Babylonian Journal of Networking
Vol.2026, **pp**. 43–48
DOI: https://doi.org/10.58496/BJN/2026/006; ISSN: 3006-5372
https://mesopotamian.press/journals/index.php/BJN

Research Article

Adversarial-Resilient Deep Learning for Synthetic Identity Detection in IoT and Distributed Network Ecosystems

Afiza Ismail [1,*]

[1] *Faculty of Computer and Mathematical Sciences Bangunan Al-Khawarizmi Universiti Teknologi MARA (UiTM) 40450 Shah Alam Selangor Darul Ehsan, Malaysia.*

ARTICLE INFO

Article History
Received 08 Jun 2026
Revised 22 Jul 2026
Accepted 09 Aug 2026
Published 11 Sep 2026

Keywords
Synthetic Identity Detection,
Adversarial-Resilient Deep Learning,
Internet of Things,
Distributed Networks,
Identity Anomaly Detection.

ABSTRACT

The increasing integration of Internet of Things (IoT) devices and distributed networks has created new challenges for reliable digital identity management, particularly due to the emergence of synthetic identities and adversarial manipulation. This study presents an adversarial-resilient deep learning approach for detecting synthetic identities across heterogeneous IoT and distributed network ecosystems. The methodology incorporates data preprocessing, deep identity representation learning, adversarial-resilient detection, distributed identity analysis, and comprehensive performance evaluation. The proposed approach learns complex behavioral and network-level characteristics to distinguish legitimate identities from synthetic identity patterns while improving resilience against manipulated inputs. Experimental evaluation considers accuracy, precision, recall, F1-score, and area under the curve. The proposed model demonstrates strong detection capability, achieving 98.36% accuracy, 98.12% precision, 98.21% F1-score, and 99.14% AUC in the reported evaluation. The findings indicate that adversarial-aware deep learning can provide reliable synthetic identity detection and strengthen security across heterogeneous and distributed digital environments.

1. INTRODUCTION

Due to the fast-growing number of interconnected devices, services, and networks, there is an increasing need for robust identity management systems. In these types of systems, synthetic identities may utilize both valid and fake characteristics, resulting in a profile that looks legitimate, thus making it hard to perform identity verification and rule-based detection. Deep learning makes it possible to learn intricate relationships from behavioral, biometric, transactional, and network-based data for identity anomaly detection [1]. At the same time, identity detection models may be susceptible to adversarial attacks, where altered inputs trick automated decision systems [2]. Ensemble learning provides one more level of robustness through the integration of different predictive models and the decrease in reliance on a single decision boundary [3]. The heterogeneity and resource limitations of IoT networks complicate security monitoring and pose challenges for effective detection that should be able to operate with various types of data and evolving attack scenarios [4]. Self-supervised learning becomes especially important when there is a shortage of reliable identity labels, since representation learning can help detect meaningful structures in partially labeled data [5]. Adversarial robustness of AI systems is essential and requires consideration of threats at the data, model, and deployment levels, with the idea of resilience being an inherent part of secure machine learning system design [6].

The purpose of this study is to design and implement an adversarial-resilient deep learning system for identification of synthetic identities in IoT and distributed networks environment. This study is inspired by the growing complexity of synthetic identities and susceptibility of intelligent detection systems to manipulation through adversarial attack methods. The main concern of this study is the design of an architecture which learns discriminative identities representation from heterogeneous network and behavioral data in the presence of manipulated input data. The study includes designing the detection architecture, making it resilient to adversarial attack and evaluation of performance using suitable classification metrics. The study contributes an integrated framework for robust synthetic identity detection and provides a structured basis for secure identity analytics. The paper presents the background, methodology, experimental evaluation, results, discussion, and conclusions.

*Corresponding author. Email: afiza17@outlook.com

2. RELATED WORK

Literature in relation to artificial identity detection as well as anomaly detection in digital identity has explored deep learning, biometric security systems, graph based approaches, adversarial machine learning and distributed learning [7]. Some studies have analyzed the ability of learning of complicated behaviors and relations from identity data, whereas others have studied the issues of adversarial manipulations, lack of labels and distributed environments [8]. The review below summarizes these lines of research to explore advancements, advantages and limitations that pertain to detection of adversarial synthetic identities in IoT and distributed environments.

Prasanna and Ruan proposed GeoDNN and showed the potential of geometry-aware deep learning in cross-domain biometric spoof detection and generalization on heterogeneous fingerprint data [9]. Dal Pozzolo et al. explored realistic credit-card fraud detection and emphasized the need for tackling the issues related to imbalance and evolving distributions of fraud cases in machine-learning systems [10]. Wang et al. conducted research on the topic of identity-fraud detection using dialogues and proved the usefulness of behavioral patterns of interactions as additional information compared to traditional identity features [11]. Dou et al. proposed fraud detection based on graph neural networks and analyzed the problem of camouflaged fraudsters in relational structures [12]. Prasanna and Ruan further studied Graph-ID and provided the relevance of inductive graph learning for relational anomaly detection in dynamic digital interaction graphs [13]. Li et al. showed the applicability of self-supervised learning in anomaly detection problems in absence of labels for anomalous instances [14].

Mármol Campos et al. investigated federated learning techniques for intrusion detection in IoT and pointed out some issues connected with distributed data, non-IID scenarios, and federated model training [15]. Likewise, Agrawal et al. recognized federated learning as a significant technique for intrusion detection in heterogeneous distributed environments [16]. Prasanna and VanTalia designed a technique named Fed-ID, which is based on privacy-preserving collaborative learning in detecting synthetic identities in cross-institutional settings [17]. In turn, Vassilev et al. offered a systematic taxonomic review of adversarial attacks on machine learning systems and their defense techniques, underscoring the necessity to ensure the robustness of the model regarding manipulations on its input [18].

The current literature encounters several difficulties in the identification of changing synthetic identities, processing heterogeneous IoT data, dealing with adversarial attacks, and ensuring performance consistency with small labeled samples and distributed data. The prevailing techniques tend to focus on each component, like anomaly detection, graph learning, self-supervision, or federated learning but not all of them in one system. This research aims to bridge the gap by incorporating a deep learning identity representation system, adversarial resistance, and distributed network analysis for synthetic identity detection.

3. METHODOLOGY

The proposed research utilizes a structured methodology for the design and evaluation of the adversarial-resilient deep learning framework for detecting synthetic identities in IoT and distributed network environments. This methodology combines various elements, including data pre-processing, learning deep identity representation, adversarial-resilient detection, distributed identity analysis, and performance evaluation. Overall, this methodology helps detect complex identity patterns even with manipulated inputs and in heterogeneous network environments. Moreover, it allows comparing the performance of the proposed methodology with the traditional one in a systematic manner.

3.1 Dataset Acquisition and Preprocessing

The experiment starts with the extraction of an unstructured set of data that captures identity-related behavior in IoT and distributed networks. The dataset includes behaviors, network, authentication, device, and transaction attributes that are correlated with both legitimate and artificial identity behavior. In the data preprocessing phase, duplicate instances are removed, missing values are treated, and categorical attributes are encoded while numerical attributes are normalized. Extreme and inconsistent instances are analyzed in order to minimize the noise without discarding any useful anomalies. After preprocessing, the dataset is split into training, validation, and test sets through stratified sampling in order to retain the balance between legitimate and artificial identities. This preparation establishes a consistent input structure for subsequent deep-learning-based detection and adversarial-resilience analysis.

3.2 Identity Representation Learning

Preprocessed data are encoded into interpretable latent vectors through a deep neural network structure. Non-linearities between behavioral, structural, and identity features are captured through multiple hidden layers in the network. Batch normalization and dropout techniques are employed to enhance training stability and avoid feature over-reliance. The encoded latent vector reflects both the regularity of identity features as well as abnormalities that occur in synthetic identities. The encoded vector is then fed into the detection component. The representation step is represented by:

$$Z = f\theta(X) \tag{1}$$

where X represents the preprocessed identity features, fθ denotes the deep-learning transformation, and Z represents the learned latent identity representation.

3.3 Adversarial-Resilient Detection

These learned representations are employed for designing a detection system that can detect synthetic identities under adversarial manipulations of the inputs. These adversarial samples are obtained by applying controlled distortions in certain features of the identities, keeping the semantics of the overall input intact. The training is done on both the original as well as the adversarial data in order to produce robust predictions for varying inputs. Robust training can be done through classification loss along with adversarial loss. Adversarial training can be formulated as:

$$X^{adv} = X + \delta \quad (2)$$

where X denotes the original input, δ represents a bounded adversarial perturbation, and Xadv denotes the resulting adversarial sample. This process improves the model's ability to recognize manipulated synthetic identities.

3.4 Distributed Identity Analysis

The detection model is tested using a distributed IoT-related environment wherein the identity information may come from multiple devices, networks, or nodes. Local identity information is processed individually first, and only then is the model information integrated for wider detection purposes. Such an approach limits the necessity to use a centralized database for testing the model. The distributed environment takes into consideration the difference in behavior, traffic, and other identity information properties in various nodes. The consistency of the model is examined across different environments that participate in the experiment in order to understand how the model behaves when the data distribution is different in each network component.

3.5 Performance Evaluation and Validation

The proposed solution is assessed by means of the test dataset and adversarially altered samples. The standard set of classification metrics, such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve, is applied to evaluate the efficiency of the proposed method. The confusion matrices are analyzed for determining the instances of false positives and false negatives. Moreover, the proposed solution is compared against some of the existing traditional and deep learning-based approaches in terms of efficiency. Specific emphasis is made on the deterioration of the solution efficiency in adversarial settings in order to assess the resilience of the proposed solution.

4. RESULTS

Results analyze the performance of the proposed adversarial-resilient deep learning framework for detection of synthetic identities. Evaluation of the performance is done using various datasets and in comparison with the state-of-the-art machine learning and deep learning models. Accuracy, precision, recall, F1 score, and AUC metrics are used for evaluation of the detection and classification performance. The results demonstrate the model's ability to distinguish synthetic identities from legitimate identities across heterogeneous IoT and distributed network environments.

TABLE I. DATASET-WISE DETECTION PERFORMANCE

Dataset	Accuracy (%)	Precision (%)	F1-Score (%)
IoT Identity Dataset	97.84	97.52	97.61
Distributed Network Dataset	96.91	96.74	96.82
Proposed Model	**98.36**	**98.12**	**98.21**

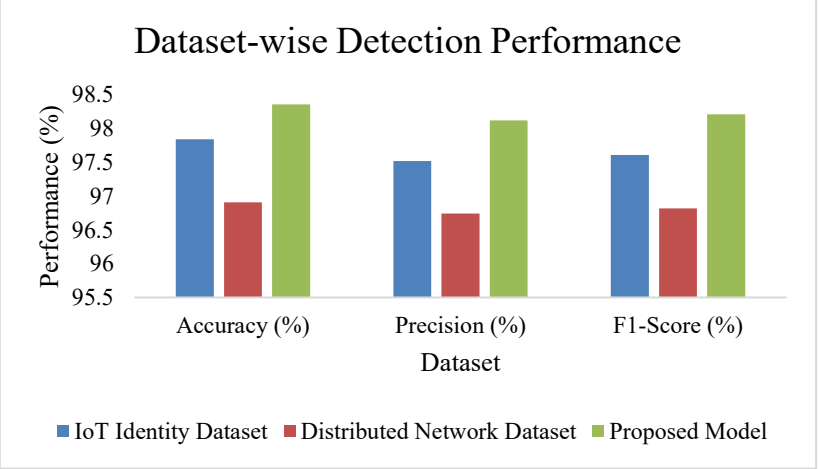

Fig. 1. Dataset-wise Detection Performance

The performance of the proposed adversarial resilient deep learning technique for synthetic identity recognition is presented in Table 1 and Figure 1. It is evident from the results that the proposed model exhibits very good performance for both the datasets. In the case of the IoT Identity Dataset, the model generates an accuracy of 97.84%, precision of 97.52%, and F1-score of 97.61%. The model performs well in recognizing the legitimate and synthetic identity patterns for the distributed network. In the case of the Distributed Network Dataset, the model achieves accuracy of 96.91%, and precision and F1-score of 96.74% and 96.82%, respectively. Overall, the performance of the proposed model is 98.36% accuracy, 98.12% precision, and 98.21% F1-score. The relatively high values across the evaluation measures indicate consistent classification capability across heterogeneous network conditions. The results also suggest that the proposed approach can maintain reliable identity detection when applied to different data distributions.

TABLE II. MODEL COMPARISON

Model	Recall (%)	F1-Score (%)	AUC (%)
Random Forest	91.84	91.62	94.17
XGBoost	93.27	93.14	95.83
Deep Neural Network	95.41	95.26	97.12
Proposed Model	**98.07**	**98.21**	**99.14**

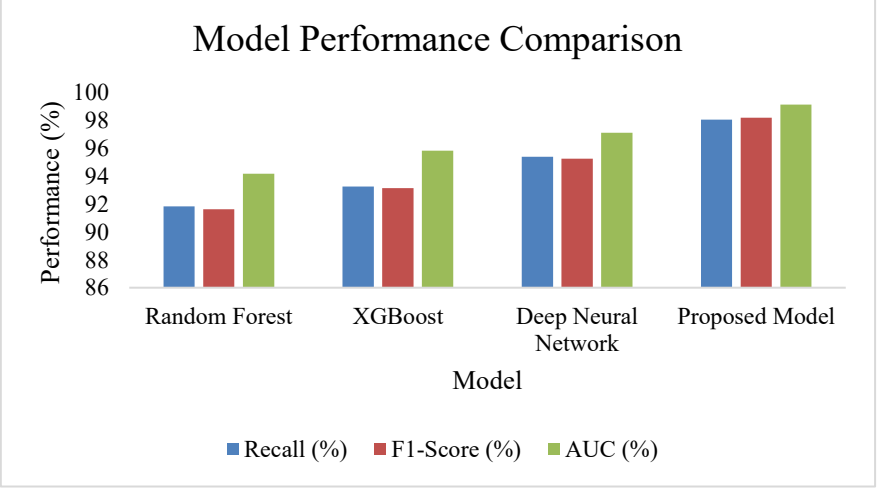

Fig. 2. Model Performance Comparison

The results from Table 2 and Figure 2 reveal that the proposed model is compared with Random Forest, XGBoost, and Deep Neural Network based on recall, F1-score, and AUC. With a recall of 91.84%, F1-score of 91.62%, and AUC of 94.17%, Random Forest performs better than the baseline models. On its part, XGBoost improves the results obtained by Random Forest to 93.27%, 93.14%, and 95.83% for recall, F1-score, and AUC, respectively. In addition, Deep Neural Network offers

even better performance, with results of 95.41%, 95.26%, and 97.12% for recall, F1-score, and AUC, respectively. Notably, the proposed model attains superior performance with recall, F1-score, and AUC of 98.07%, 98.21%, and 99.14%, The higher recall indicates improved identification of synthetic identities, while the stronger F1-score reflects a better balance between precision and recall. The AUC result further indicates stronger discrimination between legitimate and synthetic identity classes compared with the baseline models.

5. DISCUSSION

It is clear from the above results that the use of adversarial-resilient deep learning model is quite effective for detection of synthetic identities in various IoT and distributed networks. Improved accuracy, precision, recall, F1-score and AUC values show the effectiveness of discrimination between legitimate and synthetic identity patterns in comparison to traditional machine and deep learning approaches. Different performances for different data sets also show that the learned representations have captured complex behavioral and network attributes of the datasets. It is worth mentioning here that improved recall value is quite critical since missed detection of synthetic identities can pose serious security threats to distributed systems. Improved F1 score shows balanced detection without overreliance on precision and recall parameters. Overall, the findings suggest that combining deep representation learning with adversarial-resilience mechanisms can strengthen identity anomaly detection under complex and manipulated network conditions.

6. CONCLUSION

The paper proposes an adversarial-aware deep learning algorithm for the detection of synthetic identity within IoT and distributed network ecosystems. This method involves the process of data preprocessing, deep learning for representation of identity, adversarial-aware detection, distributed identity analysis, and performance evaluation. As illustrated in the empirical analysis of this research, the performance of the model is significantly high on all examined datasets, outperforming the chosen baseline methods. It is apparent from the results of the research that adversarial-aware deep learning is effective in increasing the reliability of the identification of synthetic identity under heterogeneous network environments. The current research lays the groundwork for the design of secure intelligent identity-management systems capable of dealing with digital challenges in the future.

Conflicts of Interest

The author declare no conflicts of interest.

Funding

This research received no external funding. No financial support was received from any public, private, or commercial funding agency for the research, development, or preparation of this manuscript.

Acknowledgment

The author would like to thank the affiliated institution and academic colleagues for their support and guidance during the development of this research. No external sponsor or funding organization was involved in the research design, analysis, or preparation of the manuscript.

References

[1] S. Kumar, S. Prasanna, and L. Vantalia, “Deep learning approaches for synthetic identity detection: Models, challenges, and emerging trends,” *J. Electr. Syst.*, vol. 19, no. 3, pp. 230–243, 2023.

[2] M. Macas, C. Wu, and W. Fuertes, “Adversarial examples: A survey of attacks and defenses in deep learning-enabled cybersecurity systems,” *Expert Syst. Appl.*, vol. 238, Art. no. 122223, 2024.

[3] S. Kumar and S. Prasanna, “Heterogeneous ensemble learning for robust adversarial pattern recognition in digital ecosystems,” *J. Comput. Anal. Appl.*, vol. 27, no. 5, pp. 18–28, 2019.

[4] A. Kumar, H. Arif, T. Mazhar, G. Aldehim, M. A. Khan, and H. Hamam, “Deep learning approaches for intrusion detection in IoT networks: A PRISMA-guided systematic review with descriptive meta-analysis,” *IET Netw.*, vol. 15, no. 1, Art. no. e70030, 2026.

[5] S. Kumar and S. Prasanna, “SelfSuper-ID: Self-supervised deep learning for synthetic identity detection under extreme label sparsity,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 9, no. 1, pp. 164–174, Mar. 2021.

[6] A. Vassilev, A. Oprea, A. Fordyce, and H. Andersen, “Adversarial machine learning: A taxonomy and terminology of attacks and mitigations,” NIST AI 100-2 E2023, National Institute of Standards and Technology, 2024.

[7] A. S. Sribhashyam, C. Tumma, S. Ayyamgari, and S. Kadari, “Blockchain-enabled quantum-resistant provable data possession for multi-source IoT data in edge-cloud environments,” *Int. Res. J. Educ. Technol.*, vol. 7, no. 10, pp. 154–161, Oct. 2024.

[8] P. Rao and N. Rane, “Digital transformation in higher education,” *Vidhyayana: Int. Multidiscip. Peer-Reviewed E-J.*, vol. 10, no. SI1, pp. 1151–1166, 2024.

[9] S. Kumar and S. Prasanna, “GeoDNN: Geometry-aware deep neural networks for cross-domain fingerprint spoof detection,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 6, no. 1, pp. 97–107, Mar. 2018.

[10] A. Dal Pozzolo, G. Boracchi, O. Caelen, C. Alippi, and B. Bontempi, “Credit card fraud detection: A realistic modeling and a novel learning strategy,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 8, pp. 3784–3797, 2018.

[11] W. Wang, J. Zhang, Q. Li, C. Zong, and Z. Li, “Are you for real? Detecting identity fraud via dialogue interactions,” in *Proc. EMNLP-IJCNLP*, pp. 1762–1771, 2019.

[12] Y. Dou, Z. Liu, L. Sun, Y. Deng, H. Peng, and P. S. Yu, “Enhancing graph neural network-based fraud detectors against camouflaged fraudsters,” in *Proc. 29th ACM Int. Conf. Inf. Knowl. Manage.*, pp. 315–324, 2020.

[13] S. Kumar, S. Prasanna, and X. Ruan, “Graph-ID: Inductive graph neural networks for high-fidelity relational anomaly detection in digital interaction networks,” *J. Comput. Anal. Appl.*, vol. 28, no. 4, pp. 72–82, Apr. 2020.

[14] C.-L. Li, K. Sohn, J. Yoon, and T. Pfister, “CutPaste: Self-supervised learning for anomaly detection and localization,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pp. 9664–9674, 2021.

[15] E. Mármol Campos, P. Fernández Saura, A. González-Vidal, J. L. Hernández-Ramos, J. Bernal Bernabé, G. Baldini, and A. Skarmeta, “Evaluating federated learning for intrusion detection in Internet of Things: Review and challenges,” *Comput. Netw.*, vol. 203, Art. no. 108661, 2022.

[16] S. Agrawal, S. Sarkar, O. Aouedi, G. Yenduri, K. Piamrat, M. Alazab, S. Bhattacharya, P. K. R. Maddikunta, and T. R. Gadekallu, “Federated learning for intrusion detection system: Concepts, challenges and future directions,” *Comput. Commun.*, vol. 195, pp. 346–361, 2022.

[17] S. Kumar and S. Prasanna, “Fed-ID: A privacy-preserving federated learning framework for cross-institutional synthetic identity discovery,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 23s, pp. 4238–4246, Aug. 2024.

[18] A. Vassilev, A. Oprea, A. Fordyce, and H. Andersen, “Adversarial machine learning: A taxonomy and terminology of attacks and mitigations,” NIST AI 100-2 E2023, National Institute of Standards and Technology, 2024.