ISSN 2616-5775 Vol. 8, Issue 10: 63-67, DOI: 10.25236/AJCIS.2025.081009

Design of an Artificial Intelligence-Based Privacy Security Management System

Yan Yan

Queen's University, Kingston, Ontario, Canada

Abstract: *This paper designs an artificial intelligence-based privacy security management system, focusing on analyzing data leakage and abnormal access risk prevention in enterprise IT operation and maintenance scenarios. The system adopts a modular architecture and deeply integrates technologies such as long short-term memory networks, temporal behavior modeling, convolutional neural networks, and log semantic feature extraction to intelligently analyze user operations, data flow paths, and permission invocation patterns in real time. By constructing dynamic privacy risk profiles, the system can accurately identify high-risk behaviors such as unauthorized access and covert data exfiltration. It significantly outperforms traditional rule-based methods in metrics such as anomaly detection accuracy and response latency, providing intelligent technical support for full lifecycle protection of privacy data.*

Keywords: *Artificial intelligence; Privacy security management system; Design*

1. Introduction

With the continuous advancement of enterprise digital transformation, IT system architectures have become increasingly complex, and data assets have grown exponentially, leading to a rapid increase in privacy leakage risks. Traditional operation and maintenance monitoring systems based on threshold alerts and static rules exhibit delayed detection, high false positive rates, and lack of behavioral semantic understanding when facing new types of privacy security incidents such as advanced persistent threats, internal unauthorized operations, and covert data exfiltration. In this context, deeply integrating artificial intelligence technologies into privacy security management systems has become an inevitable trend. By leveraging deep learning models for multidimensional correlation analysis of massive operation logs, user behavior sequences, and data access trajectories, a dynamic privacy risk perception mechanism can be constructed. This mechanism not only accurately identifies explicit threats such as anomalous logins and unauthorized exports but also uncovers latent and stealthy privacy violations through temporal modeling, enabling a shift from a “passive response” to an “active prediction” mode and providing an intelligent technical foundation for compliance with regulatory requirements such as the Personal Information Protection Law.

2. Major Threats to Network Privacy Security

Currently, network privacy security faces a series of high-dimensional threats, with increasingly intelligent attack methods that seriously endanger enterprise sensitive data assets. Distributed denial-of-service (DDoS) attacks manipulate large-scale botnets to launch flood requests against target systems, rapidly exhausting bandwidth and computing resources. This not only causes business interruptions but may also conceal subsequent data theft activities. Zero-day exploits bypass traditional signature-based detection and achieve remote code execution before vendors release patches, directly breaking system boundaries and obtaining access to core databases. Such attacks are characterized by sudden occurrence and high destructive potential, placing stringent demands on real-time monitoring capabilities.

Advanced persistent threat (APT) attack chains involve stages such as spear-phishing email delivery, watering hole site penetration, lateral movement, and long-term dormancy. Attackers often remain hidden within internal networks for months or even years, continuously stealing user credentials, customer privacy information, and trade secrets. APT attacks deliberately evade conventional signature detection, employing reconnaissance-avoidance strategies such as low-frequency slow communication and legitimate protocol masquerading, rendering traditional static rule-based defenses largely

ISSN 2616-5775 Vol. 8, Issue 10: 63-67, DOI: 10.25236/AJCIS.2025.081009

ineffective. These threats collectively highlight a critical issue: privacy data leakage has evolved from an occasional incident to a systemic risk. Therefore, it is imperative to build an intelligent privacy security management system with capabilities for behavioral baseline modeling, anomaly sequence recognition, and cross-domain correlation analysis. Through a deep learning-driven dynamic risk assessment mechanism, such a system can accurately detect covert data exfiltration, unauthorized access, and other privacy violation behaviors.

3. Design of an Artificial Intelligence-Based Privacy Security Management System

3.1 Overall System Architecture

The privacy security management system adopts a modular, closed-loop architecture and deeply integrates artificial intelligence technologies to achieve intelligent protection throughout the data lifecycle. The system first collects diverse data in real time through a distributed data acquisition module, including IT infrastructure logs, application logs, and user behavior audit streams. After standardization and desensitization, the data are transmitted to the central intelligent engine via a high-throughput message queue. The core intelligent analysis module integrates deep temporal models and graph neural networks to dynamically model data access paths, permission invocation sequences, and entity interaction relationships, accurately identifying privacy risk behaviors such as abnormal data export and unauthorized operations. The security response module automatically triggers tiered handling strategies based on risk levels, including session interruption, permission freezing, and data encryption isolation. The management reporting module generates compliance audit views and risk heatmaps based on analysis results, effectively supporting managerial decision-making. Through a feedback mechanism, the entire system continuously optimizes model parameters, forming an adaptive closed loop of "perception — judgment — response — learning," significantly improving the accuracy of privacy threat detection (see Figure 1).

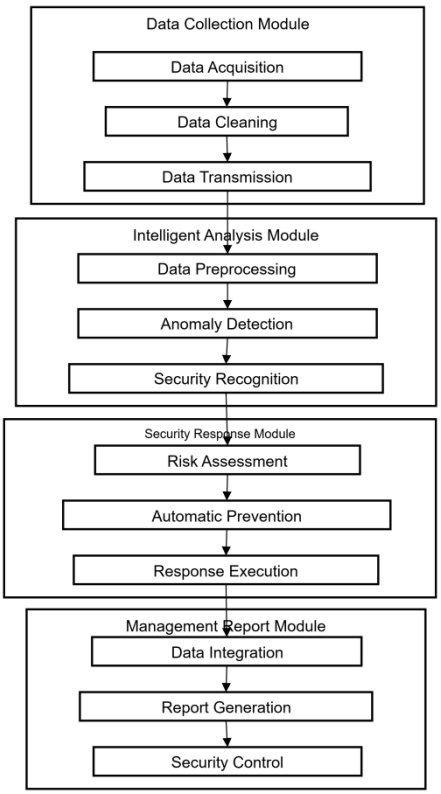

Figure 1. Overall System Architecture

3.2 System Functional Module Design

3.2.1 Data Acquisition Module

The data acquisition module serves as the information sensing front end of the privacy security

ISSN 2616-5775 Vol. 8, Issue 10: 63-67, DOI: 10.25236/AJCIS.2025.081009

management system, adopting a distributed high-availability architecture and deploying multiple lightweight collection nodes to achieve full coverage of IT infrastructure. Each node is equipped with Intel X710 series high-performance network cards, supporting 40 Gbps line-rate throughput to ensure zero packet loss and low latency under high-concurrency scenarios. During the acquisition process, a multi-protocol collaboration mechanism is comprehensively integrated. Network device traffic, port status, and session logs are obtained via SNMPv3 and NetFlow v9; Telegraf agents are employed to collect host CPU, memory, disk I/O, application-layer API calls, and database query behaviors in real time. To balance resource overhead and monitoring sensitivity, an adaptive sampling algorithm is implemented within the module. Under normal conditions, periodic polling occurs every five minutes. Once the edge detector identifies traffic anomalies, a high-frequency sampling mode (up to once per second) is immediately triggered to dynamically capture threat evolution details.

The raw data stream is processed through a preprocessing pipeline performing deduplication, missing value imputation, and timestamp alignment, and is uniformly converted into a structured JSON Schema format. Subsequently, the data are compressed using the Snappy algorithm to reduce transmission bandwidth and encrypted with AES-256 to ensure confidentiality during transit. The encrypted and compressed data stream is delivered to the central processing center via the Apache Kafka high-throughput message pipeline with millisecond-level reliability. The integrated intelligent scheduling engine dynamically optimizes task priority and CPU/memory resource allocation based on real-time system load, completely avoiding multi-task resource conflicts. Simultaneously, a built-in data quality monitoring subsystem continuously tracks metrics such as data integrity rate and field compliance rate, intelligently isolating abnormal data sources, thereby providing a high-fidelity and trustworthy data foundation for subsequent AI analysis.

3.2.2 Intelligent Analysis Module Design

The intelligent analysis module is deployed on a high-performance computing cluster equipped with NVIDIA Tesla V100 GPUs, leveraging parallel acceleration to process terabyte-scale logs. The analysis workflow first applies L1-regularized Lasso regression to select key variables sensitive to privacy leakage from the original high-dimensional feature space. Subsequently, t-SNE nonlinear dimensionality reduction maps the data to a low-dimensional manifold to support visual anomaly clustering. The core detection model adopts an LSTM-CNN hybrid architecture. A 128-unit LSTM layer captures temporal dependencies in user operations, while three multi-scale convolutional layers (3×3, 5×5, 7×7) extract local behavioral pattern shifts, enabling high-precision identification of covert data exfiltration, permission abuse, and other anomalies. At the advanced threat assessment level, the module constructs a dynamic behavior graph based on graph neural networks, abstracting hosts, accounts, files, and other entities as nodes and modeling access relationships as edges. A graph attention mechanism is used to uncover lateral movement paths in APT attacks. To enhance model applicability, the system integrates a federated learning framework, enabling cross-domain model parameter aggregation without exchanging raw privacy data, significantly improving generalization capability.

3.2.3 Security Response Module Design

The security response module adopts a Kubernetes-based microservices architecture, providing high-availability, elastically scalable automated defense capabilities. This mechanism quantitatively evaluates intelligent analysis results, strictly following the CVSS 3.1 standard by comprehensively assessing attack vectors, complexity, and confidentiality impact, outputting a dynamic risk score from 0 to 10. Gradient-based response strategies are executed according to preset thresholds: low-risk events trigger log alerts; medium-risk behaviors drive the SDN controller to perform precise traffic redirection, isolate suspicious hosts, and tighten firewall policies via OpenFlow 1.5; high-risk threats trigger emergency procedures including data snapshot backups. Critically, the response process is deeply integrated with software-defined networking technology, reconstructing network topology at millisecond-level speeds to completely block privacy data leakage channels. Meanwhile, the system integrates the Ansible engine for cross-platform automated patch deployment and leverages the Neo4j knowledge graph to preserve historical incident cases, using subgraph matching algorithms to retrieve similar attack scenarios in seconds. All response actions are fully logged and fed back to the analysis module in real time, continuously driving the evolution of the strategy library and establishing a closed-loop, self-optimizing security governance ecosystem (Table 1) [1].

ISSN 2616-5775 Vol. 8, Issue 10: 63-67, DOI: 10.25236/AJCIS.2025.081009

Table 1. Gradient-Based Response Strategies of the Security Response Module

Risk Level	CVSS 3.1 Risk Score Range	Trigger Example	Response Action	Implementation Technology/Component
Low Risk	0.1 – 3.9	Anomalous login attempts, non-critical configuration changes	Log recording, alert notification	ELK logging system, Prometheus
Medium Risk	4.0 – 6.9	Lateral movement behavior, suspicious API calls	Traffic redirection, host isolation, firewall policy tightening	SDN controller (OpenFlow 1.5)
High Risk	7.0 – 10.0	Bulk export of sensitive data, zero-day exploit	Trigger emergency plan, perform data snapshot backup, automated patch deployment, network topology reconstruction	Kubernetes, Ansible, Neo4j

3.2.4 Management Reporting Module

The management reporting module builds a distributed data warehouse based on Apache Hive, supporting efficient storage of PB-level operation and security logs. Diverse data are uniformly loaded into fact tables and dimension tables according to a star schema, significantly enhancing the performance of complex analyses. An OLAP engine is then used to perform multidimensional drilling, slicing, and roll-up operations, meeting the differentiated insight requirements on privacy risk posture for users ranging from frontline operators to senior management. The system deeply integrates natural language generation technology, leveraging a Transformer-based large language model to intelligently convert structured analysis results into narrative text that complies with audit standards. Internally, a KPI evaluation system driven by a Drools rule engine dynamically calculates metrics such as privacy incident response timeliness and data access compliance rates. Reports are made available through RESTful APIs for customized subscription and implement role-based fine-grained access control. Combined with AES-256 encryption, the module ensures the confidentiality of sensitive information during transmission, comprehensively supporting compliance audits [2].

4. Experimental Analysis

4.1 Experimental Scheme

The experiment deployed the privacy security management system in an enterprise-grade data center environment consisting of 50 servers, running the CentOS 7.9 operating system and equipped with typical business components such as Apache Web services and MySQL databases, thus constructing a test platform close to real-world scenarios. The experiment progressed in three stages: In the first stage, 30 days of normal operation logs were collected, and unsupervised learning algorithms were used to establish user behavior baselines and privacy access compliance models. In the second stage, 500 high-risk attack samples were injected, including SYN Flood denial-of-service attacks, SQL injection, and zero-day exploits, with particular emphasis on simulating typical privacy leakage paths such as unauthorized data queries and bulk sensitive information export. In the third stage, the AI-driven real-time monitoring engine was enabled. The fused model combining Long Short-Term Memory (LSTM) networks and Graph Neural Networks (GNN) dynamically identified abnormal access patterns. Evaluation metrics focused on anomaly detection accuracy, average detection latency, response execution timeliness, and false positive rate, systematically validating the system's ability to accurately perceive privacy risks in complex adversarial environments (Table 2).

Table 2. Experimental Design and Evaluation Metric Configuration

Experimental Stage	Duration	Data/Behavior Type	Injected Attack Type (Quantity)	Core Model/Method	Evaluation Metrics
First Stage	30days	Normal O&M and user operation logs	None	K-means + Isolation Forest	Baseline coverage, behavior clustering stability
Second Stage	7days	Simulated attack traffic and abnormal access behaviors	SYN Flood (150); SQL injection (120); Unauthorized export (130); Zero-day exploit (100)	—	Attack injection authenticity, scenario coverage
Third Stage	14days	Real-time mixed traffic (normal + abnormal)	Dynamic adversarial samples (including adversarial perturbations)	LSTM + GNN hybrid model	Detection accuracy (%); False positive rate (%); Average detection latency (ms); Response execution timeliness (ms)

ISSN 2616-5775 Vol. 8, Issue 10: 63-67, DOI: 10.25236/AJCIS.2025.081009

4.2 Experimental Results and Analysis

The experimental results fully validate the advantages of the privacy security management system in detection accuracy and response efficiency. As shown in Table 1, the system achieves an anomaly detection accuracy of 97.8%, representing a 14.7-percentage-point improvement over traditional rule-based engines. The mean time to detect (MTTD) is reduced from 180 seconds to 12 seconds, and the mean time to respond (MTTR) drops from 45 minutes to 3 minutes, both representing a 93.3% reduction. This performance improvement is primarily attributed to the LSTM-CNN hybrid model's ability to jointly model temporal features of user behavior and local operation patterns, enabling precise identification of covert data anomalies. Meanwhile, the SDN-based automated response mechanism achieves millisecond-level traffic scheduling, significantly shortening the threat handling closed loop. The false positive rate decreases from 8.7% to 1.2%, indicating that the system maintains high specificity while operating at high sensitivity, effectively reducing operational interference [3].

Specialized tests across different attack scenarios further analyzed the system's generalization capability. As shown in Table 3, traffic-based attacks such as SYN Flood, due to their prominent characteristics, achieved a detection rate of 99.5%, with alerts generated within an average of 8 seconds. SQL injection attacks at the application layer, leveraging semantic anomaly recognition, achieved a detection rate of 98.2%. Notably, for zero-day exploits without prior signatures, the system still reached a detection rate of 94.7%, with an average detection time controlled within 25 seconds, demonstrating the graph neural network's reasoning ability over anomaly behavior chains and the federated learning model's transfer adaptability to unknown threats. These results indicate that the system performs excellently against known threats and is capable of addressing advanced persistent privacy leakage risks, providing a reliable technical pathway for constructing an intelligent privacy security protection system [4].

Table 3. Detection Performance across Different Attack Scenarios

Attack Type	Detection Rate	Average Detection Time (s)
SYN Flood	99.50%	8
SQL Injection	98.20%	15
Zero-day Exploit	94.70%	25

5. Conclusion

This study developed an intelligent operation and maintenance management system focused on privacy security, deeply integrating an LSTM-CNN hybrid model and graph neural networks to accurately identify high-risk behaviors such as unauthorized access and abnormal data exfiltration. Experimental results demonstrate that the system significantly outperforms traditional methods in detection accuracy, average response latency, and false positive suppression, effectively supporting enterprise privacy compliance and providing feasible technical support for intelligent security operations.

References

[1] Du Zhaofang. Design and Implementation of a Privacy Security Management System Based on Artificial Intelligence Technology [J]. Information Recording Materials, 2023, 24(8): 47-49.
[2] Deng Yan. Design and Implementation of an IoT Device Control System Based on Blockchain [D]. Anhui: Anhui Normal University, 2023.
[3] Li Fang. Design and Security Management of Smart City Charging Pile Management System Based on IoT [J]. Digital Users, 2024(44): 221-222.
[4] Li Lifeng, Sun Fengna, Lü Zhuangzhuang, et al. Design of a Power Communication Security Management System Based on Blockchain Technology [J]. Mobile Information, 2025, 47(7): 62-64.