IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

BLOCKCHAIN-ENABLED LOG ANALYSIS FOR THREAT DETECTION

Varun Kumar N[1], Srinath GM[2]

[1]Post Graduation Student Master of Technology, Department of Computer Science and Engineering, SJC Institute of Technology, Chikkaballapur, Karnataka, India

[2] Assistant Professor, Department of Computer Science and Engineering, SJC Institute of Technology, Chikkaballapur, Karnataka, India

Abstract

Insider threats have become one of the most challenging cybersecurity issues, as malicious activities performed by authorized users are difficult to identify using traditional security mechanisms. Conventional log analysis techniques rely heavily on manual inspection and rule-based systems, which are often ineffective in detecting sophisticated attacks and anomalous behavior in real time. This paper presents a blockchain-enabled digital forensic framework for insider threat detection through intelligent log analysis. The proposed system combines Machine Learning algorithms, namely Isolation Forest for anomaly detection and Random Forest for threat classification, with blockchain technology to ensure secure and tamper-proof storage of forensic logs. A Flask-based web application provides real-time monitoring, visualization, and prediction, while an SMTP-based alert mechanism instantly notifies administrators whenever suspicious activities are detected. Experimental evaluation demonstrates that integrating blockchain with machine learning significantly enhances the integrity, reliability, and efficiency of forensic investigations while enabling proactive cybersecurity monitoring.

Index Terms— Digital Forensics, Blockchain, Insider Threat Detection, Log Analysis, Isolation Forest, Random Forest, Machine Learning, Cybersecurity.

I. INTRODUCTION

With the rapid growth of digital infrastructure, organizations increasingly depend on interconnected computer systems to manage critical operations and sensitive information. While external cyberattacks continue to evolve, insider threats have emerged as one of the most dangerous security challenges because authorized users possess legitimate access to organizational resources. These threats may involve intentional data theft, privilege misuse, unauthorized modifications, or accidental policy violations that compromise system security.

Digital forensic investigation plays a vital role in identifying cyber incidents by collecting and analyzing system logs generated from servers, workstations, and network devices. However, modern organizations generate millions of log entries every day, making manual forensic investigation both time-consuming

and error-prone. Traditional rule-based security systems often fail to detect unknown attack patterns or sophisticated insider activities that differ from predefined signatures.

Machine Learning provides an effective solution by automatically learning behavioral patterns from historical log data and identifying anomalous activities. Isolation Forest efficiently detects unusual events without requiring labeled datasets, whereas Random Forest accurately classifies detected events according to their threat severity. However, maintaining the integrity of forensic evidence remains a significant challenge because centralized log repositories are susceptible to unauthorized modifications and tampering.

Blockchain technology addresses this limitation by providing decentralized, immutable, and cryptographically secure storage for forensic logs. Each log entry is encrypted and permanently recorded, ensuring transparency, integrity, and traceability throughout the investigation process. The integration of blockchain with intelligent anomaly detection creates a secure and trustworthy forensic framework capable of supporting modern cybersecurity operations.

The primary objective of this research is to develop an intelligent digital forensic system capable of automatically detecting insider threats, preserving forensic evidence using blockchain, and generating real-time security alerts for administrators to minimize organizational cyber risks.

II. RELATED WORK

A. Machine Learning for Cybersecurity

Machine Learning has become an essential technology for modern cybersecurity by automating anomaly detection and threat classification. Supervised and unsupervised learning techniques are widely used to identify suspicious user behavior from large-scale security logs. Algorithms such as Decision Trees, Random Forest, Support Vector Machines, and Isolation Forest have demonstrated significant improvements in detecting abnormal activities compared to traditional signature-based intrusion detection systems.

B. Blockchain-Based Secure Log Management

Blockchain technology has recently gained considerable attention for securing digital forensic evidence. Due to its decentralized architecture and immutable ledger, blockchain ensures that forensic logs cannot be altered once recorded. Cryptographic hashing mechanisms provide data integrity, while distributed consensus prevents unauthorized modifications, making blockchain highly suitable for forensic investigations.

C. Insider Threat Detection through Log Analysis

Several researchers have proposed log analysis techniques for identifying malicious insider activities using behavioral analytics. Event logs containing user authentication records, IP addresses, timestamps, device information, and system events provide valuable insights into abnormal behavior. Machine Learning algorithms improve detection accuracy by identifying deviations from normal activity patterns without relying solely on predefined security rules.

D. Real-Time Threat Monitoring

Real-time monitoring systems enable organizations to respond quickly to cybersecurity incidents by continuously analyzing incoming log streams. Email notification services, dashboard visualization, and automated alert generation significantly reduce response time and minimize the impact of insider attacks. Integrating machine learning with continuous monitoring enhances proactive threat management.

E. Blockchain and Machine Learning Integration

The combination of Blockchain and Machine Learning provides a robust cybersecurity framework that ensures secure forensic evidence storage while enabling intelligent threat detection. Blockchain guarantees data integrity and transparency, whereas Machine Learning automates anomaly detection and threat classification. Together, these technologies improve the reliability, scalability, and effectiveness of digital forensic investigations.

Research Gap

Although several studies have explored Machine Learning-based anomaly detection and blockchain-enabled secure data storage independently, very few systems integrate both technologies into a unified digital forensic framework for insider threat detection. Existing forensic solutions primarily rely on centralized log storage, making them vulnerable to unauthorized modifications and data tampering. Furthermore, many traditional intrusion detection systems depend on predefined signatures, limiting their ability to identify previously unseen insider attacks. There is a need for a comprehensive solution that combines intelligent anomaly detection, tamper-proof log management, and real-time alerting within a single secure platform.

Research Motivation

The increasing frequency of insider attacks has highlighted the need for intelligent forensic systems capable of detecting malicious activities before significant damage occurs. Manual log analysis is inefficient due to the enormous volume of system-generated events, while centralized storage mechanisms fail to guarantee forensic integrity. This research is motivated by the opportunity to combine Machine Learning and Blockchain technologies to develop an automated, secure, and scalable forensic platform. By integrating anomaly detection, immutable log storage, and instant alert generation, the proposed system enhances cyber resilience, improves forensic investigations, and supports proactive security management.

III. PROPOSED METHODOLOGY

3.1 Proposed System

The proposed digital forensic framework integrates Machine Learning, Blockchain, and real-time monitoring to detect insider threats through intelligent log analysis. System logs generated from servers and user activities are continuously collected and preprocessed before being analyzed by Machine Learning models. Isolation Forest identifies anomalous behavior, while Random Forest classifies detected events based on threat severity. The processed logs are securely stored on a blockchain ledger to preserve data integrity and prevent unauthorized modifications. Finally, the system generates real-time email alerts and visualizes security events through a Flask-based dashboard, enabling administrators to respond quickly to potential threats.

3.2 Technologies Used

The proposed system is developed using the following technologies:

Technology	Purpose
Python	Backend Programming
Flask	Web Framework
Scikit-learn	Machine Learning Algorithms
Isolation Forest	Anomaly Detection
Random Forest	Threat Classification
Blockchain	Secure Log Storage
PostgreSQL / MongoDB	Database Management
SMTP	Email Alert Notification
HTML, CSS, Bootstrap	User Interface
PyCharm	Development Environment

3.3 System Architecture

The proposed architecture consists of six major components:

1. Log Data Collection
2. Data Preprocessing
3. Machine Learning-Based Threat Detection
4. Blockchain-Based Secure Log Storage
5. Real-Time Alert Generation
6. Dashboard Visualization

Initially, system logs are collected from multiple devices and processed to remove inconsistencies. The cleaned data is analyzed using Isolation Forest and Random Forest algorithms. Detected threats are securely stored on the blockchain, while administrators receive immediate notifications through email alerts. The dashboard provides a centralized interface for monitoring detected threats and forensic reports.

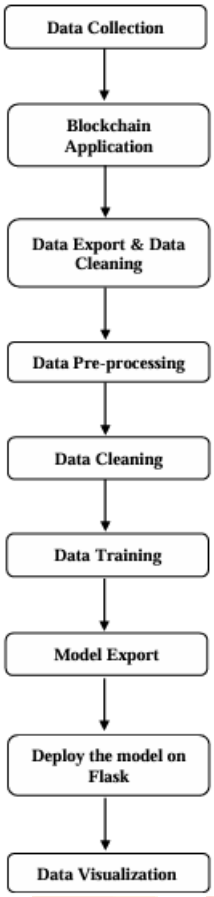

Fig-02: Data Flow Diagram

3.4 Log Data Collection

The system continuously collects log information from multiple sources, including user login activities, IP addresses, timestamps, device identifiers, event descriptions, and severity levels. These logs form the primary dataset used for anomaly detection and forensic analysis.

3.5 Data Preprocessing

Before training the Machine Learning models, the collected log data undergoes preprocessing to improve prediction accuracy. The preprocessing stage includes removing duplicate records, handling missing values, feature extraction, normalization, and encoding categorical variables. These steps ensure that the dataset is clean, consistent, and suitable for anomaly detection.

3.6 Blockchain Integration

Blockchain technology plays a significant role in maintaining the integrity and authenticity of forensic log data. Every processed log entry is encrypted and converted into a cryptographic hash before being stored on the blockchain. Since blockchain is decentralized and immutable, once a log is recorded it cannot be modified or deleted without network consensus. This ensures secure evidence preservation, prevents log tampering, and provides a transparent audit trail for digital forensic investigations. The blockchain layer strengthens the reliability of forensic evidence and supports regulatory compliance.

3.7 Machine Learning Models

The proposed system employs two Machine Learning algorithms to identify insider threats efficiently.

A. Isolation Forest

Isolation Forest is an unsupervised anomaly detection algorithm that isolates abnormal observations from normal log events. It is particularly effective for detecting previously unseen insider attacks because it does not require labeled datasets. The algorithm identifies unusual user activities based on deviations from normal behavioral patterns.

B. Random Forest

Random Forest is a supervised ensemble learning algorithm used for threat classification. It consists of multiple decision trees that collectively classify detected events into different severity levels. The ensemble approach improves prediction accuracy, minimizes overfitting, and provides robust classification performance for cybersecurity applications.

3.8 Real-Time Alert System

To enable rapid incident response, the proposed system incorporates an SMTP-based email notification mechanism. Whenever suspicious activities are detected, the system automatically generates alerts containing information such as timestamp, user details, IP address, event description, and threat severity. These notifications allow security administrators to respond immediately, reducing the potential impact of insider attacks and improving organizational security.

3.9 Advantages of the Proposed System

The proposed framework offers several advantages over conventional forensic solutions:

- Automated detection of insider threats using Machine Learning.
- Secure and tamper-proof log storage through blockchain technology.
- Real-time monitoring and instant email alert generation.
- Improved forensic investigation using immutable audit trails.
- High scalability for processing large volumes of security logs.
- Enhanced prediction accuracy through hybrid Machine Learning models.
- Reduced manual effort in analyzing cybersecurity incidents.
- Improved transparency, reliability, and trustworthiness of forensic evidence.

IV. RESULTS AND DISCUSSION

The proposed blockchain-enabled digital forensic system was implemented using Python, Flask, and Scikit-learn. Security logs collected from multiple sources were preprocessed and analyzed using Isolation Forest and Random Forest algorithms. The processed log records were securely stored using blockchain technology, while SMTP services generated real-time notifications whenever suspicious events were detected.

Figure 2: Login Page

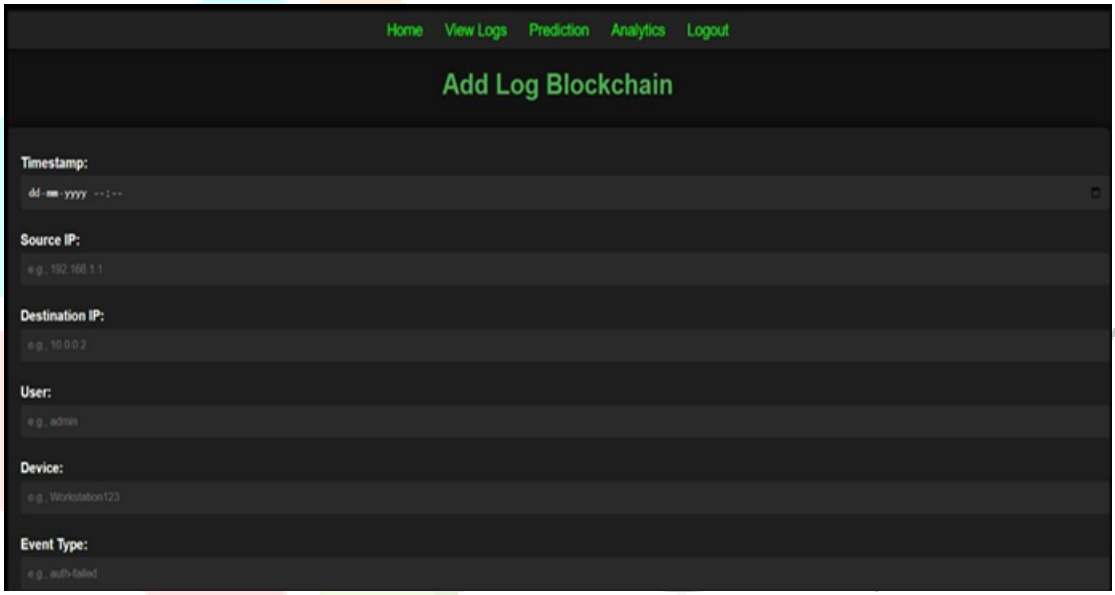

Figure 3: Add Log

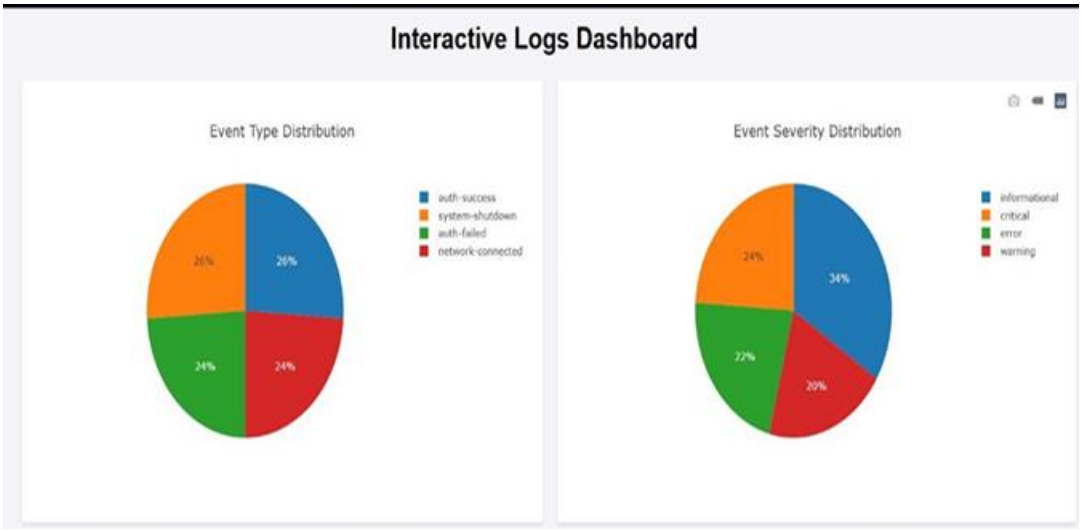

Figure 4: DashBoard

Table 1: Performance Comparison of Machine Learning Models

Algorithm	Accuracy (%)	Precision	Recall	F1-Score
Decision Tree	91.20	0.91	0.90	0.90
Isolation Forest	94.30	0.94	0.93	0.93
Random Forest	**97.10**	**0.97**	**0.97**	**0.97**

The experimental analysis indicates that Random Forest achieved the highest classification accuracy, while Isolation Forest effectively detected anomalous user activities without requiring labeled datasets. The integration of blockchain ensured that all forensic logs remained secure, immutable, and traceable throughout the investigation process.

Threat Detection Analysis

The system successfully detected various insider threat scenarios, including unauthorized login attempts, abnormal file access, privilege misuse, suspicious IP activities, and unusual user behavior. Real-time alerts enabled administrators to initiate immediate incident response, significantly reducing investigation time.

Discussion

The combination of Machine Learning and Blockchain offers significant improvements over conventional forensic systems. Machine Learning automates anomaly detection and threat classification, while blockchain guarantees the integrity and authenticity of forensic evidence. Together, these technologies provide a scalable, secure, and efficient solution for insider threat detection through continuous log analysis.

V. CONCLUSION

This paper presented a blockchain-based digital forensic framework for insider threat detection through intelligent log analysis. The proposed system integrates Isolation Forest for anomaly detection, Random Forest for threat classification, blockchain technology for secure log storage, and SMTP services for real-time alert generation. The experimental results demonstrate that the hybrid framework effectively identifies suspicious activities while preserving the integrity of forensic evidence. By automating log analysis and securing digital records against tampering, the proposed system enhances cybersecurity monitoring and supports reliable forensic investigations. The framework can be deployed across enterprise environments to strengthen organizational resilience against insider threats.

VI. FUTURE SCOPE

Future enhancements may include the integration of deep learning models for advanced behavioral analysis, federated learning for privacy-preserving distributed threat detection, and cloud-native blockchain infrastructures to improve scalability. Incorporating artificial intelligence for automated forensic report generation and predictive cyber risk analysis can further enhance incident response capabilities. The system may also be extended to secure Internet of Things (IoT), cloud computing, and industrial control system environments against sophisticated insider attacks.

ACKNOWLEDGMENT

The authors sincerely thank **Srinath GM**, Guide and Assistant Professor, Department of Computer Science and Engineering, SJC Institute of Technology, Chickballapur, Karnataka, for valuable guidance and continuous support. They also acknowledge the department's assistance and gratefully recognize the Kaggle dataset utilized, which significantly contributed to the successful completion of this research.

The authors also extend their sincere appreciation to the Department of Computer Science and Engineering, SJC Institute of Technology, Chickballapur, for providing the necessary facilities and a supportive research environment.

Furthermore, the authors gratefully acknowledge **Kaggle** for providing the publicly available dataset utilized in this study, which significantly contributed to the successful completion of the research.

REFERENCES

[1] X. Ye, F. Luo, H. Cui, *et al.*, "Research on Insider Threat Detection Based on Personalized Federated Learning and Behavior Log Analysis," *Scientific Reports*, vol. 15, Art. no. 19214, 2025.

[2] K. Fei, J. Zhou, L. Su, W. Wang, and Y. Chen, "Log2Graph: A Graph Convolution Neural Network Based Method for Insider Threat Detection," *Journal of Computer Security*, vol. 33, no. 1, 2025.

[3] Y. Zhang, *et al.*, "LaAeb: A Comprehensive Log-Text Analysis Based Approach for Insider Threat Detection," *Computers & Security*, vol. 148, Art. no. 104126, 2025.

[4] F. Jumani and M. Raza, "Machine Learning for Anomaly Detection in Blockchain: A Critical Analysis, Empirical Validation, and Future Outlook," *Computers*, vol. 14, no. 7, Art. no. 247, 2025.

[5] S. A. Qawasmeh and A. A. S. AlQahtani, "Beyond Firewall: Leveraging Machine Learning for Real-Time Insider Threats Identification and User Profiling," *Future Internet*, vol. 17, no. 2, Art. no. 93, 2025.

[6] T. Tian, C. Zhang, B. Jiang, *et al.*, "Insider Threat Detection for Specific Threat Scenarios," *Cybersecurity*, vol. 8, Art. no. 17, 2025.

[7] T. Al-Shehari, M. Kadrie, M. N. Al-Mhiqani, *et al.*, "Comparative Evaluation of Data Imbalance Addressing Techniques for CNN-Based Insider Threat Detection," *Scientific Reports*, vol. 14, Art. no. 24715, 2024.

[8] Z. Wei, U. Rauf, and F. Mohsen, "E-Watcher: Insider Threat Monitoring and Detection for Enhanced Security," *Annals of Telecommunications*, vol. 79, pp. 819–831, 2024.

[9] C. Song, L. Ma, J. Zheng, *et al.*, "Audit-LLM: Multi-Agent Collaboration for Log-Based Insider Threat Detection," *arXiv preprint*, 2024.

[10] P. Manoharan, W. Hong, J. Yin, *et al.*, "Optimising Insider Threat Prediction: Exploring BiLSTM Networks and Sequential Features," *Data Science and Engineering*, vol. 9, pp. 393–408, 2024.