Journal of Artificial Intelligence and Data Mining (JAIDM), Vol. 11, No. 4, 2023, 599-608.

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research paper

A Multi-layered Hidden Markov Model for Real-Time Fraud Detection in Electronic Financial Transactions

Abukari Abdul Aziz Danaa[1*], Mohammed Ibrahim Daabo[2] and Alhassan Abdul-Barik[3]

1. *Department of Computer Science, Tamale Technical University, Tamale, Ghana.*
2. *Department of Computer Science, C. K. Tedam University of Technology and Applied Sciences, Navrongo, Ghana.*
3. *Department of Computer Science, University for Development Studies, Tamale, Ghana.*

Article Info

Article History:
Received 19 October 2022
Revised 30 January 2023
Accepted 27 March 2023

DOI:10.22044/jadm.2023.11990.2357

Keywords:
Fraudulent, Hidden Markov Models, Optimization; Probability; Multi-layered.

**Corresponding author: azizdanaa@tatu.edu.gh (A. A. A. Danaa).*

Abstract

Hidden Markov Models (HMMs) are machine learning algorithms that have been applied to a range of real-life applications including intrusion detection, pattern recognition, thermodynamics, and statistical mechanics, among others. A multi-layered HMM for real-time fraud detection and prevention whilst reducing drastically the number of false positives and negatives is proposed and implemented in this work. The study adopts the Payment Simulator (PAYSIM) and Mobile Money Transaction (MMT), datasets and focuses on reducing the parameter optimization and detection times of the proposed models using a hybrid algorithm comprising the Baum-Welch, Genetic and, Particle-Swarm Optimization algorithms. The simulation results reveal that, for various number hidden states, our proposed model performs averagely better in terms of precision (0.984 and 0.986 for the PAYSIM and MMT datasets, respectively), recall (0.965 and 0.971 for the PAYSIM and MMT datasets, respectively), and F1-scores (0.974 and 0.978 for the PAYSIM and MMT datasets, respectively) when compared to the existing approach.

1. Introduction

Hidden Markov Models (HMMs) are very useful in modeling distributions that are probabilistic in nature with a finite number of observable outputs emanating from some number of hidden states based on an initial probability vector, transition, and emission probability matrices [1]. Due to their strong mathematical structure and sound theoretical basis, they have been applied in numerous real-life applications ranging from anomaly detection, and pattern recognition to motion analysis in videos [2]. A hidden Markov model, λ (μ, E, F) is generally represented by:

1. An initial probability vector denoted by μ = [μ_j], where:

$$\mu_j = P\left(q_1 = S_j\right), 1 \le j \le U \quad (1)$$

Where $\sum_{j=1}^{u} \mu_j = 1$ (2)

2. A state transition probability matrix denoted by E = [e_{ji}], where:

$$e_{ji} = P\left(q_{t+1} = S_i \mid q_t = S_j\right) \quad (3)$$

where $1 \le j, i \le U, t = 1,2,T-1$

$$\sum_{i=1}^{U} e_{ji} = 1, 1 \le i \le U \quad (4)$$

3. An emission probability matrix, F = [$f_i(y)$], where:

$$f_i(y) = P\left(V_y = q_t \mid S_i = q_t\right) \quad (5)$$

Where $1 \le j \le U, 1 \le y \le V$

Financial fraud refers to any intentional act of deception involving financial transactions for personal benefit [3]. The application of machine learning techniques to fraud detection by research communities has been on the rise for decades now due to their ability to detect and prevent unknown and relatively complex types of fraud [4].

The predominant approach for fraud detection and prevention using HMM is creating and maintaining a single optimized HMM for each user using their normal profile transactions [5]. This approach results in reduced system performance due to their increased parameter optimization and fraud detection times making them unsuitable for real-time fraud detection [6].

However, the process of constructing a multi-layered HMM for fraud detection involves the use of a central storage of normal transactions and then constructing a hidden Markov model corresponding to each user [4]. An incoming transaction is initially compared with the normal sequence of user transactions to detect possible mismatches before being submitted to the HMM to compute its probability of occurrence [6]. Although this approach generally enhances the training and detection times, the significant number of false negatives and positives remains a challenge [7].

In order to optimize the parameters of the proposed models within a more reasonable computing time, this research work divided the training data into sub-groups where each represents some particular user behavior large enough to efficiently optimize the parameters of a single HMM. The individually optimized models are then merged incrementally to create a final model with a main focus to drastically reduce the rate of false positives and negatives. As mentioned earlier, a modified hybrid algorithm comprising the Baum-Welch, Genetic, and Particle-Swarm Optimization algorithm was used to optimize the parameters of the proposed models.

Section 2 of this paper presents relevant literature relating to the application of some machine learning models to anomaly/fraud detection whilst the methodology adopted to develop and implement our proposed models is also outlined in Section 3. Section 4 discusses in detail the results obtained from simulations in order to establish the efficiency of our proposed models, whilst Section 5 contains the conclusions and recommendations based on the findings of the research work.

2. Related Works

An intrusion detection system based on multi-layered HMM was proposed by Hoang *et al.* [8] where system calls were utilized in constructing normal program behaviors. Significant deviations from those behaviors compared to a pre-defined threshold signified a potential intrusion. From the simulation results, their proposed approach performed better in terms of detection time with reduced false positives when compared with a single-layered hidden Markov model, although more time was required to optimize the parameters of the proposed models.

Penagarikano and Bordel [9] developed and applied multiple-layered hidden Markov models to pattern recognition with each level of knowledge represented by a layered structure comprising a number of hidden Markov models with a main focus on obtaining a significantly lower number of false positives and negatives. The positive outcome of their research further highlighted the possibility of developing enhanced architectural methods to recognition problems using a multi-layer Hidden Markov Model approach.

Abouabdalla *et al.* [10] however, incorporated correlation methods with multi-layered HMMs to drastically reduce the rate of false positives and negatives associated with fraud/intrusion detection systems. Their proposed models, however, experienced performance degradation in the presence of highly skewed datasets. Spathoulas and Katsikas [11] incorporated a filtering mechanism to handle outliers in datasets used before feeding them into a hidden Markov model for possible fraud detection and prevention. Experimental results established the capability of their system to reduce the rate of false positives to about 75% when compared with a single-layer approach but performed poorly when transaction patterns change very frequently.

An optimized multi-layered semi-HMM was proposed and implemented by Prakash and Chandrasekar [12]. Their proposed approach incorporated the cuckoo search algorithm to determine the appropriate number of hidden states and emission symbols of the proposed models. Although more training is required for training the models, the simulation results revealed the effectiveness of their proposed approach when compared with the existing techniques.

Burgio [13] utilized extreme learning machine and hidden Markov model with the main aim of achieving a significantly low number of false positives with high precision, recall, and accuracy, whilst Zegeye *et al.* [7] divided the problem into subsets to achieve better accuracy when compared with single-layer hidden Markov model.

Alarfaj *et al.* [14] developed deep learning algorithms to detect financial crimes, particularly those involving the availability of public data that

are highly class imbalanced with continuous changes in the nature of the fraudulent activities resulting in high rates of false alarm.

An analysis of their proposed approach revealed an accuracy, f1-score, and precision of 99.9%, 85.71%, and 93%, respectively. For situations involving credit card detection, the suggested model performs better than the cutting-edge machine learning and deep learning techniques.

Baghdasaryan *et al*. [15] created a machine learning-based fraud prediction model by using the universe of Armenian corporate taxpayers working under a regular tax framework, whilst primarily relying on gradient boosting. Their proposed approach was able to optimally detect fraud after successfully extracting key characteristics from tax returns with the least amount of supplementary data.

An analysis of the performances of various machine learning algorithms including decision tree, random forest, linear regression, and gradient boosting for fraud detection and prevention was performed in [16]. The simulation results revealed that an accuracy of 80% was obtained using the random forest classifier and 70% by the logistic regression technique. However, the gradient boosting algorithm produced better results by obtaining an accuracy of above 90%.

Do *et al.* [17] considered deep learning algorithms to detect phishing, and categorized them based on current literature by analyzing their benefits and drawbacks. They then suggested various challenges encountered by deep learning algorithms in fraud/phishing detection, as well as future research directions to address these problems.

Improper parameter tuning coupled with poor training times and lower detection accuracy are prevalent problems with modern deep learning systems based on the empirical experiment's findings of their study.

Fujita *et al.* [18] compared the efficiency of deep learning models with a hybrid deep learning model integrated with a hybrid parameterization model in handling complex and missing datasets as well as their performance in increasing classification. The results showed that deep learning models performed relatively better on their own with faster processing times and analyses of relatively complex datasets with significant missing values.

3. Proposed Method

3.1 Data Pre-processing

For simulation purposes, the Payment Simulator (PAYSIM) and Mobile Money Transaction (MMT) datasets as presented in Table 1 are adopted. With some fraudulent transactions introduced into it, the PAYSIM Dataset is simulated from real financial datasets which represent normal transactions performed by customers [19] whereas the MMT dataset is a mobile money transfer transactional data that includes time intervals within which those transactions took place and generated by using a multi-agent-based simulator [20].

Table 1. Details of datasets adopted for the study.

	Paysim dataset [2]			MMT dataset [1]		
Trans. type	**Genuine**	**Fraudulent**	**Total**	**Genuine**	**Fraudulent**	**Total**
Transfer	528812	4097	532909	84331	1590	85921
Cash-out	2233384	4116	2237500	115133	940	116073
Cash-in	1399284	0	1399284	52445	397	52842
Debit	41432	0	41432	81365	1626	82991
Payments	2151494	0	2151494	158391	1347	159738
Total	6354407	8213	6362620	491665	5900	497565

As presented in Algorithm 1, the density-based spatial clustering of applications with noise (DBSCAN) was adopted and modified to appropriately categorize customers into groups depending on their transaction patterns. The modification of the DBSCAN algorithm allows for the dynamic conversion of an incoming transaction into an observation symbol in order to classify it as fraud or otherwise based on the computed centroid of each cluster.

Algorithm 1: Modified DBSCAN algorithm.

STEP 1: Declare and initialize the index of cluster c to 1.

STEP 2: Visit any point, *v*, in the problem space that has not been visited already and Mark *v* as visited.

STEP 3: Determine the various points, M which are in the neighborhood of v.

SPTEP 4: If |M|>=z then M = M U M'

STEP 5: if 'M' does not belong to any available cluster, then it is automatically marked as noise.

STEP 6: For each cluster, determine its centroid, *cl*, as in (6);

$$\boldsymbol{CI} = \frac{\mathbf{1}}{\boldsymbol{m_i}} \sum_{x_j \in c_j} \boldsymbol{x_j} \qquad (6)$$

where m_i is the number of points in cluster c_j

The input to the modified DBSCAN algorithm is a set of points, x, y, and z representing the minimal number of points a cluster should contain the set of all possible points and, the threshold for the distance between any two points, respectively.
At the end of the clustering step, various groups depicting the various transaction patterns of users are determined and users are classified appropriately. The cluster that determines the transaction profile, *r* of an account holder, *a* is determined as in (7), where w_i represents the percentage of the total number of transactions of the account holder.

$$r(a) = \arg Max(w_i) \tag{7}$$

The spending profile of an accountholder refers to the cluster with the most transactions. The computed centroids are used to convert an incoming new transaction n_l into an observation symbol that is defined as in (8).

$$n_l = \nu_{arg_i} \min |1 - n_i| \tag{8}$$

3.2 Structure of proposed multi-layer HMM

The hidden states of the proposed hidden Markov model are the various transaction types, whereas the observation symbols at the various levels are represented by transaction frequencies (TF), amounts (TA), and time intervals (TI).
Let $\lambda_1, \lambda_2, ..., \lambda_p$ denote the various HMMs at each layer with corresponding hidden states ($X_1, X_2, ..., X_G$) and emission symbols ($O_1, O_2, ..., O_H$), the observation and state sequences are represented as in (9) and (10), respectively.

$$O_1^T = \left\{O_1^1, O_1^2, \ldots, O_1^T\right\}, .., O_H^T = \left\{O_H^1, O_H^2, \ldots, O_H^T\right\} \tag{9}$$

$$X_1^T = \left\{X_1^1, X_1^2, \ldots, X_1^T\right\}, \ldots, \left\{X_G^2, X_G^2, \ldots, X_G^T\right\} \tag{10}$$

The probability, δ of generating a transaction sequence by our proposed HMM is computed as in (11) and the value is compared to a threshold value to classify the transaction as suspicious or otherwise.

$$\delta = p\left(\sigma_1, \sigma_2, \sigma_3, \ldots \sigma_r | \lambda\right) \tag{11}$$

3.3 Optimizing parameters of proposed multi-layer HMM

As mentioned earlier, a hybrid optimization algorithm proposed by Danaa *et al.* [21] which comprises the BW, GA, and PSO algorithm is adopted and modified to optimize the parameters of the proposed multi-layered HMM.

The various spending profiles of accountholders as determined using (7) are used to initialize the initial probability vector, state transition, and emission symbol probability matrices. The forward and backward variables are computed as in (12) and (13), respectively.

$$\alpha_t(i) = \sum_{j=0}^{U-1} [\alpha_{t-1}(j) a_{ji}] f_i(O_t) \tag{12}$$

$$b_t(i) = \sum_{j=0}^{N-1} [\beta_{t+1}(j) a_{ij}] b_j(O_{t+1}) \tag{13}$$

The gamma and di-gamma variables are computed as in (14) and (15), respectively.

$$g_t(i) = \frac{\alpha_t(i) b_t(i)}{P(O|\lambda)} \tag{14}$$

$$g_t(i,j) = \frac{\alpha_t(i) e_{ij} f_j(O_{t+1}) b_{t+1}(j)}{P(O|\lambda)} \tag{15}$$

The values of the initial and state transition and symbol emission probability matrices are computed in (16), (17), and (18), respectively.

$$\pi_i = \gamma_0(i) \tag{16}$$

$$e_{ij} = \sum_{t=0}^{T-2} g_t(i,j) / \sum_{t=0}^{T-2} g_t(i) \tag{17}$$

$$f_j(y) = \sum_{\substack{t \hat{o} \{0,1,\ldots T-1\} \\ O_t = y}} g_t(j) / \sum_{t=0}^{T-1} g_t(i) \tag{18}$$

Feasible solutions from 100 iterations using (16), (17), and (18) are considered chromosomes for the proposed GA, where the probability of computing the probability of an observation by the HMM denoted by P(O|λ) is considered the fitness function. Multiple point crossover and mutation are then performed to select the best 50 solutions for the next generation which are transformed as particles possibly searching for a possible solution using the Particle PSO algorithm. Each particle's velocity and position are computed and updated iteratively as in (19) and (20), respectively.

$$X_i(t+1) = X_i(t) + V_i(t) \tag{19}$$

$$V_i(t+1) = wV_i(t) + r_i \upsilon([0,1])(X_i^+(t) - X_i(t)) + r_2 \upsilon([0,1])(\hat{X}_i(t) - X_i(t)) \tag{20}$$

The best solution of each particle is compared with the best position of the entire group and appropriate adjustments are made. The resulting row, R_3 for any given pair of rows, R_1 and R_2 from two HMMs, λ_1 and, λ_2 for the initial, state transition and, emission probability matrices are computed as in (21). However, the sum of each row in these three (3) matrices satisfies the row stochastic property (values must sum up to 1).

$$R_3 = \alpha R_1 + (1-\alpha)R_2 \\ = \alpha S_1 + (1-\alpha)T_1, \alpha S_2 + (1-\alpha)T_2, ..., \alpha S_N + (1-\alpha)T_N \quad (21)$$

3.4 Model evaluation metrics

As presented from (22) to (25), the study evaluated the performance of the proposed models using F1-scores(F1), precision(P) and, recall(R) metrics due to the highly skewed nature of the datasets used. The false positive (F_p)and false negative (F_n) rates as well as the receiver operating characteristics graph were also employed to establish how well our proposed model is capable of distinguishing between fraudulent and genuine transactions [22].

$$p = T_p / \left(T_p + F_p\right) \quad (22)$$

$$R / TPR = T_p / \left(T_p + F_n\right) \quad (23)$$

$$F1 = \left(2 * P * R\right) / \left(P + R\right) \quad (24)$$

$$\mathrm{FPR} = \mathrm{F_p} / (\mathrm{F_p} + \mathrm{T_n}) \quad (25)$$

4. Results and Discussion

20% of each of the datasets was retained for validation purposes, and the rest were used to optimize the parameters of the models. Simulations were done in a Python programming environment for the various numbers of hidden states and their performance compared to the multi-layered HMM approach proposed by Zegeye et al. [7] denoted as SLHMM.

4.1 Confusion matrix

With about 127 fraudulent transactions out of a total of 1,390, the confusion matrix as presented in Table 2 reveals how well our proposed model is able to correctly classify positive and negative classes (PC and NC, respectively). It is evident that, for both datasets, our proposed model obtained relatively fewer false negatives and positives as compared to the multi-layered HMM approach proposed by Zegeye et al. [7].

4.2 Recall, precision, F1-score, and ROC curve plot

Our proposed model performed better in terms of precision, recall, and F1-scores when compared with the multi-layered HMM approach proposed by Zegeye et al. [7] for both datasets using different numbers of hidden states as presented in Table 3.

As depicted in the AUC-ROC curves for both datasets presented in Figures 2 and 3, our proposed model also obtained higher AUC values (higher number of TPs and lower number of FPs) for the various numbers of hidden states.

Although our proposed model performed better on the PAYSIM dataset, the higher AUC values obtained reveal how better it distinguishes between fraudulent and genuine transactions as compared to that proposed in [7].

4.3 Computational efficiency of models

For various numbers of states, Table 4 and Figure 4 contain the time taken to optimize the proposed models as well as to detect and classify an incoming transaction as fraudulent or otherwise. Using a Microsoft Windows-based computer with an Intel Core i5 with a CPU speed of 2.30 GHz and a RAM size of 4 GB, the relatively lesser time consumed by our proposed model for both datasets was performed in python programming environment.

The results revealed that on average for the different number of hidden states, the detection times obtained by our proposed model are 0.005 s and 0.01 s and a training time of 0.032 s, and 0.022 s for the PAYSIM and MMT datasets respectively.

These values are better than those obtained by the approach proposed by Zegeye *et al.* [7], which obtained detection times of 0.019 s and 0.02 s and training times of 0.08 s and 0.079 s for the PAYSIM and MMT datasets, respectively.

The faster detection and training times of our proposed model are attributed to the fact that it is less likely to suffer from overfitting since training is performed on smaller amounts of data independently. Each data represents some particular user behavior large enough to efficiently optimize the parameters of a single HMM making it ideal for real-time fraud detection.

Table 2. Performance of models based on a confusion matrix for different numbers of states.

			Paysim		MMT dataset	
Model	Actual number	Hidden sates	P Pred	N Pred	P Pred	N Pred
SLHMM	PC	2	115	12	113	14
	NC		10	1253	11	1252
	PC	3	110	17	118	9
	NC		6	1257	9	1254
	PC	4	119	8	120	7
	NC		5	1258	10	1253
	PC	5	118	9	115	12
	NC		4	1259	7	1256
Proposed	PC	2	125	2	126	1
	NC		2	1261	1	1262
	PC	3	124	3	124	3
	NC		3	1260	2	1261
	PC	4	121	6	122	5
	NC		2	1261	1	1262
	PC	5	120	7	121	6
	NC		1	1262	3	1260

Table 3. Performance of models based on precision, recall, and F1-scores for different numbers of states.

Metric	# Hidden states	PAYSIM dataset		MMT dataset	
		Proposed	SLHMM	Proposed	SLHMM
Precision	2	0.984	0.920	0.992	0.911
	3	0.976	0.948	0.984	0.929
	4	0.984	0.960	0.992	0.923
	5	0.992	0.967	0.976	0.943
Recall	2	0.984	0.906	0.992	0.890
	3	0.976	0.866	0.976	0.929
	4	0.953	0.937	0.961	0.945
	5	0.945	0.929	0.953	0.906
F1 Score	2	0.984	0.913	0.992	0.900
	3	0.976	0.905	0.980	0.929
	4	0.968	0.948	0.976	0.934
	5	0.968	0.948	0.964	0.924

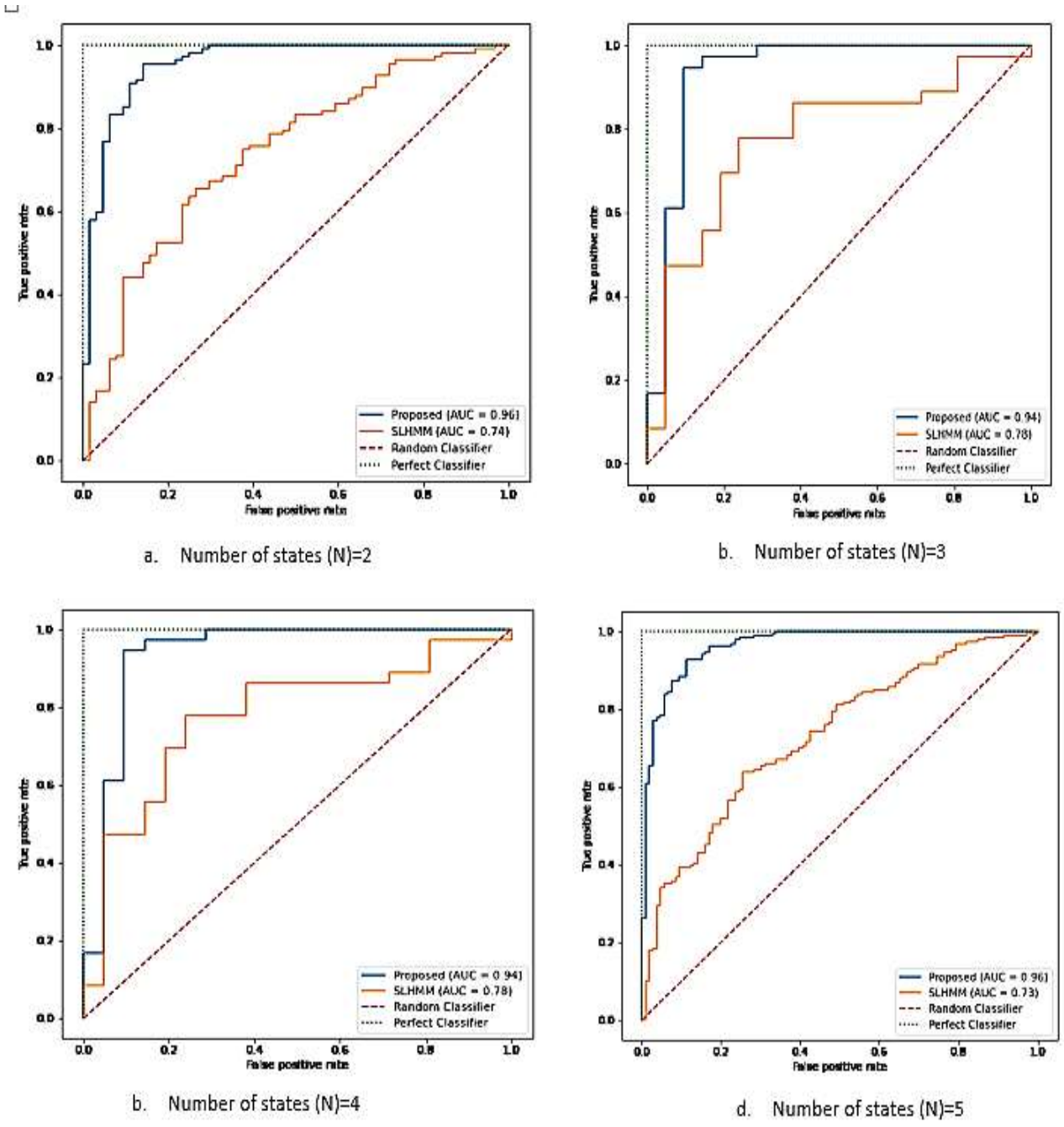

Figure 2. ROC curve plot for various numbers of hidden states using the Paysim dataset.

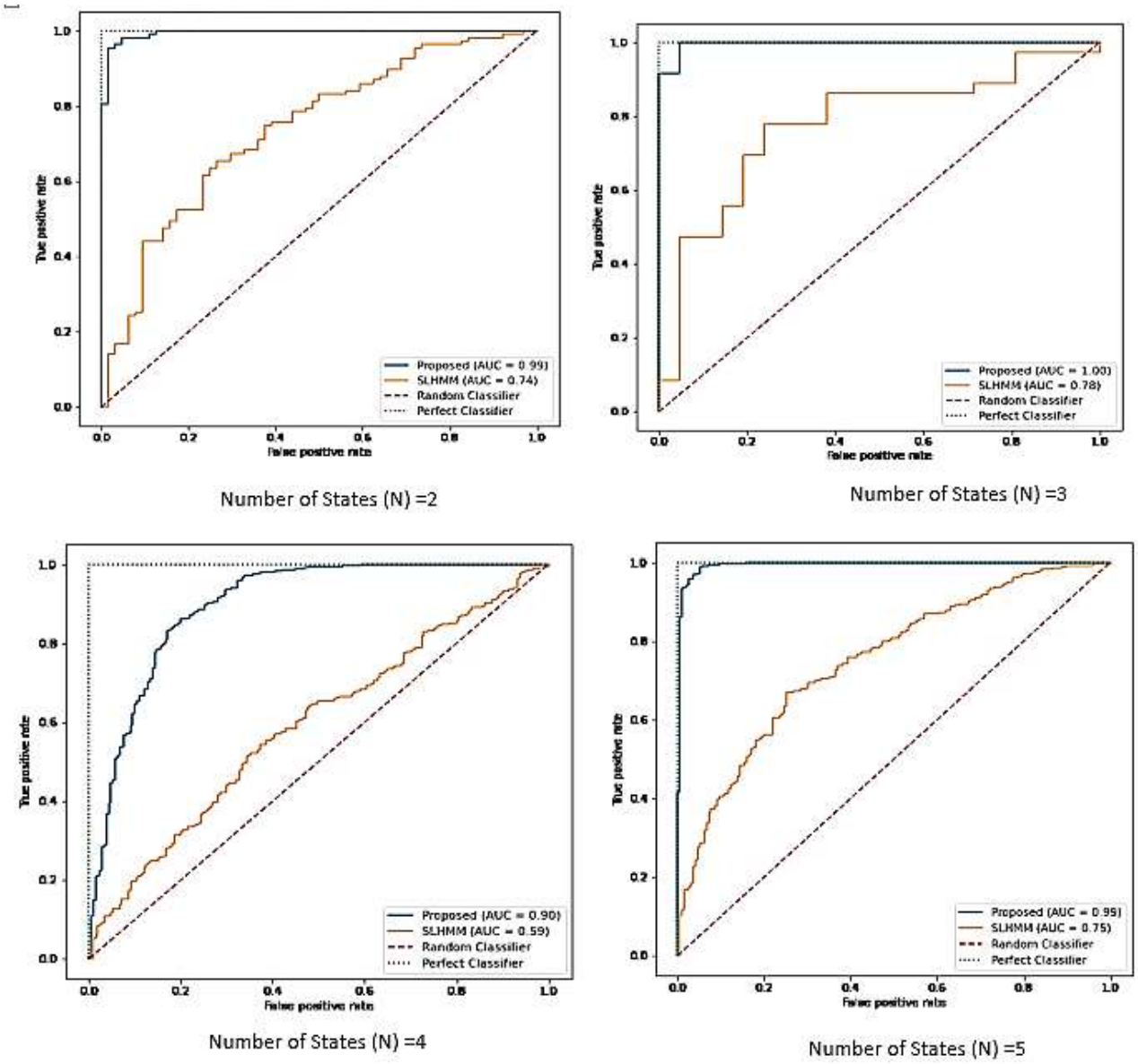

Figure 3. ROC Curve Plot for different number of states based on the MMT dataset.

Table 4. Performance of models based on training and detection times (s) for different number of states.

		PAYSIM dataset		MMT dataset	
Model	**# of hidden states**	**Training time(s)**	**Detection time(s)**	**Training time(s)**	**Detection time(s)**
SLHMM	2	0.025	0.012	0.033	0.011
	3	0.098	0.015	0.085	0.014
	4	0.099	0.023	0.098	0.022
	5	0.102	0.025	0.101	0.035
Proposed	2	0.021	0.002	0.02	0.002
	3	0.032	0.002	0.031	0.016
	4	0.045	0.005	0.012	0.004
	5	0.036	0.013	0.035	0.022

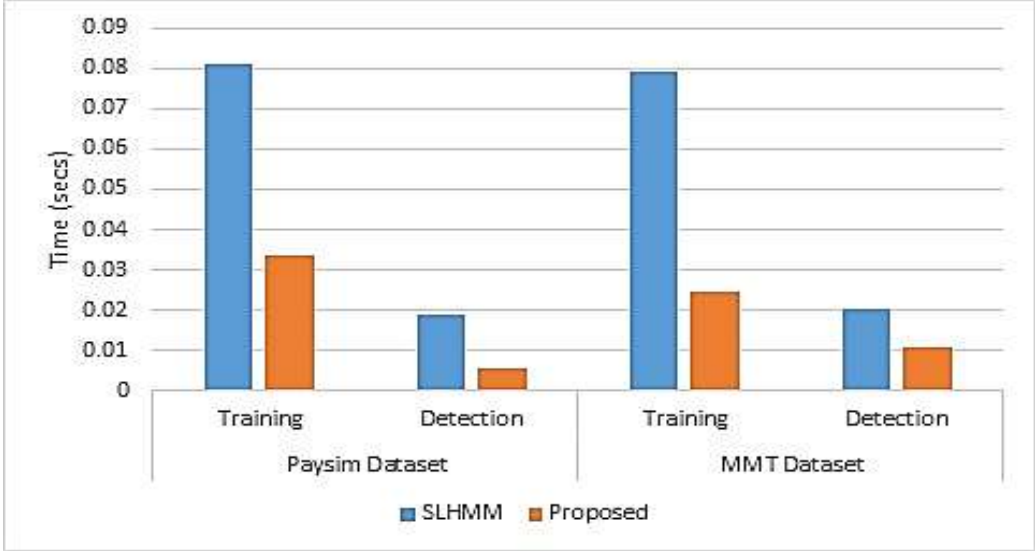

Figure 4. Models' performance based on average training and detection times (s).

5. Conclusion

A real-time financial fraud detection system based on multi-layered HMMs was proposed and implemented in this paper. A number of samples were created from the training dataset, where each reflected a specific customer transaction behaviour and enough to optimize the parameters of a single hidden Markov model. An incoming transaction was compared with a database containing normal user transaction behaviour before it was fed into the model in order to compute its probability of occurrence. It is evident from the simulation results that for the various number of hidden states, our proposed multi-layered HMM obtained a precision of 0.984 and 0.986, recall of 0.965 and 0.971, as well as an F1-score of 0.974 and 0.978 using the PAYSIM and MMT datasets, respectively. These values are better than those obtained by using the approach proposed by [14], which obtained an average precision of 0.949 and 0.927, recall of 0.91 and 0.918, and an f1-score of 0.929 and 0.922 on the PAYSIM and MMT datasets, respectively, for the various number of hidden states. It implies that our proposed model is able to better classify both genuine and fraudulent transactions with enhanced optimization and detection times.

References

[1] L.R.Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition", *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257-286, 1989.

[2] L. Duan, L. Xu, F. Guo, J. Lee, and B. Yan, "A local-density based spatial clustering algorithm with noise", *Information systems*, vol. 32, no. 7, pp. 978-986, 2007.

[3] A. Abdallah, A.M. Mohd, and Z. Anazida, "Fraud detection system: A survey" *Journal of Network and Computer Applications*, Vol. 68, pp.90-113, 2016

[4] A. Srivastava, K. Amlan, S. Shamik and Arun Majumdar, "Credit card fraud detection using hidden Markov model." *IEEE Transactions on dependable and secure computing*, vol. 5, no. 1, pp. 37-48, 2008.

[5] R. Ahmadian Ramaki, A. Rasoolzadegan, and A. Javan Jafari, "A systematic review on intrusion detection based on the Hidden Markov Model." *Statistical Analysis and Data Mining: The ASA Data Science Journal*, vol. 11, no. 3, pp. 111-134, 2018.

[6] B. Mor, S. Garhwal, and A. Kumar. "A systematic review of hidden Markov models and their applications." *Archives of computational methods in engineering*, *vol.* 28, pp. 1429-1448, 2021.

[7] W. K. Zegeye, R.A. Dean and F. Moazzami, F, "Multi-layer hidden Markov model-based intrusion detection system" Machine *Learning and Knowledge Extraction*, vol.*1 no. 1*, pp.265-286, 2018.

[8] X. D Hoang, J. Hu and P. Bertok, "A multi-layer model for anomaly intrusion detection using program sequences of system calls" *11th IEEE International Conference on Networks, 2003. ICON2003*, pp. 531-536, 2003.

[9] M. Penagarikano, and B. German "Layered Markov models: A New architectural approach to automatic speech recognition." In *Proceedings of the 2004 14th IEEE Signal Processing Society Workshop Machine*

Learning for Signal Processing pp. 305-314, IEEE, 2004.

[10] O. Abouabdalla, H. El-Taj, A. Manasrah and S. Ramadass, "False positive reduction in intrusion detection system: A survey" In *2009 2nd IEEE International Conference on Broadband Network & Multimedia Technology*, pp. 463-466, IEEE, 2009

[11] G.P. Spathoulas and S.K. Katsikas, "Reducing false positives in intrusion detection systems." *computers & security*, vol.*29, no.* 1, pp. 35-44, 2010.

[12] A. Prakash and C. Chandrasekar, "A novel hidden Markov model for credit card fraud detection." *International Journal of Computer Applications*, vol. 59, no. 3, pp.35-41, 2012.

[13] D.A. Burgio, "R*eduction of False Positives in Intrusion Detection Based on Extreme Learning Machine with Situation Awareness*", PhD diss., Nova Southeastern University, 2020.

[14] F.K. Alarfaj, I. Malik, H.U. Khan, N. Almusallam, M. Ramzan and M. Ahmed, "Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms" *IEEE Access*, vol.*10*, pp. 39700-39715, 2022.

[15] V. Baghdasaryan, H. Davtyan, A. Sarikyan and Z. Navasardyan," Improving tax audit efficiency using machine learning: The role of taxpayer's network data in fraud detection" *Applied Artificial Intelligence*, vol. *36*, no. 1, pp. 2012002, 2022.

[16] M. Valavan, and S. Rita. "Predictive-Analysis-based Machine Learning Model for Fraud Detection with Boosting Classifiers." *Computer Systems Science & Engineering*, vol. 45, no. 1, pp. 232-245, 2023.

[17] N.Q. Do, A. Selamat, O. Krejcar, E. Herrera-Viedma and H. Fujita, "Deep learning for phishing detection: Taxonomy, current challenges and future directions". *IEEE Access*, vol.*10*, pp. 36429-36463, 2022.

[18] H. Fujita, "Effectiveness of a hybrid deep learning model integrated with a hybrid parameterization model in decision-making analysis." In *Knowledge innovation through intelligent software methodologies, tools and techniques: proceedings of the 19th international conference on new trends in intelligent software methodologies, tools and techniques (SoMeT_20)*, vol. 327. 2020.

[19] E. Lopez-Rojas, A. Elmir and S. Axelsson, "PaySim: A financial mobile money simulator for fraud detection". In *28th European Modeling and Simulation Symposium, EMSS, Larnaca*, pp. 249-255. Dime University of Genoa, 2016.

[20] A. dedoyin, "*Predicting fraud in mobile money transfer* (Doctoral dissertation, University of Brighton), 2018.

[21] A.AA. Danaa, M.I. Daabo and A. Abdul-Barik, "An Improved Hybrid Algorithm for Optimizing the Parameters of Hidden Markov Models" *Asian Journal of Research in Computer Science*, vol.*10, no.* 1, pp. 63-73, 2021.

[22] R. Wedge, J.M. Kanter, K. Veeramachaneni, S.M. Rubio and S.I Perez, "Solving the false positives problem in fraud prediction using automated feature engineering in *Machine Learning and Knowledge Discovery in Databases" European Conference, ECML PKDD 2018, Dublin, Ireland,* pp. 372-388, 2019.

یک مدل مارکوف پنهان چند لایه برای تشخیص تقلب در زمان واقعی در معاملات مالی الکترونیکی

Abukari Abdul Aziz Danaa [1,*] و Mohammed Ibrahim Daabo [2] و Alhassan Abdul-Barik [3]

[1] گروه علوم کامپیوتر، دانشگاه فنی تاماله، تامال، غنا.

[2] گروه علوم کامپیوتر، دانشگاه فناوری و علوم کاربردی C. K. Tedam ، نارونگو، غنا.

[3] گروه علوم کامپیوتر، دانشگاه مطالعات توسعه، تامال، غنا.

ارسال ۲۰۲۲/۱۰/۱۹؛ بازنگری ۲۰۲۳/۰۱/۳۰؛ پذیرش ۲۰۲۳/۰۳/۲۷

چکیده:

مدل‌های پنهان مارکوف (HMMs) الگوریتم‌های یادگیری ماشینی هستند که برای طیف وسیعی از برنامه‌های کاربردی واقعی از جمله تشخیص نفوذ، تشخیص الگو، ترمودینامیک و مکانیک آماری و غیره استفاده شده‌اند .یک HMM چند لایه برای تشخیص و پیشگیری از تقلب در زمان واقعی و در عین حال کاهش شدید تعداد مثبت و منفی کاذب در این کار پیشنهاد و اجرا شده است .این مطالعه از شبیه‌ساز پرداخت (PAYSIM) و تراکنش پول موبایل(MMT) ، مجموعه داده‌ها استفاده می‌کند و بر کاهش بهینه‌سازی پارامتر و زمان‌های تشخیص مدل‌های پیشنهادی با استفاده از یک الگوریتم ترکیبی متشکل از الگوریتم‌هایBaum-Welch ، ژنتیک و بهینه‌سازی ازدحام ذرات تمرکز می‌کند. نتایج شبیه‌سازی نشان می‌دهد که برای حالت‌های پنهان اعداد مختلف، مدل پیشنهادی ما از نظر دقت (به ترتیب ۰,۹۸۴ و ۰,۹۸۶ برای مجموعه داده‌های PAYSIM وMMT)، یادآوری(۰,۹۶۵ و ۰,۹۷۱ برای مجموعه داده‌های PAYSIM وMMT ، به‌ترتیب بهتر عمل می‌کند). و امتیازاتF1 (به ترتیب ۰,۹۷۴ و ۰,۹۷۸ برای مجموعه داده‌های PAYSIMو MMT)در مقایسه با رویکرد موجود.

کلمات کلیدی: متقلبانه، مدل‌های پنهان مارکوف، بهینه‌سازی، احتمال، چند لایه.