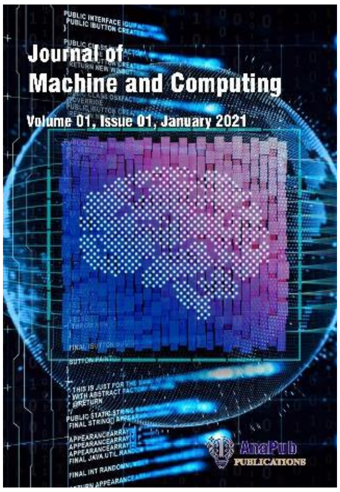
Journal Pre-proof

Securing Voice Software Applications Using 5G + WSN + AI-Driven Privacy Preservation Protocols



Hayder M A Ghanimi, Swaroopa K, Amit Mishra, Anusha Papasani, Kolluru Suresh Babu and Vivekanandhan Vijayarangan

DOI: 10.53759/7669/jmc202505142

Reference: JMC202505142

Journal: Journal of Machine and Computing.

Received 30 March 2025

Revised from 26 April 2025

Accepted 16 June 2025

Please cite this article as: Hayder M A Ghanimi, Swaroopa K, Amit Mishra, Anusha Papasani, Kolluru Suresh Babu and Vivekanandhan Vijayarangan, "Securing Voice Software Applications Using 5G + WSN + AI-Driven Privacy Preservation Protocols", Journal of Machine and Computing. (2025). Doi: https:// doi.org/10.53759/7669/jmc202505142.

Securing Voice Software Applications Using 5G + WSN + AI-Driven Privacy Preservation Protocols

Hayder M. A. Ghanimi[1,2], K. Swaroopa[3], Amit Mishra[3], Anusha Papasani[4], Kolluru Suresh Babu[5], Vivekanandhan Vijayarangan[6*]

[1]Department of Information Technology, College of Science, University of Warith Al-Anbiyaa, Karbala, 56001, Iraq.

[2]Department of Computer Science, College of Computer Science and Information Technology, University of Kerbala, Karbala, 56001, Iraq. *Email: hayder.alghanami@uowa.edu.iq*

[3]Department of Computer Science and Engineering (Data Science), Aditya Institute of Technology and Management, Tekkali, Srikakulam, Andhra Pradesh, 532201, India. *Email: drksp.cse@gmail.com*

[4]Department of Computer Science and Applications, Dr. Vishwanath Karad MITWPU, Pune, Maharashtra 411038, India. *Email: i.amitmishra@gmail.com*

[5]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram 522502, Andhra Pradesh, India. *Email: anoosha.papasani@gmail.com*

[6]Department of Computer Science and Engineering, Vasireddy Venkatadri Institute of Technology, Guntur, Andhra Pradesh, 522508, India. *Email: kollurusuresh@gmail.com*

[7]Department of Computer Science and Engineering, Malla Reddy College of Engineering, Secunderabad, 500100, India.

**Corresponding Author Email: cevivek7677@gmail.com*

Abstract

The reality-based, dynamic, and context-aware user experiences provided by voice software applications have contributed to their common acceptance. But, problems with data privacy and computer performance are challenges. In order to process voice data reliably, the present research proposes a secure integrated model of 5G-Wireless Sensor Networks with Artificial Intelligence (5G + WSN + AI) to apply privacy preservation protocols. To train decentralized models, the model used Federated Learning (FL). To prevent unauthorized inference, it deployed Secure Multi-Party Computation (SMPC). In the end, to secure sensitive data, it applied adaptive encryption methods. Word Error Rate (WER), Feature Extraction Accuracy (FEA), End-to-End Delay (EED), Network Throughput (NT), Packet Loss Rate (PLR), and Encryption Overhead (EO) represent several of the key performance measures that the model is considered superior to conventional networks such as SVPS, BDPS, GACS, and cloud-based centralized models. Additionally, it proved that next-generation Voice Learning Systems (VLS) are reliable, leveraging AI + 5G setup and maintaining robustness against privacy breaches in real-world asymmetric scenarios.

Keywords: Wireless Sensor Networks, Artificial Intelligence, 5G, Voice Software Applications, Security, Federated Learning.

1. Introduction

The method learners use to communicate with online material has been altered by the increasing adoption of voice-activated software applications in the field of virtual information technology, particularly within Voice Learning Systems (VLS) [1-4]. The primary objective of such technologies is to enhance engagement, knowledge, and spoken language by modifying voice data. However, there are also significant privacy and efficiency concerns with voice data, considering its increasing importance [5-6]. Even more so currently, when attackers and data thefts have become more intelligent [7-9], it is vital to ensure the privacy and security of this data. For real-time virtualized software, it is crucial to have an accurate model that integrates high security with low End-to-End Delay (EED) [10].

5G-Wireless Sensor Networks (WSN) provide novel chances to improve the functioning of VLS through improved connectivity, low EED, and high NT [11, 12]. The security risks in conventional networks can be solved by combining 5G with privacy protocols based on Artificial Intelligence (AI) [13–5]. Edge computing in 5G enables localized data processing, thereby minimizing Response Times (RT) and reducing the risk of security attacks [16-19]. This is in contrast to centralized cloud-based models, which experience high EED and security problems.

Technologies such as Cloud-Based Centralized Models (CBCM) and Standard Voice Processing Systems (SVPS) currently face several problems [20]. SVPS is vulnerable to data breaches because it fails to implement security models. The centralized processing of data in CBCM, on the other hand, causes conjunction and EED. While other decisions, such as GACS and the Basic Differential Privacy System (BDPS), provide precise improvements, they fail to provide complete security. While GACS uses predictable authentication methods that attackers can access, BDPS employs noise to ensure data privacy, which typically results in reduced accuracy.

The research presented here recommends an innovative model to secure VLS by proposing an integrated model of 5G-Wireless Sensor Networks with Artificial Intelligence (5G + WSN + AI). Using Federated Learning (FL), Secure Multi-Party Computation (SMPC), and dynamic encryption methods, the proposed approach addresses significant problems with accuracy, EED, and privacy. FL ensures that voice data is sustained on local devices and complete training, which reduces the risk of attacks. SMPC enables secure group computations without compromising private data, and adaptive encryption adjusts to evolving types of attacks in real-time.

Word Error Rate (WER), Feature Extraction Accuracy (FEA), End-to-End Delay (EED), Network Throughput (NT), Packet Loss Ratio (PLR), and Encryption Overhead (EO) represent a few of the most significant metrics evaluated to find the 5G + WSN + AI's efficiency. The performance of this 5G + WSN + AI has been verified through analyses with baseline models (SVPS, CBCM, BDPS, and GACS). With significantly better outcomes than these baselines, the recommended model provides lower WER, reduced EED, higher NT, and minimal PLR, as indicated by the results. Additionally, even as the size of malicious attacks increases, the model maintains minimal Privacy Leakage Rates (PrLR).

The rest of the paper is organized as follows: Section 2 presents the model and methodology, Section 3 presents the experimental set-up, Section 4 presents the results and analysis, and Section 5 concludes the paper.

2. Proposed Model

2.1. System Overview

Figure 1 presents the recommended model, which integrates 5G + WSN + AI into VLS applications. Securing voice data while maintaining virtual performance in VLS is a significant challenge, and this detailed model addresses it effectively. The network provides adaptive confidentiality, improved network operation, and continuous real-time communication by integrating security systems at every level.

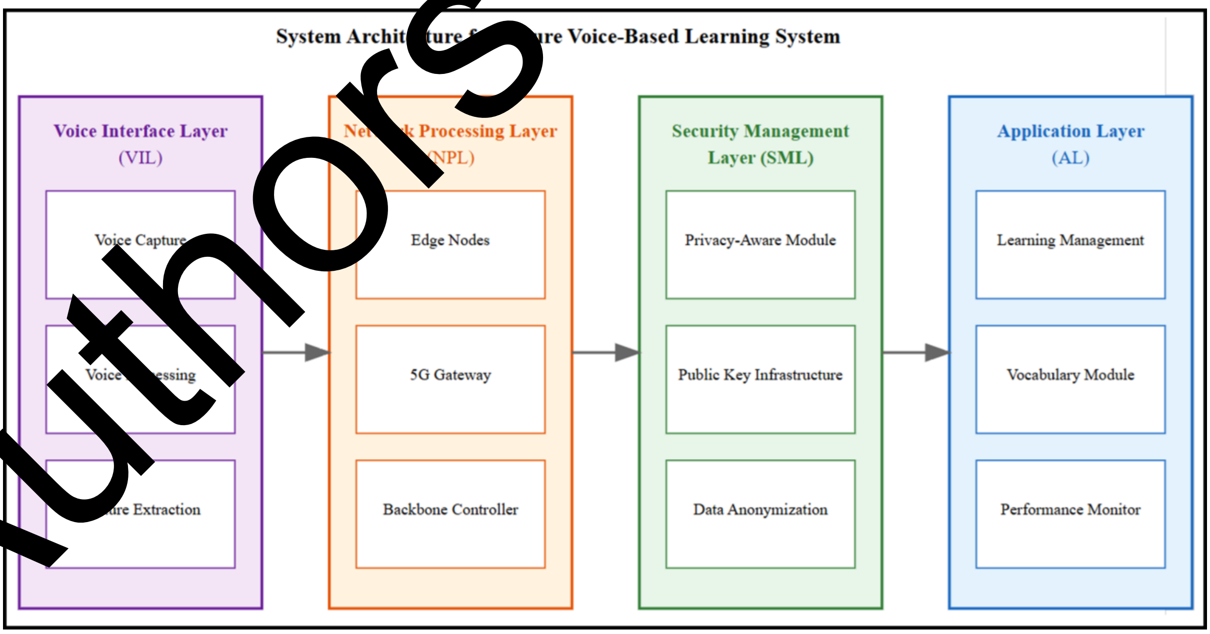


Figure 1: The proposed 5G + WSN + AI architecture

2.1.1 Network Components

The Application Layer (AL), Voice Interface Layer (VIL), Network Processing Layer (NPL), and Security Management Layer (SML) form a set of interlinked layers that comprise the system's hierarchical design. Processing data, implementing security measures, and providing higher education are tasks that these linked layers perform. The VIL performs any communications using voice and first processing at the basic level of responsibility. The module's within the VIL, as V_i, voice capture modules (C_v), processing units (P_v), and Feature Extraction (FE) as (F_v). The voice input I_v from the user's experience, the initial preprocessing involves converting raw audio signals into feature sets appropriate for study. This layer ensures real-time openness and high-fidelity data capture.

The NPL as N_iIt combines the operations of 5G + WSN + AI. This layer comprises distributed edge nodes. (E_n) for localized processing, a central gateway (G_c) for managing data traffic and a backbone controller (B_c) for coordinating edge nodes. The NPL leverages the high speed and low EED of 5G + WSN + AI to optimize data routing (R_d) and maintain network constancy (S_n) under variable load settings.

The connection between edge nodes and the gateway can be stated as Eq. (1)

$$N_i = \{E_n, G_c, B_c\}. \quad (1)$$

Where,

- SML as S_i, forms the core of the security model.
- This layer integrates a privacy-aware module (P_m), public key set-up (K_p)
- Data anonymization engine (A_d)
- The SML enforces security policies (S_p)
- adapts to emerging attacks (T_e) using a dedicated privacy implementation module (E_p).

The measured symbol of the SML can be summarized as Eq. (2)

$$S_i = \{P_m, K_p, A_d, E_p\} \quad (2)$$

Where,

- AL as A_i manages the educational features of the system.
- This layer comprises the VLS (L_m)
- vocabulary modules (V_m)
- performance monitoring tools (M_p).

The AL delivers personalized learning experiences while ensuring security compliance. The learning modules can be expressed as Eq. (3)

$$A_i = \{L_m, V_m, M_p\} \quad (3)$$

2.1.2 Data Flow Design

The data flow within the system follows a structured sequence that balances processing efficiency and security implementation. Voice input $'I_v'$ is taken at the edge nodes (E_n), where initial preprocessing and FE as (F_v) ensue. This distributed network decreases the load on the primary network by handling initial processing nearby. The pre-processed data, signified by $D_{p,}$ experiences initial security transmission (S_s) at the edge level to filter out anomalies and probable attacks.

The data is then routed by the 5G + WSN as (N_i),

Where,

- Edge nodes coordinate to maintain data integrity (I_d)
- Optimize resource allocation (R_a).
- The central gateway (G_c) manages the overall data flow while the backbone controller (B_c) performs adaptive load balancing (L_b) to mitigate EED and network congestion.

The data flow can be expressed as Eq. (4)

$$D_f = (E_n \rightarrow G_c \rightarrow B_c) \times L_b \tag{4}$$

Where,

- SML (S_i)
- Encryption (E_c)
- Anonymization (A_d) are applied to protect privacy.
- Real-time attack detection (T_d) ensures the data remains secure in transmission.
- The secure data packet is $'D_s'$, is then delivered to the AL (A_i) for learning processing and feedback generation.
- The overall data flow maintains a continuous balance between efficiency, security, and learning performance.

2.1.3. Security Model Integration

The security paradigm has been fully integrated with the network, demonstrating that security measures are essential, not mandatory. Security measures are detailed and efficient because of this integration, which reduces runtime EED. The security model (S_f) incorporates multi-layer authentication (A_m), dynamic encryption protocols (E_d), and advanced privacy controls (P_c). Multi-layer authentication mechanisms (A_m) combine Role-Based Access Control (RBAC) and Multi-Factor Authentication (MFA) to ensure that only authorized users access sensitive data. The authentication method can be expressed as Eq. (5).

$$A_m = \text{RBAC} + \text{MFA} \tag{5}$$

Where,

- Dynamic encryption protocols (E_d) adapt to real-time attack levels (T_r) by adjusting encryption strength.
- *e.g.,* encryption (E_c) using AES-256 ensures data security during transmission and storage.

The encryption process is defined as Eq. (6)

$$E_d = \text{AES} - 256 \times T_r \tag{6}$$

Where,

- Privacy controls (P_c) leverage automated policy enforcement (P_e)
- SMPC to protect user data.
- Anonymization (A_d) ensure that voice data retains its learning value while securing user identities.

The privacy control equation can be summarized as Eq. (7).

$$P_c = P_e + \text{SMPC} + A_d \tag{7}$$

Where,

- The model employs adaptive resource allocation (R_a)
- Continuous monitoring (M_c).
- Security parameters dynamically adjust based on resource availability (R_v)
- Detected attacks (T_d), ensuring the model remains resilient without compromising efficiency.

This adaptive method represents a significant improvement over traditional static security measures.

2.2. Voice Processing Module

The Voice Processing Module (VPM) plays a key role in capturing, processing, anonymizing, and expressing FE from voice input within the model. This module ensures that the integrity, accuracy, and security of voice data are maintained throughout the VLS. The VPM operates within the VIL and interfaces closely with the NPL and SML, providing a seamless and secure voice-based interaction experience.

2.2.1 Voice Capture Mechanisms

The voice capture mechanisms are designed to ensure the high-fidelity acquisition of the user's voice input in several environmental backgrounds. These mechanisms use advanced hardware and signal processing methods to minimize noise, distortion, and EED, capturing clear and accurate voice signals for further processing.

1 **Microphone Arrays (M_a):** The network uses directional and omnidirectional microphone arrays intentionally positioned to capture voice input accurately while mitigating background noise. Each microphone element in the array captures voice signals, and beamforming methods combine these signals to focus on the primary audio system. The mathematical symbol of the beamformed signal 'V_b' as Eq. (8)

$$V_b(t) = \sum_{i=1}^{N} w_i \cdot M_i(t - \tau_i) \tag{8}$$

Where,

- M_i → The signal from the i^{th} microphone
- w_i → The weight assigned to the i^{th} microphone
- τ_i → The time delay applied to align the signals.

2 **Noise Reduction Filters (N_r) :** To improve the precision of the captured voice signal, noise reduction filters such as spectral subtraction and Wiener filters are applied. Spectral subtraction computes the noise spectrum and subtracts it from the captured signal, as shown in Eq. (9).

$$V_c(f) = V_b(f) - N(f) \tag{9}$$

Where,

- $V_b(f)$ → The beamformed signal in the frequency domain
- $N(f)$ → The estimated noise spectrum.

3 **Automatic Gain Control (AGC):** AGC dynamically adjusts the amplitude of the incoming voice signal to ensure consistent loudness regardless of the user's distance from the microphone. The AGC function $G(t)$ modifies the signal amplitude $A(t)$ as, Eq. (10)

$$G(t) = \frac{A_{\text{target}}}{A(t)} \cdot V_c(t) \tag{10}$$

Where,

- A_{target} → The desired amplitude level.

4 **Latency Reduction Methods:** To meet real-time processing requirements, EED is minimized using the use of hardware-based signal buffering and parallel processing at the edge nodes. The sum of capture EED is 'L_c' is Eq. (11).

$$L_c = \frac{1}{f_s} + T_p \tag{11}$$

Where,

- f_s → The specimen frequency
- T_p → The processing time for noise reduction and gain control.

2.2.2 Real-Time Processing Requirements

The Voice processing module adheres to stringent real-time constraints to ensure a seamless user experience. Processing voice signals in real-time involves multiple steps, including filtering, FE, and security checks, all of which require to be executed with minimal EED.

1 **EED Constraints:** The EED is $'L_t'$ must remain under 50 *ms* to provide immediate feedback. This EED includes the sum of the capture time (T_c), preprocessing time (T_p), and network transmission time (T_n), Eq. (12)

$$L_t = T_c + T_p + T_n \leq 50 \text{ ms} \quad (12)$$

2 **Edge-Level Processing:** Distributed edge nodes (E_n) handle initial processing tasks such as noise reduction, preliminary FE, and basic Anomaly Detection (AD). This decentralized method minimizes the load on central servers and reduces EED by processing data closer to the source.

3 **Parallel Processing Pipelines:** The module implements parallel processing pipelines for different phases of voice processing. Each pipeline handles a specific task, such as filtering, segmentation, and FE, ensuring that multiple processes are performed concurrently.

The total processing time T_p can be expressed as Eq. (13)

$$T_p = \text{Max}(T_f, T_s, T_e) \quad (13)$$

Where,

- T_f → The filtering time
- T_s → The segmentation time
- T_e → The FE time.

4. Adaptive Load Balancing: To handle variable user loads, the network employs adaptive load balancing systems. The load $'L'$ is dynamically distributed across edge nodes based on their current capacity C_i, Eq. (14)

$$L = \sum_{i=1}^{N} \frac{W_i}{C_i} \quad (14)$$

Where,

- W_i → The workload assigned to the i-th node
- C_i → Its processing capacity.

2.2.3 Data Anonymization

Ensuring privacy during voice signal transmission is critical. The Voice Processing Module uses multiple anonymization to secure user identity while preserving the integrity of the voice data for educational purposes.

1. **Voice Data Masking:** Voice data masking alters identifiable voice features such as pitch and tone while maintaining the linguistic content.

The masked voice signal V_m can be defined as Eq. (15)

$$V_m(t) = T_m(V_c(t)) \tag{15}$$

where

- T_m → A transformation function that replaces the original pitch and tone with neutralized values.

2. **Feature-Level Anonymization:** Before transmitting the data, identifiable features (*e.g.,* speaker-specific features) are obfuscated to ensure anonymity. Let $'F_v'$ represents the feature vector extracted from the voice signal. The anonymized feature vector F_a is, Eq. (16).

$$F_a = F_v \setminus \{f_s\} \tag{16}$$

Where,

- f_s → Speaker-specific features.

3. **Differential Privacy:** The Security against re-identification and controlled noise $'\epsilon'$ is added to the data, Eq. (17)

$$V_d = V_c + \epsilon,\ \epsilon \sim \mathcal{N}(0, \sigma^2) \tag{17}$$

Where,

- σ→ Controls the level of privacy protection.

4. **SMPC:** During distributed processing, SMPC enables different nodes to compute functions on encrypted data without requiring access to the raw data.

The computation of $f(V_c)$ across n nodes is, Eq. (18)

$$f(V_c) = \sum_{i=1}^{n} f_i(V_c^i) \tag{18}$$

Where,

- V_c^i → The encrypted data fragment at node i.

2.[illegible].4 Voice Feature Extraction

FE is the process of transforming raw voice data into a set of measurable features that can be used for analysis and learning tasks. The FE captures the temporal and spectral features of the voice signal.

1. **Mel-Frequency Cepstral Coefficients (MFCC):** MFCC is the spectral envelope of the voice signal.

The MFCC vector F_m is computed as Eq. (19).

$$F_m = \text{DCT}\left(\log\left(|\text{FFT}\,(V_c)|\right)\right) \tag{19}$$

Where,

- DCT→ The discrete cosine transform
- FFT→ The fast frontier transform.

2. **Pitch and Tone Analysis:** Pitch $P(t)$ is predicted using the autocorrelation method, Eq. (20).

$$P(t) = \text{Arg}\ \max_{\tau} \sum_{t=0}^{T} V_c(t) V_c(t+\tau) \tag{20}$$

3. **Spectrogram Analysis:** A spectrogram $S(t, f)$ validates how the frequency content of the voice signal changes over time, Eq. (21)

$$S(t, f) = |\text{STFT}\,(V_c(t))| \tag{21}$$

4 **Zero-Crossing Rate (ZCR):** ZCR counts the rate of sign changes in the signal, Eq. (21)

$$\text{ZCR} = \frac{1}{N-1} \sum_{n=1}^{N-1} |\text{sgn}\,(V_c[n]) - \text{sgn}\,(V_c[n-1])| \tag{22}$$

5 **Energy-Based Features:** The short-term energy $E(t)$ is specified by Eq. (23).

$$E(t) = \sum_{n=0}^{N-1} V_c[n]^2 \tag{23}$$

These FE form a complete vector F_v used for anomaly detection, security implementation, and educational feedback.

2.3 5G + WSN Implementation

The proposed architecture focuses on the 5G + WSN + AI, allowing the VLS to share data at high speeds with minimal energy consumption. Using 5G, the network ensures secure data transmission, effective processing, and direct communication. The 5G + WSN + AI provides error-free reliability by operating within the NPL and integrating directly with VIL and SML. To maximize the reliability and effectiveness of the network, this section describes the layout and implementation factors that must be considered.

2.3.1. Network Topology Design

Data accuracy, trustworthiness, and sustainability are key features of the 5G + WSN + AI. To find a balance between availability and accuracy, the network deploys to a hybrid model that integrates star and mesh topologies at various levels of the framework.

1 **Star Topology for Edge-Level Nodes:** At the edge level, voice capture devices and edge nodes (E_n) are arranged in a star topology, with each node connected to a central

gateway (G_c). This arrangement simplifies data aggregation and minimizes connection overhead.

The edge node communication can be represented as Eq. (24)

$$E_n = \{N_1, N_2, \dots, N_k\} \rightarrow G_c \quad (24)$$

Where,

- N_i→ Individual edge nodes
- k→ The number of nodes connected to the gateway.

2 **Mesh Topology for Core-Level Nodes:** At the core network level, gateways and backbone controllers (B_c) are connected in a mesh topology, ensuring multiple redundant paths for data transmission. This enhances network reliability and fault tolerance.

The core communication paths are expressed as Eq. (25).

$$G_c = \{B_{c1}, B_{c2}, \dots, B_{cn}\} \quad (25)$$

Where,

- B_{ci}→ Backbone controllers
- n → The number of controllers forming the mesh

3 **Hierarchical Network Structure:** Combining these topologies results in a two-tier hierarchical network. Layer two controls the rapid transfer of data between the gateways and the controllers of the backbone, while layer one is molded of edge nodes that send voice data to the gateway. This layered network provides optimum load distribution and adaptability.

4 **Redundancy and Fault Tolerance:** The connection of redundant paths and backup mechanisms in the model improves the model's reliability. If an edge node or gateway fails, data traffic is rerouted by different paths in the mesh network, minimizing disruptions.

2.3.2 Bandwidth Optimization

Efficient use of bandwidth is vital for maintaining the performance of voice applications over the 5G + WSN + AI, particularly when multiple users interact simultaneously.

The system employs several methods to optimize bandwidth utilization.

1 **Adaptive Bitrate Control:** The network dynamically adjusts the bitrate of voice data streams based on network conditions. *e.g.,* The network performs congestion, the bitrate $'B_a'$ is reduced to maintain seamless data flow, Eq. (26).

$$B_a = \mathrm{Max}(B_{\mathrm{Min}}, B_{\mathrm{Ideal}} \cdot C) \quad (26)$$

Where,

- B_{Min} → The minimum allowable bitrate
- B_{Ideal} →the optimal bitrate
- C → The current network capacity factor.

2 **Compression Techniques:** Advanced voice compression, such as Opus and AMR-WB (Adaptive Multi-Rate Wideband), reduce the size of voice packets without cooperating quality.

The compression function $'C_v'$ applied to raw voice data $'V_c'$ generates, Eq. (27)

$$V_c' = C_v(V_c) \tag{27}$$

Where,

- V_c' → The compressed voice data

3 **Quality of Service (QoS) Prioritization:** The network assigns higher priority to real-time voice traffic to ensure low EED and minimal PLR.

QoS policies prioritize voice packets over other data types, Eq. (28).

$$P(V_t) > P(D_t) \tag{28}$$

Where,

- $P(V_t)$ → The priority of voice traffic.
- $P(D_t)$ → The priority of general data traffic.

4 **Packet Aggregation:** To reduce overhead, multiple small voice packets are aggregated into larger frames before transmission.

The aggregated packet P_a is defined as, Eq. (29)

$$P_a = \sum_{i=1}^{m} P_i \tag{29}$$

Where,

- P_i → Individual voice packets
- m → The number of packets aggregated.

2.3.3 EED Management

Maintaining low EED is vital for real-time VLS. The system employs several methods to minimize EED and ensure immediate feedback during VLS activities.

1 **Edge Computing:** Processing tasks are offloaded to edge nodes (E_n) close to the user, reducing the distance data must travel.

The EED as L_e for edge-level processing is given by Eq. (30)

$$L_e = T_c + T_p \tag{30}$$

Where,

- T_c → The capture time
- T_p → The edge processing time.

2 **Network Slicing:** The 5G + WSN employs slicing to allocate dedicated bandwidth and processing resources to VLS. A network slice $'S_v'$ for voice traffic ensures consistent low-EED performance, Eq. (31)

$$S_v = \{B_s, R_s, Q_s\} \tag{31}$$

Where,

- B_s → The allocated bandwidth,
- R_s → The reserved resources
- Q_s → The QoS policy for the slice.

3 **Ultra-Reliable Low-Latency Communication (URLLC):** The network impacts URLLC size of 5G to achieve EED as low as 1 *ms*. URLLC ensures high reliability and low EED for critical voice data transmissions.

4 **Dynamic Latency Control:** Real-time monitoring adjusts EED parameters in response to network load and application requirements.

The dynamic EED as L_d is expressed as Eq. (32)

$$L_d = L_{\text{base}} + \Delta L \tag{32}$$

Where,

- L_{base} → He baseline EED,
- ΔL → The adjustment factor based on current conditions.

2.3.4 Edge Node Deployment

Edge nodes (E_n) are intentionally deployed to balance processing efficiency, EED reduction, and network coverage. The deployment method studies factors such as user density, geographical distribution, and hardware capabilities. Edge nodes are deployed in locations with high user activity, such as classrooms, libraries, and study centers.

The deployment density D_e can be expressed as Eq. (33).

$$D_e = \frac{N_u}{A} \tag{33}$$

Where,

- N_u → The number of users in a region
- A → The area covered by the edge nodes.

Each edge node is equipped with high-performance processors, memory, and specialized hardware tools for real-time voice processing and encryption tasks.

The computational capacity C_e of an edge node is defined as Eq. (34).

$$C_e = f(C_p, C_m, C_a) \tag{34}$$

Where,

- C_p → Processing power
- C_m → Memory capacity
- C_a → The capability of accelerators.

To ensure efficient processing, edge nodes dynamically share workloads based on their real-time size. The load $'L_i'$ on an edge node $'i'$ is Eq. (35).

$$L_i = \frac{W_i}{C_i} \tag{35}$$

Where,

- W_i → The current workload
- C_i → The node's capacity.

Each edge node has backup nodes to ensure constant operation in the event of a failure. Redundant nodes (R_n) activate automatically when a primary node (P_n) Fails, Eq. (36).

$$R_n = \text{Failover } (P_n) \tag{36}$$

2.4. AI-Based Privacy Model

Designed to protect private voice signals in real-time VLS, the AI-Based Privacy Model is a vital module of the network. Securing user privacy and following privacy laws is the highest priority for this model, which is why it incorporates privacy-preserving VLS, Secure Multiparty Computation (SMPC), data minimization, and robust access control mechanisms. The resulting sections validate how these modules work using complete operational measures and models.

2.4.1 Privacy-Preserving Learning Algorithms

To train algorithms on voice signals while securing user privacy, privacy-preserving systems for learning are essential. The network solves this through the use of FL and Differential Privacy.

1. **FL:** FL allows models to be trained on user devices or edge nodes rather than transferring raw voice data to a central server.

The process involves the following steps:

a. **Local Model Initialization:** Each edge node $'E_i'$ sets a local model $'M_i^0'$ based on the global model M_{global}^0.

b. **Local Training:** The edge node $'E_i'$ trains the local model M_i^t using the local dataset $'D_i'$ (containing voice features). The model update ΔM_i^t is Eq. (37).

$$\Delta M_i^t = M_i^t - M_{\text{global}}^t \tag{37}$$

c. **Secure Transmission:** The local update ΔM_i^t is encrypted and sent to the central server.

d. **Global Aggregation:** The central server aggregates the local updates using a weighted average, Eq. (38)

$$M_{\text{global}}^{t+1} = M_{\text{global}}^t + \eta \sum_{i=1}^{N} w_i \Delta M_i^t \tag{38}$$

Where

- η → The learning rate
- N → The number of participants
- w_i → The weight for node $'i'$.

e. **Model Distribution:** The updated global model M_{global}^{t+1} is sent back to all edge nodes for the next round of training.

This decentralized method ensures that raw voice data remains on local devices, reducing privacy risks.

2 **Differential Privacy:** Differential privacy ensures that individual voice signals cannot be reverse-engineered from network updates. The process involves adding controlled noise to the model outputs.

For a function $'f(D)'$ on dataset $'D'$, the device $'M'$ with noise $'N'$ provides ϵ-differential privacy, Eq. (39)

$$M(D) = f(D) + N,\ N \sim \mathcal{N}(0, \sigma^2) \tag{39}$$

- **Noise Calibration:** The standard deviation $'\sigma'$ of the noise is standardized based on the privacy $'\epsilon'$ and sensitivity $'S_f'$ of the function $'f'$, Eq. (40).

$$\sigma = \frac{S_f}{\epsilon} \tag{40}$$

- **Clipping Gradients:** To limit sensitivity, model gradients are clipped before noise addition, Eq. (41)

$$g_i' = \frac{g_i}{\text{Max}\left(1, \frac{\|g_i\|}{C}\right)} \tag{41}$$

Where,

- g_i → The gradient
- C → The clipping threshold

2.4.2. SMPC

SMPC enables multiple entities to collaboratively compute a function over their private inputs without revealing these inputs to one another.

The system employs SMPC for distributed voice data processing and training.

1 **Secret Sharing:** Voice data $'D'$ is divided into $'n'$ shares $D_1, D_2, \ldots, D_n$ such that no single share reveals data about $'D'$.

The shares data as Eq. (42)

$$D = \sum_{i=1}^{n} D_i \text{ Mod} p \quad (42)$$

Where

- p→ A large prime number.
- Each share $'D_i'$ is distributed to a different party.

2 **Computation on Shares:** Each party performs computations on their shares. For a function $f(D)$, the parties compute $f(D_1), f(D_2), \ldots, f(D_n)$. The results are combined to rebuild the output, Eq. (43)

$$f(D) = \sum_{i=1}^{n} f(D_i) \text{ Mod} p \quad (43)$$

3 **Garbled Circuits:** Garbled circuits are used for the secure evaluation of Boolean functions. The process involves:

- **Circuit Generation:** One party (Garbler) generates an encrypted version of the computation circuit.
- **Input Encryption:** Each input bit is assigned a pair of encrypted values (wire labels).
- **Evaluation:** The other party (Evaluator) evaluates the garbled circuit without seeing the actual inputs, obtaining the final result securely.

4 **Oblivious Transfer (OT):** OT enables a party to securely select one of multiple data pieces without the sender being aware of which piece was selected. For input bits $'b'$ and choices m_0, m_1, the receiver attains $'m_b'$ without revealing $'b'$.

2.4.3 Data Minimization Approaches

Data Minimization aims to limit the collection, processing, and storage of voice data to only what is required for the learning application. This reduces exposure to probable attacks and enhances compliance with privacy laws. The system employs a multi-layered method of data minimization involving precise FE, controlled maintenance policies, on-device processing, and adaptive anonymization methods.

1 **FE Process:** Instead of retaining raw voice recordings, the network extracts vital features that are sufficient for learning and analysis tasks.

The process can be broken down as follows:

- **Preprocessing Stage:** Raw voice input $V_c(t)$ is denoised and normalized using filters like Wiener Filtering.

The denoised signal $V_d(t)$ is expressed as, Eq. (44)

$$V_d(t) = V_c(t) - \hat{N}(t) \tag{44}$$

Where,

- $N(t) \rightarrow$ The estimated noise.
  - **Segmentation:** The denoised signal $V_d(t)$ is divided into overlapping frames $F_i(t)$ of length T_f (*e.g.,* 25 *ms*) with a stride of T_s (*e.g.,* 10 *ms*), Eq. (45)

$$F_i(t) = V_d(t + i \cdot T_s),\ i = 0,1,\dots,N_f \tag{45}$$

Where,

- $N_f \rightarrow$ The sum of frames.
  - **Feature Calculation:** From each frame $'F_i'$, relevant features like MFCC, pitch, and energy are extracted, Eq. (46)

    MFCCs: $\text{MFCC}_k = \sum_{j=1}^{M} F_i(j) \cdot \cos\left(\frac{k(j-0.5)\pi}{M}\right)$ (46)

  Where,

  - $k \rightarrow$ The number of cepstral coefficients
  - $M \rightarrow$ The number of Mel filter banks.
    - **Pitch P_i :** Computed using the autocorrelation method for each frame, Eq. (47)

$$P_i = \arg\max_{\tau} \sum_{n=0}^{N_f} F_i(n) F_i(n+\tau) \tag{47}$$

- **Energy E_i:** The total energy in each frame is given by Eq. (48)

  $E_i = \sum_{n=0}^{N_f} F_i(n)^2$ (47)

  The FE as F_v for each segment is then, Eq. (48)

$$F_v = \{\text{MFCC}_k, P_i, E_i\} \tag{48}$$

2. **On-Device Processing and Storage Control:**
   - **Edge-Level Processing:** The FE is processed on edge nodes (E_n) Rather than using central servers, this method minimizes data transfer.
   - **Local Storage Limitations:** Voice data and features are stored temporarily on the user's device. Retention policies enforce automatic deletion after a predefined time (T_{Max}), Eq. (49)

$$D_{\text{retention}} = \{D \mid t \leq T_{\text{MAX}}\} \tag{49}$$

3. **Adaptive Anonymization:** Anonymization is applied dynamically based on the context of data usage.

- **Masking Identifiable Features:** Voice features that may reveal user identities, such as pitch and tone, are masked or randomized while retaining linguistic data. The masked feature vector F_a, Eq. (50)

$$F_a = \{\text{Mask}\ (P_i), E_i, \text{MFCC}_k\} \tag{50}$$

- **Pseudonymization:** User identifiers are replaced with pseudonyms ID_p mapped via secure lookup tables:

$$V_p = \left(F_v, ID_p\right),\ ID_p = \text{Hash}\ (ID_{\text{user}}) \tag{51}$$

These methods collectively ensure that only the minimal and required data is processed and stored, reducing the attack layer and privacy risks.

2.4.4. Access Control Mechanisms

Access control mechanisms enforce strict policies to regulate who can access voice data and system resources, ensuring that only authorized entities interact with sensitive data. The network includes Role-Based Access Control (RBAC), Attribute-Based Access Control (ABAC), and Multi-Factor Authentication (MFA).

1 **RBAC:** RBAC assigns permissions based on predefined roles (R). The access control matrix defines what actions each role can perform on specific resources.

Roles:

- **Student ($\boldsymbol{R_s}$) :** Can access their own voice data and learning progress.
- **Instructor ($\boldsymbol{R_i}$) :** Can access aggregated class performance data but not individual voice data.
- **Administrator ($\boldsymbol{R_a}$) :** Manages system configurations and user roles.
- **Permission Matrix, Eq. (52)**

$$\text{Permission}\ (R, \text{Res}) = \begin{cases} \text{Read} & \text{If } R = R_s \text{ and Res } = \text{ Self Data} \\ \text{Read Aggregate} & \text{If } R = R_i \text{ and Res } = \text{ Class Data} \\ \text{Full Control} & \text{If } R = R_a \end{cases} \tag{52}$$

2 **ABAC:** ABAC evaluates the user as $'A_u'$ and resource as $'Ar'$ to generate dynamic access decisions.

- *Policy Rule, Eq. (53)*

$$\text{Access}\ (A_u, A_r) = \text{ True If } A_u.\text{Role } = \text{ Instructor} \wedge A_r.\text{Type } = \text{ Aggregate Data} \tag{53}$$

3 **MFA:** MFA enhances security by demanding users to provide multiple verification factors.

The authentication process involves:

- **Step 1:** Password entry $'A_p'$.

- **Step 2:** Biometric verification $'A_b'$ (*e.g.,* Voiceprint/Fingerprint).
- **Step 3:** One-time code $'A_o'$ sent to a registered device.

The authentication function $'A_m'$ is expressed as Eq. (54)

$$A_m = \left(A_p \wedge A_b \wedge A_o\right) \tag{54}$$

4 **Audit Logging and Monitoring:** All access attempts and actions are logged in secure, immutable audit trails.

Each log entry $'L_a'$, Eq. (55)

$$L_a = (U_i, A_i, \mathrm{Res}_i, T_i, \text{Status}) \tag{55}$$

Where,

- U_i→ The user
- A_i→ The action,
- Res_i→The resource
- T_i → The timestamp
- Status indicates success or failure.

3. Experimental Set-up

Secure hardware, cutting-edge software, optimized network settings, and a controlled testing environment are all components of the research setup for voice security applications in VLS. This setup enables the 5G + WSN + AI to properly record, analyze, and secure audio data while being reactive in real-time and maintaining compliance with privacy standards.

3.1. Hardware Configuration

The setup of the hardware enables secure communication and distributed processing through the use of user-end devices, edge nodes, and central servers. In order to collect voice input with minimal intrusion, devices like smartphones and computers with focused microphone sets and high-quality audio connectors are used at the user's side. To facilitate local preprocessing tasks, these systems have been equipped with processors such as the Intel Core i7-1165G7 and 16 GB of RAM. To manage FE, preliminary verification of security, and real-time voice data processing, each edge node is provided with 64 GB of RAM, 1 TB of storage space on an SSD, and an Intel Xeon E5-2670 CPU. An NVIDIA DGX Station, equipped with four NVIDIA Tesla V100 GPUs, 128 GB of RAM, and a 4 TB NVMe SSD, is designed for demanding AI training and FL aggregation and is assigned to model aggregation and high-level data management.

3.2. Software Components

To implement voice capture, processing, and privacy measures, the system relies on a collection of software tools and frameworks. Researchers use Python and various library resources, such as SciPy for signal processing and Librosa for FE, to analyze audio signals. Also, deploy TensorFlow 2.5 and PyTorch 1.9 for developing Machine Learning (ML), which involves FL and Differential Privacy. The PySyft library, a network for encrypted processes across multiple nodes, has been integrated into the SMPC. The encryption techniques use Elliptic Curve Cryptography (ECC) for safe key exchange and the PyCryptodome library for AES-256 encryption. In order to guarantee continuous use and scalability, the model's backend uses Flask for API development and Docker containers. System functionality and health can be monitored in real-time with Grafana and Prometheus.

3.3 Network Parameters

The 5G + WSN + AI setup supports high-speed, low-EED communication required for real-time voice data processing. The network operates in the 3.5 GHz (C-band) frequency band, with a bandwidth of 100 MHz, to store multiple users. The peak data rate for uplink and downlink is set at 1 and 10 Gbps, respectively. Edge nodes are deployed within a 50 *m* radius of user devices to minimize latency. The average round-trip latency between the user device and the edge node is 5 *ms*., while the latency between edge nodes and the central server is stopped at 20 *ms*. To ensure Quality of Service (QoS), network slicing is implemented, reserving dedicated bandwidth for voice data traffic. Adaptive bitrate control dynamically adjusts data rates in response to varying network congestion levels.

3.4 Testing Environment

The testing environment is set up in a controlled lab environment, simulating real-world classroom and home-learning conditions. The lab is equipped with acoustic panels to control noise levels and ensure consistent audio quality. Tests are conducted with a sample size of 50 users, each performing VLS over several network conditions, including high-load scenarios to assess scalability. The environment includes edge computing nodes strategically positioned to simulate varying distances and network conditions.

Three test scenarios are implemented:

(1) Ideal network conditions with minimal EED and no PLR,

(2) Congested network conditions with 10% PLR and 100 *ms* EED

(3) Edge node failure scenarios to evaluate system robustness and failover mechanisms.

The primary performance measures, such as EED, NT, PLR, and voice processing accuracy, are monitored in real-time using specialized tools. The primary aim of security evaluation is to assess the value of secure user data mechanisms, such as encryption, anonymization, and access control. To ensure the network continues to function correctly and securely under various use cases, results have been recorded and analyzed for changes in network parameters.

3.5 Metrics and Baseline

Several metrics are provided to evaluate the proposed 5G + WSN + AI for secure VLS, focusing on efficiency, security, and performance. Integrated with these metrics are features like precision, computational speed, security reliability, and network performance.

Additionally, baseline models are developed to provide comparative analysis and highlight the advantages of the proposed system.

i) Metrics

1 **Accuracy Metrics:** These metrics assess the network's ability to capture and process voice data for VLS responsibilities accurately.

- **WER:** The WER quantifies the transcription accuracy of the VLS.

It is calculated as Eq. (56).

$$\text{WER} = \frac{S+D+I}{N} \quad (56)$$

Where,

- $S \rightarrow$ The number of substitutions
- $D \rightarrow$ The number of deletions
- $I \rightarrow$ The number of insertions
- $N \rightarrow$ The sum of words in the reference transcript.

2. **FEA:** This metric evaluates the accuracy of voice FE (MFCC, diameter, and energy) to data from ground-based sensors. The ratio of authentic FE is the standard deviation of accuracy.

3. **Network Performance Metrics:** These metrics assess the efficiency of the 5G + WSN + AI in handling voice data transmission.

- **NED:** The round-trip time for voice data to travel from the user device to the edge node and back. Measured in milliseconds (*ms*), it should be minimized for real-time feedback, Eq. (57)

$$L = T_{\text{Transmission}} + T_{\text{Processing}} \quad (57)$$

- **NT:** The rate at which voice data is successfully transmitted by the network, measured in bits per second (*bps*).
- **PLR:** The percentage of packets lost during transmission, computed as Eq. (58).

$$\text{PLR} = \frac{\text{Number of Lost Packets}}{\text{Total Packets Sent}} \times 100\% \tag{58}$$

4. **Computational Efficiency Metrics:** These metrics evaluate the system's ability to process voice data efficiently and promptly.
   - **Processing Time (PT):** The time reserved to process a single voice input, including feature extraction and encryption.
   - **Edge Node Utilization (ENU):** The percentage of computational resources used by edge nodes during processing, Eq. (59)

$$\text{ENU} = \frac{\text{Active Processing Time}}{\text{Total Available Time}} \times 100\% \tag{59}$$

4 **Security Metrics:** These metrics measure the effectiveness of the security mechanisms.
   - **Encryption Overhead (EO):** The additional processing time incurred due to encryption, expressed as Eq. (60)

$$\text{EO} = \frac{T_{\text{encrypted}} - T_{\text{unencrypted}}}{T_{\text{unencrypted}}} \times 100\% \tag{60}$$

   - **Privacy Leakage Rate:** The probability of rebuilding the original voice data from anonymized or encrypted data.
   - **Access Control Effectiveness (ACE):** The percentage of unauthorized access attempts successfully blocked by the system.

ii) Baseline Models

The proposed 5G + WSN + AI is evaluated against four baseline models to highlight its advantages in security, privacy, and performance.

- The SVPS uses traditional voice capture and processing without encryption, FL, or anonymization, making it vulnerable to data breaches. It serves as a benchmark for assessing security enhancements.
- The CBCM processes voice data on a central server, leading to high latency, privacy risks, and potential network congestion. It shows the benefits of edge-based FL in reducing latency and enhancing privacy.
- The [illegible]DPS applies differential privacy during training but lacks FL and SMPC, thereby risking accuracy degradation and data exposure. It helps evaluate the combined effectiveness of FL and differential privacy.
- The GACS implements basic RBAC without MFA or real-time monitoring, making it susceptible to unauthorized access. It benchmarks the robustness of the proposed access control mechanisms.

- These models provide a comparative model to prove the proposed 5G + WSN +AI superiority in secure, privacy-preserving voice processing.

4. Results and Analysis

The WER comparison (Figure 2) across different numbers of voice samples illustrates the superior performance of the proposed model over baseline models, including SVPS, CBCM, BDPS, and GACS. Maintaining values between 0.0598 and 0.0802, the recommended approach sustains the lowest WER as the sum of voice recordings increases from 10 to 200. As the test set size increases, the model continues to capture and process voice data, demonstrating its resilience accurately.

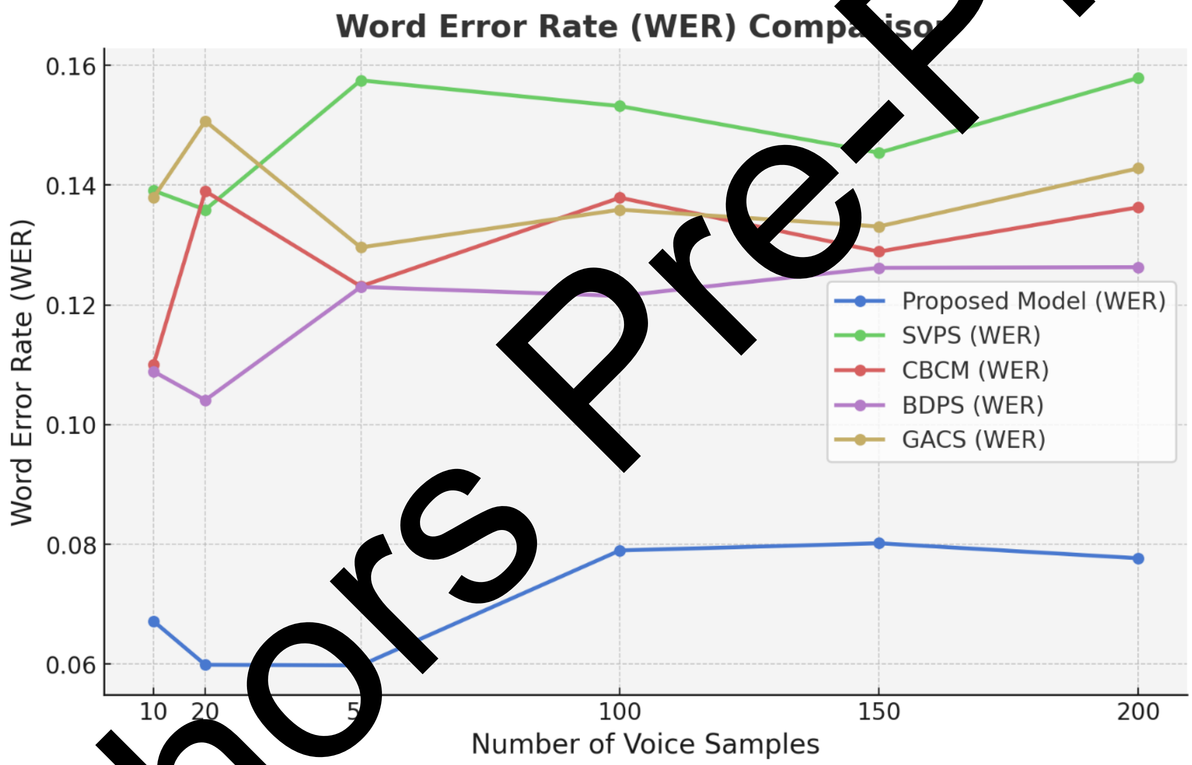


Figure 2: WER Comparison

In contrast, the SVPS exhibits higher and more variable WER, ranging from 0.1359 to 0.1579, due to the lack of privacy-preserving methods and optimized processing. The CBCM exhibits moderate performance, with WER values ranging from 0.1100 to 0.1390, indicating the impact of network EED and centralized data processing on accuracy.

The BDPS maintains WER values between 0.1041 and 0.1263, indicating that while differential privacy protects data, the added noise affects accuracy, especially with larger sample sizes.

The GACS follows a similar trend, with WER values ranging from 0.1296 to 0.1507, highlighting the limitations of traditional access control mechanisms without advanced optimization methods.

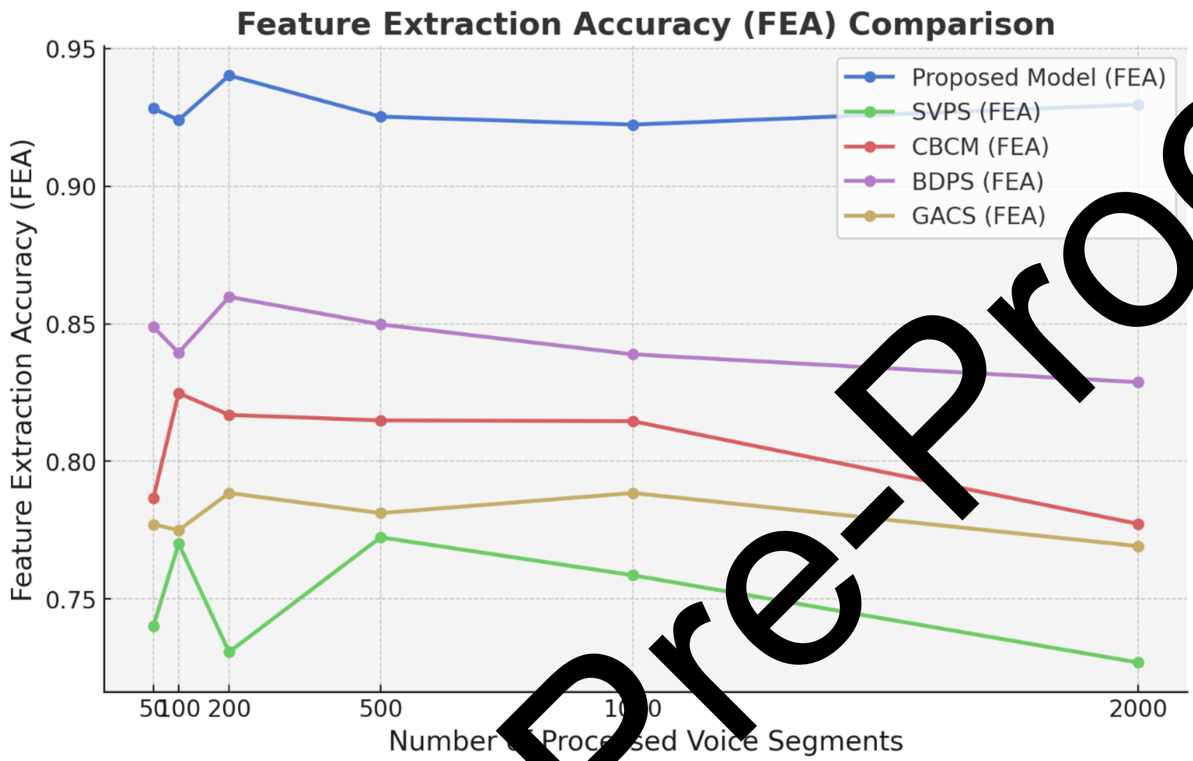


Figure 3: FEA Comparison

The FEA comparison (Figure 3) across varying numbers of processed voice segments demonstrates the superior performance of the proposed 5G + WSN + AI over the baseline models, including SVPS, CBCM, BDPS, and GACS. The proposed 5G + WSN + AI consistently maintaining FEA ranging between 0.9224 and 0.9403, signifying its ability to accurately predict FE from voice data, even as the number of processed segments increases from 50 to 2000. This indicates that the advanced proposed 5G + WSN + AI ensures robust performance and scalability.

In contrast, the SVPS challenges lower FEA values, ranging from 0.7269 to 0.7724, due to the lack of optimization and privacy-preserving mechanisms.

The CBCM shows moderate accuracy, ranging from 0.7773 to 0.8248, but is delayed by centralized processing constraints and latency, which impede the timely extraction of accurate features.

The BDPS achieves slightly better accuracy, ranging from 0.8288 to 0.8598, but the addition of noise for privacy protection slightly degrades feature quality.

The GACS maintains FEA between 0.7692 and 0.7885, reflecting the limitations of traditional access control systems without advanced optimization methods.

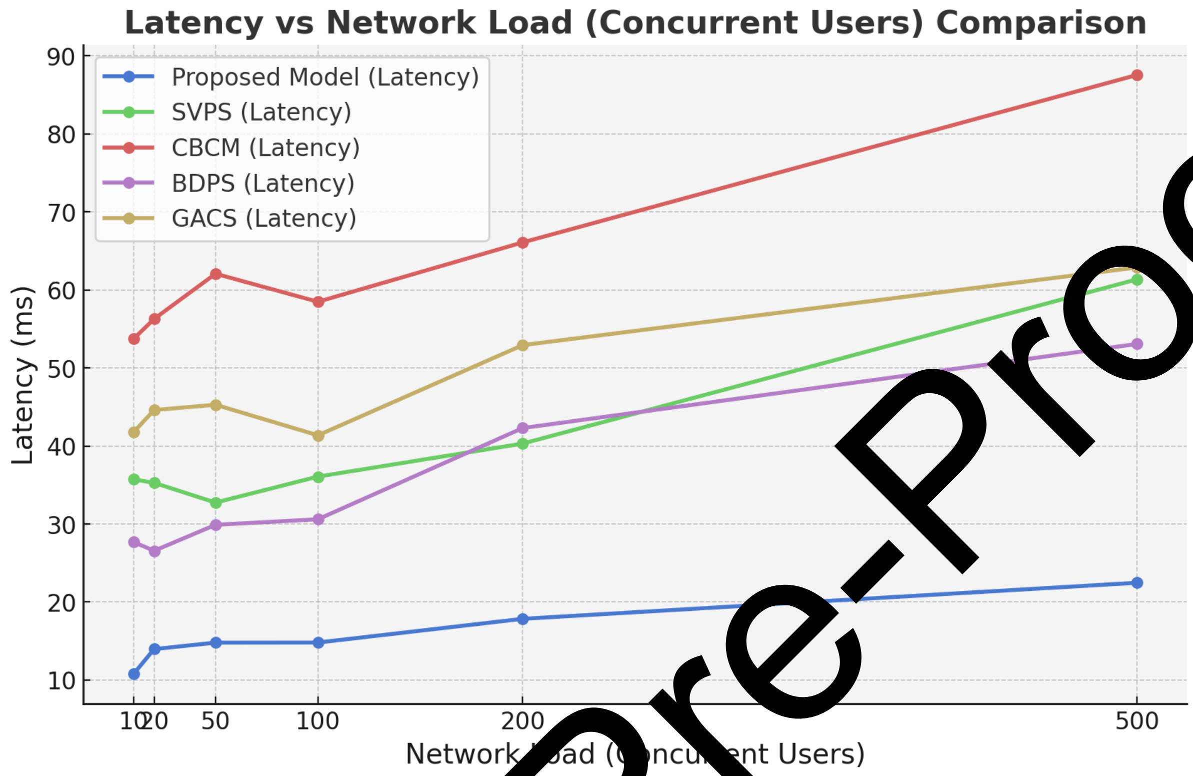


Figure 4: EED vs Network Load (NL) (Concurrent Users)

The comparison of EED (Figure 4) across different NL reveals the superior efficiency of the proposed 5G + WSN + AI in handling concurrent users compared to baseline models such as SVPS, CBCM, BDPS, and GACS. As the number of concurrent users increases from 10 to 500, the proposed 5G + WSN + AI maintains a significantly lower EED, ranging from 10.80 to 22.47 *ms*. This proves the effectiveness of edge-based processing and optimized 5G integration in reducing EED, ensuring real-time responsiveness even under high user loads.

In contrast, the SVPS exhibits higher EED, ranging from 32.75 to 61.36 *ms*. The lack of optimization and reliance on traditional processing methods result in EED, particularly as network load increases.

The CBCM exhibits the highest EED, ranging from 53.74 to 87.56 *ms*. This is due to centralized data processing and the inherent network congestion that arises when handling large numbers of concurrent users, making it unsuitable for real-time applications.

The BDPS maintains moderate EED, ranging from 26.57 to 53.08 *ms*. The privacy-preserving noise addition and centralized processing contribute to these EEDs, which increase significantly as the number of users increases.

The GACS also suffers from high EED, varying between 41.37 and 62.89 ms, primarily due to the lack of optimization in managing large-scale user loads and the reliance on basic access control mechanisms.

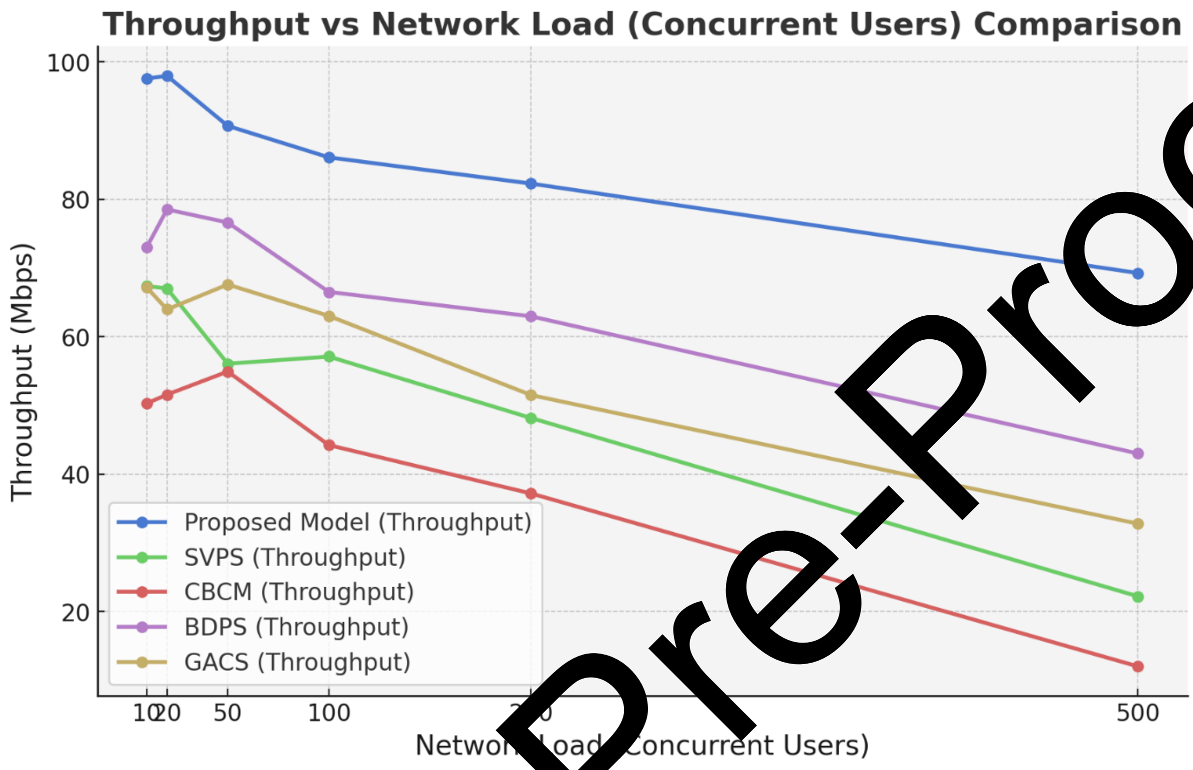


Figure 5: NT vs NL (Concurrent Users)

The comparison of NT (Figure 5) across different NL levels (*i.e.,* the number of concurrent users) highlights the effectiveness of the proposed 5G + WSN + AI in maintaining high data transmission rates under increasing user demand. The proposed 5G + WSN + AI consistently achieves the highest NT, starting at 97.57 *Mbps* with 10 concurrent users and gradually decreasing to 69.21 *Mbps* with 500 users. This performance proves the effectiveness of the proposed 5G + WSN + AI's integration with 5G + WSN, optimized data handling, and edge-based processing, enabling it to manage higher loads without significant degradation in NT.

In contrast, the SVPS begins with an NT of 67.38 *Mbps* at 10 users but drops sharply to 22.22 *Mbps* at 500 users. This decline is attributed to the lack of optimization in handling concurrent connections and traditional processing methods, resulting in network congestion.

The CBCM expressions the poorest performance, with NT starting at 50.30 *Mbps* and plummeting to 12.03 *Mbps* as the network load increases. The centralized nature of this proposed 5G + WSN + AI generates bottlenecks and EED, making it unsuitable for real-time applications with high user demand.

The BDPS maintains practical NT, beginning at 73.04 *Mbps* and decreasing to 43.01 *Mbps*. The privacy-preserving noise addition and EO donate to the decline of NT as the number of users increases.

The GACS exhibits NT values starting at 67.15 *Mbps* and falling to 32.79 *Mbps* with 500 users, reflecting the limitations of traditional access control mechanisms and a lack of optimization for large-scale concurrent processing.

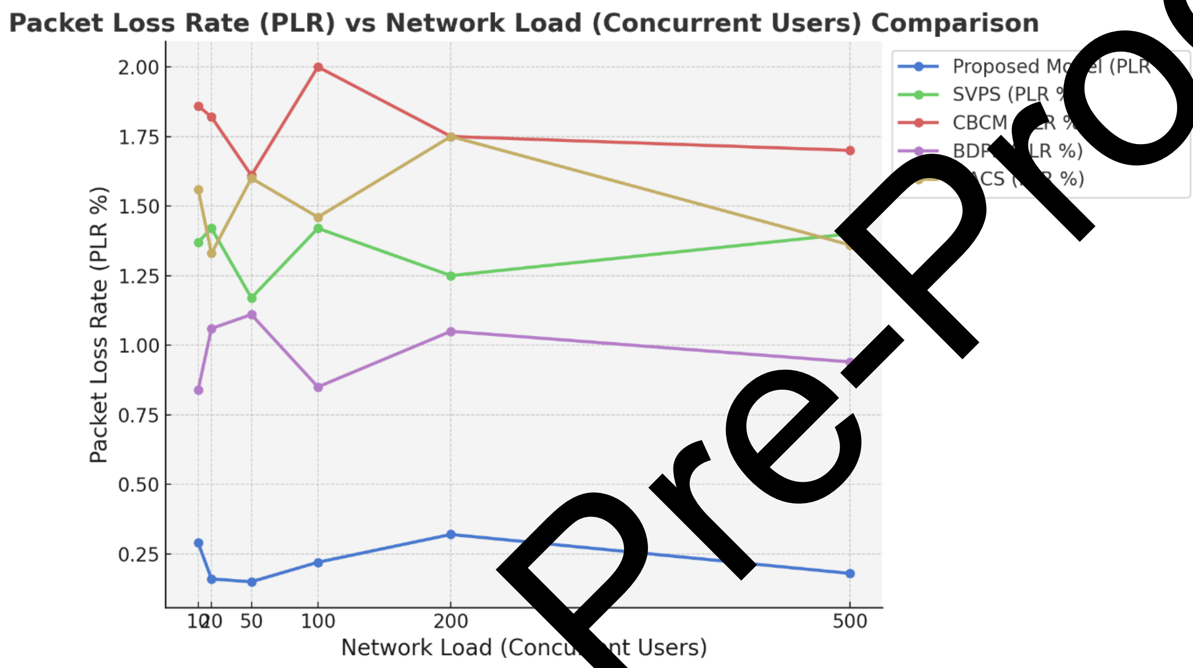


Figure 6. PLR vs NL (Concurrent Users)

The comparison of PLR across different NL (Figure 6) loads highlights the resilience and efficiency of the proposed 5G + WSN + AI in maintaining data integrity as the number of concurrent users increases. The proposed 5G + WSN + AI consistently achieves the lowest PLR, starting at 0.29% for 10 concurrent users and maintaining a range between 0.15% and 0.32% as the load increases to 500 users. This low PLR is a result of optimized 5G + WSN + AI, efficient data routing, and edge-based processing, which minimize congestion and PLR even under high user loads.

In contrast, the SVPS experiences higher PLR values, ranging from 1.17% to 1.42%, reflecting the inefficiencies in traditional processing networks, which lack optimization for concurrent connections.

The CBCM exhibits the highest PLR, starting at 1.86% and reaching 2.00% as the load increases. The centralized nature of this model creates significant bottlenecks and network congestion, resulting in higher PLR and making it unsuitable for real-time applications.

The BDPS maintains moderate PLR values, ranging from 0.84% to 1.11%. The added encryption and noise for privacy protection contribute to slight PLR, particularly as the number of concurrent users increases.

The GACS also consistently shows high PLR values, fluctuating between 1.33% and 1.75%, which reflects the limitations of traditional access control mechanisms and the lack of efficient load-balancing methods.

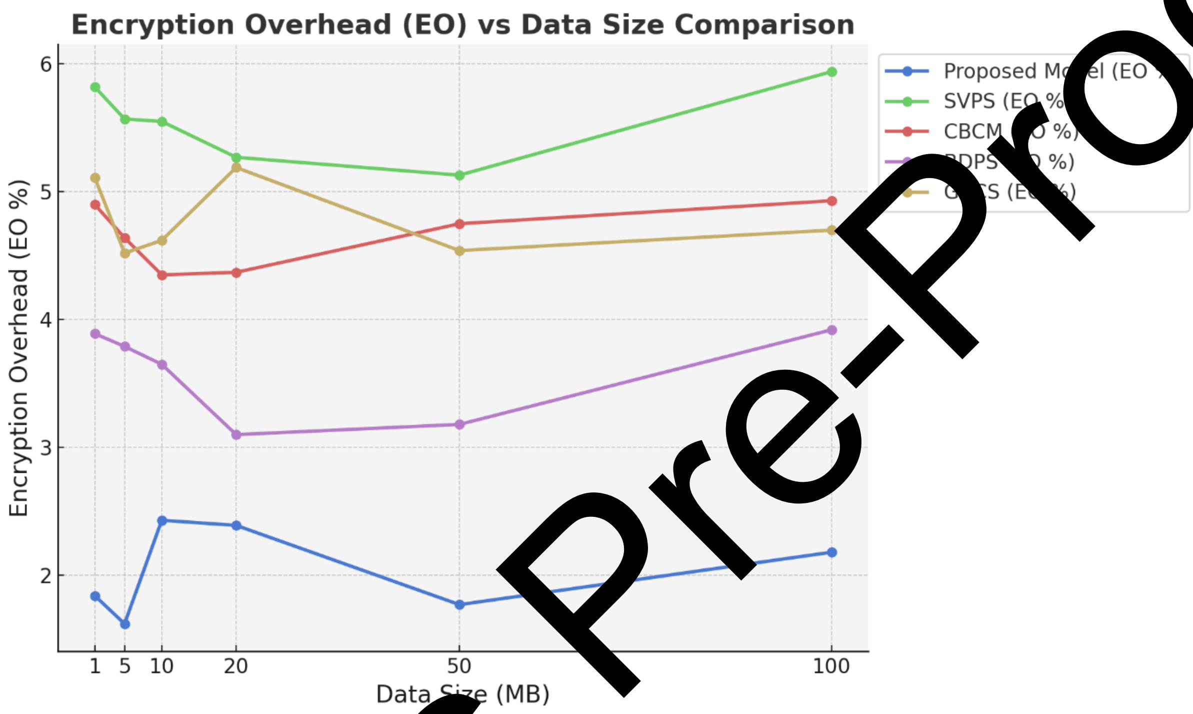


Figure 7: EO vs Data Size

The comparison of EO across different data sizes (Figure 7) proves the efficiency of the proposed 5G + WSN + AI in minimizing the additional computational burden associated with encryption. The proposed 5G + WSN + AI consistently achieves the lowest EO, ranging between 1.62% and 2.43%, even as data sizes increase from 1 to 100 MB. This efficiency is due to optimized encryption methods, such as lightweight AES-256 implementations and streamlined edge-based processing, which minimize processing EED while maintaining security.

In contrast, the SVPS shows the highest EO, ranging from 5.13% to 5.94%. The lack of optimization and reliance on conventional encryption methods result in significant computational costs, particularly as data sizes increase.

The CBCM exhibits moderate overhead values, ranging from 4.35% to 4.93%. The centralized encryption processes incur EED due to data transmission and processing bottlenecks, resulting in higher EO compared to the proposed 5G + WSN + AI.

The BDPS achieves EO values between 3.10% and 3.92%. The additional noise injection for privacy preservation contributes to a moderate increase in processing EED, particularly for larger data sizes.

The GACS shows an increase in EO from 4.52% to 5.19%, reflecting the inefficiencies of basic access control mechanisms combined with standard encryption methods.

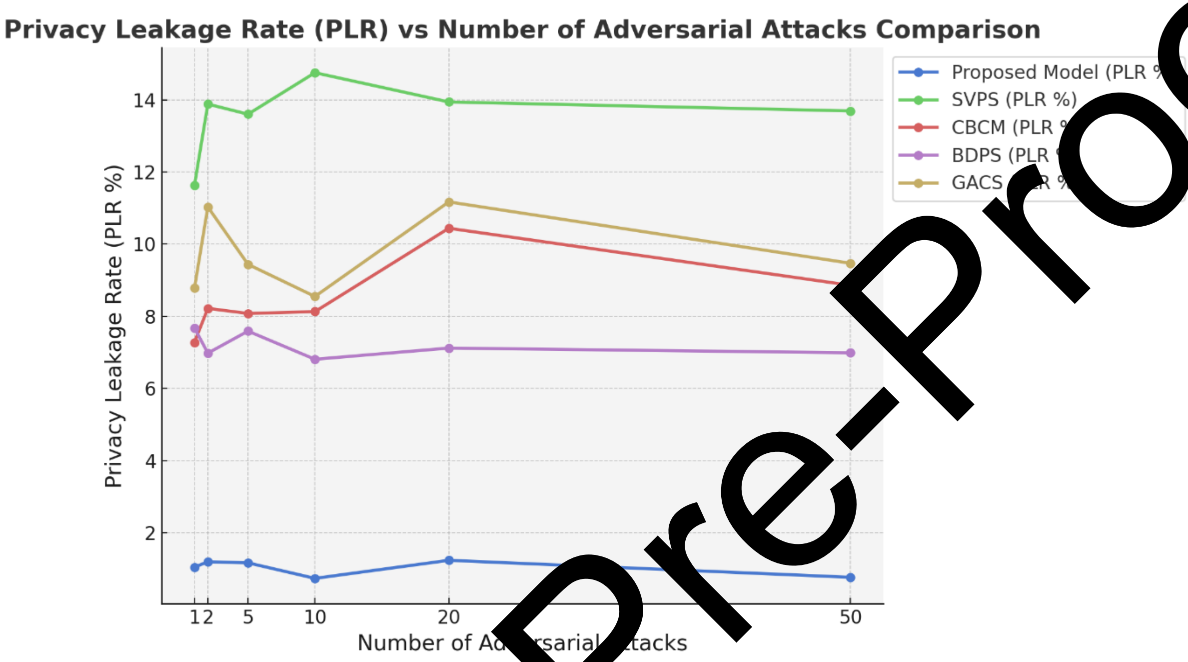


Figure 8: Privacy Leakage Rate (PrLR) vs Number of Adversarial Attacks

The comparison of PrLR (Figure 8) across variable numbers of adversarial attacks highlights the robustness of the proposed 5G + WSN + AI in maintaining data privacy, outperforming the baseline models: SVPS, CBCM, BDPS, and GACS. The proposed model consistently achieves the lowest PLR, with values ranging between 0.74% and 1.24%, even as the number of adversarial attacks increases from 1 to 50. This proves the effectiveness of the proposed 5G + WSN + AI's privacy-preserving mechanisms, including FL, SMPC, and dynamic anonymization, in safeguarding sensitive data against privacy attacks.

In contrast, the SVPS shows the highest PLR, starting at 11.63% for a single attack and rising to 14.75% with 10 adversarial attacks. The lack of encryption, anonymization, and other privacy-preserving methods makes this model particularly vulnerable to data leakage.

The CBCM exhibits moderate PLR, with values ranging from 7.28% to 10.44%. The centralized processing of data increases the risk of privacy breaches, particularly as the number of adversarial attacks increases.

The BDPS performs better than SVPS and CBCM, with PLR ranging from 6.81% to 7.68%. The differential privacy methods add noise to secure data, but this method offers limited resilience against higher numbers of attacks.

The GACS shows PLR values between 8.55% and 11.17%, reflecting the vulnerabilities of basic access control mechanisms, which lack advanced privacy-preserving measures.

5. Conclusion and Future Work

This study presents a comprehensive network to enhance the performance and security of VLS by integrating 5G + WSN + AI. The model provides the efficient and secure analysis of sensitive voice data through the integration of FL, SMPC, and dynamic encryption. To overcome the limitations of conventional models, 5G, WSN + AI processing reduces EED and network load. Several performance metrics validate that the recommended system is superior to baseline models, including SVPS, CBCM, BDPS, and GACS. Consistently low WER and high FEA performance are achieved by the proposed model, regardless of the volume of data or user load. Because it maintains its EED at a low level and its NT at a high level, it works exceptionally well for RL, which requires real-time virtual learning. The method's effectiveness in managing secure data communications is reinforced by its minimal PLR and low EO. Even when attacked by an increasing amount of attackers, the network maintained a low PrLR, proving its resilience. This work addresses the immediate demand for VLS applications that are secure, have low EED, and perform highly. In addition to increasing security, scalability and real-time adaptability are ensured by integrating 5G + WSN + AI-driven privacy methods. These findings lead to the method for more trustworthy and fetching online learning by providing substantial proof for the application of secure VLS in virtual classrooms.

Research in the future may explore the merits of integrating additional privacy-protecting methods and developing the application of this model to include other domains in the context of language study.

References:

1. Li, Y., & Wu, F. (2023). Design and Application Research of Embedded Voice Teaching System Based on Cloud Computing. *Wireless Communications and Mobile Computing*, *2023*(1), 7873715.
2. Essel, H. B., Vlachopoulos, D., Nunoo-Mensah, H., & Amankwa, J. O. (2024). Exploring the impact of VoiceBots on multimedia programming education among Ghanaian university students. *British Journal of Educational Technology*.
3. Saadia, K. H. (2023). Assessing the Effectiveness of Text-to-Speech and Automatic Speech Recognition in Improving EFL Learner's Pronunciation of Regular Past-ed.

4. Karatay, Y., & Hegelheimer, V. An overview of new technologies in English language teaching. *Educational technology in English Language Teaching. Eğiten Kitap (2023, Preprint).*
5. Li, J., Chen, C., Azghadi, M. R., Ghodosi, H., Pan, L., & Zhang, J. (2023). Security and privacy problems in voice assistant applications: A survey. *Computers & Security*, *134*, 103448.
6. Raja, V. (2024). Exploring challenges and solutions in cloud computing: A review of data security and privacy concerns. *Journal of Artificial Intelligence General Science (JAIGS) ISSN: 3006-4023*, *4*(1), 121-144.
7. Chukwunweike, J. N., Yussuf, M., Okusi, O., & Oluwatobi, T. (2024). The role of deep learning in ensuring privacy integrity and security: Applications in AI-driven cybersecurity solutions. *World Journal of Advanced Research and Reviews*, *23*(2), 2550.
8. JosephNg, P. S., EricMok, Z. C., Phan, K. Y., Sun, J., & Wei, Z. (2025). Mitigating Social Media Cybercrime: Revolutionising with AES Encryption and Generative AI. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *46*(2), 124-154.
9. Jalal Mohammed Hachim Altmemi et al., A Software-Centric Evaluation of the VEINSFramework in Vehicular Ad-Hoc Networks, Journal of Robotics and Control (JRC), Volume 6, Issue 2, 2025.
10. Burhan, M., et al., (2023). A comprehensive survey on the cooperation of fog computing paradigm-based IoT applications: layered architecture, real-time security issues, and solutions. *IEEE Access*.
11. Abdulrazzaq, A. Z., Ali, Z. G., Al-Ani, A. R. M., Khaleel, B. M., Alsalame, S., Snovyda, V., & Kanbar, A. B. (2024). Evaluation of Voice Interface Integration with Arduino Robots in 5G Network Frameworks. *36th Conference of Open Innovations Association, 44-53.*
12. Khedkar, A., Musale, S., Padalkar, G., Suryawanshi, [illegible], & Sa[illegible]e, S. (2023). An Overview of 5G and 6G Networks from the Perspective of AI Applications. *Journal of The Institution of Engineers (India): Series B*, *104*(6), 1329-1341.
13. Fakhouri, H. N., Alawadi, S., Awaysheh, F. M., Hani, I. B., Alkhalaileh, M., & Hamad, F. (2023). A comprehensive study on the role of machine learning in 5G security: challenges, technologies, and solutions. *Electronics*, *12*(22), 4604.
14. Abdullah Lakhan et al., Secure blockchain assisted Internet of Medical Things architecture for data fusion enabled cancer workflow, Internet of Things, Vol. 24, 2023, 100928.
15. Alsadie, D. (2024). Artificial Intelligence Techniques for Securing Fog Computing Environments: Trends, Challenges, and Future Directions. *IEEE Access*.
16. Rupanetti, D., & Kaabouch, N. (2024). Combining Edge Computing-Assisted Internet of Things Security with Artificial Intelligence: Applications, Challenges, and Opportunities. *Applied Sciences*, *14*(16), 7104.
17. Rafique, W., Barai, J., Fapojuwo, A. O., & Krishnamurthy, D. (2024). A survey on beyond 5g network slicing for smart cities applications. *IEEE Communications Surveys & Tutorials*.
18. Nazrin Haziq Jemaludin et al., Compact Physical and Electrical Patch Antenna Engineered for 5G Mobile Devices and Multiband Systems," Progress In Electromagnetics Research B, Vol. 111, 71-81, 2025.
19. J. K. Madhloom, et al., An Information Security Engineering Framework for Modeling Packet Filtering Firewall Using Neutrosophic Petri Nets. *Computers* 2023, *12*, 202.
20. Nazrin Haziq Jemaludin et al., A Miniaturized Horizontally Oriented MIMO Antenna for 5G and Wireless Communication Systems, International Journal of Intelligent Engineering and Systems, Vol.18, No.5, 2025.