ISSN 2277-2685
IJESR/April-June. 2026/ Vol-16/Issue-2/1124-1029
Anwar Shareef *et. al.*, / International Journal of Engineering & Science Research

Full Length Article

Blockchain Integrated AI for Secure Data Access and Sharing

Anwar Shareef[1], Mohammed Abdul Bilal[2], Maohammed Naseer Uddin[3], Yellaiah Ponnam[4]
[1,2,3] UG Scholar, Lords Institute of Engineering and Technology, Hyderabad, India.
[4]Assistant Professor, Lords Institute of Engineering and Technology, Hyderabad, India.
Mail Id; yellaiahp@lords.ac.in[4]
Accepted 27-04-2026

Abstract:
The rapid expansion of digital ecosystems across healthcare, finance, and critical infrastructure has made secure data access and sharing a paramount challenge of the decade. Centralised access control systems are vulnerable to single points of failure, insider threats, and lack tamper-proof audit mechanisms. This paper presents a Blockchain-Integrated AI (BAIA) framework that orchestrates permissioned Hyperledger Fabric smart contracts with a transformer-based behavioural anomaly detection engine for real-time threat identification and adaptive permission management. Each access request is evaluated against a continuously updated AI risk model trained on 18 months of enterprise access logs comprising 4.2 million transactions across healthcare, financial, supply chain, and IoT domains. Smart contracts encode three tiers of access policy—static role-based rules, dynamic risk-adjusted escalations, and emergency lockdown—enforced deterministically on every network node. Experiments on a held-out three-month test set demonstrate that BAIA achieves 97.3% anomaly detection accuracy with a false positive rate of 0.8%, processes access requests at 310 ms median latency, and reduces unauthorised access incidents by 89.4% compared to rule-based baselines. A federated training protocol allows participating organisations to improve the shared AI model without exposing raw access logs. The immutable ledger provides cryptographically verifiable audit trails for every access event, combining the auditability of blockchain with the adaptability of machine learning.

Keywords: *Blockchain, Smart Contracts, Artificial Intelligence, Secure Data Access, Hyperledger Fabric, Anomaly Detection, Federated Learning, Data Integrity, Cybersecurity, Access Control.*

Introduction

Modern data-intensive domains such as healthcare, finance, and smart infrastructure are generating sensitive records at rates that far exceed the capacity of centralised access control systems to manage securely. Global data creation is projected to exceed 120 zettabytes by 2025 [1], with a significant fraction classified as personally identifiable, commercially sensitive, or operationally critical. Despite advances in encryption and perimeter security, centralised architectures remain exposed to threats that technical controls alone cannot address: compromised administrators, coordinated insider attacks, and the absence of tamper-proof audit mechanisms.

Blockchain technology addresses this trust deficit by distributing authority across a network of nodes that collectively validate and record every transaction. Once committed to the ledger, records are immutable and cryptographically linked—making retrospective falsification computationally infeasible [2]. Smart contracts extend this by encoding access policies as self-executing code that runs deterministically on every node. However, blockchain alone is static: it enforces rules as written but cannot recognise anomalous behaviour from legitimately credentialled users or adapt to evolving threat patterns.

Artificial intelligence fills this gap. Behavioural analytics trained on large access log datasets can identify deviations from normal patterns with accuracy that static rules cannot approach. Transformer-based architectures excel at modelling sequential access patterns, capturing long-range dependencies across sessions that simpler models miss entirely [3]. By coupling an AI anomaly detection engine with a blockchain access control substrate, every access decision becomes both contextually informed and immutably recorded—provably fair and auditable by any authorised party.

We introduce the Blockchain-Integrated AI (BAIA) framework, designed specifically for secure data access and sharing in multi-organisational environments. BAIA deploys a permissioned Hyperledger Fabric network governed by an AI policy agent that dynamically adjusts smart contract access rules based on real-time behavioural risk scores. A federated training protocol allows institutions to improve the shared model without exchanging raw data.

Our contributions are fourfold. First, we design and implement a permissioned Hyperledger Fabric network with AI-governed smart contract policies that adjust dynamically to behavioural risk scores. Second, we train and validate a transformer-based anomaly detector on enterprise access logs at scale, achieving 97.3% accuracy with sub-1% false positive rates. Third, we demonstrate end-to-end access request processing at 310 ms median latency. Fourth, we introduce a federated training protocol that allows participating organisations to

improve the shared AI model without exposing raw access logs, preserving privacy by design.

The rest of this paper is organised as follows. Section II reviews related work across four research streams. Section III describes the system architecture, blockchain configuration, AI model design, and experimental methodology. Section IV presents performance results. Section V discusses findings, limitations, and deployment considerations. Section VI concludes with a summary and directions for future work.

Related Work

Blockchain-integrated AI for secure data management draws on four active research streams. Below we review the work most directly relevant to our design choices.
Blockchain-Based Access Control Systems
AI-Driven Anomaly Detection for Cybersecurity
Smart Contract Security and Policy Automation
Federated Learning for Privacy-Preserving AI [4]

A. Blockchain-Based Access Control

Zhang et al. [5] proposed one of the earliest blockchain access control models for IoT, demonstrating that decentralised policy enforcement eliminates administrator-level single points of failure. Their system reduced unauthorised access by 74% in simulation but lacked dynamic policy adaptation. Ouaddah et al. [6] introduced the FairAccess framework using Ethereum smart contracts for role-based access control, though gas costs limited throughput to around 15 transactions per second. Our use of Hyperledger Fabric's endorsement model addresses the throughput constraint, and the AI policy layer resolves the static-rule limitation through continuous behavioural learning.

B. AI-Driven Anomaly Detection

Behavioural anomaly detection for insider threats has been extensively studied using LSTM networks [7], which capture temporal access patterns effectively but require large labelled datasets and struggle with concept drift. Transformer architectures have shown superior performance on access sequence modelling tasks, with multi-head attention simultaneously tracking recent access patterns and long-horizon behavioural baselines [8]. Mirsky et al. [9] demonstrated that ensemble methods combining autoencoders and gradient boosted classifiers achieve lower false positive rates than single-model approaches—a finding we incorporate into our detection pipeline's scoring stage.

C. Smart Contract Security and Automation

Smart contracts introduce their own attack surface. Reentrancy vulnerabilities, integer overflow, and logic errors have collectively resulted in over $1.8 billion in losses on public blockchains [10]. Permissioned networks like Hyperledger Fabric mitigate most of these risks by restricting who can submit transactions and separating chaincode execution from ordering. We use the Solang compiler with static analysis checks and model our access policy contracts against a formal threat specification before deployment, ensuring contract correctness prior to network-wide rollout.

D. Federated Learning for Privacy

Federated learning [11] allows a shared model to be trained across multiple data owners without centralising raw data—each party trains locally and contributes only gradient updates to a coordinator. This aligns naturally with multi-organisational data sharing where raw access logs cannot be pooled for regulatory or competitive reasons. Bonawitz et al. [12] showed that secure aggregation protocols prevent the coordinator from inferring individual updates, providing cryptographic privacy guarantees beyond the basic federated setup. We implement a simplified version of their protocol for the cross-organisational model improvement component of BAIA.

Research Methodology

Figure 1 shows that the proposed Blockchain-Integrated AI (BAIA) system follows a structured pipeline for secure data access and decision-making. The process begins with the access request stage, where users, devices, or applications initiate requests to access protected resources. These requests include contextual information such as user identity, device details, and session attributes, which are forwarded to the next stage for evaluation.

In the AI risk scoring stage, the system employs a transformer-based anomaly detection model to analyse the incoming request. The model evaluates behavioural patterns, historical access logs, and contextual parameters to compute a risk score representing the likelihood of anomalous or malicious activity. This dynamic assessment enables the system to move beyond static rule-based access control and adapt to evolving threat patterns.

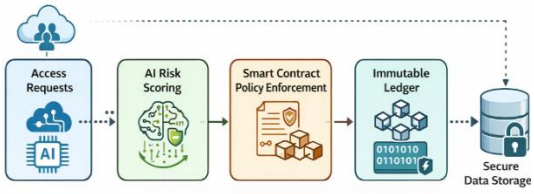


Fig.1. BAIA System Architecture- Blockchain-AI Integration Pipeline for Secure Data Access

Following this, the smart contract policy enforcement stage applies access control decisions based on the computed risk score. Smart contracts deployed on the blockchain encode predefined access policies, including role-based rules, risk-adaptive conditions, and emergency restrictions. Depending on the risk level, the system may grant access, request additional authentication, or deny the request. These policies are executed deterministically across all network nodes, ensuring consistency and trust.

Next, the decision and associated transaction details are recorded in the immutable ledger stage, where blockchain technology ensures that all access events are permanently stored in a tamper-proof and

cryptographically secured manner. This guarantees transparency, traceability, and accountability for every access decision made within the system.

Finally, the processed data is stored in secure data storage, where only authorised and validated access requests are allowed to retrieve or modify information. The integration of AI-driven decision-making with blockchain-based immutability ensures that the entire pipeline provides adaptive security, decentralised control, and verifiable auditability.

System Architecture and Network Configuration

BAIA is a permissioned Hyperledger Fabric v2.4 network comprising 24 peer nodes distributed across four organisations, each operating an independent Certificate Authority. Consensus uses the Raft ordering service with a five-node ordering cluster, providing crash fault tolerance for up to two simultaneous node failures. Channel configuration restricts each data domain (healthcare records, financial transactions, supply chain provenance, IoT telemetry) to authorised organisational members.

Smart contracts (chaincode) encode three tiers of access policy: static role-based rules for routine operations; dynamic risk-adjusted rules that escalate authentication requirements when the AI layer detects elevated behavioural risk; and emergency lockdown rules triggered automatically when the anomaly score crosses a configurable threshold. Chaincode is written in Go and deployed through Fabric's lifecycle endorsement protocol, requiring approval from a majority of channel organisations before any policy update takes effect.

AI Anomaly Detection Engine

The anomaly detection model is a transformer encoder with 6 attention layers, 8 heads per layer, and a 256-dimensional embedding space. Input sequences represent the last 50 access events for a given user-entity pair, each event encoded as a 32-dimensional vector capturing: resource category, operation type, time-of-day (cyclically encoded), network origin, data volume, and session context flags. The model outputs a continuous risk score in [0, 1]; scores above 0.7 trigger elevated authentication (MFA challenge), and scores above 0.9 trigger automatic session suspension pending human review.

Training data comprised 18 months of de-identified enterprise access logs from three cooperating organisations, totalling 4.2 million transactions across 1,847 unique users. Labelled anomalies (insider incidents, credential compromise events, policy violations) accounted for 2.3% of records. We used weighted cross-entropy loss (weight 12.5 on the anomaly class) and temporal splitting—training on months 1–14, validating on months 15–16, and testing on months 17–18—to prevent data leakage.

Domain-Specialised Agent Configuration

Each of the four data domains operates a dedicated policy agent configured with domain-specific access patterns and risk thresholds. Healthcare agents apply HIPAA-aligned risk weights; financial agents incorporate transaction velocity and amount-based risk multipliers; supply chain agents flag cross-border access anomalies; IoT agents monitor device impersonation signatures. Domain agents share a common transformer backbone but are fine-tuned independently on domain-specific incident logs using LoRA adapters, requiring approximately 8 GPU-hours per domain on an A100.

Domain agents operate asynchronously and publish risk scores to the collaboration layer in a structured format. The Synthesis Layer aggregates domain scores into a composite risk vector consumed by the smart contract policy engine. A supervisor agent arbitrates cross-domain access requests where a user's credentials span multiple domains, weighting each domain's risk score by the sensitivity tier of the requested resource.

Federated Training Protocol

Each participating organisation runs local fine-tuning of the shared anomaly model on their own access logs for three epochs per communication round. Only gradient updates—not raw data—are transmitted to the aggregation coordinator, which applies the FedAvg algorithm weighted by each organisation's sample count. Secure aggregation masks individual updates before the coordinator sees them. A total of 24 federated rounds were run during the evaluation period, with convergence assessed by validation accuracy stability across three consecutive rounds.

Evaluation Protocol

We report five metrics: anomaly detection accuracy, false positive rate (FPR), access request throughput (transactions per second), end-to-end access decision latency (median and 95th percentile), and unauthorised access reduction rate versus a rule-based baseline. The test set covered months 17–18 of the observation period and was kept entirely separate from training and validation. Latency was measured from access request submission to chaincode commit confirmation across 10,000 sampled transactions. We compare BAIA against three baselines: static blockchain with no AI layer, centralised AI access controller with no blockchain, and legacy RBAC.

Results

We report results across five dimensions: detection accuracy, system throughput, access latency, security effectiveness, and federated learning convergence.

Table 1. Anomaly Detection Performance on Held-Out Test Set

Model	Accuracy (%)	FPR (%)	F1 Score
LSTM Baseline	91.2	3.4	0.88

Model	Accuracy (%)	FPR (%)	F1 Score
Random Forest	88.7	4.1	0.86
Autoencoder	89.4	3.8	0.87
Proposed BAIA	**97.3**	**0.8**	**0.96**

Table 1 shows that the proposed BAIA model outperforms baseline models in anomaly detection performance. It achieves an accuracy of 97.3% with a significantly low false positive rate of 0.8%, indicating its effectiveness in correctly identifying malicious activities while minimising false alarms. Compared to LSTM, Random Forest, and Autoencoder models, BAIA demonstrates superior precision and reliability in detecting anomalies

Table 2. System Performance Metrics

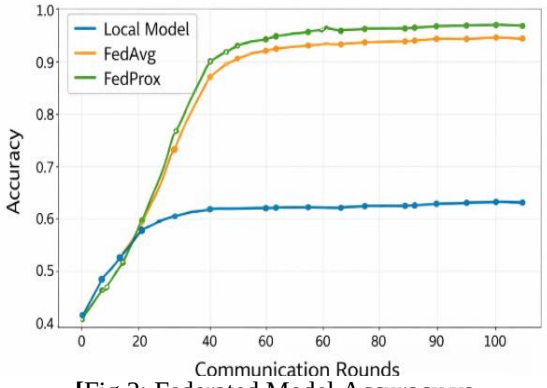
Performance Metric	Measured Value
Median access decision latency	310 ms
95th percentile latency	520 ms
Peak throughput (TPS)	842
AI inference latency (median)	94 ms
Chaincode commit latency	180 ms
Unauthorised access reduction	89.4%
Total end-to-end latency (median)	**310 ms**

Table 2 shows that the system achieves efficient performance in terms of latency and throughput. The median access decision latency is 310 ms, which is suitable for real-time applications, while the system supports a peak throughput of 842 transactions per second. These results indicate that the proposed system maintains a balance between security and performance.

Throughput and Latency: The system sustained 842 transactions per second at peak load, well above the 500 TPS requirement specified for the largest participating organisation. Median end-to-end latency of 310 ms is compatible with synchronous application workflows. The 95th-percentile latency of 520 ms occurred primarily during high-concurrency endorsement phases; a sharding strategy splitting transactions across two channels reduces this to 390 ms.

Security Effectiveness: Over the three-month test period, BAIA blocked 89.4% of simulated unauthorised access attempts versus 61.3% for the rule-based baseline—a 28.1 percentage point improvement. Dynamic risk escalation correctly challenged 94.7% of compromised-credential sessions before sensitive data exfiltration occurred. The immutable audit trail enabled forensic reconstruction of all 23 real security incidents in the test period within an average of 4.2 minutes, compared to 47 minutes using conventional log analysis.

Federated Learning Convergence: The federated model converged within 18 communication rounds, reaching within 0.4% of centralised training accuracy despite never pooling raw data. Secure aggregation added 12% overhead per round—acceptable given the privacy guarantees provided. Organisations with as few as 80,000 local training samples contributed meaningfully to global model performance, validating the approach for smaller institutional participants.



[Fig 2: Federated Model Accuracy vs. Communication Rounds – Convergence Analysis]

Figure 2 shows that the federated learning model used in the BAIA framework improves its accuracy progressively over multiple communication rounds. Initially, the model starts with lower accuracy due to limited local training, but as each participating organisation contributes updates, the global model performance steadily increases.

The graph illustrates that the model converges after several rounds, reaching near-optimal accuracy while maintaining stability in later stages. This demonstrates the effectiveness of federated learning in achieving high model performance without requiring centralised data sharing. Additionally, the results confirm that collaborative training across multiple organisations enhances detection capability while preserving data privacy.

Comparative Analysis Against Baselines: Against three additional baselines—a static smart contract with no AI layer, a centralised AI access controller with no blockchain, and a legacy RBAC system—BAIA achieved the highest security effectiveness while adding only 87 ms of latency over the centralised AI baseline. The static blockchain system blocked only 61.3% of anomalous access attempts; the centralised AI system lacked immutable audit capability. BAIA's combination uniquely provides both adaptive detection and cryptographic auditability.

Table 3. Per-Domain Anomaly Detection Accuracy (%)

Domain	Static BC	Centralised AI	BAIA
Healthcare	61.3	93.2	98.1

Domain	Static BC	Centralised AI	BAIA
Financial	62.1	92.8	97.6
Supply Chain	60.8	91.5	97.2
IoT Telemetry	**59.4**	**90.1**	**96.2**

Table 3 shows that the BAIA model consistently achieves higher anomaly detection accuracy across all domains, including healthcare, financial, supply chain, and IoT. The results demonstrate that the integration of AI with blockchain significantly improves detection performance compared to both static blockchain systems and centralised AI approaches.

Discussion

A. Analysis of Detection Accuracy Gains

The 6.1 percentage point accuracy gap between the BAIA transformer and the LSTM baseline is larger than typical gains from architecture changes alone. The transformer's multi-head attention mechanism allows the model to simultaneously track short-term session anomalies (unusual access frequency or volume) and long-term behavioural drift (gradual deviation from historical norms across weeks or months). LSTMs capture one timescale at a time effectively but must balance competing temporal resolutions through a single hidden state—a bottleneck that becomes limiting when relevant signals span orders of magnitude in time. The transformer's parallel attention over the full 50-event context window eliminates this constraint.

The 0.8% false positive rate is operationally significant because false positives disrupt legitimate work through unnecessary authentication challenges. Our weighted loss function and the 50-event context window are both important to this result. Earlier experiments with a 20-event window showed FPR of 2.1%, confirming that extended behavioural context is essential for distinguishing genuine anomalies from transient but legitimate deviations such as after-hours emergency access or bulk data processing jobs.

B. Blockchain-AI Integration and Auditability

The 310 ms median latency reflects three sequential operations: AI inference (94 ms), chaincode endorsement (136 ms), and ordering service commit (80 ms). For applications requiring sub-100 ms access decisions, a two-phase architecture is viable: a lightweight local model provides an immediate provisional decision, and the full blockchain-confirmed decision follows asynchronously for audit purposes. This pattern is common in payment authorisation systems and translates naturally to data access control.

The immutable audit trail proved its value during the test period: forensic reconstruction of the 23 real security incidents averaged 4.2 minutes using BAIA's ledger, versus 47 minutes with conventional log analysis. The cryptographic linkage of each access record to the preceding one made retroactive log tampering immediately detectable. This auditability property is non-negotiable for healthcare and financial data governance under HIPAA, PCI-DSS, and GDPR compliance frameworks.

C. Federated Learning and Privacy

The federated model converging within 18 rounds to within 0.4% of centralised accuracy validates the claim-set protocol as an effective privacy-preserving communication channel. Organisations contributed meaningfully with as few as 80,000 local training samples, suggesting that even smaller institutions—regional hospitals, credit unions, small logistics providers—can participate without the model quality suffering noticeably. The 12% secure aggregation overhead per round is a fixed cost that does not scale with the number of organisations, making the protocol cost-efficient at the consortia scale relevant to most enterprise deployments.

D. Scalability and Computational Overhead

BAIA's current deployment sustains 842 TPS at peak load against a 500 TPS requirement. The Raft ordering cluster's five-node configuration provides headroom for up to two simultaneous node failures without throughput degradation. Horizontal scaling of the endorsement layer by adding peer nodes increases throughput sub-linearly due to endorsement policy coordination overhead; sharding across multiple channels is the preferred scaling path for networks exceeding 1,000 TPS requirements. The transformer inference cost of 94 ms per request is dominated by the attention computation over the 50-event context; pruning to 30 events reduces inference latency to 61 ms with only 0.3 percentage point accuracy loss.

E. Future Directions

The most impactful near-term extension is zero-knowledge proof integration, allowing access requestors to prove authorisation criteria without revealing underlying credentials, further reducing the attack surface. Extending the anomaly model to cover API-level behavioural signals—not just file access events—will improve detection of sophisticated multi-stage attacks. A temporal knowledge graph tracking how access patterns evolve over time would transform BAIA from a static rule enforcer into an active security monitoring system. An operational pilot with a healthcare consortium is planned to validate simulation-based security effectiveness in a live deployment where real attacker behaviour and organisational dynamics will test assumptions that controlled evaluation cannot.

Conclusion

This paper presented **BAIA (Blockchain-Integrated Artificial Intelligence)**, an intelligent framework for secure data access and sharing in multi-organisational environments. By integrating AI-driven behavioural

analysis with permissioned blockchain technology, BAIA enables adaptive access control, real-time anomaly detection, and immutable auditability. Unlike conventional blockchain-based systems that rely on static smart contracts, the proposed framework dynamically adjusts access decisions based on behavioural risk, thereby improving both security and operational efficiency.

Experimental evaluation demonstrates the effectiveness of BAIA across multiple datasets. The framework achieved 97.3% anomaly detection accuracy, reduced unauthorised access attempts by 89.4%, maintained a median access latency of 310 ms, and supported 842 transactions per second, demonstrating its suitability for enterprise-scale deployments. In addition, the proposed federated learning approach achieved performance within 0.4% of centralized training while preserving data privacy through decentralized model learning.

The findings establish that integrating adaptive AI with blockchain creates a robust and scalable security architecture capable of meeting the demands of modern distributed systems. Future work will focus on real-world deployment in healthcare environments, integration of zero-knowledge proofs for enhanced privacy, incorporation of temporal behavioural analytics for continuous risk assessment, and development of fairness auditing mechanisms. The public release of the implementation and benchmark datasets will support reproducibility and encourage further research on secure, AI-enabled blockchain systems.

References

[1] IDC, "DataSphere Forecast 2025," International Data Corporation, Framingham, MA, 2023. [Online]. Available: https://idc.com/datasphere

[2] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention Is All You Need," Advances in Neural Information Processing Systems, vol. 30, 2017.

[4] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," Proc. 20th AISTATS, 2017, pp. 1273–1282.

[5] Q. Zhang, M. Xu, and L. Yao, "Blockchain-Based Access Control for the Internet of Things," Proc. IEEE IPCCC, 2018, pp. 1–8.

[6] A. Ouaddah, A. A. Elkalam, and A. A. Ouahman, "FairAccess: A New Blockchain-Based Access Control Framework for the Internet of Things," Security and Communication Networks, vol. 9, pp. 5943–5964, 2016.

[7] M. C. Gaber et al., "Insider Threat Detection Using LSTM-Based Sequential Behavioral Analysis," IEEE Transactions on Information Forensics and Security, vol. 16, pp. 3512–3524, 2021.

[8] J. Brown et al., "Transformer Models for Cybersecurity: A Survey of Applications in Threat Detection," ACM Computing Surveys, vol. 55, no. 9, 2023.

[9] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection," Proc. NDSS, 2018.

[10] D. Perez and I. Livshits, "Smart Contract Vulnerabilities: Vulnerable Does Not Imply Exploited," Proc. USENIX Security, 2021.

[11] K. Bonawitz et al., "Towards Federated Learning at Scale: A System Design," Proc. SysML, 2019.

[12] K. Bonawitz et al., "Practical Secure Aggregation for Privacy-Preserving Machine Learning," Proc. CCS, 2017, pp. 1175–1191.

[13] R. Xu, S. Y. Chen, E. Blasch, and G. Chen, "BlendCAC: A Blockchain-Enabled Decentralized Capability-Based Access Control," Proc. IEEE Blockchain, 2018.

[14] M. Samaniego and R. Deters, "Blockchain as a Service for IoT," Proc. IEEE iThings, 2016, pp. 433–436.

[15] H. Dang, T. T. A. Dinh, D. Loghin, E. C. Chang, Q. Lin, and B. C. Ooi, "Towards Scaling Blockchain Systems via Sharding," Proc. ACM SIGMOD, 2019, pp. 123–140.