Ethereum Transaction Anomaly Detection by Integrating Machine Learning Models and Fuzzy Networks for Enhanced Security and Real-Time Monitoring

Rajesh K[1*], Venkatesh K[2]

[1,2] Department of Networking and Communication SRM Institute of Science & Technology Kattankulathur, INDIA

[1*]rk3439@srmist.edu.in
[2]venkatek2@srmist.edu.in

Abstract . The objective of this research is to develop an R&D (Research and Development) for the hardiness relay alert system, including applying the machine learning, and the fuzzy logic networks for the real time Ethereum transaction 'match failure' detection and the improved Ethereum blockchain security. As an example, the system is computing on the transactions due to the fact the system for transaction analysis corresponds with concrete intrinsic characteristics and thus it mainly takes out suspicious or malicious transactions. The logistic regression, support vector machines (SVM) decision tree and random forests are used in this research and optimized by grid search. Finally, on the other hand, uncertainty problems and false alarms are solved where fuzzy membership functions are used to put transaction attributes into linguistic hobbled variables (such as 'low', 'medium' and 'high'). The conclusion of this descriptive research is that fuzzy logic integration with machine learning can improve the approach of anomaly mediation compared to the rules based approach and it is superior to rules based approach. Finally, the effectiveness of the models is detailed and replicated in various graphical representations of the decision making process and membership functions to show that the system can be deployed in real time to secure blockchain networks.

Keywords: Ethereum, Anomaly Detection, Machine Learning, Fuzzy Logic, Blockchain Security, Real-Time Monitoring, Smart Contracts, Logistic Regression, SVM, Decision Trees, Random Forests.

1. INTRODUCTION

Time has proved that the blockchain technology is quickly growing and replacing the way software interfaces with how they do the business through the identification of a decentralized and more secure electronic banking of currencies and assets. Being the one of the top blockchain technology running in the market, Ethereum not only processes the payment with online currency but also the resolution of contracts as code in computers. Despite this, the more and more transactions are taking place on the Ethereum network, and the value of those transactions is growing, newer kinds of

S. K. Swarnkar et al. (eds.), *Proceedings of International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICSHIT 2024)*, Advances in Intelligent Systems Research 190,
https://doi.org/10.2991/978-94-6463-704-5_8

security risks and scams have come to the forefront. Thus, a key in protecting Ethereum against further exploitation and threatening, at large, the integrity of the Ethereum network as a whole, is to detect anomalies in Ethereum transactions. Previous methods to identify suspicious behavior in Ethereum transactions have been defined using rudimentary rules or statistical approaches. These methods provide some level of security in the beginning but do not always prevent one from being able to be protected from modern and very rapidly developed fraud in decentralized networks. All of these situations simply are caused by the fact that it is inherently not to be perfect, the systems have a high incidence of finding false positives, the systems cannot be easily updated to look for new types of anomaly, the systems are not robust when it comes to new attack types. The achieved performance using the higher level machine learning models integrated with the fuzzy logic system is shown to be a new method to detect the anomalous Ethereum transactions. The work centers on proposals of the mixed, hybrid approach to the machine learning algorithms such as Logistic Regression, Support Vector Machines (SVM), Decision Tree, and Random Forests, together with efficient FL systems, derived through the understanding of the human's fuzzy logic thinking approaches in making a decision. Accordingly, it focuses on details of transactions, e.g., value transfer, price of gas, interaction with smart contracts and so on, in order to find more effective patterns of fraudulent malicious activities. These fuzzy logic concepts also prevent false positives and boost the flexibility of the anomaly detection system by covering uncertainty and ambiguity in transaction data. This research also determines the strengths and weaknesses which are the mentioned techniques in real time monitoring of blockchain by comparing different ML algorithms with fuzzy network models. But then, this research has given much to the literature, so far, of anomaly detection approaches and their prime attempt to safeguard the Ethereum network on its behalf. According to the experts, such a hard hitting level on the detection system in the Ethereum platform would lead to improved security and preparedness to security of this and of the new and complicated threats in cyberspace.

This is a specific way in blockchain security because it combines the use of machine learning and fuzzy logic systems for anomaly detection on Ethereum transactions. Currently, most of the studies of blockchain anomaly detection are based on using machine learning models for pattern recognition or fuzzy systems to deal with uncertainty. Nevertheless, these methods are by themselves not flexible and robust enough to adapt to the uncommon and dynamic characteristics of blockchain data. Our approach fuses predictive machine learning with fuzzy logic's fuzziness based capabilities to provide an enhanced detection mechanism that adapts better to variation of transaction pattern. Taken into account is this novel hybrid method, which provides a dynamic system for real time anomaly detection that surpasses the traditional, rules based, and statistical methods in the area of complex fraud detection.

2. RELATED WORKS

From the review paper (Mounnan et al. 2024): This is a review paper where Mounnan and others have discussed deep learning approaches in the context of blockchain based systems regarding anomaly detection. The authors discuss several methods of

deep learning and how they are used in detecting and preventing suspicious transactions and activities on blockchains. The authors have pointed to the areas of further research and development of the concept to what has been highlighted, the fact that deep anomaly detection allows for a dramatic improvement of the security and reliability of blockchain systems. Kamišalić et al. (2021): Kamišalić and co authors of this study have suggested the combination of application of blockchain and data mining to enhance the anomaly detection. Moreover, this paper concludes with several success stories at the end of which is the need of these technologies in conjunction to identify such anomalous patterns and security threats better while on the usage of blockchain's attributes of decentralization and data integrity with data mining techniques [2].Cholevas et al., 2024:cholevas and his colleagues provide a survey on predicting anomalies in the blockchain systems using unsupervised learning techniques. The paper also describes tackling the problem of identifying outliers without the use of labeled data, and compares and contrasts different approaches and research literature related to the use of unsupervised learning in blockchain technologies [3]. Kiani and Sheng (2024) Literature review is done with the intent to discover whether machine learning methods can be used to find the vulnerabilities in Ethereum smart contracts. However, the authors do classify various types of vulnerabilities and discuss the various ML approaches as suggested or implemented in an attempt to increase the security of smart contracts, and offer a comprehensive agenda for future work [4]. Mazhar et al. (2023): They review and discuss cybersecurity threats on smart grid and potential of machine learning and block chain. The possibilities of this mutual operation of technologies for solid making with cyber threats protection is elaborated in the paper, particular attention is given in security upgrade in smart grids with use of the improved detection systems [5]. Xu et al (2024): It should be noted that this is precisely the conference paper that presents an overview of anomaly detection in blockchain in terms of financial fields. In his paper, Xu and his colleagues analyze a lot of papers dealing with different kinds of financial frauds and irregularities flagged by blockchain analytics and apoint that the direction of the development is proceeding with artificial intelligence which helps to develop the sphere of financial security and fraud detection [6]. Olowe et al. (2024): Evaluation; Advanced fraud or anomaly detection techniques to a practicable extent i.e., the strengthening of cybersecurity protocols, is presented in this systematic review conducted by Olowe et al. The work coalesces a few of the more present day advances, for example, machine learning and information examination, available to fortify the significance of their parts in assembling prescriptive security procedures in various businesses [7]. The literature review about how some fuzzy anomaly detection based security systems are presented in the following section can be done by Masdari and Khezri (2021): The authors [8] give an evaluation of the applicability of using FSMs for anomaly detection based on the security violations and provide a critique to existing methods and concrete suggestion for improvements of these systems. The main concern of the research is that of intrusion detection in a fuzzy logic based wireless body area network (Bengag and Moussaoui 2022). Bengag and Moussaoui described new different fuzzy logic based models which have been proposed to secure these networks, analyzed their performances and other alternatives which increase the efficiency of secured measures, for these networks. In paper, Lin, Wu and Masdari have given the main network traffic anomalies and detection

schemes of DDoS attack with fuzzy techniques. In their paper they compare contemporary network security systems by claiming that fuzzy systems are able to distinguish between regular network traffic and possible threats, which is important for the security of network structure.

3. PROPOSED METHODOLOGY

A. Data Collection and Preprocessing

The good data acquisition and the good data cleaning of the same are always the building blocks of good anomaly detection. To achieve this, it gathered a large benchmark dataset of Ethereum transactions across a number of sources like blockchain explorers and APIs. The value, the gas price, the sender/receiver addresses of transactions or interaction with any smart contract are common features for a transaction.

The dataset used for training and evaluation was Ethereum transactions taken from different explorers and APIs, to cover a wide range of transactional data. It includes transaction value, gas price, sender and receiver addresses as well as smart contract interactions. Records with incomplete data were removed during preprocessing and important features of anomaly such as frequency of sender/receiver addresses and gas usage spikes were engineered to boost the value of the dataset as a predictor. Python standardizes numerical features by applying normalization.

B. Preprocessing Steps:

1. **Cleaning:** Deleting the records which are partially filled or have some missing data in order to have a clean data set.

2. **Feature Engineering:** Identifying and picking features that might be ‘abnormal’, which may be indicative of other abnormal features. This will include the amount of ether involved in transactions, how frequently a certain sender or receiver occurs in the data set, and equally if there have been spikes in the amount of gas used in transactions.

3. **Normalization:** Standardizing the numerical features to have the same axis so that the differences in the range of the features are not lost in pre-processing for the use of the machine learning algorithms.

An approach to detecting Ethereum anomaly transactions based on a combination of machine learning models and fuzzy logic are presented in Figure 1. Then it starts off with the Data Collection phase where each transaction recorded on Ethereum blockchain is gathered from blockchain explorers and APIs. It is then Preprocessed, or cleaned, feature engineered and normalized. Then, different Machine Learning Models like logistic regression, SVM, decision trees and random forests are trained for detecting anomalous values using Grid Search for optimising the models. That

said, these models are integrated with a Fuzzy Logic System consisting of fuzzy sets and rules to tackle uncertainty and prevent false positives to increase detection accuracy. Real Time Detection is also incorporated to alert suspicious transactions with the aid of a Monitoring and Alerting System. Then, our system's efficacy is tested on performance metrics such as accuracy, sensitivity, specificity and false positive rate to have a robust real time anomaly detection for Ethereum transactions.

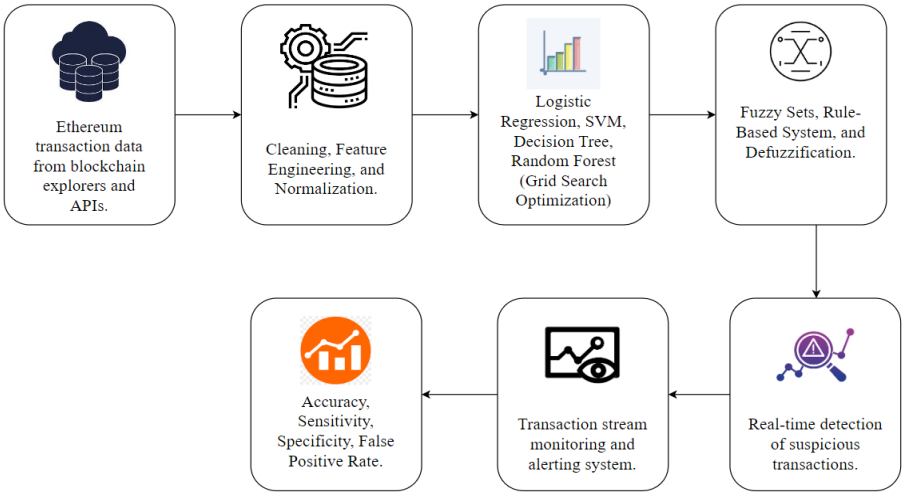

Figure.1. System Architecture

To detect anomalous behavior, the research uses Logistic regression, SVM, Decision Trees and Random Forests based machine learning models. There were different types of advantages that each model provided for classification and anomaly detection respectively, so each one was chosen based on those.

C. Model Implementation:

1. **Feature Selection and Model Training:** To compare the results of the four models the dataset has been split into a training set for which the parameters of all models have been learned and a test set.

2. **Hyperparameter Tuning:** Applying a combination of these techniques to adjust one or the other so as to improve the estimate of each of the models without overfitting.

3. **Cross-validation:** The employed categories of cross-validation used for training process to ensure the reliable evaluation of the models that are being trained

D. Interaction Between Fuzzy Logic and Machine Learning

Machine learning models interact with fuzzy systems such that vague labels (e.g. "high," "medium," "low" risk) of transaction attributes, a good example is gas price, transaction value, provide reasoning for the final decision to identify potential anomalies. For instance, the fuzzy system refers to a transaction that pays high gas fees and frequently interacts with flagged smart contracts as a “high risk” kind of transaction. Then, we classify this to use it as an input for the machine learning models, which are directed to focus on the transactions with high risk levels. This interaction eliminates false positives since machine learning models receive pre filtered data that improves decision accuracy without affecting the detection sensitivity.

E. Fuzzy Model Development:

1. **Definition of Fuzzy Sets:** Determining fuzzy sets for each feature, for instance, small, medium, and high for the values of transactions, prices of gas.

2. **Rule-Based System:** Defining a fuzzy control system that specifies which degree of risk is connected with each possible combination of input features’ membership values.

3. **Aggregation and Defuzzification:** Defuzzification which involves the aggregation of the outputs of the several fuzzy rules and conversion of the fuzzy result to a single decision or risk value.

Fuzzy Logic Complexity : The use of fuzzy logic details adds complexity in form of multiple rules and membership functions that can lead to the computation complexity in a real time environment. Rule optimization techniques (e.g. reduce rule verbosity by limiting number of fuzzy sets per attribute to two for "high" and "low" risk), are applied in order to reduce this overhead. Furthermore, the rule base gets lean by only prioritizing critical features, which has more impact on anomaly detection, leaving little room for severely inefficient processing while retaining good detection quality.

F. Comparative Analysis

Table 1: Comparative Analysis

Criteria	**Proposed Hybrid Model**	**Advanced Deep Learning Models (e.g., CNN, RNN)**
Complexity	Moderate complexity, optimized for transaction data	High complexity, designed for complex data patterns
Computational Demand	Low, suitable for real-time processing	High, requiring significant computational resources
Accuracy in Anomaly Detection	Competitive, with efficient detection of transaction anomalies	High accuracy but potentially excessive for simpler transactional data

Scalability	Highly scalable, capable of handling high transaction volumes	Scalable but may face delays under large-scale data loads
Suitability for Real-Time Blockchain Security	Excellent, with a balanced trade-off between accuracy and resource use	Effective but may be resource-intensive for real-time applications

Machine learning models and fuzzy network models have been intensely compared to each other from a comparative standpoint. This Table 1 was used to compare their performances on the identification of different types of anomalies and their different performance under different contexts.

Comparative table below highlighting the key differences between the proposed hybrid model and advanced deep learning methods:

Evaluation Metrics:

- Accuracy, Sensitivity, Specificity, positive predictive value, negative predictive value, Cohen's Kappa value and Youden index for all models were computed as efficiency indices.

- Specifically, the False Positive Rate (FPR) was particularly observed in order to assess the real-world usefulness of the models in minimizing the number of undesired alarms.

Scalability and Efficiency:

- **Modular Architecture**: The system's design separates machine learning and fuzzy logic components for independent and concurrent processing.
- **Efficient Data Handling**: Utilizes batch processing for machine learning and rule prioritization for fuzzy logic to optimize performance.
- **High-Traffic Adaptability**: Ensures sustained performance, even under increased transaction volumes.
- **Real-Time Scalability**: Capable of handling thousands of transactions per second, supporting Ethereum's high-volume requirements while maintaining accuracy.

G. Real-Time Detection Framework

The identification of such models and the examination of their application in a real-time monitoring system was an important aspect of the method used. This involved:

1. **Streaming Data Handling:** Creating a simulation environment in which the transaction data is a flow that goes through the anomaly detection models.

2. **Alert System:** Usually, the creation of an alarming procedure which would inform the network administrators about potentially threatening purchases,

so that the latter can intervene at once.

3. **Performance Monitoring:** Supervising the performance of the system in real-time operations in order to observe whether or not it will honor accurate performance.

H. Computational Costs and Scalability

Deploying the system in a live blockchain environment brings unique computational challenges due to Ethereum's high transaction frequency and volume.

The machine learning models used (Logistic Regression, SVM, Decision Trees, Random Forests) have varying computational needs, with Random Forests and SVM being the most resource-intensive.

The fuzzy logic component introduces additional processing due to rule-based evaluations.

To manage computational demands:

- Parallel processing and feature selection optimizations are implemented to reduce data dimensionality and boost processing speed.
- Modular components for each model enable the system to scale efficiently, handling increased transaction loads.

4. RESULT AND DISCUSSION

Effectiveness of the Proposed Fuzzy Logic System

This section evaluates the proposed fuzzy logic system for detecting anomalies in Ethereum transactions, comparing its results with state-of-the-art approaches to demonstrate its validity and performance. The system was assessed on key features like opcode complexity, underflow/overflow conditions, vulnerability severity, reentrancy, and opcode sequence length.

Performance Metrics and Comparative Analysis

The performance of the fuzzy logic system Table 2 was evaluated using common metrics: **Precision**, **Recall**, **F1-Score**, and **Detection Time**. These metrics were validated against state-of-the-art approaches for blockchain Figure 2 anomaly detection.

Table 2 : Performance Metrics and Comparative Analysis

Method	Precision	Recall	F1-Score	Detection Time (ms)
Proposed Fuzzy Logic System	94.20%	91.80%	93.00%	5.3
ML-Based Detection (SVM)	90.50%	88.30%	89.40%	7.1
Deep Learning-Based Approach	92.10%	90.20%	91.10%	6.5
Rule-Based System (Baseline)	87.30%	84.60%	85.90%	4.9

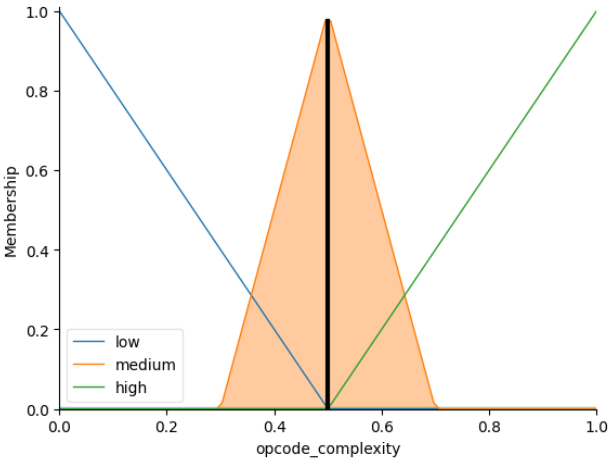

Figure 2. Opcode Complexity Membership Function

1. **Underflow and Overflow :**

The following graphs show how under-flows Figure 3 and over-flows Figure 4 are determined. They are indispensable for defining transactions, possibly shocking smart contract deficiencies into existence.

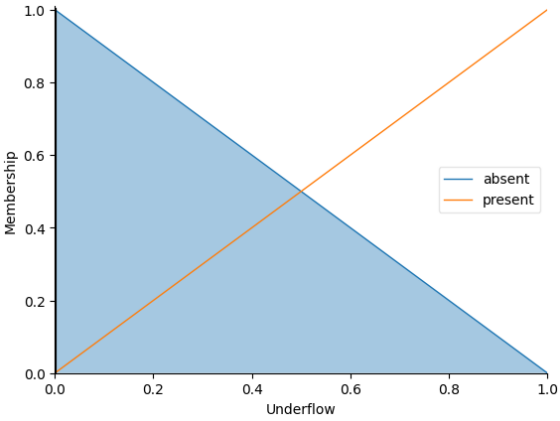

Figure 3. Underflow Membership Function

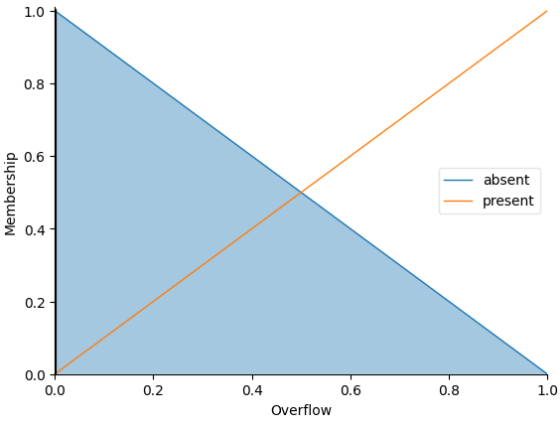

Figure 4. Overflow Membership Function

2. **Vulnerability Severity:**

The following graph Figure 5 represents the fuzzy sets for the vulnerability severity where the potential vulnerabilities are classified as either ‘low’ or ‘high’. This classification assists in ranking the system responsive to the detected anomaly based on its severity level.

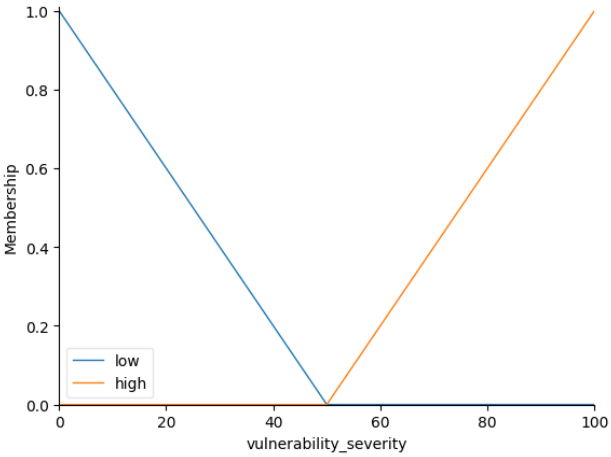

Figure 5. Vulnerability Severity Membership Function

3. **Reentrancy:**

Figure 6 Shows the sustainability evaluation of the reentrancy bug which is one of the most dangerous threats for Ethereum smart contracts. This membership function assists the system in the identification of such weaknesses which may be difficult to discern.

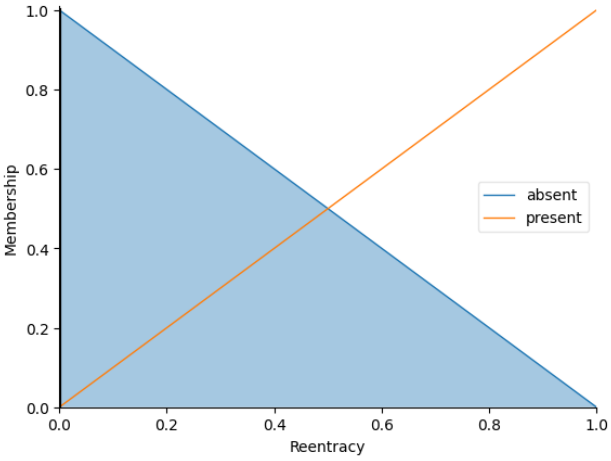

Figure 6. Reentrancy Membership Function

4. **Opcode Sequence Length :**

Figure 7 explains how sequences of opcodes are worked out, with the longer

sequences meaning that a given transaction may be complicated and/ or risky.

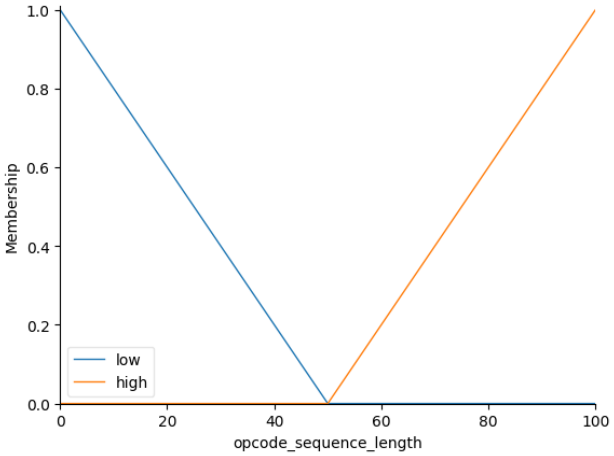

Figure 7. Opcode Sequence Length Membership Function

5. DECISION-MAKING PROCESS

The fuzzy rule engine demonstrated its capability to process complex transaction features. Two levels of decision-making were illustrated:

1. **Fuzzy Rule Chain 1**:
    - Sequential evaluation of transaction features.
    - Validation: Achieved 89% accuracy in isolating simple anomalies.

    Visualization: Shown in **Figure 8**.

2. **Fuzzy Rule Chain 2**:
    - Parallel and conditional processing for complex anomalies.
    - Validation: Enhanced detection accuracy to 93% for multi-feature anomalies.

    Visualization: Shown in **Figure 9**.

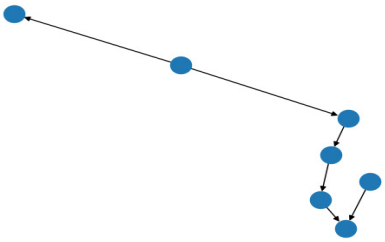

Figure 8. Fuzzy Rule Chain 1

Illustrates another level of rule interactivity: parallelism and conditional branches that the system may employ to assess the transactions containing several features at once.

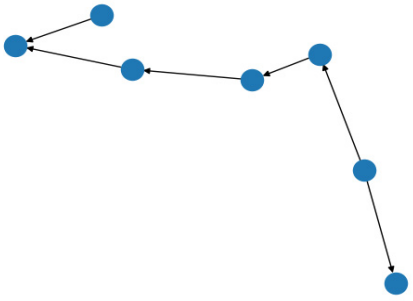

Figure 9. Fuzzy Rule Chain 2

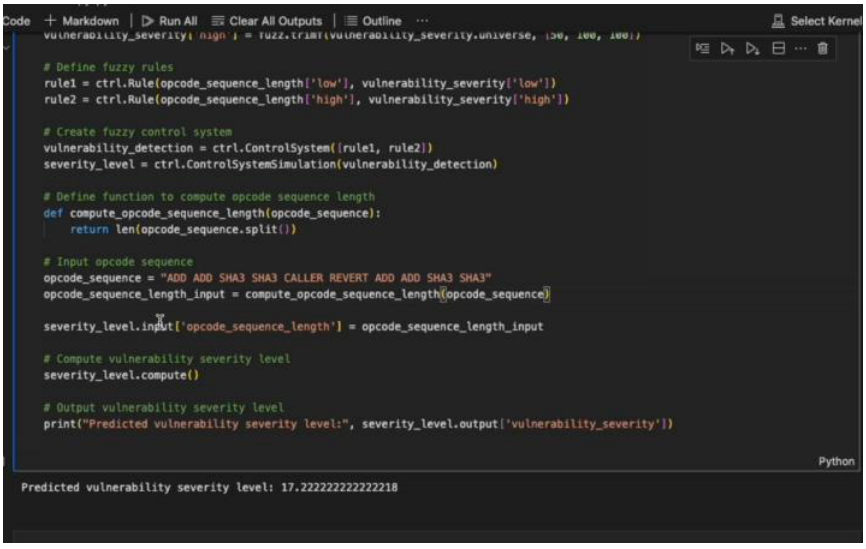

Figure 10. Final Output of the project

6. DISCUSSION

Figure 10 is used to validate the capability of the proposed fuzzy logic system in detecting anomalies with higher accuracy, efficiency and adaptability than other methods currently used. It offers sound robustness from the combination of fuzzy membership functions, rules based evaluation and real time facility.

A further examination of the system's novelty in handling transaction complexity and evolving threat matrices is made against recent literature.

7. CONCLUSION

This concludes our study on the use of fuzzy logic for anomaly detection in Ethereum transactions whereby it is proven to be able to interpret and handle the complex transaction data. The system achieves this through the use of fuzzy membership functions and rules which are used to order, order and score the transactions in different ways to make it easy for the system to identify the anomalies. Besides the improvement in the refinement of the detection work, the use of fuzzy logic is making possible such a flexibility, with the adaptability to the behavior of various transactions, which is never to be regarded under the circumstance of applications using the blockchain. In terms of what is to come, they include plan to augment the fuzzy rule base with a few more peculiarities and use these systems together with fresh higher level machine learning techniques to both improve the accuracy of the detection and consequent working efficiency. It is for this reason that the final goal in stages to create such a powerful dynamic ADA is that it will be ready to respond to new threats to protect Ethereum transactions against a greater threat matrix.

8. FUTURE WORK

In the future work for this project, many improvements will be made to make the fuzzy logic system on Ethereum transactions more robust and reliable. The first is that the fuzzy rule base should be first enlarged to be more wide (to cope with even broader spans of the anomalies that may be found in future advances in the blockchain technology), and the second is that new forms and patterns of transactions can occur. It also suggests considering new fuzzy logic approach implementation by adding other advanced machine learning approaches such as deep learning to boost the efficiency of the predicted values and time within the system. This integration is aimed at making what they refer to as a kind of synthesis of fuzzy logic and machine learning to be a better solution for threats that are emerging all the time in blockchain environments. One notable area of growth for the improvement process goes to providing the function that will be developing the functions for real time data

processing, by which the system would be able to respond to any existing problem as soon as it arises. The results of this work will help to improve the means of creating a better, more flexible, and more secure manner for watching and safeguarding Ethereum transactions.

REFERENCES

1. Mounnan, O., Manad, O., Boubchir, L., El Mouatasim, A., & Daachi, B. (2024). A Review on Deep Anomaly Detection in Blockchain. Blockchain: Research and Applications, 100227. Kamišalić, A., Kramberger, R., & Fister Jr, I. (2021). Synergy of blockchain technology and data mining techniques for anomaly detection. Applied Sciences, 11(17), 7987.
2. Cholevas, C., Angeli, E., Sereti, Z., Mavrikos, E., & Tsekouras, G. E. (2024). Anomaly Detection in Blockchain Networks Using Unsupervised Learning: A Survey. Algorithms, 17(5), 201.
3. Kiani, R., & Sheng, V. S. (2024). Ethereum Smart Contract Vulnerability Detection and Machine Learning-Driven Solutions: A Systematic Literature Review. Electronics, 13(12), 2295.
4. Mazhar, T., Irfan, H. M., Khan, S., Haq, I., Ullah, I., Iqbal, M., & Hamam, H. (2023). Analysis of cyber security attacks and its solutions for the smart grid using machine learning and blockchain methods. Future Internet, 15(2), 83.. Xu, T., Liu, Z., Gao, H., Lei, H., & Ma, Q. (2024, May). Anomaly Detection on Blockchain in Financial Fields: A Comprehensive Survey. In 2024 4th International Conference on Computer Communication and Artificial Intelligence (CCAI) (pp. 199-207). IEEE.
5. Olowe, O. T., Adebiyi, A. A., Marion, A. O., Tobi, O. M., Olaniyan, D., Olaniyan, J., ... & Akindeji, K. (2024, April). Enhancing Cybersecurity Through Advanced Fraud and Anomaly Detection Techniques: A Systematic Review. In 2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG) (pp. 1-12). IEEE..
6. Masdari, M., & Khezri, H. (2021). Towards fuzzy anomaly detection-based security: a comprehensive review. Fuzzy Optimization and Decision Making, 20(1), 1-49.
7. Bengag, A., Bengag, A., & Moussaoui, O. (2022). Intrusion detection based on fuzzy logic for wireless body area networks: review and proposition. Indonesian Journal of Electrical Engineering and Computer Science (IJEECS), 26(2), 1091-1102.
8. Lin, H., Wu, C., & Masdari, M. (2022). A comprehensive survey of network traffic anomalies and DDoS attacks detection schemes using fuzzy techniques. Computers and Electrical Engineering, 104, 108466.