isrjournals

International Journal of Advanced Research in Computer Science Engineering and Information Technology

Volume: 6, Issue: 3,Special Issue: 2 ,Apr,2024 ,ISSN_NO: 2321-3337

EFFECTIVE TRACKING OF FRAUDULENT ACTION FROM MULTI FUND TRANSACTION USING BIG DATA AND MACHINE LEARNING

Sathish Kiran.V[1], Mutharasan. R[2], Venkatesh Reddy.V[3], Mr.Vinayagam T. A[4*]

[123] UG Scholar-Dept.CSE, GRT Institute of Engineering and Technology Tiruttani,India.

[4*]Assistant Professor-Dept.CSE, GRT Institute of Engineering and Technology Tiruttani, India.

joulasathish@gmail.com , mutharasan9345@gmail.com ,victoryvenky3012@gmail.com

***Corresponding Author: vinayagam.ta@grt.edu.in**

Abstract

In scenarios involving cloud services, a key challenge arises from the potential differences in how cloud vendors isolate and connect virtual machines across various cloud networks. Our approach addresses this challenge by prioritizing the tenant's needs, allowing them to dictate their preferred connectivity methods. We integrate Blockchain technology into this project, enabling secure and transparent data management. Our implementation encompasses both Public and Private cloud data storage, with Private storage reserved for sensitive data while Public storage caters to regular data needs. Specifically tailored for the banking sector, our system aims to analyze user behavior while safeguarding personal information. It achieves this by aggregating and monitoring all user transactions, including banking activities, land registrations, gold purchases, and any cash transactions exceeding Rs. 20,000.

Keywords: *Cloud services, virtual machines, cloud networks, tenant-driven approach, Block chain technology, banking system, user behavior, personal identification, transactions, integration, sensitive data, transparency.*

1. Introduction

Dirty money is cash earned or acquired through illegal means, often tied to criminal activities. Tax evasion involves the process of making this dirty money appear legitimate, or "clean." In many countries, combatting crime includes prioritizing efforts against tax evasion. Governments have historically relied on enacting specific laws, such as the Tax Evasion Control Act and the Proceeds of Crime Act, to deter tax evasion. However, while these laws may increase the difficulty of tax evasion and provide some preventive measures, they are not always effective at detecting tax evasion in action. Consequently, the task of anti-tax evasion often demands significant manual labor from specialists trained in this field. Today, with the advancements in machine learning and big data analysis, there are new opportunities to streamline anti-tax evasion efforts. However, to date, no existing methods adequately address this issue. This work aims to fill that gap by developing an intelligent approach to identify suspicious tax evasion activities, thereby reducing the manual effort required for anti-tax evasion measures.

2. Related Work

To explores different crimes associated with Bitcoin. It particularly focuses on detecting Tax Evasion by examining mixing services, which are often utilized for Tax Evasion. This detection method is an integral part of the anti-money laundering (AML) strategy. By analysing transaction sample data involving mixers, we can identify if these services are employed in specific transactions. Tax Evasion with Bitcoin is commonly employed to evade fund tracking in illicit activities, making it crucial to analyse for situational awareness in fund tracking [1].

Tax evasion is when people avoid paying taxes, often by doing illegal things. Many countries think it's important to stop tax evasion. But it's hard because money launderers split their dirty

money and move it around through different bank transfers or buying and selling things. Finding tax evasion manually is tough. So, this project creates a smart way using machines and data analysis to find suspicious accounts that might be evading taxes. First, it finds all the suspicious accounts. Then, it looks deeper to find the most suspicious ones. When tested with Bank Sino Pac's data, this method found 26.3% of the tax evasion cases in the first phase, which is three times better than what the law in Taiwan could do (8.6%). Later, in the second phase, it could be even more accurate, up to 87.04% [2].

Virtual currencies, like Bitcoin, are becoming more popular, but they're also linked to tax evasion. Governments worldwide are trying to fight tax evasion, but many worry about how virtual currencies contribute to the problem. Some cryptocurrencies, such as Bitcoin, make it easier for criminals to avoid paying taxes online because they're decentralized, anonymous, and transactions can't be reversed. This paper explores how efforts to stop tax evasion clash with the rise of cryptocurrencies like Bitcoin. It also examines the case of Liberty Reserve to show the difficulties these currencies bring [3].

This paper discusses using computer methods to process financial data as a practical way to reduce various crimes in the financial sector. It introduces a new method called Adaptive Neuro-Fuzzy Inference System (ANFIS) to identify Tax Evasion in banks and currency exchanges. This method, implemented through MATLAB software, offers an alternative to traditional approaches for detecting Tax Evasion in suspicious banking transactions. It can also be used online within banking systems to analyze customer account data and monitor the risk of Tax Evasion in currency exchanges. One key advantage is its ability to categorize different types of customers. [4].

Tax evasion, also known as money laundering (ML), involves the process of cleaning illicit funds to obscure their origin. Detecting instances of tax evasion poses a significant challenge due to the vast volume of financial transactions occurring daily in the global market. This paper introduces a new approach to identifying tax evasion transactions within extensive financial datasets efficiently and accurately. Our proposed framework employs case reduction techniques to progressively shrink the dataset size. Subsequently, it examines the reduced data to identify pairs of transactions sharing common attributes and behaviours indicative of potential involvement in money laundering activities. Furthermore, the framework utilizes clustering methods to uncover potential money laundering groups. Preliminary experimental results are presented to illustrate the efficacy of our proposed framework [5].

3. Objective

The main goal of our project is to detect any signs of money laundering, which is when people try to disguise illegal money by making it seem like it comes from legal sources. We'll carefully examine all their financial activities, including their transactions in banks and other financial institutions. If we notice anything unusual or suspicious, we'll mark those individuals as potentially involved in illegal activities. Our next step will be to inform the government about these findings so they can conduct further investigations. By doing this, we aim to play our part in preventing financial crimes and maintaining the trust and security of the financial system.

4. Proposed System

Our proposed system combines public and private cloud data storage for enhanced security and accessibility. The private cloud is reserved for storing sensitive information, while the public cloud handles regular data. This system is tailored for banking operations, where we aim is to analyse user the behaviour comprehensively, if utilizing personal identification. By integrating transactions such as banking activities, land registrations, gold purchases, and cash transactions exceeding Rs. 20,000, we can provide users with a holistic

view of their financial interactions. The system offers several advantages, including a tracking system for property purchases, notifications for income tax payments, and secure cloud storage for both public and private information. This approach ensures transparency, compliance, and data protection in banking operations, ultimately benefiting both users and financial institutions.

5. Architecture Diagram

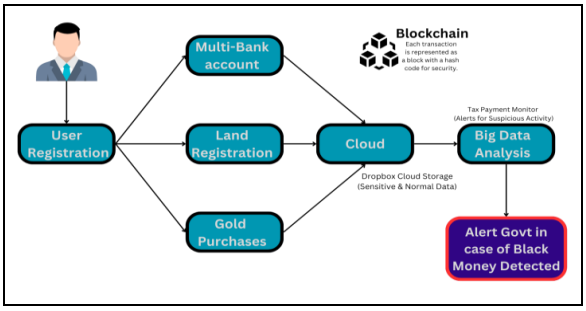

Fig.5.1.Architecture diagram

The overall architecture of the project begins with users registering their basic information along with their Aadhar number. Once registered, users can carry out transactions using their bank accounts to purchase gold or land. These transaction details are stored securely on both the block chain and cloud platforms. In the cloud, the information is organized into three folders: gold, land, and bank transactions. Utilizing Big Data technology, the stored data is analysed using Map Reduce to identify transactions exceeding Rs. 20,000. These findings are then forwarded to the Reserve Bank of India (RBI) for further scrutiny and action.

6. Algorithm

6.1 Asymmetric Key Encryption

Asymmetric key encryption is a way to keep information secret using two different keys. One key is public, meaning everyone can see it. The other key is private, meaning only one person knows it. When someone wants to send secret information, they use the receiver's public key to encrypt it. The receiver then uses their private key to decrypt and read the message. This method ensures that only the intended recipient can read the encrypted message, even though the public key used to encrypt the message is available to everyone. It's like having a special lock that only one key can open, even though everyone has the other key.

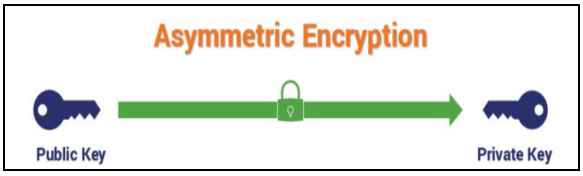

Fig.6.1.Asymmetric key encryption Diagram

6.2 Digital Signature

A digital signature algorithm is a method used to create a unique electronic signature for digital documents or messages. It works by generating a cryptography code based on the content of the document or message, as well as a private key held by the signer. This code, known as the digital signature, verifies the authenticity and integrity of the document or message. When someone receives a digitally signed document, they can use the signer's public key to verify that the signature matches the content and that the document has not been altered since it was signed. Digital signature algorithms ensure the security and trustworthiness of electronic communications and transactions.

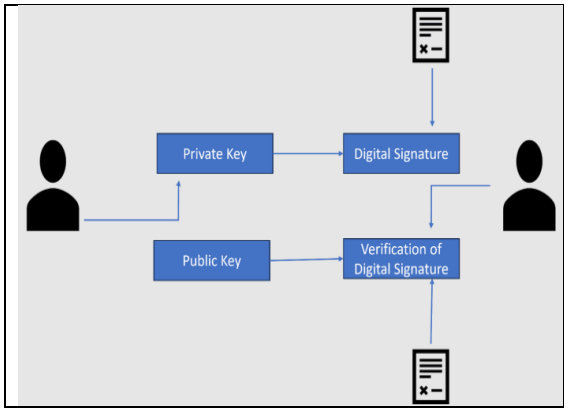

Fig.6.2.Digital Signature Diagram

isrjournals

International Journal of Advanced Research in Computer Science Engineering and Information Technology

Volume: 6, Issue: 3,Special Issue: 2 ,Apr,2024 ,ISSN_NO: 2321-3337

6.3 SHA 256

The SHA-256 algorithm is a type of cryptography hash function that generates a fixed-size output (256 bits) from input data of any size. It is widely used in digital security applications to ensure data integrity and authentication. The algorithm processes input data through a series of mathematical operations, producing a unique hash value that serves as a digital fingerprint for the original data. This hash value is extremely difficult to reverse-engineer, making it suitable for securely verifying the integrity of files and messages, as even a small change in the input data will result in a vastly different output hash.

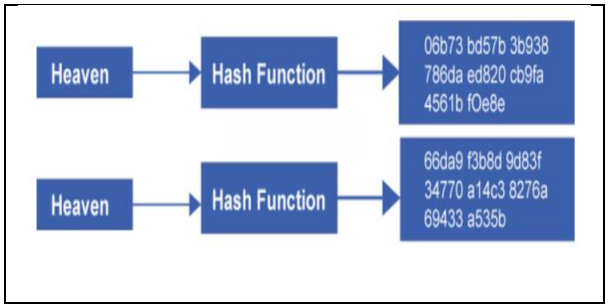

Fig.6.3.SHA 256

6.4 Merkle Hash Tree

A Merkle Hash Tree is a cryptography algorithm used to efficiently verify the integrity of large data sets or files. It works by organizing the data into a tree structure, where each leaf node represents a small piece of the data, and each non-leaf node contains a hash value computed from the hash values of its child nodes. These hash values are then recursively combined until a single hash value, called the root hash or Merkle root, is obtained. This root hash serves as a unique fingerprint for the entire data set, allowing for quick and secure verification of its integrity. If any part of the data is altered, even a small change, it will result in a different root hash, signaling potential tampering. This makes Merkle Hash Trees particularly useful in ensuring the authenticity and reliability of large-scale data storage and transmission systems.

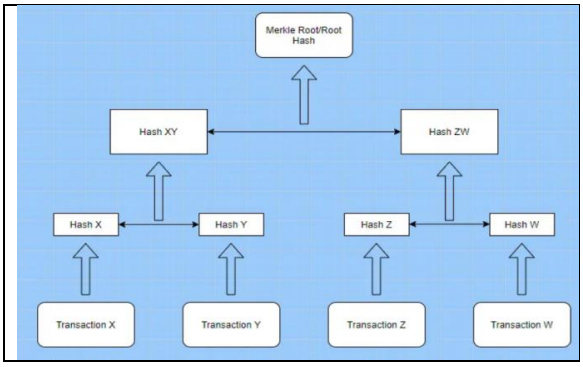

Fig.6.4.Merkle Hash Tree

7. Implementation

7.1 User Registration

In this module we are going to create a User application by which the User is allowed to access the data from the Server. Here first the User wants to create an account and then only they are allowed to access the Network. Once the User create an account, they are to login into their account and request the Job from the Server. Based on the User's request, the service Provider will process the User requested Job and respond to them. All the User details will be stored in the Database.

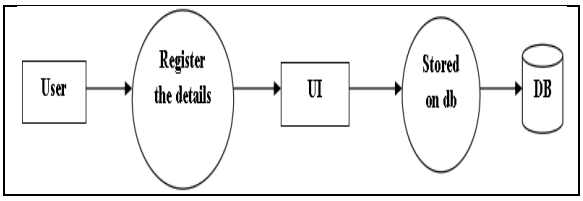

Fig.7.1.User Registration

7.2 Bank Server

Bank Service Provider will contain information about the user in their Data Storage. Also the Bank Service provider will maintain the all the User information to authenticate when they want to login into their account. The User information will be stored in the Database of the Bank Service Provider. To communicate with the Client and with the other modules of the

Company server, the Bank Server will establish connection between them. For this Purpose we are going to create a User Interface Frame.

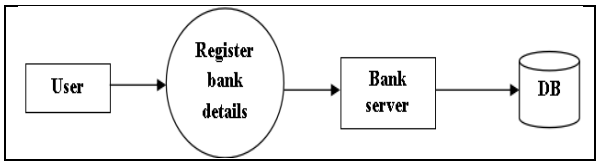

Fig.7.2.Bank Server

7.3 Land Registration and Gold Purchase

In this module we implement land registration and purchased details to be monitor. Here, user name, land documents, price and selling price land. And also we monitor the gold purchase of every user and all other property details will be monitored based on user' Id.

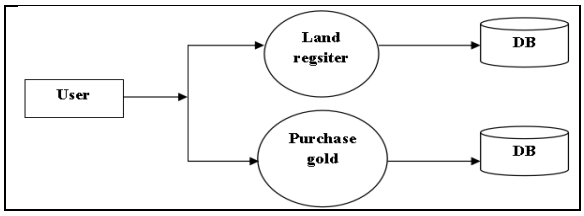

Fig.7.3.Land Registration and Gold Purchase

7.4 Cloud Deployment

User will upload their data to the cloud server and request for a particular file is send to cloud server. To deploy our system we use drop box cloud storage to store our details. Here we store sensitive and normal information on private and public cloud server respectively.

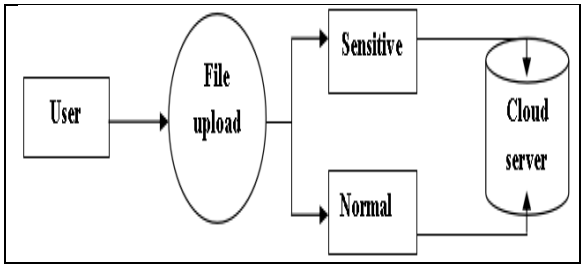

Fig.7.4.Cloud Deployment

7.5 Block chain Deployment

A block is a container data structure. The average size of a block seems to be 1MB (source). Here every certificates number will be created as a block. For every block an hash code will generate for security. Here we store all transaction information like land purchase, gold purchase and all other purchasing details will stored on block chain. For every transaction we a block will create with hash code to refer the other block. Transaction detail will be more secure on block chain.

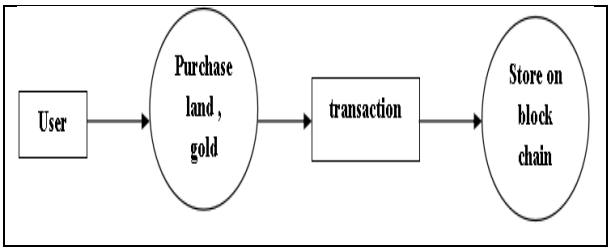

Fig.7.5.Block chain Deployment

7.6 Big data Analysis & Black Money Notification

Throughout all transaction here we monitor proper payment of tax payment. Because, more number of forgeries were made on purchasing of land, people shows a fake price for land purchase and gold purchase. So , in this module we get the details of purchasing rate more than 20K . If user purchasing rate is increased more than 20K, system will alert the income tax notification to the user. Using aadhar number we can monitor all bank transaction also.

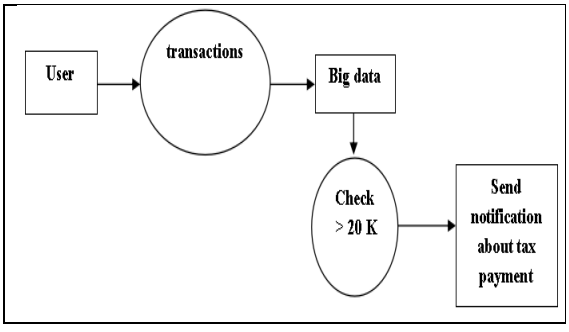

Fig.7.6.Big data Analysis & Black Money Notification

isrjournals

International Journal of Advanced Research in Computer Science Engineering and Information Technology

Volume: 6, Issue: 3,Special Issue: 2 ,Apr,2024 ,ISSN_NO: 2321-3337

8. Experimental Results

Fig 8.1. Shows the Admin login page

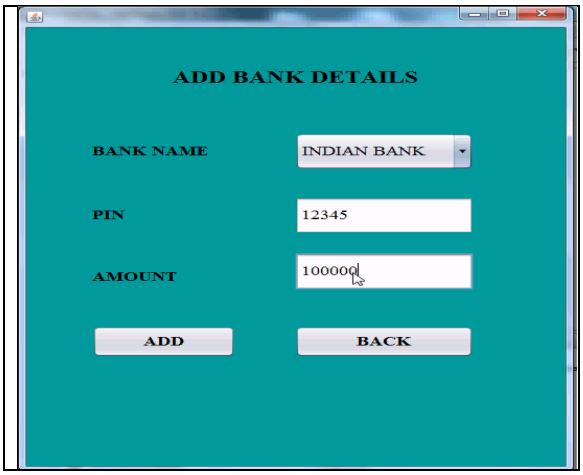

Fig 8.4. Add Bank Details

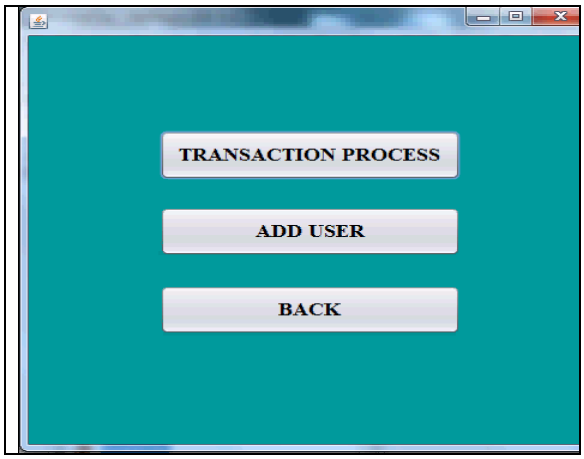

Fig 8.2. Table for Transaction Process

Fig 8.5. Add Land Registration

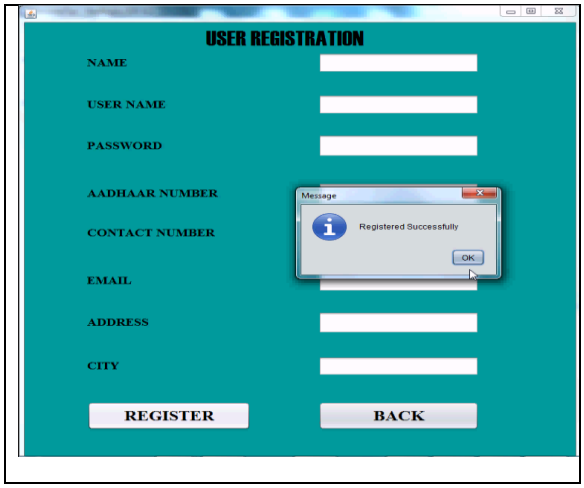

Fig 8.3. Shows Registered Successfully

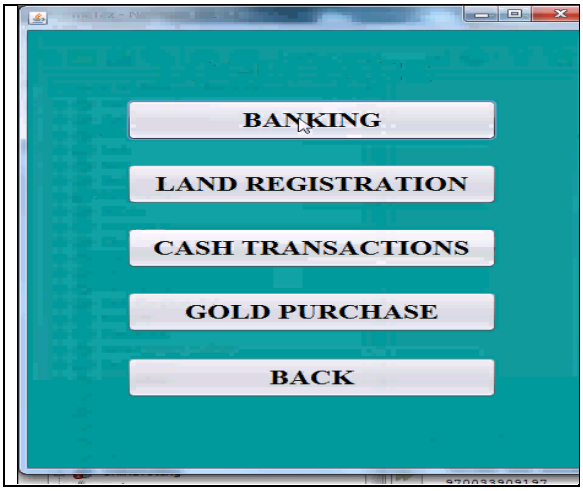

Fig 8.6. Shows the Admin Policy

9. Conclusion & Future Work

Thus the paper infer that we provide an tracking system while purchasing gold or any asset above 20k. Now a day's forgeries level is increasing in smarter way so to provide security we track the money using blockchain technology.

10. Reference

[1] "History of Anti-Money Laundering Laws". United States Department of the Treasury. 30 June 2015. Retrieved 30 June 2015. https://www.fincen.gov/history-anti-money-laundering-laws

[2] "Money Laundering Act 2012 amended". Resource Portal of OGR Legal. OGR Legal. Retrieved in month 2 November 2015. https://resource.ogrlegal.com/money-laundering-act-2012-amended/

[3] "OPSI: Proceeds of Crime Act". Retrieved 14February2009. http://www.opsi.gov.uk/acts/acts2002/ukpga_20020029_en_1

[4] "Suspicious Money Laundering Transaction: Q&A". Anti-Money Laundering Division, Investigation Bureau, Ministry of Justice, Taiwan. Retrieved 20 November 2017. https://goo.gl/jbu6Ln

[5] "Money Laundering". Wikipedia. Retrieved 18January2018. https://en.wikipedia.org/wiki/Money_laundering

[6] "Suspicious Money Laundering or Terrorist Trading" Anti-Money Laundering Division, Investigation Bureau, Ministry of Justice, Taiwan. Retrieved 30 January 2018. https://goo.gl/SfDrvC

[7] "Don't Trust Your Instincts-Benford's law" PanSci. Retrieved 25 December 2017. https://pansci.asia/archives/54264

[8] "Money Laundering Control Act". Wikipedia. Retrieved 25 May 2019. https://en.wikipedia.org/wiki/Money_Laundering_Control_Act

[9] Zhou, Y., Liu, Y., Zhao, D., & Yu, Y. (2019). A Review of Big Data Analytics for Fraud Detection in E-Commerce. IEEE Access, 7,Money Laundering 167459-167476. doi:10.1109/ACCESS.2019.2950072

[10] Zhang, X., Jiang, S., & Wang, B. (2019). A survey on big data-driven fraud detection in finance service industry. Electronic Commerce Research,19(2), 283-305. doi:10.1007/s10660-018-9302-3

[11] Hassan, M. M., Chowdhury, M. A., Nasser, M., & Alelaiwi, A. (2019). Big Data Analytics for Financial Fraud Detection: A Review. IEEE Access, 7, 18496-18516. doi:10.1109/ACCESS.2019.2893693

[12] Wu, D., Wang, Z., & Li, J. (2018). An Overview of Big Data Analytics in Finance. IEEE Access, 6, 18491-18509. doi:10.1109/ACCESS.2018.2803059

[13] Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management, 35(2), 137-144. doi:10.1016/j.ijinfomgt.2014.10.007

[16] Zheng, M., Cai, Y., Xu, H., & Ye, N. (2019). A Survey of Big Data Technologies for Cybersecurity. IEEE Access, 7, 115137 115157. doi:10.1109/ACCESS.2019.2935457

[17] Sari, D. K., & Supriyanto, A. (2020). Big Data for Financial Fraud Detection: A Systematic Literature Review. Journal of Theoretical and Applied Information Technology, 98(1), 150-164.

[18] Yue, X., Wang, F., Jin, D., & Ma, J. (2016). Application and development. Frontiers of Computer Science, 10(4), 691-702. doi:10.1007/s11704-016-5314-z

Volume: 6, Issue: 3,Special Issue: 2 ,Apr,2024 ,ISSN_NO: 2321-3337