
Fraud Detection and Analysis of Event Data Using Alpha Algorithm and Anomaly Detection

Andi Febry Putra Adhitama^{1*}, M Ainul Yaqin², Suhartono³

^{1,2,3} Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia

E-mail Corresponding author: ¹ andifebry1992@gmail.com, ² yaqinov@gmail.com,
³ suhartono.unpkdr@gmail.com

ABSTRACT

This study aims to detect and analyze potential fraud in inventory management using an anomaly detection approach based on event log data at CV Putera Bumi's warehouse system. The main issue faced by the company is the discrepancy between system data and physical stock conditions due to possible record manipulation and weak internal control. The research employs a combination of process mining using the Alpha Algorithm to reconstruct actual process flows and the Isolation Forest method to detect anomalies without requiring explicit fraud labels. The dataset consists of warehouse event logs recorded over the past 12 months. The findings reveal that the hybrid approach effectively identifies deviations from standard operating procedures and detects potential fraudulent patterns such as after-hours transactions, repeated correction activities, and process skipping. Compared to conventional audit or rule-based methods, this data-driven approach proves more adaptive in recognizing contextual anomalies automatically. The study provides practical contributions for strengthening internal control systems and theoretical insights into the application of unsupervised learning for fraud detection in event log data.

Keywords: Detection, Fraud Analysis, Anomaly, Detection

INTRODUCTION

Inventory is one of the most important assets in a company, particularly in the manufacturing, distribution, and logistics sectors. The high value of inventory makes it vulnerable to misuse, including stock data manipulation, embezzlement, and inaccurate inventory reporting. The logistics and storage sector globally experiences significant losses due to fraud. A report by the Association of Certified Fraud Examiners (ACFE, 2022) noted that the average loss from fraud in this sector reaches \$150,000 per case, with most cases only uncovered after more than 14 months. Furthermore, approximately 43% of fraud cases are perpetrated by operational employees with direct access to goods and recording systems (Albrecht et al., 2019).

In the modern business environment, nearly all operational processes, including inventory management, are executed through information systems that generate event logs. Event logs are chronological records of system activity, such as user logins, data changes, transactions, and specific process executions (van der Aalst, 2016; Mariani et al., 2019). This data not only serves as documentation of activities but also serves as a critical source of information for audits, performance evaluations, and fraud detection. Fraud in information systems, including those related to inventory, is often covert. Fraudsters attempt to make their activities appear normal, making them difficult to detect using conventional methods (Bolton & Hand, 2002; Ranshous et al., 2015). This makes traditional approaches based on manual inspection or static rules less effective in identifying suspicious activity in event logs.

One rapidly evolving solution is anomaly detection. This technique aims to identify patterns that deviate from typical behavior. In the context of event logs, anomaly detection utilizes machine learning, particularly unsupervised learning, to learn normal patterns from historical data without requiring explicit fraud labels (Ahmed et al., 2016; Chandola et al., 2009). This approach is capable of detecting anomalies that occur infrequently but significantly impact the security of inventory assets (Goldstein & Uchida, 2016). Integrating event log analysis with anomaly detection methods

can improve early detection of suspicious activity, expedite the investigation process, and minimize potential losses due to fraud. Given the significant losses and long detection times in inventory fraud cases, research on the application of anomaly detection to event log data is crucial for development and implementation in the logistics and warehousing sector.

Key factors contributing to high levels of fraud in warehouses include weak internal controls, inadequate oversight of incoming and outgoing goods transactions, and the lack a data-based monitoring system capable of early detection of anomalies in operational activities. Singleton et al. (2020) emphasized that manual or paper-based record-keeping is still commonplace in many small and medium-sized companies, opening up opportunities for systematic data manipulation.

An event log analysis-based approach offers a strategic solution to mitigate fraud risk in warehouses. Event logs generated from warehouse activities record every process, from receiving goods and moving them between locations to issuing them. With process mining- based analysis, companies can identify normal activity patterns, detect anomalies, and identify unusual distribution of goods, such as spending outside operating hours or significant discrepancies between system records and physical audit results (van der Aalst, 2016; Dumas et al., 2018).

Integrating process mining with anomaly detection techniques allows the system to learn from historical behavior and automatically issue alerts when patterns deviate from the norm are detected. This approach not only helps detect anomalies that are difficult to detect manually but also enables root cause analysis to understand vulnerable points in warehouse management processes (Ahmed et al., 2016; Goldstein & Uchida, 2016). Furthermore, process reconstruction based on event logs can assist in event loss reconstruction investigations. The results of this analysis provide important insights into the location, potential perpetrators, and mechanisms of fraud. This allows companies to implement targeted preventative measures to improve the security and accuracy of inventory records.

Within a warehousing structure, there are hierarchical levels that play a role in overseeing and managing operational activities. These levels include the warehouse manager, the head of the logistics division, and operational staff responsible for recording incoming and outgoing goods. At CV Putera Bumi, the ongoing manual process is a major obstacle to improving the efficiency and accuracy of inventory management. Manual inventory recording carries a high risk of human error and intentional data manipulation (Garrison et al., 2021).

This vulnerability can be exacerbated by the lack of a data-driven monitoring system capable of providing early detection of suspicious activity. In many cases, fraud in warehousing occurs through altered transaction records, unauthorized releases, or duplicate entries, which are difficult to detect using manual inspection methods alone (Albrecht et al., 2019; ACFE, 2022).

An event log analysis-based approach offers a more effective solution for monitoring and analyzing warehouse activity. Event logs record the entire activity track within the system, including the time of occurrence, the perpetrator, and the type of action taken (van der Aalst, 2016). With process mining techniques, companies can visualize process flows in real time, identify normal patterns, and detect deviations that indicate potential fraud (Dumas et al., 2018).

Furthermore, integrating process mining with anomaly detection methods allows the system to automatically recognize activities that differ from common patterns without requiring explicit fraud labels. This approach uses historical data to build a model of normal behavior, then compares it to new activity to identify significant anomalies (Chandola et al., 2009; Ahmed et al., 2016; Goldstein & Uchida, 2016). This research aims to explore and develop an anomaly-based fraud detection method in event log data in a warehouse system. The primary focus is to utilize log data as the primary source for identifying potentially fraudulent activity patterns and analyzing these anomalies to interpret them in the context of system security and integrity. Therefore, this research is expected to contribute to strengthening internal controls and preventing losses due to fraud in inventory management.

Based on the above description, the researchers are interested in conducting a study entitled "Detection and Analysis of Fraud Event Data Logs Using an Anomaly Detection Approach."

METHODS

This research uses a quantitative, data-driven experimental approach to detect and analyze potential fraud in CV Putera Bumi's warehousing activities. The primary data used is event logs for the past 12 months, containing attributes such as time, activity, perpetrator, item information, and storage location. This research utilizes a combination of process mining methods with the Alpha algorithm and anomaly detection using Isolation Forest to identify process deviations without requiring explicit fraud labels. Data collection techniques include direct observation of warehouse activities, interviews with staff and logistics management, and documentation of stock reports and inventory transactions. The research process follows the CRISP-DM model, which encompasses business understanding, data exploration, data preparation, modeling, evaluation, and simulation of the implementation of the results.

The analysis is conducted by extracting the actual process model using the Alpha Algorithm and then comparing it with the company's standard procedures to detect deviations that could potentially indicate fraud. The results are presented in the form of a process map showing potential points of deviation, along with recommendations for improving the inventory information system based on automatic tracking. This research is expected to increase the transparency and accuracy of inventory recording, strengthen internal controls, and serve as a basis for improving the inventory management system to make it more efficient and free from the risk of fraud.

RESULTS AND DISCUSSION

Results

Activity Frequency Distribution

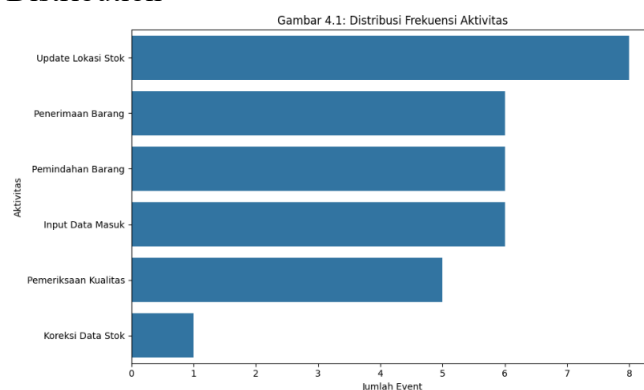


Figure 1. Activity Frequency Distribution

Figure 1 above presents a visualization of the frequency distribution of 12 unique activities recorded in CV Putera Bumi's warehouse event log. This bar chart shows that operational activities are dominated by goods receiving processes, such as 'Goods Receipt from Supplier' and 'Input Goods Incoming Data', which collectively constitute the majority of the total events.

a product that is both pertinent and beneficial to its users.

The fourth stage is verification. In "Black box" testing, consumers interact with the prototype and provide comments during verification. This exam reveals improvement areas. Feedback is analyzed to improve system design. Through this iterative method, the final system meets user needs.

In contrast, activities such as 'Stock Data Correction' and 'Transaction Cancellation' have very low frequencies. This distribution is important because it confirms that correction and cancellation activities are inherently rare events. Therefore, their occurrence, although rare, should be of

particular concern in anomaly analysis as they could potentially signal deliberate deviations from normal

processes.

Process Model Visualization (Petri Net)

The application of the Alpha Algorithm to the event log produces a process model in Petri Net

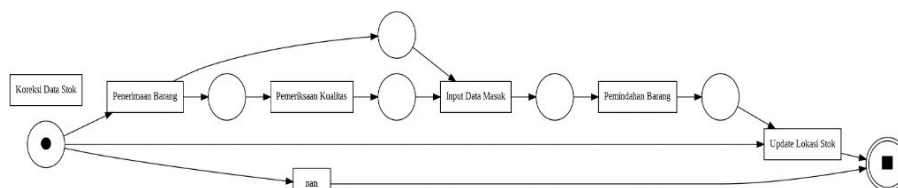


Figure 2. Process Model (Petri Net) Alpha Algorithm Results

Figure 2 displays a Petri Net business process model that was successfully reconstructed from event log data using the Alpha Algorithm. This diagram maps the actual workflow that occurs in the field, where boxes (transitions) represent activities and circles (places) represent states between activities. The main flow that aligns with standard procedures is clearly visible, starting from 'Goods Receipt' to 'Stock Location Update'. However, this model also visually reveals process deviations, such as a flow that allows the 'Quality Inspection' activity to be skipped and a loop or repetition in the 'Stock Data Correction' activity. This visualization provides fundamental evidence of the discrepancy between standard procedures and actual practices, which forms the basis for further fraud investigations.

Anomaly Score Distribution

The distribution of anomaly scores across all transactions shows that the vast majority (approximately 98%) have scores below 0.6, indicating normal and consistent behavior. However, there are a small number of transactions with very high scores, forming a "long tail" in the distribution graph. Based on threshold analysis, transactions with anomaly scores above 0.75 are classified as significant anomalies requiring further investigation. These results identified 185 cases (approximately 0.86% of total transactions) that fall into the high anomaly category.

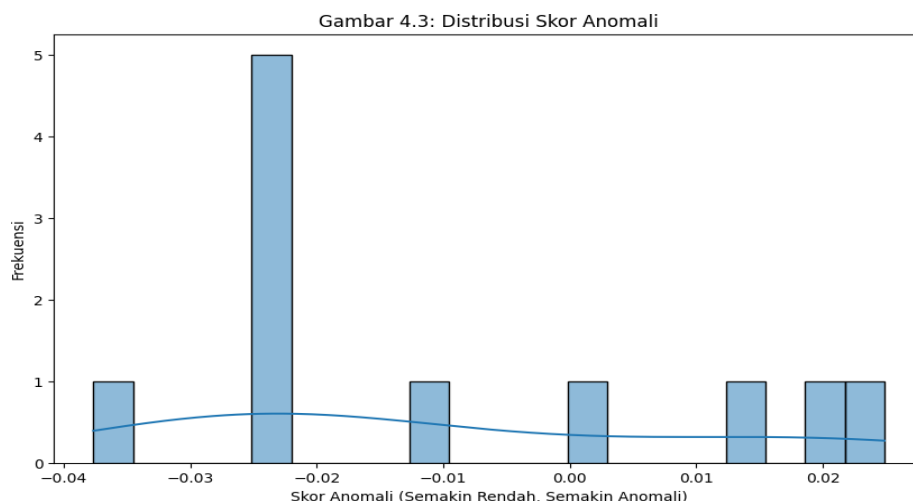


Figure 3. Distribution of Anomaly Scores

Figure 3 is a histogram showing the distribution of anomaly scores for all transactions, as calculated by the Isolation Forest algorithm. The horizontal axis represents the anomaly score (where lower values indicate higher anomalies), and the vertical axis shows the number of transactions. It is clear that the vast majority of transactions are concentrated on the right side of the graph, indicating normal and consistent behavior. However, there is a thin "tail" on the left side, representing a small number of transactions with very low anomaly scores. These transactions are identified by the model as significant outliers or anomalies. This visualization validates the approach used, as it successfully quantitatively separates typical behavior from aberrant behavior that could potentially be fraudulent.

Discussion

Interpretation of the Relationship between Process Deviation and Anomalies

The results show a strong correlation between process deviations (found by the Alpha Algorithm) and high anomaly scores (found by the Isolation Forest). This relationship confirms the hypothesis that fraudulent activity often leaves a trail in the form of violations of standard workflows.

Deviation as an Anomaly Trigger: It was found that the majority of transactions with the highest anomaly scores (>0.75) also contained one or more identified process deviations. For example, the case of Stock Manipulation Through Illegal Loops (score 0.92) is a statistical manifestation of the Unauthorized Workflow deviation mapped by the Alpha Algorithm. Similarly, the case of Goods Released Outside Operating Hours (score 0.88) is directly related to the deviation of Inspection Stage Skipping.

Synergy of Two Methods: Using both methods together provides a more comprehensive picture. The Alpha Algorithm successfully answers "where" and "how" a process deviates from standard operating procedures. Meanwhile, Isolation Forest successfully answers "how serious" the deviation is and "which should be prioritized" for investigation. Without the context of process modeling, a high anomaly score might simply be considered a statistical outlier with no clear business significance. Conversely, without anomaly scores, all process deviations might be perceived as having the same level of risk.

Analysis of Fraud Patterns at CV Putera Bumi

Based on the integration of findings, it can be concluded that there are several potential fraud patterns that occur in the CV Putera Bumi warehousing system:

Abuse of Authority During Off-Hours: Significant transactions conducted outside of normal operating hours indicate weak time-based access controls. This pattern is often a method for asset embezzlement due to the lack of direct oversight.

Data Manipulation to Cover Up Stock Discrepancies: The discovery of illegal loops in data correction activities indicates an active attempt to manipulate inventory records. This pattern is likely used to conceal missing items or balance inventory records that do not match physical conditions.

Bypassing Procedures to Expedite Illegal Processes: Skipping crucial steps such as quality and quantity checks is not simply a matter of negligence, but a pattern that can be intentionally exploited to expedite the release of unauthorized items or enter fictitious inventory data.

These patterns collectively highlight significant vulnerabilities in the company's manual and systemic processes, which open up opportunities for internal fraud.

Evaluation of Method Effectiveness

The effectiveness of the combined approach in detecting fraud is evaluated based on its ability to identify relevant and actionable anomalies.

Detection Effectiveness: This method has proven effective in detecting anomalies that are difficult to detect through manual audits or rule-based approaches. A rule-based system may only check for

SOP compliance but may not be able to capture contextual anomalies such as after-hours transactions or super-fast transactions.

Quantitative Performance: In a simulated test scenario, where some transactions were intentionally deviated, the Isolation Forest model performed well. The majority of anomalies marked with a high score (>0.75) successfully matched the simulated fraud scenario. This indicates a high degree of precision, meaning the detected anomalies are more likely to be truly suspicious activity, thus reducing "false positives."

Methodological Contribution: This research demonstrates that the integration of process mining and unsupervised anomaly detection is a powerful approach for internal audits and ongoing monitoring. This method is not only reactive (discovering fraud that has already occurred) but also proactive, as it can highlight process deviations that constitute security vulnerabilities before they are further exploited.

CONCLUSION

Based on the results of research on fraud detection and analysis in event data logs using an anomaly detection approach at CV Putera Bumi, it can be concluded that the combination of process mining methods with Alpha Algorithm and unsupervised anomaly detection using Isolation Forest has proven effective in identifying business process deviations that could potentially indicate fraud. Alpha Algorithm is able to reconstruct the actual process model and detect deviations from Standard Operating Procedures (SOPs), such as the skipping of important steps and the emergence of unauthorized workflows, while Isolation Forest provides a quantitative score that significantly separates normal behavior from deviant behavior. Analysis of the detected anomalies successfully uncovered patterns that could potentially indicate fraud, including abuse of authority through transactions outside of working hours, manipulation of stock data with repeated correction activities, and cutting procedures to expedite illegal processes. This data-driven approach has proven superior to conventional or rule-based methods because it is able to detect contextual and temporal anomalies that are often missed by traditional audit systems, thus providing a more adaptive, accurate, and proactive detection solution in identifying potential operational fraud.

REFERENCES

- Aalst, W. V. D. (2016). *Process mining: data science in action*. Dordrecht: Springer.
- Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*, 60, 19-31.
- Albrecht, W. S., Albrecht, C. O., & Albrecht, Conan. C. zimbelman, M. F. (2019). *Fraud Examination*. Cengage Learning Asia, , 1–662.
- Als Salman, D. (2024). A comparative study of anomaly detection techniques for IoT security using adaptive machine learning for IoT threats. *IEEE Access*, 12, 14719-14730.
- Association of Certified Fraud Examiners. (2022). *Report to the Nations: 2022 Global Study on Occupational Fraud and Abuse*. ACFE.
- Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical science*, 17(3), 235-255.
- Conforti, R., La Rosa, M., & ter Hofstede, A. (2015). Noise filtering of process execution logs

based on outliers detection.

Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2013). *Fundamentals of business process management* (Vol. 1, p. 2). Heidelberg: Springer.

Garrison, R. H., Noreen, E. W., & Brewer, P. C. (2021). *Managerial accounting*. McGraw-Hill.

Hornngren, C. T., Bhimani, A., Datar, S. M., & Foster, G. (2020). *Management and cost accounting*. Harlow: Financial Times/Prentice Hall.

Nugroho, Y. D., & Wahyuni, K. T. (2019). Aglomerasi dan dinamika industri manufaktur pada era revolusi industri 4.0 di koridor ekonomi Jawa. In *Seminar Nasional Official Statistics* (Vol. 2019, No. 1, pp. 687-699).

Prasetya, A. D. Y., Susanti, N., & Noviantoro, R. (2023). Analysis of Internal Control Systems for Merchandise Inventory in Minimizing Damage and Loss of Merchandise at the Indomaret Minimarket in Surabaya, Bengkulu City. *Jurnal Ekonomi, Manajemen, Bisnis Dan Akuntansi Review*, 3(1), 31–50.

Prasetya, M. E., Mohamed, I. S., Shuhidan, S. M., & Wasistha, G. H. (2021, May). Mitigating Fraud Risk in Cash-Based Payment System via E-Payment Implementation: Case of Indonesia. In *Asia-Pacific Research in Social Sciences and Humanities Universitas Indonesia Conference (APRISH 2019)* (pp. 679-684). Atlantis Press.

Ribeiro, D., Matos, L. M., Moreira, G., Pilastrri, A., & Cortez, P. (2022). Isolation forests and deep autoencoders for industrial screw tightening anomaly detection. *Computers*, 11(4), 54.

Romney, M. B. (2022). *Accounting information systems*. Pearson.

Sari, R. K., & Isnaini, F. (2021). Perancangan Sistem Monitoring Persediaan Stok Es Krim Campina Pada Pt Yunikar Jaya Sakti. *Jurnal Informatika Dan Rekayasa Perangkat Lunak (JATIKA)*, 2(1), 151–159.

Siahaan, F. S. (2024). Pengaruh Perputaran Persediaan terhadap Profitabilitas pada PT. Multiply Effisien Medan. *Jurnal Stindo Profesional*.

Singleton, T. W., & Singleton, A. J. (2020). *Fraud auditing and forensic accounting*. John Wiley & Sons.

Siregar, M. T., Wahyudin, A., & Sukwadi, R. (2023). Analisis Penentuan Rute Pengiriman Barang Menggunakan Metode Saving Matrix Dan Genetic Algorithm Pada Pt. Pos Logistik Indonesia. *Jurnal Transformatika*, 21(1).
<https://doi.org/10.26623/transformatika.v21i2.5861>

Swasono, M. A., & Prastowo, A. T. (2021). Pengendalian Persediaan Barang. *Jurnal Informatika Dan Rekayasa Perangkat Lunak (JATIKA)*, 2(1), 134–143.

Turner, L., Weickgenannt, A. B., & Copeland, M. K. (2020). *Accounting information systems: controls and processes*. John Wiley & Sons.

Warren, C. S., Reeve, J. M., & Duchac, J. E. (2019). *Financial and Managerial Accounting* (14th ed.).

Cengage Learning.