work

Although this system possesses high security and feasibility in theory, it still faces challenges in actual large-scale deployment. Future work will focus on the following two aspects:

5.2.1 Performance Optimization

Introduce Layer 2 scaling technologies (such as ZK-Rollup) to reduce the Gas fees of on-chain verification and improve the system's high-concurrency processing capability.

5.2.2 Post-Quantum Upgrade

Further explore post-quantum signature schemes based on lattice cryptography (Module-SIS) [2] to cope with potential threats of future quantum computing to existing cryptographic systems.

References

[1] R.Zhang, R. Xue, and L. Liu, "Security and Privacy on Blockchain," *ACM Computing Surveys*, vol. 52, no. 3, Article 51, 2019.

[2] P.-W. Chi, Y.-H. Lu, and A. Guan, "A Privacy-Preserving Zero-Knowledge Proof for Blockchain," IEEE Access, vol. 11, pp. 85108-85117, 2023.

[3] A. D. Dwivedi, R. Singh, U. Ghosh, et al., "Privacy preserving authentication system based on non-interactive zero knowledge proof suitable for Internet of Things," Journal of Ambient Intelligence and Humanized Computing, vol. 13, pp. 4639–4649, 2022.

[4] Y. Song and R. Feng, "A survey on applications of zero-knowledge proof in blockchain," Journal of Guangzhou University (Natural Science Edition), vol. 21, no. 4, pp. 21-36, 2022.

[5] C. P. Schnorr, "Efficient identification and signatures for smart cards," in *Advances in Cryptology—CRYPTO'89*, New York, NY, USA: Springer, 1990, pp. 239–252.

[6] H. Krawczyk and T. Rabin, "Chameleon hashing and signatures," in *Network and Distributed System Security Symposium (NDSS)*, 2000, pp. 143–154.

[7] A. Fiat and A. Shamir, "How to prove yourself: Practical solutions to identification and signature problems," in *Advances in Cryptology—CRYPTO'86*, Berlin, Heidelberg: Springer, 1987, pp. 186–194.

[8] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," Decentralized Business Review, 2008.

[9] V. Buterin, "Ethereum: A next-generation smart contract and decentralized application platform," 2014.

[10] J. Benet, "IPFS-content addressed versioned P2P file system," arXiv preprint arXiv:1407.3561, 2014.

[11] S. Goldwasser, S. Micali, and C. Rackoff, "The knowledge complexity of interactive proof-systems," in *Proceedings of the 17th Annual ACM Symposium on Theory of Computing*, 1985, pp. 291–304.

[12] E. Ben-Sasson et al., "Zerocash: Decentralized anonymous payments from Bitcoin," in *2014 IEEE Symposium on Security and Privacy*, 2014, pp. 459–474.