DETECTION OF FRAUDULENT PHONE CALLS DETECTION IN MOBILE APPLICATIONS

Dr K. BHARGAVI[1], B. MITHILA SHIVANI[2]

[1]Associate Professor & HOD Dept Of C.S.E, PVKK Institute of Technology Anantapur, Andhra Pradesh- 515001

[2]PG Scholar, Dept Of C.S.E, PVKK Institute of Technology Anantapur, Andhra Pradesh- 515001

ABSTRACT: The primary challenge faced over the course of this decade-long endeavour is the difficulty in devising effective features without direct access to telephony network infrastructure. we conducted an extensive three-month measurement study using these call logs, which encompassed a staggering 9 billion records. Based on the insights gleaned from this study, we identified and designed 29 features that could be used by machine learning algorithms to predict malicious calls. Fraudulent phone calls or scams and spam s via telephone or mobile phone have become a common threat to individuals and organizations. Artificial Intelligence (AI) and Machine Learning (ML) have emerged as powerful tools in detecting and analyzing fraud or malicious calls. This paper presents an overview of AI-based fraud or spam detection and analysis techniques, along with its challenges and potential solutions. The novel fraud call detection approach is proposed that achieved high accuracy and precision. The outcomes revealed that the most effective approach could reduce unblocked malicious calls by up to 90%, while maintaining a precision rate exceeding 93.79% for benign call traffic. Moreover, our analysis demonstrated that these models could be implemented efficiently without incurring significant latency overhead.

KEYWORDS: Malicious, Artificial Intelligence, Fraudulent, Machine learning.

1. INTRODUCTION

An ever-evolving danger that affects people, businesses, and the government is phone-based spam or scams [7]. The Federal Trade Commission (FTC) in the United States got over 3 million reports of fraud in 2021, resulting in a $3billion-dollar loss overall. Spammers use a variety of ploys, including impersonation, spoofing, and digital manipulation, to access private information, steal money, or harm a person's image. Around the globe, it resulted in financial and information losses. Inherently, fraud phone calls are designed to cause stress and anxiety. The traditional methods [14] of detecting malicious phone calls involve manual review of call details and recordings and identifying fraudulent patterns. However, these methods are time-consuming, expensive, and may not give accurate results or effective in identifying new types of scams. Therefore, there is a need for a good technique that can detect and analyse fraud phone calls accurately and efficiently.

The key contributions of our work are summarized as follows:

- We design and construct a classifier based on Calling Detail Records (CDR) for fraudulent phone call recognition. The classifier only uses the CDR as input data, so it can be constructed easily, quickly, and efficiently. It provides a basic framework for recognition task and defines the main steps of the task.
- Our study provides the first systematic exploration of state-of-the-art machine learning algorithms applied to fraudulent phone call recognition, namely, we design, tune, and evaluate three models—the (RNN), SVM Our ML models are capable of automatically learning phone number features and call behavior features for fraudulent phone call recognition. We demonstrate that our ML-based approach achieves a higher accuracy rate than the state-of-the-art approaches.
- We reevaluate previous work on our new real-world datasets. As a result of a systematic comparison of our novel ML-based approach to previous fraudulent phone call recognition approaches, we demonstrate comparable recognition results with slight improvements of up to 3.0%-4.7% on average. Furthermore, our ML models reveal more general and stable phone number features and call behavior features of fraudulent phone calls than the state-of-the-art approaches, which make them more robust to concept drift caused by a highly dynamic fraudulent phone number and its call behavior.
- We make the generated dataset publicly available, allowing researchers to replicate our results and systematically evaluate new approaches to fraudulent phone call recognition.

2. LITERATURE SURVEY

Artificial intelligence techniques such as machine learning, deep learning, neural language processing have been applied to detect and analyse fraud phone calls. These techniques can provide accurate results for already existing models and systems. The authors [3] introduces a solution based on machine learning for telecommunication without making any harm to telephone network infrastructure. The experiment was stimulated by mat plot lib. The findings show that the method proposed is 87% accurate. The paper [5] introduces a mining-based phone call recognition framework. The experiment show that the technique can achieve the high reputation precision regarding 97.6% [5] which exhibits that the proposed methods has a brilliant execution with best draws. The research proposed the solution of detecting fraud phone calls by using historical data to pre process the data [6]. And as a result, artificial neural network is better method for detecting telephone frauds due to speed and accuracy.

3. SYSTEM ANALYSIS

3.1EXISTING SYSTEM

The existing system for "Detection and Analysis of Fraud Phone Calls using Artificial Intelligence" primarily relies on traditional methods and rule-based systems, often struggling to keep pace with evolving fraud tactics. Conventional call filtering techniques exhibit limitations in accurately distinguishing between legitimate and fraudulent calls. The absence of advanced machine learning models hinders the system's ability to adapt to dynamic fraud patterns. Additionally, the lack of comprehensive analysis tools results in a limited understanding of fraudsters' evolving strategies. This underscores the need for a more sophisticated approach, prompting the integration of Artificial Intelligence and Machine Learning to enhance fraud detection accuracy and provide valuable insights into the techniques employed by malicious actors.

3.2 LIMITATIONS OF EXISTING SYSTEM

Rule-Based Approach Limitation: The existing system relies on rule-based approaches, making it less adaptive to emerging and sophisticated fraud techniques that often evolve beyond predefined rules.
False Positive Challenges: The current system may generate false positives, inaccurately flagging legitimate calls as fraudulent. This can lead to user frustration and decreased trust in the effectiveness of the fraud detection system.

3.3 PROPOSED SYSTEM

The proposed system for "Detection and Analysis of Fraud Phone Calls using Artificial Intelligence" represents a paradigm shift, leveraging advanced machine learning algorithms for dynamic adaptation to evolving fraud tactics. By integrating cutting-edge anomaly detection techniques, the system aims to overcome the limitations of rule-based approaches, enhancing accuracy and significantly reducing false positives. This solution incorporates a comprehensive learning model, continuously evolving through real-time data analysis to identify emerging patterns of fraudulent behaviour. Implementation of deep learning algorithms allows for a nuanced analysis of complex fraud patterns, providing a more robust and efficient detection mechanism. Additionally, the proposed system prioritizes user feedback and employs a feedback loop mechanism to further refine its accuracy and reduce false alarms. Scalability is a core focus, ensuring the system's effectiveness in handling the escalating volume of calls while maintaining real-time responsiveness. Overall, the proposed system aims to establish a state-of-the-art framework, offering heightened accuracy, adaptability, and insights into evolving fraud strategies in the realm of phone calls.

4. SYSTEM ARCHITECTURE

Fig 4: System Architecture

The architecture relies on a comprehensive dataset and metadata for analysis. Machine learning algorithms are deployed to detect patterns in the data. The output provides real-time insights into the authenticity of phone calls.

4.1 METHODOLOGY

- First remove duplicate data and missing values from the set.
- Transform categorical features such as call type, caller ID into numerical features. Using label encoding.
- Normalization of the numerical features, such as call duration, frequency, is done by using z-score normalization.
- Selection of suitable artificial intelligence-based model such as RNN, support vector machine (SVM), decision tree etc. Implement the model and train the pre-processed data. Here, RNN is selected for training and testing of the dataset.

5. MODULES

i)Data collection: The dataset of phone call recordings, along with metadata such as call duration, location, phone numbers, etc. is collected from real world source like. Dataset is taken from real world sources such as Kaggle. The dataset contains 1000 genuine and fraudulent calls shown in fig 6.1 and the following features such as State, area code, a phone number, date and time, IP address, code, etc. this dataset is divided into two parts training set and testing set.

ii)Data Pre-processing: Data cleaning and pre-processing is used for dataset to remove noise, distortions, and irrelevant information.

iii) Model Evaluation: Selection of suitable artificial intelligence-based model such as RNN, support vector machine (SVM), decision tree, etc. Implement the model and train the pre processed data. Here, support vector machine and recurrent neural network is selected for training and testing of the dataset.

iv)User Interface (UI) Development: Constructs an intuitive and interactive user interface (UI) for CADM, facilitating user interaction and visualization of results generated algorithms.

6. RESULT

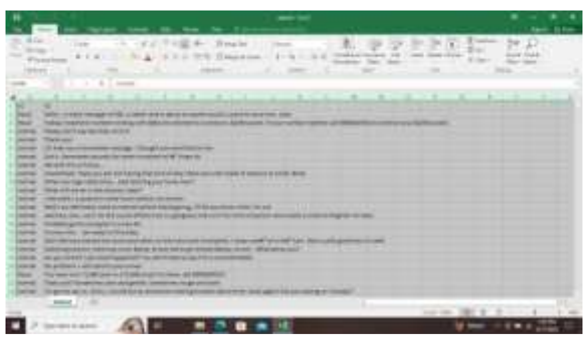

Fig 6.1: Phone Call database from Kaggle

Fig 6.2: Phone Call Prediction by using Machine Learning

The proposed approach was evaluated on a dataset of 1000 genuine and fraudulent calls. the approach achieved a high accuracy of 90% and precision of 93.79%, outperforming the existing approaches. Hence, the approach was able to detect fraudulent phone calls brilliantly. Below figures shows the results expected and actual result.

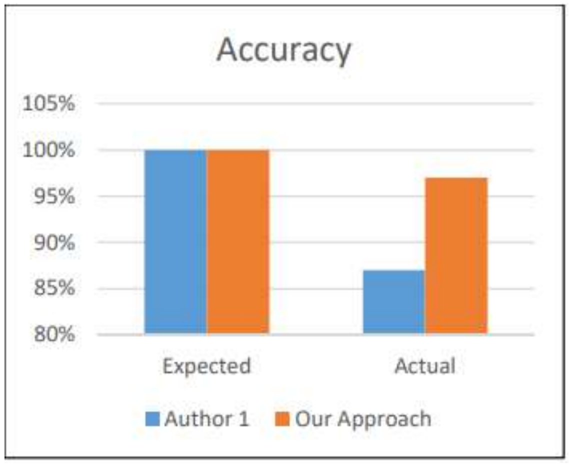

Fig 6.3: Accuracy of Fraudulent Calls Detection

In above, fig 6.3 it shows the accuracy of the approach applied by other author and the approach applied in this paper. The figure is shown in the form of expected result versus actual result. The expected accuracy is 100% and actual accuracy is 87% for author 1, whereas, the expected accuracy is 100% and actual accuracy is 97% for our approach. Hence, our approach is better.

7. CONCLUSION & FUTURE SCOPE

Fraudulent phone calls are a growing concern that affects individuals as well as organizations worldwide. The main purpose of this paper is to detect and analyze fraud phone calls using artificial intelligence. For achieving this goal, RNN, support vector machine (SVM), algorithms are used. The approach achieved a high accuracy and precision. In future To achieved a high accuracy and precision Implementing block chain for storing call records can provide a tamper-proof and transparent system for tracking calls, thereby reducing the chances of fraud. Hence, it will be a good solution to detect and analyze fraud or malicious calls.

REFERENCES

[1] P. Sornsuwit, and S. Jaiyen, “A new hybrid machine learning for cybersecurity threat detection based on adaptive boosting,”Applied Artificial Intelligence, 33(5), pp.462-482, 2019.

[2] K.Shaukat, S. Luo, S.Chen, and D. Liu, “Cyber threat detection using machine learning techniques: A performance evaluation perspective,”in IEEE international conference on cyber warfare and security. IEEE, October2020,pp. 1-6.

[3] S. M. Gowri, G. Sharang Ramana, M. Sree Ranjani and T. Tharani," Detection of Telephony Spam and Scams using Recurrent Neural Network (RNN) Algorithm," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2021, pp. 1284-1288, doi: 10.1109/ICACCS51430.2021.9441982.

[4] Abidogun, Olusola Adeniyi. "Data mining, fraud detection and mobile telecommunications: call pattern analysis with unsupervised neural networks." PhD diss., University of the Western Cape, 2005.

[5] S. Sandhya, N. Karthikeyan, R. Sruthi “Machine learning method for detecting and analysis of fraud phone calls datasets” International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878 (Online), Volume-8 Issue-6, March 2020

[6] Mohammad Iquebal Akhter, Dr. Mohammad Gulam Ahamad “Detecting Telecommunication fraud using neural networks through data mining” international Journal of Scientific & Engineering Research, Volume 3, Issue 3, March-2012.

[7] I. Murynets, M. Zabarankin, R. P. Jover and Panagia, "Analysis and detection of SIMbox fraud in mobility networks," IEEE INFOCOM 2014 - IEEE Conference on Computer Communications,Toronto,ON,Canada,2014,pp.1519-1526,doi: 10.1109/INFOCOM.2014.6848087.

[8] Crawford, M., Khoshgoftaar, T.M., Prusa,J.D. et al. Survey of review spam detection using machine learning techniques. Journal of Big Data 2, 23 (2015). doi:10.1186/s40537-015-0029- 9.

[9] Marzuoli A, Kingravi H, Dewey D and Pienta R. (2016). Uncovering the Landscape of Fraud and Spam in the Telephony Channel 2016 15th IEEE International Conference on Machine Learning and Applications(ICMLA).10.1109/ICMLA.2016.0 153. 978-1-5090-6167-9. (853- 858).

[10] B. Teh, M. B. Islam, N. Kumar, M. K. Islam and U. Eaganathan,"Statistical and Spending Behavior based Fraud Detection of Card-based Payment System,"2018 International SSConference on Electrical Engineering and Informatics (ICELTICs), Banda Aceh, Indonesia, 2018, pp. 78-83, doi:10.1109/ICELTICS.2018.8548878.

[11] H. Tu, A. Doupe, Z. Zhao, and G.-J. Ahn, " Sok: Everyone hates 'robocalls: A survey of techniques against telephone spam," 2016 IEEE Symposium on Security and Privacy (SP), pp. 320 338, 2016.

[12] M. Crawford, T.M. Khoshgoftaar, J.D Prusa, A.N. Richter, H. Al Najada, "Survey of review spam detection using machine learning techniques", Journal of Big Data, 2, pp. 1-24, 42015.