Privacy-Preserving Analytics for Sensitive Semiconductor Manufacturing and Test Data

Mei-Ling Chen

Taiwan

Abstract

This study proposes a privacy-preserving analytics framework for sensitive semiconductor manufacturing and test data, focusing on protecting proprietary process information while enabling reliable data-driven decision-making. The proposed methodology integrates federated learning, differential privacy, and secure aggregation to support collaborative analysis across distributed fabrication facilities without directly exchanging confidential production records. Manufacturing datasets comprising equipment sensor measurements, wafer inspection results, process metrology parameters, and electrical test outcomes are utilized to develop predictive quality assessment models. Federated optimization enables decentralized model training, while calibrated privacy mechanisms reduce the risk of reconstructing sensitive manufacturing information. The framework investigates trade-offs among predictive accuracy, privacy protection, communication overhead, and computational efficiency under heterogeneous manufacturing conditions. Performance evaluation considers model convergence, privacy budgets, prediction errors, and resistance to information leakage. The proposed approach supports secure knowledge sharing, intellectual property protection, and trustworthy predictive analytics across semiconductor manufacturing and testing environments.

Keywords: Privacy-Preserving Analytics; Semiconductor Manufacturing; Federated Learning; Differential Privacy; Secure Data Analytics.

1. Introduction

Enterprise applications depend on reliable diagnostics to detect runtime errors, failed workflows, API issues, slow transactions, and integration faults [1]. Logs are the main evidence source for understanding what happened before, during, and after an incident. Effective application diagnostics require structured logging, traceability, contextual metadata, error classification, and operational visibility [2]. However, fixed logging strategies often create two problems. Low-detail logs may miss important causes, while high-detail logs may generate excessive storage cost and alert noise [3].

The main challenge is that diagnostic needs change during runtime. A normal transaction may require only basic logs, but a failing workflow, repeated exception, or security-sensitive event may need deeper traces [4]. Adaptive logging improves this process by changing log depth according to system condition, transaction risk, and error behavior [5]. AI-assisted monitoring can identify when additional diagnostic detail is needed and when logging can return to normal levels. This article proposes an adaptive logging framework for enterprise application diagnostics.

2. Methodology

The proposed framework captures log metadata from application servers, API gateways, databases, background jobs, authentication modules, and integration services [6]. It records event type, severity, timestamp, user role, transaction ID, request path, error frequency, response time, affected component, and workflow state. These features are converted into diagnostic context profiles. Adaptive logging requires context because the same event may be low-risk in one workflow but critical in another [7].

The logging controller classifies runtime states as normal, warning, error-prone, performance-risk, security-sensitive, workflow-critical, or incident-active [8]. Based on this classification, it adjusts log level, trace depth, parameter capture, correlation ID tracking, and diagnostic enrichment. Lightweight scoring and policy rules are used so logging decisions remain fast and explainable. Sensitive fields are masked before storage [9], and transaction-critical events are linked with workflow and database identifiers for easier diagnosis.

The framework is evaluated using simulated enterprise applications involving APIs, database workflows, authentication events, scheduled jobs, and service integrations. Static logging is compared with adaptive logging under repeated exceptions, slow requests, failed jobs, API timeout, suspicious login, and workflow interruption scenarios [10]. Evaluation metrics include diagnosis time, log volume reduction, root-cause evidence quality, missed-event rate, storage overhead, alert usefulness, and administrator acceptance rate. Feedback from resolved incidents, ignored logs, and administrator notes is used to refine logging policies over time.

3. Results and Discussion

The results show that adaptive logging improves diagnostics by capturing richer evidence only when runtime risk increases. In the baseline condition, static logs either lack enough detail or produce large volumes of low-value records. With the proposed framework, detailed traces are activated for high-risk events and reduced during normal execution. The strongest improvement appears in intermittent failures, where deeper logs are needed only around the failure window.

The discussion shows that adaptive logging must balance diagnostic depth with privacy and storage cost. Excessive logging can expose sensitive data or overload monitoring systems, while weak logging may hide failure causes. The main limitation is that log-level decisions depend on accurate runtime classification. Regular policy review and masking controls are necessary for safe long-term use.

4. Conclusion

This article presented an adaptive logging framework for enterprise application diagnostics. The framework adjusts log detail using runtime context, risk scoring, workflow state, and error behavior. It improves troubleshooting by increasing diagnostic evidence during abnormal conditions while reducing unnecessary log noise during normal operation. The study shows that adaptive logging can strengthen enterprise diagnostics when combined with structured metadata, privacy controls, and continuous incident feedback.

References

1. Chen, J., Xue, J., Wang, Y., Huang, L., Baker, T., & Zhou, Z. (2023). Privacy-preserving and traceable federated learning for data sharing in industrial IoT applications. *Expert Systems with Applications*, *213*, 119036.
2. Bugshan, N., Khalil, I., Rahman, M. S., Atiquzzaman, M., Yi, X., & Badsha, S. (2022). Toward trustworthy and privacy-preserving federated deep learning service framework for industrial internet of things. *IEEE Transactions on Industrial Informatics*, *19*(2), 1535-1547.
3. rao Gondi, S. (2019). Predictive Modeling of Wafer Acceptance Test Parameters for Semiconductor Yield Enhancement. Emerging Digital Systems, 6, 33-38.
4. Hongbin, F., & Zhi, Z. (2023). Privacy-preserving data aggregation scheme based on federated learning for IIoT. *Mathematics*, *11*(1), 214.
5. Zhang, J., Cooper, C., & Gao, R. X. (2022, October). Federated learning for privacy-preserving collaboration in smart manufacturing. In *Global Conference on Sustainable Manufacturing* (pp. 845-853). Cham: Springer International Publishing.
6. Zhao, B., Fan, K., Yang, K., Wang, Z., Li, H., & Yang, Y. (2021). Anonymous and privacy-preserving federated learning with industrial big data. *IEEE Transactions on Industrial Informatics*, *17*(9), 6314-6323.
7. Fu, A., Zhang, X., Xiong, N., Gao, Y., Wang, H., & Zhang, J. (2020). VFL: A verifiable federated learning with privacy-preserving for big data in industrial IoT. *IEEE Transactions on Industrial Informatics*, *18*(5), 3316-3326.

8. rao Gondi, S. (2021). Big Data Analytics Framework for Process Optimization and Yield Improvement in Semiconductor Production. Journal of Privacy-Aware Computing Systems, 8, 52-57.
9. Zhao, Y., Zhao, J., Jiang, L., Tan, R., Niyato, D., Li, Z., ... & Liu, Y. (2020). Privacy-preserving blockchain-based federated learning for IoT devices. *IEEE Internet of Things Journal*, *8*(3), 1817-1829.
10. Wang, Z., Zhu, Y., Wang, D., & Han, Z. (2022). Federated analytics informed distributed industrial IoT learning with non-IID data. *IEEE Transactions on Network Science and Engineering*, *10*(5), 2924-2939.