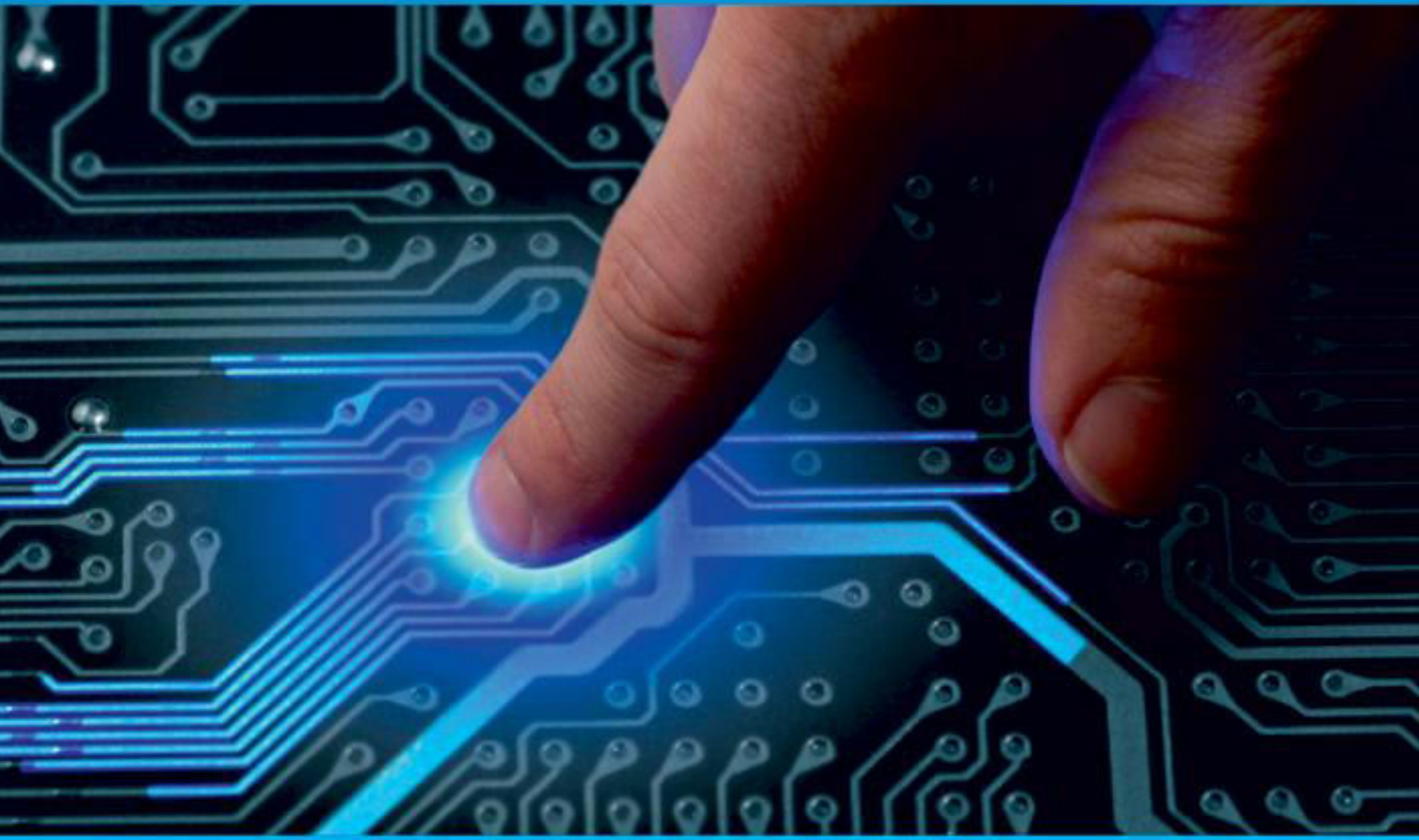
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 6, June 2021

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.542

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906165 |

Advanced Reverification of Credit Card Analysis Using Machine Learning

Mamatha A, Dr. Deepak G

PG Student, Department of CSE, Dayananda Sagar College of Engineering, VTU, Bengaluru, Karnataka, India

Assistant Professor, Department of CSE, Dayananda Sagar College of Engineering, VTU, Bengaluru, Karnataka, India

ABSTRACT: Detecting frauds in credit card transactions is probably one of the most effective test for the process of intelligence algorithms. In fact, this drawback involves variety of relevant challenges, namely: conception implication (customers' habits evolve and fraudster's modification their ways over time), category imbalance (genuine transactions way come frauds), and authentication latency (only a tiny low set of transactions are timely checked by investigators). However, the overwhelming majority of learning algorithms that are planned for fraud detection suppose assume that hardly hold during a real-world fraud-detection system (FDS). This lack of realism considerations 2 main aspects: 1) the means and temporal order with that supervised data is provided .2) the measures accustomed assess fraud-detection performance. Here it has 3 major contributions. First, we propose, with the assistance of our industrial partner, a rationalization of the fraud-detection drawback that realistically describes the operative conditions of FDS that everyday analyse huge streams of credit card transactions. Additionally illustrate the foremost acceptable performance measures to be used for fraud-detection functions. Second, to style and assess a completely unique learning strategy that effectively addresses category imbalance, conception implication, and authentication latency. Third, in this demonstration the impact of sophistication unbalance and conception drift during a real-world information stream containing over seventy five million transactions, licensed over a time window of 3 years.

KEYWORDS: Credit Card; Fraud detection; machine learning; Decision tree classifier; AES encryption.

I. INTRODUCTION

Credit and line data law-breaking is the oldest type of crime identification. It is one of the oldest method to identify the hackers. Hackers sometimes will plan to steal the shopper information at the point of sale systems, so the vendor will acquire shopper data. The new POS system was economical advanced with secures code. Moreover sometimes, user details cannot browse throughout the method. From this way, hackers will hack credit or debit card information as shortly as the devices has read. But due to disconnection of the network for user and trafficker as a result of this there is no any security once the system is offline. Here it describes the semi-offline payment to avoid the card information hacking by the hackers. So here there is more security and flexibility, usually this can be often the first answer that is ready to relinquish secure completely and semi off-line payment way is secured and avoid the stealing of the data. Especially, when the complete system style is detailed, this particular system comprises of components and protocols moreover here will be implementing admin module and will be providing all the information regarding the card details and admin will check all the transaction details once the transaction is verified by PUF only then the vendor will receive the amount.

POS systems act as a gateways and wish to quite network affiliation so they can contact external master card processors. This is often obligatory to validate transactions. To scale back price and change administration and maintenance, POS devices may even be remotely managed over these internal networks.Mobile payment solutions projected so far are usually classified as absolutely on-line, semi off-line, weak off-line or absolutely off-line.The previous work known as force that is, equally to admin, was engineered using a PUF based design. Force provided a difficulty strategy supported information complication and did not address the foremost relevant attacks aimed toward threatening client sensitive information, therefore being at risk of several advanced attack techniques.

All the POS system needs online resources to make the transaction between vendor and user. So there are many possibilities of hacking the card information, all the existing system depends on internet resources.so there is no proper admin monitoring in the previous system so there is so much of chance to steal the user card details, all the payment should be online or semi online there is no semi offline payment system, so the system follows either online or semi online the system has weak interface and the security of the previous system is very less so there may be a huge change to hack the information. Admin is particularly designed to be a secure and reliable encapsulation theme of digital coins,and admin conjointly applicable to multiple-bank events. Indeed, as for credit and debit cards wherever trustworthy third parties (for short, TTPS) like card issuers guarantee the validity of the cards, some common customary convention will be utilized in admin to make banks able to manufacture and sell their own coin part. The coin part area unit usually thought of tamper-proof devices with a secure storage and execution setting for sensitive information.

II. RELATED WORK

In [1] Pay word and micro mint system two simple micropayment schemes where the necessary Pepper coin method are implemented during an identical variation of ways, to maximize easy use for the customer during a tremendously given situation. While the simple pepper coin method requires that all consumer have digital signature capability, one can easily eliminate this requirement by having a celebration trusted by the customer sign payments for him as a proxy, this might rather be a natural approach in an exceedingly web services environment. The pepper coin method is additionally implemented so as that it feels to the patron as a natural extension of his existing credit-card processing procedure, further increasing consumer receiving and modest use.

In [7] Secure POS & kiosk supports the Limited interfaces and placement within local networks, supporting kiosks and point of sale (POS) terminals are visiting be challenging. Often they're located on networks that don't seem to be connected to the net, making direct access impossible for several remote support tools. And even when an employee is present at the terminal, access restrictions and/or lack of technical knowledge makes communicating the answer to an argument difficult. To feature complications, hackers are ramping up their efforts to steal payment card data by gaining access to POS systems and kiosks.

In [3] the author have discussed about Reliable OSPM schema for secure transaction using mobile agent in micropayment system, here it presents an expansive unique offline expense administration in mobile business using the situation learning of micro-payments. this is an extension version of our earlier study addressing on suggestion of secure micropayment system installing process oriented physical design in mobile network. The earlier system has wide application of SPKI and hash chaining to provide trustworthy and protected offline transaction in mobile commerce. However, this work has attempted to meet the expense of rather more light weight secure offline payment system in micro-expenses by designing a bright schema termed as Offline Secure Payment in Mobile Commerce (OSPM).
The empirical operation are distributed on three methods of transaction process considering maximum scenario of real time offline cases. Therefore, the present idea introduces two new constraints i.e. mobile agent and mobile token which is organized to check better security and moderately less network overhead.

In [6] the author have presented a Lightweight and secure PUF key storage using limits of machine learning scheme using silicon Physical Unclonable Functions (PUFs) is defined. To progress a constant PUF bits from chip developed variations, a light-weight inaccuracy correction code (ECC) encoder / decipherer is hired. With an index total of 69, this codec core doesn't use any current error correction methods and is 75% smaller than a previous provably secure implementation, and yet achieves strong eco-friendly presentation in 65nm FPGA and 0.13µ ASIC implementations. The protection of the pattern bits uses a current security disagreement that relies on what cannot be learned from a machine learning viewpoint. The capacity of Leaked Bits is complete for all Disorder Word, reducible using Syndrome Distribution Shaping. the appearance is secure from a min-entropy position against a machine-learning-furnished contender that, given a ceiling of leaked bits, contains a classification error bounded by ε. Mathematical illustrations are given using modern machine learning results.

IN [4] the “Building robust m-commerce payment system on offline wireless network, the where Mobile market is one of the upcoming research area with target mobile expense systems. Unfortunately, the present payment systems is directly entranced with fixed infrastructure of network (cellular network), which fails to facilitate optimal level of security for the payment system. The proposed system highlights a really unique approach for building a secure, scalable, and useful e-payment systems within the circulated development of wireless ADHOC network in offline mode of communication for improved security on transaction and payment process. The proposed system uses Simple Public Key Structure for providing the protection in payment processes. The performance analysis of the proposed model shows that the system is enormously strong and secure ensuring anonymity, privacy, non-repudiation offline payment system over wireless ADHOC network.

III. PROPOSED ALGORITHM

1 .Given training vectors xi∈Rn, i=1,….l and a label vector y∈Rl, a decision tree recursively partitions the feature space such that the samples with the same labels or similar target values are grouped together.

2. Let the data at node m be represented by Qm with Nm samples. For each candidate split θ=(j,tm) consisting of a feature j and threshold tm, partition the data into Qmleft(θ) and Qmright(θ) subsets.
$Qmleft(\theta)=\{(x,y)|xj<=tm\}Qmright(\theta)=Qm\setminus Qmleft(\theta)$.
The quality of a candidate split of node m is then computed using an impurity function or loss function H(), the choice of which depends on the task being solved (classification or regression)
$G(Qm,\theta)=NmleftNmH(Qmleft(\theta))+NmrightNmH(Qmright(\theta))$

3. Select the parameters that minimises the impurity
$\theta*=\arg\min \theta\ G(Qm,\theta)$
Recuse for subsets Qmleft(θ∗) and Qmright(θ∗) until the maximum allowable depth is reached, Nm<minsamples or Nm=1.

Step 1: Classification criteria

If a target is a classification outcome taking on values 0,1,…,K-1, for node m, let
$Pmk =1/Nm \sum y\in Qm\ I(y=k)$
Be the proportion of class k observations in node m. If m is a terminal node, predict probability for this region is set to pmk. Common measures of impurity are the following.
Gini:
$H (Qm) =\sum kpmk(1-pmk)$
Entropy:
$H (Qm) =-\sum kpmk \log (pmk)$
Misclassification:
$H (Qm)=1-\max(pmk)$

Step 2: Regression criteria

If the target is a continuous value, then for node m, common criteria to minimize as for determining locations for future splits are Mean Squared Error (MSE or L2 error), Poisson deviance as well as Mean Absolute Error (MAE or L1 error). MSE and Poisson deviance both set the predicted value of terminal nodes to the learned mean value y¯m of the node whereas the MAE sets the predicted value of terminal nodes to the median median(y) m.

Mean Squared Error:
$y\bar{}m =1Nm\sum y\in QmyH(Qm)=1Nm\sum y\in Qm(y-y\bar{}m)2$
Half Poisson deviance:
$H(Qm)=1Nm\sum y\in Qm(y \log yy\bar{}m-y+y\bar{}m)$
Setting criterion="poison" might be a good choice if target is a count or a frequency (count per some unit). In any case, y>=0 is a necessary condition to use this criterion. Note that it fits much slower than the MSE criterion.

Mean Absolute Error:
$Median (y) m=mediany\in Qm(y)\ H (Qm) =1Nm\sum y\in Qm|y-median(y) m|$
Note that it fits much slower than the MSE criterion.
Minimal Cost-Complexity Pruning
Minimal cost-complexity pruning is an algorithm used to prune a tree to avoid over-fitting, \\ this algorithm is parameterized by α≥0 known as the complexity parameter. The complexity parameter is used to define the cost-complexity measure, Rα(T) of a given tree T:

Rα(T)=R(T)+α|T~|
Where |T~| is the number of terminal nodes in T and R(T) is traditionally defined as the total misclassification rate of the terminal nodes. Alternatively, scikit-learn uses the total sample weighted impurity of the terminal nodes for R(T). As shown above, the impurity of a node depends on the criterion. Minimal cost-complexity pruning finds the subtree of T that minimizes Rα(T).

The cost complexity measure of a single node is Rα(t)=R(t)+α. The branch, it, is defined to be a tree where node t is its root. In general, the impurity of a node is greater than the sum of impurities of its terminal nodes, R(Tt)<R(t). However, the cost complexity measure of a node, t, and its branch, it, can be equal depending on α. We define the effective α of a node to be the value where they are equal, Rα(Tt)=Rα(t) or αeff(t)=R(t)−R(Tt)|T|−1. A non-terminal node with the smallest value of αeff is the weakest link and will be pruned. This process stops when the pruned tree's minimal αeff is greater than the ccp_alpha parameter.

AES Encryption

1) Generation of RSA Key Pair

Each person or a party who desires to participate in communication using encryption needs to generate a pair of keys, namely public key and private key. The process followed in the generation of keys is described below.

- Generate the RSA modulus (n)
  - Select two large primes, p and q.
  - Calculate n=p*q. For strong unbreakable encryption, let n be a large number, typically a minimum of 512 bits.
- Find Derived Number (e)
  - Number **e** must be greater than 1 and less than (p − 1)(q − 1).
  - There must be no common factor for e and (p − 1)(q − 1) except for 1. In other words two numbers e and
  - (p-1)(q – 1) are comprised.
- Form the public key
  - The pair of numbers (n, e) form the RSA public key and is made public.
  - Interestingly, though n is part of the public key, difficulty in factorizing a large prime number ensures that attacker cannot find in finite time the two primes (p & q) used to obtain n. This is strength of RSA.
- Generate the private key
  - Private Key d is calculated from p, q, and e. For given n and e, there is unique number d.
  - Number d is the inverse of e modulo (p - 1)(q – 1). This means that d is the number less than (p – 1) (q - 1) such that when multiplied by e, it is equal to 1 modulo (p - 1)(q - 1).
  - This relationship is written mathematically as follows −ed = 1 mod (p − 1)(q − 1)
  - The Extended Euclidean Algorithm takes p, q, and e as input and gives d as output.

IV. MPLEMENTATION

AES
AES-encrypts input text and returns the encrypted text. Accepts any AES-based algorithm specified by Open SSL.
Unless otherwise indicated, field type is text.
Required fields are indicated by a red asterisk.

Input Fields
Algorithm (dropdown): options for Open SSL encryption algorithms.
Key: text to use as the encryption key; the same key will be needed, along with the algorithm, to decrypt.
Data: text to encrypt.

Output Fields

Output: the encoding of your input data.
The value of the output field can be sent wherever text can be sent. In order to decipher it, you will need to have the correct algorithm and key.

Coin Element

In coin element, the card number is been converted into binary code so if any one hacks the system even then they will not get the card number. The hackers will hack the customer details at the point of sale so in the POS System we are converting card number in the form of binary code. This binary code is generated by using Key Generator and Cryptographic Element. The Key Generator and the binary code will be generated by admin.Once the admin generates the binary code admin PUF will verify the code if the code matches only then transaction will be done otherwise transaction gets failed and seller will not receive the amount from the customer. The binary code should not be the same for all the transaction, as it gets varied for each transaction.

Identity Element

In Identity Element, the identity module develops functionalities. Admin don't need any hardware component apart from the identity and the coin element which can be either plugged into the customer device or directly connected into the device. Similarly to secure elements, both the identity and the coin element are often considered as tamperproof devices with a secure storage and execution environment for sensitive data. Thus, as defined within the ISO7816-4 standard, both of them are often accessed via some APIs while maintaining the specified security and privacy level. Such software components (i.e., APIs) aren't central to the security of our solution and should be easily and constantly updated. This renders infrastructure maintains easier.

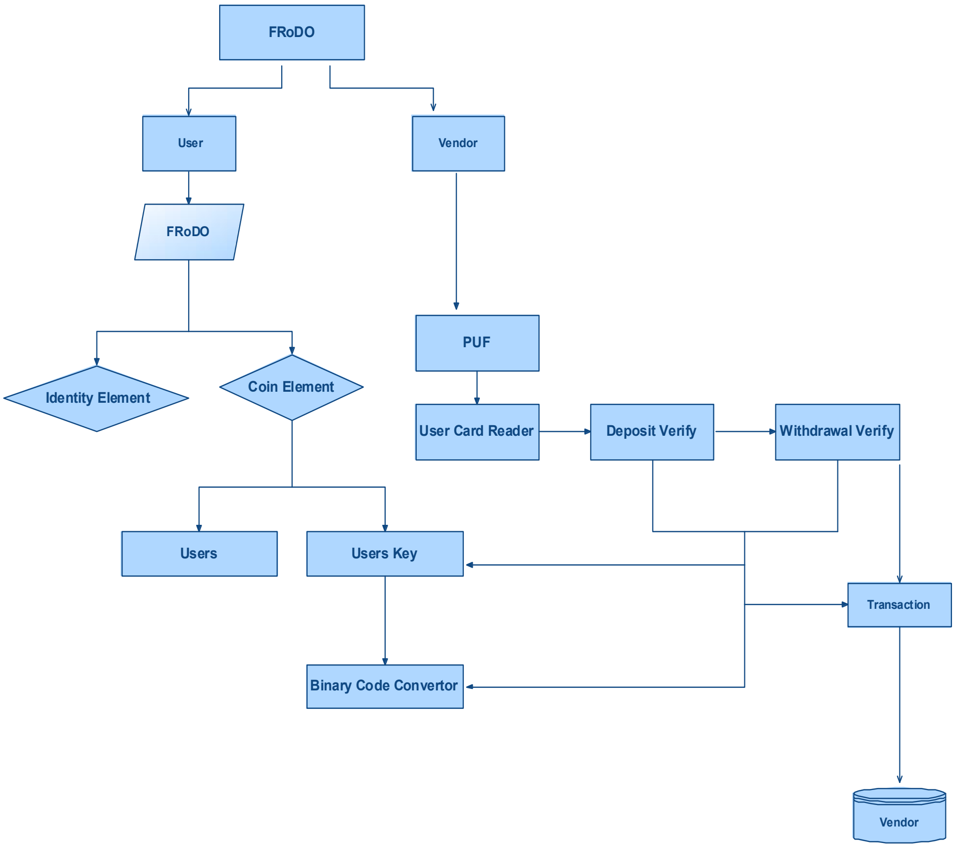


Fig 1 Data Flow Diagram

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906165 |

Data Sets

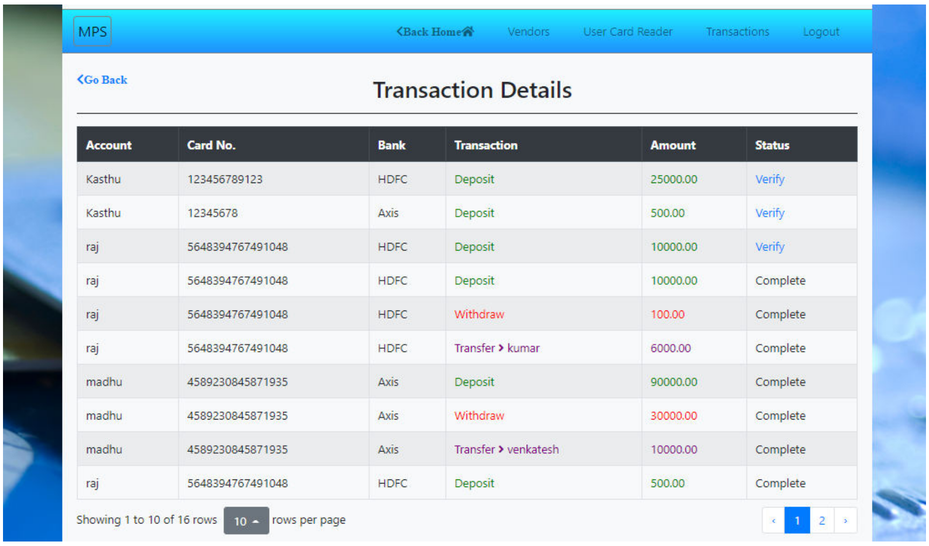
Table 1

V1	V2	V3	V4	V5	V6	V7	V8	Amount	Class
-1.35981	-0.07278	2.536347	1.378155	-0.33832	0.462388	0.239599	0.098698	149.62	0
1.191857	0.266151	0.16648	0.448154	0.060018	-0.08236	-0.0788	0.085102	2.69	0
-1.35835	-1.34016	1.773209	0.37978	-0.5032	1.800499	0.791461	0.247676	378.66	0
-1.81328	4.917851	-5.92613	5.7015	1.204393	-3.03514	-1.7134	0.561257	1	1
-0.25147	4.313523	-6.89144	6.796797	0.616297	-2.96633	-2.43665	0.489328	1	1
0.314597	2.66067	-5.92004	4.5225	-2.31503	-2.27835	-4.68405	1.20227	1	1

KNN is a non-parametric and lazy learning algorithm. Non-parametric means there is no assumption for underlying data distribution. In other words, the model structure determined from the dataset. This will be very helpful in practice where most of the real world datasets do not follow mathematical theoretical assumptions. Lazy algorithm means it does not need any training data points for model generation. All training data used in the testing phase. This makes training faster and testing phase slower and costlier. Costly testing phase means time and memory. In the worst case, KNN needs more time to scan all data points and scanning all data points will require more memory for storing training data.In KNN, K is the number of nearest neighbors. The number of neighbors is the core deciding factor. K is generally an odd number if the number of classes is 2. When K=1, then the algorithm is known as the nearest neighbor algorithm. This is the simplest case. Suppose P1 is the point, for which label needs to predict. First, you find the one closest point to P1 and then the label of the nearest point assigned to P1.

V.RESULTS

KNN performs better with a lower number of features than a large number of features. By this when the number of features increases than it requires more data. Increase in dimension also leads to the problem of over fitting. To avoid over fitting, the needed data will need to grow exponentially as you increase the number of dimensions. This problem of higher dimension is known as the Curse of Dimensionality.To deal with the problem of the curse of dimensionality, we perform principal component analysis before applying any machine learning algorithm, or use feature selection approach. Research has shown that in large dimension Euclidean distance is not useful anymore. Therefore, prefer other measures such as cosine similarity, which get decidedly less affected by high dimension.

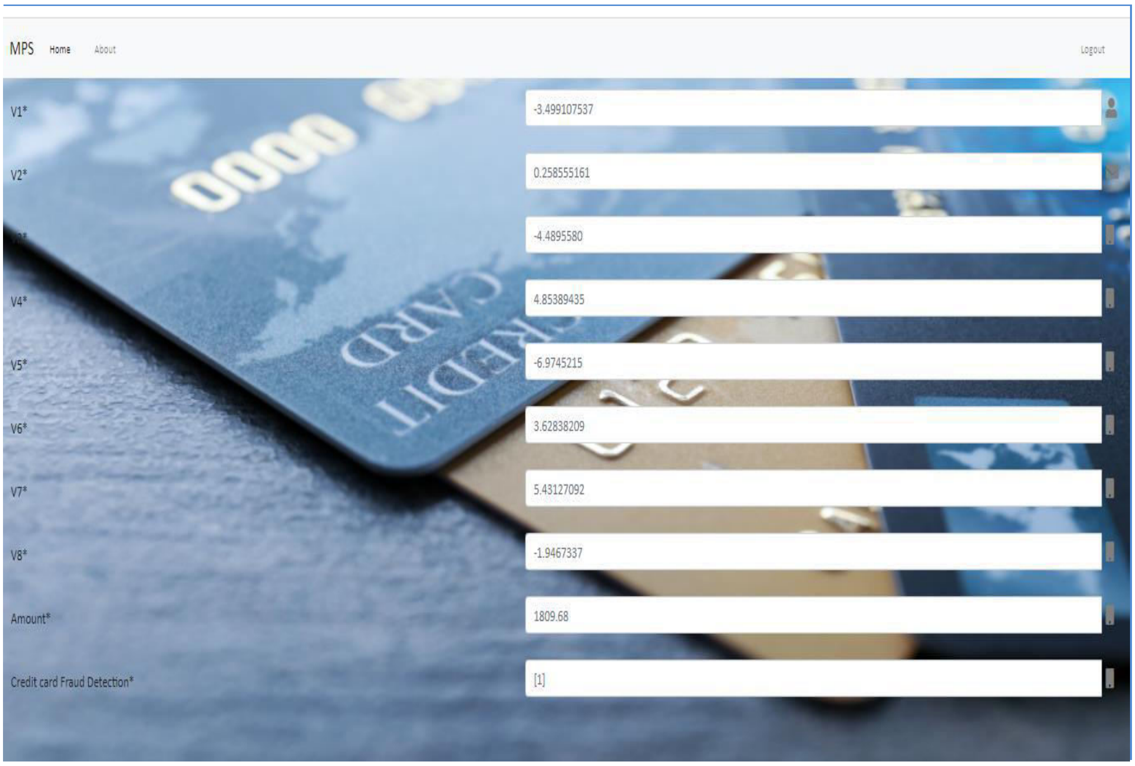


Transaction Details

Account	Card No.	Bank	Transaction	Amount	Status
Kasthu	123456789123	HDFC	Deposit	25000.00	Verify
Kasthu	12345678	Axis	Deposit	500.00	Verify
raj	5648394767491048	HDFC	Deposit	10000.00	Verify
raj	5648394767491048	HDFC	Deposit	10000.00	Complete
raj	5648394767491048	HDFC	Withdraw	100.00	Complete
raj	5648394767491048	HDFC	Transfer › kumar	6000.00	Complete
madhu	4589230845871935	Axis	Deposit	90000.00	Complete
madhu	4589230845871935	Axis	Withdraw	30000.00	Complete
madhu	4589230845871935	Axis	Transfer › venkatesh	10000.00	Complete
raj	5648394767491048	HDFC	Deposit	500.00	Complete

Showing 1 to 10 of 16 rows 10 rows per page

Figure 2: Transaction Details



V1* -3.499107537

V2* 0.258555161

V3* -4.4895580

V4* 4.85389435

V5* -6.9745215

V6* 3.62838209

V7* 5.43127092

V8* -1.9467337

Amount* 1809.68

Credit card Fraud Detection* [1]

Figure 3: Credit Card Fraud Detection

V. Conclusion and Future Work

By introduced Offline payment that is to the simplest of our data, the primary data-breach-resilient absolutely offline micro-payment approach. The protection analysis shows that doesn't impose trait assumptions. Further, it's additionally the primary answer within the literature wherever no client device information attacks is exploited to compromise the system. This has been achieved in the main by investing a totally distinctive effaceable PUF design and a totally distinctive protocol style. Moreover, this technique has been totally mentioned and compared against the state of the art. Our analysis shows that it is the unique technique that enjoys all the properties needed to a secure payment answer, whereas additionally introducing flexibility once considering the payment medium (types of digital coins). Finally, some open problems area unit known that area unit left as future work. Specifically, we tend to area unit investigation the chance to allow digital modification to be spent over multiple off-line transactions whereas maintaining a similar level of security and utility.

References

[1] R. L. Rivest, “Payword and micromint: Two simple micropayment schemes,” in Proc. Int. Workshop Security Protocols, 1996, pp. 69–87.
[2] J. Lewandowska. (2013). http://www.frost.com/prod/servlet/press-release.pag?docid=274238535.
[3] N. Kiran and G. Kumar, “Reliable OSPM schema for secure transaction using mobile agent in micropayment system,” in Proc. 4th Int. Conf. Computer., Communication. Network. Technology July. 2013, pp. 1–6.
[4] N. Kiran and G. Kumar, “Building robust m-commerce payment system on offline wireless network,” in Proc. IEEE 5th Int. Conf. Adv. Network Telecommunication. Syst., Dec. 2011, pp. 1–3.
[5]Mandiant, “Beyond the breach,” Mandiant, 2014, https://dl.mandiant.com/EE/library/WP_M-Trends2014_140409.pdf.
[6] M.-D. Yu, D. MRaihi, R. Sowell, and S. Devadas, “Lightweight and secure PUF key storage using limits of machine learning,” in Proc. Int. Workshop Cryptographic Hardware Embedded Syst., 2011, vol. 6917, pp. 358–373.
[7]Bogmar,”Secure POS & kiosk support,” Bogmar, 2014, http://www.bomgar.com/assets/documents/Bomgar_Remote_Support_for_POS_Systems.pdf

| DOI: 10.15680/IJIRCCE.2021.0906165 |

[8] T. Nishide and K. Sakurai, “Security of offline anonymous electronic cash systems against insider attacks by untrusted authorities revisited,” in Proc. 3rd Int. Conf. Intell. Network. Collaborative Syst., 2011, pp. 656–661.
[9 W.-S. Juang, “An efficient and practical fair buyer-anonymity exchange scheme using bilinear pairings,” in Proc. 8th Asia Joint Conf. Inf. Security, Jul. 2013, pp. 19–26.
[10] M. A. Salama, N. El-Bendary, and A. E. Hassanien, “Towards secure mobile agent based e-cash system,” in Proc. Int. Workshop Security Privacy Preserving e-Soc., 2011, pp. 1–6.
[11 R. Battistoni, A. D. Biagio, R. Di Pietro, M. Formica, and L. V. Mancini, “A live digital forensic system for Windows networks,” in Proc. 20th IFIP TC Int. Inf. Security Conf., 2008, vol. 278, pp. 653–667.
[12] G. Hong and J. Bo, “Forensic analysis of skimming devices for credit fraud detection,” in Proc. 2nd IEEE Int. Conf. Inf. Financial Eng., Sep. 2010, pp. 542–546.
[13] C. R. Group, “Alina & other POS malware,” Cymru, 2013,
https://www.teamcymru.com/ReadingRoom/Whitepapers.
[14] W. Whitteker, “Point of sale (POS) systems and security,” SANS Inst., Fredericksburg, VA, USA, 2014, http://www.sans.org/reading-room/whitepapers/bestprac/point-sale-pos-systemssecurity-35357.
[15] U. Reuhrmair, F. Sehnke, J. Seolter, G. Dror, S. Devadas, and J. Schmidhuber, “Modeling attacks on physical unclonable functions,” in Proc. 17th ACM Conf. Computer. Community. Security, 2010, pp. 237–249.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

☎ 9940 572 462 ✆ 6381 907 438 ✉ ijircce@gmail.com

www.ijircce.com

Scan to save the contact details