INTERNATIONAL JOURNAL OF APPLIED METHODS IN ELECTRONICS AND COMPUTERS

www.ijamec.org

International Open Access

Volume 12
Issue 04

December, 2024

Research Article

An Artificial Neural Network-Based Caller Authentication and Identification Algorithm in Cellular Communication Networks

Bem Sombo [a,*], ***Simon Tooswem Apeh*** [a], ***Isi Arthur Edeoghon*** [a]

[a] *Department of Computer Engineering, University of Benin, Benin City, Nigeria*

ARTICLE INFO

Article history:
Received 18 October 2024
Accepted 16 November 2024

Keywords:
Artificial Neural Networks,
Caller Authentication,
Cellular Communication,
Prevention of Black Market SIMs,
SIM User Identification

ABSTRACT

Cellular communication companies have created tens of thousands of jobs within the industry. However, while business opportunities have expanded for legitimate enterprises, criminal gangs have also exploited the new business environment to further their illegal activities. Weaknesses in the authentication algorithm allow criminals to commit fraud while remaining anonymous, thereby facilitating wireless crime. This paper presents the development of an artificial neural network-based caller authentication and identification algorithm for cellular communication networks that addresses these weaknesses. Voice data were collected from volunteer participants and the Mozilla Common Voice (MCV) database. Voice feature vectors were extracted from the voice data using the Mel Frequency Cepstral Coefficient (MFCC) technique. The extracted voice feature vectors were used to train a Multilayer Perceptron (MLP) neural network for voiceprint generation. The MLP neural network architecture was optimized through Neural Architecture Search (NAS) using the Neural Network Intelligence (NNI) toolkit. The optimized MLP neural network architecture was trained with the extracted voice feature vectors to generate a voiceprint generation model. The developed voiceprint model was then deployed to create an artificial neural network-based caller authentication and identification algorithm in cellular communication networks. The developed algorithm was evaluated in the cellular communication emulation setup using accuracy, false acceptance rate (FAR), and false rejection rate (FRR) as metrics. The results of the performance evaluation of the developed algorithm for authenticating first-time users of a new SIM and identifying third-party SIM users both showed 100% accuracy, 0% FAR, and 0% FRR. These results indicate that the developed algorithm has strong potential to prevent the use of black market SIMs and deter perpetrators from using third-party SIMs to access network services while remaining anonymous.

1. Introduction

The SIM authentication and subscriber identification algorithms are used in cellular communication systems to prevent criminals from committing fraud through cellular services. The SIM authentication algorithm verifies the identity of SIM subscribers before granting access to network services, while the identification algorithm helps security agencies trace criminal activities involving cellular communication. However, criminals have increasingly exploited cellular services to maintain anonymity and evade detection (Okonji, 2022) due to vulnerabilities in these algorithms.

Cryptographic challenge-response authentication algorithms are employed in cellular networks, spanning 2G to 5G technologies (Alezabi et al., 2014; Nakarmi, 2021). These algorithms have evolved from one-way authentication in 2G, which verifies only the SIM, to mutual authentication in later generations, where both the SIM and network are verified (Tsay & Mjolsnes, 2012). Despite these improvements, authentication remains limited to the legitimacy of the SIM and network, allowing users with valid SIM cards to remain anonymous.

Measures such as caller identification apps like Truecaller, mobile phone security features, and SIM registration policies have been implemented to enhance

* Corresponding author. E-mail address: *sombopeter@gmail.com*
DOI: 10.58190/ijamec.2024.111

security (Asda Mobile, 2022; Douglas & Mumbi, 2018; GSMA, 2016; Indriyani & Aprinia, 2022; Mohammed, 2023). Truecaller, for example, helps users identify unknown callers by displaying their names using a large phone number database (Truecaller, 2021). While these measures add some level of protection, they can be circumvented by criminals using unregistered SIM cards, stolen devices, or the phones of victims to preserve their anonymity.

SIM registration requires collecting and verifying a user's personal information linked to their SIM card to establish identity, prevents anonymous usage, and combat criminal activities like terrorism, kidnapping, and fraud (GSMA, 2016; Okonji, 2022). However, this measure is often undermined when criminals use SIM cards that are not registered under their personal details. For instance, during the March 2022 Abuja-Kaduna train attack in Nigeria, kidnappers used victims' phones, which had registered SIM cards, to contact families while maintaining anonymity (Okonji, 2022). This situation highlighted a significant weakness in relying solely on SIM registration for security.

A review of existing authentication algorithms reveals vulnerability in not verifying the actual user behind the SIM. Current solutions primarily focus on securing devices and encrypting data rather than addressing the gap in user verification. To overcome this challenge, this paper presents a novel artificial neural network-based caller authentication and identification algorithm that aims to prevent the misuse of unregistered and third-party SIMs while enhancing user identification in cellular networks.

The paper is organized as follows: Section 2 reviews related works, Section 3 details the methodology, including the technical background and development workflow for the proposed algorithm, Section 4 presents the results and discussion, and Section 5 concludes the paper.

2. Related Works

Luo and Zhang (2023) proposed the secure data transmission method of a cellular communication network based on the asymmetric encryption algorithm. The research implemented an asymmetric encryption algorithm for secure data transmission in cellular networks. It analyzed the structure of cellular networks, established a network architecture model, and constructed a public key system. The encryption process was optimized for efficiency. The effectiveness of the method was evaluated by comparing performance metrics with existing methods, while also analyzing its security benefits and impact on transmission energy consumption. The study yielded significant results demonstrating the efficacy of the proposed method for secure data transmission in cellular networks. It showed a marked improvement in both throughput and reduction in bit error rate compared to existing methods.

Kazmi et al. (2023) put forward security concepts in emerging 6G communication: threats, countermeasures, authentication techniques, and research directions. The methodology of the research involved a comprehensive literature review on the landscape of 6G cellular communication and its security aspects. From this review, a taxonomy of the threat model in 6G communication was developed, focusing on five key security concepts: Confidentiality, Integrity, Availability, Authentication, and Access control (CIA3). The result of the study presents a structured framework for understanding and addressing security challenges in the context of 6G communication networks.

Zukarnain et al. (2022) present authentication securing methods for mobile identity: issues, solutions and challenges. The methodology of the study involved a comprehensive review of existing literature on smartphone security and authentication methods and identifies common issues and challenges. Through this review, various authentication methods commonly used by mobile users were identified, including passwords, short messages, biometric techniques, and identity-based cryptography. The study then conducted an analysis of the security vulnerabilities inherent in these methods and evaluated the effectiveness and feasibility of different authentication solutions in overcoming the identified challenges. Based on this analysis and evaluation, the study proposed the best authentication solution or combination of solutions to address the identified issues and enhance smartphone security.

Khan et al. (2022) put forward lightweight multifactor authentication scheme for NextGen cellular networks. The methodology for the study involved identifying the security challenges arising from the integration of intelligent relays and quantum computing support in 5G/6G cellular networks, focusing on device-to-device (D2D) communication. A comprehensive literature review was conducted to understand the existing research landscape and identify gaps. Subsequently, the researchers developed a Lightweight ECC-based Multifactor Authentication Protocol (LEMAP) tailored for miniaturized mobile devices to address the identified challenges. The protocol design incorporated multifactor authentication mechanisms to mitigate potential attacks while maintaining low computation and communication costs.

Ibrahim and Jauro (2021) presented biometric encryption of data using voice recognition. The work proposed a software-based architecture solution for biometric encryption of data using voice recognition that employed dynamic time warping (DTW) technique to solve the problem of speech biometric duration varying with non-linear expansion and contraction. A database was

employed to store the monolithically bind cryptography key with the equivalent biometric hardened template of the user in such a manner that identity of the key will stay hidden unless there is a successful biometric authentication by intended user. The proposed solution was evaluate and the result shows that it offers a better substitute method of user authentication than the traditional pre-shared keys for benefit of protecting secrete keys.

Fan et al. (2021) proposed cross-network-slice authentication scheme for the 5th generation mobile communication system. The methodology of the study involved conducting a comprehensive literature review on the characteristics and challenges of fifth-generation mobile networks (5G), focusing on authentication procedures. Identified challenges related to traditional authentication procedures in 4G networks led to the development of a novel authentication scheme tailored for 5G networks. This scheme aimed to decentralize authentication to edge clouds to achieve low latency and reduce reliance on the core network, thus decreasing time latency.

A meta-analysis table that summarizes the critique of the literature review is presented in Table 1.

1. Methodology

1.1. Background

1.1.1. Voice Feature Extraction

The voice signal contains some irrelevant features that are not useful in recognizing a speaker's voice. In order to facilitate the processing of voice signals for speaker voice recognition purposes, some feature extraction techniques have to be deployed to extract the relevant features in the voice signal that can facilitate the learning of unique voice features for speaker recognition task. The Mel-Frequency Cepstral Coefficients (MFCC) is one of the most commonly used feature extraction techniques.

1.1.1.1. Mel Frequency Cepstral Coefficient (MFCC)

The Mel Frequency Cepstral Coefficient (MFCC) is a technique that captures the spectral characteristics of voice signals in alignment with human auditory perception. The process begins with pre-emphasis, which enhances high-frequency components of the signal to improve auditory analysis, as described in Equation (1). The signal is then segmented into short frames and windowed to reduce spectral leakage. Each frame undergoes transformation to the frequency domain using the Fast Fourier Transform (FFT), as indicated in Equation (3). The linear frequency spectrum is then mapped onto the Mel scale, approximating human pitch perception, according to Equation (4). The frequency spectrum is processed through triangular filters, and logarithmic compression is applied to the filter outputs to model human loudness perception. The energies are then transformed using the Discrete Cosine Transform (DCT) to produce compact cepstral coefficients, referenced in Equation (5). Liftering is subsequently applied to emphasize higher-order MFCCs, capturing spectral slope information, as shown in Equation (6). This extraction process enhances audio data representation, which is crucial for speaker recognition and audio analysis applications.

$$y(t) = x(t) - \alpha x(t-1) \ldots\ldots\ldots\ldots\ldots \quad (1)$$

where $y(t)$ is the pre-emphasized signal, $x(t)$ is the original signal and α is the pre-emphasis coefficient

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) \ldots\ldots\ldots \quad (2)$$

where n is the sample number and N is the total number of samples

$$X(k) = \sum_{n=0}^{N-1} x(n)w(n)e^{-j2\pi kn/N} \ldots\ldots \quad (3)$$

where $X(k)$ is the k-th frequency component of the Fourier transform, $x(n)$ is the n-th sample of the windowed frame, $w(n)$ is the window function value at sample n and N is the number of samples in the frame.

$$Mel(f) = 2595log_{10}^{(1+f/700)} \quad (4)$$

where $Mel(f)$ is the Mel frequency corresponding to linear frequency f

$$C(m) = \sum_{k=0}^{k-1} \log\big(S(k)\big)\cos\left(\frac{\pi m(k+0.5)}{K}\right). \quad (5)$$

where $C(m)$ is the m-th MFCC, $S(k)$ is the k-th Mel spectrum coefficient, and K is the number of Mel spectrum coefficient.

$$C'(m) = C(m) \times \left(1 + \left(\frac{L}{2}\right) \times \sin\left(\frac{\pi m}{L}\right)\right) \quad (6)$$

where $C'(m)$ is the lifted MFCC and $L$is the liftering coefficient

1.1.2. Multi-Layer Perceptron (MLP) Networks

The Multi-Layer Perceptron (MLP) is a cornerstone in neural network technology, distinguished by its structured architecture comprising input, hidden, and output layers connected by weighted connections. Recognized for its adaptability, MLPs excel in various tasks like classification, regression, and pattern recognition, leveraging their capability to discern intricate relationships between input data and desired output. Moreover, MLPs offer interpretability, providing insights into decision-making processes. Despite challenges in training deep architectures due to the vanishing gradient problem, MLPs find extensive application across diverse domains including speaker recognition. In this domain, MLPs contribute to tasks such as accurately identifying speakers by analyzing their speech patterns. The general

architecture of MLP neural network is shown in Figure 1.

1.1.3. Cosine Similarity

Cosine similarity is an essential concept in the fields of data science, text analysis, machine learning, and much more; it is a mathematical way to measure how similar two sets of data, which are converted into vectors, are (Miesle, 2023). Cosine similarity quantifies the similarity between two vectors by measuring the cosine of the angle between them. It focuses on the angle between two vectors, thus making cosine similarity more robust than other similarity measures in capturing the pattern similarities between two sets of data, even if their magnitudes differ.

Table 1. Meta analysis table of related works

S/N	Author and Publication Date	Title	Method	Strength	Remark
1.	Luo and Zhang (2023)	The secure data transmission method of a cellular communication network based on the asymmetric encryption algorithm	The research implemented an asymmetric encryption algorithm for secure data transmission in cellular networks.	It showed a marked improvement in both throughput and reduction in bit error rate compared to existing methods.	The proposed method cannot stop perpetrators from using cellular communication services while maintaining their anonymity.
2	Kazmi et al. (2023)	Security concepts in emerging 6G communication: threats, countermeasures, authentication techniques, and research directions	A taxonomy of the threat model in 6G communication was developed, focusing on five key security concepts: Confidentiality, Integrity, Availability, Authentication, and Access control (CIA3).	The study presents a structured framework for understanding and addressing security challenges in the context of 6G communication networks.	The study does not propose a solution to address third-party SIM user anonymity in cellular communication systems.
3	Zukarnain et al. (2022)	Authentication securing methods for mobile identity: issues, solutions and challenges.	An analysis of the security vulnerabilities inherent in authentication methods and evaluation of the effectiveness and feasibility of different authentication solutions in overcoming the identified challenges.	The study identifies vulnerabilities in common authentication methods and proposed effective solutions.	The proposed solutions in the study does not improve the authentication mechanism in the cellular communication system
4	Khan et al. (2022)	Lightweight multifactor authentication scheme for NextGen cellular networks	The methodology for the study involved identifying the security challenges arising from the integration of intelligent relays and quantum computing support in 5G/6G cellular networks, focusing on device-to-device (D2D) communication.	The study offers a viable solution that effectively mitigates potential attacks while minimizing computation and communication costs.	It leaves the challenge of perpetrator anonymity unaddressed.
5	Ibrahim and Jauro (2021)	Biometric encryption of data using voice recognition.	The approach involved a software-based architecture solution for biometric encryption of data using voice recognition that employed dynamic time warping (DTW) technique	The security and convenience of using the proposed solution will promote widespread attention for cryptographic systems.	The system can not reveal the identity of the individual authorized which can be useful during investigation purposes.
6	Fan et al. (2021)	Cross-network-slice authentication scheme for the 5th generation mobile communication system.	The identified challenges related to traditional authentication procedures in 4G networks led to the development of a novel authentication scheme tailored for 5G networks.	The study offers a practical solution to improve authentication procedures, reduce latency, and enhance network reliability.	The novel authentication scheme proposed does not deter perpetrators from exploiting third-party SIM user anonymity in cellular communication systems.

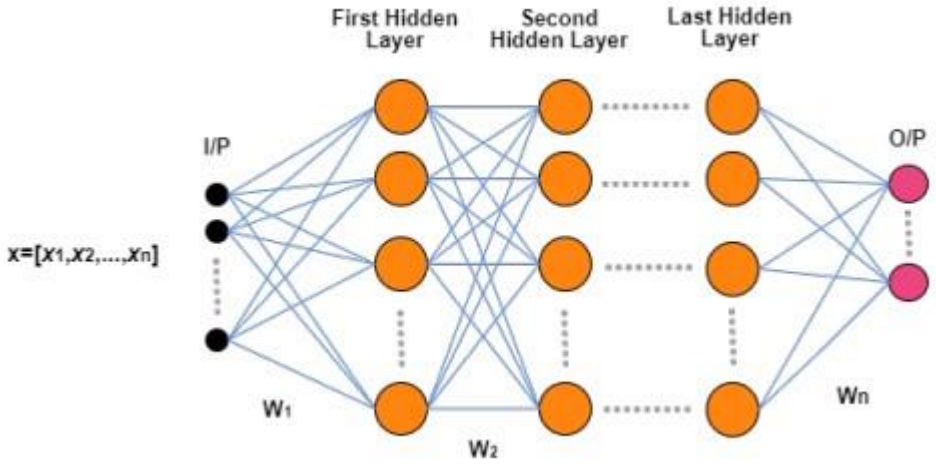


Figure 1. General architecture of MLP neural network architecture (Botalb et al.,2018)

It is a versatile metric widely employed to quantify the similarity between vectors in multidimensional spaces by measuring the cosine of the angle between them, making it particularly suited for scenarios where vector magnitudes vary. Cosine similarity finds application in voiceprint recognition, where voiceprints are represented as numerical vectors based on extracted acoustic features. By comparing the cosine similarity between voiceprint vectors, one can assess the resemblance between different voiceprints, facilitating tasks like speaker verification and identification. Cosine similarity's focus on vector direction, coupled with its insensitivity to scale variations, renders it advantageous in voiceprint comparison, contributing to advancements in biometric technologies. The mathematical expression for cosine similarity between two vectors A and B is given in (7).

$$Cosinesimilarity = cos\emptyset = \frac{A.B}{||A||.||B||} \quad (7)$$

where

A·B is the dot product of the vectors A and B

||A|| is the magnitude (or norm) of vector A, calculated as $\left||A|\right| = \sqrt{\sum_{i=1}^{n} A_i^2}$

||B|| is the magnitude (or norm) of vector B, calculated as $\left||B|\right| = \sqrt{\sum_{i=1}^{n} B_i^2}$

The cosine similarity will return a value between -1 and 1; a value closer to 1 indicates greater similarity.

1.1.4. Microsoft Neural Network Intelligence (NNI) Toolkit

The Neural Network Intelligence (NNI) toolbox, developed by Microsoft, is an open-source toolkit designed to automate and optimize the process of neural network model development. It simplifies neural architecture search (NAS) and hyperparameter optimization (HPO) by offering modularity, autoML capabilities, experiment management, distributed training, visualization and analysis tools, seamless integration with popular deep learning frameworks, and support for cloud services. NNI empowers researchers, data scientists, and machine learning practitioners to efficiently explore and optimize neural network architectures and hyperparameters, accelerating the development of cutting-edge AI solutions.

The Neural Network Intelligence (NNI) experiment is designed for determining optimal neural network architectures (NAS). I designing this experiment, users define a search space specifying parameters and their ranges, create model definitions and training scripts, and configure experiments. NNI automates the model selection, hyperparameter tuning, and architecture exploration, running multiple trials in parallel and managing them efficiently. With support for popular deep learning frameworks, distributed training, and comprehensive experiment management, NNI accelerates the development of high-performance neural networks. Users can monitor experiments through a web-based interface, facilitating the analysis and comparison of different architectures to identify the best-performing model.

1.1.5. Performance Measures

Performance evaluation measures for a developed system are specific criteria or metrics used to assess the system's functionality, efficiency, reliability, and overall effectiveness. These measures provide insights into how well the system meets its intended objectives and requirements. The commonly used performance measures for evaluating biometric authentication systems include the False Acceptance Ratio (FAR) and the False Rejection Ratio (FRR) (Yang et.al., 2019). Along with Recognition Accuracy, these metrics are also suitable for assessing the performance of a biometric-based user verification and identification systems. These performance measures are explained as follows:

A. False Acceptance Rate (FAR): The False Acceptance Rate (FAR) is a critical performance metric used in biometric systems, security applications, and other

verification and authentication processes. It measures the likelihood that an unauthorized user is incorrectly accepted as an authorized user by the system. FAR is essential for evaluating the security and accuracy of systems that rely on biometric or other forms of authentication. FAR is a key indicator of a system's security. A high FAR indicates that the system is more prone to security breaches, as it frequently allows unauthorized access. In contrast, a low FAR suggests that the system effectively prevents unauthorized access. The FAR can be expressed as given in (8).

$$FAR = \frac{Number\ of\ false\ acceptances}{Totla\ number\ of\ unauthorized\ attempts} \times 100\% \quad (8)$$

Where $Numberoffalseacceptances$ is the count of instances where the system incorrectly accepts unauthorized users, and $Totlanumberofunauthorizedattempts$ is the total number of attempts made by unauthorized users to access the system.

B False Rejection Rate (FRR): The False Rejection Rate (FRR) is a significant performance metric like False Acceptance Rate (FAR) used in biometric systems, security applications, and various verification and authentication processes. It measures the likelihood that an authorized user is incorrectly rejected by the system. FRR is crucial for evaluating the usability and convenience of systems that rely on biometric or other forms of authentication. FRR is a key indicator of a system's usability. A high FRR can cause frustration among legitimate users, as they are frequently denied access. Conversely, a low FRR enhances user experience by minimizing the chances of legitimate users being rejected. The FRR can be expressed as given in (9).

$$FRR = \frac{Number\ of\ false\ rejections}{Total\ number\ of\ authorized\ attempts} \times 100\% \quad (9)$$

Where Number of False Rejections is the count of instances where the system incorrectly rejects authorized users, and Total number of authorized attempts is the total number of attempts made by authorized users to access the system.

C. Recognition Accuracy: Recognition accuracy is a critical performance metric used to evaluate the effectiveness of biometric systems, machine learning models, and other pattern recognition systems. It measures the proportion of correctly identified instances out of the total number of instances tested. High recognition accuracy indicates that the system can accurately identify or classify inputs, making it reliable and effective for its intended purpose. In other words, recognition accuracy is the percentage of correct identifications or classifications made by a system out of all the attempts. It reflects how well the system performs its recognition task. It is a fundamental measure of the performance and reliability of recognition systems. High accuracy is crucial for applications where precise identification or classification is essential, such as security, medical diagnosis, and automated decision-making. The recognition accuracy can be expressed as given in (10).

$$Recognitionacc = \frac{Numberofcorrectrecognitions}{Totalnumberofattempts} \times 100\% \quad 10$$

Where Number of correct recognitions is the count of instances where the system correctly identifies or classifies the input, and Total number of attempts is the total number of recognition attempts made by the system.

1.2. Workflow Diagram

The various activities undertaken to develop the artificial neural network-based caller authentication and identification algorithm are clearly formulated and structured in the workflow diagram presented in Figure 2.

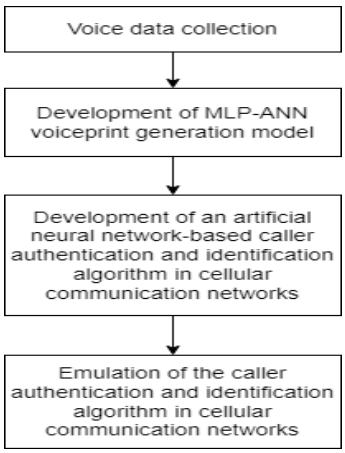


Figure 2. Workflow diagram of the algorithm development

1.2.1. Collection of Voice Data

A clean and diverse Voice data was collected from volunteer participants using Audacity software and high-quality microphones to minimize background noise. Each of the 20 volunteers contributed 50 recordings. To ensure sufficient representation for each speaker, data augmentation techniques such as time stretching and pitch shifting were applied as needed. A MATLAB script was developed to label all voice samples with their respective speaker identities, assigning labels like "1," "2," "3," and so on for the first, second, and third speakers, respectively. This organized dataset of labeled voice recordings served as the foundation for training neural network models for voiceprint generation.

1.2.2. Development of MLP-ANN-Based Voiceprint Generation Model

Mel Frequency Cepstral Coefficients (MFCC) were extracted from the voice data and used to train Multilayer Perceptron (MLP) neural networks for voiceprint generation. The extracted features served as inputs for training the MLP neural networks. The choice of MLP was based on its ability to learn intricate patterns and relationships in the input data. The underlying MLP neural

network architecture for the voiceprint generation model was optimized through neural architecture search (NAS) using the Neural Network Intelligence (NNI) toolkit. The optimized neural network architectures from the NNI experiment, as displayed in the NNI web interface, are shown in Figure 3.

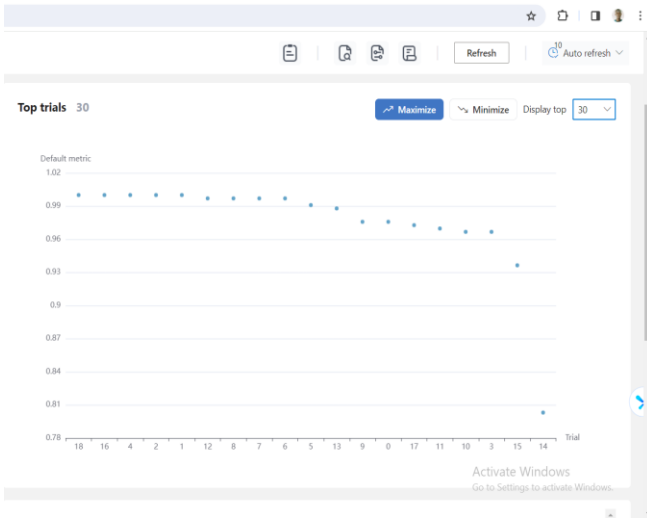


Figure 3. Neural architectures predicted from NNI Experiment

The detailed architecture of the optimized neural network, with a predicted accuracy of 100%, is shown in Figure 4.

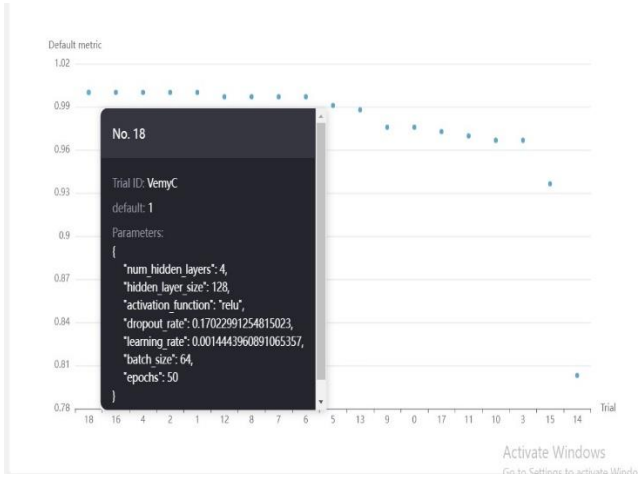
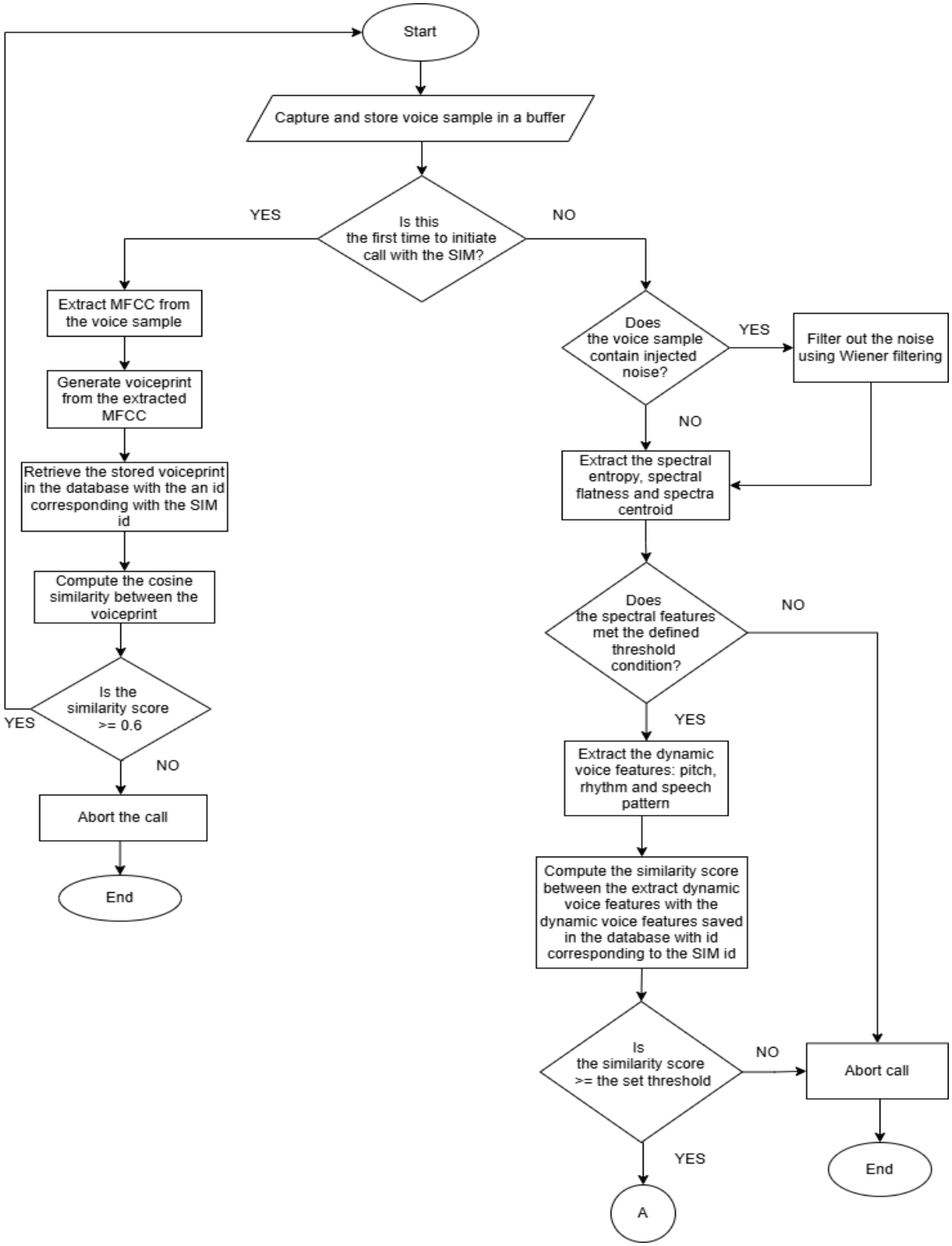


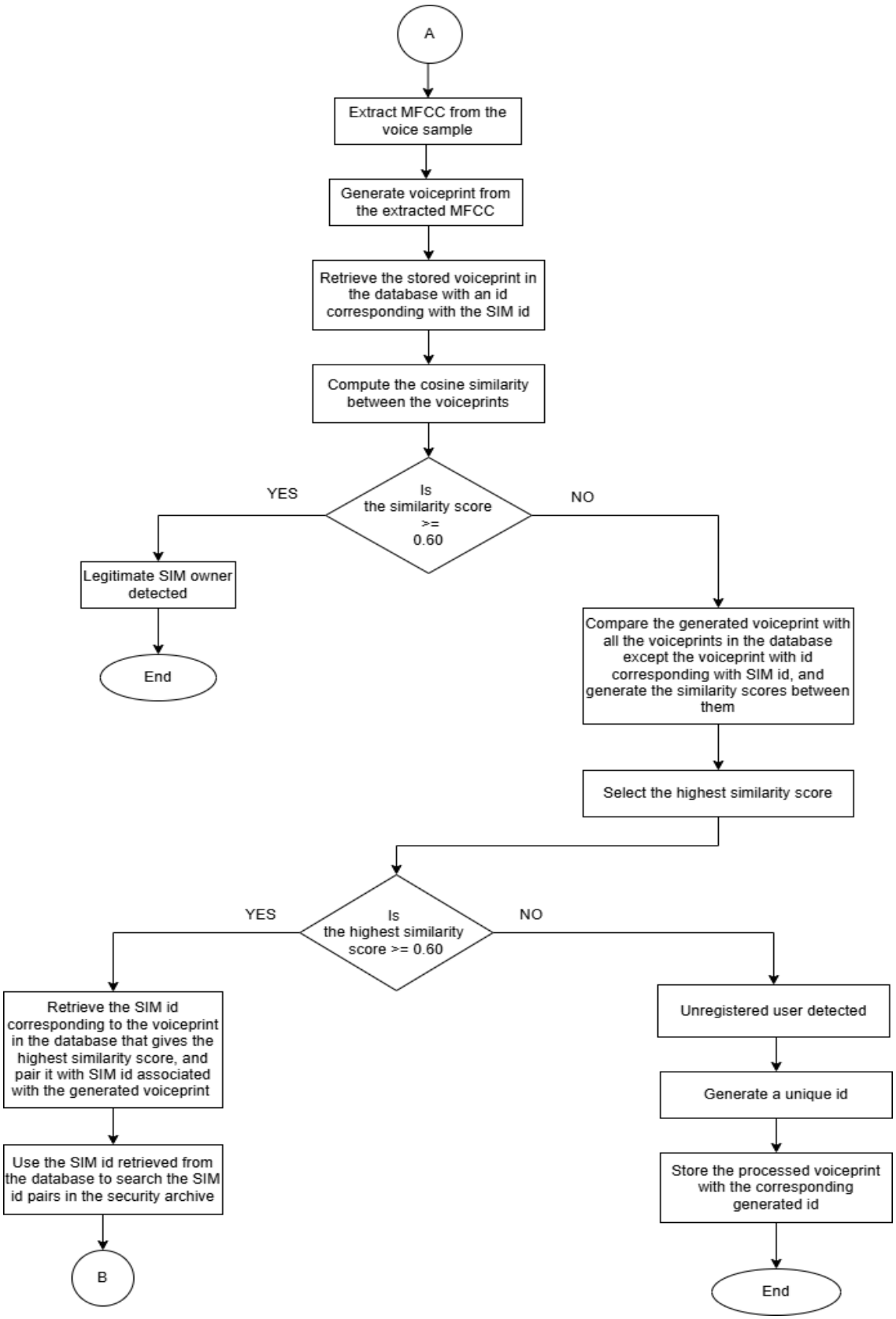
Figure 4. Details of Trial 18 MLP Neural Network Architecture

The optimized optimize neural network architecture shown in Figure 4 was subsequently trained to produce a voiceprint generation model.

1.2.3. Development of an Artificial Neural Network-Based Caller Authentication and Identification Algorithm in Cellular Communication Networks

The MLP-ANN-based voiceprint generation model was deployed as an engine to develop a neural network-based caller authentication and identification algorithm in cellular communication networks. The algorithm detects whether the SIM through which a call is initiated is a new SIM or a SIM that has been in use. When a new SIM is detected, the algorithm verifies the SIM user by processing raw voice input from the user, converting it to its equivalent Mel-Frequency Cepstral Coefficients (MFCC), and then feeding the MFCC feature vector into the MLP-ANN-based voiceprint generation model to generate a unique voice signature (voiceprint) of the user. The generated voiceprint is compared with the voiceprint of the legitimate owner of the SIM already stored in a database, producing a cosine similarity score between the voiceprints. If the similarity score exceeds 0.60, the user is verified as the owner of the SIM; otherwise, the user is declared not to be the owner of the SIM. When the SIM that has been in use is detected, the algorithm identifies the SIM user seamlessly. It periodically takes voice samples from the user, ensuring the voice sample does not contain injected noise, and confirms that the voice sample is not synthetic or mimicking a voice before converting it to an MFCC feature vector. This vector is then fed into the MLP-ANN-based voiceprint generation algorithm to generate a unique voice signature (voiceprint) of the user. The generated voiceprint is compared with all the voiceprints stored in a database, producing similarity scores between the voiceprints, in order to identify the actual identity of the user. The revealed identity of the user is stored in the security archive for accountability and traceability of the SIM user. The algorithm prevents a third party from using a SIM card for more than three consecutive times. If a user is not found in the voiceprint database, that is, an individual who has no registered SIM, the algorithm creates a unique identity and saves the voiceprint of the individual with the corresponding identity in the voiceprint database for subsequent usage. The flowchart for this algorithm is presented in Figure 5.


Start
Capture and store voice sample in a buffer
Is this the first time to initiate call with the SIM?
YES
NO
Extract MFCC from the voice sample
Generate voiceprint from the extracted MFCC
Retrieve the stored voiceprint in the database with the an id corresponding with the SIM id
Compute the cosine similarity between the voiceprint
Is the similarity score >= 0.6
YES
NO
Abort the call
End
Does the voice sample contain injected noise?
YES
Filter out the noise using Wiener filtering
NO
Extract the spectral entropy, spectral flatness and spectra centroid
Does the spectral features met the defined threshold condition?
NO
YES
Extract the dynamic voice features: pitch, rhythm and speech pattern
Compute the similarity score between the extract dynamic voice features with the dynamic voice features saved in the database with id corresponding to the SIM id
Is the similarity score >= the set threshold
NO
Abort call
End
YES
A


A
Extract MFCC from the voice sample
Generate voiceprint from the extracted MFCC
Retrieve the stored voiceprint in the database with an id corresponding with the SIM id
Compute the cosine similarity between the voiceprints
Is the similarity score >= 0.60
YES
NO
Legitimate SIM owner detected
End
Compare the generated voiceprint with all the voiceprints in the database except the voiceprint with id corresponding with SIM id, and generate the similarity scores between them
Select the highest similarity score
Is the highest similarity score >= 0.60
YES
NO
Retrieve the SIM id corresponding to the voiceprint in the database that gives the highest similarity score, and pair it with SIM id associated with the generated voiceprint
Use the SIM id retrieved from the database to search the SIM id pairs in the security archive
B
Unregistered user detected
Generate a unique id
Store the processed voiceprint with the corresponding generated id
End

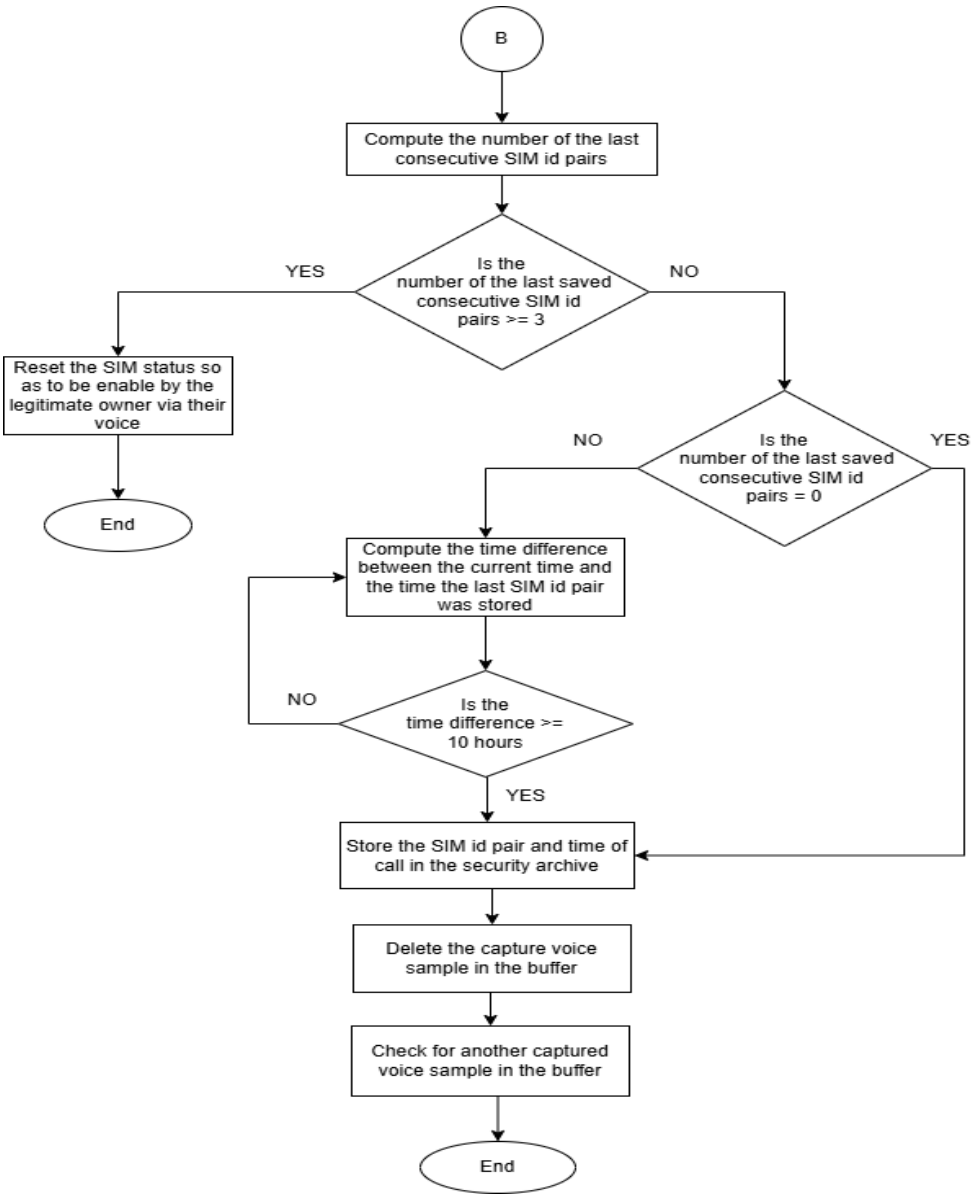


Figure 5. Flowchart of the Artificial Neural Network-Based Caller Authentication and Identification Algorithm in Cellular Communication Networks

1.2.4. Emulation of the Caller Authentication and Identification Algorithm in Cellular Communication Networks

The developed artificial neural network-based caller authentication and identification algorithm in cellular communication networks was demonstrated to showcase its application in a cellular communication system for verifying the legitimacy of SIM users via their voice and identifying third-party SIM users. The emulation consists of three units which shares the role of the infrastructures involved in call establishment in a cellular communication system. These units were linked together via the wireless LAN dongle. The block diagram of this emulation is presented in Figure 6.

Figure 6. Emulation of the artificial neural network-based caller authentication and identification algorithm in cellular communication networks

The various units in the block diagram were realized as follows:

1.2.4.1. Mobile Station (MS)

The mobile station 1 (MS1) and mobile station 2 (MS2) represent mobile phones equipped with Subscriber Identification Module (SIM) cards, facilitating their connection to a mobile network. While both units operate similarly, they are distinguished by different SIM numbers associated with each. The mobile stations were configured such that when the call/receive button of MS1 is pressed

to initiate a call establishment between MS1 and MS2, the content of a directory named SIM status is accessed. If the content of the directory is '0', it indicates that the SIM in MS1 is being used for the first time. Thus, a record icon prompt appears on MS1 for the user to say their passphrase. Once the user finishes saying their passphrase, the call/receive button is pressed again. The recorded voice sample is taken, used in conjunction with the content of the SIM status directory and the SIM Id associated with MS1 to generate a signal, which is transmitted to the MSC/AuC for further processing. The format of the generated signal is: 'ContentOfSimStatusDirectory_SimId_Audio' where ContentOfSimStatusDirectory is either 0 or 1. If the content of the SIM status directory is '1', it indicates that the SIM in MS1 has been in use. Thus, a record icon will not be prompted to capture the user's voice sample. Here, the content of the directory named Proximity sensor status will be checked. When the content of the directory becomes '1', the voice sample of the SIM user will be seamlessly captured periodically. Each periodic voice sample will be used in conjunction with the content of the Proximity sensor status directory and the SIM ID associated with MS1 to generate a signal, which is transmitted to the MSC/AuC for further processing. The format of this signal is: 'ContentOfProximitySensorStatusDirectory_SimId_Audio', where ContentOfProximitySensorStatusDirectory is either 0 or 1. The value '1' indicates that the user has started speaking on the phone, and the value '0' indicates that the user has not started speaking on the phone. The same process occurs when MS2's call/receive button is pressed to initiate a call establishment between MS2 and MS1.

The emulated view of the mobile station, prompting users voice samples to be captured when the content of the SIM status directory is '0', is depicted in Figure 7. Figures 7 (a) and (b) show the first and second times the call/receive button is pressed, respectively.

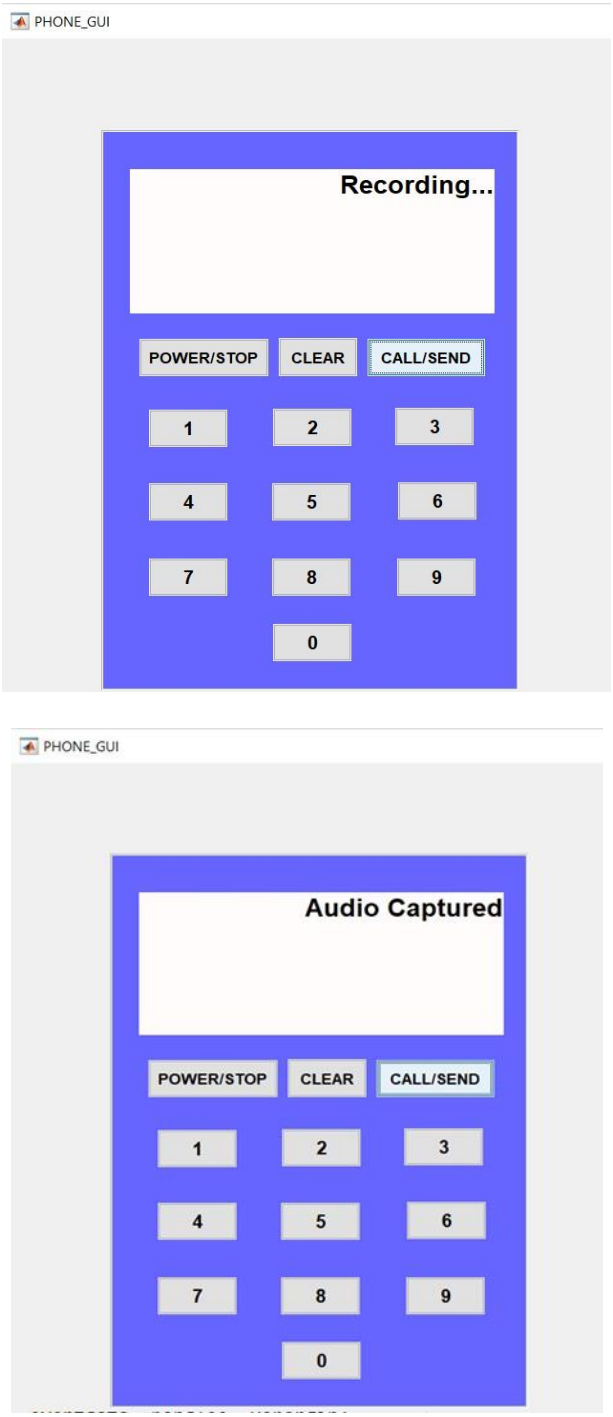


Figure 7. Emulated view of the mobile station capturing user voice when the SIM card resident in the mobile station is being used for the first time.

1.2.4.2. Mobile Switching Centre (MSC)

The computer system representing the mobile switching centre (MSC) was configured for the registration of SIM users' voiceprints, hosting security archive which holds the revealed identity of the SIM users, and for performing verification and identification of SIM users. These functionalities were achieved as follows:

A. Registration of SIM Users Voiceprint

The 'Register Voiceprint for New User' button was created on the mobile switching centre GUI (Graphical User Interface). The button was configured so that when pressed, it opens an interface containing a keypad used for

entering the user's ID (SIM number). Upon pressing the 'Take Voice Print' key on the interface, the user is prompted for voice capturing. The captured voice is then processed using the developed MLP-ANN voiceprint generation to generate the user's voiceprint, which is then appended with the user's ID in the database.

The emulated view of the registration of new users in the voiceprints database is as shown in Figure 8. Figure 8 (a) and (b) depict the registration button on the MSC GUI and the registration interface initiated by pressing the registration button. Figure 8(c) and (d) depict the status on the interface when the 'Take_Voice_Print' and 'Stop' buttons are pressed, respectively.

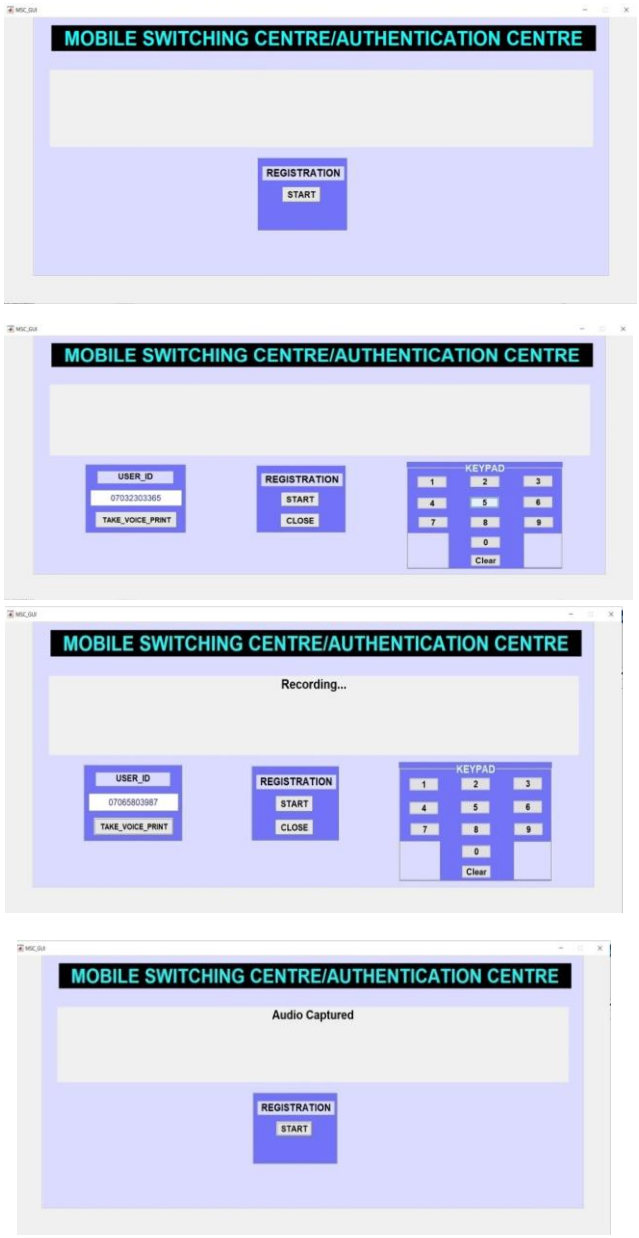


Figure 8. Emulation view for the registration of new users in the voiceprint database

B. Security Archive

The security archive serves as the database designed for storing the revealed identities of third-party SIM users. The reveal of the actual SIM ID of the third-party SIM user from the voiceprint database is paired with the ID of the SIM being used, along with the time of use. This information pair is stored within the security archive to ensure accountability and traceability of the SIM user. The format in which the information is paired and stored is presented as follows: 'SimIdOfSIMBeingUsed_SimIdOfPartyUser_CallTime'. Figure 9 illustrates the content of the developed security archive.

Figure 9. Information Pair Stored in the Security Archive

C. Verification and Identification of Callers in the Emulated Cellular Communication Network

The developed caller verification and identification algorithm in cellular communication network was deployed at the mobile switching center in the emulation. The mobile switching center (MSC) in the emulation setup was configured to receive signals from the mobile station (MS). The signal received in the MSC from the MS resides in a particular directory at the MSC named the signal directory where the MSC continually checks for signals. The signal in this directory is removed to give room for the subsequent incoming signal. The removed signal is decoded into its three constituent elements: the first, second, and third elements represent the signal status, SIM id, and voice sample respectively. The value of the signal status embedded in the signal determines how the caller verification and identification algorithm will run. If the value of the signal status is '0', it indicates that the SIM is being used for the first time to initiate a call, and if the value of the signal status is '1', it indicates that the SIM has been in use. The deployed caller verification and identification algorithm works in line with the algorithm's flowchart in Figure 5.

2. Deployment and Implementation of the Developed Caller Authentication and Identification Algorithm in Cellular Networks.

The developed artificial neural network-based caller authentication and identification algorithm in cellular communication networks will be integrated by deploying it at the Mobile Switching Center (MSC), the unit in the cellular communication system that ensures smooth call establishment and routing. Network operators will modify the signaling protocols so that when a caller initiates a conversation, their voice is recorded on the mobile station

for a duration of 10 seconds and sent to the MSC for caller authentication when the SIM card is detected to be in use for the first time after registration, thus deterring perpetrators from using black-market SIMs in the network. If the SIM is detected to have been used previously, the caller's voice is recorded for 10 seconds at regular 10-second intervals and sent to the MSC for SIM user identification, thereby preventing perpetrators from using third-party SIMs to remain anonymous.

To efficiently collect and manage large-scale voice data, the existing SIM registration database will be updated by network operators to store users' voiceprints. During SIM registration, each user's voiceprint is captured and generated by a trained and deployed artificial neural network (ANN) model for voiceprint generation. This database of voiceprints will serve as the reference for caller authentication and identification. The system is designed to update each user's voiceprint in the database whenever their voiceprint identity is successfully verified up to set threshold for user verification, allowing the voiceprint to reflect changes in the user's voice over time. This process ensures that aging or other variations in voice do not affect the accuracy of the algorithm. Data augmentation or regularization techniques were also used to improve robustness against varied conditions associated with voice. To protect the voiceprint database from unauthorized access and potential manipulation, it will be secured by the same robust security measures already in place for the SIM registration database. These measures may include encryption, access control, and regular monitoring to safeguard sensitive voiceprint data and prevent unauthorized access, tampering, or breaches.

The algorithm preprocesses each input voice sample to detect whether it contains injected noise, which may be deliberately added to bypass the algorithm, or if it is a mimicked or synthetic voice. If any of these conditions are detected, the algorithm treats the call as suspicious and aborts it. The developed algorithm has a fast voiceprint generation rate. The computation time for user authentication (verification) is very short, as it is one-to-one (1:1) verification. For user identification, the computation time varies depending on the size of the database used in the identification process; however, this computation time does not interrupt communication as it runs in the background during call establishment. The performance evaluation of the algorithm was conducted using 20 volunteer participants, a reasonable sample size for establishing proof of concept. The algorithm was tested with the volunteer voice samples, including those containing injected noise and mimicked voices. This number allows for an initial assessment of the algorithm's core functionalities.

The comparison of the cryptographic challenge-response-based authentication algorithm in cellular communication networks with the proposed algorithm shows that the existing algorithm verifies the legitimacy of the SIM and network (Alezabi et. a., 2014; Tsay and Mjolsnes, 2012; Nakarmi, 2021), whereas the proposed algorithm authenticates the SIM user and also reveals the true identity of a third-party SIM user, thus providing a security edge over the existing algorithm, as it cannot verify or identify third-party SIM users. This allows perpetrators to use black market SIM cards or third-party SIMs while remaining anonymous.

The SIM registration approach ensures that network users' identities are traceable (GSMA, (2016), thereby preventing perpetrators from using their SIMs to engage in fraudulent activities while remaining anonymous. However, this security approach is circumvented by perpetrators through the use of black market SIMs or third-party SIMs, whereas the proposed algorithm deters such usage.

The True Caller application reveals the true identity of the SIM owner of the caller (Truecaller, 2021; Indriyani and Aprinia, 2022), thus deterring perpetrators from using their own SIM while hiding their identity associated with the SIM. This security approach is circumvented by perpetrators through the use of a third-party SIM, consequently implicating an innocent owner of the SIM. In contrast, the proposed algorithm reveals the true identity of a third-party SIM user, thus providing security over the True Caller application.

The proposed algorithm provides a security edge over biometric and AI-based approaches deployed for securing mobile devices (Patel et. al., 2016; Jung and Hong, 2015; Pocovnicu , 2009; Chintalapati et. al. 2023). These security approaches ensure that only the legitimate owner of a mobile device can access or grant access to the device. Perpetrators circumvent these security measures by coercing the rightful owner to grant them access to the mobile device, whereas the proposed algorithm reveals the true identity of the perpetrator using a third-party mobile device (SIM).

The comparison of the proposed algorithm with the security enhancement of voice over Internet Protocol using the speaker recognition technique put forward by Ibrahim et al. (2012) shows that the securing approach involves verifying a caller through their passphrase before call establishment, thus ensuring the legitimacy of the caller (user). However, perpetrators can coerce a legitimate caller to establish a call and hand over to them in order to maintain their anonymity. The proposed algorithm provides a security edge over the security provided by Ibrahim et al. (2012) by revealing the true identity of a perpetrator using a third-party SIM card.

The proposed algorithm does not require any action from the user except during SIM registration, as it based on voice (speaking) which is natural and does not require the user to position themselves in a certain way or touch a device, making it straightforward and convenient. This

algorithm exhibits ease of use and consequently has a high potential for user acceptance and adoption of the new system.

3. Results and Discussion

The false acceptance ratio, false rejection ratio and accuracy were the performance metrics used to evaluate the performance of the ANN-based caller authentication and identification algorithm in cellular communication networks. Table 2 shows results of the performance evaluation of the algorithm.

Table 2. Performance evaluation of the ANN-based caller authentication and identification algorithm in cellular communication networks

Performance evaluation of the algorithm for authenticating first-time users of a new SIM					
S/N	Test Scenario	Total number of attempts	Number of users correctly granted access	Number of users correctly denied access	Performance Metric
1	Authentication of 10 legitimate and 10 illegitimate SIM owners initiating call with new registered SIM	20	10	10	Accuracy (100%)
2	Authentication of 20 illegitimate SIM owners initiating call with new registered SIM	20	0	20	FAR (0%)
3	Authentication of 20 legitimate SIM owners initiating call with new registered SIM	20	20	0	FRR (0%)
Performance evaluation of the algorithm for identifying third-party SIM users					
S/N	Test Scenario	Total number of attempts	Number of correct identified third-party registered SIM users	Number of correct recognized third-party unregistered SIM users	Performance Metric
1	Identification of 10 registered and 10 unregistered SIM users making call with third-party registered SIM	20	10	10	Accuracy (100%)
2	Identification of 20 registered SIM users making call with third-party registered SIM	20	20	0	FRR (0%)
3	Identification of 10 unregistered SIM users making call with third-party registered SIM	10	0	10	FAR (0%)

In table 2, 20 SIM users comprised of legitimate and illegitimate users who made a first call with a registered SIM were subjected to the developed algorithm for verifying SIM user legitimacy. The algorithm successfully identified legitimate SIM users (users assigned as the SIM owner) from illegitimate SIM users. This depicts the algorithm's accuracy in granting network services only to legitimate users in their first attempt of making a call with a new registered SIM.

The algorithm's ability to incorrectly grant network access to an illegitimate SIM user during their first call attempt with a new registered SIM was tested. 20 illegitimate SIM users tried to initiate a first call with a new registered SIM, and the algorithm denied them network access. This demonstrates the algorithm's ability to avoid incorrectly accepting an illegitimate user as legitimate. None of the tested illegitimate users were falsely granted access as legitimate SIM users

The algorithm was also tested for its ability to incorrectly reject a legitimate SIM user trying to initiate a first call with their newly registered SIM. 20 legitimate SIM users made repeated attempts to initiate a first call with their new registered SIMs; the algorithm successfully verified them for network access. This demonstrates the algorithm's reliability in granting access to authorized SIM users in their first attempt to call with a new registered SIM. The algorithm showed a False Rejection Ratio (FRR) of 0% among the tested legitimate users.

The identification of subscribed SIM users by the algorithm was tested. 20 SIM users, comprised of subscribed and unsubscribed users, were involved. The algorithm was able to accurately identify all the subscribed SIM users through their voice samples during a call; it also identified all the unsubscribed SIM users. This demonstrates the algorithm's ability to identify subscribed SIM users when they make a call. An accuracy of 100% was achieved in the conducted tests.

The algorithm's tendency to incorrectly identify a subscribed SIM user was tested. 20 subscribed SIM users were tested with the algorithm to see how well it could correctly identify them. The algorithm correctly identified all the SIM users subjected to the test. This shows that the False Rejection Ratio (FRR) of the algorithm was 0% among the tested subscribed SIM users.

The algorithm's ability to avoid incorrectly identifying unsubscribed SIM users was tested. Ten users were

assigned SIM numbers (subscribed), while ten others were not (unsubscribed). The unsubscribed SIM users were then tested with the algorithm to see how well it could identify them correctly. The algorithm correctly identified all the unsubscribed SIM users, demonstrating a False Acceptance Ratio (FAR) of 0% for the tests conducted.

4. Conclusion

The study successfully developed an artificial neural network-based caller authentication and identification algorithm for cellular communication networks, utilizing voice data from 20 volunteers and 2,118 speakers from the Mozilla Common Voice database. Voice samples were preprocessed, and relevant features were extracted using Mel-frequency cepstral coefficients (MFCC), then used to train an optimized multilayer perceptron (MLP) neural network for voiceprint generation. The algorithm demonstrated promising accuracy when evaluated in an emulated cellular network system, indicating its potential to deter black market SIM usage and facilitate reliable caller identification.

While the results are encouraging, future research could explore expanding the model to handle larger, more diverse datasets, including voices affected by illness or environmental noise, to enhance robustness. Further development could also refine the algorithm for use in broader applications, such as secure online transactions, remote banking, and IoT device authentication. These enhancements would strengthen the model's effectiveness across diverse applications, ensuring broader impact in securing communication systems and enhancing user trust in authentication technologies.

References

[1] Alezabi, K., A., Hashim, F., Hashim, S.J., & Ali, B.M. (2014). An Efficient Authentication and Key Agreement Protocol for 4G (LTE) Networks. 2014 IEEE Region 10 Symposium. Pp. 502 – 507. https://doi.org/10.1109/TENCONSpring.2014.6863085

[2] Asda Mobile. (2022). *How to unlock your phone with a PUK code*. *https://mobile.asda.com/scoop/how-to-unlock-your-phone-with-a-puk-code*

[3] Botalb, A., Moinuddin, M., Al-Saggaf, U. M., & Ali, S. S. A. (2018). Contrasting Convolutional Neural Network (CNN) with Multi-Layer Perceptron (MLP) for Big Data Analysis. 2018 International Conference on Intelligent and Advanced System (ICIAS). doi:10.1109/icias.2018.8540626

[4] Chintalapati, P.V., Babu, G.R., Sree, P.K., Kode, S.K., Kumar, G.S. (2023). Usage of AI Techniques for Cyberthreat Security System in Android Mobile Devices. In: Hassanien, A.E., Castillo, O., Anand, S., Jaiswal, A. (eds) International Conference on Innovative Computing and Communications. ICICC 2023. Lecture Notes in Networks and Systems, vol 703. Springer, Singapore. https://doi.org/10.1007/978-981-99-3315-0_33

[5] Douglas, K., & Mumbi, C. (2018). A Survey of Android Mobile Phone Authentication Schemes. Mobile Networks and Applications. doi:10.1007/s11036-018-1099-7

[6] Fan, C. I., Shih, Y. T., Huang, J. J. & Chiu, W. R. (2021). Cross-Network-Slice Authentication Scheme for the 5th Generation Mobile Communication System. IEEE Transactions on Network and Service Management, 18(1). Pp. 701-712. doi:10.1109/TNSM.2021.3052208.

[7] GSMA. (2016). *Mandatory registration of prepaid SIM cards: Addressing challenges throughbestpractice*. *https://www.gsma.com/publicpolicy/wp*content/uploads/2016/04/GSMA2016_Report_MandatoryRegistrationOfPrepaidSIMCards.pdf

[8] Ibrahim, A. J., &Jauro, U. A. (2021). Bio-metric encryption of data using voice recognition. Autom. Control Intell. Syst, 9 (3).

[9] Ibrahim, Q., Abdulghani, N. (2012). Security enhancement of voice over Internet protocol using speaker recognition technique. IET Communications. 6 (6), pp. 604 – 612.

[10] Indriyani, & Aprinia, D. (2022). Role of Truecaller application in preventing phone call and text message scams. *Jurnal Mantik*, *6*(2), 1475-1483. *https://doi.org/10.35335/mantik.v6i2.2533*

[11] Jung, E., & Hong, K. (2015). Biometric verification based on facial profile images for mobile security. Journal of Systems and Information Technology, 17(1), 91–100. doi:10.1108/jsit-03-2014-0020

[12] Kazmi, S. H. A., Hassan, R., Qamar, F., Nisar, K. & Ibrahim, A. A. A. (2023). Security Concepts in Emerging 6G Communication: Threats, Countermeasures, Authentication Techniques and Research Directions. MDPI, 15(6). https://doi.org/10.3390/sym15061147

[13] Khan, A. S., Javed, Y., Saqib, R. M., Ahmad, Z., Abdullah, J., Zen, K., Abbasi, I. A., Khan, N. A. (2022). Lightweight Multifactor Authentication Scheme for NextGen Cellular Networks. IEEE Access, vol. 10. doi: 10.1109/ACCESS.2022.3159686

[14] Luo, Q. & Zhang, Z. (2023). The Secure Data Transmission Method of a Cellular Communication Network Based on the Asymmetric Encryption Algorithm. Journal of Communications, 18(2). Pp 82 – 88. doi:10.12720/jcm.18.2.82-88

[15] Miesle, P. (2023, September). What is cosine similarity: A comprehensive guide. *Datastax*. *https://www.datastax.com/guides/what-is-cosine-similarity*

[16] Mohammed, A. (2023). How to enable SIM lock on your Android phone: Prevent malicious SIM card use by adding an extra layer of security. *Android Police*. *https://www.androidpolice.com/enable- sim-lock-android-phone-protection/*

[17] Nakarmi, P. K. (2021). *Cheatsheets for authentication and key agreement in 2G, 3G, 4G, and 5G*. Ericsson. *https://arxiv.org/pdf/2107.07416v1.pdf*

[18] Okonji, E. (2022, March 04). Terrorism: No Unregistered SIM Cards on Operators' Network. Telcos Insist. Thisday. *https://www.thisdaylive.com/index.php/2022/04/03/terrorism-no-* unregistered-sim-cards-on-operators-network-telcos-insist/

[19] Patel K., Han H. & Jain A.K. (2016). Secure Face Unlock: Spoof Detection on Smartphones. IEEE Transactions on Information Forensics and Security, 11(10), pp. 1 – 16. *https://doi.org/10.1109/TIFS.2016.2578288*

[20] Pocovnicu A. (2009). Biometric Security for Cell Phones. Informatica Economică , 13(1), pp. 57 – 63.

[21] Truecaller. (2021). About Truecaller. *https://www.truecaller.com/*

[22] Tsay, J. K., & Mjolsnes, S. F. (2012). A vulnerability in the UMTS and LTE authentication and key agreement protocols. In *Advances in Information Security.* Springer. Vol. 55, pp. 77–90. *https://doi.org/10.1007/978-3-642-33704-8_6*

[23] Yang, W., Wang, S., Hu, J., Zheng, G., & Valli, C. (2019). Security and accuracy of fingerprint-based biometrics: A review. *Symmetry*, *11*(2), 141–160. *https://doi.org/10.3390/sym11020141*

[24] Zukarnain, Z. A., Muneer, A., & Aziz, M. K. (2022). Authentication Securing Methods for Mobile Identity: Issues, Solutions and Challenges. MDPI, 14(4). *https://doi.org/10.3390/sym14040821*.