



UNIVERSITÀ DI PISA

Pesaresi Seminar Series

Blockchain-based Content Attribution for Multimedia

Evaluation and Open Challenges

Syed Muzakkir Reza Sabri

Dipartimento di Informatica · Università di Pisa, Italy

muzakirr.sabri@phd.unipi.it



About Me



muzakkirr.sabri@phd.unipi.it



Office: Room 301



Who I Am

PhD Student in Computer Science
Università di Pisa, Italy



Research Focus

Blockchain & Distributed Ledger Technology (DLT)



Interests

AI in Blockchain
Digital Sovereignty · Data Privacy

Talk Overview

01 The Problem

Why multimedia security matters — recent breach landscape

02 Why Blockchain?

Centralized vs. decentralized systems — key dimensions

03 Proposed Framework

Architecture, dual-layer encryption, zero-trust & key rotation

04 Performance Results

Throughput, latency, stability — empirical evaluation

05 Open Challenges

Scalability, PQ cryptography, AI, cross-chain, GDPR & UX

06 Roadmap & Conclusions

Future directions and research opportunities

The Problem: Why Multimedia Security?

Recent high-profile breaches (2020–2025)

2.27 TB

Medical records stolen

HCRG Care Group, 2025

3.2B

Records leaked

National Public Data, 2023

81M+

Patient records exposed

Change Healthcare, 2024

815M

Personal records breached

ICMR India, 2023

Root Causes of the Crisis

- Centralized storage = single point of failure and attack
- Third-party platforms aggregate data without owner consent
- Lack of fine-grained access control and auditability
- No verifiable content authenticity or tamper-detection

Refs: [6] PhotoSquared 2020 · [7] NPD 2023 · [8] ICMR 2023 · [9] Change Healthcare 2024 · [10] HCRG 2025

Also Noteworthy:

PhotoSquared App Breach (2020)

Thousands of private user photos exposed in plain text

T-Mobile Data Breach (2021)

54M+ customer records exfiltrated

LastPass Vault Theft (2022)

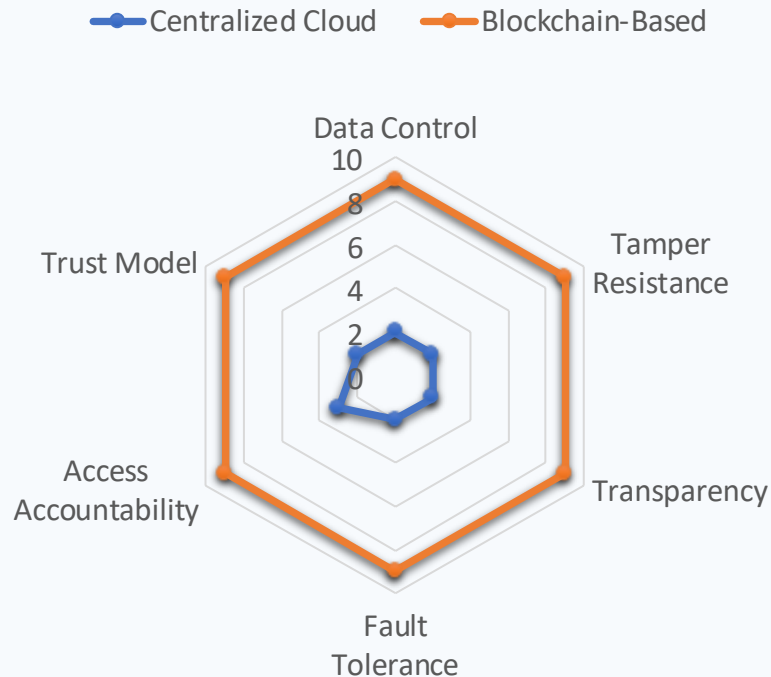
Encrypted password vaults & metadata stolen

MOVEit Transfer Attack (2023)

Multi-org breach via zero-day, 60M+ affected

Why Blockchain for Multimedia?

Multi-dimensional comparison: Centralized Cloud vs. Blockchain-Based



Data Control

Provider-owned

User-sovereign

Tamper Resistance

Low (mutable DB)

High (immutable ledger)

Transparency

Opaque

Fully auditable

Trust Model

Trust the provider

Trustless / Zero-Trust

Access Accountability

Limited logs

On-chain audit trail

Centralized

Blockchain

Research Overview



Hyperledger Fabric

Permissioned blockchain for metadata, access control & immutable audit logs



Zero-Trust Security

Dynamic trust scoring — never trust, always verify every request



IPFS Storage

Decentralized off-chain storage for encrypted multimedia with CID linking



Periodic Key Rotation

Forward secrecy via time-bounded cryptoperiods & automated re-encryption



Dual-Layer Encryption

AES-256-GCM + ChaCha20-Poly1305 for robust end-to-end confidentiality



SHA3-512 Integrity

Hash-Before-Flight protocol to detect tampering at ingestion and retrieval

System Architecture

End-to-end data flow across layered components

Frontend

User Interface — upload, retrieve, manage media

Backend

Zero-Trust Access Gateway · Encryption/Decryption · Session Management · Key Rotation Service

Blockchain Layer

Smart Contracts · Access Control · Metadata: CID, Hash, Owner DID, Timestamps · ACL · Rotation Events

Storage Layer

Encrypted Multimedia Objects (CID-addressed) · Session Tokens · Activity Logs · Operational Monitoring

All layers communicate via authenticated calls enforced by Zero-Trust principles

Dual-Layer Encryption & Integrity

Why encrypt twice? Defense-in-depth cryptographic strategy



✓ Advantages of Dual Encryption

- Independent cipher layers — if one is broken, data stays protected
- Block cipher (AES) + Stream cipher (ChaCha20): complementary attack resistance
- Authenticated encryption: both layers include MAC tags preventing tampering
- Effective security: $s_{\text{eff}} = \min\{s_{\text{AES}}, s_{\text{ChaCha}}\}$ — conservatively bounded
- Ideal for heterogeneous environments (hardware-constrained & high-throughput)

⚠ Trade-offs to Consider

► Computational overhead

Two encryption passes increase CPU usage — mitigated by ChaCha20's speed

► Key management complexity

Two key pairs (K₁, K₂) require secure storage & rotation — handled by KMS

► Larger key rotation cost

Re-encryption on rotation processes double the layers — bounded by design

Note: Advantages far outweigh costs in high-risk decentralized environments.

Zero-Trust Security Architecture

"Never trust, always verify" — coined by Stephen Paul Marsh, 1994

Trust Score Computation

$$\text{TrustScore}(u,d,r) = \alpha C(u) + \beta D(d) + \gamma A(r) - \delta R(u,d,r)$$

$$\text{Grant}(u,d,r) = 1 \text{ [TrustScore}(u,d,r) \geq \theta \text{]}$$

$C(u)$ = Credential assurance · $D(d)$ = Device posture

$A(r)$ = Access-context quality · R = Estimated risk · θ = Tunable threshold



MFA at Login

Multi-factor auth
establishes node
identity



Time-Bound Tokens

Session tokens guard
against replay attacks



Micro- Segmentation

Need-to-know access
per file/record



Forensic Logging

All grants & denials
recorded on-chain

Probabilistic Access Model

$$P(\text{Grant}=1) = 1 - F_N(\theta; \mu, \sigma^2)$$

Tunable threshold θ balances usability vs. strictness in enforcement. Higher θ = stricter security. Lower θ = more accessible but higher risk.

Smart Contract Access Control

$\text{Permit}(u,r,f) = 1$ iff:

- ① $\text{role}(u) \in R_f$ (role check)
- ② $A_u \models P_f$ (attribute check)
- ③ $\tau < T_{\text{valid}}$ (token valid)

All denials & grants logged on-chain → tamper-evident audit trail

Hash-based Integrity



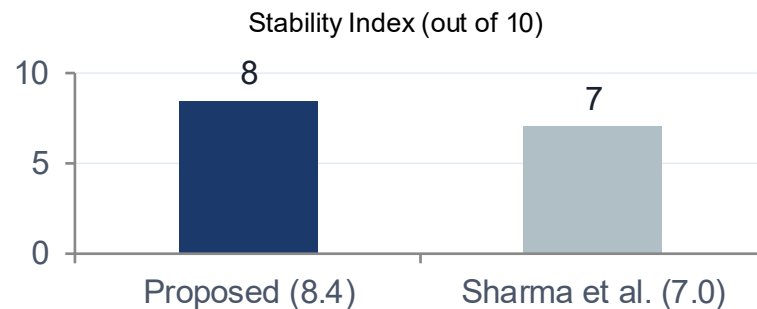
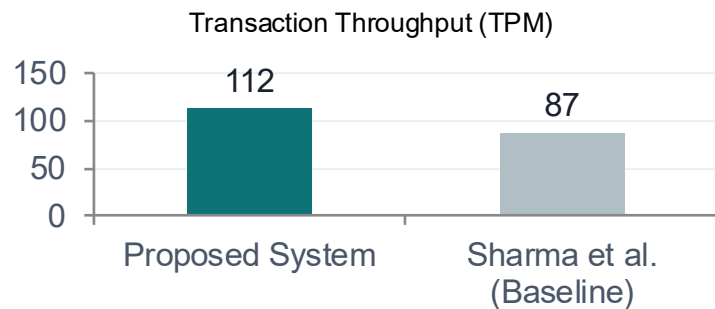
$$H(m) = \text{SHA3-512}(m)$$

Computed BEFORE upload
Stored ON-CHAIN
Recomputed on RETRIEVAL
Mismatch → Integrity Alert

$$H_o = \text{SHA3-512}(m_{\text{raw}}) \leftarrow \text{Hash-Before-Flight}$$

Performance Evaluation

Empirical benchmarks under controlled multimedia workloads



- ✓ **1.2–1.3×** Throughput improvement over Sharma et al. baseline
- ✓ **~20%** Stability index improvement (8.4 vs 7.0 / 10)
- ✓ **25%** Reduction in peak CPU utilization
- ✓ **1.5–2.2s** Bounded commit latency across 50–500 MB file sizes
- ✓ **400 TPS** Token verification peak throughput (vs. 335 TPS by Liu et al.)

Ref: Baseline: Sharma et al. [11] · Liu et al. [13] · Chen et al. [12]

Literature: Gaps & Open Challenges

From the comparative analysis — what the field is STILL missing:



Scalability Bottlenecks

Most works (Sharma, Liu, Nalina) face scalability walls under high concurrent loads or large peer counts



Limited Real-World Deployment

Architectures are evaluated in controlled lab conditions — geographically distributed production deployments remain untested



No Temporal Security Vectors

Liu et al., FileWallet and others lack periodic key rotation or any forward-secrecy mechanism



Computational Overhead (SMPC/SSI)

Liu et al.'s SSI/DID approach introduces prohibitive on-chain overhead unfit for multimedia workloads



No Unified Security + Performance Metrics

Few works report quantitative stability indices alongside cryptographic depth simultaneously



Static Key Management

14 out of 16 surveyed frameworks use static key management — no rotation, no forward secrecy

Challenge 1: Scalability & Dynamic Load Balancing

Current system bottlenecks and proposed solutions

⚡ *Current State: Latency grows from ~35s → ~70s as peer count triples (1→3 peers). IPFS performance is hardware-dependent and degrades with payload size.*

Kubernetes Peer Orchestration

Auto-scale endorsing peers based on transaction pressure. Reduces latency spikes during multimedia upload bursts.

Research Needed

Channel Sharding

Partition Fabric into domain-specific channels (healthcare, education, DRM) to isolate transaction processing.

Partially Explored

Adaptive IPFS Pinning

Predictive prefetching based on access analytics — pre-pin hot CIDs to geographically close nodes before demand peaks.

$A_CID = 1 - (1 - P_node)^n$

Open Problem

Hierarchical Ledger Caching

Off-chain LRU-cache for recently accessed metadata reduces on-chain query pressure. CouchDB/LevelDB optimization.

Open Problem

Challenges 2: Post-Quantum Cryptography & Cross-Chain Interoperability

Two frontier challenges in blockchain-based multimedia security

Post-Quantum Cryptographic Agility

Shor's Algorithm poses a long-term threat to classical ciphers. NIST (2024) standardized ML-KEM & ML-DSA — but blockchain integration is wide open.

- Replace AES key exchange with CRYSTALS-Kyber (ML-KEM-768) — adds ~1.1 KB overhead per transaction
- Hash-based signatures for chaincode — 8-50× larger but quantum-resistant
- Hybrid classical-PQ migration: retain ChaCha20+AES for performance, add Kyber key-wrapping layer
- Quantum-resistant ZK proofs (STARKs) for privacy-preserving access without exposing identity

Cross-Chain Interoperability & DRM

Multimedia rights span multiple organizations and blockchains. A closed Fabric network cannot operate cross-jurisdictionally without bridges.

- Atomic cross-chain swaps (HTLCs) to transfer ownership rights between Fabric and Ethereum/Polygon
- Publish SHA3-512 fingerprints to a public chain for third-party verifiable provenance
- DID-anchored DRM: creator DID + content DID + license DID form a verifiable credential chain
- Smart contract royalty automation: on-chain payments split among co-creators per pre-agreed rules

Challenge 3: AI-Driven Threat Detection & Adaptive Access

Intelligent, context-aware security for blockchain-multimedia systems



AI Threat Detection

- LSTM / Transformer models trained on on-chain access patterns to detect anomalous credential usage in real-time
- Graph Neural Networks on transaction graphs — identify colluding insiders attempting lateral movement
- Federated learning across Fabric orgs: improve threat models without sharing raw audit data
- Auto-recalibrate trust threshold θ based on detected threat level — closes the manual threshold gap



Context-Aware Access Control

- Dynamic ABAC: access rights adapt based on role + location + time-of-day + behavioral history
- TrustScore(u, d, r) recomputed continuously — not just at login — addresses session hijacking
- Risk-proportional MFA: additional auth only when risk exceeds threshold — reduces friction
- ZK-SNARK attribute proofs: prove an attribute without revealing identity or content

Challenges 5: Regulatory Compliance & User-Centric Deployment

Bridging the gap between technical security and real-world usability

GDPR & Right to Be Forgotten

Core tension: Blockchain immutability vs. GDPR Art. 17 Right to Erasure

► Cryptographic Erasure

Encrypt permanently; destroy decryption key on deletion — renders data unreadable

► Redactable Blockchain

Chameleon hash-based mutable blocks for authorized on-chain metadata redaction

► GDPR Smart Contracts

Auto-trigger re-encryption or key destruction when consent expires

► Jurisdictional Residency

Enforce CIDs only on IPFS nodes in compliant EU geographic regions

Ref: GDPR Art. 17

User-Centric Metrics & Real-World Deployment

Technical benchmarks are necessary but insufficient — UX is where most systems fail

► Perceived Latency

2.2s lab ≠ user experience. Partial failures push perceived latency to 10–30s. User studies needed.

► Key Management UX

DID wallets for Hyperledger are nascent. Passkey/WebAuthn integration is an open engineering challenge.

► Partial Network Resilience

Graceful degradation when orderers or Pinata pins fail — no formalized availability SLA yet.

► Total Cost of Ownership

IPFS pinning + Kaleido + key rotation compute costs at millions-of-files scale remain unmodeled.

References

- [1] Kim, L. (2014). iCloud hack exposes private photos of celebrities.
- [2] Smith, R.H. (2013). Experian data breach exposes 15 million SSNs. Reuters.
- [3] Zetter, A. (2013). Hackers steal 38 million Adobe accounts. Wired.
- [4] Greenberg, A. (2016). Hackers breach FriendFinder Network, 412 million accounts.
- [5] Rawat, P. (2018). India's Aadhaar system exposed 1.1 billion biometric records.
- [6] Seals, A. (2020). Data breach at PhotoSquared exposes thousands of private photos.
- [7] BBC News (2023). National Public Data leak exposes millions of records.
- [8] Rath, M. (2023). ICMR data breach exposes sensitive medical records.
- [9] Gupta, M. (2024). Ransomware attack on Change Healthcare.
- [10] Page, C. (2025). HCRG confirms hack — 2.27 TB sensitive data stolen. TechCrunch.
- [11] Sharma et al. (2020). Blockchain technology for cloud storage: A systematic review.
- [12] Chen et al. (2022). FileWallet: File management system based on IPFS and Hyperledger Fabric. CMES.
- [13] Liu et al. (2021). Blockchain platform for multimedia data management. Multimedia Tools & Applications.
- [14] Dobre et al. (2018). Authentication of JPEG images on the blockchain. ICCAIRO.
- [15] Sharma et al. (2024). Blockchain-based distributed application for multimedia. Multimedia Tools & Applications.
- [16] Qureshi & Jiménez (2021). Blockchain-based multimedia content protection: Review and open challenges.
- [17] Nalina et al. (2024). Decentralized file storage using IPFS and blockchain. ICETCS.
- [18] Awale & Gaikwad (2023). Zero Trust Architecture using Hyperledger Fabric. ICCCNT.
- [19] Bhowmik & Feng. The multimedia blockchain: A distributed and tamper-proof media transaction framework.
- [20] Zyskind et al. (2015). Decentralizing privacy: Using blockchain to protect personal data. IEEE SPW.
- [21] Taloba et al. (2023). Blockchain-based hybrid platform for multimedia in IoT-healthcare. Alexandria Eng. J.
- [22] Ali et al. (2018). A blockchain-based decentralized data storage framework. IEEE TrustCom.
- [23] Vishwa & Hussain (2018). A blockchain-based approach for multimedia privacy protection. IEEE.
- [24] Kumar et al. (2024). Blockchain: From Cryptocurrency to Decentralized Revolution.
- [25] Khan et al. (2021). MF-ledger: Blockchain Hyperledger Sawtooth for multimedia forensics. IEEE Access.
- [26] Hasan & Salah (2018). Proof of delivery of digital assets using blockchain. IEEE Access.
- [27] Huang et al. (2022). A secure file sharing system based on IPFS and blockchain.



UNIVERSITÀ DI PISA

Thank You

Questions & Discussion

Syed Muzakkir Reza Sabri

Dipartimento di Informatica · Università di Pisa, Italy

muzakkir.sabri@phd.unipi.it

Office: Room 301 · Pesaresi Seminar Series · University of Pisa, Department of Computer Science

Presented in memory of Mauriana Pesaresi — PhD Student, Dept. Computer Science, Università di Pisa