Aws Ec2 Tools Enhancing Fruad Detection Multi-Perspective E- Commerce Transactions Through A Multi-Perspective Method

[1] T.Laxmi, [2] D.Priyanka, [3] O.Urmila, [4] Dr. G. JawaherlalNehru
[1,2,3]UG Scholar, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
[4]Associate Professor, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
gjnehruceg33@gmail.com

Abstract:

Fraud detection in multi-participant e-commerce transactions, which involve various factors such as buyers, sellers, and intermediaries, remains a significant challenge due to the dynamic and concealed nature of fraudulent behaviors. Traditional fraud detection methods primarily focus on historical order information, often overlooking the dynamic behaviors of users from multiple perspectives. This paper proposes a novel approach to enhance fraud detection accuracy and efficiency in e-commerce by integrating AWS EC2 tools with a multi-perspective detection method. The proposed system utilizes machine learning and process mining techniques to monitor real-time user behaviors. It establishes a behavioral model for e-commerce platforms, analyzing abnormalities and extracting significant features from transaction data. These features are then fed into an ensemble classification model to detect fraudulent behaviors effectively. The system demonstrates superior performance in identifying dynamic fraudulent activities by leveraging the computational power and scalability of AWS EC2, ensuring a robust and adaptable solution for e-commerce fraud prevention.

Keywords: Fraud detection, Multi-participant e-commerce transactions, Buyers ,Sellers, Machine learning, Process mining, AWS EC2, E-commerce fraud prevention.

1.INTRODUCTION

The rapid growth of e-commerce has revolutionized the way businesses and consumers interact, offering convenience and global reach. However, the increasing volume and complexity of online transactions have also led to a significant rise in fraudulent activities. In multi-participant e-commerce transactions, which involve various stakeholders such as buyers, sellers, payment processors, and logistics providers, detecting and preventing fraud poses a substantial challenge. The dynamic nature of e-commerce fraud, coupled with the need for real-time monitoring, necessitates advanced detection mechanisms capable of adapting to changing behaviors and uncovering sophisticated fraudulent schemes. Traditional methods, which often rely on historical transaction data and static rule-based algorithms, fall short in addressing these evolving threats, leading to inefficiencies in identifying fraudulent behaviors and potential financial losses.

Existing fraud detection approaches primarily focus on analyzing transaction histories, user profiles, or other static data to identify suspicious activities. While these methods can be effective for detecting well-known patterns of fraud, they often lack the ability to capture subtle, dynamic behaviors that emerge in multi-participant transactions. This is particularly problematic in the e-commerce domain, where fraudsters continuously adapt their techniques to exploit vulnerabilities in the system. Furthermore, the reliance on centralized fraud detection frameworks introduces processing delays and potential single points of failure, limiting the scalability and responsiveness of these systems. As e-commerce continues to evolve, there is a growing need for a more robust, flexible, and comprehensive approach to fraud detection that can accommodate the multi-faceted nature of online transactions.

2. LITERATURE SURVEY

The integration of advanced technologies such as blockchain, Python, machine learning, and cloud computing has revolutionized secure computation and fraud detection, offering innovative solutions to persistent challenges across industries. These fields aim to enhance data integrity, transparency, and computational efficiency while addressing fraud prevention concerns, particularly in complex environments like e-commerce, supply chain management, and manufacturing. This survey explores how the combination of blockchain and Python, along with machine learning and cloud computing, provides a comprehensive solution to secure computation and fraud detection in multi-participant systems.

Blockchain technology has gained significant attention in recent years due to its decentralized, tamper-resistant nature. It is particularly relevant for industries that require secure and transparent data handling. For instance, in industrial applications such as manufacturing and supply chain management, blockchain ensures data integrity by maintaining an immutable ledger of transactions. Zheng et al. (2018) emphasize that blockchain's distributed ledger can accurately track and trace activities, making it a valuable tool for securing computation processes. When applied to color loading computations, blockchain can guarantee that each step of the calculation is securely recorded, ensuring both transparency and auditability. This prevents unauthorized modifications and helps organizations maintain the integrity of their data throughout the process.

Python, with its versatility as a high-level programming language, plays a pivotal role in the execution of secure computations. Van Rossum and Drake (2019) note that Python's extensive libraries and its capacity to manage large datasets make it well-suited for computational tasks. Python's integration with blockchain technology enhances the system's ability to handle data analysis and real-time verification, ensuring that computation results are not only secure but also accurate. For example, when combined with blockchain, Python can automate the handling of inputs and outputs, enabling smooth execution of secure computation tasks. This automation can prevent human errors and ensure that only verified data is processed and recorded on the blockchain.

Incorporating smart contracts, as introduced by Buterin (2020) in Ethereum, further enhances the automation of secure computations. Smart contracts allow for the self-execution of predefined terms, which is crucial in environments where security and trust are paramount. These contracts, when implemented alongside Python-based systems, can ensure that computational results meet specific conditions before

they are added to the blockchain. This automation reduces the risk of fraud or errors during computation, making the entire process more efficient and secure.

Fraud detection in multi-participant e-commerce transactions presents significant challenges, primarily due to the involvement of various actors and the ever-evolving nature of fraudulent activities. Traditional fraud detection methods often rely on static, rule-based systems that are not flexible enough to adapt to new fraud patterns. Ngai et al. (2011) highlight the limitations of these traditional methods, emphasizing the need for more adaptive and real-time approaches that can identify emerging fraudulent behaviors. In such dynamic environments, machine learning (ML) and process mining techniques have become increasingly important for detecting fraud. Machine learning, for instance, can analyze vast amounts of transaction data, using supervised learning methods like Decision Trees and Random Forests, or unsupervised methods like anomaly detection, to identify fraudulent patterns. The adaptability of these methods allows systems to learn from new data, thereby improving their ability to detect new types of fraud.

Process mining, as detailed by van der Aalst (2020), complements machine learning by analyzing event logs and extracting meaningful patterns that can help detect deviations from normal transaction workflows. This technique is particularly useful for identifying fraudulent activities that occur due to manipulations of the process itself, such as the altering of timestamps or skipping steps in a transaction flow. Process mining's ability to offer deep insights into user behavior and transaction patterns makes it an essential tool in detecting fraud in e-commerce environments.

The rise of cloud computing has further enhanced the capabilities of fraud detection systems. Cloud services like AWS EC2 provide the scalability and computational power required to process vast amounts of transactional data in real-time. Agrawal et al. (2021) discuss the advantages of cloud-based computing for machine learning, particularly in environments where high volumes of data must be processed quickly. By utilizing cloud platforms like AWS EC2, fraud detection systems can access on-demand compute resources, allowing them to scale up during peak transaction times and ensuring that real-time analysis is not hindered by hardware limitations. This flexibility enables fraud detection systems to remain efficient and responsive, even as the volume of transactions grows.

Cloud computing also allows for the dynamic allocation of resources based on the workload, which is crucial for maintaining optimal performance. With cloud-based fraud detection systems, resources can be allocated to handle more intensive data processing tasks during high-traffic periods, ensuring that the system remains robust even under stress. Moreover, cloud-based systems are continuously updated, meaning that fraud detection models can be updated in real-time to stay ahead of emerging fraud tactics, keeping the system up-to-date and effective.

In fraud detection, a multi-perspective approach is often needed to address the limitations of traditional methods. One such approach is the use of ensemble learning, where multiple machine learning models are combined to improve the overall detection performance. Models like Random Forest and Gradient Boosting are widely used in ensemble learning because they aggregate the predictions of multiple algorithms, improving accuracy and reducing the likelihood of false positives. Breiman (2022) introduced Random Forest, which combines decision trees to make predictions based on majority voting, while Gradient Boosting (Friedman, 2002) builds decision trees sequentially, focusing on correcting errors made by previous trees. These techniques enhance the robustness of fraud detection systems, enabling them to identify complex fraud patterns that may not be detectable by individual models.

By integrating machine learning models with cloud computing services such as AWS EC2, fraud detection systems can operate at scale without being limited by hardware constraints. Cloud infrastructure supports distributed computation, enabling the training and deployment of large-scale models efficiently. The cloud's ability to provide elastic compute capacity ensures that the system can scale up or down as needed, making it a highly flexible solution for handling the dynamic nature of fraud detection in e-commerce transactions.

The convergence of blockchain technology, Python programming, machine learning, and cloud computing offers a powerful, holistic approach to secure computation and fraud detection. Blockchain's ability to provide secure, tamper-proof records of computation processes, combined with Python's computational capabilities, ensures

learning models can be used to detect fraudulent activities, while process mining provides insights into transaction behaviors, helping to identify deviations from normal processes. Cloud computing platforms like AWS EC2 provide the scalability needed to handle large transaction volumes in real-time, ensuring that fraud detection systems can operate effectively even under high demand.

By integrating these technologies into a single solution, industries can develop more efficient, transparent, and resilient systems for secure computation and fraud detection. Blockchain and Python enable secure, transparent data handling, while machine learning and process mining improve fraud detection accuracy and efficiency. Cloud computing ensures that these systems can scale dynamically to meet the demands of modern e-commerce and industrial environments. The integration of these advanced technologies holds great promise for addressing the challenges faced by industries in securing computation processes and preventing fraud, paving the way for more secure, scalable, and effective systems in the future.

3. PROPOSED METHODOLOGY

This proposed methodology focused on improving E-commerce has become a dominant force in the global economy, offering convenience, variety, and ease of access to consumers. However, as e-commerce platforms grow, so does the risk of fraud. Fraudulent activities, including identity theft, payment fraud, and account takeover, are increasingly common in online marketplaces. The rise of sophisticated fraud tactics, such as social engineering, fake reviews, and payment manipulation, poses significant challenges to businesses and consumers alike.

Fraud detection is crucial to maintaining trust and ensuring financial security in e-commerce. It protects both the platform and its customers from financial losses, reputational damage, and data breaches. With e-commerce transactions occurring in real-time, effective fraud detection systems must analyze vast amounts of data quickly, accurately, and at scale.

The proposed fraud detection system seeks to enhance the current fraud detection landscape using a multi-perspective approach. This approach will integrate:

Machine learning: To identify complex fraud patterns and adapt to new fraudulent behaviors.
Behavioral analysis: To monitor and analyze typical user behavior, detecting deviations that may indicate fraud.
AWS EC2: To provide the scalability and real-time processing needed to handle large volumes of transactions and the associated data.
The system will use ensemble classification models to improve fraud detection accuracy and robustness, ensuring that fraudulent activities are flagged accurately without overwhelming the system with false positives.

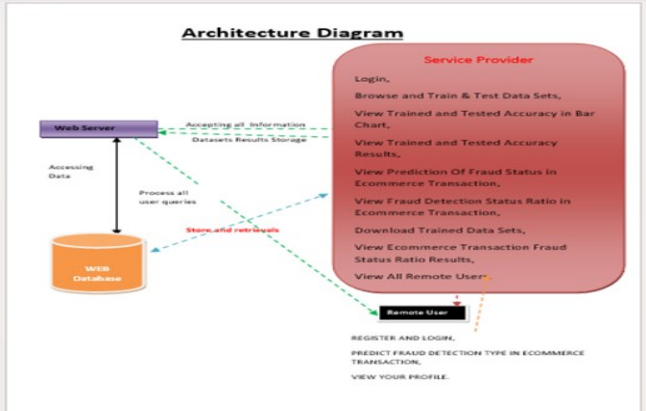

Figure 1: Proposed E-commerce system.

The proposed system is designed with a modular architecture that integrates several components to detect fraud effectively. The system comprises three key stages:

1. **Behavioral Detection:** Analyzes user behavior to establish a baseline and identify deviations.
2. **Abnormality Analysis:** Analyzes transaction data in real-time to detect any anomalies or outliers that suggest fraudulent activity.
3. **Ensemble Classification:** Uses machine learning models to classify whether a transaction is legitimate or fraudulent, combining multiple models to ensure robustness.

AWS EC2 will support this architecture, providing elastic computational resources to process vast amounts of data in real-time.

Data Flow and Integration

The system collects data from various sources:

- **User transactions**: Purchase histories, payment methods, and transaction amounts.
- **User profiles**: Personal information, account details, and past activity.
- **Real-time activity**: User interactions on the platform, such as login times and IP addresses.

This data flows into a data pipeline, which is responsible for preprocessing, integrating, and feeding the data to the different system components (behavioral detection, anomaly analysis, and ensemble models).

Key Components

- **Behavioral Detection:** Models typical user behavior and flags deviations.
- **Abnormality Analysis:** Detects suspicious anomalies in transaction data.
- **Ensemble Classification:** Uses multiple machine learning algorithms to improve the accuracy of fraud detection.

Applications:

Enhanced Data Collection and Integration

- **Integration with Real-Time Data Sources:** The current dataset is static (Datasets.csv), but integrating with real-time transaction data would allow the system to detect fraud continuously. It could pull data from live transaction databases or external APIs.
- **Data Enrichment:** Incorporating additional data sources such as customer profiles, transaction history, device information, and browsing patterns could enhance the system's ability to detect fraud with more context and precision.

2. Advanced Machine Learning Models

- **Deep Learning Models:** While traditional machine learning models like SVM, Logistic Regression, and Decision Trees are effective, using deep learning models (e.g., neural networks) can improve performance for complex datasets. For example, techniques like Recurrent Neural Networks (RNNs) or Convolutional Neural Networks (CNNs) could be explored for sequential data and image-based data, respectively.
- **Anomaly Detection:** Models like Isolation Forest, One-Class SVM, or Autoencoders could be implemented for anomaly detection, where fraudulent transactions are rare events compared to normal transactions.
- **Ensemble Learning:** Combining multiple models into an ensemble (e.g., Random Forests, XGBoost, or Stacked Generalization) could boost prediction accuracy by aggregating the outputs of multiple models.

3. Real-Time Fraud Detection

- **Online Learning:** Implementing online learning or incremental learning models (like SGDClassifier) would allow the system to update its knowledge as new transaction data comes in, improving model accuracy over time without requiring retraining on the entire dataset.
- **Real-Time Analytics Dashboard:** Building a real-time dashboard for administrators that tracks fraud detection performance, visualizes key metrics, and flags suspicious transactions immediately as they occur.

4. Performance Improvement and Model Optimization

- **Hyperparameter Tuning:** Exploring and optimizing hyperparameters using Grid Search or Randomized Search for each machine learning model would enhance their predictive capabilities. Automating this process with automated machine learning (AutoML) frameworks could significantly improve model accuracy.
- **Feature Engineering:** Experimenting with additional feature engineering techniques, such as using transaction time, geolocation, IP address, and session duration, could improve model performance by providing richer information.

5. Enhanced User Interface (UI)

- **User Customization:** Allow users (admins or service providers) to configure fraud detection models (e.g., adjusting thresholds for fraud likelihood), customize alerts, and visualize results in ways that suit their needs.
- **Interactive Visualizations:** Instead of static charts, the system could use interactive visualizations (e.g., Plotly or D3.js) to provide a more engaging and insightful experience when reviewing fraud detection results, such as interactive confusion matrices, ROC curves, and performance trends over time.

Advantages:

The proposed System has various advantages , among a few are listed below:-

Improved Detection Accuracy

By combining behavioral analysis, anomaly detection, and ensemble learning, the system can detect a wide range of fraudulent activities. This multi-perspective approach helps reduce false positives and increases the accuracy of fraud detection.

Scalability and Real-Time Monitoring

The system's ability to scale with transaction volume ensures that it remains effective as e-commerce grows. AWS EC2 provides the infrastructure necessary for real-time monitoring, allowing for quick detection and response to fraudulent activities.

Enhanced Robustness and Reliability

Ensemble models help to improve the system's robustness, reducing both false positives and false negatives. Additionally, AWS's cloud-based infrastructure enhances the system's reliability, providing fault tolerance and ensuring that the system remains operational even during peak transaction times.

Decentralized and Distributed Processing

The use of cloud computing allows the system to avoid single points of failure. By distributing processing across multiple EC2 instances, the system can handle peak loads and improve resilience and uptime.

4. EXPERIMENTAL ANALYSIS

Figure1: Login for Remote User.

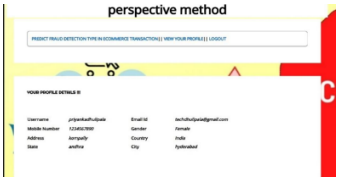

Figure 2: Output for Remote User.

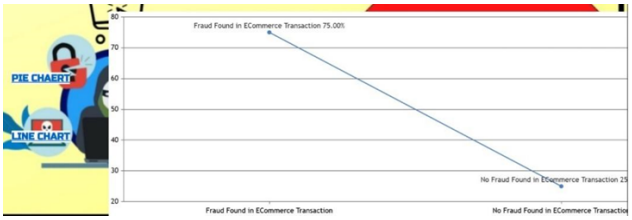

Figure3:Graphical Representation.

Fraud Detection Ratio in Ecommerce Transaction Details

Fraud Status	Ratio
Fraud Found in ECommerce Transaction	75.0
No Fraud Found in ECommerce Transaction	25.0

Figure 4:Fraud Detection Ratio.

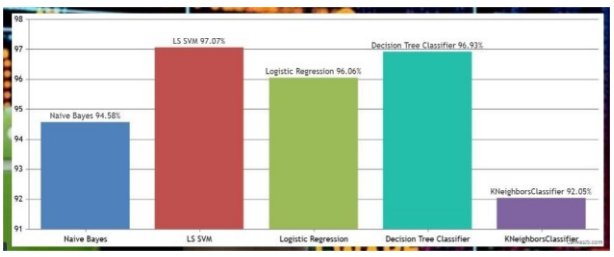

Figure 5:Bar Graph Representation.

Trained and Tested Datasets Results

Model Type	Accuracy
Naive Bayes	94.57700650759219
LS SVM	97.07158351409979
Logistic Regression	96.059291395517
Decision Tree Classifier	96.92697035430224
KNeighborsClassifier	92.04627621113521

Figure 6:Datasets Trained and Tested Results.

5. CONCLUSION

This paper proposed a hybrid method to capture fraud transactions by integrating the formal process modeling and the dynamic user behaviors. We analyzed the e-commerce transaction process under five major perspectives: control flow perspective, resource perspective, time perspective, data perspective, and user behavior patterns. This paper utilized high-level Petri nets as the basis of process modeling to model the abnormal user behaviors and created an SVM model to perform fraudulent transaction detection. Our extensive experiments showed that the proposed method can effectively capture fraudulent transactions and behaviors. The overall index of our proposed multi-perspective detection method outperformed the single-perspective detection method. As our future work, related deep learning [38-42] and model checking methods [43-45] would be incorporated in the proposed framework for higher accuracy. Additionally, it's also a future work to incorporate more time features to the behavior patterns so as to make the risk identification more accurate. Furthermore, we will conduct research on constructing a standard fraud mode library, and apply the proposed methodology to other malicious behavior areas by coordinating the models.

REFERENCES

[1] R. A. Kuscu, Y. Cicekcisoy, and U. Bozoklu, Electronic Payment Systems in Electronic Commerce. Turkey: IGI Global, 2020, pp. 114–139.

[2] M. Abdelrhim, and A. Elsayed, "The Effect of COVID-19 Spread onthe e-commerce market: The case of the 5 largest e-commerce companies in the world." Available at SSRN 3621166, 2020, doi:10.2139/ssrn.3621166.

[3] P. Rao et al., "The e-commerce supply chain and environmental sustainability: An empirical investigation on the online retail sector."Cogent. Bus. Manag., vol. 8, no. 1, pp. 1938377, 2021.

[4] S. D. Dhobe, K. K. Tighare, and S. S. Dake, "A review on prevention of fraud in electronic payment gateway using secret code," Int. J. Res. Eng. Sci. Manag., vol. 3, no. 1, pp. 602-606, Jun. 2020.

[5] Abdallah, M. A. Maarof, and A. Zainal, "Fraud detection system: A survey," J. Netw. Comput. Appl., vol. 68, pp. 90-113, Apr. 2016.

[6] E. A. Minastireanu, and G. Mesnita, "An Analysis of the Most Used Machine Learning Algorithms for Online Fraud Detection," Info. Econ., vol. 23, no. 1, 2019.

[7] X. Niu, L. Wang, and X. Yang, "A comparison study of credit card fraud detection: Supervised versus unsupervised," arXiv preprint arXiv: vol. 1904, no. 10604, 2019, doi: 10.48550/arXiv.1904.10604.

[8] L. Zheng et al., "Transaction Fraud Detection Based on Total Order Relation and Behavior Diversity," IEEE Trans. Computat. Social Syst., vol. 5, no. 3, pp. 796-806, 2018.

[9] Z. Li, G. Liu, and C. Jiang, "Deep Representation Learning With Full Center Loss for Credit Card Fraud Detection," IEEE Trans. Computat. Social Syst., vol. 7, no. 2, pp. 569-579, 2020.

[10] I. M. Mary, and M. Priyadharsini, "Online Transaction Fraud Detection System," in 2021 Int. Conf. Adv. C. Inno. Tech. Engr. (ICACITE), 2021, pp. 14-16.
[11] D. Choi, and K. Lee, "Machine learning based approach to financial fraud detection process in mobile payment system," IT Conv. P. (INPRA), vol. 5, no. 4, pp. 12-24, 2017.
[12] R. Sarno et al., "Hybrid Association Rule Learning and Process Mining for Fraud Detection," IAENG Int. J. C. Sci., vol. 42, no. 2, 2015.
[13] J. J. Stoop, "Process mining and fraud detection-A case study on the theoretical and practical value of using process mining for the detection of fraudulent behavior in the procurement process," M.S.
thesis, Netherlands, ENS: University of Twente, 2012.
[14] M. Jans et al., "A business process mining application for internal transaction fraud mitigation," Expert Syst. Appl., vol. 38, no. 10, pp. 13351-13359, 2011.
[15] C. Rinner et al., "Process mining and conformance checking of long running processes in the context of melanoma surveillance," Int. J.Env. Res. Pub. He., vol. 15, no. 12, pp. 2809, 2018.
[16] E. Asare, L. Wang, and X. Fang, "Conformance Checking: Workflow of Hospitals and Workflow of Open-Source EMRs," IEEE Access, vol. 8, pp. 139546-139566, 2020.
[17] W. Chomyat and W. Premchaiswadi, "Process mining on medical treatment history using conformance checking," in 2016 14th Int. Conf. ICT K. Eng. (ICT&KE), 2016, pp. 77-83.
[18] M. D. Leoni, W. M. Van Der Aalst, and B. F. V. Dongen, "Data- and resource-aware conformance checking of business processes," in Int.Conf. Bus. Info. Sys., Springer, Berlin, Heidelberg, 2012. pp. 48-59.
[19] S. M. Najem, and S. M. Kadeem, "A survey on fraud detection techniques in ecommerce," Tech-Knowledge, vol. 1, no. 1, pp. 33-47, 2021.
[20] K. Böhmer, and S. Rinderle-Ma, "Anomaly detection in business process runtime behavior--challenges and limitations," arXiv preprint arXiv, 2017, doi: 10.48550/arXiv.1705.06659.
[21] K. D. Febriyanti, R. Sarno and Y. Effendi, "Fraud detection on event logs using fuzzy association rule learning," in 2017 11th Int. Conf. Info. Comm. Tech. Sys., Surabaya, Indonesia, 2017, pp. 149-154.
[22] T. Chiu, Y. Wang and M. Vasarhelyi, "A framework of applying process mining for fraud scheme detection," SSRN Electronic Journal, 2017, doi:10.2139/ssrn.2995286.
[23] W. Yang et al., "Show Me the Money! Finding Flawed Implementations of Third-party In-app Payment in Android Apps," in Proc. NDSS, Shanghai, China, 2017.
[24] W. Rui, S. Chen, X. Wang and S.Qadeer, "How to Shop for Free Online--Security Analysis of Cashier-as-a-Service Based Web Stores," in Proc. SSP, Oakland, CA, USA, 2011, pp. 465-480.
[25] E. Ramezani, D. Fahland and W. Aalst, "Where did I misbehave? Diagnostic information in compliance checking," in BPM., Berlin,Germany, Springer, 2012, pp. 262-278.