Infocommunications Journal

A PUBLICATION OF THE SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS (HTE)

March 2026 Volume XVIII Number 1 ISSN 2061-2079

Technically Co-Sponsored by

IEEE ComSoc™ IEEE Communications Society

hte

IEEE HUNGARY SECTION

Indexing information

Infocommunications Journal is covered by Inspec, Compendex and Scopus.
Infocommunications Journal is also included in the Thomson Reuters – Web of ScienceTM Core Collection, Emerging Sources Citation Index (ESCI)

Infocommunications Journal

Technically co-sponsored by IEEE Communications Society and IEEE Hungary Section

Supporters

FERENC VÁGUJHELYI – president, Scientific Association for Infocommunications (HTE)

MTA200

HTE Infocommunications Journal welcomes the 200th anniversary of the Hungarian Academy of Sciences and gratefully acknowledges its continuous support of the journal in recent years.

nmhh
National Media and Infocommunications Authority | Hungary

The publication was also produced with the support of the NMHH.

Editorial Office (Subscription and Advertisements)**:**
Scientific Association for Infocommunications
H-1051 Budapest, Bajcsy-Zsilinszky str. 12, Room: 502
Phone: +36 1 353 1027 • E-mail: info@hte.hu • Web: www.hte.hu

Articles can be sent also to the following address:
Budapest University of Technology and Economics
Department of Telecommunications and Media Informatics
Phone: +36 1 463 4189 • E-mail: pvarga@tmit.bme.hu

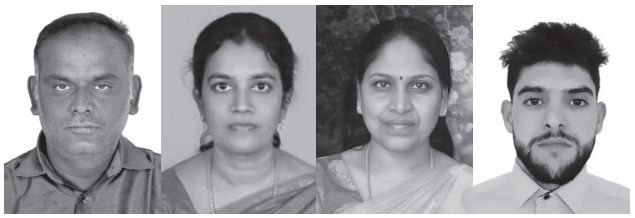
Subscription rates for foreign subscribers: 4 issues 13.700 HUF + postage

Publisher: PÉTER NAGY

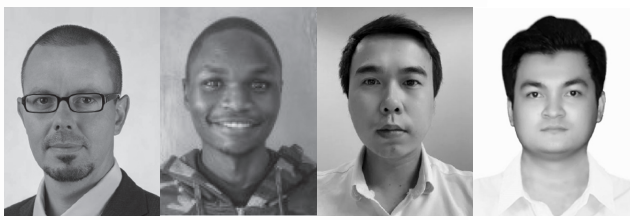
HU ISSN 2061-2079 • Layout: PLAZMA DS • Printed by: FOM Media

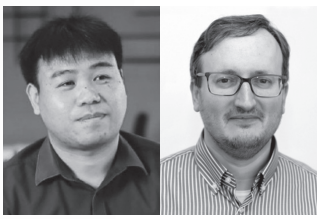
Authors, co-authors of the March 2026 issue

M. Rajakani, R. Beaulah Jeyavathana,
R. J. Kavitha, Oussama Lagnfdi,
Marouane Myyara, Anouar Darif,
Hussein Tuama Al-kinani, Sándor Imre,
Márton Pál Lipcsey-Magyar,
Attila Ármin Madarász, Adrian Pekar,
Nour Ammar, László T. Kóczy,
Doaa Mostafa, Sally S. Ismail,
Mostafa Aref, György Wersényi,
Victor Koech, Sang-Quang Nguyen,
Chi-Bao, John Baghous,
Mohamed Khaled Chahine,
Angel D, Robin Rohit Vincent,
Aphilak Lonklang, János Botzheim









Relay Pursuit-Vathana: A Novel Optimization Approach for Feature Selection in Software Defect Prediction

Rajakani M.[1], Beaulah Jeyavathana R.*[2], and Kavitha R. J.[3]

***Abstract*—Software defect prediction plays a crucial role in ensuring the quality and reliability of software systems. Feature selection, the process of identifying the most relevant features from a large set of potential features which is essential for building effective defect prediction models. In this paper, we propose a novel feature selection model based on the RelayPursuit-Vathana (RP-Vathana) optimization algorithm, inspired by relay races and pursuit dynamics in biological systems. The proposed model aims to identify an optimal subset of features for software defect prediction, maximizing the predictive performance of the resulting classification model. The RP-Vathana algorithm was integrated with a Naïve Bayes classifier and benchmarked on three datasets (PC5, JM1, KC2) to validate its effectiveness in feature selection for defect prediction. The results show that RP-Vathana significantly outperforms existing wrapper-based methods, obtaining mean accuracies of 94.28%, 93.69%, and 96.35% on PC5, JM1, and KC2, respectively, compared to the 83–90% range of rival techniques. While the parameter-free design improves usability, the algorithm's performance on highly noisy or very small datasets warrants future investigation into hybrid extensions for enhanced robustness.**

***Index Terms*—Software defect prediction, Feature selection, Bio-inspired algorithm, RP-Vathan**

I. INTRODUCTION

In the software development landscape, ensuring product quality and reliability is critical. Software flaws frequently appear despite careful planning and thorough testing, presenting serious difficulties for stakeholders and developers. Software defect prediction(SDP) model is essential for developing better software by enabling early defect identification, optimizing resource allocation, improving software quality, mitigating risks, enhancing decision-making, and fostering continuous improvement in the software development process.[1-2] Fixing errors caused by logic mistakes, insufficient requirement analysis and inaccurate system design takes time and resources, and it can negatively impact user experience and accurate solution.[3-4] As a result, proactive defect prediction techniques are gaining popularity as a means of identifying possible problems before they materialise in operational environments. Machine learning (ML) techniques have emerged as powerful tools for software defect prediction due to their ability to analyze vast amounts of data, extract patterns, and make predictions with remarkable accuracy [5]. By leveraging historical project data, such as code metrics, version control information, and defect reports, ML models can learn to identify patterns indicative of defect-prone areas within software systems.[6] In this study, the objective is to develop a defect prediction model for software development using supervised learning algorithms.
Various software metrics are utilized as decision variables to predict faults within the software. Given the vast dimensionality of the feature set, the aim is to employ feature selection techniques to identify the optimal subset of features that contribute most significantly to defect prediction accuracy.

The wrapper method stands out as the most commonly used approach for feature selection. It selects optimal features based on the performance of the predictive model, potentially enhancing its effectiveness. Moreover, wrapper methods can adapt to various performance metrics tailored to specific problem domains. These methods evaluate individual features while considering their interactions, thereby capturing complex relationships among them, which cannot be assessed by evaluating features individually. The commonly used wrapper approaches for software defect prediction model are: Particle Swarm Optimization (PSO) [7], Boosted Whale Optimization (BWO) [8], Ant Colony Optimization (ACO) [9], Genetic Algorithm (GA) [10], Firefly Algorithm (FA) [11] and Hybrid Grey Wolf Optimization (HGWO) [12].

In recent years, there has been a growing interest among researchers in developing optimization algorithms that operate without the need for control parameters because of the reduced complexity and improved convergence rates. Rao [32] developed the Rao Optimization algorithm, which updates the population set by leveraging the features of the best and worst individuals, without requiring any control parameters. This approach yields improved results. For instance, a parameter-free optimization algorithm [31]] has been successfully applied to the document classification problem demonstrating its robustness across domains. This highlights the overall effectiveness of such algorithms, while our work specifically targets their application in software defect prediction. Although promising, these approaches rely solely on the best and worst individuals in the population for updates, neglecting the contributions of intermediate solutions, which can restrict exploration of the search space.

[1, *2,3] Assistant Professor
Department of Data Science and Business Systems Department of Computational Intelligence
Department of Electronics and Communication Engineering
[1, *2] SRM Institute of Science and Technology, University College of Engineering Panruti, Panruti, India
[1] (e-mail: rajakanijes@gmail.com)
[*2] (e-mail: rbvathanaj@gmail.com)
[3] (e-mail: rjkresearch@gmail.com)

DOI: 10.36244/ICJ.2026.1.1

To address these shortcomings, the present study introduces a novel parameter-free and deterministic optimization algorithm, named RP-Vathana, for feature selection in SDP. Unlike conventional metaheuristics, RP-Vathana eliminates the reliance on randomness and control parameters by adopting a deterministic update strategy based on competitive interactions among all superior individuals. This ensures stable, consistent performance while preserving the balance between exploration and exploitation. The detailed motivation behind the algorithm is presented in Section II, while its working principle and computational framework are explained step by step in Section III.

The study design involves three major steps: (i) extraction of features from benchmark software defect datasets (PC5, JM1, KC1), (ii) application of the RP-Vathana algorithm for optimal feature selection, and (iii) classification of defect-prone modules using Naïve Bayes, with extensive comparisons against state-of-the-art metaheuristic algorithms across multiple independent runs.

The principal contributions of this study are:

- **A Novel Parameter-Free Algorithm**: We introduce RP-Vathana, a deterministic optimization algorithm for feature selection. This design effectively overcomes the inherent limitations of randomness and manual parameter tuning required in conventional methods.
- **Demonstrated Superiority**: We prove that RP-Vathana yields superior, more stable performance (higher classification accuracy and robustness) compared to current state-of-the-art algorithms on standard Software Defect Prediction (SDP) datasets.
- **New Direction in Optimization:** Our work highlights deterministic optimization as a promising methodological alternative to randomness-driven metaheuristics, offering new insights for SDP and broader machine learning applications.

II. RELATED WORKS

Software defect prediction (SDP) has been studied extensively, with applications in within-project [13–15], cross-project [16–18], and heterogeneous settings [19–20]. Most approaches employ supervised learning, where prediction accuracy depends heavily on the choice of features [21].

A. Feature Selection Approaches: Filter, Wrapper and Hybrid

Feature selection methods are generally classified as filter, wrapper, and hybrid. Filter methods apply statistical criteria, wrappers evaluate subsets using classifiers, and hybrids combine both. Alsaeedi et al. [26] reported that machine learning–based models generally outperform traditional filter-only approaches, while Aleem et al. [27] benchmarked 11 algorithms across 15 NASA datasets and found NB and SVM to be consistently effective. This indicates that careful feature subset selection remains a key factor for reliable SDP.

B Parameter-Dependent vs Parameter-Free Metaheuristics Approaches

Metaheuristic search has also been widely adopted for feature selection. Parameter-dependent methods (e.g., GA, PSO) require careful tuning and are often vulnerable to getting trapped in local optima, which limits their global search ability. Despite these drawbacks, they have shown effectiveness in domains such as image classification [28], tuberculosis detection [29], and SDP [30]. In contrast, parameter-free approaches reduce reliance on parameter tuning and improve robustness, though their application in SDP remains limited. This suggests an opportunity to explore parameter-free metaheuristics for more stable and scalable feature selection in SDP.

C Benchmarking Approaches and Limitations

Benchmarking studies play a central role in comparing SDP methods. Aleem et al. [27] demonstrated that NB and SVM often outperform other classifiers across datasets, while further evaluations confirm that SVM, MLP, and Bagging consistently achieve high accuracy (~89%), low error (MAE ≈ 0.10–0.14), and strong F-measure (~0.94). Conversely, KNN frequently underperforms, with accuracy below 75% and higher error rates. These findings reinforce the effectiveness of ensemble and margin-based methods, but also highlight the variability of results across datasets.

Beyond conventional classifiers, neuro-fuzzy techniques [22-24] and software agent–based approaches [25] have been explored to enhance test-case generation and improve defect detection. Hybridizing these approaches with machine learning methods offers potential for more accurate and adaptable prediction models.

However, benchmarking efforts face several limitations, including dataset imbalance, lack of diversity across repositories, and inconsistencies in evaluation protocols. Therefore, there is a need for more comprehensive benchmarking strategies that can ensure fair comparisons and generalizable conclusions.

D. Proposed Method: RP-Vathana Feature Selection

To address the limitations identified in prior studies, we propose RP-Vathana, a parameter-free optimization algorithm for feature selection in software defect prediction. Unlike parameter-dependent metaheuristics, RP-Vathana requires no control parameters, which enhances adaptability across diverse optimization tasks and eliminates the risk of poor performance due to improper tuning.

The algorithm is inspired by relay-race dynamics, where team members adjust their pace in response to one another. In this framework, candidate solutions cooperate and refine their search strategies relative to their peers, enabling the algorithm to escape local optima and move toward globally competitive solutions.

RP-Vathana has been applied to decision-metric selection in software defect prediction and has demonstrated effectiveness across three benchmark datasets, confirming its robustness and potential as a scalable feature selection approach.

E. Background and Motivation

In a relay race, teamwork and cooperation among team members are vital for success. Typically, teams follow a structured approach to maximize efficiency and speed. In an individual race, each runner focuses on matching the fitness

level of the fastest participant without assistance. However, in a relay race, teams consist of four participants who must collaborate and adjust their fitness levels based on each other's abilities. This means that a runner can receive support from their teammates even if they underperform at certain stages of the race.

To win a race, it's essential to adjust your pace according to the faster runners. You must increase your speed gradually, starting by overtaking those ahead of you. In many population-based optimization algorithms, individuals' positions are typically updated relative to the best individual in the population. However, in this research, a novel optimization algorithm called Relay Pursuit - Vathana (RP-Vathana) is introduced, which draws inspiration from relay races. In RP-Vathana, a relay race structure is utilised to enhance the position updates of weaker individuals within the population.

The remainder of the manuscript is organized as follows: Section III introduces the theoretical concept of the RP-Vathana algorithm, while Section IV details its implementation. Section V covers the experimental setup, parameter settings for benchmark algorithms, evaluation metrics, and an in-depth analysis of the results. Section VI addresses potential threats associated with our proposed algorithm, and Section VII and VIII provides the Limitations/Future scopes and conclusion respectively.

III. Working principle of Relay Pursuit-Vathana algorithm

This section describes the working principle of the algorithm followed by the mathematical model of the algorithm.

The algorithm begins by generating a random population of N individuals, each with a unique fitness evaluated using a Bayesian Information Criteria (BIC). These individuals are then sorted based on their fitness values in descending order. Next, the population is divided into 'm' groups, with each group containing an equal number of individuals. The highest-fitness individual overall is selected as the best individual, while the top N/m individuals in the sorted population become the leaders of each group.

Let us consider the initial population set $POP= (I_1, I_2, I_3,..I_N)$, where I_j represents j-th individuals. $Sorted\_POP = [I_{R1},I_{R2},..,I_{RN}]$ where I_{Rj} represents individuals with rank j. $POPgroup = [POP_1,POP_2,POP_3,..,POP_{N/m}]$ where POP_i represents the i-th population subset. Let N denote the population size and m the number of subsets, where the first m ranked individuals act as leaders. The population is partitioned into disjoint subsets based on modular arithmetic.

For each leader $l \in \{1,2,3,..,m\}$, the corresponding subset is defined as

$$s_l = \{i \in \{1,2,\dots,N\}: i \equiv l(mod\ m)\} = \left\{l + km : k = 0,1,2,\dots,\left\lfloor\frac{N-l}{m}\right\rfloor\right\}$$

In other words, if the individuals are denoted as, x_1,x_2,x_N sorted in the ascending order of the rank (where x_i corresponds to the i-th ranked individual), then the l-th subset can be expressed as

$$POP_l = \{x_i : i \in s_l\} = \left\{x_{l+km} : k = 0,1,\dots,\left\lfloor\frac{N-l}{m}\right\rfloor\right\}$$

For instance, if the population size is 20 and m is 4, each group will have 5 individuals. The individuals with rank *1* to *4* are selected as leaders. The subsets are organised such that individuals with rank *1,5,9,13,17* form the first subset, individuals with rank *2,6,10,14,18* form the second subset, and so on. The leader of each subset corresponds to the individual with the lowest index within that subset.

After organizing the individuals into groups, the velocity of each individual, except the leader candidate is update based on the velocity of best performing candidates within the same group. Let v_{ij} denote the velocity of the j-th variable in the i-th individual, and v_{kj} represent the velocity of the j-th variable of the k-th best candidate in the population g.

The velocity is updates using Eq.1 and Eq.2:

$$v_{ij}^g = v_{ij}^g + \gamma.\left(v_{kj}^g - v_{ij}^g\right) \quad (1)$$

$$x_{ij}^g = x_{ij}^g + v_{ij}^g \quad (2)$$

This process continues for a maximum number of iterations or until the fitness function converges. Subsequently, the candidate with the maximum fitness value is selected as the best candidate.

Algorithm 1 RP-Vathana optimization algorithm

```
procedure RP-VATHANA(featureVector)
    Design objective function
    Generate initial population set POP_1
    Calculate the fitness of each candidate using BIC
    x_Best: candidate x with best fitness in POP_1
    while T<T_Max do
        Sort all the candidates in POP_{T-1} in descending order
        for each Group g from 1 to m do
            for each candidate l from g to N do
                Append candidate X_l to group 'g'
                l:=l+m
            end for
        end for
        for each Group g from 1 to m do
            for each candidate i in the group g do
                X_new : Move the candidate X towards all the better
                candidates in the group 'g' using Equations 3 and 4
                X_New1 : Move X_New towards X_Best using equations 1 and 2
                Keep best among X,X_New,X_New1 in the POP_T
            end for
        end for
        Update X_Best in POP_T
    end while
    return X_Best
```

A. Mathematical model of RP-Vathana optimization:

This section describes about the mathematical model of proposed RP-Vathana optimization algorithm

1. Population initialization:

Assume that the size of the population is N, and each candidate has d number of variable. The i-th candidate is represented as :

$x_i=\{x_{i1},x_{i2},..,x_{id}\}$ where x_{ij} represents the j-th variable of the candidate

2. Fitness Evaluation:

The Bayesian Information Criteria is employed as the fitness function to evaluate the fitness of each candidate. This is a minimization problem so that the candidate with minimum value is selected as best candidate, represented as x_{best}.

3. Grouping and Movement:

Assume that the entire population set is divided in to m number of sub groups. Each group has *N/m* candidates. The initial population set is sorted in descending order and divided into m groups and candidates are selected based on their rank.

In each group the velocity of the candidates is updated based on the better performing candidate in the same group. For each individual x_i in each group g, the position of variable *j* is updated as follows.

$$v_{ij}^{g} = v_{ij}^{g} + \gamma . \sum_{k=1}^{K} \left(v_{kj}^{g} - v_{ij}^{g} \right) \quad (3)$$

$$x_{ij}^{g} = x_{ij}^{g} + v_{ij}^{g} \quad (4)$$

Where,

v_{ij}^{g} is the velocity of the j-th variable of the i-th candidate in the group 'g'

γ is a random value ranging from 0 to 1. The introduction of the random variable *r* in the modification step adds stochasticity to the algorithm. This stochastic element facilitates exploration of the search space, mimicking the inherent randomness observed in biological processes.

v_{kj}^{g} is velocity the j-th variable of the k-th better performing candidate in the group 'g'

x_{ij}^{g} is the j-th variable of the i-th candidate in the group 'g'

K is the number of better performing candidates in the group 'g'

In the Eq.3, if the difference between velocity of the j-th variable of candidate *i* and better candidate *k* is positive, encourages the individual *i* to move towards including the j-th variable. If it is negative, encourages the individual *i* to move towards excluding the j-th variable.

Repeat the steps 2 and 3 for maximum number of iterations. (T_{Max}) and the candidate with minimum BIC is selected as optimal candidate. Algorithm 1, shows the working principle of the RP-Vathana optimization algorithm.

Fig.1 shows the flow chart of the RP-Vathana algorithm. The candidates are ranked by the fitness values and the best candidate is selected as X_{best}. The population set is divided into m subsets. Candidates are placed into the subsets based on their ranks. First '*N/m* ranked candidates are set as leader of the respective subset and the leader candidates are moved towards the X_{best} candidate. All the remaining candidates *X* in the group g are moved towards the better performing candidate in g. If the new candidate X_{new} is better than *X* then X_{new} is moved towards X_{best} and kept as X_{new1}. The better candidate among *X*, X_{new} and X_{new1} is retained in the population set.

IV. Implementation of RP-Vathana optimization algorithm

This section describes the implementation details of the proposed algorithm

A. Data Preprocessing using RobustScaler

To mitigate the effect of outliers on feature scaling, we applied the RobustScaler method. Unlike standard scaling, which relies on the mean and standard deviation, RobustScaler uses the median and interquartile range (IQR) to transform the features. The scaling formula is given by:

$$x' = \frac{x - median(x)}{IQR(x)} \quad (5)$$

where *x* is the original feature value, median(x) is the median of the feature and $IQR(x) = Q_3 - Q_1$ represent the interquartile range. This approach ensures that the transformation is robust to outliers, improving the stability and performance of subsequent algorithms.

B. Population representation scheme

Each candidate in the population set is represented as binary feature set of size n_f where n_f is the number of design metrics in the software metrics data set. This feature set is represented as binary values (0,1) where the value of 1 represents the inclusion of the design variable and value of 0 represents its absence. Suppose if the size of feature set is 10 and the i-th candidate is represented as x_i= {1,0,1,1,0,0,1,1,0,1}. It indicates that the features *f1,f3,f4,f7,f8* and *f10* are present in the candidate solution.

C. Initialize the population table:

To generate the initial population set with a size specified by N, which is calculated as

$N = n_f \,.\, 0.10$, each candidate solution undergoes the following process:

Firstly, the continuous search space is transformed into a discrete search space using the sigmoid function. Random values are generated for each feature within the defined range of that feature. These randomly generated values represent the initial state of each feature in the solution. Subsequently, the sigmoid function is applied using Eq.6 to these random values, converting them into discrete values within the range of 0 to 1.

$$Sigmod(x) = \frac{1}{e^{-x} + 1} \quad (6)$$

Each candidate solution, denoted as xi, is represented as a vector $\{v_1, v_2, .. v_d\}$, where each v_j represents a random value generated for the j-th variable in the i-th solution, *r(i,j)*. Initially, the value of v_j is assigned randomly within the defined range of the variable *r(i,j)*. Subsequently, the sigmoid function is applied to map the value v_j to a value between 0 and 1. Finally, the Eq.7 is utilized to convert the resulting continuous value into a binary value.

$$x_{ij} = floor(Sigmod(x) + rand()) \; mod \; 2 \quad (7)$$

D. Fitness evaluation

The fitness of candidates is assessed using the BIC (Eq.8) as an objective function. BIC uses error rate and number of features used as the key for evaluating the performance of the candidate. Since it is a penalty-based evaluation method, the candidate which shows less error rate with minimum number of features gets a low penalty.

$$BIC = n \cdot \log MSE + k \cdot \log n \quad (8)$$

were n denotes the number of instances in the dataset, MSE represents the mean square error rate, k is the number of features.

MSE is calculated using the Eq.9

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad (9)$$

were N is the number of instances in the data set , y_i is the actual value of the instance i, $\hat{y}_i$ is the predicted value of the instance i.

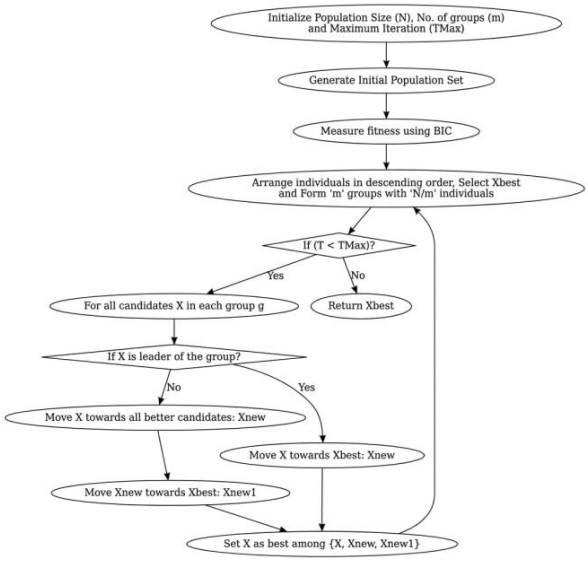


Fig.1 RP-Vathana optimization algorithm flow chart

V. EXPERIMENT AND VALIDATION

A. Experimental set up

Navie Bayes (NB) classifier are used for classification purpose. The model is trained for 80% of the samples and remaining samples are used for validation purpose. 5-fold cross validation process is used for classification. The entire samples are divided in to five equal folds. On each iteration, one fold is used for training the model and the remaining folds are used for testing. This process is repeated for all five possible combinations of training and testing folds. Highest accuracy achieved across the five iteration of 5-fold cross validation is selected as overall accuracy. To compare the effectiveness of the proposed approach, we evaluate it alongside several state-of-the-art algorithms, including Particle Swarm Optimization (PSO) [7], Boosted Whale Optimization (BWO) [8], Ant Colony Optimization (ACO) [9], Genetic Algorithm (GA) [10], Firefly Algorithm (FA) [11], and Hybrid Grey Wolf Optimization [12]. These algorithms are parameter-dependent, meaning their performance heavily relies on parameter values. Table.1 shows the parameters used for the benchmark algorithms to be compared.

Each algorithm is independently executed 30 times, and the average results are documented for analysis and comparison.

TABLE I
ALGORITHM SPECIFIC PARAMETER VALUES

Algorithm	Parameters	Value(s)
GA	C – Crossover method	One-point
	Pc – Crossover probability	0.5
	M – Mutation method	Swap
	PM – Mutation probability	0.3
PSO	W – Inertia weight	0.1
	C1 – Local learning coefficient	0.4
	C2 – Global learning coefficient	0.9
FA	α – Randomization parameter	0.1
	B0 – Base attraction	1
	Γ – Absorption coefficient	1
BWO	α – Linearity decreasing parameter	α ∈ [0,2]
	b – Constant	1
	l – Randomization parameter	[-1,1]
HGWO	α – Controlling parameter	α ∈ [0,2]

B. Performance metrics used:

The following well known metrics are used for evaluating the performance of the proposed RP-Vathana algorithm: accuracy, recall, precision and F1 score.

To calculate the above metrics True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) values are calculated from the predicted values of the testing instances.

TP indicates the number of correctly classified positive instances, TN indicates number of correctly classified negative instances, FP indicates number of negative instances classified positively and FN indicates number of positive instances classified negatively.

Accuracy is the percentage of samples that are correctly classified. (Eq.10)

$$ACC = \frac{(TP+TN)}{N} \qquad (10)$$

Precision is the ratio between the correctly classified positive values to the number of instances that are classified as positive. (Eq.11)

$$Precision = \frac{TP}{(TP+FP)} \qquad (11)$$

Recall is the percentage of positive values that are correctly classified.(Eq.12)

$$Recall = \frac{TP}{TP+FN} \qquad (12)$$

F1 score combines the precision and recall values (Eq.13)

$$F1 - Score = \frac{TP}{\left(TP + \frac{1}{2} * (TP + TN)\right)} \qquad (13)$$

C.Result and discussion

The effectiveness of the proposed software defect prediction method is assessed using three well-known benchmark datasets [PC5, JM1, KC1], which are accessible through the NASA-MDP public dataset repositories. Table 2 provides summaries of these benchmark datasets for reference.

TABLE II
NASA DATASET DESCRIPTION

Dataset	No. of features	Total no. of instances	No of defective instances	No of non-defective instances
PC5	39	17186	516	16670
JM1	21	10885	8779	2106
KC1	21	2109	326	1783

The comparison results are presented in Table 3, depicting the performance of our proposed method alongside other algorithms using NB classifier on the PC5 dataset. The proposed RP-Vathana achieves 94.28% accuracy with precision: 0.973, recall: 0.9347 and F1-Score:0.9504, which are comparatively better than other algorithms.

TABLE III
ACCURACY COMPARISON ON PC5 DATA SET USING NB CLASSIFIER

	ACC %	ERR	Precision	Recall
GA	85.79%	14.21%	0.9697	0.8579
PSO	88.71%	11.29%	0.9712	0.8771
FA	89.14%	10.86%	0.9715	0.8814
BWO	90.44%	9.56%	0.9736	0.9044
HGWO	92.94%	8.06%	0.974	0.9094
RP-Vathana	94.28%	5.72%	0.967	0.9109

Table 4 presents the comparative results for defect prediction rates such as accuracy, error rate, precision, recall, and F1-score on the JM1 dataset using NB classifier. Notably, the proposed feature selection method, RP-Vathana, has a macro-averaged F-score of 0.9142 for the NB classifier. Our RP-Vathana-based software defect prediction method achieves an impressive 93.69% classification accuracy. These results highlight the superiority of our proposed feature selection method on the publicly available JM1 dataset.

TABLE IV
ACCURACY COMPARISON ON JM1 DATA SET USING NB CLASSIFIER

	ACC %	ERR	Precision	Recall
GA	85.78%	14.22%	0.89	0.8578
PSO	86.70%	13.30%	0.8958	0.867
FA	87.13%	12.87%	0.8986	0.8713
BWO	89.94%	10.06%	0.9139	0.8944
HGWO	91.93%	8.07%	0.9243	0.9093
RP-Vathana	93.69%	6.31%	0.9177	0.9107

Classification accuracy, error rate, precision, recall, and F1-score are the evaluation metrics shown in Table 5 that are obtained from the proposed RP-Vathana method and other feature selection algorithms that use NB classifier on KC1 data set. Compared to other approaches, our proposed algorithm notably achieves high classification accuracy. In particular, RP-Vathana software defect prediction model outperforms other feature selection methods with a better success rate of 96.35%. Moreover, the proposed model performs better in terms of recall, F1-score, and macro-averaged precision.

TABLE V
ACCURACY COMPARISON ON KC1 DATA SET USING NB CLASSIFIER

	ACC %	ERR	Precision	Recall
GA	86.72%	13.62	0.9062	0.8672
PSO	88.63%	11.37%	0.9058	0.8663
FA	89.05	10.95%	0.9191	0.8905
BWO	91.97%	8.03%	0.9267	0.9037
HGWO	94.84%	5.16%	0.9488	0.9384
RP-Vathana	96.35%	3.65%	0.9795	0.9635

TABLE VI
PERFORMANCE METRICS COMPARISON OF PRO-POSED ALGORITHM AND BASELINE ALGORITHMS ACROSS 30 RUNS

		ACCURACY (%)				
		Maximum	***Minimum***	***Mean***	***Median***	***Std.Dev***
PC5 Data Set	***RP-Vathana***	94.28	89.32	91.43	91.34	1.27
	HGWO	92.94	85.73	89.63	89.36	1.9
	BWO	90.44	83.75	87.31	86.72	2.08
	FA	89.14	82.98	86.37	86.95	1.98
	PSO	88.71	83.89	86.62	86.7	1.28

	GA	85.79	81.58	83.81	83.95	1.59
JM1 Data Set	***RP-Vathana***	93.69	88.7	91.02	91.29	1.38
	HGWO	91.93	87.08	89.28	88.96	1.45
	BWO	89.94	85.22	87.51	87.59	1.52
	FA	87.13	82.14	84.75	84.99	1.61
	PSO	86.7	82.07	84.19	84.33	1.49
	GA	85.78	80.04	82.73	82.68	1.68
KC2 Data Set	***RP-Vathana***	96.35	92.57	94.25	94.17	1.16
	HGWO	94.84	90.69	92.85	92.83	1.18
	BWO	91.97	88.24	90.06	90.2	1.21
	FA	89.05	84.77	86.81	86.86	1.2
	PSO	88.63	84.46	86.09	85.95	1.21
	GA	85.93	81.94	83.88	83.72	1.3

Table 6 summarizes the maximum, minimum, mean, median, and standard deviation of the accuracy scores obtained from 30 runs of our proposed algorithm compared to other state-of-the-art algorithms. The proposed algorithm shows competitive performance with a higher mean accuracy and lower standard deviation, indicating both effectiveness and stability. The detailed statistical measures provide a clear and comprehensive understanding of how the proposed method performs relative to others.

TABLE VII
PERFORMANCE EVALUATION OF THE PROPOSED MODEL USING MCC AND F1-SCORE ACROSS ALL THREE DATASETS

Dataset	MCC	F1-Score
PC5	0.928	0.938
KC1	0.824	0.970
JM1	0.894	0.914

As summarized in Table 7, the proposed model consistently achieves high MCC and F1-score values across all three datasets. The MCC values indicate a strong correlation between the predicted and actual defect labels, while the F1-scores reflect a balanced trade-off between precision and recall. Together, these metrics confirm the model's robustness and effectiveness in handling the imbalanced class distributions present in the datasets.

A. Comparative Analysis of Selected Features with Existing Studies

The study utilized three benchmark datasets (PC5, JM1, and KC2) that feature a diverse set of software metrics. These metrics capture the essential characteristics of source code modules, including size (e.g., **LOC_TOTAL**), complexity (e.g.,**CYCLOMATIC_COMPLEXITY, DECISION_COUNT**), and maintainability as measured by Halstead metrics (e.g., **HALSTEAD_EFFORT**). Additional features, such as **PARAMETER_COUNT** and **GLOBAL_DATA_COMPLEXITY**, provide structural and modularity insights. These comprehensive metrics serve as the input features for the defect prediction model. To evaluate the effectiveness of RP-Vathana in identifying informative features, we compared the feature subsets selected from the PC5, JM1, and KC2 datasets against those reported in prior studies. Table VIII summarizes the top features obtained by RP-Vathana alongside rankings from existing feature selection approaches.

TABLE VIII
COMPARISON OF TOP SOFTWARE METRICS SELECTED BY RP-VATHANA AND PRIOR STUDIES ACROSS DATASETS

Dataset	Top Features (RP-Vathana)	Top Features (Prior Studies)
PC5	CYCLOMATIC_COMPLEXITY, HALSTEAD_EFFORT, PARAMETER_COUNT, ESSENTIAL_COMPLEXITY, LOC_EXECUTABLE	PARAMETER_COUNT, NUM_OPERANDS, NUM_OPERATORS, NUM_UNIQUE_OPERANDS, NUM_UNIQUE_OPERATORS, HALSTEAD_CONTENT, HALSTEAD_DIFFICULTY [33]
JM1	LOC_EXECUTABLE, DECISION_COUNT, HALSTEAD_DIFFICULTY, NODE_COUNT, MAINTENANCE_SEVERITY	LOC_TOTAL, MAINTENANCE_SEVERITY, ESSENTIAL_DENSITY, CYCLOMATIC_DENSITY, HALSTEAD_LEVEL [34]
KC2	BRANCH_COUNT, HALSTEAD_VOLUME, NUM_UNIQUE_OPERANDS, ESSENTIAL_COMPLEXITY, DESIGN_COMPLEXITY	NUM_OPERANDS, HALSTEAD_DIFFICULTY, HALSTEAD_EFFORT, LOC_EXECUTABLE, DECISION_COUNT [35]

Our findings show that RP-Vathana consistently emphasizes complexity and effort-related metrics, such as **CYCLOMATIC_COMPLEXITY, ESSENTIAL_COMPLEXITY, HALSTEAD_EFFORT**,

and **PARAMETER_COUNT**. This agrees with earlier studies that also highlighted the importance of Halstead and complexity measures for defect prediction

VI. VALIDITY THREATS

This section addresses the validity challenges posed to the proposed feature selection scheme. Although RP-Vathana feature subset selection method outperforms other algorithms in terms of performance, it requires more computational time to execute.

A. Complexity Analysis

The computational complexity of the proposed algorithm is derived by analyzing its major components as presented in Algorithm~1.

Initialization (Lines 2–5): The initial population of N candidates is generated and their fitness values are computed. This requires $O(N)$ operations.

Sorting (Line 7): In each iteration, the population is sorted according to fitness values. This requires $O\ (\ N \log N)$ time.

Group Assignment (Lines 8–13): Each candidate is assigned to one of the mmm groups. Since all NNN candidates are assigned exactly once, this step takes $O(N)$ time.

Candidate Updates (Lines 14–21): Within each group, every candidate is updated with respect to all better candidates in that group. For a group size of approximately N/m, each update requires $\frac{N}{m}$ operations. Since there are N candidates in total, the overall cost of updates per iteration is:

$$O\left(N \cdot \frac{N}{m}\right) = O\left(\frac{N^2}{m}\right)$$

Additional operations, such as updating with respect to the global best and performing comparisons, incur constant-time overheads and do not affect asymptotic complexity.

Iteration over TMax Cycles (Lines 6–23): The above steps are repeated for a maximum of TMax iterations.

By combining the above components, the total time complexity of RP-Vathana is given as:

$$O\left(T_{Max} \cdot \left\{N + N \log N + N + \frac{N^2}{m}\right\}\right)$$

For moderate values of m, the quadratic term $\frac{N^2}{m}$ dominates, and the worst-case asymptotic complexity simplifies to:

$$O(N^2 \cdot T_{Max})$$

This result highlights that the higher computational overhead of RP-Vathana arises primarily from its update mechanism, which considers interactions among all candidates rather than updating with respect to only the global best solution. Although this results in a higher runtime compared to conventional algorithms with time complexity $O(N \cdot T_{Max})$ that update only based on the global best, it enhances exploration capability and reduces the risk of premature convergence.

VII. LIMITATIONS AND FUTURE SCOPE

The parameter-free design of RP-Vathana ensures robustness but limits user control over exploration-exploitation balance. In highly multimodal landscapes, averaging influences from multiple candidates (Centroid Effect) can slow convergence and cause suboptimal solutions. Additionally, the need to aggregate forces from all candidates increases computational overhead, particularly in high-dimensional problems. Future work could address these issues via a dynamically weighted candidate pool to enhance exploration and convergence. Future work can focus on integrating RP-Vathana with other supervised learning models (e.g., SVM, Random Forest, MLP) to validate robustness and to optimize its time complexity for faster execution on large datasets. Hybrid feature selection strategies and selective parameter tuning could further enhance performance in challenging scenarios. To mitigate limitations such as the Centroid Effect and slow convergence in multimodal landscapes, a dynamically weighted candidate pool can be introduced. Additionally, benchmarking against emerging optimization and deep learning approaches can provide a more comprehensive evaluation of generalization and predictive capabilities.

VIII. CONCLUSION

In this study, we presented a novel software defect prediction model called RP-Vathana which is a parameter-free optimization algorithm. We showcased the superior performance of our proposed model with three benchmark datasets (PC-5, JM-1, and KC-1) through extensive experimentation and validation. Our results show that the software defect prediction model based on RP-Vathana consistently outperformed other meta-heuristic algorithms that require control parameters. One of the key strengths of our approach lies in its parameter-free nature, eliminating the need for parameter tuning. Furthermore, the robust performance of our model proves its potential for real-world applications in software defect prediction. In future work, our parameter-free RP-Vathana algorithm combined with auxiliary filter-based methods could potentially improve the accuracy of predictive models. Moreover, investigating parallel computing paradigms may optimise our algorithm's runtime performance considering the computational overhead it carries.

REFERENCES

[1] S. Hamayun, & L.F. Calvo-Flores, "Interpretable Software Defect Prediction from Project Effort and Static Code Metrics." *Applied Sciences*, vol. 14, no. 4, p. 52, 2024.

[2] G. Giray, K.E. Bennin, Ö. Köksal, Ö. Babur, & B. Tekinerdogan, "On the use of deep learning in software defect prediction". *The Journal of Systems & Software*, vol. 195, p. 111 537, 2023. DOI: 10.1016/j.jss.2022.111537.

[3] Z. Li, J. Niu, & X.Y. Jing, "Software defect prediction: future directions and challenges." *Automated Software Engineering*, vol. 31, no. 1, p. 19, 2024.

[4] Y. Jiang, B. Shen, & X. Gu, "Just-In-Time Software Defect Prediction via Bi-modal Change Representation Learning". *The Journal of Systems & Software*, p. 112 253, 2024. DOI: 10.1016/j.jss.2024.112253

[5] Karpagalingam Thirumoorthy, J. A., "Jerold John Britto feature selection model for software defect prediction using binary Rao optimization algorithm," *Applied Soft Computing*, vol. 131, p. 109 737, 2022, ISSN 1568-4946, DOI: 10.1016/j.asoc.2022.109737.

[6] Rohit Vashisht, Abhinav Juneja, Gagan Thakral & Sonam Gupta. "An empirical study of just-in-time-defect prediction using various machine learning techniques", *International Journal of Computers and Applications*, 2024. DOI: 10.1080/1206212X.2024.2328489

[7] R. Malhotra, N. Nishant, S. Gurha, V. Rathi, "Application of particle swarm optimization for software defect prediction using object oriented metrics," in: *2021 11th International Conference on Cloud Computing, Data Science Engineering (Confluence)*, pp. 88–93, 2021, http://dx.doi.org/10.1109/Confluence51648.2021.9377116.

[8] S.A.H. El-Kenawy, M.M. Eid, M.H.S. El-Shafai, A.M.H.E. Gamil, and D.S. El-Dahshan, "Hybrid binary whale optimization algorithm based on taper shaped transfer function for software defect prediction," *Comput. Mater. Continua*, vol. 78, no. 3, pp. 3177–3199, 2024.

[9] A. Hashemi, M.B. Dowlatshahi, "Exploring Ant Colony Optimization for Feature Selection: A Comprehensive Review." In: *Dey, N. (eds) Applications of Ant Colony Optimization and its Variants. Springer Tracts in Nature-Inspired Computing*. Springer, Singapore, 2024. DOI: 10.1007/978-981-99-7227-2_3

[10] M. Azzeh, R. Al-Sayyed, & F. Al-Tahir, "Software Defect Prediction Using Non-Dominated Sorting Genetic Algorithm and k-Nearest Neighbour Classifier." *e-Informatica Software Engineering Journal*, vol. 18, no. 1, p. 240 103, 2024.

[11] N. Monga and P. Sehgal, "A Framework of Software Defect Prediction using Machine Learning with Updated Firefly Algorithm and Neural Networks," *Journal of Information Systems Engineering and Managemen*t, vol. 10, no. 3, 2025.

[12] S. Sangeetha, & D. S.Rajakumari, "A Hybrid Genetic Based Grey Wolf Optimized Sophisticated Support Vector Machine (SSVM) Model for Software Defect Prediction". *Journal of Advanced Technology and Innovative Research*, vol. 101, no. 19, pp. 164–177, 2023. (Published late 2023, relevant to current work).

[13] H. Chen, X. Li, W. Shi, B. Zhang, & K. Wang, "Cross-Project Defect Prediction Using Transfer Learning with Long Short-Term Memory Networks." *Complexity*, 2024.

[14] A. Al-Sabaawi, B. Al-Khateeb, M.S. Al-Saeed, & N.N. Al-Taisan, "Software Defect Prediction Using an Intelligent Ensemble-Based Model". *IEEE Access*, vol. 12, pp. 3681–3694, 2024.

[15] B. Ghotra, S. McIntosh, A.E. Hassan, "Revisiting the impact of classification techniques on the performance of defect prediction models". In: *ICSE'15. IEEE*, pp. 789–800, 2015.

[16] T. Zimmermann, N. Nagappan, H. Gall, et al. "Cross-project defect prediction: a large-scale experiment on data vs. domain vs. process". In: *FSE/ESEC'09. ACM*, pp. 91–100, 2009.

[17] Z. Li, J. Niu, XY. Jing, et al. "Cross-project defect prediction via landmark selection-based kernelized discriminant subspace alignment." *IEEE Trans. Reliab*. vol. 70, no. 3, pp. 996–1013, 2021.

[18] Z. Li, H. Zhang, X.Y. Jing, et al. "Dssdpp: data selection and sampling-based domain programming predictor for cross-project defect prediction." *IEEE Trans. Softw. Eng*. vol. 49, no. 4, pp. 1941–1963, 2023

[19] X. Jing, F. Wu, X. Dong, et al. "Heterogeneous cross-company defect prediction by unified metric representation and CCA-based transfer learning." In: *FSE'15. ACM*, pp. 496–507, 2015.

[20] Z. Li, X.Y. Jing, X. Zhu, "Heterogeneous fault prediction with cost sensitive domain adaptation". *Softw. Test. Verif. Reliab*. vol. 28, no. 2, 1–22, 2018.

[21] D. Al-Fraihat, Y. Sharrab, A.R. Al-Ghuwairi, H. Alshishani and A. Algarni, "Hyperparameter Optimization for Software Bug Prediction Using Ensemble Learning," in *IEEE Access*, DOI: 10.1109/ACCESS.2024.3380024

[22] Ömer Faruk Arar, Kürşat Ayan, A feature dependent Naive Bayes approach and its application to the software defect prediction problem, *Applied Soft Computing*, vol. 59, 2017, pp. 197-209, ISSN 1568-4946, DOI: 10.1016/j.asoc.2017.05.043.

[23] S. Pandey, R. Mishra, A. Tripathi, "Software bug prediction prototype using Bayesian network classifier: A comprehensive model", *Procedia Comput. Sci*. vol. 132, 2018, pp. 1412–1421, http://dx.doi.org/10.1016/j.procs.2018.05.071.

[24] X. Rong, F. Li, Z. Cui, "A model for software defect prediction using support vector machine based on CBA", *Int. J. Intell. Syst. Technol. Appl*. vol. 15, 2016, p. 19, http://dx.doi.org/10.1504/IJISTA.2016.076102.

[25] S. Rathore, S. Kumar, "A decision tree logic-based recommendation system to select software fault prediction techniques", *Computing*, vol. 99, 2017, pp. 255–285, http://dx.doi.org/10.1007/s00607-016-0489-6.

[26] A. Alsaeedi, M. Khan, "Software defect prediction using supervised machine learning and ensemble techniques: A comparative study", *J. Softw. Eng. Appl*. vol. 12, 2019, pp. 85–100, http://dx.doi.org/10.4236/jsea.2019.125007

[27] S. Aleem, L. Capretz, and F. Ahmed, "Benchmarking Machine Learning Technologies for Software Defect Detection." *International Journal of Software Engineering & Applications*, vol. 6, pp. 11–23. DOI: 10.5121/ijsea.2015.6302

[28] M.R., RJ, K. "Invasive weed optimization with deep transfer learning for multispectral image classification model." *Multimed Tools Appl*., 2023. DOI: 10.1007/s11042-023-17429-9

[29] R.B. Jeyavathana and R. Balasubramanian, "Automatic detection of tuberculosis based on adaboost classifier and genetic algorithm", International Journal of Biomedical Engineering and Technology, vol. 36, no. 3, pp. 203–219, 2021.

[30] M. Anbu, G.S. Anandha Mala, "Feature selection using firefly algorithm in software defect prediction". *Cluster Comput* vol. 22, Suppl 5, pp. 10 925–10 934, 2019. DOI: 10.1007/s10586-017-1235-3

[31] K. Thirumoorthy, K. Muneeswaran, "Optimal feature subset selection using hybrid binary Jaya optimization algorithm for text classification", *Sadhan*., vol. 45, no. 1, 2020, pp. 1–13, http://dx.doi.org/10.1007/s12046-020-01443-w

[32] R. Rao, "Rao algorithms: Three metaphor-less simple algorithms for solving optimization problems." *International Journal of Industrial Engineering Computations*, vol. 11, no. 1, pp. 107–130, 2020.

[33] Y. Wang, A. Patel, & J. Thomas, "Selection of test case features using fuzzy entropy measure and Random Forest for software defect prediction." International Journal of Information and Systems, vol. 24, no. 3, pp. 85–96, 2023. DOI: 10.18280/isi.240306

[34] R. Sharma, M. Gupta, & P. Singh, "Analysis of important features in software defect prediction using SMOTE, RFE and Random Forest." *Journal of Software Engineering Research and Development*, vol. 12, no. 2, pp. 145–160, 2024. DOI: 10.1007/s40595-024-00123-x

[35] S. Lee, & H. Kim, "Feature importance analysis for NASA defect datasets using Random Forest and information gain." *Journal of Computer Science and Technology*, vol. 37, no. 5, pp. 1024–1038, 2022. DOI: 10.1007/s11390-022-2567-5

M. Rajakani is an Assistant Professor at SRM Institute of Science and Technology, Kattankulathur, Chennai. He received his Ph.D. in optimization techniques from Anna University. His research interests include Satellite Image Processing, optimization algorithms, machine learning, and intelligent communication systems. He has published several research articles in high-impact international journals and international Conferences.

R. Beaulah Jeyavathana is an Associate Professor at SRM Institute of Science and Technology, Kattankulathur, Chennai. She received her Ph.D. in Medical Image Analysis from MS University. Her research focuses on artificial intelligence-based medical image analysis, machine learning models for healthcare systems, and optimization techniques. She has authored several research articles published in reputed international journals and conference proceedings.

R. J. Kavitha is an Assistant Professor at University College of Engineering, Panruti. She received her Ph.D. in Wireless Communications from Anna University. Her research interests include signal processing, wireless communications, and machine learning. She is currently supervising five research scholars and has successfully guided four research scholars to completion under Anna University.

A New Deep Learning-Based Approach for IoT Task Offloading in Multi-access Edge Computing

Oussama Lagnfdi, Marouane Myyara, and Anouar Darif

***Abstract*—The exponential growth of Internet of Things (IoT) devices and the growing demand for resource-intensive applications have introduced significant challenges in computation, storage, and network efficiency. Although cloud computing provides partial relief, its centralized nature leads to unacceptable latency for delay-sensitive applications. Multi-access Edge Computing (MEC), especially with the advent of 5G, has emerged as a compelling solution by relocating computation closer to data sources, thereby reducing latency and improving responsiveness in applications such as smart agriculture, autonomous vehicles, augmented reality, and telemedicine. However, efficient workload offloading in MEC environments remains complex due to system heterogeneity, varying application requirements, and limited edge resources. This paper proposes a novel neural network-based approach to computation offloading in MEC, integrating workload allocation and resource management while accounting for application delay sensitivity, processing capacity, and communication constraints. The proposed model enables driving offloading decisions, adapting to fluctuating system states without relying on complex mathematical formulations. Simulation results demonstrate that the approach significantly reduces service time and enhances resource utilization, ensuring responsiveness for modern IoT applications. This research underscores MEC's potential to meet the rising computational and latency demands of next-generation IoT infrastructure.**

***Index Terms*—Multi-access Edge Computing Network, IoT, Task offloading, Deep Learning, Neural Network, Service time, Processing time**

I. INTRODUCTION

THE Internet of Things (IoT) is a rapidly developing ecosystem of distinct physical items connected through various wired and wireless networks [1]. It has enabled different industries to innovate new products, systems, and service offerings. This integration across sectors has led to a significant increase in computational load, especially in terminal devices that support AI-enabled functions [2]. Constraints such as manufacturing costs, limited battery capacity, and restricted processing capabilities prevent these devices from executing complex tasks efficiently.

With the advancement of high-speed internet and communication technologies, large volumes of data are generated by computation-heavy use cases such as augmented reality, online gaming, and video streaming [3]. This growth demands a platform that can effectively collect, manage, and process data from a growing number of IoT devices. An IoT device does not always have the computational resources to handle high-demand applications, both in terms of CPU power and memory size. Offloading these computations to more powerful devices becomes essential to meet communication and processing requirements [4].

The rise of Multi-Access Edge Computing (MEC), especially with 5G, has made it a key solution for offloading tasks from resource-limited IoT devices. By processing data closer to the source, MEC reduces latency and network congestion, which is crucial in low-bandwidth environments [5]. Unlike cloud computing, which introduces delays due to its centralized nature, MEC supports real-time, high-QoS applications more effectively [6]. However, MEC still faces challenges. Devices often depend on nearby edge servers due to limited local resources, and large-scale IoT deployments add complexity due to hardware and software heterogeneity [7]. Furthermore, the strict and varied resource needs of applications in fields like healthcare and smart agriculture make task migration more difficult.

Existing task offloading strategies using heuristic or rule-based methods often fail to adapt effectively to such complexity and real-time fluctuations in network and resource conditions. These approaches typically rely on static assumptions, which limit their scalability and responsiveness. In contrast, deep learning offers a data-driven solution that can model complex relationships between IoT task characteristics and available edge resources. However, there remains a lack of comprehensive approaches that apply deep learning models specifically to the joint optimization of task placement and execution decisions in MEC-enabled IoT systems. This paper proposes a deep learning-based offloading model that predicts the optimal execution point—local, edge, or cloud—for incoming IoT tasks based on system state and task attributes, including VM utilization, network delay, and workload. By continuously updating its decision policies according to the current environment, the model balances resource usage and improves performance under varying IoT loads.

The structure of this paper includes a review of related work (Section II), a detailed system model and problem formulation (Section III), the proposed deep learning-driven offloading strategy (Section IV), performance evaluation through simulation results (Section V), and a conclusion with future research perspectives (Section VI).

II. RELATED WORK

The Internet of Things (IoT) enables large-scale connectivity among heterogeneous devices [8], but limited computa-

The authors are with the LIMATI Laboratory, Department of Mathematics and Computer Science, Polydisciplinary Faculty, Sultan Moulay Slimane University, PO Box 592, Beni Mellal, 23000, Morocco. (E-mail: lagnfdi.o@gmail.com, marouane.myyara@usms.ac.ma, anouar.darif@gmail.com)

DOI: 10.36244/ICJ.2026.1.2

tional and energy resources require efficient task offloading. Mobile Edge Computing (MEC) brings computation closer to devices, reducing latency and improving Quality of Service (QoS). Surveys highlight the need for adaptive, scalable offloading strategies in edge and cloud environments [9].

Early task offloading solutions primarily relied on heuristic and rule-based approaches due to their low complexity. A representative example is the latency-classification-based deadline-aware (LCDA) task offloading algorithm proposed in [10], where tasks are classified as latency-sensitive or latency-tolerant based on deadlines. A weight-based formulation considering task size, execution time, and urgency is used to prioritize execution at edge servers. Similar heuristic-driven strategies focusing on reducing service time and improving QoS are presented in [11], [12]. Although efficient, these approaches depend on predefined rules and exhibit limited adaptability under dynamic IoT workloads. Several studies formulate task offloading as a mathematical optimization problem to achieve optimal decisions. In [13], a decentralized Lagrangian-based approach is proposed for online edge task scheduling, jointly optimizing latency, offloading cost, and resource utilization. Mixed-integer programming and greedy heuristics are adopted in [14] to maximize user service satisfaction in dense and delay-sensitive IoT environments. Joint optimization of computation offloading, software caching, and communication resources is investigated in [15], demonstrating notable latency and energy efficiency gains. Energy-aware optimization models are further explored in [16], [17], where power consumption and resource utilization are jointly optimized in cellular and NOMA-enabled IoT networks.

Game-theoretic approaches have been employed to model competition among IoT devices for limited edge resources. In [18], the task offloading problem is formulated as a non-cooperative game, where each device independently selects its offloading strategy to minimize latency. The existence of a Nash equilibrium enables decentralized decision-making without a central controller. However, such approaches may suffer from convergence delays and scalability issues as the number of devices increases. To address the NP-hard nature of task offloading, meta-heuristic and swarm intelligence methods have been widely adopted. Particle Swarm Optimization (PSO) is used in [19] to minimize execution delay and energy consumption in industrial IoT scenarios. Genetic algorithms are applied in [20] to optimize offloading decisions under bandwidth and computing constraints. Despite their effectiveness, these methods require iterative search processes, leading to high computational overhead and limited suitability for real-time deployment.

Control-theoretic frameworks have also been proposed to handle dynamic system behavior. Lyapunov-based optimization techniques are employed in [21], [22] to jointly manage task offloading and resource allocation while ensuring system stability and energy efficiency, particularly in UAV-enabled and energy-harvesting MEC systems. Although theoretically robust, these approaches often involve complex mathematical modeling, limiting practical implementation. Fuzzy-logic-based offloading mechanisms provide an alternative for handling uncertainty in MEC environments. In [23], a fuzzy workload orchestration framework is proposed to dynamically distribute workloads across edge nodes. A flexible fuzzy-based mobile edge orchestrator is further introduced in [24], enabling adaptive offloading decisions based on system context. Large-scale offloading challenges are also explored in [25], highlighting the need for scalable orchestration mechanisms.

Recent studies have explored hybrid and collaborative architectures to enhance MEC capabilities. Task offloading in cloud–edge collaboration environments is investigated in [6], while comprehensive discussions on challenges and future research directions are presented in [26]. These works emphasize the increasing complexity of modern MEC-enabled IoT systems.

Despite advances, most approaches rely on static heuristics or complex models, limiting scalability. Learning-based methods can predict offloading decisions from real-time system states. Motivated by this, we propose a neural network–based framework that uses task characteristics and VM utilization to enable adaptive, low-latency offloading and improved QoS.

III. SYSTEM MODEL AND PROBLEM FORMULATION

Figure 1 illustrates a multi-tier Mobile Edge Computing (MEC) architecture supporting a set of Internet of Things (IoT) devices. Let $\mathcal{I}$ denote the set of devices, where each device $i \in \mathcal{I}$ generates a set of tasks $\{\tau_{ij}\}$, with j indexing the tasks. Due to limited computational and energy resources, each task τ_{ij} must be offloaded to one of three execution tiers: (i) *local edge servers*, co-located with IoT devices and offering low latency but limited computing power; (ii) *remote edge servers*, distributed MEC nodes accessible via the network with moderate latency; or (iii) *cloud servers*, centralized data centers providing abundant computing resources at higher latency. Each task is executed entirely at a single tier, as determined by the offloading decision.

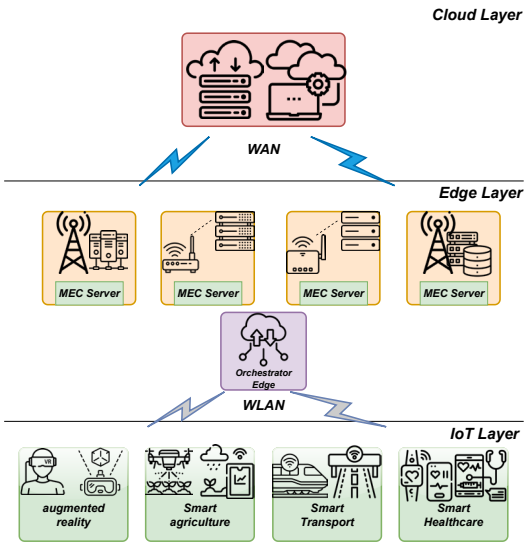


Fig. 1. An Overview of Multi-Access Edge Computing Systems

A. Base Notation

We consider a multi-tier Mobile Edge Computing (MEC) environment with multiple IoT devices, where each device generates tasks τ_{ij} characterized by computational workload L_{ij} (MI), input size D_{ij} (MB), maximum delay $T_{ij}^{\max}$ (ms), and memory requirement R_{ij}^{req} (MB). Tasks can be offloaded to three execution tiers: local edge servers ($k = 0$) close to devices with low delay but limited resources, remote edge servers ($k = 1$) with moderate latency, and cloud servers ($k = 2$) with high computational capacity but larger latency. Offloading decisions are denoted $x_{ij} \in \{0, 1, 2\}$. Each tier k is described by processing capacity f_k, bandwidth BW_k, maximum CPU $CPU_k^{\max}$, available memory R_k^{avail}, and current utilization U_k^{current}, with available capacity $f_k^{\text{avail}} = f_k(1 - U_k^{\text{current}})$.

B. Task and Offloading Model

The total execution time of a task τ_{ij} generated by device i depends on the selected offloading tier $k \in \{0, 1, 2\}$ (local edge, remote edge, or cloud) and comprises the service time and queuing delay:

$$T_{ij}^{\text{serv}} = T_{ij}^{\text{exec}} + T_{ij}^{\text{queue}}.$$

The service time T_{ij}^{serv} captures both computation and communication delays, influenced by processing capacity, bandwidth, and current utilization. For local execution ($k = 0$), no network transfer is needed, and the service time reduces to

$$T_{ij}^{\text{exec}} = \frac{L_{ij}}{f_0(1 - U_0^{\text{current}})}.$$

For remote edge ($k = 1$) and cloud ($k = 2$) execution, it includes both processing and network transmission delays:

$$T_{ij}^{\text{exec}} = \frac{L_{ij}}{f_k(1 - U_k^{\text{current}})} + \frac{D_{ij}^{\text{up}} + D_{ij}^{\text{down}}}{BW_k^{\text{avail}}},$$

where L_{ij} is the computational workload, f_k the processing speed, U_k^{current} the CPU utilization, D_{ij}^{up} and D_{ij}^{down} the input/output data sizes, and BW_k^{avail} the available bandwidth. While T_{ij}^{total} accounts for queuing delays.

C. Problem Formulation

The objective of the optimization is to **minimize the average service time** for all tasks generated by the IoT devices. Each task τ_{ij} can be processed locally, at a remote edge, or in the cloud, depending on the offloading decision $x_{ij} \in \{0, 1, 2\}$. The total service time includes both computation and transmission delays:

$$\min \frac{1}{N} \sum_{i,j} T_{ij}^{\text{serv}},$$

where N is the total number of tasks.

Subject to:

$$x_{ij} \in \{0, 1, 2\} \quad \forall \tau_{ij} \tag{C1}$$

$$T_{ij}^{\text{serv}} = \frac{L_{ij}}{f_{x_{ij}}(1 - U_{x_{ij}}^{\text{current}})} + \delta_{x_{ij}} \leq T_{ij}^{\max} \quad \forall \tau_{ij} \tag{C2}$$

$$U_{x_{ij}}^{\text{current}} + \frac{L_{ij}}{f_{x_{ij}} T_{ij}^{\max}} \leq CPU_{x_{ij}}^{\max} \quad \forall \tau_{ij} \tag{C3}$$

$$R_{x_{ij}}^{\text{avail}} \geq R_{ij}^{\text{req}} \quad \forall \tau_{ij} \tag{C4}$$

Here, $\delta_{x_{ij}} = 0$ for local execution ($x_{ij} = 0$) and $\delta_{x_{ij}} = \frac{D_{ij}^{\text{up}} + D_{ij}^{\text{down}}}{BW_{x_{ij}}^{\text{avail}}}$ for edge/cloud ($x_{ij} \in 1, 2$). Tasks are assigned sequentially in batches, updating system state ($U_k^{\text{current}}, R_k^{\text{avail}}$) after each task. Only offloading choices x_{ij} are decision variables. Constraints (C1)–(C4) enforce valid offloading, respect deadlines, and ensure CPU and memory capacities are not exceeded, allowing online allocation decisions to reflect the most recent resource availability. This design enforces sequential, state-aware allocation and aligns with the Problem Formulation. Overall, it minimizes the average service time across all tasks.

IV. Proposed Deep Learning-Driven Offloading Strategy

Figure 2 illustrates a Deep Learning (DL)-based offloading framework for a three-tier IoT–MEC–Cloud architecture. The system employs Neural Networks (NN) to optimize task allocation using real-time parameters such as link quality, task complexity, and latency requirements. The DL model dynamically selects the optimal execution tier—local edge, remote edge, or cloud—to **minimize service time**. By processing tasks in batches, the neural network efficiently handles resource contention and multi-user scheduling, ensuring high responsiveness and scalability for time-sensitive IoT applications, including autonomous systems and smart healthcare, while maintaining balanced load distribution across the infrastructure.

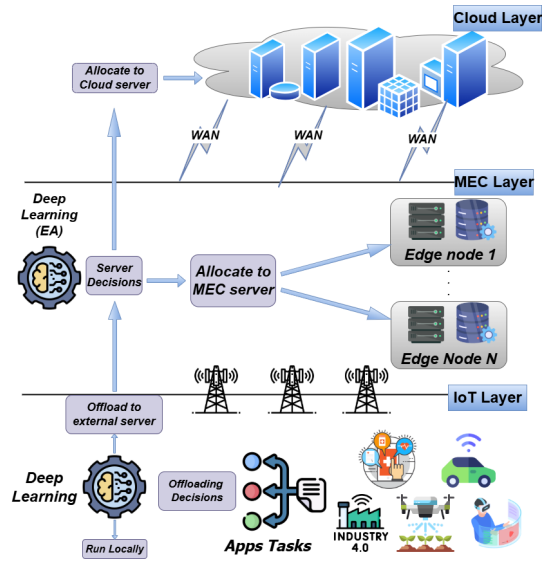


Fig. 2. An Illustration of Computation Offloading type in MEC networks

A. Neural Network Training Methodology

The task offloading neural network (NN) is trained in a supervised manner using datasets generated from diverse MEC simulation scenarios, varying the number of IoT devices, task arrival rates, CPU utilization, and network bandwidth. Each instance includes task-specific features (workload L_{ij}, input/output sizes $D_{ij}^{\text{up}}, D_{ij}^{\text{down}}$, delay sensitivity T_s) and system-level features (current VM utilization U_k^{current}, available memory R_k^{avail}, bandwidth BW_k^{avail}).

Training labels are generated using expert-defined rules:

- **Local Edge Assignment:** Tasks with high delay sensitivity ($T_s > 0.7$) and small workload ($L_{ij} < 10$ GI) are assigned locally if $U_0^{\text{current}} < 0.6$.
- **Remote Edge Assignment:** Tasks with moderate delay sensitivity ($0.3 \leq T_s \leq 0.7$) and medium workload ($10 \leq L_{ij} \leq 30$ GI) are assigned to the remote edge when $BW_1^{\text{avail}} > 50$ Mbps.
- **Cloud Assignment:** Tasks with low delay sensitivity ($T_s < 0.3$) or large workload ($L_{ij} > 30$ GI) are assigned to the cloud.

These rules ensure feasible and diverse training examples for the NN, capturing both task characteristics and dynamic system states. This process enables the NN to learn realistic offloading decisions for heterogeneous MEC scenarios.

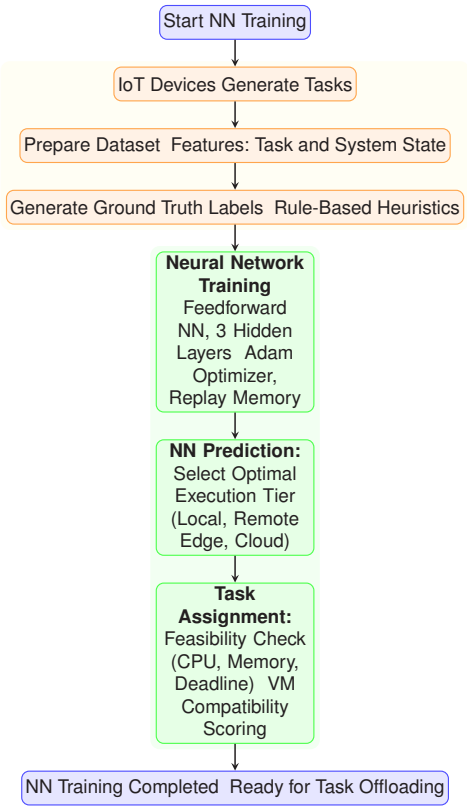


Fig. 3. Neural network-based IoT task offloading process in MEC, from dataset preparation to tier selection and task assignment.

The NN-based tier selection and VM assignment uses a feedforward network with three hidden layers (64, 32, 16 neurons, ReLU) and a softmax output for tier prediction. It is trained with Adam (learning rate 0.001), batch size 32, 10,000 samples, 100 episodes, 3,000 steps, discount 0.9, and soft replacement 0.01.

Algorithm 1 presents a hierarchical deep learning framework for tier selection and task-to-VM assignment in multi-tier MEC systems. Tasks are processed sequentially by delay sensitivity. For each task, the TierSelectionNN predicts the most suitable execution tier using task-specific features (workload, delay constraints, input/output sizes) and system-level parameters (available CPU, memory, bandwidth, and utilization). Within the selected tier, candidate VMs are scored by the VMCompatibilityNN to assess task–VM compatibility. The task is assigned to the highest-ranked feasible VM that meets CPU, memory, and delay constraints; otherwise, it is deferred to an overflow queue. The system state is updated after each assignment, ensuring efficient, feasible task placement and optimal resource utilization across heterogeneous MEC tiers.

Algorithm 1 DL-Based Tier selection and VM assignment

Require: Task batch τ_{batch}, system state S, TierSelectionNN, VMCompatibilityNN
Ensure: Task-to-(tier, VM) assignments and overflow queue

```
Sort τ_batch by delay sensitivity (ascending)
Initialize temporary state S_temp ← S
Initialize assignments ← ∅, overflow ← ∅
for each task τ in τ_batch do
    Extract task features F_τ and system features F_S
    Predict tier scores p ← TierSelectionNN(F_τ, F_S)
    Rank tiers by descending p
    assigned ← false
    for each tier k in ranked tiers do
        if tier k is not feasible then
            continue
        end if
        Extract VM set V_k in tier k
        Compute compatibility scores for all v ∈ V_k using VMCompatibilityNN
        Rank V_k by descending compatibility score
        for each VM v in ranked V_k do
            if v satisfies resource and deadline constraints then
                Assign τ → (k, v)
                Update S_temp
                assigned ← true
                break
            end if
        end for
        if assigned then
            break
        end if
    end for
    if not assigned then
        Add τ to overflow queue
    end if
end for
return assignments, overflow
```

V. PERFORMANCE EVALUATION

This section evaluates the performance of the proposed deep learning offloading framework in a MEC environment, modeling IoT devices, edge servers, and cloud infrastructure to assess scalability, responsiveness, and robustness under varying workloads.

A. Simulation Setup

The offloading strategy was evaluated using a Python 3.7 simulation framework with TensorFlow 2.6 and Tkinter 8.6 on Windows 10. Simulations ran on a PC with a 1.9 GHz Intel i5-8365U CPU and 16 GB RAM, emulating MEC-based task offloading with variable network, computation, and task conditions.

The simulated MEC system comprises 200–2000 IoT devices, each generating tasks at random intervals of 0.3–0.7 s. Task types—*Heavy*, *Infotainment*, *AR/VR*, and *Health*—represent diverse real-world applications, including video analytics, media streaming, augmented reality, and remote healthcare. Task parameters, summarized in Table I, are adapted from EdgeCloudSim to ensure realistic workload sizes, latency sensitivity, and data volumes.

TABLE I
APPLICATION CHARACTERISTICS

Property	Heavy	Infotainment	AR/VR	Health
Task Length (GI)	45	15	9	3
Delay Sensitivity (T_s)	0.1	0.3	0.9	0.7
Upload Data (KB)	2500	25	1500	20
Download Data (KB)	200	1000	25	1250

The infrastructure includes three MEC hosts (8 VMs each) and one cloud host (4 VMs), with VMs provisioned with fixed CPU cores and processing capacities (Table II). Network bandwidth varies between 10–100 Mbps to emulate real-time fluctuations. Each task is processed sequentially: the neural networks predict the optimal execution tier and score candidate VMs, and resource constraints (CPU, memory, and bandwidth) are updated dynamically and provided as input to the neural offloading model.

TABLE II
INFRASTRUCTURE PARAMETERS

Parameter	MEC	Cloud
Number of Hosts	3	1
VMs per Host	8	4
Cores per VM	2	4
VM CPU Speed (MIPS)	10,000	20,000
VM Storage (MB)	50,000	100,000

B. Simulation Results

This section evaluates and compares the performance of five resource management strategies—Neural Network, Fuzzy Logic, Utilization-Based, Sonmez, and Flores—across key quality of service (QoS) metrics under varying IoT device loads. In the Neural Network-based model, task execution is selected among edge (MEC) and cloud resources based on system state, task requirements, and network conditions.

- **Utilization-Based Approach** [24]: Tasks are offloaded to the least-loaded server to balance edge resources and prevent overutilization, but application-specific delays and communication demands are not considered.
- **Flores Approach** [25]: Uses fuzzy logic for offloading without considering resource consumption, which can overload VMs and increase latency under peak loads.
- **Sonmez Approach** [23]: Fuzzy logic-based offloading evaluates application features and resource use but assumes homogeneous resources, limiting adaptability in heterogeneous environments.
- **Fuzzy Logic Approach** [12]: Assigns tasks based on multiple criteria, including VM utilization, task length, network demand, and delay sensitivity, adapting to resource heterogeneity to reduce failures and improve service time.

The Neural Network-based algorithm considers task delay sensitivity, workload, and system resource utilization to optimize offloading and VM assignment. Compared to existing methods, it provides more efficient resource allocation and improved QoS for IoT applications.

C. Modeling of Service Time

Figure 4 illustrates the variation in service time as the number of IoT devices increases. At 200 devices, the Neural Network achieves the lowest service time of 0.738 s, outperforming Fuzzy Logic (1.5 s) and the Utilization-Based method (1.8 s). At 1000 and 2000 devices, it maintains 0.848 s and 0.963 s, respectively, while Flores degrades to 7.5 s and 12 s. Sonmez and Utilization-Based methods also increase to 4.0 s at 2000 devices. The Neural Network's consistent performance arises from its ability to learn from historical system data and dynamically predict optimal execution sites. By prioritizing low-latency tasks for local or MEC nodes, it minimizes wait times and queue lengths. Unlike Flores and Utilization-Based methods, it adapts in real-time to avoid overloaded nodes, ensuring lower and more stable service times across varying workloads.

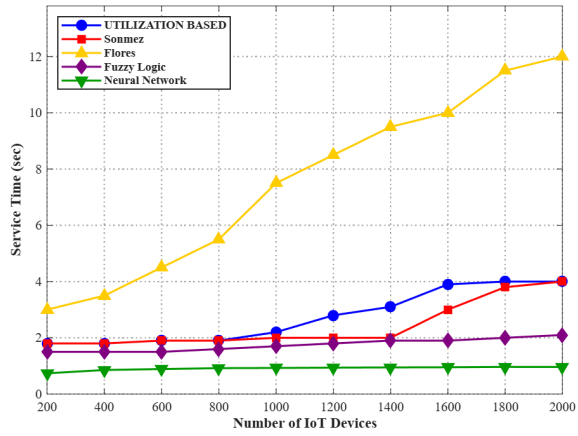


Fig. 4. The service time of all approaches VS the number of IoT devices

1) Modeling of Network Delay: Figure 5 illustrates the effect of increasing IoT devices on average network delay. For the Neural Network approach, the reported delay includes both transmission and queuing time, reflecting end-to-end latency. At 200 devices, all methods maintain delays below 0.21 ms. As device density increases, performance gaps widen: the Neural Network achieves 0.262 ms at 2000 devices, outperforming Fuzzy Logic (0.32 ms), Sonmez (0.35 ms), and Flores (0.35 ms), while the Utilization-Based approach remains stable at 0.20 ms. The Neural Network's advantage stems from its ability to learn real-time network conditions and predict optimal task placement. By recognizing congestion patterns, it dynamically offloads tasks closer to data sources, reducing transmission and queuing delays. Unlike static or rule-based methods, it adapts to system variations, ensuring consistently low latency and scalability across varying workloads.

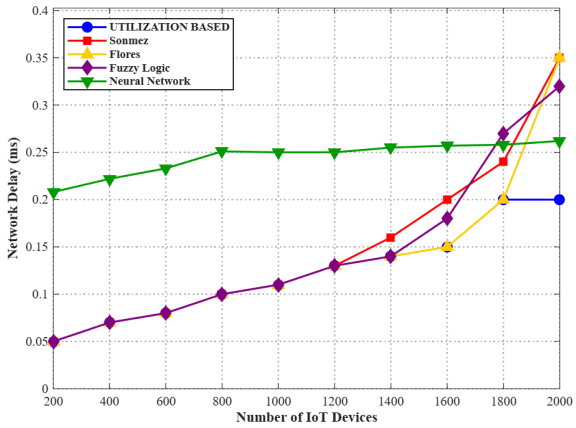


Fig. 5. The network delay of all approaches VS the number of IoT devices

2) Modeling of Processing Time: As illustrated in Figure 6, the Neural Network consistently achieves the shortest processing durations across all IoT device densities. At 200 devices, it records 0.530 s, outperforming Fuzzy Logic (0.6 s) and the Utilization-Based method (0.8 s).

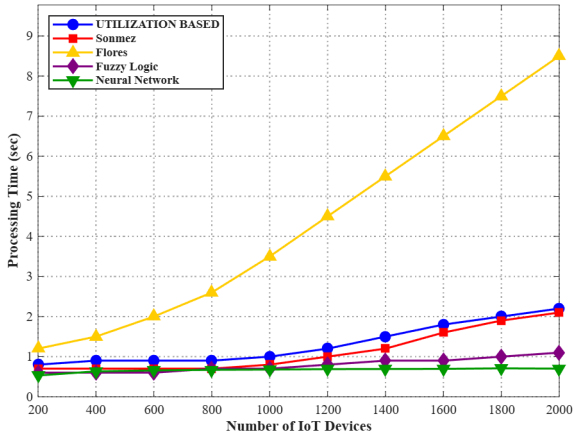


Fig. 6. The processing time of all approaches VS the number of IoT devices

The advantage becomes more pronounced at higher scales: at 2000 devices, the Neural Network maintains 0.701 s, while Flores escalates dramatically to 8.5 s, and Sonmez and the Utilization-Based method reach 2.1 s and 2.2 s, respectively. This performance stems from dynamic task allocation based on both node capacity and task complexity. By assigning tasks to nodes with optimal resources and processing capabilities, the system prevents CPU bottlenecks and reduces execution delays. Predictive analytics anticipate resource contention, allowing proactive redistribution of workloads before performance degradation occurs. Unlike static approaches such as Flores and Sonmez, which lack context-aware distribution, the Neural Network adapts in real time, balancing load effectively and ensuring efficient resource utilization even under high-demand scenarios, resulting in consistently low and stable processing times.

3) Modeling of Task Failure Rate: Figure 7 illustrates task failure rates across varying IoT device counts. The Neural Network consistently achieves the lowest rates, with 0.3% at 200–800 devices, 1.0% at 1200–1400, and 2.0% at 2000 devices. By comparison, Sonmez, Utilization-Based, and Flores reach 27%, 24%, and 22%, while Fuzzy Logic records 19%. This highlights the Neural Network's robust load-handling and fault-tolerance. It predicts potential overloads and avoids assigning tasks to resource-constrained nodes. Dynamic task reallocation using real-time feedback further reduces failures, preventing node saturation. Fixed-rule approaches lack such adaptability, often overcommitting resources and triggering cascading failures. The Neural Network maintains high reliability even as device density and system stress increase, making it particularly suitable for mission-critical IoT deployments that require consistent task completion and low failure rates.

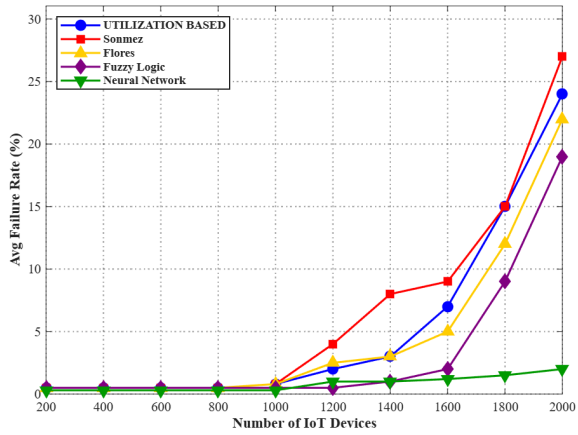


Fig. 7. Task Failure Rate of All Approaches vs. Number of IoT Devices

4) Modeling of VM Resource Utilization: Figure 8 shows VM CPU utilization patterns as IoT devices increase. At 200 devices, all methods exhibit low utilization (¡2.5%). At 2000 devices, the Neural Network remains efficient at 8.5%, whereas Sonmez reaches 97%, Flores stabilizes at 42%, Fuzzy Logic at 22%, and the Utilization-Based approach at 60%.

The Neural Network achieves this efficiency by learning from historical VM performance and using intelligent task allocation to balance loads. Reinforcement learning optimizes

resource use while avoiding VM overload. In contrast, static methods like Sonmez and Flores do not account for holistic VM usage or long-term resource trends, causing uneven distribution and bottlenecks. The Neural Network maintains balanced utilization across servers, supporting higher workloads and demonstrating scalability and efficiency for large-scale IoT deployments.

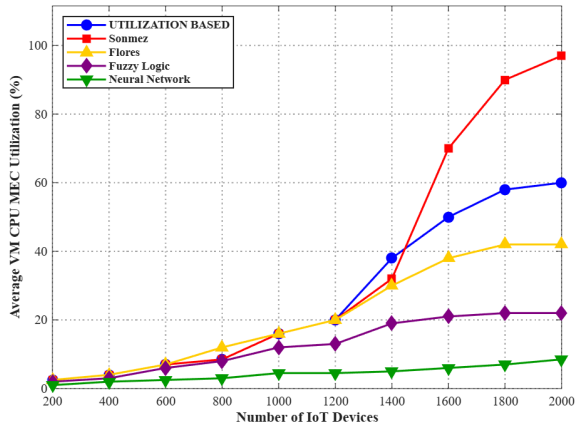


Fig. 8. Average CPU Utilization of MEC VMs of all approaches VS the number of IoT devices

Discussion

The experimental results demonstrate that the proposed Neural Network-based approach markedly outperforms traditional task offloading methods in Multi-access Edge Computing (MEC) environments for IoT applications. Across all evaluated metrics—including service time, network delay, processing time, task failure rate, and VM resource utilization—the Neural Network consistently achieves superior performance compared to Fuzzy Logic, Sonmez, Utilization-Based, and Flores approaches. Leveraging a learning-based architecture, the framework dynamically identifies the optimal execution location—local, edge, or cloud—based on real-time system conditions. This adaptability translates into reduced service and processing times, minimal network delays, and lower task failure rates under diverse workloads and device densities. In contrast to Flores, which exhibits limited scalability under high IoT loads, the Neural Network maintains robustness through intelligent VM management and real-time load balancing, achieving efficient resource utilization while sustaining high task completion rates.

VI. CONCLUSION

This study introduces a comprehensive neural network-based task offloading framework for Multi-access Edge Computing (MEC), addressing computational, networking, and scheduling challenges in large-scale IoT deployments. The proposed approach consistently outperforms traditional methods, including Fuzzy Logic and Utilization-Based strategies, by improving service times, task completion rates, and VM resource utilization. Its adaptive architecture dynamically manages fluctuating IoT workloads and heterogeneous application requirements, effectively balancing CPU, network, and latency demands while maintaining strong scalability. Future research will focus on predictive resource management, integrating workload forecasting and demand-aware scheduling for resource-intensive applications such as augmented reality, real-time gaming, and industrial IoT. Emphasis will also be placed on energy optimization for battery-powered devices and on integrating the framework into real-world IoT deployments. To validate simulation fidelity, we plan to reproduce baseline scenarios in Simu5G and EdgeCloudsim for per-task service time, network delay, and VM utilization comparisons, and to run small-scale Akraino IoT deployments to measure end-to-end latency and task completion. By combining intelligent learning mechanisms, predictive orchestration, and real-world validation, this framework aims to establish a foundation for efficient, reliable, scalable, and context-aware next-generation MEC-based IoT systems.

REFERENCES

[1] G. Paolone, D. Iachetti, R. Paesani, F. Pilotti, M. Marinelli, and P. Di Felice, "A holistic overview of the internet of things ecosystem," *IoT*, vol. 3, no. 4, pp. 398–434, 2022, DOI: 10.3390/iot3040022.

[2] B. Gupta and M. Quamara, "An overview of internet of things (IoT): Architectural aspects, challenges, and protocols," *Concurrency Computat.: Pract. Exper.*, vol. 32, no. 21, p. e4946, 2020, DOI: 10.1002/cpe.4946.

[3] Q. Chen, Z. Guo, W. Meng, S. Han, C. Li, and T. Q. Quek, "A survey on resource management in joint communication and computing-embedded SAGIN," *IEEE Commun. Surv. Tutor.*, 2024, DOI: 10.1109/COMST.2024.3421523.

[4] K. Bakar, F. Zuhra, B. Isyaku, and S. Sulaiman, "A review on the immediate advancement of the internet of things in wireless telecommunications," *IEEE Access*, vol. 11, pp. 21 020–21 048, 2023.

[5] X. Jin, W. Hua, Z. Wang, and Y. Chen, "A survey of research on computation offloading in mobile cloud computing," *Wirel. Netw.*, vol. 28, no. 4, pp. 1563–1585, 2022, DOI: 10.1007/s11276-021-02702-6.

[6] C. Wang, R. Guo, H. Yu, Y. Hu, C. Liu, and C. Deng, "Task offloading in cloud-edge collaboration-based cyber physical machine tool," *Robot. Comput.-Integr. Manuf.*, vol. 79, p. 102 439, 2023, DOI: 10.1016/j.rcim.2022.102439.

[7] S. Azizi, M. Othman, and H. Khamfroush, "Deco: A deadline-aware and energy-efficient algorithm for task offloading in mobile edge computing," *IEEE Syst. J.*, vol. 17, no. 1, pp. 952–963, 2023, DOI: 10.1109/JSYST.2022.3215789.

[8] A. Khanna and S. Kaur, "Internet of things (IoT), applications and challenges: A comprehensive review," *Wirel. Pers. Commun.*, vol. 114, no. 2, pp. 1687–1762, 2020, DOI: 10.1007/s11277-020-07363-y.

[9] K. Lone and S. A. Sofi, "A review on offloading in fog-based Internet of Things: Architecture, machine learning approaches, and open issues," *High-Confidence Comput.*, vol. 3, no. 2, p. 100 124, 2023, DOI: 10.1016/j.hcc.2023.100124.

[10] H. Choi, H. Yu, and E. Lee, "Latency-classification-based deadline-aware task offloading algorithm in mobile edge computing environments," *Appl. Sci.*, vol. 9, no. 21, p. 4696, 2019, DOI: 10.3390/app9214696.

[11] M. Myyara, O. Lagnfdi, A. Darif, and A. Farchane, "Enhancing QoS for IoT devices through heuristics-based computation offloading in multi-access edge computing," *Infocommun. J.*, vol. 16, no. 4, 2024, DOI: 10.36244/ICJ.2024.4.2.

[12] J. Almutairi and M. Aldossary, "A novel approach for IoT tasks offloading in edge-cloud environments," *J. Cloud Comput.*, vol. 10, p. 28, 2021, DOI: 10.1186/s13677-021-00241-3.

[13] Q. Peng, C. Wu, Y. Xia, Y. Ma, X. Wang, and N. Jiang, "Dosra: A decentralized approach to online edge task scheduling and resource allocation," *IEEE Internet Things J*., vol. 9, no. 6, pp. 4677–4692, 2022, DOI: 10.1109/JIOT.2022.3152310.

[14] J. Li, "Maximizing user service satisfaction for delay-sensitive IoT applications in edge computing," *IEEE Trans. Parallel Distrib. Syst*., vol. 33, no. 5, pp. 1199–1212, 2022, DOI: 10.1109/TPDS.2021.3107137.

[15] W. Wen, Y. Cui, T. Q. S. Quek, F.-C. Zheng, and S. Jin, "Joint optimal software caching, computation offloading and communications resource allocation for mobile edge computing," *IEEE Trans. Veh. Technol*., vol. 69, no. 7, pp. 7879–7894, 2020, DOI: 10.1109/TVT.2020.2993359.

[16] Y. Cheng, H. Zhao, and W. Xia, "Energy-aware offloading and power optimization in full-duplex mobile edge computing-enabled cellular IoT networks," *IEEE Sens. J*., vol. 22, no. 24, pp. 24 607–24 618, 2022, DOI: 10.1109/JSEN.2022.3218584.

[17] B. Liu, C. Liu, and M. Peng, "Resource allocation for energy- efficient MEC in NOMA-enabled massive IoT networks," *IEEE J. Sel. Areas Commun*., vol. 39, no. 4, pp. 1015–1027, 2021, DOI: 10.1109/JSAC.2020.3018809.

[18] J. Luo, Q. Qian, L. Yin, and Y. Yiao, "A game-theoretical ap- proach for task offloading in edge computing," in *Proc. 16th Int. Conf. Mobility, Sens. Netw. (MSN)*, 2020, pp. 756–761, DOI: 10.1109/MSN50589.2020.00129.

[19] Q. You and B. Tang, "Efficient task offloading using particle swarm optimization algorithm in edge computing for industrial internet of things," *J. Cloud Comput*., vol. 10, no. 1, p. 41, 2021, DOI: 10.1186/s13677-021-00212-8.

[20] M. Myyara, O. Lagnfdi, A. Darif, and A. Farchane, "A new approach based on genetic algorithm for computation offloading optimization in multi-access edge computing networks," *IAES Int. J. Artif. Intell*., vol. 13, no. 4, pp. 4186–4194, 2024.

[21] Z. Yu, Y. Gong, S. Gong, and Y. Guo, "Joint task offloading and resource allocation in UAV-enabled mobile edge computing," *IEEE Internet Things J*., vol. 7, no. 4, pp. 3147–3159, 2020, DOI: 10.1109/JIOT.2020.2965898.

[22] X. Song, Q. Ma, G. Zheng, L. Li, P. Cong, and J. Zhou, "Dynamic task offloading and resource allocation for energy-harvesting end-edge- cloud computing systems," *J. Syst. Archit.,* p. 103 469, 2025, DOI: 10.1016/j.sysarc.2025.103469.

[23] C. Sonmez, A. Ozgovde, and C. Ersoy, "Fuzzy workload orchestration for edge computing," *IEEE Trans. Netw. Serv. Manag*., vol. 16, no. 2, pp. 769–782, 2019, DOI: 10.1109/TNSM.2019.2911615.

[24] V. Nguyen, T. T. Khanh, T. D. Nguyen, C. S. Hong, and E. N. Huh, "Flexible computation offloading in a fuzzy-based mobile edge orchestrator for IoT applications," *J. Cloud Comput*., vol. 9, no. 1, pp. 1–18, 2020, DOI: 10.1186/s13677-020-0170-8.

[25] H. Flores, X. Su, V. Kostakos, A. Y. Ding, P. Nurmi, S. Tarkoma, P. Hui, and Y. Li, "Large-scale off loading in the internet of things," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops)*, 2017, pp. 479–484, DOI: 10.1109/PERCOMW.2017.7917638.

[26] M. Akhlaqi and Z. M. Hanapi, "Task offloading paradigm in mobile edge computing: Current issues, adopted approaches, and future directions," *J. Netw. Comput. Appl*., vol. 212, p. 103 568, 2023, DOI: 10.1016/j.jnca.2022.103568.

Oussama Lagnfdi received his B.Sc. in Physical Matter Science in 2020 and M.Sc. in Telecommunications Systems and Computer Networks in 2022 from Sultan Moulay Slimane University, Beni Mellal, Morocco. Currently, he is a Ph.D. candidate at the Laboratoire d'Innovation en Mathématiques et Applications et Technologies de l'Information (LIMATI), Polydisciplinary Faculty, Sultan Moulay Slimane University, Morocco. His ongoing research is focused on enhancing the performance of Internet of Things (IoT) and Mobile Edge Computing (MEC), Artificial Intelligence, Deep Learning, and Fuzzy Logic.

Marouane Myyara received his B.Sc. in Electronic and Telecommunication Engineering in 2019 and M.Sc. in Telecommunication Systems and Computer Networks in 2021 from Sultan Moulay Slimane University, Beni Mellal, Morocco. He is currently a Ph.D. candidate at the Laboratory of Innovation in Mathematics, Applications, and Information Technology (LIMATI), Polydisciplinary Faculty, Sultan Moulay Slimane University, Morocco. His current research focuses on improving the performance of Multi-access Edge Computing networks (MEC), Cloud Computing, Computation Offloading, and the Internet of Things (IoT).

Anouar Darif received the bachelor in IEEA (Informatique Électrotechnique, Électronique and Automatique) from Dhar El Mahraz Faculty of Sciences at Mohamed Ben Abdellah University Fez, Morocco in 2005. He received the Diplôme d'Etudes Supérieures Approfondies in Computer Sciences and Telecommunications from the Faculty of Sciences Rabat in 2007. He received the Ph.D. degree in Computer Sciences and Telecommunications from the Faculty of Sciences of Rabat in 2015. He is currently a Research and Teaching Associate in the Multidisciplinary Faculty at the University of Sultan Moulay Slimane Beni Mellal, Morocco. His research interests include Wireless Sensor Networks (WSN), Mobile Edge Computing (MEC), Internet of Things (IoT), Cloud Computing, and Neural Networks.

Enhancing Quantum State Transmission Fidelity through Quantum Orthogonal Frequency Division Multiple Access

Hussein Tuama[1,2], and Sándor Imre[1]

***Abstract*—In this paper, we propose quantum orthogonal frequency division multiple access (Q-OFDMA), a novel quantum communication scheme designed to overcome the fidelity limitations imposed by noise in multi-user quantum networks. Inspired by its classical counterpart, Q-OFDMA employs the quantum Fourier transform (QFT) and its inverse (IQFT) to encode and decode information across quantum channels. We evaluate our model under both a depolarization channel and a generalized noise model that interpolates between depolarizing and phase-damping noises. The simulation results conducted on Qiskit platform demonstrate that Q-OFDMA outperforms the reference model, achieving superior average fidelity across varying qubit counts and noise levels.**

***Index Terms*—Quantum Communication, State Fidelity, OFDMA, Quantum Fourier Transform, Quantum Channel.**

I. INTRODUCTION

QUANTUM communication promises revolutionary advancements in secure information transfers, distributed computing, and quantum networking. The fundamental unit of information in quantum communication is the quantum bit (qubit). Unlike classical bits, which are constrained to be either 0 or 1, qubits can be in coherent superpositions of these states [1]–[3]. Quantum communication is rooted in the fundamental principles of quantum mechanics, most notably entanglement [4]. This phenomenon allows for the instantaneous correlation of quantum states regardless of distances, forming the basis for communication protocols that can achieve unprecedented levels of security [5], [6].

Qubit-based communication technologies have attracted a great deal of interest in view of their ability to offer dramatic performance gains over classical communication schemes [7]. One striking potential of quantum computers is that they can generate algorithms that are exponentially faster for certain problems in computation, optimization, and search [8].In the networking side, a quantum internet is expected to provide communications that are fundamentally secure and more robust than current classical infrastructure [9].

Although quantum communication holds tremendous promise and potential, it has challenges, especially as related to managing error and noise. These errors are a significant obstacle to the realization of scalable quantum devices. Errors accumulate and are significantly detrimental to performance in complex quantum networks constructed by the interconnection of many components [10], [11]. These problems are exacerbated in a quantum multiple-access scenario, where multiple users share a quantum channel simultaneously. Noise in quantum communication systems is commonly due to gate imperfections and different channel noises [12]. The latter are fundamentally due to the inherent interactions between the quantum system and its environment. This procedure severely undermines the integrity and fidelity of quantum state transmission.

The general model of quantum communication is illustrated in Figure 1. The encoder prepares the quantum state in some encoding way according to the probability distribution of the source message. The input to a quantum channel is a quantum state, encoding information into a physical property. This quantum state travels through a quantum communication channel. To retrieve information, the quantum state must be measured at the receiver's end. The measurement outcome, which may be altered, depends on the quantum channel's transformation, which can be either fully probabilistic or deterministic. A quantum channel modifies information encoded in quantum states, such as the spin state of a particle, the ground and excited states of an atom, or other physical properties [13]. There are several important challenges in designing scalable quantum multiple-access systems. One of the main challenges is that quantum states from multiple users must be transmitted simultaneously, yet remain intact [14]. Similarly, quantum channels are also subject to inherent noise sources. In practical quantum state transmission, the ideal quantum communication channel will always stumble due to its environmental noise. Such interaction changes the transmission fidelity from the idealized to the practical.

There is a definition of fidelity, which provides a mathematical formality of how similar two quantum states are. This measure is actually useful in many experimental situations. Its main application is to verify quantum state preparation, which is always prone to noise and imperfections in the process. Here, fidelity quantifies how close the experimentally realized state is to the desired target state [15], [16]. It is a central issue in many fields related to quantum communications or quantum computing, in which quantum states need to be generated and transmitted with exact fidelity but are inevitably subject to decoherence and error.

The quantum Fourier transform and its inverse are the main parts that power our model as QFT and IQFT, respectively.

[1] Department of Networked Systems and Services, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Budapest, Hungary.
[2] College of Engineering, University of Misan, Amarah, Iraq.
[3] E-mail: {ahtuama, imre}@hit.bme.hu

DOI: 10.36244/ICJ.2026.1.3

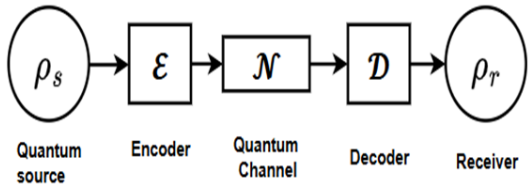


Fig. 1. Basic model of quantum communication.

QFT is an important quantum computing building block, and it is a key subroutine in well-known quantum algorithms, such as Shor's and phase estimation [17]–[20]. A unitary operation that is exponentially faster than its classical implementation, the Fast Fourier Transform (FFT) [21], [22]. It's involved in the encoding of our quantum states. The IQFT, on the other hand, is its mathematical opposite. We use it for decoding. It carries the data back from the frequency domain so we can read it properly. These two transforms are what make our Q-OFDMA functional together; they ensure that we can process and retrieve our data in an efficient manner. This work introduces a new quantum multiple-access method, which we term Quantum-Orthogonal Frequency Division Multiple Access (Q-OFDMA). This approach is motivated by the classical Orthogonal Frequency Division Multiple Access (OFDMA) widely used in LTE and 5G. Classical OFDMA uses approaches such as the Fast Fourier Transform (FFT) to apply an efficient transformation between the time and frequency domains, where, in turn, our Q-OFDMA system uses its theoretical quantum counterpart, the Quantum Fourier Transform (QFT). In our system, the QFT occupies a central role similar to that of the FFT in the standard classical OFDMA, enabling the unique advantages of our quantum-based model.

Our proposed model is a quantum communication scheme from the generation of quantum states to sending them over a quantum channel. While our model applies a centralized QFT operation acting on jointly prepared user qubits, such a model aligns well with centralized quantum repeater nodes or trusted node configurations commonly envisioned in the first-generation quantum internet. These centralized setups can support co-located quantum registers and enable multi-user access via global unitary transformations. Additionally, Q-OFDMA could complement hybrid classical quantum network control planes, serving as a quantum-layer multiplexing technique beneath classical routing. The fidelity of the transmitted quantum state is a key figure of performance; it is involved since this measurement provides the verification and the characterization of the output state of quantum communication. Its measurement is extensively used in quantum data processing, quantum engineering, and quantum machine learning.To evaluate our system's performance, we used two distinct quantum channel models to simulate noise. The first is a standard depolarization channel, which serves as a benchmark for uniform noise. We also considered a second, more complex scenario using a generalized noise model that interpolates between depolarizing and phase-damping noise [23]. This latter scenario was chosen to reflect the inherent uncertainty in noise characterization and allows for a tunable bias toward specific error types, providing a more realistic test of our system's robustness.

The rest of the paper is organized as follows: Section II reviews related work on quantum multiple-access schemes and fidelity analysis, establishing the context for our contribution. In Section III, we present the system model, which includes a description of Q-OFDMA and introduces the two quantum noise channels used for assessment. Section IV presents the mathematical framework for system fidelity under these noise models. To validate the performance of the Q-OFDMA model, the simulation results are introduced in Section V. Finally, Section VI concludes the paper by summarizing the findings and suggesting directions for future work.

II. Related Work

Building upon the foundational context outlined in the introduction, the following review synthesizes the most pertinent related work that contextualizes and motivates the present study. In wireless communication systems to provide several multiple access techniques for efficient usage of the radio spectrum [24]–[26]. These include Time Division Multiple Access (TDMA), Orthogonal Frequency Division Multiple Access (OFDMA), Code Division Multiple Access (CDMA), and Non-Orthogonal Multiple Access (NOMA) [27]. Different strategies that offer unique ways for sharing the scarce spectral resource among multiple users. TDMA allocates available time slots to different users so that collisions are avoided and access to the channel is sequential and organized. OFDMA achieves the diversity gain for multipath fading, and robust to co-channel interference by using differently allocated orthogonal subcarriers. CDMA, on the other hand, allows users to transmit at the same time over the entire spectral bandwidth by using user-specific spreading codes that spread signals in the code domain. The NOMA, as one of the recent techniques, can partially overlap some resources between users; following dividing power multiplexing and intelligent signal processing, it can decode superimposed signals, such as successive interference cancellation (SIC) [28], [29]. The main advantages of these multiple access techniques are efficient interference rejection, better spectrum utilization, and support for high Quality of Service (QoS). By allowing improved exploitation and sharing of the wireless channel, such techniques significantly increase the performance, reliability, and scalability of current communication systems.

The challenge of scalable, trust-free multi-user was addressed by [30], which provided an analysis of time and code division multiple access (TDMA/CDMA) protocols within a passive star network topology. This work is distinguished by its systematic comparison of these multiple access techniques specifically for the demanding low-photon-number regime of quantum key distribution (QKD) [31].In this context, authors in [32] propose a quantum code division multiple access (QCDMA) network architecture that leverages direct sequence spread spectrum (DSSS) techniques at the single-photon level. Their work is distinguished by its use of practical, commercially available optical components to construct an add-drop multiplexer capable of combining multiple single-photon channels into a single optical fiber. while the work in [33]

introduces a QCDMA communication system. Their model uses spectral phase encoding to apply pseudorandom barcodes to quantum light pulses.

In the pursuit of enhancing the scalability and key rates of QKD networks, orthogonal frequency-division multiplexing (OFDM) has emerged as a promising spectrally efficient multiplexing technique. Work in [34] demonstrated the feasibility of OFDM for frequency-coded QKD, showing that multiplexing multiple subcarriers could achieve a practical raw key rate across long-distance links, with orthogonal subcarriers and guard bands reducing noise and quantum bit error rate (QBER). Authors in [35] proposed two all-optical OFDM-QKD schemes for trusted-node quantum networks. Their work identified a key limitation of passive OFDM decoders offer no secret key rate gain due to loss. They suggested an active decoder with an optical switch. This design can linearly increase the key rate with subcarriers. Previous studies have explored quantum OFDM-like or frequency-multiplexed communication techniques; however, the proposed Q-OFDMA framework offers a fundamentally different operational and architectural perspective. Existing approaches, such as those in [19] and [20], primarily emphasize frequency-domain multiplexing-based transmission for individuals. In contrast, Q-OFDMA introduces a global QFT-based encoding that jointly processes all user qubits within the quantum register, enabling orthogonal multi-user access through a centralized multiplexing operation. This joint encoding approach distinguishes Q-OFDMA from earlier schemes by emphasizing scalable multi-user access under noisy conditions and by systematically evaluating performance using fidelity-based metrics rather than solely focusing on bit error rates.

Research on quantum state fidelity has been extensively analyzed under various noisy conditions. Authors in [36] highlight that when the entangled resource is subject to correlated noise, the performance degrades, although they derive analytical expressions showing a threshold memory coefficient beyond which teleportation remains viable. In another detailed analysis, authors in [37] demonstrate that the achievable fidelity and the range of teleportable states depend critically on both the type of noise and its point of introduction within the teleportation protocol.

While previous studies have demonstrated the benefits of quantum techniques in communication contexts, they often overlook scalable integration within orthogonal multi-user frameworks. To address this, the present paper proposes Q-OFDMA model using quantum encoding and QFT techniques. The study examines how system fidelity changes with different user counts and depolarizing noise. It focuses on channels with varying depolarization levels. The goal is to improve multi-user communication and support future progress in quantum communication systems.

III. System Description

This section describes in detail the Q-OFDMA model and its elements. We express the Q-OFDMA and a conventional reference model that is used as a benchmark of performance. The quantum channel is a basic barrier to information transmission based on our model. It is a source of error that degrades the fidelity of the transmitted state. Here, we introduce the common noise model illustrated by the depolarization channel model to serve as a baseline for unitary noise. For a more realistic analysis, we also made use of a generalized noise model that combines the effect of phase flip and depolarizing noises.

A. Q-OFDMA scheme

The Q-OFDMA model extends conventional OFDMA into the quantum domain by employing the quantum Fourier transform (QFT) and its inverse (IQFT) [38]. OFDMA functions as an essential multi-user communication technology for current broadband wireless networks, based on which computationally-efficient orthogonal subcarrier assignments are implemented. The main principle behind it is to divide the wideband radio spectrum into a number of narrowband sub-carriers, the frequencies of which must meet stringent orthogonality conditions [39]. It is implemented using DFT and IDFT operations. At the transmitter, user data streams undergo IDFT-based conversion to a time-domain signal via an inverse fast Fourier transform (IFFT), while the receiver employs FFT-based DFT processing to recover the transmitted signals [40], [41]. In contrast, our Q-OFDMA model replaces this classical framework with a quantum analogue that uses QFT and IQFT to encode and decode data transmitted through quantum channels. The QFT, as employed in quantum computing, transforms the amplitudes of quantum states, thereby enabling quantum-parallel processing of information.

The Q-OFDMA shown in Figure 2 depicts a simplified and typical quantum communication system with M users. It is composed of M quantum transmitters (QT) and M quantum receivers (QR). In this work, we assume the transmitters emit pure states $|\Psi\rangle$. Each user sends their quantum state into a QFT encoding operator unit. The QFT fundamentally differs from its classical DFT counterpart in its input–output representation, operational framework, and computational efficiency. While the classical discrete Fourier transform (DFT) processes a complex-valued vector $(x_0, x_1, \ldots, x_{N-1}) \in \mathbb{C}^N$ and produces a new vector $(y_0, y_1, \ldots, y_{N-1}) \in \mathbb{C}^N$ defined by the relation:

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \, \omega_N^{-jk}, \quad \text{for } k = 0, 1, \ldots, N-1, \tag{1}$$

where $\omega_N = e^{2\pi i/N}$ denotes a primitive N-th root of unity. The quantum Fourier transform (QFT) acts on a quantum superposition state:

$$|x\rangle = \sum_{j=0}^{N-1} x_j \, |j\rangle \, . \tag{2}$$

And transforms it into the state $\sum_{k=0}^{N-1} y_k |k\rangle$, where the amplitudes are given by:

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \, \omega_N^{-jk}, \quad \text{for } k = 0, 1, \ldots, N-1. \tag{3}$$

When the input is a basis state $|x\rangle$, the quantum Fourier transform can be expressed as:

$$\text{QFT} : |x\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \omega_N^{-xk} |k\rangle. \tag{4}$$

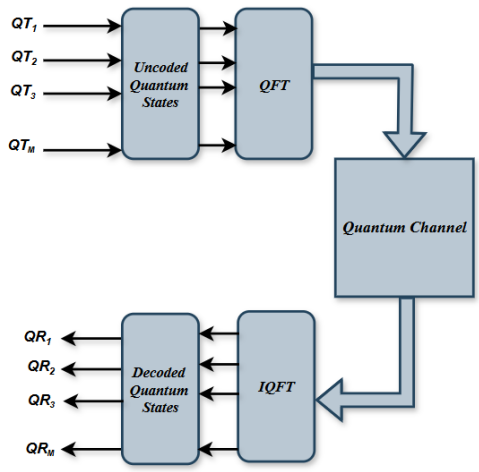


Fig. 2. quantum orthogonal frequency division multiple access scheme

The QFT output normalization includes a $1/\sqrt{N}$ factor to preserve unitarity and uses a complex exponential $\exp(2\pi ikj/N)$ consistent with quantum phase conventions. This enables the QFT to process $N = 2^n$ components simultaneously via quantum parallelism, achieving a gate complexity on the order of $O(\log^2 N)$ compared with the classical FFT's $O(N \log N)$, while preserving reversibility through unitary evolution. Following encoding, the quantum state propagates through the quantum channel, where it is inevitably subjected to a variety of quantum noise processes.

Our proposed Q-OFDMA framework can be also viewed as a centralized quantum multiplexing approach, where multiple users' data is prepared into a single quantum register. This register is processed using a QFT, which acts jointly across all qubits, enabling orthogonal allocation of the quantum spectrum to each user. The IQFT is then applied to recover each user's information with minimal interference.

B. Quantum Noise Model

In our analysis, we investigate two quantum channel models to simulate the effects of noise on quantum systems. This provides a comprehensive framework for understanding fidelity degradation in quantum communication systems. The first model is the depolarizing channel, which captures symmetric noise by replacing the input quantum state with a maximally mixed state with a specific probability, effectively eroding information uniformly across all basis states.

This channel is mathematically described by the quantum operation:

$$\mathcal{E}(\rho) = (1-p)\rho + p\frac{I}{2}, \tag{5}$$

where ρ represents the input density matrix, p is the depolarizing probability indicating the likelihood of the state being replaced by the maximally mixed state, and I is the identity operator, with the factor $I/2$ representing the maximally mixed state for a single qubit.

This model bridges a critical gap in quantum information degradation. Symmetric errors happen in all Pauli operators for isotropic noise. It offers a simple and solid tool to determine the resilience of a quantum system in the presence of unitary noise.

In the second model, we consider a scenario where we have introduced the phase-damping and depolarizing channel with random effects included to represent the uncertainty existing in practical experiments where the dominant noise mechanism could not be known exactly. This generalized noise model interpolates between phase-damping and depolarizing noise and is expressed as:

$$\mathcal{E}(\rho) = (1-p_x-p_y-p_z)\rho + p_x X\rho X + p_y Y\rho Y + p_z Z\rho Z, \tag{6}$$

where p_x, p_y, and p_z denote the probabilities of applying Pauli-X, Pauli-Y, and Pauli-Z errors, respectively, and the term $1 - p_x - p_y - p_z$ corresponds to the probability of the identity operation. The behavior of this noise model is highly flexible:

- When $p_x = p_y = p_z = p/3$, it reduces to the standard depolarizing channel, mimicking isotropic noise;
- When $p_x = p_y = 0$ and $p_z = p$, it simplifies to pure phase-damping noise, which primarily affects the phase of the quantum state (phase-flip errors);
- For intermediate values, the model allows a bias toward Z-type (phase-flip) errors while still permitting X-type (bit-flip) and Y-type (combined bit-and-phase-flip) errors, thereby capturing a broad spectrum of noise characteristics that may arise in real-world quantum implementations.

At the receiver, the IQFT is designed to reverse the effects of the QFT. Specifically, the IQFT maps a quantum state from the frequency-domain representation back to the computational-basis representation, yielding the original qubit information in the appropriate basis for subsequent measurement or further quantum processing. This reverse transformation is essential for interpreting the results of quantum computations that have been processed through the QFT. The final step in the Q-OFDMA system is a quantum process to evaluate system fidelity, which is described in the next section.

IV. SYSTEM FIDELITY

The fidelity of a quantum channel is a cornerstone for assessing the quality of quantum state transmission, quantifying the similarity between an initial state and its received counterpart after undergoing a noisy evolution. In this section, we analyze the fidelity for a bipartite quantum state under the two distinct noise models explained in the previous section. We consider a bipartite quantum state shared between the transmitted and received systems:

$$|\Psi\rangle = \sum_{ij} C_{ij}|\psi_i\rangle \otimes |\phi_j\rangle, \tag{7}$$

where C_{ij} are complex coefficients satisfying $\sum_{ij} |C_{ij}|^2 = 1$. The states $\{|\psi_i\rangle\}$ and $\{|\phi_j\rangle\}$ form orthonormal bases for the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ of dimensions d_A and d_B, respectively, yielding a total dimension $d = d_A d_B$.

The fidelity between the initial pure state $|\Psi\rangle$ and the output state $\rho_{\text{out}} = \mathcal{E}(|\Psi\rangle\langle\Psi|)$, where $\mathcal{E}$ is a quantum channel

(completely positive trace-preserving map), is given by the overlap:

$$F(|\Psi\rangle, \rho_{\text{out}}) = \langle\Psi|\rho_{\text{out}}|\Psi\rangle. \tag{8}$$

To evaluate the performance of a channel for communication, we often consider the average fidelity over an ensemble of input states. For an ensemble of pure states $\{|\psi_j\rangle\}$ occurring with probabilities $\{P_j\}$, the average fidelity is defined as:

$$\langle F\rangle = \sum_j P_j \langle\psi_j| \mathcal{E}(|\psi_j\rangle\langle\psi_j|) |\psi_j\rangle. \tag{9}$$

This measure is closely related to conditional information measures used in communication theory. The depolarizing channel is defined by its action on a density matrix ρ is given in equation 5.

For any transmitted pure state $|\psi_j\rangle$, the received state is $\sigma_j = \mathcal{E}_{\text{depol}}(|\psi_j\rangle\langle\psi_j|)$. The fidelity for a transmitted state is:

$$F(|\psi_j\rangle, \sigma_j) = \langle\psi_j| \left((1-p)|\psi_j\rangle\langle\psi_j| + p\frac{I}{d}\right) |\psi_j\rangle \tag{10}$$

$$= (1-p) + \frac{p}{d}\langle\psi_j|I|\psi_j\rangle \tag{11}$$

This indicates that the average fidelity under depolarizing noise depends solely on the error probability p and the system dimension d, a direct consequence of the channel's symmetry. For the second noise model, described in the equation 6. The total channel is the tensor product $\mathcal{E}_{\text{Pauli}} \otimes \mathcal{E}_{\text{Pauli}}$. The output state for the input $|\Psi\rangle$ is given by:

$$\rho_{\text{out}} = \sum_{k,l\in\{I,X,Y,Z\}} (p_k p_l)\,(k\otimes l)|\Psi\rangle\langle\Psi|\,(k\otimes l)^\dagger, \tag{12}$$

where $p_I = 1 - p_x - p_y - p_z$, and k, l are Pauli operators (including the identity I). The fidelity is then:

$$\begin{aligned} F &= \langle\Psi|\rho_{\text{out}}|\Psi\rangle \\ &= \sum_{k,l} p_k p_l \,\langle\Psi|(k\otimes l)|\Psi\rangle\langle\Psi|(k\otimes l)^\dagger|\Psi\rangle \\ &= \sum_{k,l} p_k p_l \,|\langle\Psi|(k\otimes l)|\Psi\rangle|^2 . \end{aligned} \tag{13}$$

To evaluate this expression, we analyze the expectation values $\langle\Psi|(k\otimes l)|\Psi\rangle$. The Pauli operators (excluding identity) are traceless and have eigenvalues ± 1. For a general state $|\Psi\rangle$ that is not an eigenstate of $k\otimes l$, this expectation value can be non-zero. However, a critical observation simplifies the calculation: the Pauli operators are unitary and the set $\{k\otimes l\}$ forms an orthogonal basis for operators on the Hilbert space. The expectation value $\langle\Psi|(k\otimes l)|\Psi\rangle$ represents the coefficient of the operator $k\otimes l$ in the Pauli transfer matrix representation of the state $|\Psi\rangle$. For many important states (e.g., maximally entangled states), these coefficients are non-zero only for specific k, l.

V. Simulation Results

In this section, we show the simulation results for the performance analysis of the proposed Q-OFDMA system. A thorough simulation investigation was performed on the Qiskit platform, which provides a reliable environment for modeling quantum circuits.

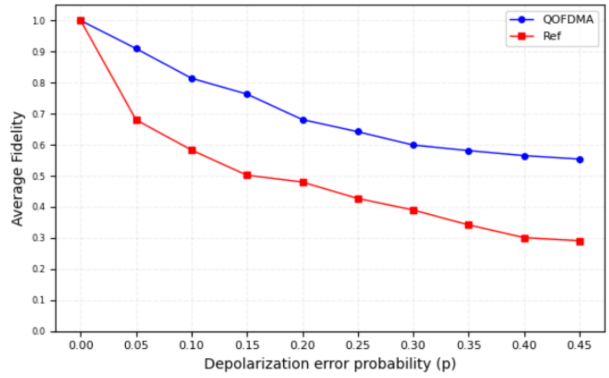


Fig. 3. Average fidelity through the depolarization channel with a fixed number of qubits.

In order to illustrate the benefits of the proposed system, a comparison was made between our Q-OFDMA model and a reference model. The Q-OFDMA model accounts for QFT before the quantum channel and IQT after it. On the other hand, in the reference model, we consider a standard scheme in which the transmission proceeds directly from the encoding stage to the channel without being converted via the QFT and IQFT. This comparative approach aims to highlight the unique characteristics of the Q-OFDMA model and evaluate its potential advantages in enhancing information fidelity compared to the conventional model.

In our work, we consider two different quantum channel models, highlighted in previous sections, which are used for simulating noise in quantum systems and offering an ample universe for understanding information loss in quantum communication systems.

In in first simulation scenario, we consider a depolarizing channel that acts as the benchmark simulating quantum information loss in an isotropic noise framework. In that model, errors are symmetric with respect to the possible Pauli operators, hence providing an easy but robust tool for the assessment of the resilience of a quantum system when subjected to uniform noise.

Figure 3 shows the performance comparisons between the proposed Q-OFDMA model and the reference model for the average fidelity, with respect to the depolarization error probability. The simulation was performed with a fixed number of qubits. It is found that in both systems, the average fidelity decreases as the probability of depolarization error increases, and this is an expected behavior as the channel noise becomes more intense. Nonetheless, the Q-OFDMA model always results in much better mean fidelity over the error probability for the given range of error probabilities that we have considered. For instance, the Q-OFDMA model here could produce an average fidelity of about 0.55 with a high error probability, whereas the reference model's fidelity is around 0.29. This great difference in performance indicates that Q-OFDMA can better tolerate depolarizing noise, illustrating the excellent performance of the introduced QFT-IQFT block to maintain quantum information fidelity in noisy channels.

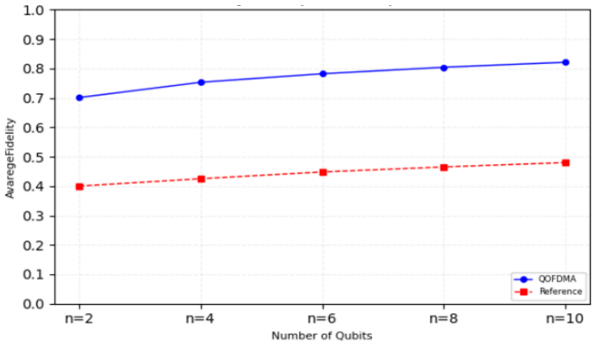


Fig. 4. Average fidelity through the depolarization channel with a fixed error probability

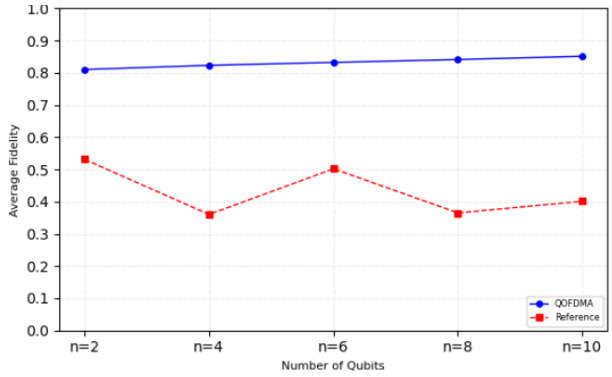


Fig. 5. Average fidelity through the second noise model

The QFT redistributes the information content of the quantum state across a superposition of basis states, preventing the concentration of information in individual qubits. Since depolarizing noise acts locally by introducing random bit and phase flips, this redistribution reduces the impact of such errors on the global quantum state. The subsequent IQFT coherently reconstructs the encoded information, leading to an effective averaging of noise effects and, consequently, higher average fidelity compared to the reference model.

Figure 4 shows the average fidelity of the Q-OFDMA model and its reference model versus the number of qubits n and the depolarization error probability fixed at 0.3. The findings confirm a definite course. For both systems, as the number of qubits increases, the average fidelity gets better.This is primarily because depolarizing noise acts locally on individual qubits, meaning each qubit is independently affected. The information is distributed across a higher-dimensional Hilbert space, which helps dilute the effects of noise. Furthermore, the use of the QFT and IQFT for encoding and decoding distributes quantum information across orthogonal basis states. This structured spreading and recovery process enhances the stability and resilience of quantum transmission. By dispersing the information throughout the quantum system, the scheme naturally averages out the effects of local noise, resulting in improved fidelity under noisy conditions.

In the second simulation scenario, we seek to test our system under more realistic, noisy conditions. This situation gives rise to a more complicated model that incorporates phase damping and depolarizing noise randomly to capture the practical uncertainties of an experimental quantum system, where we do not always have a clear idea of the dominant noise mechanism. The results shown in figure 5 indicate opposite behaviour of the two systems. The fidelity of the reference model shows an erratic behavior as a function of the number of qubits. This jittery characteristic reveals its great sensitivity to unpredictable noise types.By contrast, the QOFDMA model remains fairly consistent and consistently better than the reference model for any number of qubits. As the number of qubits increases, the fidelity of the QOFDMA system is gradually better and remains about 0.85 for n = 10. This stability demonstrates that the quantum Fourier transform (QFT)-based suppression strategy is robust against complex and uncertain noise environments. The approach is particularly valuable when the exact noise type is unknown but suspected to be phase-dominant. The underlying noise model incorporates randomness by representing errors as a tunable mixture of Pauli channels, with a bias that can be adjusted toward Z-type errors or toward a more balanced, depolarizing-like profile.

Evaluating the model's robustness across varied noise types is essential, as it confirms strong performance even under non-idealized, non-depolarizing conditions. This provides a practical and adaptable framework for real-world experiments where noise characteristics are often unknown. It would also apply when qubit counts are moderate. The model shows a consistent 30-35% fidelity advantage over conventional methods. This advantage holds across all tested conditions. It represents a significant advancement toward fault-tolerant quantum information processing.

While the proposed Q-OFDMA model demonstrates improved performance under simulated conditions, it is important to recognize its current hardware limitations. Implementing a global QFT operation over multiple user qubits requires a high degree of coherence across multi-qubit registers, which is technically demanding in today's noisy intermediate-scale quantum (NISQ) systems. Additionally, the increased gate depth introduced by QFT and IQFT operations can exacerbate gate noise accumulation, reducing fidelity in practical setups. Cross-talk between qubits and limitations in error correction further challenge scalability. Therefore, while the simulation results validate the model's theoretical performance, near-term physical realization would require significant advances in quantum hardware stability and error mitigation techniques.

VI. CONCLUSION

This work presents Quantum-Orthogonal Frequency Division Multiple Access (Q-OFDMA), a new approach to enhance the fidelity and scalability of quantum communication systems. By integrating the Quantum Fourier Transform (QFT) and its inverse (IQFT) for encoding and decoding, Q-OFDMA effectively mitigates the impact of noise and errors inherent in quantum channels. Simulations using the Qiskit platform

demonstrate that Q-OFDMA outperforms the reference model across various noise scenarios, including depolarizing and generalized phase-damping channels. The system exhibits superior average fidelity, reaching up to 0.85 under complex noise conditions, and shows improved resilience as the number of qubits increases.

The findings validate the efficacy of QFT-based techniques in addressing the challenges of quantum state transmission, paving the way for practical advancements in quantum networking and computing. For future work, it would be interesting to integrate quantum error correction codes, such as surface codes and tailored codes, into the Q-OFDMA framework to enhance system fidelity under high noise conditions.

ACKNOWLEDGMENTS

The research was supported by the European Union's Digital Europe Programme under Grant Agreement No. 101081247 (QCIHungary) and by the Parliamentary State Secretariat of the Ministry of Public Administration and Regional Development.

REFERENCES

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, Jun. 2012. [Online]. Available: http://dx.doi.org/10.1017/cbo9780511976667

[2] G. Arun and V. Mishra, "A review on quantum computing and communication," in *2014 2nd International Conference on Emerging Technology Trends in Electronics, Communication and Networking*. IEEE, Dec. 2014, p. 1–5. [Online]. Available: http://dx.doi.org/10.1109/et2ecn.2014.7044953

[3] L. Bacsardi, "Satellite communication over quantum channel," *Acta Astronautica*, vol. 61, no. 1–6, p. 151–159, Jun. 2007. [Online]. Available: http://dx.doi.org/10.1016/j.actaastro.2007.01.024

[4] S. Imre and M. Berces, "Entanglement-based competition resolution in distributed systems," *IEEE Access*, vol. 9, p. 10 253–10 262, 2021. [Online]. Available: http://dx.doi.org/10.1109/access.2021.3050271

[5] P. Zhang, N. Chen, S. Shen, S. Yu, S. Wu, and N. Kumar, "Future quantum communications and networking: A review and vision," I*EEE Wireless Communications*, vol. 31, no. 1, p. 141–148, Feb. 2024. [Online]. Available: http://dx.doi.org/10.1109/mwc.012.2200295

[6] L. Gyongyosi and S. Imre, "A survey on quantum computing technology," *Computer Science Review*, vol. 31, p. 51–71, Feb. 2019. [Online]. Available: http://dx.doi.org/10.1016/j.cosrev.2018.11.002

[7] J. Chen, "Review on quantum communication and quantum computation," *Journal of Physics: Conference Series*, vol. 1865, no. 2, p. 022 008, Apr. 2021. [Online]. Available: http://dx.doi.org/10.1088/1742-6596/1865/2/022008

[8] A. M. A. Sabaawi, M. R. Almasaoodi, S. El Gaily, and S. Imre, "Quantum genetic algorithm for highly constrained optimization problems," *Infocommunications Journal*, vol. 15, no. 3, p. 63–71, 2023. [Online]. Available: http://dx.doi.org/10.36244/icj.2023.3.7

[9] L. Gyongyosi and S. Imre, "Networked quantum services†," *Quantum Information amp; Computation*, vol. 25, no. 2, p. 97–140, May 2025. [Online]. Available: http://dx.doi.org/10.2478/qic-2025-0006

[10] A. Mihály and L. Bacsárdi, "Effects of selected noises on the quantum memory satellite based quantum repeaters," *Infocommunications Journal*, vol. 13, no. 2, p. 19–24, 2021. [Online]. Available: http://dx.doi.org/10.36244/icj.2021.2.3

[11] L. Gyongyosi, S. Imre, and H. V. Nguyen, "A survey on quantum channel capacities," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1149–1205, 2018.

[12] K. Bu, D. E. Koh, L. Li, Q. Luo, and Y. Zhang, "Effects of quantum resources and noise on the statistical complexity of quantum circuits," *Quantum Science and Technology*, vol. 8, no. 2, p. 025 013, 2023.

[13] S. Imre and L. Gyongyosi, *Advanced quantum communications: an engineering approach*. John Wiley & Sons, 2012.

[14] P. T. Kolusu and M. Anand, *Qubit Share Multiple Access Scheme (QSMA)*. Springer International Publishing, 2021, p. 91–104. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-79276-3_8

[15] S. T. Flammia and Y.-K. Liu, "Direct fidelity estimation from few pauli measurements," *Physical Review Letters*, vol. 106, no. 23, Jun. 2011. [Online]. Available: http://dx.doi.org/10.1103/physrevlett.106.230501

[16] T. Tanizawa, Y. Takeuchi, S. Yamashika, R. Yoshii, and S. Tsuchiya, "Fidelity-estimation method for graph states with depolarizing noise," *Physical Review Research*, vol. 5, no. 4, Dec. 2023. [Online]. Available: http://dx.doi.org/10.1103/physrevresearch.5.043260

[17] D. Jiang, X. Liu, H. Song, and H. Xie, "An survey: Quantum phase estimation algorithms," in *2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, vol. 5, 2021, pp. 884–888.

[18] T. Monz, D. Nigg, E. A. Martinez, M. F. Brandl, P. Schindler, R. Rines, S. X. Wang, I. L. Chuang, and R. Blatt, "Realization of a scalable shor algorithm," *Science*, vol. 351, no. 6277, pp. 1068–1070, 2016. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.aad9480

[19] A. M. A. Sabaawi, M. R. Almasaoodi, and S. Imre, "Advancing quantum communications: Q-ofdm with quantum fourier transforms for enhanced signal integrity," in *2024 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. IEEE, Sep. 2024, p. 1–6. [Online]. Available: http://dx.doi.org/10.23919/softcom62040.2024.10721632

[20] M. R. Almasaoodi, A. M. A. Sabaawi, and S. Imre, "Quantum ofdm: A novel approach to qubit error minimization," in *2024 14th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*. IEEE, Jul. 2024, p. 53–58. [Online]. Available: http://dx.doi.org/10.1109/csndsp60683.2024.10636648

[21] M. Rezai and J. A. Salehi, "Fundamentals of quantum fourier optics," *IEEE Transactions on Quantum Engineering*, vol. 4, pp. 1–22, 2023.

[22] M. R. Almasaoodi, A. M. A. Sabaawi, S. E. Gaily, and S. Imre, "New quantum genetic algorithm based on constrained quantum optimization," *Karbala International Journal of Modern Science*, vol. 9, no. 4, Oct. 2023. [Online]. Available: http://dx.doi.org/10.33640/2405-609x.3325

[23] S. Dutta, S. Banerjee, and M. Rani, "Qudit states in noisy quantum channels," *Physica Scripta*, vol. 98, no. 11, p. 115 113, Oct. 2023. [Online]. Available: http://dx.doi.org/10.1088/1402-4896/ad0006

[24] H. Mathur and T. Deepa, "A survey on advanced multiple access techniques for 5g and beyond wireless communications," *Wireless Personal Communications*, vol. 118, no. 2, p. 1775–1792, Jan. 2021. [Online]. Available: http://dx.doi.org/10.1007/s11277-021-08115-w

[25] Y. Cai, Z. Qin, F. Cui, G. Y. Li, and J. A. McCann, "Modulation and multiple access for 5g networks," *IEEE Communications Surveys Tutorials*, vol. 20, no. 1, pp. 629–646, 2018.

[26] A. S. Mohamed and E. Udvary, "Mode selection in mode division multiple access system for in building solution in mobile networks," *Infocommunications Journal*, vol. 17, no. 2, p. 83–88, 2025. [Online]. Available: http://dx.doi.org/10.36244/icj.2025.2.10

[27] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5g," *IEEE Communications Surveys Tutorials*, vol. 20, no. 3, pp. 2294–2323, 2018.

[28] Y. Yu, X. Fang, J. Mao, L. Zhang, X. Gao, D. Ding, and L. Liu, "Novel noma system for optical communications based on ofdm/oqam," in *2019 18th International Conference on Optical Communications and Networks (ICOCN)*, 2019, pp. 1–3.

[29] A. M. A. Sabaawi, M. R. Almasaoodi, and S. Imre, "Exploiting ofdm method for quantum communication," *Quantum Information Processing*, vol. 23, no. 7, Jun. 2024. [Online]. Available: http://dx.doi.org/10.1007/s11128-024-04465-z

[30] M. Razavi, "Multiple-access quantum key distribution networks," *IEEE Transactions on Communications*, vol. 60, no. 10, pp. 3071–3079, 2012.

[31] L. Gyongyosi, L. Bacsardi, and S. Imre, "A survey on quantum key distribution," *Infocommunications Journal*, no. 2, pp. 14–21, 2019. [Online]. Available: http://dx.doi.org/10.36244/icj.2019.2.2

[32] V. Sharma and S. Banerjee, "Quantum communication using code division multiple access network," *Optical and Quantum Electronics*, vol. 52, no. 8, Aug. 2020. [Online]. Available: http://dx.doi.org/10.1007/s11082-020-02494-3

[33] M. Rezai and J. A. Salehi, "Quantum cdma communication systems," *IEEE Transactions on Information Theory*, vol. 67, no. 8, pp. 5526–5547, 2021.

[34] M. Anandan, S. Choudhary, and K. Pradeep Kumar, "Ofdm for frequency coded quantum key distribution," in *2012 International Conference on Fiber Optics and Photonics (PHOTONICS)*, 2012, pp. 1–3.

[35] S. Bahrani, M. Razavi, and J. A. Salehi, "Orthogonal frequency-division multiplexed quantum key distribution," *Journal of Lightwave Technology*, vol. 33, no. 23, pp. 4687–4698, 2015.

[36] Y.-n. Guo, Q.-l. Tian, K. Zeng, and P.-x. Chen, "Fidelity of quantum teleportation in correlated quantum channels," *Quantum Information Processing*, vol. 19, no. 6, May 2020. [Online]. Available: http://dx.doi.org/10.1007/s11128-020-02675-9

[37] S. Oh, S. Lee, and H.-w. Lee, "Fidelity of quantum teleportation through noisy channels," *Physical Review A*, vol. 66, no. 2, Aug. 2002. [Online]. Available: http://dx.doi.org/10.1103/physreva.66.022316

[38] Y. Wang, L. Li, P. Zhang, and Z. Liu, "Dft-based channel estimation with symmetric extension for ofdma systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, no. 1, Dec. 2008. [Online]. Available: http://dx.doi.org/10.1155/2009/647130

[39] P. M. A. Shah, S. S. Qureshi, R. A. Butt, S. Mahdaliza Idrus, and J. Mirza, "Design and analysis of 5g network architecture with orthogonal frequency division multiple access based passive optical network," *Optical Fiber Technology*, vol. 67, p. 102 678, Dec. 2021. [Online]. Available: http://dx.doi.org/10.1016/j.yofte.2021.102678

[40] C.-M. Chen, C.-C. Hung, and Y.-H. Huang, "An energy-efficient partial fft processor for the ofdma communication system," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 57, no. 2, pp. 136–140, 2010.

[41] A. Sadat and W. Mikhael, "Fast fourier transform for high speed ofdm wireless multimedia system," in *Proceedings of the 44th IEEE 2001 Midwest Symposium on Circuits and Systems*. MWSCAS 2001 (Cat. No.01CH37257), vol. 2, 2001, pp. 938–942 vol.2.

Hussein Tuama Al-kinani Ph.D. student at the Department of Networked Systems and Services at the Budapest University of Technology and Economics (BME). He obtained his master's degree in telecommunication engineering from the University of Pune, India, in 2015. He obtained his bachelor's degree in electrical engineering from Basrah University in 2011 in Basrah, Iraq. His current research interests include wireless communication, quantum computing, quantum communication, quantum computing-based algorithms, and telecommunication systems.

Sándor Imre [M'93] professor and Head of the Department of Networked Systems and Services at the Budapest University of Technology and Economics (BME). He obtained dr. univ. degree in probability theory and statistics in 1996, a Ph.D. degree in 1999, and DSc degree from the Hungarian Academy of Sciences in 2007. He was elected a corresponding member of HAS in 2019. He has acted as supervisor in the High-Speed Networks Laboratory since 1999. He is a member of the Doctoral Council of HAS. He was invited to join the eMobile Innovation Center of BME as an R&D director in 2005. His research interests include mobile and wireless systems, quantum computing, and communications. Especially, he has contributions to different wireless access technologies, mobility protocols and their game-theoretical approaches, reconfigurable systems, quantum communication, computing-based algorithms, and protocols.

Beyond JA4+: Flow Statistics vs. TLS Fingerprinting for Encrypted Malware Detection

Márton Pál Lipcsey-Magyar, Attila Ármin Madarász, and Adrian Pekar

***Abstract*—The deployment of Encrypted Client Hello (ECH) challenges TLS fingerprinting, a widely used approach for encrypted malware detection, by encrypting the handshake fields these methods rely on. This paper presents a systematic evaluation of flow-based statistical features as a handshake-independent alternative to fingerprinting. Through validation against the official JA4+ implementation, we establish limitations in fingerprinting approaches for this corpus: only 64.9% of malware families possess unique signatures, placing an inherent ceiling on achievable recall in our evaluation. We evaluate flow-level features—packet counts, timing patterns, and size distributions—across 27 experimental configurations on a dataset of 16,542 flows spanning 101 families (59 malware and 42 benign applications). Random Forest classifiers using combined flow statistics and sequential packet length features achieve 98.11% F1-score for binary malware detection with 97.22% recall, substantially exceeding fingerprinting's theoretical recall bound of 64.9%. For fine-grained family identification, we obtain 54.81% macro F1 across 101 classes and 48.71% macro F1 for malware-only attribution, demonstrating that flow-based methods retain meaningful discriminative power where fingerprinting abstains. Across all tasks, Random Forest consistently outperforms neural networks and k-NN, with performance gaps widening in complex multiclass scenarios. These findings highlight flow-based classification as a practical and reproducible approach that can help maintain network security visibility as ECH deployment progresses, showing that behavioral traffic patterns are expected to provide durable signals for detection even as handshake fields become encrypted.**

***Index Terms*—JA4+ fingerprints, malware classification, flow statistics, encrypted client hello, TLS fingerprinting, network security**

I. INTRODUCTION

The analysis of encrypted network traffic is a critical component of modern network security. With most internet traffic now protected by TLS, the ability to identify malicious activity without relying on payload inspection has become essential. One widely deployed approach is TLS fingerprinting, which leverages the fact that different software applications and malware families often produce distinctive handshake signatures. A recent example is the *JA4+ suite* [1], which derives fingerprints from the Client Hello (*JA4*) and Server Hello (*JA4S*), optionally enriched with metadata such as the Server Name Indication (SNI). By making classification decisions at the very beginning of the connection, such fingerprints enable *early detection* and near real-time blocking.

Despite its utility, TLS fingerprinting faces two fundamental challenges. First, fingerprinting is limited by signature coverage: numerous distinct malware families share identical TLS handshakes, and only a subset are uniquely identifiable [2], [3]. Second, fingerprinting assumes visibility into the plaintext handshake. This assumption is being challenged by the deployment of Encrypted Client Hello (ECH), which encrypts the inner Client Hello while exposing only an outer, potentially generic version. As major browsers adopt ECH and CDNs facilitate rollout, the visibility on which client-side fingerprinting relies is expected to diminish [4], [5]. Together, these trends highlight the need for complementary approaches that are less sensitive to changes in TLS protocol visibility.

Flow-level statistical features offer one such alternative. By focusing on observable characteristics such as packet counts, timing patterns, and size distributions, prior work has shown that encrypted malware traffic can be distinguished from benign traffic without access to payloads or handshake fields [6], [7]. These features remain largely observable under encryption enhancements such as ECH, making them a more durable basis for traffic analysis.

In this paper, we systematically evaluate flow-level statistical features as an ECH-resilient complement to TLS fingerprinting. Using a validated malware dataset containing 16,542 flows across 101 families, we demonstrate that flow-based classifiers provide robust detection in binary tasks and meaningful discriminative power for multi-family attribution, substantially exceeding the theoretical recall limits of TLS fingerprinting.

Our contributions are as follows:

- A systematic evaluation of flow-based features for ECH-resilient malware detection, demonstrating that flow statistics achieve 98.11% F1 in binary detection without requiring handshake inspection—a capability expected to persist as ECH adoption renders traditional fingerprinting increasingly limited.
- A quantitative comparison framework establishing theoretical upper bounds for TLS fingerprinting (precision $\approx$99.62%, recall $\leq$64.9%, F1 $\leq$78.6%) and demonstrating that flow-based methods exceed fingerprinting's recall bound by 32+ percentage points (97.22% vs. 64.9%).
- Feature importance analysis across all three tasks revealing task-dependent patterns—packet sizes dominate binary detection while temporal features dominate family attribution—with all top features remaining ECH-resilient

M. P. Lipcsey-Magyar, A. Á. Madarász, and A. Pekar are with the Budapest University of Technology and Economics, Hungary
(e-mail: lipcsey-magyarmartonpal@edu.bme.hu, madarasz.attila@edu.bme.hu, apekar@hit.bme.hu).
M. P. Lipcsey-Magyar, A. Á. Madarász, and A. Pekarare also with HUN-REN Office for Supported Research Groups, Hungary.
A. Pekar is also with CUJO LLC, Budapest, Hungary.

DOI: 10.36244/ICJ.2026.1.4

and cross-validation (5-fold, F1 std <0.002 for binary) confirming result stability.
- A validated pipeline where JA4/JA4S extraction conforms to the official JA4+ suite (99.98% flow matching; 100% hash conformance) and a feature-rich dataset of 16,542 flows across 101 families with 89 properties enabling reproducible research.

The remainder of this paper is organized as follows. Section II reviews related work in encrypted traffic analysis and TLS fingerprinting. Section III describes our methodology, including dataset validation against official JA4+ implementations, flow feature extraction, and the classification framework across three tasks and three algorithms. Section IV presents experimental results for binary malware detection, multiclass family identification, and comparative analysis with TLS fingerprinting baselines. Section V discusses operational implications for ECH-resilient detection, explains performance patterns, and identifies limitations and future research directions. Section VI concludes the paper.

II. Related Work

Network security in encrypted environments has driven extensive research into methods for identifying malicious traffic without payload inspection. We organize existing work into three primary categories: TLS fingerprinting approaches, statistical flow-based detection methods, and hybrid contextual techniques.

A. TLS Fingerprinting Methods

TLS fingerprinting leverages unique signatures generated during the TLS handshake process for encrypted traffic analysis. The seminal JA3 method [8] extracts fingerprints from Client Hello messages by hashing TLS version, cipher suites, extensions, elliptic curves, and formats. Complementing this, JA3S fingerprints capture Server Hello responses, enabling bidirectional analysis that significantly improves malware detection accuracy over unidirectional approaches.

Recent advances culminated in the JA4+ suite [1], [9], which addresses several limitations of JA3 through improved hash stability and expanded coverage. Matoušek et al. [10] conducted a comprehensive evaluation of JA4+ fingerprints across 64 malware families, reporting that combining JA4, JA4S, and SNI achieves 87% uniqueness with approximately 80% family coverage on their dataset. However, they also report residual overlap between malware and benign fingerprints (e.g., certificate hashes) and cases where distinct families share identical handshake signatures.

The reliability challenges are further exposed by Matoušek et al. [11] in their analysis of mobile applications, where JA3 fingerprints alone prove insufficient due to high collision rates among different applications. Similarly, Siwakoti and Rawat [12] highlight that while machine learning can extend beyond known fingerprint signatures, achieving 96.4% accuracy in their evaluation, the fundamental coverage problem persists—only a fraction of malware families possess unique TLS signatures.

Beyond JA3/JA4-style techniques, recent work explores broader fingerprinting strategies. Ede et al. [13] proposed *FLOWPRINT*, a semi-supervised fingerprinting framework that demonstrated how mobile apps with distinct behaviors may nonetheless collide under identical TLS handshakes, underscoring coverage limitations. On the server side, Theofanous et al. [14] presented *Fingerprinting the Shadows*, which unmasked malicious servers behind CDNs and proxies. Their work shows that while advanced server fingerprinting remains feasible, visibility is increasingly constrained by encrypted handshake fields and middleboxes.

B. Statistical Flow-Based Detection

An alternative paradigm leverages statistical characteristics of network flows, offering potential resilience to encryption enhancements. Anderson and McGrew [15] pioneered this approach by combining TLS metadata with DNS and HTTP contextual features, achieving 99.978% accuracy on their dataset via contextual correlation across protocols. Their "data omnia" philosophy demonstrates that unencrypted metadata can effectively identify malicious encrypted communications.

Piskozub et al. [16] advanced purely flow-based detection with MalAlert, which aggregates flows into "flowsets" and extracts 441 statistical features for malware classification. Their approach achieves 90% F1-score for malware type identification in their experiments while maintaining privacy through byte-level statistics rather than content inspection. Crucially, this method operates independently of TLS handshake visibility.

Yeo et al. [17] explored deep learning approaches using convolutional neural networks on 35 flow statistical features, achieving over 85% accuracy across multiple malware families in their evaluation. Their comparison with traditional machine learning methods (Random Forest, SVM) showed that both CNN and Random Forest exceed 93% performance across all evaluation metrics on their dataset.

Barut et al. [7] provided a comprehensive performance study comparing machine learning and deep learning approaches on flow features, uniquely addressing computational efficiency alongside accuracy. Their analysis revealed that traditional machine learning methods often outperform deep learning while requiring fewer computational resources, with acceleration libraries providing up to 68.6× speedup.

Building on these foundations, Fu et al. [18] recently demonstrated that unknown encrypted malicious traffic can be detected in real time using flow statistics and side-channel features, without relying on handshake visibility. Collectively, these works show that flow-based methods can achieve strong accuracy with modest computational overhead, making them practical for real-time monitoring deployments.

C. Hybrid and Advanced Approaches

Recent work has explored techniques that combine multiple signal sources or advanced pattern recognition. Kim et al. [19] revisited Markov chain-based fingerprinting as an alternative to ClientHello parsing, achieving 88.1% accuracy for malware family classification in their experiments. Their

approach offers advantages including resilience to ECH and generalizability for approximate matching.

Yu and Won [20] provide a comprehensive taxonomy of encrypted traffic fingerprinting methods, categorizing approaches into fingerprint collection techniques and AI-based classification methods. Their analysis identifies the growing need for hybrid approaches that combine the speed of traditional fingerprinting with the accuracy and content detection capabilities of machine learning techniques. However, hybrid methods remain less mature than either pure fingerprinting or flow-based statistics, with limited validation in operational settings.

D. Limitations and Research Gaps

Despite significant progress, existing work exhibits several limitations that motivate our approach. *Reproducibility gaps* affect fingerprinting validation: few studies validate extraction against official JA4+ implementations, and processing pipelines vary across research groups, limiting reproducibility. *Evaluation methodology gaps* for fingerprinting are widespread: most studies lack train/test splits, systematic comparison frameworks with theoretical bounds, and fail to report coverage or abstention rates.

The *ECH constraint* challenges handshake-dependent methods: with major browsers implementing ECH support and CDN providers beginning rollout, methods that inspect ClientHello/ServerHello fields may lose visibility into the attributes they depend on. This motivates approaches that do not require handshake field inspection. Finally, *metric limitations* affect multiclass evaluation: many studies report accuracy or weighted averages rather than macro-averaged metrics that appropriately handle long-tail class distributions common in malware datasets.

In this work, we directly address these gaps through a systematic evaluation of ECH-resilient flow features, a principled comparison against TLS fingerprinting using both theoretical bounds and empirical results, and a unified dataset foundation that establishes clear baselines for future research.

III. METHODOLOGY

Our research employs the public dataset provided by Matoušek et al. [10], which contains authenticated network traces for both malware communications and benign mobile and desktop applications. The dataset is distributed in two formats: labeled CSV files containing pre-extracted TLS fingerprinting metrics and PCAP network traces. The dataset authors noted that the public repository contains a subset of PCAP files due to size limitations, while the CSV files represent the complete experimental record. Through direct communication with the authors, we obtained the full PCAP collection, ensuring our analysis uses the complete dataset for comprehensive evaluation of ECH-resilient classification approaches.

A. Flow Feature Extraction and Validation

Our analysis required comprehensive flow-level statistical features that were not available in the author-provided CSV files, which contained exclusively TLS fingerprinting metrics. To obtain the necessary features for ECH-resilient classification, we extracted flow records directly from the complete PCAP collection using NFStream [21], a high-performance network flow analysis framework. This extraction process enabled us to capture bidirectional flow characteristics that remain observable under encrypted communication channels, including those protected by ECH deployment.

1) NFStream Configuration and Feature Extraction: We configured NFStream to extract bidirectional flows using standard 5-tuple identification (source IP, destination IP, source port, destination port, protocol) with industry-standard timeout parameters: 120 seconds for idle flows and 1800 seconds for active connections. The framework extracted comprehensive flow statistics, providing coarse- and fine-grained traffic signatures.

To enable direct comparison with existing TLS fingerprinting approaches, we developed an NFStream plugin implementing the official JA4+ specification [1]. This plugin generates both JA4 (client) and JA4S (server) fingerprints in strict conformance with the published standard, ensuring reproducibility and enabling systematic comparison between fingerprinting and flow-based approaches.

2) Validation Against Established Baselines: To validate our NFStream extraction methodology, we performed systematic comparative analysis using the author-provided datasets. As the initial comparison revealed several important variations in data formatting and filtering approaches that made a direct, one-to-one analysis challenging, we developed a *unified filtering pipeline* that ensures consistent, reproducible comparisons between the author-provided CSVs and our NFStream-generated data. Our pipeline implements four systematic filtering steps:

1) Removing entries with missing JA4 hashes, as our research focuses exclusively on TLS connections.
2) Filtering flows labeled as "Unknown" applications to ensure classification accuracy.
3) Applying the author-provided SNI-based filtering list to remove advertising and tracking traffic that introduces classification noise.
4) Removing families with only single samples.

The SNI filtering list proved particularly critical, as it included not only advertising services but also ubiquitous domains that appear across multiple malware families and benign applications. These overlapping domains can yield identical fingerprints that significantly impact classification performance, making their removal essential for accurate evaluation.

3) Comparative Analysis of Fingerprinting Metrics: Table I presents our comparative analysis of key TLS fingerprinting metrics after applying our unified filtering pipeline. We evaluate four critical metrics following the definitions from [10]:

- *Coverage*: The percentage of flows for which a specific fingerprint type can be extracted. JA4 requires only the client hello message, while JA4S additionally requires the server response, making it more susceptible to incomplete handshakes.

TABLE I: TLS Fingerprinting Performance After Unified Filtering

Metric	Author CSV (Filtered)	NFStream (Filtered)
Dataset Characteristics		
Total Flows	15,157	16,566
JA4 Coverage	15,157 (100%)	16,566 (100%)
JA4S Coverage	13,934 (91.9%)	7,821 (47.2%)
SNI Coverage	14,732 (97.2%)	15,144 (91.4%)
JA4+JA4S+SNI Performance		
Total Unique Fingerprints	2,423	1,363
Uniqueness	89.4%	89.9%
Malware Family Coverage	75.0%	64.9%
Application Coverage	94.2%	93.1%
Malware-Benign Overlap	1.23%	0.38%

- *Uniqueness*: The percentage of fingerprints that uniquely identify a single malware family or application, indicating the discriminative power of the fingerprinting method.
- *Family Coverage*: The percentage of malware families or applications that possess at least one unique fingerprint, representing the method's ability to identify distinct behavioral patterns.
- *Overlap*: The percentage of fingerprints shared between malware and benign traffic, directly impacting false positive rates and classification ambiguity.

The comparative analysis reveals a critical divergence in JA4S coverage between our NFStream extraction (47.2%) and the author-provided CSVs (91.9%). Despite this substantial difference in server-side fingerprint availability, the metrics most critical for security applications demonstrate remarkable consistency. Uniqueness remains high at approximately 90% in both approaches, application coverage exceeds 93%, and crucially, the malware-benign overlap remains minimal (0.38–1.23%). These consistent patterns suggest that while the two extraction methods capture different numbers of complete handshakes, the fundamental discriminative power of TLS fingerprinting is preserved when combining JA4, JA4S, and SNI attributes.

4) Implementation Validation Against Official JA4+ Suite: The significant discrepancy in JA4S coverage between our NFStream extraction (47.2%) and the author-provided CSVs (91.9%) necessitated independent validation to establish ground truth. We processed the complete PCAP collection using the official JA4+ suite implementation, generating authoritative fingerprints. This validation serves two critical purposes: verifying the correctness of our extraction methodology and understanding the source of coverage discrepancies.

Table II presents the conformance analysis results. We evaluated two key metrics: *Flow Match Rate* (the percentage of flows identified by the official tool that were also present in each dataset) and *Fingerprint Conformance* (the percentage of matched flows for which the JA4/JA4S hashes were identical to those generated by the official implementation).

The validation results are conclusive and reveal important methodological insights. Our NFStream pipeline successfully matches and processes 99.98% of flows identified by the official JA4+ tool, with every extracted fingerprint conforming perfectly to the official specification (100% conformance for both JA4 and JA4S). This complete conformance validates the correctness of our implementation and provides high confidence in subsequent analyses.

TABLE II: Conformance with Official JA4+ Suite Implementation

Metric	Author CSV	NFStream
Total Flows in Dataset	33,589	163,897
Official Tool Flows (PCAPs)	43,132	43,345
Matched Flows (5-tuple)	33,577	43,337
Flow Match Rate	77.85%	99.98%
Among Matched Flows:		
JA4 Conformance	30,077/33,577 (89.58%)	32,787/32,787 (100%)
JA4S Conformance	15,924/31,846 (50.00%)	15,470/15,470 (100%)

In contrast, the author-provided CSVs exhibit significant discrepancies: missing over 22% of TLS flows present in the PCAPs and achieving only 50% JA4S conformance among matched flows. This strongly suggests that the high JA4S coverage (91.9%) reported in the author-curated CSVs results from a non-standard extraction process rather than actual server response availability in the network traces. The discrepancy likely stems from different interpretations of incomplete handshakes or alternative fingerprint computation methods not aligned with the official specification. Thus, earlier reports of ∼92% JA4S coverage appear inflated due to non-standard extraction pipelines, whereas our validation against the official JA4+ suite indicates that the true coverage in real PCAP traces is closer to 47%.

5) Implications for Flow-Based Analysis: These validation findings have several important implications for our research:

1) *Methodological Rigor*: The complete conformance with official JA4+ specifications ensures that our comparisons between fingerprinting and flow-based approaches are based on accurate, standardized implementations rather than potentially biased extraction methods.
2) *Coverage Limitations*: The actual JA4S coverage of 47.2% in real network traces highlights a fundamental limitation of server-dependent fingerprinting methods. In contrast, our flow-based features can be extracted from any TLS connection regardless of handshake completion.
3) *Reproducibility*: Our NFStream-based extraction pipeline, validated against the official implementation, provides a reproducible foundation for future research. *All extraction code and processed datasets are made available to ensure scientific reproducibility [22].*
4) *Performance Baselines*: The consistent performance patterns across key security metrics (uniqueness 90%, minimal overlap 0.38%) despite different extraction methods demonstrate that our dataset captures the essential characteristics needed for meaningful classification experiments.

The systematic filtering pipeline ensures that all subsequent comparisons between fingerprinting and flow-based approaches are conducted on a consistent, reproducible basis.

While absolute coverage values vary between extraction methods, the fundamental ability to discriminate between malware and benign traffic remains intact, providing a solid foundation for evaluating ECH-resilient alternatives to traditional TLS fingerprinting.

B. Feature Engineering

Our feature engineering approach leverages three complementary sets of traffic characteristics that remain observable under encryption and, critically, are resilient to ECH deployment. Unlike TLS fingerprinting methods that rely on plaintext handshake fields susceptible to privacy-enhancing technologies, our features capture fundamental communication patterns that persist regardless of encryption enhancements.

1) Core Flow Statistics: We extracted 33 bidirectional flow metrics from NFStream, capturing fundamental behavioral characteristics of network communications. These features are computed across three directional perspectives to capture both asymmetric communication patterns and aggregate behavior:

- *Source-to-destination (src2dst)*: Metrics computed exclusively for packets traveling from the flow initiator to the responder, capturing client request patterns and command transmission characteristics.
- *Destination-to-source (dst2src)*: Metrics for response traffic, revealing server behavior patterns and data exfiltration characteristics particularly relevant for malware detection.
- *Bidirectional*: Cumulative aggregation of both directions, where values represent the sum of src2dst and dst2src metrics (*e.g.*, if src2dst contains 10 packets and dst2src contains 15 packets, the bidirectional packet count equals 25).

The comprehensive feature set encompasses:

- *Volumetric metrics* (9 features): Packet counts and byte volumes computed separately for each directional perspective, capturing traffic intensity and data transfer patterns characteristic of different malware families.
- *Temporal characteristics* (3 features): Flow duration in milliseconds for each perspective, revealing communication session patterns and distinguishing between ephemeral connections and persistent command-and-control channels.
- *Packet size distributions* (12 features): Minimum, mean, standard deviation, and maximum packet sizes for each directional perspective. These statistical moments capture protocol-specific behaviors that remain consistent even under encryption, such as characteristic message sizes in malware communication protocols.
- *Packet inter-arrival times (PIAT)* (12 features): Minimum, mean, standard deviation, and maximum inter-arrival times in milliseconds for each perspective. PIAT distributions reveal timing patterns indicative of automated malware behavior versus human-driven benign applications, providing crucial discriminative signals for classification.

These 33 features capture *macro-level communication patterns* that persist regardless of encryption enhancements. The bidirectional perspective provides both directional asymmetry information (through comparison with unidirectional metrics) and aggregate flow behavior, enabling robust traffic classification even when individual packet contents are completely opaque.

2) Sequential Packet Length (SPL) Features: To capture *micro-level communication patterns* that characterize specific application protocols, we extracted Sequential Packet Length (SPL) features consisting of the sizes of the first 25 packets in each flow, ordered by arrival time. This approach is motivated by the observation that many application protocols, including malware command-and-control communications, exhibit characteristic packet size sequences during connection establishment and initial data exchange.

The 25-packet window was selected based on empirical analysis showing that most distinctive protocol behaviors manifest within this initial exchange, while longer sequences provide diminishing returns for classification accuracy.

3) Combined Feature Set: We constructed a hybrid feature representation incorporating both the 33 core flow statistics and 25 SPL values, resulting in a 58-dimensional feature vector. This combined approach aims to leverage complementary information from both feature types:

- *Macro-level patterns* from flow statistics capture overall communication behavior, session characteristics, and traffic intensity patterns that distinguish malware families with different operational profiles.
- *Micro-level signatures* from SPL sequences identify specific protocol implementations and message exchange patterns unique to particular malware variants or benign applications.

The synergy between these feature types addresses limitations of each individual approach: flow statistics may miss subtle protocol variations while SPL features alone may not capture broader behavioral patterns. The combined representation provides comprehensive traffic profiles that remain robust under current and future encryption enhancements, including ECH deployment.

C. Classification Framework

We evaluated three machine learning algorithms across three distinct classification tasks to comprehensively assess the effectiveness of flow-based features for ECH-resilient malware detection.

1) Classification Tasks: We designed three classification tasks to address different operational requirements in malware detection systems:

1) *Binary Classification*: Distinguishing malware from benign traffic, representing the most critical security task where minimizing false negatives is paramount for preventing successful attacks.
2) *Full Multiclass Classification*: Identifying specific families among all classes (malware families and benign applications), enabling fine-grained threat attribution and targeted response strategies.
3) *Malware-Only Multiclass Classification*: Discriminating between malware families exclusively, isolating the challenge of malware family attribution without the simplifying presence of benign traffic patterns.

Throughout this paper, we use the term *class* to denote either a malware family or a benign application family, depending on task context.

2) Machine Learning Algorithms: We selected three complementary algorithms representing different learning paradigms and operational trade-offs:

Random Forest (RF): We configured ensemble classifiers with 300 decision trees, a parameter selected through empirical convergence analysis showing negligible improvement beyond this threshold. The maximum tree depth is set to 15 for binary classification and 20 for multiclass tasks, providing sufficient model capacity while preventing overfitting. Regularization parameters include minimum samples split of 5 and minimum samples leaf of 2, determined through preliminary experiments to prevent memorization while maintaining discriminative power. To address class imbalance inherent in malware datasets, we employ balanced class weighting that inversely weights classes proportional to their frequency. This ensures minority malware families receive appropriate attention during training, crucial for detecting rare threats.

Neural Networks (NN): We implemented fully-connected architectures with batch normalization and dropout regularization, tailored to task complexity. For binary classification, we employ a funnel architecture (64→32→16→1 neurons) that progressively compresses feature representations toward the decision boundary. Multiclass tasks utilize a wider architecture (256→128→64→*num_classes*) to accommodate the increased complexity of distinguishing between up to 101 distinct traffic patterns. ReLU activations after batch normalization address vanishing gradient problems while accelerating convergence. Dropout rates of 0.3 and 0.2 at different layers prevent neuron co-adaptation, with the Adam optimizer (learning rate 10^{-3}) providing adaptive per-parameter learning rates crucial for features with different scales. Early stopping with patience of 3 epochs prevents overfitting while ensuring convergence, monitoring validation loss to restore optimal weights.

FAISS k-Nearest Neighbors (k-NN): We implemented Facebook AI Similarity Search (FAISS) [23] for scalable nearest neighbor classification, using `IndexFlatIP` (inner product similarity) on L2-normalized features to approximate cosine similarity. This choice enables efficient similarity search in high-dimensional feature spaces while maintaining geometric interpretability. We selected $k = 7$ for binary classification and $k = 5$ for multiclass tasks, representing an empirically-determined trade-off between local smoothness and robustness to outliers. Odd values prevent tied votes in classification decisions. These parameters were optimized based on expected class density in the feature space, with larger k for binary classification increasing stability in the simpler two-class problem while smaller k for multiclass tasks preserves local neighborhood structure necessary for fine-grained discrimination.

3) Training and Evaluation Methodology: All models were trained using stratified 80-20 train-test splits with fixed random seeds (42) to ensure complete reproducibility. Stratification maintains proportional representation of all classes in both training and test sets, crucial for unbiased performance estimation in imbalanced datasets.

Building upon our initial filtering (*cf.* Section III-A2) that removed singleton samples, we applied an additional minimum support threshold of 5 samples per family. This threshold is mathematically necessary for stratified sampling: classes with fewer than 5 samples would yield zero test samples after the 80-20 split ($\lfloor 4 \times 0.2 \rfloor = 0$), violating the requirement for representative evaluation. With exactly 5 samples, we guarantee at least one test sample per class ($\lfloor 5 \times 0.2 \rfloor = 1$), ensuring all families are represented in both training and evaluation sets.

This additional filtering eliminated 8 underrepresented families (24 samples total), yielding a final dataset of 16,542 records spanning 101 unique families, with 59 malware families and 42 benign application families.

4) Evaluation Metrics: We report standard classification metrics including *accuracy*, *precision*, *recall*, and *F1-score*. For multiclass tasks, we report macro-averaged metrics to avoid bias toward majority classes, ensuring that performance on rare malware families is appropriately weighted in aggregate scores. This is particularly important given that rare malware variants often represent emerging threats requiring immediate detection capability.

Additionally, we report *ROC-AUC* (Receiver Operating Characteristic - Area Under Curve) scores to characterize classifier confidence and decision robustness beyond threshold-dependent metrics. For binary classification, ROC-AUC measures the probability that a randomly chosen positive sample ranks higher than a randomly chosen negative sample. For multiclass tasks, we compute macro-averaged ROC-AUC using the One-vs-Rest strategy, providing insight into per-class discriminability.

5) Cross-Validation for Stability Assessment: While our primary results use a single stratified 80-20 split for direct comparability with prior work, we additionally perform 5-fold stratified cross-validation to assess model stability and transferability. Cross-validation provides variance estimates that quantify how sensitive performance is to the particular train-test partition, addressing concerns about result generalizability.

We employ `StratifiedKFold` with 5 folds and a fixed random seed (42) to ensure reproducibility. For each fold, models are trained on 80% of the data and evaluated on the held-out 20%, with metrics aggregated as mean ± standard deviation across folds. This protocol enables assessment of whether the single-split results are representative or anomalous, and provides confidence intervals for deployment planning.

D. Comparison Methodology with TLS Fingerprinting

To establish the relative merits of our ECH-resilient approach against traditional TLS fingerprinting, we developed a systematic comparison framework addressing the fundamental differences between deterministic fingerprint matching and probabilistic machine learning classification.

1) Deriving Comparable Metrics: TLS fingerprinting is a deterministic lookup: a flow either matches a known fingerprint or it does not. In contrast, our flow-based approach is probabilistic classification trained and evaluated on stratified splits. To enable a fair, apples-to-apples comparison

without overclaiming, we derive optimistic upper bounds for fingerprinting performance from two corpus-level quantities measured on our NFStream-cleaned dataset: *(i)* the percentage of fingerprint keys shared between malware and benign traffic ("overlap"), and *(ii)* the fraction of malware families that have at least one identifying fingerprint ("malware family coverage").

We derive optimistic upper bounds under the assumption that fingerprint keys are equally likely across flows:

- $P_{\max} = 1 - \text{overlap}$,
- $R_{\max} = \text{family coverage}$, and
- $F1_{\max} = \frac{2 \cdot P_{\max} \cdot R_{\max}}{P_{\max} + R_{\max}}$.

Using our NFStream-cleaned dataset (Table I), overlap is 0.38% and malware family coverage is 64.9%, yielding $P_{\max} \approx 99.62\%$, $R_{\max} \approx 64.90\%$, and $F1_{\max} \approx 78.6\%$.

These bounds favor fingerprinting and represent optimistic estimates: overlap is computed on unique fingerprints, not weighted by flow frequency; if shared fingerprints are high-frequency, real false positive rates could be higher and precision lower. Similarly, the recall bound ignores per-flow extraction failures and incomplete handshakes that limit coverage in practice.

IV. Experimental Results

We present a comprehensive evaluation of flow-based statistical features for handshake-independent malware detection, systematically exploring 27 configurations across three classification tasks, three feature sets, and three machine learning algorithms, demonstrating the viability of flow-based approaches as practical complements to traditional TLS fingerprinting.

A. Binary Classification: Malware Detection

Binary classification represents the most critical security task: distinguishing malware from benign traffic. This fundamental capability determines whether a network security system can identify threats regardless of their specific family or variant. Table III presents comprehensive results across all feature sets and models.

TABLE III: Binary Classification Performance (Malware vs. Benign). Random Forest with combined features achieves the best performance (F1=0.9811), with all configurations exceeding 92.5% F1-score.

Feature Set	Model	Accuracy	Precision	Recall	F1-Score
Core	NN	0.8870	0.9530	0.9000	0.9258
	RF	0.9655	0.9866	0.9691	0.9778
	FAISS	0.9099	0.9338	0.9525	0.9431
SPL	NN	0.8876	0.9305	0.9255	0.9280
	RF	0.9665	0.9874	0.9695	0.9784
	FAISS	0.8849	0.9102	0.9464	0.9279
Combined	NN	0.9003	0.9439	0.9278	0.9358
	RF	**0.9707**	**0.9902**	**0.9722**	**0.9811**
	FAISS	0.8906	0.9148	0.9487	0.9314

Random Forest consistently achieves superior performance across all feature combinations, with the combined feature set reaching 97.07% accuracy and 98.11% F1-score. Recall rates across all configurations range from 90.00% to 97.22%.

Several key patterns emerge from the binary classification results. Core flow statistics achieve 97.78% F1-score with Random Forest, while SPL features achieve 97.84%—nearly identical performance when used independently. The combined feature set reaches 98.11% F1-score (+0.33 percentage points over core features alone).

Random Forest outperforms neural networks by 4.5–5.2 percentage points in F1-score across all feature configurations. Even the lowest-scoring configuration—FAISS k-NN with SPL features—achieves 92.79% F1-score, and all nine binary configurations achieve F1-scores above 92.5%.

Precision-recall analysis reveals distinct model characteristics: Random Forest models maintain high precision (98.66–99.02%) and the highest recall (96.91–97.22%), while neural networks show higher variance with precision ranging from 93.05% to 95.30% and recall from 90.00% to 92.78%. FAISS k-NN achieves strong recall (94.64–95.25%) at the cost of lower precision (91.02–93.38%).

ROC-AUC Analysis: To characterize classifier confidence beyond threshold-dependent metrics, we computed ROC-AUC scores for all binary classification configurations. Table IV presents the results, demonstrating that Random Forest achieves near-perfect discrimination with ROC-AUC scores exceeding 0.99 across all feature sets.

TABLE IV: ROC-AUC Scores for Binary Classification. Random Forest achieves near-perfect discrimination (0.99+) across all feature sets, indicating robust ranking of malware above benign traffic regardless of threshold selection.

Model	Core	SPL	Combined
Neural Network	0.9422	0.9288	0.9420
Random Forest	**0.9933**	**0.9947**	**0.9949**
FAISS k-NN	0.9479	0.9212	0.9270

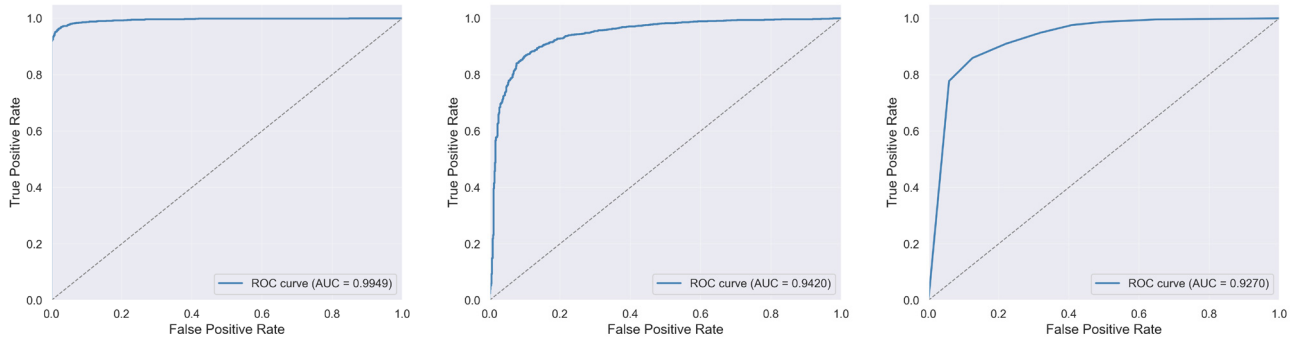
The ROC-AUC results reinforce the F1-score findings: Random Forest demonstrates exceptional discriminative ability, correctly ranking malware flows above benign flows with 99.49% probability when using combined features (99.61% in 5-fold CV). This high ROC-AUC indicates robust performance across all classification thresholds, not merely at the default 0.5 boundary. Fig. 1 visualizes the ROC curves for all three models using combined features, illustrating the near-perfect discrimination achieved by Random Forest.

B. Multiclass Classification: Complete Family Identification

The full multiclass task—distinguishing between all 101 classes (59 malware families and 42 benign application families)—represents the most challenging classification scenario, requiring models to capture subtle differences between semantically similar traffic patterns across 101 classes. Table V presents the performance across all experimental configurations.

Random Forest with combined features achieves the highest performance at 61.62% accuracy and 54.81% F1-score. This represents a substantial drop relative to binary classification (98.11% vs. 54.81% F1-score).

Fig. 2 shows the normalized confusion matrix for the best multiclass configuration across 101 families. The matrix



(a) Random Forest (b) Neural Network (c) FAISS k-NN

Fig. 1: ROC curves for binary classification using combined features. Random Forest achieves near-perfect discrimination, substantially outperforming Neural Network and FAISS k-NN across all decision thresholds.

TABLE V: Full Multiclass Classification Performance (101 Families). Performance drops substantially from binary detection, with Random Forest (combined features) achieving 54.81% macro F1—still meaningful for family attribution.

Feature Set	Model	Accuracy	Precision	Recall	F1-Score
Core	NN	0.4113	0.2685	0.2254	0.2053
	RF	0.5966	0.5191	0.5559	0.5239
	FAISS	0.4642	0.3966	0.3772	0.3743
SPL	NN	0.4092	0.2780	0.2447	0.2353
	RF	0.5561	0.5066	0.5055	0.4923
	FAISS	0.4180	0.3180	0.2981	0.2948
Combined	NN	0.4527	0.3482	0.2895	0.2889
	RF	**0.6162**	**0.5526**	**0.5738**	**0.5481**
	FAISS	0.4397	0.3794	0.3387	0.3430

exhibits a strong diagonal pattern with off-diagonal confusion clusters where related families are misclassified into each other.

Per-class analysis of the same model reveals that perfect accuracies are concentrated in classes with very small test support (≤ 12 flows), while several moderate-support families exhibit systematic confusion (e.g., *mega-wins-slot*: 31 test samples, 0% accuracy; *jelly-connect*: 27, 0%; *njrat*: 15, 0%).

Several important patterns emerge from the multiclass results. Core flow statistics outperform SPL features in isolation (52.39% vs. 49.23% macro F1 with Random Forest; +3.12 percentage points), contrasting with binary classification where the two feature types achieve nearly identical performance (97.78% vs. 97.84%). The combined feature set improves macro F1 by 2.42 percentage points over core flow statistics (54.81% vs. 52.39%).

Neural networks achieve only 28.89% macro F1 with combined features compared to Random Forest's 54.81% (gap: 25.92 percentage points). The dataset exhibits substantial imbalance (median 91 samples per family; min 5, max 1845; 18 classes with < 20 samples and 32 with < 50). Unlike binary classification where models maintain balanced precision and recall, multiclass results show significant divergence: Random Forest with combined features achieves 55.26% precision and 57.38% recall, while neural networks exhibit both low precision (34.82%) and low recall (28.95%).

Performance scaling analysis reveals substantial degradation from binary to multiclass tasks. Random Forest degrades by 43.30 percentage points (from 98.11% to 54.81% macro F1), FAISS k-NN by 58.84 percentage points (from 93.14% to 34.30%), and neural networks by 64.69 percentage points (from 93.58% to 28.89%).

ROC-AUC Analysis: Despite the lower F1-scores in multiclass classification, ROC-AUC scores remain high, suggesting strong class separability at the probability level on this dataset. Table VI presents ROC-AUC scores for the full multiclass task (101 classes), computed using macro-averaged One-vs-Rest (OvR) scoring.

TABLE VI: ROC-AUC Scores for Multiclass Classification (101 Classes). Despite lower F1-scores, Random Forest achieves 0.9768 ROC-AUC, indicating strong per-class discriminability.

Model	Core	SPL	Combined
Neural Network	0.9527	0.9521	0.9629
Random Forest	**0.9761**	**0.9562**	**0.9768**
FAISS k-NN	0.8113	0.7518	0.7714

The high ROC-AUC scores (0.9768 for Random Forest) contrast with the moderate F1-scores (0.5481), suggesting that the classifier assigns useful probability rankings even when hard predictions are incorrect. This is consistent with misclassifications occurring among closely related families with similar probabilities, rather than confident errors. Fig. 3 visualizes the macro-averaged and micro-averaged ROC curves for all three models.

C. Malware-Only Multiclass Classification

To isolate the challenge of discriminating between malware families without the simplifying presence of benign traffic, we evaluated classification performance across the 59 malware families exclusively. Table VII presents the results for this focused classification task.

The malware-only classification achieves a maximum F1-score of 48.71% with Random Forest on combined features.

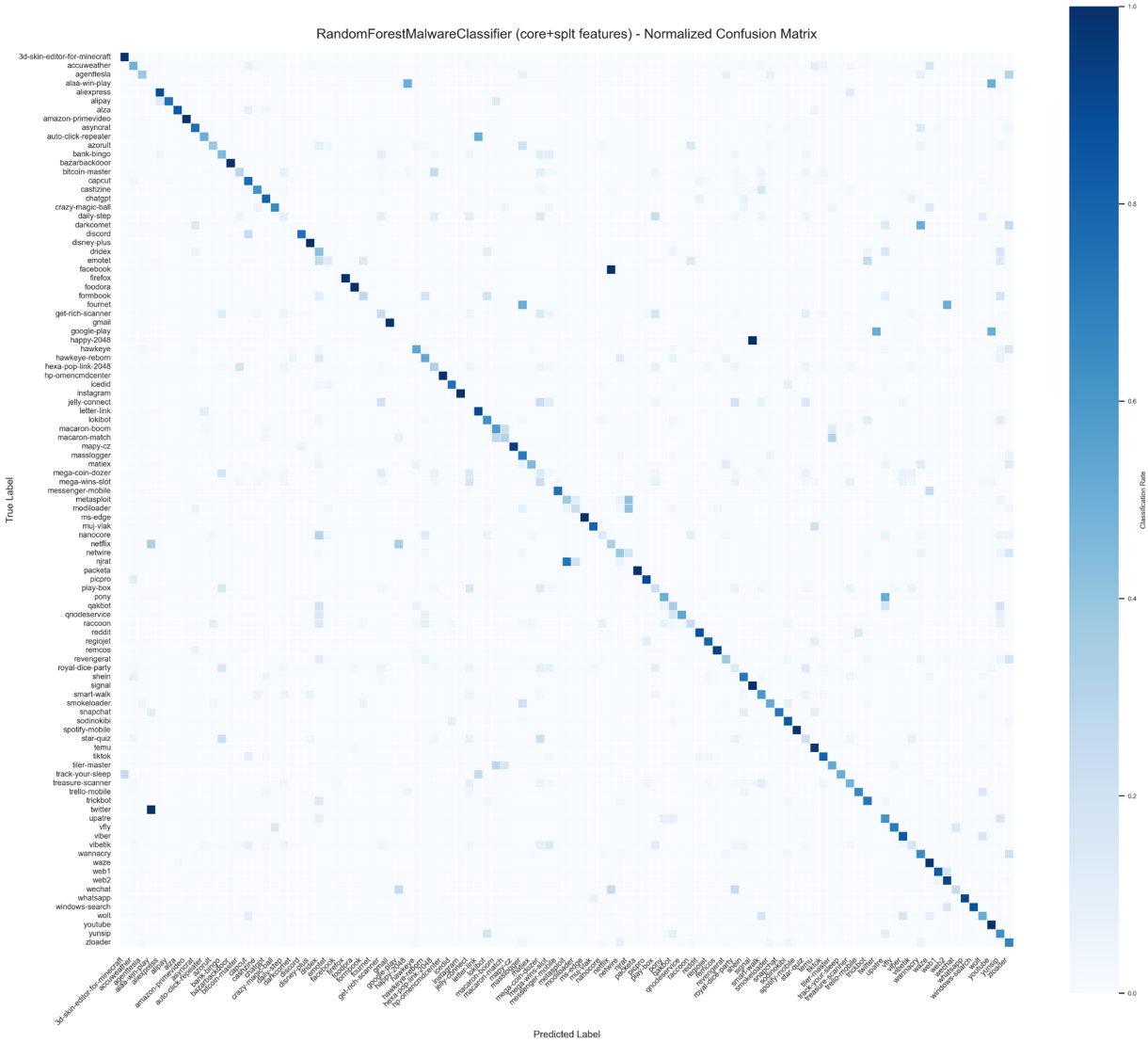


Fig. 2: Normalized confusion matrix for full multiclass classification (Random Forest, combined features). Darker diagonal entries indicate successful classification, while off-diagonal patterns reveal systematic confusions between related families.

TABLE VII: Malware-Only Multiclass Classification Performance (59 Families). Excluding benign classes reduces macro F1 (48.71% vs. 54.81%), reflecting inter-malware similarity and the loss of easily separable benign traffic.

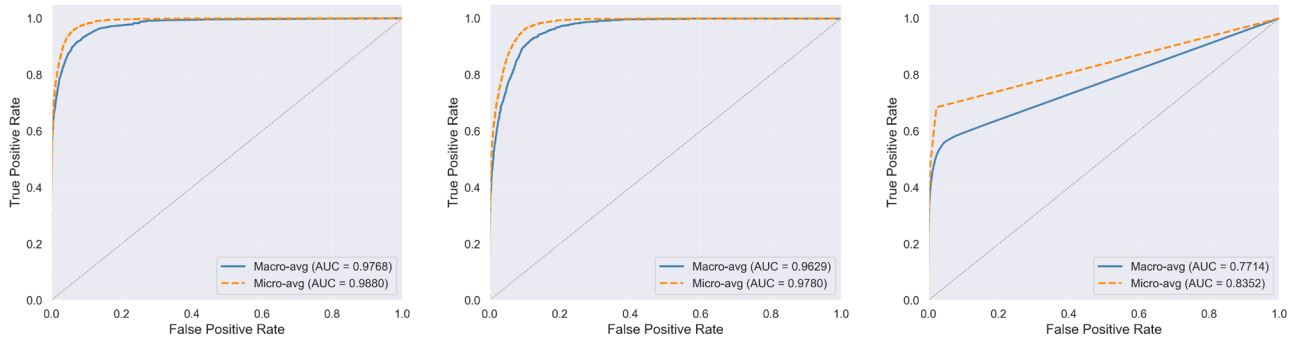
Feature Set	Model	Accuracy	Precision	Recall	F1-Score
Core	NN	0.4052	0.2573	0.2278	0.2121
	RF	0.5886	0.4843	0.4960	0.4764
	FAISS	0.4821	0.3535	0.3571	0.3412
SPL	NN	0.4099	0.2641	0.2243	0.2198
	RF	0.5083	0.3809	0.3944	0.3716
	FAISS	0.3998	0.2736	0.2495	0.2494
Combined	NN	0.4384	0.3288	0.2567	0.2621
	RF	**0.5998**	**0.4958**	**0.5126**	**0.4871**
	FAISS	0.4481	0.3074	0.3008	0.2916

Malware-only macro F1 is lower than full multiclass (48.71% vs. 54.81%).

Fig. 4 shows the normalized confusion matrix for malware-only classification, revealing the patterns of inter-malware confusion across 59 families.

Per-class analysis of the same model shows that perfect accuracies are concentrated in families with very small test support, while several moderate-support families exhibit systematic confusion (e.g., *mega-wins-slot*: 31 test samples, 0% accuracy; *jelly-connect*: 27, 11.1%; *njrat*: 15, 0%). In contrast, well-represented families such as *sodinokibi* (369 test samples) exceed 82% accuracy.

The counterintuitive result—that malware-only classification performs worse than full multiclass—presents several notable patterns. Malware-only classification achieves 48.71% macro F1 compared to 54.81% in full multiclass, representing



(a) Random Forest (b) Neural Network (c) FAISS k-NN

Fig. 3: ROC curves for multiclass classification (101 classes) using combined features. Macro-averaged (solid) and micro-averaged (dashed) curves show Random Forest maintains strong discriminability despite lower F1-scores.

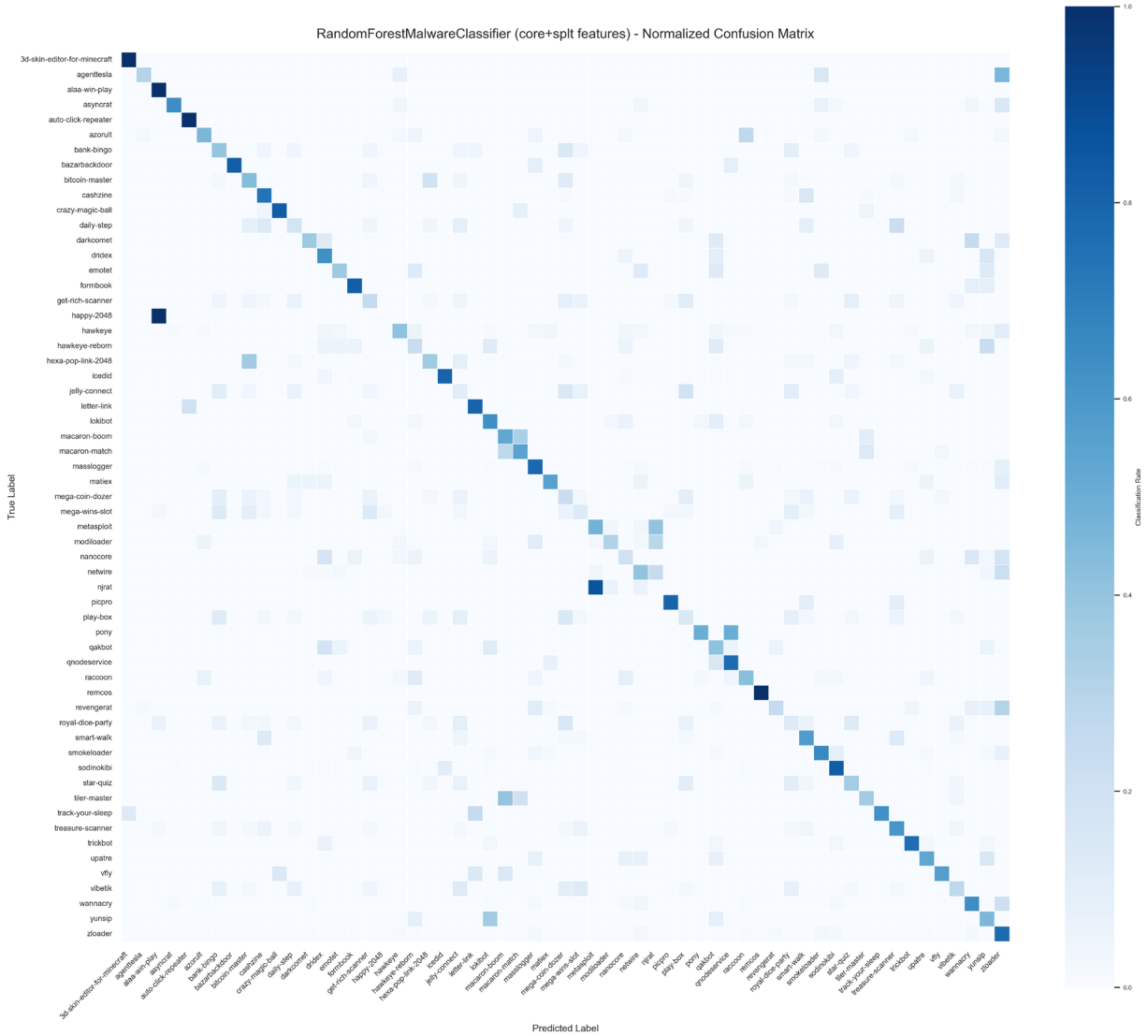
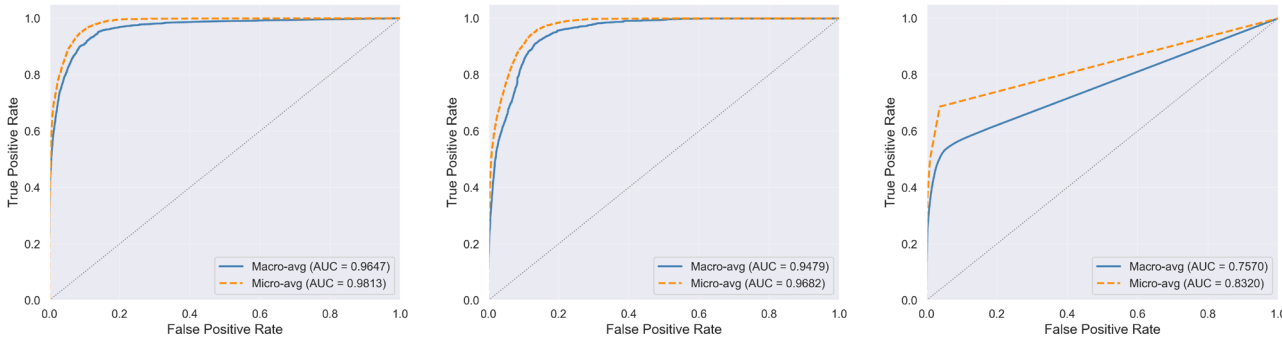


Fig. 4: Normalized confusion matrix for malware-only multiclass classification (Random Forest, combined features) across 59 malware families.



(a) Random Forest (b) Neural Network (c) FAISS k-NN

Fig. 5: ROC curves for malware-only classification (59 classes) using combined features. Macro-averaged (solid) and micro-averaged (dashed) curves show Random Forest maintains strong discriminability for distinguishing between malware families.

a 6.1 percentage point disadvantage. Core flow statistics significantly outperform SPL features in the malware-only context (47.64% vs. 37.16% macro F1 with Random Forest; +10.48 percentage points), a larger gap than observed in full multiclass (+3.12 pp).

Model degradation from binary to malware-only classification is substantial across all approaches: Random Forest drops from 98.11% to 48.71% macro F1 (−49.40 pp), FAISS from 93.14% to 29.16% (−64.00 pp), and Neural Networks from 93.58% to 26.21% (−67.47 pp). The combined feature set improves macro F1 by +1.07 percentage points over core features (48.71% vs. 47.64%), a smaller gain than observed in full multiclass (+2.42 pp).

ROC-AUC Analysis: Table VIII presents ROC-AUC scores for malware-only classification, computed using macro-averaged One-vs-Rest scoring across 59 malware families.

TABLE VIII: ROC-AUC Scores for Malware-Only Classification (59 Classes). Random Forest achieves 0.9647 ROC-AUC, indicating strong discriminability despite lower F1-scores.

Model	Core	SPL	Combined
Neural Network	0.9355	0.9389	0.9478
Random Forest	**0.9614**	**0.9430**	**0.9647**
FAISS k-NN	0.8146	0.7376	0.7570

Similar to full multiclass, the high ROC-AUC (0.9647) contrasting with moderate F1-score (0.4871) indicates that misclassifications occur among similar malware families with comparable probability scores rather than confident errors. Fig. 5 visualizes the macro-averaged and micro-averaged ROC curves for all three models.

D. Comparison with TLS Fingerprinting

Table IX contrasts theoretically derived bounds for TLS fingerprinting with empirical flow-based classification results on the same corpus. TLS entries are upper bounds computed from our NFStream-cleaned dataset: precision bound is 1 − overlap with overlap = 0.38%, recall bound equals malware family coverage (64.9%), and F1 bound is their harmonic mean. Flow-based entries are test-set metrics for Random Forest with combined features. The key distinction: TLS fingerprinting requires handshake inspection and abstains when no match exists, while flow-based classification uses only traffic statistics and produces a prediction for every flow. Consequently, per-instance accuracy cannot be computed for fingerprinting without instance-level predictions.

These results highlight that flow-based methods substantially extend coverage and recall, while fingerprinting remains fundamentally constrained by family uniqueness and handshake visibility.

E. Random Forest: Interpretability and Stability

The preceding results establish Random Forest as the consistently best-performing model across all tasks and feature configurations. We therefore focus our interpretability and stability analyses on Random Forest, which additionally provides natural feature importance measures through its ensemble structure.

1) Feature Importance Analysis: To provide interpretability into the classification decisions, we extracted feature importances from the Random Forest models across all three classification tasks. Table X present the top 10 features for each task using combined features, averaged across 5-fold cross-validation.

The feature importance analysis reveals task-dependent patterns. For *binary classification*, packet size features dominate: the top three features (bidirectional_max_ps, dst2src_max_ps, ps_4) each contribute 10.7–10.9% importance, collectively accounting for 32% of total importance. Both aggregate statistics and early sequential packet sizes appear in the top 10.

For *multiclass classification* (101 families), importance is more evenly distributed across features, with duration and timing features gaining prominence alongside packet sizes. The top feature (ps_4) contributes only 3.8% importance compared to 10.9% in binary classification, indicating that fine-grained family discrimination requires a broader feature combination.

For *malware-only classification* (59 families), temporal features dominate: duration and packet inter-arrival time (PIAT)

TABLE IX: TLS Fingerprinting (bounds) vs. Flow-Based Classification (empirical). Flow-based ML achieves 97.22% recall compared to fingerprinting's 64.9% upper bound—a 32+ percentage point improvement in detection coverage.

Characteristic	TLS Fingerprinting (JA4+JA4S+SNI)	Flow-Based ML (RF + Combined)
Fundamental Capabilities		
ECH-Resilient	No	**Less affected (handshake-independent)**
Handshake Inspection[a]	Requires	**Not required**
Requires Training Data	No	Yes
Handles Unknown Patterns	No (no match)	**Yes (prediction)**
Binary Detection Performance		
Precision (estimated)[b]	∼99.62%	99.02%
Recall (maximum)[c]	≤64.9%	**97.22%**
F1-Score	≤78.6%	**98.11%**
Family Attribution (Multiclass, 101 families)		
Malware Coverage	64.9%	**100%**[d]
Benign Coverage	93.1%	**100%**[d]
Accuracy	N/A[e]	**61.62%**

[a] TLS fingerprinting depends on ClientHello/ServerHello fields; flow-based classification does not inspect handshake fields.
[b] Based on 0.38% overlap between malware and benign fingerprints.
[c] Limited by 64.9% family coverage; 35.1% of families lack unique fingerprints.
[d] Can attempt classification for all families, though accuracy varies.
[e] Cannot compute without instance-level predictions.

TABLE X: Top 10 Feature Importances (Random Forest, Combined Features) across all classification tasks. Binary classification is dominated by packet size features (top 3 = 32%), multiclass shows more even distribution with timing features gaining prominence, and malware-only is dominated by temporal features (8 of top 10).

(a) Binary

Rank	Feature	Imp.
1	bidirectional_max_ps	0.109
2	dst2src_max_ps	0.108
3	ps_4	0.107
4	ps_6	0.066
5	ps_7	0.056
6	src2dst_max_ps	0.051
7	src2dst_stddev_ps	0.025
8	ps_8	0.023
9	ps_2	0.023
10	ps_1	0.019

(b) Multiclass (101)

Rank	Feature	Imp.
1	ps_4	0.038
2	src2dst_max_ps	0.034
3	src2dst_stddev_ps	0.029
4	bidirectional_duration_ms	0.029
5	src2dst_duration_ms	0.028
6	bidirectional_max_ps	0.028
7	src2dst_bytes	0.027
8	dst2src_duration_ms	0.027
9	src2dst_max_piat_ms	0.026
10	dst2src_mean_ps	0.026

(c) Malware-Only (59)

Rank	Feature	Imp.
1	src2dst_duration_ms	0.037
2	bidirectional_duration_ms	0.037
3	src2dst_max_ps	0.032
4	bidirectional_max_piat_ms	0.032
5	src2dst_max_piat_ms	0.032
6	dst2src_max_piat_ms	0.032
7	dst2src_duration_ms	0.031
8	dst2src_mean_piat_ms	0.030
9	bidirectional_mean_piat_ms	0.029
10	src2dst_bytes	0.029

features occupy 8 of the top 10 positions. This shift suggests that distinguishing between malware families relies heavily on communication timing patterns—likely reflecting differences in command-and-control polling intervals, data exfiltration rates, and protocol-specific timing behaviors.

Critically, these top features are largely observable without TLS handshake fields, since they capture transport-layer characteristics rather than ClientHello/ServerHello parameters. The task-dependent feature importance patterns explain why combined features consistently outperform single feature sets: binary detection benefits from strong packet size signals, while family attribution requires the complementary timing information.

2) Stability Analysis: To assess result stability and transferability across different data partitions, we performed 5-fold stratified cross-validation for the best-performing model (Random Forest) across all three classification tasks. Table XI presents the cross-validation results, showing consistent performance across folds.

The cross-validation results demonstrate stability across all tasks. For binary classification, F1-score standard deviations are at most 0.0011, indicating highly stable results. The multiclass tasks show higher variance (F1 std 0.011–0.021), reflecting sensitivity to class imbalance and fold composition—this is expected given the long-tailed distribution with 18 classes having fewer than 20 samples. Importantly, the 5-fold mean F1-scores closely match single-split results (e.g., binary combined: 0.9836 CV vs. 0.9811 single-split), confirming the reliability of our primary evaluation protocol.

F. Summary of Experimental Findings

Across all experiments, several consistent patterns emerge. Flow-based classification achieves near-perfect binary detection, with Random Forest models exceeding 98% F1 (98.36% in 5-fold CV) and 99.6% ROC-AUC, while multiclass family attribution remains more challenging, reaching 54.81% F1 across 101 classes and 48.71% across 59 malware families. Cross-validation confirms result stability across all tasks, with binary F1 standard deviations below 0.002 and multiclass standard deviations of 0.011–0.021. Random Forest consistently outperforms neural networks and k-NN, particularly

TABLE XI: 5-Fold Cross-Validation Results (Random Forest). Low standard deviations confirm stable, representative results across all tasks. Binary classification shows F1 std $\leq$0.0011; multiclass tasks show higher variance (std 0.011–0.021) reflecting class imbalance sensitivity.

Task	Features	Accuracy	F1-Score	ROC-AUC
Binary	Core	0.9682±0.0016	0.9796±0.0010	0.9942±0.0007
	SPL	0.9719±0.0016	0.9819±0.0011	0.9961±0.0005
	Combined	**0.9744±0.0011**	**0.9836±0.0007**	**0.9961±0.0002**
Multiclass (101 classes)	Core	0.6080±0.0028	0.5448±0.0111	0.9711±0.0046
	SPL	0.5587±0.0067	0.4939±0.0108	0.9634±0.0038
	Combined	**0.6268±0.0075**	**0.5694±0.0181**	**0.9751±0.0035**
Malware-only (59 classes)	Core	0.5819±0.0106	0.4806±0.0203	0.9687±0.0027
	SPL	0.5014±0.0102	0.3759±0.0164	0.9440±0.0029
	Combined	**0.5911±0.0094**	**0.4888±0.0200**	**0.9694±0.0024**

under class imbalance, confirming the suitability of tree ensembles for tabular flow features. Feature importance analysis reveals task-dependent patterns: packet size statistics dominate binary detection, while temporal features (duration, inter-arrival times) become more important for family attribution—all features remain observable under ECH deployment. Most importantly, flow-based methods substantially extend detection coverage and recall compared to TLS fingerprinting, which remains fundamentally constrained by family uniqueness and handshake visibility. These findings establish flow features as a robust, ECH-resilient foundation for network-based malware detection.

V. DISCUSSION

A. Performance Analysis and Operational Implications

Our results highlight a clear distinction on this corpus: binary malware detection reaches very high performance, while fine-grained family attribution remains challenging. The performance patterns across binary, multiclass, and malware-only tasks (Tables III, V and VII) reveal important insights about feature complementarity, model suitability, and deployment considerations. Binary detection achieves consistently high F1-scores across all models (>92.5%) with Random Forest reaching 98.11% using combined features, indicating that flow-based features are well-suited for malware vs. benign detection. For binary classification, core flow statistics and SPL features achieve nearly identical performance (97.78% vs. 97.84% F1 with Random Forest), with combined features providing minimal improvement (+0.33 percentage points).

This contrasts with multiclass scenarios where core features consistently outperform SPL features—by +3.12 points in full multiclass and +10.48 points in malware-only classification. Combined features yield only modest gains in multiclass settings (+2.42 points and +1.07 points respectively), suggesting limited complementarity for family discrimination beyond core flow statistics.

Neural networks consistently underperform Random Forest across all tasks, with performance gaps widening as task complexity increases: from 4.5–5.2 points in binary classification to 25.92 points in full multiclass (28.89% vs. 54.81%). This is consistent with overfitting risks under high class counts (101), tabular feature regimes (33–58 input features: 33 core + 25 SPL), and pronounced class imbalance (median 91 samples/class; min 5, max 1845; 18 classes with <20 samples, 32 with <50). Tree ensembles benefit from implicit feature selection, robustness via bagging, and non-linear interactions without heavy regularization. Practical mitigations for neural models include class-weighted loss, mild L2/dropout or label smoothing, and careful width/depth tuning for shallow MLPs; we applied early stopping, and leave broader NN ablations to future work.

For operational deployment, model selection depends on performance requirements: Random Forest maintains both high precision (98.66–99.02%) and high recall (96.91–97.22%) for binary detection, while FAISS k-NN offers strong recall (94.64–95.25%) at the cost of lower precision (91.02–93.38%) in settings prioritizing threat detection over false positive minimization. To ensure methodological rigor, we validated JA4/JA4S extraction against the official JA4+ suite (99.98% flow matching, 100% hash conformance) and used a unified filtering/evaluation pipeline with fixed seeds for full reproducibility.

Malware-only family attribution (59 classes) reaches 48.71% macro F1 with Random Forest, lower than full multiclass (54.81%). This counterintuitive result likely stems from: (1) removal of easily separable benign classes reducing macro-F1 averages; (2) inter-malware similarity and long-tail class imbalance causing systematic confusion in moderate-support families (e.g., *mega-wins-slot*: 31 test samples, 0%; *jelly-connect*: 27, 11.1%; *njrat*: 15, 0%), as shown in Figs. 2 and 4. The similarity among malware families reflects shared operational requirements—many malware types exhibit similar communication patterns when sharing command-and-control infrastructure, common development frameworks, or similar operational objectives, making flow-based discrimination inherently challenging.

B. Dataset Suitability and Generalizability

The dataset provided by Matoušek et al. [10] represents one of the most comprehensive publicly available corpora for TLS-based malware detection research, encompassing 101 families across desktop/mobile malware and benign applications. However, several characteristics warrant discussion regarding result generalizability.

Class Distribution. The dataset exhibits a long-tailed distribution typical of real-world malware collections: median sup-

port is 91 samples per family, but 18 families have fewer than 20 samples and 32 have fewer than 50. This imbalance reflects operational reality—some malware families generate abundant traffic while others are rare—but complicates evaluation. We address this through macro-averaged metrics that weight all families equally, ensuring rare families are not masked by high-frequency ones. The cross-validation results (Table XI) demonstrate that performance is stable across different data partitions, with binary F1 standard deviations below 0.002 and multiclass standard deviations of 0.011–0.021 reflecting sensitivity to fold composition under class imbalance.

Temporal Scope. The dataset captures malware behavior at specific collection periods. Malware evolves continuously; new variants may exhibit different communication patterns than those in the training corpus. Our feature importance analysis suggests that the dominant discriminative features—packet size distributions—reflect fundamental protocol-level behaviors that may be more stable than application-specific patterns. However, longitudinal validation on temporally disjoint datasets remains important future work.

Traffic Mix. The corpus contains controlled capture conditions that may not reflect operational network diversity (e.g., NAT traversal, middlebox interference, variable network conditions). Flow-level statistics may exhibit different distributions in high-latency or lossy environments. Deployment in diverse network conditions should be validated empirically.

Generalization to Unseen Families. Our evaluation uses closed-set classification where all test families appear in training. Open-set scenarios—detecting malware from previously unseen families—require different evaluation protocols. The high ROC-AUC scores (0.99+) for binary detection suggest strong separation between malware and benign distributions, which may support open-set generalization, but explicit evaluation is needed.

Despite these considerations, we believe the dataset is appropriate for this study's primary contribution: demonstrating that flow-level features provide handshake-independent detection capability that substantially exceeds TLS fingerprinting bounds on this corpus. The methodological framework—validated extraction pipelines, cross-validation stability assessment, and systematic comparison—establishes reproducible baselines that future work can extend with additional datasets.

C. Future Work and System Enhancements

To improve family-level attribution beyond the current 48.71–54.81% F1-score range, several enhancement strategies merit investigation. Flow features should be augmented with complementary signals such as DNS query patterns, TLS certificate characteristics, and temporal behavioral sequences that capture attack progression over time.

Methodological improvements should address class imbalance through targeted strategies: class-weighted loss functions for neural networks, intelligent resampling techniques, and flow-weighted evaluation metrics that account for traffic volume differences between families. Hybrid detection pipelines combining fast fingerprint lookups with ML-based fallback classification could leverage the strengths of both approaches.

For a directly comparable TLS fingerprinting baseline, one should construct the fingerprint database from the training split only and evaluate deterministic lookup on the test split, reporting coverage (abstention rate) alongside precision/recall and top-k accuracy. Open-set handling (predicting "Unknown" for unseen families) should be included for realistic deployment scenarios.

Future ablation studies should employ rigorous cross-validation to quantify performance variance and assess the stability of family attribution across different temporal periods. A systematic analysis of confusion matrix clusters can guide feature engineering specifically tailored to problematic family pairs that exhibit persistent misclassification patterns. Taken together, these directions suggest that while binary detection is largely solved, advancing reliable multi-family attribution will require richer features, better imbalance handling, and hybrid pipelines.

D. Limitations

Our evaluation reports macro-averaged metrics, which weigh classes equally and can be sensitive to long-tailed class distributions and very small supports. While we supplement the primary 80-20 split results with 5-fold cross-validation to assess stability (showing binary F1 standard deviations below 0.002 and multiclass standard deviations of 0.011–0.021), the cross-validation was performed for Random Forest across all three tasks; broader CV analysis across all models would strengthen generalizability claims. We applied balanced class weighting in Random Forest but did not systematically evaluate other rebalancing strategies (e.g., SMOTE, class-weighted losses for neural networks), leaving potential performance gains unexplored.

We did not evaluate adversarial robustness (e.g., malware authors deliberately shaping traffic to mimic benign patterns), which remains an important open question for deployment. The dataset captures malware at specific temporal snapshots; concept drift over time may degrade performance on newer malware variants. Finally, compute/runtime characteristics are not benchmarked here and are left to future system evaluations.

E. TLS Fingerprinting vs. Flow ML

Table IX contrasts theoretically derived bounds for TLS fingerprinting (precision bound from 0.38% overlap, recall bound from 64.9% family coverage, F1 bound 78.6%) with empirical flow-based classification results (RF + combined features). The key limitation of fingerprinting is structural: roughly one-third of malware families lack unique TLS signatures, so recall cannot exceed 64.9% even under ideal conditions. In contrast, flow ML attains 97.22% recall, reducing missed detections by over 32 percentage points.

ECH resilience is a critical differentiator. As ECH adoption advances in major browsers and CDNs, methods that depend on inspecting ClientHello/ServerHello fields may lose visibility into the very attributes they use. This trend has been highlighted in multiple surveys [4], [5], which point out that fingerprinting approaches are increasingly brittle as TLS evolves. In contrast, flow statistics remain largely observable

without handshake fields, which should help preserve efficacy as privacy technologies evolve. We do not claim fingerprinting is immediately obsolete under ECH, but its dependence on handshake fields creates a structural limitation that flow-based methods avoid.

Adaptability also differs. Fingerprinting is a deterministic lookup that abstains on previously unseen patterns, whereas learned models may generalize to novel variants that retain behavioral characteristics even when TLS signatures change. Similar observations appear in recent NDSS work [18], which shows that unknown malicious traffic can be detected in real time from flow characteristics even without handshake visibility. This flexibility carries trade-offs: ML requires labeled training data, periodic retraining to handle concept drift, and higher computational resources than hash lookups. For an apples-to-apples baseline, we outline in Future Work a train/test split evaluation for fingerprinting that reports coverage (abstention rate) alongside precision/recall and top-k accuracy, including open-set handling.

VI. Conclusion

This paper presented a systematic evaluation of flow-based statistical features as a handshake-independent complement to TLS fingerprinting for malware detection in encrypted traffic. Across 27 configurations, Random Forest models with combined flow statistics and sequential packet lengths achieved 98.11% F1 in binary detection and 54.81% macro F1 in full 101-class family attribution—substantially exceeding the theoretical recall bound of 64.9% imposed by fingerprinting coverage limits. These results indicate that while binary detection reaches very high performance on this corpus, flow-based features also retain meaningful discriminative power for fine-grained attribution and are less dependent on handshake visibility.

Our evaluation emphasized reproducibility and methodological rigor: JA4/JA4S extraction was validated against the official JA4+ suite, fingerprinting bounds were quantified, and all experiments were performed within a unified pipeline with fixed seeds. These contributions provide reliable performance baselines and support flow-level statistics as a practical foundation for network security monitoring as encryption evolves. Future work will explore hybrid pipelines, integration of complementary side-channel signals, and longitudinal studies to strengthen family-level attribution and operational deployment.

Code Availability

The complete source code, experimental configurations, and reproducibility pipelines supporting this work are publicly available at [22].

Dataset Availability

The complete dataset (16,542 flows, 89 features) is publicly available at [22]. The dataset is an enriched reprocessing of the malware traffic captures originally collected by Matoušek et al. [10], extending the original 15 TLS handshake-focused features with 58 flow-level classification features (33 statistical flow metrics and 25 sequential packet length features), plus 24 TCP flag features and metadata columns extracted using NFStream.

A key distinguishing feature of our dataset is the *temporal flexibility* provided by sequential packet-level (SPLT) features, not typically available in flow datasets. While the 33 core flow statistics represent aggregate measurements from complete flows, the SPLT features preserve packet sizes, directions, and inter-arrival times for the first 25 packets in temporal order. This enables reconstruction of partial flows at any cutoff point (k=1 to 25) without accessing original PCAPs. Researchers can simulate early detection scenarios—computing volumetric, temporal, and statistical features from only the first k packets—to study detection accuracy versus latency trade-offs, progressive classification strategies that adapt observation windows based on confidence, and inline blocking feasibility for network security appliances.

This capability supports a growing research direction in real-time malware detection where classification decisions must be made within milliseconds of connection establishment, and distinguishes our dataset from traditional flow collections that provide only aggregate statistics from completed flows.

Acknowledgment

This work has been part of Celtic-Next project RAI-6Green: Robust and AI Native 6G for Green Networks with project-id: C2023/1-9 funded by 2024-1.2.6-EUREKA-2024-00009.

References

[1] FoxIO-LLC, JA4+: A suite of network fingerprinting methods, https://github.com/FoxIO-LLC/ja4, Accessed: 2025-09-05.

[2] G. Gomez *et al.*, "Unsupervised detection and clustering of malicious tls flows," *Security and Communication Networks*, vol. 2023, pp. 1–17, 2023. DOI: 10.1155/2023/3676692.

[3] B. Anderson *et al.*, "Deciphering malware's use of tls (without decryption)," *Journal of Computer Virology and Hacking Techniques*, vol. 14, no. 3, pp. 195–211, 2017. DOI: 10.1007/s11416-017-0306-6.

[4] C. Oh *et al.*, "A survey on tls-encrypted malware network traffic analysis applicable to security operations centers," *Applied Sciences*, vol. 12, no. 1, p. 155, 2021. DOI: 10.3390/app12010155.

[5] Z. Wang *et al.*, "Machine learning for encrypted malicious traffic detection: Approaches, datasets and comparative study," *Computers & Security*, vol. 113, p. 102 542, 2022. DOI: 10.1016/j.cose.2021.102542.

[6] C. Novo and R. Morla, "Flow-based detection and proxy-based evasion of encrypted malware c2 traffic," in *Proceedings of the 13th ACM Workshop on Artificial Intelligence and Security*, ser. CCS '20, 2020, pp. 83–91. DOI: 10.1145/3411508.3421379.

[7] O. Barut *et al.*, "Machine learning based malware detection on encrypted traffic: A comprehensive performance study," in *Proceedings of the 7th International Conference on Networking, Systems and Security (NSysS '20)*, 2020, pp. 45–55. DOI: 10.1145/3428363.3428365.

[8] J. Althouse. "Tls fingerprinting with ja3 and ja3s." Salesforce Engineering blog post.

[9] J. Althouse. "Ja4+ network fingerprinting." Blog post, FoxIO.

[10] P. Matoušek *et al.*, "Experience report: Using ja4+ fingerprints for malware detection in encrypted traffic," in *2024 20th International Conference on Network and Service Management (CNSM)*, 2024, pp. 1–5. DOI: 10.23919/cnsm62983.2024.10814358.

[11] P. Matoušek *et al.*, "On reliability of ja3 hashes for fingerprinting mobile applications," in *Digital Forensics and Cyber Crime*. 2021, pp. 1–22. DOI: 10.1007/978-3-030-68734-2_1.

[12] Y. R. Siwakoti and D. B. Rawat, "Detecting malicious traffic using ja3 fingerprints attributed ml approach," in *2024 IEEE 44th International Conference on Distributed Computing Systems Workshops (ICDCSW)*, 2024, pp. 128–133. DOI: 10.1109/ICDCSW63686.2024.00024.

[13] T. van Ede *et al.*, "Flowprint: Semi-supervised mobile-app fingerprinting on encrypted network traffic," in *Proceedings 2020 Network and Distributed System Security Symposium*, ser. NDSS 2020, 2020. DOI: 10.14722/ndss.2020.24412.

[14] A. Theofanous *et al.*, "Fingerprinting the shadows: Unmasking malicious servers with machine learning-powered tls analysis," in *Proceedings of the ACM Web Conference 2024*, ser. WWW '24, 2024, pp. 1933–1944. DOI: 10.1145/3589334.3645719.

[15] B. Anderson and D. McGrew, "Identifying encrypted malware traffic with contextual flow data," in *Proceedings of the 2016 ACM Workshop on Artificial Intelligence and Security (AISec)*, 2016, pp. 35–46. DOI: 10.1145/2996758.2996768.

[16] M. Piskozub *et al.*, "Malalert: Detecting malware in large-scale network traffic using statistical features," *ACM SIGMETRICS Performance Evaluation Review*, vol. 46, no. 3, pp. 151–154, 2019. DOI: 10.1145/3308897.3308961.

[17] M. Yeo *et al.*, "Flow-based malware detection using convolutional neural network," in *2018 International Conference on Information Networking (ICOIN)*, 2018, pp. 910–913. DOI: 10.1109/ICOIN.2018.8343255.

[18] C. Fu *et al.*, "Detecting unknown encrypted malicious traffic in real time via flow interaction graph analysis," in *Proceedings 2023 Network and Distributed System Security Symposium*, ser. NDSS 2023, 2023. DOI: 10.14722/ndss.2023.23080.

[19] H. Kim et *al.*, "Revisiting tls-encrypted traffic fingerprinting methods for malware family classification," in *2022 International Conference on Information and Communication Technology Convergence (ICTC)*, 2022, pp. 1273–1278. DOI: 10.1109/ICTC55196.2022.9952872.

[20] S. Yu and Y. Won, "A survey of methods for encrypted network traffic fingerprinting," *Mathematical Biosciences and Engineering*, vol. 20, no. 2, pp. 2183–2202, 2022. DOI: 10.3934/mbe.2023101.

[21] Z. Aouini and A. Pekar, "Nfstream: A flexible network data analysis framework," *Computer Networks*, vol. 204, p. 108 719, 2022. DOI: 10.1016/j.comnet.2021.108719.

[22] FlowFrontiers, *ECH-Resilient Malware Detection via Flow-Level Statistical Features - Digital Artifacts*, https://github.com/FlowFrontiers/MalwareDet-JA4vsFlowStats, 2025.

[23] M. Douze *et al.*, "The faiss library," 2024. *arXiv: 2401.08281* [cs. LG].

Márton Pál Lipcsey-Magyar earned his BSc degree from the Department of Networked Systems and Services at the Budapest University of Technology and Economics, Hungary, in 2026. He is currently pursuing his MSc degree in the same department while also serving as a Research Fellow. His research interests include network and service management and the application of machine learning to various networking domains.

Attila Ármin Madarász earned his BSc degree from the Department of Networked Systems and Services at the Budapest University of Technology and Economics, Hungary, in 2026. He is currently pursuing his MSc degree in the same department while also serving as a Research Fellow. His research interests include network and service management and the application of machine learning to various networking domains.

Adrian Pekar is currently a Senior Data Scientist at CUJO AI, where he develops ML-powered solutions for home networks, focusing on attack detection and encrypted traffic analytics. Previously, he held the position of Associate Professor at Budapest University of Technology and Economics, where he continues to teach part-time. His research interests encompass network traffic flow measurement, machine learning for traffic analytics, federated learning for traffic classification, and cybersecurity applications.

Fuzzy Linguistic Signatures

Nour Ammar and László T. Kóczy

***Abstract*—Fuzzy Linguistic Signatures (FLS) extend the concept of Fuzzy Signatures (FSigs) by introducing linguistic variables as qualitative descriptors within a hierarchical fuzzy structure. Although fuzzy signatures have been successfully applied in various domains, their reliance on numerical membership degrees limits their ability to model subjective or linguistically defined information. This paper establishes a formal mathematical frame-work for FLS by defining a family of fuzzy linguistic signatures equipped with suitable linguistic aggregation operators and a partial ordering relation among linguistic values. Furthermore, meet-and-join operators are introduced to demonstrate that FLS satisfies the properties of a lattice as an algebraic structure. Consequently, fuzzy linguistic signatures provide an expressive representational framework capable of handling qualitative, human-like reasoning.**

***Index Terms*—Fuzzy Sets; Fuzzy Signature; Fuzzy Linguistic Signature**

I. INTRODUCTION

Fuzzy Signatures are multi-component fuzzy descriptors, extensions of the original concept of fuzzy set [1] and of Vector Valued Fuzzy Sets [5], with multi-level nested structure, where sub-components may be arranged in sub-signatures, according to closer interdependence or other ways of connectedness. Fuzzy sets are defined as follows:

$$A_f : \langle X, \mu_A : X \to [0,1] \rangle. \tag{1}$$

A vector valued fuzzy set (VFF set) is an extension of the above:

$$A_f^v : \langle X,\ \boldsymbol{\mu}^v : X \to [0,1]^n \rangle. \tag{2}$$

Here, the membership function maps each element of X into an n-component vector, where each component is an element of [0, 1]. This type of extended fuzzy set can be used when there is a multitude of properties for the same elements, and the i^{th} component of the membership vector expresses the degree of belonging of the VFF set in the sense of the i^{th} property.

Fuzzy Signatures (FSigs) are further extensions, namely, where the components of the fuzzy membership degree vector may be arbitrary multi-dimensional vectors themselves, thereby constructing a multilevel nested hierarchy of membership degrees. FSigs may be conveniently represented either by the mentioned nested vectorial structure or by a rooted tree graph, where each nested vector corresponds to a subtree, and the actual membership degrees are assigned to the leaves. Equation (3) defines fuzzy signatures recursively: each component μ_i may be either a scalar membership degree in $[0,1]$ or a vector-valued membership function of the same form, thereby inducing a finite rooted tree structure.:

$$A_f^{\text{sig}} : \langle X \to \boldsymbol{\mu}^{\text{sig}} \rangle, \quad \boldsymbol{\mu}^{\text{sig}} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \dots \\ \mu_n \end{bmatrix}. \tag{3}$$

Where:

$$\mu_i = \left\{ \begin{array}{l} \mu_i \\ \text{or} \\ \begin{bmatrix} \mu_{i1} \\ \mu_{i2} \\ \mu_{i3} \\ \vdots \\ \mu_{im} \end{bmatrix} \end{array} \right\} \quad \text{and so on, recursively.}$$

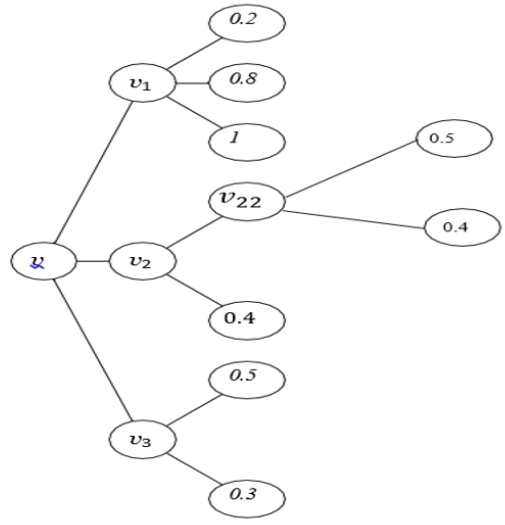


Fig. 1: Rooted tree for a FSigs

At the leaves, there may be membership functions of the respective universe rather than a single membership degree. For an exact mathematical definition and the algebraic structure of fuzzy signatures, a special case of L-fuzzy objects/sets, see [1]. Some aspects of the potential applications may be found in [2], and the specific problem of modeling the traffic situation in road networks was proposed in [3]. Linguistic variables have been applied by Wong [2] to describe uncertain values where even the exact fuzzy membership functions are hard to

N. Ammar is with the Department of Telecommunications and Artificial Intelligence, Budapest University of Technology and Economics, Budapest, Hungary (e-mail: nour.ammar@edu.bme.hu).

L. T. Koczy is with the Department of Information Technology, Szechenyi Istvan University, Gyor, Hungary (e-mail: koczy@tmit.bme.hu)

DOI: 10.36244/ICJ.2026.1.5

determine. The idea was first proposed by Zadeh [4] but was later discussed by numerous publications [16].

However, in some cases, assigning membership values and membership functions to the leaves is not feasible, thus they require an extension of this structure. For instance, in [7] the study aimed to analyze employee behavior, including both organizational citizenship behavior (OCB) and counter-productive work behavior (CWB). The study includes many subjective components that cannot be easily represented in a traditional, quantitative manner (e.g. altruism, courtesy and complaints).

Thus, we assigned linguistic labels to these components rather than membership degrees, where linguistic terms can characterize the values of these variables, will be more suitable to tackle this problem. This is the motivation to introduce Fuzzy Linguistic Signatures (FLS).

II. THE CONCEPT OF LINGUISTIC FUZZY SIGNATURES

Fuzzy systems were originally introduced to model systems and situations with non-probabilistic uncertainty (vagueness). However, in some areas, even the determination of fuzzy membership degrees and membership functions are not possible as inputs and outputs of such a system. So, they are formulated in natural language terms and expressions. To address this issue, Zadeh first proposed the concept of fuzzy sets and fuzzy systems, later introduced an even more human-friendly modeling technique, namely linguistic variables [4].

According to Zadeh, a linguistic variable is a variable whose values are words or sentences in a natural language rather than numbers. It can be characterized by the quintuple-tuple $\langle X, T(X), U, G, M\rangle$, where X is the variable; $T(X)$ is the term set of X; U is the universe of discourse; G is a syntactic rule set that generates the terms in $T(X)$; and M is a semantic rule set that associates a meaning with each linguistic value l (i.e., $M(l)$ denotes a fuzzy subset of U). For example, the term set of the linguistic variable Age may be $T(\mathrm{Age}) = \{\text{very young, young, not young, extremely old}\}$ [4]. Such linguistic-variable-based representations have been successfully employed in practical optimization and decision models, including fuzzy assignment models formulated using linguistic variables [17], [18].

Definition 1: Linguistic Variable

The above definition of linguistic variable may be simplified by omitting G from the model, as the syntax of the model will be restricted to a set of production rules, which may be defined and represented in the form of lookup tables (multidimensional tables) for associating output values to the input combinations. Thus, the model may be reduced to a quadruple $\langle X, T(X), U, M\rangle$.

Let us present a simple example. Let X be the linguistic variable "Weather" with the three-value term set $T(X) = \{\text{rainy, overcast, sunny}\}$. In this context, U may be the universe containing the percentage of the sky covered by clouds in Cartesian product with the amount of precipitation. This seems to be a reasonable formalization that helps map U to the unit interval and thus generate a set of fuzzy membership functions, formally by the mapping M. The meaning of a linguistic value l is characterized by a membership function, which associates its degree of compatibility with each $u \in U$ with the linguistic value l [14].

In the proposed structure, the structure of the rooted tree itself represents the hierarchical organization of fuzzy linguistic signatures (FLS) and is intentionally kept identical to that of fuzzy signatures (FSigs). This design ensures structural compatibility and allows direct comparison with the existing fuzzy signature model. The novelty of FLS lies in the semantic and algebraic nature of the linguistic values assigned to the leaves and propagated through the hierarchy. Another essential component of the FLS structure is the set of linguistic aggregation operators assigned to the internal nodes.

The paper proposes a definition of fuzzy linguistic aggregation operators that differs from earlier approaches in the literature. In FLS, numerical membership degrees are replaced by linguistic labels, which constitute a partially ordered set and cannot, in general, be embedded into a numerical scale without loss of meaning. To ensure this, aggregation at internal nodes must be defined in a manner that is entirely linguistic and does not rely on any numerical interpretation or mapping.

In the next section, a representation and reasoning framework is presented that is fundamentally different from FSigs, opening the door to computing with words [8] ,[15] within a hierarchical structure. The key question that follows is what kind of aggregation operators are capable of combining such linguistic values into a single linguistic result while satisfying the exact mathematical requirements?

A. Aggregation Operators for Fuzzy Linguistic Signatures

As mentioned above, Fuzzy Linguistic Signatures (FLS) are represented by a tree structure where the leaves are assigned linguistic values or labels, and intermediate nodes represent the aggregation result for the leaves. In this paragraph, we will review the definition of fuzzy aggregation operators and partial order among linguistic values and then will address the concept of linguistic aggregation as it appears in the literature and attempt to formulate the potential approaches for linguistic aggregations within FLS.

Definition 2: Aggregation Operators

Let $(P, \leq, \bot, \top)$ be a bounded partially ordered set (poset), where $\bot$ and $\top$ denote the lower bound and upper bound elements of P, respectively, such that

$$\bot \leq p \leq \top \quad \text{for all } p \in P.$$

An n-argument aggregation operator $a : P^n \to P$ is an order-preserving operator satisfying the following conditions:

1) $a(\top, \ldots, \top) = \top$.
2) $a(\bot, \ldots, \bot) = \bot$.
3) If $x_i \leq y_i$ for all i, then

$$a(x_1, \ldots, x_n) \leq a(y_1, \ldots, y_n).$$

The first and second conditions are called boundary conditions, and the third resembles the monotonicity property of the operator.

Definition 3: Partial Order of Linguistic Values

Let us assume we have a set of linguistic labels $L = \{L_1, L_2, \ldots, L_n\}$, where each label is represented by a vector:

$$\begin{aligned} L_1 &= \{L_{11}, L_{12}, \ldots, L_{1i}\} \\ L_2 &= \{L_{21}, L_{22}, \ldots, L_{2j}\} \\ L_3 &= \{L_{31}, L_{32}, \ldots, L_{3k}\} \\ &\vdots \\ L_n &= \{L_{n1}, L_{n2}, \ldots, L_{nm}\} \end{aligned}$$

The set L is partially ordered if there exists a relation $\leq$ for all pairs within L, and this relation satisfies the following properties:

1) $(i_1, j_1, k_1, \ldots, n_1) \leq (i_2, j_2, k_2, \ldots, n_2)$ $\Leftrightarrow i_1 \leq i_2$, $j_1 \leq j_2$, and $k_1 \leq k_2, \ldots, n_1 \leq n_2$.

2) They have an upper bound $(i_1, j_1, k_1, \ldots, n_1)$, $(i_2, j_2, k_2, \ldots, n_2) \leq (\max\{i_1, i_2\}, \max\{j_1, j_2\}, \max\{k_1, k_2\}, \ldots, \max\{n_1, n_2\})$

3) and lower bound: $(i_1, j_1, k_1, \ldots, n_1), (i_2, j_2, k_2, \ldots, n_2) \geq (\min\{i_1, i_2\}, \min\{j_1, j_2\}, \min\{k_1, k_2\}, \ldots, \min\{n_1, n_2\})$

Let us discuss a simple example for illustrating the partial order where not all the labels are directly comparable. In this context, we use the supremum and infimum operators to combine linguistic labels. Throughout this section, we assume that the set of linguistic labels is finite and equipped with a bounded componentwise partial order. Under this assumption, every finite subset admits a well-defined infimum and supremum, given by the componentwise minimum and maximum, respectively. Consider two linguistic sets describing temperature and humidity where Tem ={"Cold", "Medium", "Warm"} and Humidity={"Dry", "Moderate", "Humid"}.

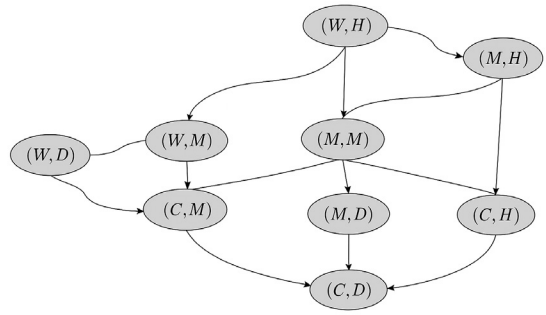


Fig. 2: Partial order of humidity and temperatures

Figure 2 illustrates the partial order defined on the Cartesian product of the linguistic variables temperature and humidity. The tuple (C, D), corresponding to {Cold, Dry}, represents the lower bound of the set, while the tuple (W, H), corresponding to {Warm, Humid}, represents the upper bound. Tuples such as (W, D), corresponding to {Warm, Dry}, and (M, M), corresponding to {Medium, Moderate}, are incomparable, although they share the same upper and lower bounds, because the partial order on the Cartesian product is defined component-wise and neither tuple dominates the other in all dimensions.

Linguistic aggregations have a very wide literature; an excellent overview is given in [9]. Numbers are assigned to the linguistic labels; the 2-tuple over weighted aggregation (TOWA), which is based on aggregation, utilizes the Extension Principle rather than relying solely on ordinal numbers [9]. Further extensions of the OWA by Xu [10], among others, follow this principle. All these operators share two common features: they are based on the ordinal numbers assigned to the linguistic labels, and they are related to the fuzzy OWA operators by Yager [11].

While the importance of OWA aggregation is undeniable, it should be noted that fuzzy aggregation includes several other important types. Fuzzy t-norms (intersections), t-conorms (unions), and various types of mean operations (arithmetic, geometric, harmonic, power means, etc.) also fall under this category, with the OWA being a special case of these means. Additionally, all kinds of hybrids may be defined, with the only condition being that they satisfy the original definition of fuzzy aggregations. Let us here recall the definition of fuzzy aggregation.

Definition 4: Fuzzy Aggregation Operators

Let $a(x_1, x_2, \ldots, x_n)$, where $x_i \in [0, 1]$, be a fuzzy function. Then a is an aggregation if and only if

$$a(0, 0, \ldots, 0) = 0, \quad a(1, 1, \ldots, 1) = 1,$$

$a(x_1, x_2, \ldots, x_n) \geq a(y_1, y_2, \ldots, y_n)$ iff $x_i \geq y_i$ for all i.

So, $a(x_1, x_2, \ldots, x_n)$ is a fuzzy aggregation operator over $x_1, x_2, \ldots, x_n$.

There is no argument against defining a concrete aggregation operations in a way that utilizes ordinal numbers (possibly tuples), but it is not a necessity, even in the case of finite and fixed sets of linguistic labels, and other ways, as e.g., by definition using rule-based system, defining the finite possible combinations, is an alternative. However, such ordinal number-based definitions cease to be applicable if the set of possible linguistic values is not yet fixed or the linguistic labels satisfy only partial ordering (no linear ordering). In natural languages, there are almost infinite possibilities to express fine differences among values, as e.g., linguistic hedges may modify the meaning of a basic label. If "expensive" is defined, "very expensive" may be considered a separate linguistic value, but "very-very expensive" has no meaning in this closed system. If there are fuzzy sets (in practice, convex and normal fuzzy sets, such as fuzzy numbers or fuzzy intervals) behind each label, representation, and calculations may become more complicated, but the set of possible labels may be kept "open" for possible additional linguistic values occurring during the operation of the modeled system, or when specifications and requirements are changed. A new, more “linguistic” definition could include such possibilities as well and also multi-component linguistic descriptors like "fast and expensive" or "slow and cheap". This way, much wider application fields for computing with words are opened [9], [15].

However, before discussing this new class of operations, it should be stated that it is reasonable to keep the signatures within a similar mathematical framework as FSigs of a family were defined in [1], as this would guarantee that signatures can be combined with each other by fuzzy (or similar) operations, that they can be compared, and that other manipulations necessary for further processing the given (linguistic fuzzy) data may be executed while preserving the exact mathematical justification for why and how to do this. As FSigs of the same family form a lattice, it would be advantageous to keep this essential property also for FLSs as well.

In order to determine the lattice structure of FLSs, the main point is to define the concept of linguistic aggregations in a way that preserves the properties forming the base of FSigs of a family being a lattice. In the next section, a simple extension for linguistic aggregations is given.

Definition 5: Linguistic Aggregation

Let

$$a_L(x_1, x_2, \ldots, x_n) \in L = \{L_\downarrow, L_1, L_2, \ldots, L_m, L_\uparrow\}$$

be a mapping

$$a_L : L^n \to L,$$

where $x_1, x_2, \ldots, x_n \in L$ are linguistic variables. Furthermore,

$$L_\downarrow \leq L_i \leq L_\uparrow, \quad \text{for all } i = 1, 2, \ldots, m,$$

and $\leq$ denotes a partial ordering on L.

L is the set of linguistic labels where $L_\downarrow$ is the lower bound and $L_\uparrow$ is the upper bound of the elements of L in the sense of $\leq$. Then a_L is a linguistic aggregation over L if and only if:

$$a_L(L_\downarrow, L_\downarrow, \ldots, L_\downarrow) = L_\downarrow,$$
$$a_L(L_\uparrow, L_\uparrow, \ldots, L_\uparrow) = L_\uparrow,$$
$$a_L(x_1, x_2, \ldots, x_n) \leq a_L(y_1, y_2, \ldots, y_n),$$
$$\text{for all } x_i \leq y_i \in L,$$

in the sense of the partial ordering.

Definition 6: Family of Linguistic Aggregation Operators

Let a family of linguistic aggregations over L, denoted by A_L, be defined as a set of all aggregation operators:

$$A_L = \{a_1, a_2, a_3, \ldots, a_n \mid a_i : L^n \to L\}.$$

Here, A_L is partially ordered by $\leq$, where for any two operators $a_1, a_2 \in A_L$:

$$a_1 \leq a_2 \;\Leftrightarrow\; a_1(x_1, \ldots, x_n) \leq a_2(x_1, \ldots, x_n),$$
$$for \quad all(x_1, \ldots, x_n).$$

With this ordering $(A_L, \leq)$, this pair is called the family of aggregators, and there exist a lower and upper bound (infimum and supremum) associated with this ordering. These are denoted as $\inf_L$ and $\sup_L$, respectively.

One of the simplest and non-trivial families is given by the set of aggregation operators:

$$A_L = \{a_{\inf}, a_x, a_y, a_{\sup}\},$$

where:

- $a_{\inf}$: returns the greatest lower bound (infimum) of the inputs (e.g., “low”);
- $a_{\sup}$: returns the least upper bound (supremum) of the inputs, according to the partial order over L (e.g., “high”);
- a_x and a_y: identity operators that simply return the linguistic values x and y.

To establish a suitable aggregation operation for linguistic variables within the fuzzy linguistic signature (FLS) framework, it is essential to define a partial ordering relation among the linguistic variables from the same family. This ordering provides a structured way to rank linguistic terms (such as “low,” “medium,” and “high”) so that aggregation operators can meaningfully combine them.

Figure 3 shows an example of how these operators can be organized in a complete lattice, allowing linguistic terms to be ordered hierarchically.

Let the linguistic term set be:

L = {hot, warm, cold}, where L is partially ordered as

$$\text{cold} \leq \text{warm} \leq \text{hot}.$$

Let us define a family of linguistic aggregation operators A_L, acting on pairs of inputs from L.

We define the following three operators:

- $a_{\inf}$: returns cold, which is the minimum of cold and hot;

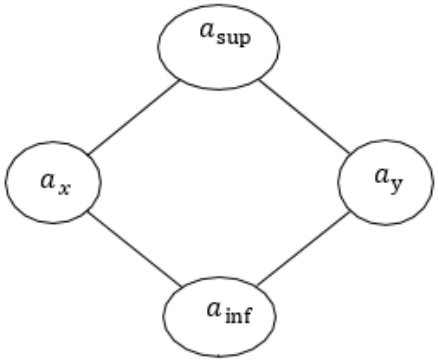


Fig. 3: Hasse diagram of the ordering relation defined on A.

- a_{avg}: returns warm, the middle value between cold and hot;
- $a_{\sup}$: returns hot, the maximum of cold and hot.

Thus, the aggregation family is:

$$A_L = \{a_{\inf}, a_{\text{avg}}, a_{\sup}\},$$

where $\leq$ is a partial order such that

$$a_{\inf} \leq a_{\text{avg}} \leq a_{\sup}.$$

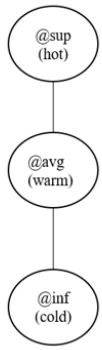


Fig. 4: Hasse diagram for the temperature example

Remark. The aggregation operator a_{avg} is defined so that, when the input linguistic terms are immediate successors with respect to the partial order, it returns the smaller term.
Fuzzy Linguistic Signatures (FLS) extend the concept of fuzzy Signatures (FSigs) by allowing the leaves of the signature tree to be assigned linguistic values or labels instead of fuzzy numbers. This structure enables the modeling of systems where qualitative, human-like reasoning is preferred over precise numerical values.

Definition 7: Fuzzy Linguistic Signatures

An FLS is represented as a tree linguistic structure G, where:

- Leaf nodes are assigned linguistic labels (e.g., low, medium, high).
- Internal nodes represent aggregation results of their child nodes using linguistic aggregation operators.

The formal definition of a fuzzy linguistic signature is

$$S_L = \langle N_I,\ N_L,\ \{a_1, \ldots, a_n\},\ \{L_1, L_2, \ldots, L_m\}\rangle.$$

Where:

S_L : Fuzzy linguistic signature,
N_I : Set of internal nodes of the graph G,
N_L : Set of leaf nodes of the graph G,
$V = N_I \cup N_L$ (Set of all vertices in G),
$\{a_1, \ldots, a_n\}$: Set of linguistic aggregation operators,
$\{L_1, L_2, \ldots, L_m\}$: Set of linguistic labels.

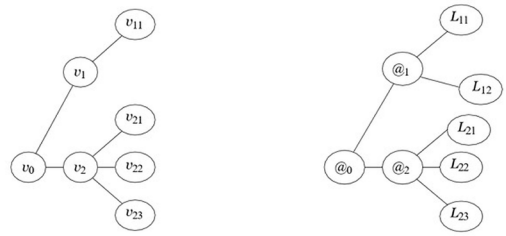


Fig. 5: Rooted tree G (left) and structure S_L (right)

An Example for Fuzzy Linguistic Signature Aggregation In the next, let us consider a very simple example. Let the traffic situation be described by the following set of linguistic expressions:

$$T_r = \big(\text{(Approximately) Zero},\ \text{Very Low},\ \text{Low},\ \text{Medium},\ \text{High},\ \text{Very High (Total Jam)}\big)$$

This set is linearly ordered, with $L_{\downarrow}$ = Zero and $L_{\uparrow}$ = Very High (see Figure 6. Let us assume now that the traffic intensity in a simple intersection is described by four-tuples of the traffic arriving from North, East, South, and West, denoted by

$$T_{r1} = (T_{rN}, T_{rE}, T_{rS}, T_{rW}),$$

a four-component vector-valued linguistic descriptor. This latter is analogous to and representable by a four-component vector-valued fuzzy descriptor [7]. However, in a real-life situation, the model is more adequate if the traffic is separately labeled in every possible incoming–outgoing direction, thus forming a subgraph with 12 leaves indexed by NW, NS, NE, EN, EW, etc.
For the simplest four-input example, if the traffic intensity situation is described by

$$T_r = (\text{Very Low, Low, High, Very High}),$$

where obviously, the EW and WE traffic is rather high, while the intersecting road has fewer vehicles approaching.
In the set of four-component linguistic traffic intensity descriptors, there exists a partial ordering defined based on the linear orderings of the component sets, namely:

$T_{r1} \leq T_{r2} \Leftrightarrow$
$T_{rN1} \leq T_{rN2},\ T_{rE1} \leq T_{rE2},\ T_{rS1} \leq T_{rS2},\ T_{rW1} \leq T_{rW2}.$

Here, $L_{\downarrow} =$ (Zero, Zero, Zero, Zero),
$L_{\uparrow} =$ (Very high, Very high, Very high, Very high).
The descriptor of a single intersection becomes more complex and requires a full fuzzy linguistic descriptor if, for example, for the intelligent traffic light control of the given intersection, a component describing the potential approach of one or more emergency vehicles is added. This necessitates a priority assignment to the given direction. Let us define the set of this additional feature by:

$$T_e = \big(\text{(None, Single, Multiple, Many)}\big)$$

Then the traffic intensity with added priority has two subtrees in every direction, both being single nodes describing T_e and T_r arriving from each direction. Figure 6 depicts the graph of such a simple linguistic signature describing a four-input intersection. Partial ordering can easily be defined here, too, as T_e is a linearly ordered set, and:

$$\begin{aligned}&(T_{r1}, T_{e1}) \leq (T_{r2}, T_{e2}) \ \Leftrightarrow\\ &(T_{rN1}, T_{eN1}) \leq (T_{rN2}, T_{eN2}),\\ &(T_{rE1}, T_{eE1}) \leq (T_{rE2}, T_{eE2}),\\ &(T_{rS1}, T_{eS1}) \leq (T_{rS2}, T_{eS2}),\\ &(T_{rW1}, T_{eW1}) \leq (T_{rW2}, T_{eW2}),\end{aligned}$$

$$\text{and } (T_{rJ1}, T_{eJ1}) \leq (T_{rJ2}, T_{eJ2}) \ \Leftrightarrow\ \begin{cases} T_{rJ1} \leq T_{rJ2}, \\ T_{eJ1} \leq T_{eJ2}. \end{cases}$$

where $\mathrm{J} \in \{N, E, S, W\}$.

Obviously, here:

$$\begin{aligned}L_{\downarrow} = \big(&\text{(Zero, None), (Zero, None),}\\ &\text{(Zero, None), (Zero, None)}\big),\\ L_{\uparrow} = \big(&\text{(Jam, Many), (Jam, Many),}\\ &\text{(Jam, Many), (Jam, Many)}\big).\end{aligned}$$

In [13], a simple traffic system was modeled, where traffic intensity and potential emergency vehicle appearance in the various directions, along with the waiting time of the queue (maybe consisting only of a single car), and the emergency vehicle priority in every incoming and outgoing direction, were taken into consideration. Let us define the following labels (which may be extended to all 12 directions):
W_t = {None, Short, Medium, Long, Very Long}, for the queues.
W_e = {None, Short, Long} , for potentially present emergency vehicles.
Here again, there is linear ordering in both sets, and there is a partial ordering in the set $W_t \times W_e$ so that $L_{\downarrow} =$ (None, None) and $L_{\uparrow} =$ (Very Long, Long).

It depends on the application of a given FLS and how the aggregations of the model structure can be defined. In [13], a fuzzy rule-based strategy was proposed, and in this illustrative example, a similar but more compact model and algorithm may be constructed based on FLS as above. The real-life application of such intelligent traffic control algorithms becomes interesting when a (maybe rather large) system of interconnected intersections is modeled as a single object. Traffic itself generates the information flow connecting the whole system, and it may be expected that some general features or patterns in the behavior of the traffic lights emerge as the result of the entire system adapting to the input "signals," namely, the traffic intensities and emergency vehicle appearances in the intersection system.

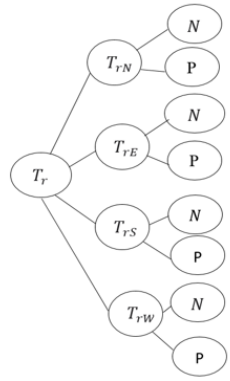


Fig. 6: Fuzzy linguistic tree for the traffic example

In the above paragraphs, it was presented how the graph structure of an FLS can be built up, and how the leaves of the trees can be linguistic values assigned. However, the problem of aggregations in the intermediate nodes (including the root itself) is just as important when the structure of the FLS is determined.

Let us continue the above example and focus on the traffic light control signal for the NS direction in a given intersection. The control output may be:

- G = Switch immediately to/keep on Green,
- GD = Switch with Delta delay to/keep on Green,
- RD = Switch with Delta delay to/keep on Red,
- R = Switch immediately to/keep on Red.

Then, in the simplest four-direction example, the linguistic aggregation for the FLS determining the traffic light of the North incoming street may be constructed as follows:

- The traffic light must be set to G if

$$(T_{rN}, T_{eN}) \geq \max\{(T_{rE}, T_{eE}),\ (T_{rW}, T_{eW})\}.$$

In a refined model, (T_{rS}, T_{rN}) is also taken into consideration.

The aggregation

$$A_N = A_N(T_{rN}, T_{eN}, T_{rE}, T_{eE}, T_{rW}, T_{eW}, W_{tE}, W_{tW})$$

(maybe also including T_{rS} and T_{eS}) should be defined by a lookup table, where the size grows exponentially with the number of inputs. Thus, the vector-valued approach is not feasible; instead, a real (multiple-level) FLS should be constructed with subtrees representing the "strength" or degree of each direction pushing the control decision towards a given value.

III. Definition of Meet and Join for Fuzzy Linguistic Signatures

Fuzzy Linguistic Signatures (FLS) are designed to represent uncertain and imprecise data in situations where exact fuzzy membership functions and degrees cannot be precisely specified. As an extension of fuzzy signatures (FSigs), it is reasonable to maintain FLS within a similar mathematical framework as FSigs. Previous research has shown that FSigs form a lattice structure [1], so it would be valuable to explore whether FLSs exhibit the same behavior, under what conditions this occurs, and to examine all potential cases. To demonstrate that FLSs form a lattice, we will discuss the meet and join operators and investigate if FLS satisfies the properties of idempotency, commutativity, and associativity. This proof will build on the findings in [1], while taking into account the unique requirements of linguistic labels. We further assume that the sets of linguistic labels and linguistic aggregation operators form partially ordered sets that admit infimum and supremum. For any given application, it is necessary that the set of linguistic labels contains a supremum and an infimum element, which semantically and intuitively bound all other linguistic labels from above and below, respectively.

Definition 8: Fuzzy Linguistic Signatures Generated from G and A_L

Let $G = (V, E)$ be a tree with root v_0, whose set of internal vertices is given by $N_I = \{v_0, v_1, v_2, \ldots, v_n\}$, and let $A_L = \{A_{L0}, A_{L1}, A_{L2}, \ldots, A_{Ln}\}$, be a set of families of aggregation operators. The family of fuzzy linguistic Signatures generated from G and A_L, which will be denoted as $F(G, A_L)$, is defined as follows:

$$F = \Big\{ S_{Lk} = \langle N_{Ik},\, N_{Lk},\, \{a_{i1}, \ldots, a_{ip}\}, L_1, \ldots, L_q\}\rangle \;\Big|\; G_k \subseteq G,\; a_{ij} \in A_{Lij},\; \text{for all } v_{ij} \in N_{Ik} \Big\}.$$

where $G_k = (N_{Ik} \cup N_{Lk}, E_k)$ is a subgraph of G, satisfying that $v_0 \in N_{Ik}$, and S_{Lk} is a fuzzy linguistic signature associated with G_k.

We are also interested in defining the family of fuzzy linguistic signatures generated from a given fuzzy signature. Notice that such a concept of family will be suitable for formalizing the information corresponding to situations similar to the ones explained at the beginning of this section.

Definition 9: The Family of Fuzzy Linguistic Signatures Generated from S_L

Let S_L be a fuzzy linguistic signature associated with the rooted tree G_s and the set of families of aggregation operators A_s. The family of fuzzy linguistic signatures generated from S_L is defined as the family of fuzzy linguistic signatures $F(G_s, A_s)$.

Therefore, given a fuzzy linguistic signature S_L, the family $F(G_s, A_s)$ is the set of fuzzy linguistic signatures S'_L obtained from S_L by omitting any number of leaves, or leaf subtrees (except the root itself), with the necessary corresponding modifications to maintain the definition of fuzzy linguistic signature. The labels of the internal vertices v of S'_L are associated with aggregation operators in the same family as the label of this vertex v in S_L. When a whole subtree has been removed, the root of this subtree becomes a leaf instead of its original non-leaf position in the original fuzzy linguistic signature and so, a linguistic label will be assigned to it, instead of a linguistic aggregation operator.

Definition 10 (Join of the Fuzzy Linguistic Signatures).

Let S_L be a fuzzy linguistic signature associated with the rooted tree G_s and the family of linguistic aggregation operators A_s. Let $S_{L_1}, S_{L_2} \in \mathcal{F}(G_s, A_s)$ be fuzzy linguistic signatures associated with the rooted trees $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, respectively, where N_{L_1}, N_{L_2} denote the sets of leaves and N_{I_1}, N_{I_2} denote the sets of internal vertices of G_1 and G_2.

The join of the fuzzy linguistic signatures S_{L_1} and S_{L_2}, denoted by $S_{L_1} \cup S_{L_2}$, is the fuzzy linguistic signature associated with the rooted tree

$$G_1 \cup G_2 = (V_1 \cup V_2, E_1 \cup E_2),$$

where $N_I(G_1 \cup G_2)$ denotes the set of internal vertices and $N_L(G_1 \cup G_2)$ denotes the set of leaves.

The linguistic aggregation operator assigned to each internal vertex $v \in N_I(G_1 \cup G_2)$ is defined by the following cases:

1. Internal vertices (Linguistic aggregation operators).

1) If $v \in N_{I_1}$ and $v \in N_{I_2}$, then
$$a_v = \sup\{a_v^1, a_v^2\},$$
where a_v^1 and a_v^2 are the linguistic aggregation operators assigned to v in S_{L_1} and S_{L_2}, respectively.
2) If $v \in N_{I_1}$ and $v \notin V_2$, then
$$a_v = a_v^1.$$
3) If $v \notin V_1$ and $v \in N_{I_2}$, then
$$a_v = a_v^2.$$
4) If $v \in N_{I_1}$ and $v \in N_{L_2}$, then
$$a_v = a_v^1.$$
5) If $v \in N_{L_1}$ and $v \in N_{I_2}$, then
$$a_v = a_v^2.$$

Remark. Since S_{L_1} and S_{L_2} belong to the same family of fuzzy linguistic signatures, their structures may not be identical. Consequently, it may occur that a given vertex is not present in one of the two signatures, or that a given vertex is a leaf in one signature and an internal vertex in the other. The above rules explicitly define how the join operation is computed in all these cases.

2. Leaves (Linguistic Labels):

If $v \in N_L\{G_1 \cup G_2\}$ (a leaf), the linguistic label assigned to v is computed considering the following cases:

1) If $v \in N_{L1}$ and $v \in N_{L2}$, the linguistic label assigned to v is:
$$L_v = \sup\{L_v^1, L_v^2\}$$
where L_v^1 is the linguistic label assigned to $v \in N_{L1}$ and L_v^2 is the linguistic label assigned to $v \in N_{L2}$.
2) If $v \in N_{L1}$ and $v \notin N_{L2}$, the linguistic label assigned to v is $L_v = L_v^1$, where L_v^1 is the linguistic label assigned to $v \in N_{L1}$.
3) If $v \notin N_{L1}$ and $v \in N_{L2}$, the linguistic label assigned to v is $L_v = L_v^2$, where L_v^2 is the linguistic label assigned to $v \in N_{L2}$.

Mixed cases in which a vertex is internal in one fuzzy linguistic signature and a leaf in the other are resolved according to the convention stated in the internal-vertex cases above.

Definition 11: Meet of the Fuzzy Linguistic Signatures

Let S_L be a fuzzy linguistic signature associated with the rooted tree G_s and the family of linguistic aggregation operators A_s. Let $S_{L_1}, S_{L_2} \in \mathcal{F}(G_s, A_s)$ be fuzzy linguistic signatures associated with the rooted trees $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, respectively, where N_{L_1}, N_{L_2} denote the sets of leaves and N_{I_1}, N_{I_2} denote the sets of internal vertices of G_1 and G_2. For each tree G_i, the vertex set satisfies $V_i = N_{L_i} \cup N_{I_i}$.

The meet of the fuzzy linguistic signatures S_{L_1} and S_{L_2}, denoted by $S_{L_1} \cap S_{L_2}$, is the fuzzy linguistic signature associated with the rooted tree
$$G_1 \cap G_2 = (V_1 \cap V_2, E_1 \cap E_2),$$
where $N_I(G_1 \cap G_2)$ denotes the set of internal vertices and $N_L(G_1 \cap G_2)$ denotes the set of leaves.

The linguistic aggregation operator assigned to each vertex in $N_I(G_1 \cap G_2)$ and the linguistic label assigned to each vertex in $N_L(G_1 \cap G_2)$ are defined as follows:

1. Internal vertices (Linguistic aggregation operators).

If $v \in N_I(G_1 \cap G_2)$, then the linguistic aggregation operator assigned to v is
$$a_v = \inf\{a_v^1, a_v^2\},$$
where a_v^1 and a_v^2 are the linguistic aggregation operators assigned to v in S_{L_1} and S_{L_2}, respectively.

The treatment of vertices that are internal in one fuzzy linguistic signature and leaves in the other follows the same convention as described in the remark following the definition of the join operation.

2. Leaves (Linguistic labels).

If $v \in N_L(G_1 \cap G_2)$, the linguistic label assigned to v is defined by the following cases:

1) If $v \in N_{L_1}$ and $v \in N_{L_2}$, then
$$L_v = \inf\{L_v^1, L_v^2\},$$
where L_v^1 and L_v^2 are the linguistic labels assigned to v in S_{L_1} and S_{L_2}, respectively.
2) If $v \in N_{L_1}$ and $v \notin N_{L_2}$, then
$$L_v = L_v^1.$$
3) If $v \notin N_{L_1}$ and $v \in N_{L_2}$, then
$$L_v = L_v^2.$$

IV. Lattice Properties of Fuzzy Linguistic Signatures

A. Structural Compatibility Assumption

Throughout this section, we restrict our attention to the family $\mathcal{F}(G_s, A_s)$ of fuzzy linguistic signatures defined over an identical rooted tree $G_s = (V_s, E_s)$. Consequently, for any two signatures $S_{L_1}, S_{L_2} \in \mathcal{F}(G_s, A_s)$, each vertex $v \in V_s$ is either a leaf in both signatures or an internal vertex in both. Thus, mixed cases where a vertex is a leaf in one signature and an internal vertex in the other do not occur.

We further assume that the sets of linguistic labels and linguistic aggregation operators are partially ordered and admit infimum and supremum. Under these assumptions, the meet ($\cap$) and join ($\cup$) operations defined in Definitions 10 and 11 are well defined and operate node-wise on G_s. If the underlying rooted trees are not identical, the proofs can be formulated in a very similar manner by a case-by-case analysis, following the constructions given in Definitions 10 and 11.

B. Lattice Properties of Fuzzy Linguistic Signatures

Let $S_{L_1}, S_{L_2}, S_{L_3} \in \mathcal{F}(G_s, A_s)$. We verify that the structure
$$(\mathcal{F}(G_s, A_s), \cap, \cup)$$
satisfies the lattice axioms.

a) Commutativity.:

$$S_{L_1} \cap S_{L_2} = S_{L_2} \cap S_{L_1}, \qquad S_{L_1} \cup S_{L_2} = S_{L_2} \cup S_{L_1}.$$

Proof. By Definition 11, for any internal vertex $v \in N_I(G_s)$,
$$a_v = \inf\{a_v^1, a_v^2\} = \inf\{a_v^2, a_v^1\},$$
and for any leaf $v \in N_L(G_s)$,
$$L_v = \inf\{L_v^1, L_v^2\} = \inf\{L_v^2, L_v^1\}.$$
Thus, the resulting aggregation operators and linguistic labels are independent of the order of the operands. The same argument applies to the join operation using the supremum. *End of proof.*

b) Associativity.:

$$(S_{L_1} \cap S_{L_2}) \cap S_{L_3} = S_{L_1} \cap (S_{L_2} \cap S_{L_3}),$$

$$(S_{L_1} \cup S_{L_2}) \cup S_{L_3} = S_{L_1} \cup (S_{L_2} \cup S_{L_3}).$$

Proof. For any internal vertex $v \in N_I(G_s)$,

$$\inf\{\inf\{a_v^1, a_v^2\}, a_v^3\} = \inf\{a_v^1, \inf\{a_v^2, a_v^3\}\},$$

by associativity of the infimum. Similarly, for any leaf $v \in N_L(G_s)$,

$$\inf\{\inf\{L_v^1, L_v^2\}, L_v^3\} = \inf\{L_v^1, \inf\{L_v^2, L_v^3\}\}.$$

Hence, aggregation results coincide at every vertex. The same reasoning applies to the join operation using supremum. *End of proof.*

c) Idempotency.:

$$S_{L_1} \cap S_{L_1} = S_{L_1}, \qquad S_{L_1} \cup S_{L_1} = S_{L_1}.$$

Proof. For any internal vertex $v \in N_I(G_s)$,

$$\inf\{a_v^1, a_v^1\} = a_v^1,$$

and for any leaf $v \in N_L(G_s)$,

$$\inf\{L_v^1, L_v^1\} = L_v^1.$$

Thus, applying meet or join to a signature with itself leaves it unchanged. *End of proof.*

d) Absorption.:

$$S_{L_1} \cap (S_{L_1} \cup S_{L_2}) = S_{L_1}, \qquad S_{L_1} \cup (S_{L_1} \cap S_{L_2}) = S_{L_1}.$$

Proof. For any internal vertex $v \in N_I(G_s)$,

$$\inf\{a_v^1, \sup\{a_v^1, a_v^2\}\} = a_v^1,$$

and for any leaf $v \in N_L(G_s)$,

$$\inf\{L_v^1, \sup\{L_v^1, L_v^2\}\} = L_v^1,$$

by the absorption property of infimum and supremum. The dual equality follows analogously. *End of proof.*

C. Resulting Lattice Structure of Fuzzy Linguistic Signatures

Since the meet and join operations satisfy commutativity, associativity, idempotency, and absorption, the structure

$$(\mathcal{F}(G_s, A_s), \cap, \cup)$$

forms a lattice of fuzzy linguistic signatures.

V. CONCLUSION

Linguistic variables are very useful in describing and handling everyday situations, especially those involving subjectivity and the terminology of natural language. When linguistic features can be structured in a multi-component hierarchy, this motivates the introduction of fuzzy linguistic signatures (FLS), which are particularly important for modeling such subjective cases.

FLS are similar in structure to fuzzy signatures. However, instead of fuzzy membership degrees and fuzzy membership functions, linguistic values are applied on the leaves. In the internal nodes, linguistic aggregation must appear; thus, we define the concept of linguistic aggregation.

We have shown that FLS can be handled if they belong to the same family of FLS that can be derived from a fuzzy linguistic mother signature by omission of the edges and the corresponding nodes. Such FLS belonging to the same family can be compared and aggregated.

We have shown that FLS of a family form a lattice, and in this way, they can be considered as an extension of the concept of L-fuzzy sets.

We have shown an example of decision-making in the context of a traffic situation, and we have presented that linguistic variables and FLS are suitable for modeling and making decisions in this application area.

REFERENCES

[1] L. T. Kóczy, M. E. Cornejo, and G. Medina, "Algebraic structure of fuzzy signatures," *Fuzzy Sets and Systems*, vol. 418, pp. 25–50, 2021, DOI: 10.1016/j.fss.2020.09.020.

[2] K. W. Wong, T. D. Gedeon, and L. T. Kóczy, "Fuzzy signature and cognitive modelling for complex decision model," in *Theoretical Advances and Applications of Fuzzy Logic and Soft Computing*, O. Castillo, P. Melin, O. M. Ross, R. Sepúlveda Cruz, W. Pedrycz, and J. Kacprzyk, Eds., ser. Advances in Soft Computing, vol. 42, Springer, 2007, pp. 380–389, DOI: 10.1007/978-3-540-72434-6_39.

[3] G. Mikulás and L. T. Kóczy, "Macro-level road network evaluation by fuzzy signature rule bases," *Hungarian Statistical Review: Journal of the Hungarian Central Statistical Office*, vol. 4, no. 1, pp. 3–16, 2021.

[4] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning–I," *Information Sciences*, vol. 8, pp. 199–249, 1975, DOI: 10.1016/0020-0255(75)90036-5.

[5] L. T. Kóczy, "Vector valued fuzzy set," *BUSEFAL (Université Paul Sabatier, Toulouse)*, pp. 41–57, 1980.

[6] Z. Xu, "Linguistic aggregation operators: An overview," in *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*, ser. Studies in Fuzziness and Soft Computing, vol. 220, Springer, 2007, pp. 163–181, DOI: 10.1007/978-3-540-72434-6_7.

[7] T. Calvo, A. Kolesárová, M. Komorníková, and R. Mesiar, "Aggregation operators: Properties, classes and construction methods," in *Aggregation Operators, ser. Studies in Fuzziness and Soft Computing*, vol. 97, Heidelberg: Physica, 2002, pp. 3–104, DOI: 10.1007/978-3-7908-1781-1_1.

[8] F. Herrera and J. L. Verdegay, "Linguistic assessments in group decision," in *Proc. 1st European Congress on Fuzzy and Intelligent Technologies*, Aachen, 1993, pp. 941–948.

[9] F. Herrera and L. Martínez, "A 2-tuple fuzzy linguistic representation model for computing with words," *IEEE Transactions on Fuzzy Systems*, vol. 8, pp. 746–752, 2000, DOI: 10.1109/91.890332.

[10] Z. Xu, "On generalized induced linguistic aggregation operators," *International Journal of General Systems*, vol. 35, pp. 17–28, 2006, DOI: 10.1080/03081070500226824.

[11] R. R. Yager, "On the ordered weighted averaging operators in multicriteria decision making," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 18, pp. 183–199, 1988, DOI: 10.1109/21.87068.

[12] T. D. Chala and L. T. Kóczy, "Intelligent fuzzy traffic signal control system for complex intersections using fuzzy rule base reduction," *Symmetry*, vol. 16, no. 9, p. 1177, 2024, DOI: 10.3390/sym16091177.

[13] T. D. Chala and L. T. Kóczy, "A novel, three-stage intelligent fuzzy traffic signal control system," *Acta Polytechnica Hungarica*, vol. 21, no. 8, 2024.

[14] N. Ammar and L. T. Kóczy, "Decision-Making Based on Fuzzy Linguistic Signatures," in *Proc. 4th Int. Conf. on Communications, Information, Electronic and Energy Systems (CIEES)*, Nov. 2023, pp. 1–5.

[15] L. A. Zadeh, "Fuzzy logic = computing with words," *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 2, pp. 103–111, May 1996, DOI: 10.1109/91.493904.

[16] N. Ammar and L. T. Kóczy, "Fuzzy Linguistic Signatures and Their Applications," in *Proc. 28th IEEE Int. Conf. on Intelligent Engineering Systems (INES)*, July 2024, pp. 27–30.

[17] Z. H. Radhy, F. H. Maghool, and N. K. Hady, "Fuzzy-assignment model by using linguistic variables," *Baghdad Science Journal*, vol. 18, no. 3, p. 5, 2021.

[18] D. Liu, Y. Liu, and X. Chen, "Fermatean fuzzy linguistic set and its application in multicriteria decision making," *International Journal of Intelligent Systems*, vol. 34, no. 5, pp. 878–894, 2019, DOI: 10.1002/int.22079.

Nour Ammar is a PhD candidate at the Department of Telecommunications and Artificial Intelligence, Budapest University of Technology and Economics, Hungary, where she also obtained her Master's degree. Her research focuses on fuzzy systems and fuzzy linguistic signatures, particularly their role in decision-making.

László T. Kóczy is a Professor at the De- partment of Information Technology, Széchenyi István University; Foreign Member of the Polish Academy of Sciences; Fellow of the International Fuzzy Systems Association; LIFE Honorary President of the Hungarian Fuzzy Association.

A Hybrid Syntactic–Statistical–Semantic Framework for Detecting AI-Generated Text Across Domains

Doaa Mostafa, Sally S. Ismail, and Mostafa Aref

***Abstract*—Recent advances in large language models (LLMs) have enabled highly human-like text generation, raising concerns related to misinformation, authorship verification, and academic integrity. Current approaches for detecting LLM-generated text suffer from several limitations, including limited robustness to linguistic diversity, sensitivity to text length variations and paraphrasing, weak domain generalization, and high computational cost. To address these challenges, this paper proposes a hybrid framework for detecting LLM-generated text that integrates syntactic and statistical features with deep semantic representations learned using GloVe embeddings, Convolutional Neural Networks (CNNs), and Bidirectional Long Short-Term Memory (BiLSTM) networks. By combining linguistic cues with contextual semantics, the proposed model captures both structural and semantic patterns to distinguish human-written text from LLM-generated content. Experiments conducted on the ChatGPT Research Abstracts and ElectAI datasets demonstrate strong cross-domain generalization and robustness to text length variations and paraphrasing. The proposed framework achieves an accuracy of 98.63%, an F1-score of 98.66%, and a minimum false positive rate (FPR) of 0.01. These results indicate the effectiveness, stability, and reliability of the framework for detecting LLM-generated text.**

***Index Terms*—Large language models (LLMs), AI-generated text, Text generation, Word Embedding, Feature Extraction.**

I. INTRODUCTION

Recent advances in natural language generation (NLG) have significantly improved the fluency, coherence, and diversity of the text produced by LLMs. Advanced generative models such as GPT-4 [1], Claude [2], and Gemini [3] now generate writing that is becoming sufficient to achieve human-level performance across tasks such as question answering, email drafting, news article composition, scientific writing, story generation, and code generation. However, alongside these capabilities, significant concerns regarding the potential misuse in areas such as phishing [4], misinformation [5], and academic integrity [6-8]. Current approaches for identifying LLM-generated text generally fall into four categories: watermarking, feature-based, neural-based, and human-assisted approaches [9]. Watermarking approaches embed hidden signals into text during the generation process; however, they require full access to the underlying model, making them impractical for most modern LLMs. Also, watermarking methods are sensitive to minor text modifications, such as adding space, substituting smaller words with larger words, using similar words, which can significantly reduce detection accuracy [10]. Feature-based approaches rely on linguistic indicators such as syntax, grammar, and other stylistic features. They can also suffer degradation in effectiveness when applied to advanced LLMs[10]. Neural-based approaches use pre-trained transformer models [11,12] to classify LLM-generated content. Although these methods achieve strong performance, they require substantial computational resources and large annotated datasets for training [9,10]. Finally, human-assisted approaches rely on human judgment, however, as LLM-generated text becomes increasingly indistinguishable from human writing, reliable manual detection becomes progressively more challenging [10].

To address the limitations of existing detection approaches, this paper proposes a hybrid framework for identifying LLM-generated content that combines handcrafted syntactic, statistical features and deep semantic representations derived from GloVe [13], CNNs [14], and BiLSTM networks together. By using a hybrid framework with interpretable linguistic features and high-level context-based semantics, the proposed framework achieves improved reliability and generalizability across multiple domains and diverse LLMs.

The structure of this paper is as follows. Section 2 reviews the most recent literature on detecting text generated by LLMs. Section 3 presents the proposed framework. Section 4 describes the datasets, evaluation metrics, experimental results, and the discussion of results. Finally, Section 5 will conclude the paper and provide potential avenues for future research.

II. RELATED WORK

This section provides background on the most recent research on detecting LLM-generated text. Recent approaches to detecting LLM-generated text can be broadly categorized into watermarking-based, neural-based, and feature-based detection methods.

Doaa Mostafa, Sally. S. Saad, and Mostafa Aref are affiliated with Computer Science Department, Faculty of Computer and Information science, Ainshams University, Cairo, Egypt. (E-mails: doaa.ahmed74@yahoo.com, sallysaad@cis.asu.edu.eg, Mostafa.aref@cis.asu.edu.eg).

DOI: 10.36244/ICJ.2026.1.6

Watermarking Framework for LLMs (WLLM) [15] embeds imperceptible signals into generated text by biasing token selection during generation. Although WLLM is lightweight and efficient in controlled environments, it essentially depends on access to the generation process and is extremely susceptible to paraphrasing or post-editing, which can greatly weak or eliminate the watermark [9,16]. This limits its applicability in real-world, open-text scenarios.

REMARK-LLM [17] increases the robustness of watermarks by embedding identifiers into internal semantic representations and decoding them via a retrieval-based mechanism. Although this enhances resistance to mild paraphrasing, the approach incurs higher computational costs and still degrades under stronger paraphrasing attacks. Moreover, it requires specialized retrieval and decoding components, reducing scalability.

In DNA-GPT [18], genetic signatures are incorporated in the text generation process in the form of marks, which supports traceability. In spite of achieving a high level of detection accuracy, this model is difficult to implement and requires access to the generation pipeline. The model is also affected by fluctuations in text length.

Giant Language Model Test Room (GLTR) [19] in order to identify LLM-generated text by analyzing token-level statistics like likelihood, rank, and entropy. Although GLTR is interpretable and efficient for older language models, it is less reliable for advanced LLM-generated text, edited text, and paraphrased text and requires access to model probability outputs. AuthentiGPT [20] uses a feature-based, multi-stage detection approach that integrates watermarking signals, semantic embeddings, and linguistic cues. Despite its effectiveness, the method's practicality for large-scale or real-time deployment is limited by the need for substantial computational resources and large labeled datasets.

SeqXGPT[21] is a sentence-level approach that employs sequential models and contextual embeddings to identify linguistic and semantic inconsistencies between human-written and LLM-generated text. This approach generalizes well across various LLM types due to its use of sentence-based architecture. However, the approach is computationally expensive, as processing sentences individually increases resource demands, and its accuracy may drop when sentences are edited or when stylistic cues become harder to distinguish.

In summary, despite recent advances in current detection approaches, they face important restrictions, including dependence on access to the source generation model, prohibitive computation cost, and weak generalizability across models and domains. Many existing detectors also struggle with paraphrased input and short text segments. These issues demonstrate the need for more generalizable, interpretable, and text- level detection. The proposed framework, which utilizes a combination of crafted linguistic and statistical features alongside deep semantic representations, addresses these gaps in performance and detection mode through robust and reliable detection performance.

III. PROPOSED ARCHITECTURE

The proposed framework for detecting LLM-generated text consists of six stages: (1) Preprocessing, in which the input text is cleaned by removing irrelevant elements. (2) Handcrafted statistical and syntactic feature extraction, in which statistical and syntactic features are computed from pre-processed text. (3) Text representation using GloVe embeddings, which convert each token into a fixed-length dense vector that captures semantic relationships through word co-occurrence patterns. (4) Semantic feature extraction, where a CNN captures local contextual patterns and a BiLSTM models long-range dependencies. (5) Feature fusion, in which handcrafted statistical and syntactic features are concatenated with the semantic features produced by the CNN and BiLSTM to form a unified feature representation. (6) Classification was then performed by taking the combined feature vector and passing it through a fully connected layer to finally predict the label. The architecture of the proposed framework is illustrated in Figure 1.

A. Phases of the Proposed Framework

1) Text Preprocessing

The goal of preprocessing is to remove task-irrelevant content (e.g., URLs, emails, symbols, hashtags, numbers) that do not contribute to the linguistic, grammatical, or semantic features used in this study [22]. Such tokens were found to have minimal impact on detection performance and may introduce noise, particularly in short or informal texts, as the MFAD framework primarily relies on stylistic and semantic cues to identify AI-generated text. Therefore, preprocessing uses lowercasing, normalization, lemmatization, and tokenization to get the text ready for trustworthy feature extraction.

2) Statistical and Syntactic Features Extraction

Statistical and syntactic features capture the underlying structure of the text and reflect key indicators such as readability and writing style, which serve as strong cues for assessing textual originality and coherence. Lexical diversity is quantified through measures of vocabulary richness. The statistical features represent measurable aspects of the textual structure, style, and complexity. Part-of-speech (POS) tag frequency, and bigram frequency are used to measure syntactic complexity, which expose syntactic and lexical patterns indicative of human-written versus LLM-generated text [23]. A detailed list of the statistical and syntactic features used in this study is provided in Table I.

3) Text Representation

The final step of text preprocessing is tokenization, which the text is split into individual tokens. These tokens are then passed to GloVe, which models both global statistical relationships and local contextual meanings among words in the dataset. GloVe represents each word as a dense vector in continuous space, typically with 50, 100, 200 or 300 dimensions, where semantically similar words are positioned closer together.

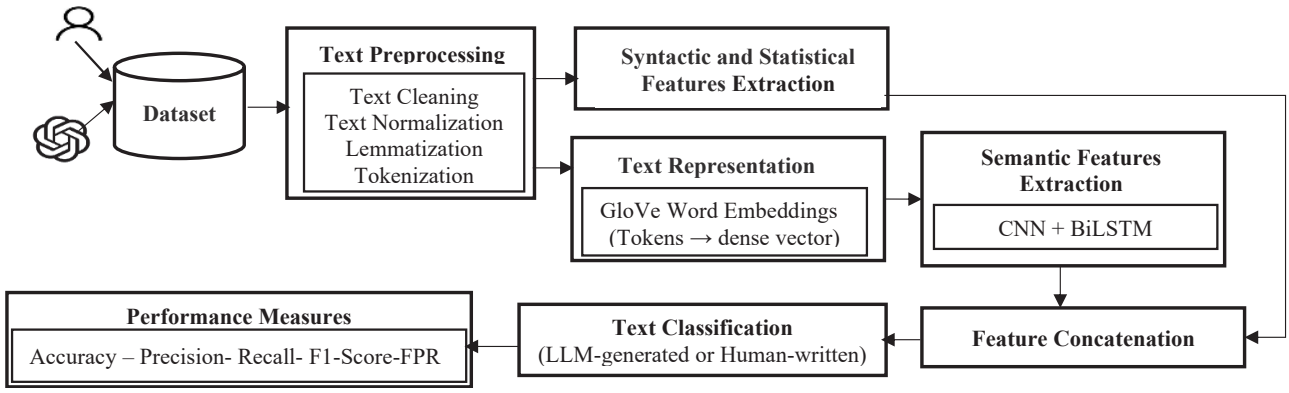


Fig. 1. Architecture of the proposed framework

These embeddings, derived from word co-occurrence statistics in a large corpus, enable GloVe to model semantic relationships very well. In contrast to context-dependent models such as BERT, GloVe embeddings are context-independent in nature, and the model assigns every word the same vector without any additional context-based meaning, preserving semantic proximity based on global co-occurrences. The numerical vectors generated by GloVe are then passed to the CNN for higher-level feature extraction and semantic representation learning. The embedding matrix is constructed according to equations (1–3).

Given a sequence of T tokens:

$$X = [w_1, w_2, \dots, w_T] \quad (1)$$

Each token w_t is mapped to its GloVe embedding:

$$\mathbf{e}_t = \text{GloVe}(w_t) \in \mathbb{R}^d \quad (2)$$

The embedding matrix is constructed as follows:

$$E = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_T]^\top \in \mathbb{R}^{T \times d} \quad (3)$$

where T is the length of the input sequence, w_t is the token at position (t), et is the GloVe embedding vector for the token, d is the embedding dimensionality, and E is the sequence embedding matrix.

4) Semantic Feature Extraction

High-level semantic features were extracted using a combination of CNN and BiLSTM architectures. The CNN consists of convolutional and pooling layer. The convolutional layer learns patterns within the local context and the pooling layer reduces the dimensionality of the features for computational efficiency and robustness. As convolutional filters slide across the text, the CNN captures local n-grams and phrase-level contextual representations, with max pooling selecting the most informative activations from each filter. The obtained local feature maps are then passed into a Bidirectional Long Short-Term Memory (BiLSTM) layer to model long-range dependencies by processing the sequence in both forward and backward directions. This allows each word to be interpreted within its full context. By integrating CNN for capturing local patterns and BiLSTM for modeling global semantic connections into one structure, the framework produces a rich and comprehensive contextual representation.

5) Feature Concatenation

The numerical vectors representing syntactic and statistical features were fused with the semantic feature vectors generated by the CNN+BiLSTM model to form a unified feature vector for each text sample. These feature types provide complementary information: syntactic and statistical features measure structural and quantifiable properties of writing, and the semantic features describe meaning and contextual relations between words. Together, these features yield a richer and more comprehensive representation of the text, thereby enhancing the model's ability to distinguish human-written content from LLM-generated text.

6) Text Classification

Subsequently, the concatenated feature vector is fed into a classifier implemented using the TensorFlow framework. The classifier consists of a fully connected (dense) neural network with a sigmoid activation function in the output layer. This activation function maps the network output to a value between 0 and 1, corresponding to the probability of being in either of two classes, which is well suited for binary classification tasks, predicting whether a text is written by a human or generated using an LLM.

B. Rationale for Model Architecture.

We adopt a GloVe + CNN + BiLSTM architecture rather than fine-tuned Transformer-based models like BERT and DistilBERT for several reasons. Primarily, our aim is to propose an efficient and interpretable detection mechanism that can generalize across domains and LLMs without invoking the need to perform heavy fine-tuning. Transformer-based detectors often involve substantial computational cost involved in training and inference processes of transformer models, as well as sensitivity to domain-specific data distributions.

In contrast, GloVe provides stable corpus-level semantic representations. CNNs can also adequately represent local n-gram information, and BiLSTMs can represent long-range dependencies. This provides effective modeling of semantics but also with a lower computational cost. Additionally, separating handcrafted features from deep semantic

representations allows clearer attribution of performance gains, which supports more interpretable analysis and ablation studies

TABLE I
STATISTICAL AND SYNTACTIC FEATURES

Type	Feature	Description
Statistical [23]	Total Words	The total count of words.
	Total Sentences	The total number of sentences.
	Total Unique Words	The total count of unique words.
	Type-Token Ratio	TTP evaluates vocabulary diversity.
	Total Stop Words	The total count of stop words.
	Total Punctuation	The total count of punctuation marks (commas, periods, exclamation marks, etc.).
	Total Discourse Markers	The total number of discourse markers.
	Total Spelling Errors	The total count of spelling errors.
	Total Grammar Errors:	The total count of grammar errors.
	Readability Scores	The ease of reading the text was evaluated; higher scores indicated easier readability.
	Syllable Count	Total number of syllables.
	Average-Sentence Length	The average number of words per sentence.
	Average word length	The average number of characters per word.
Syntactic [23]	Part-of-Speech (POS) Tag Distributions	Frequencies of noun singular (NN), noun plural (NNS), verb base-form (VB), verb past tense (VBD), adjectives (JJ), adverbs (RB), personal pronoun (PPR), preposition (IN), verb present participle (VBG), contracting conjunction (CC), and determiners (DT).
	Complexity of Sentence	Tree depth: The depth of the syntactic tree reflects sentence complexity, such as the average dependency tree, depth, maximum of dependency tree depth, and number of subordinate clauses.
	Top Bigram/Trigram Frequency	Calculate the maximum number of two- or three-word combinations.

IV. RESULTS

A. Datasets

The ChatGPT Research Abstracts [24] and ElectAI [25] datasets were used for the experiments. The ChatGPT Research Abstracts dataset contains 10,000 titles for papers with a combination of Human and ChatGPT-generated (GPT 3.5) abstracts for each title, and can be used to differentiate between human and AI-generated text. The ElectAI dataset is a collection of English tweets that refer to elections and political claims, tagged as human-written or AI generated, and contains approximately 9,400 tweets. The dataset has been produced using numerous LLMs including Llama-2-7B [26], Mistral-7B [27] and Falcon-7B [28]. Table II provides an overview of the dataset metadata and composition.

TABLE II
METADATA AND DATASET COMPOSITION OF THE CHATGPT RESEARCH ABSTRACTS AND ELECTAI DATASETS.

Dataset	Generation Model	Domian	LLM generated articles Count	Real Data Count (Human)
ChatGPT Research Abstracts	GPT 3.5	Scientific Writing	10000	10000
ElectAI	Llama-2-7B	Political Tweets	2350	2350
	Mistral-7B		2350	
	Falcon-7B		2350	

B. Experimental Setup

Each experiment was conducted on a Lenovo laptop equipped with 12 GB of RAM, an Intel Core i5 processor, and a 64-bit operating system. The proposed framework was implemented in Python 3.10.5 using several libraries: scikit-learn [29] for computing evaluation metrics, NLTK [30] for preprocessing, such as tokenization, stemming, lemmatization, and stop-word removal, and TensorFlow [31] for constructing and training deep neural networks. Additional NLP libraries were used to extract the statistical and syntactic features listed in Table I. In particular, advanced syntactic analyses, including syntactic complexity and dependency-based features, were performed using spaCy [32]. Text complexity and readability metrics were computed with TextStat [33]. Grammatical errors were detected using LanguageTool [34]. The number of spelling mistakes in the text is calculated using PySpellChecker [35]. The hyperparameter configuration is provided in Table III, and includes batch size, number of epochs, optimizer, and dropout rate.

C. Performance Measures

Five classification metrics were used to evaluate the performance of the proposed framework: accuracy, precision, recall, F1-score, and false positive rate (FPR) [9]. Accuracy measures the proportion of texts that are classified correctly, as shown in Eq. (4). Precision represents the proportion of texts classified as AI-generated, that are truly AI-generated, as shown in Eq. (5). Recall (True Positive Rate) measures the proportion of AI-generated texts that were correctly classified as AI-generated, as defined in Eq. (6). The F1 score is the harmonic mean of precision and recall and defined as a uniform indicator of both recall and precision, as given in Eq. (7). Finally, the FPR quantifies the proportion of human-written texts that are incorrectly classified as AI-generated, as defined in Eq. (8).

$$\text{Accuracy} = \frac{(\text{Tp}+\text{TN})}{(\text{TP}+\text{FP}+\text{TN}+\text{FN})} \quad (4)$$

$$\text{Precision} = \frac{\text{Tp}}{(\text{TP}+\text{FP})} \quad (5)$$

$$\text{Recall} = \frac{\text{Tp}}{(\text{TP}+\text{FN})} \quad (6)$$

$$\text{F1-score} = 2\times\frac{(\text{Precision}*\text{Recall})}{(\text{Precision}+\text{Recall})} \quad (7)$$

$$\text{FPR} = \frac{\text{FP}}{(\text{FP}+\text{TN})} \quad (8)$$

Here, TP (True Positive) denotes to the number of LLM-generated texts that are correctly classified as LLM-generated. TN (True Negative) represents the number of human-written texts that are correctly classified as human-written. FP (False Positive) refers to human-written texts that are incorrectly classified as LLM-generated and FN (False Negative) refers to LLM-generated texts that have been incorrectly classified as human written.

TABLE III
THE HYPERPARAMETER SETTINGS

Parameters	Value
Activation Function (feature extraction)	ReLU
Activation function (classification)	Sigmoid
Optimizer	Adam
Loss Function	binary_crossentropy
Batch size	32
No. of. Epochs	5
Dropout-rate	0.5
Glove Embedding dimensions	100
Pooling Type	Max
Kernel-Size	3
Dataset splits	70% → training, 10→ validation, and 20% → testing

D. Experimental Results

This section presents the performance of the proposed framework on the ChatGPT Research Abstracts and ElectAI datasets, including detection performance across different dataset categories, robustness to varying text lengths, resilience to paraphrasing, and comparison with baseline approaches.

1) Evaluation of the Proposed Framework on the Two Datasets

The proposed framework is evaluated on the ChatGPT Research Abstracts and ElectAI datasets, as illustrated in Figures 2–3.

These figures illustrate the effectiveness of the proposed framework in distinguishing human-written text from LLM-generated text on both the ChatGPT Research Abstracts and ElectAI datasets. In the ChatGPT Research Abstracts dataset, the model achieves 96.63% accuracy with exceptionally high precision (96.52%) and recall (97.61%), resulting in a strong F1-score of 97.06% and an exceptionally low false positive rate (FPR) of .02. In the ElectAI dataset, the results indicate the model's efficacy across various generation models. In the Human vs Falcon category, the model reaches 96.65% accuracy with 95.60% precision and 97.53% recall. Notably, this performance increases when detecting text generated by Mistral (98.3%) and LlaMA (98.63%). These results confirm the framework's ability to generalize across domains and various LLMs, along with reasonable levels of precision and excellent recall rates, and very few false positives.

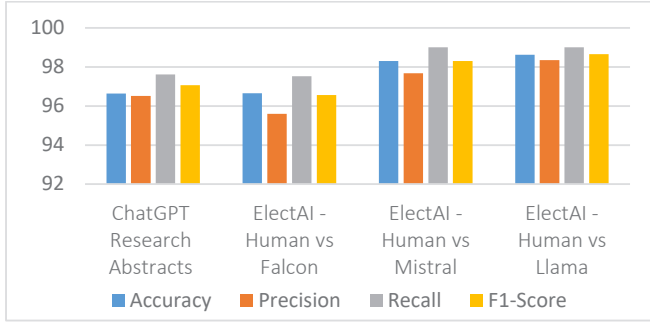


Fig. 2. Evaluation metrics of the proposed framework on the ChatGPT Research Abstracts and ElectAI datasets.

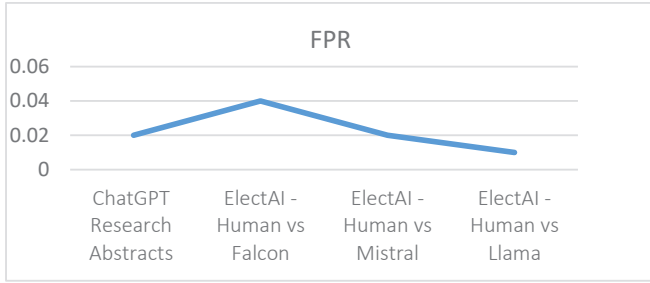
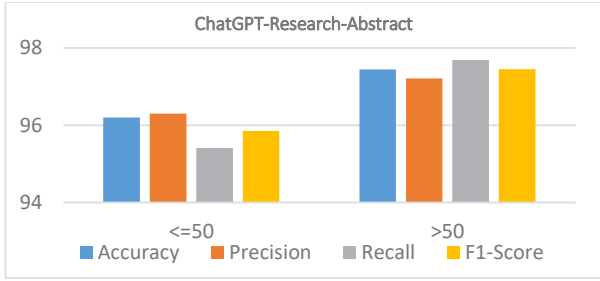


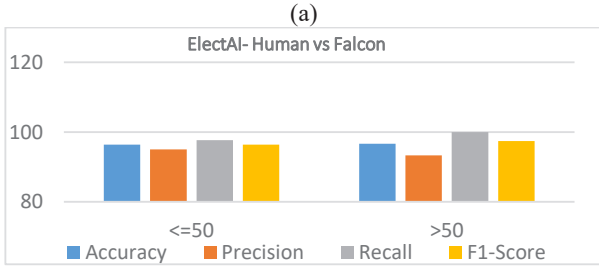
Fig. 3. FPR of the proposed framework on the ChatGPT Research Abstracts and ElectAI datasets.

2) Evaluation of the Proposed Framework on the Two Datasets Across Different Text Lengths

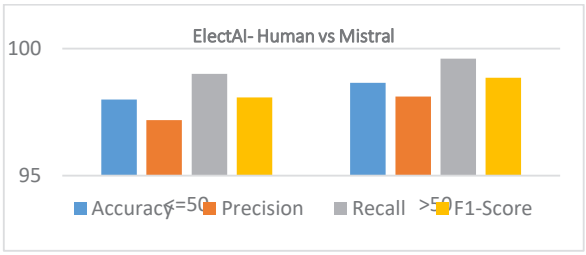
In this section, the proposed framework is evaluated under varying text lengths by partitioning each dataset into two token-count categories: texts with fewer than 50 tokens and texts with more than 50 tokens. The corresponding evaluation results are presented in Figure 4.



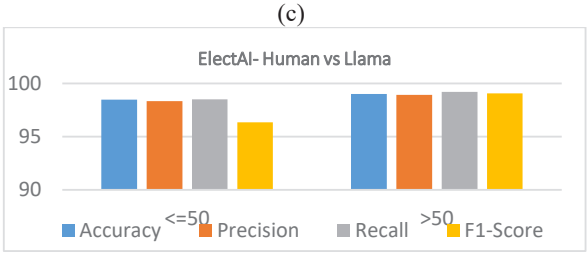
(a)



(b)



(c)



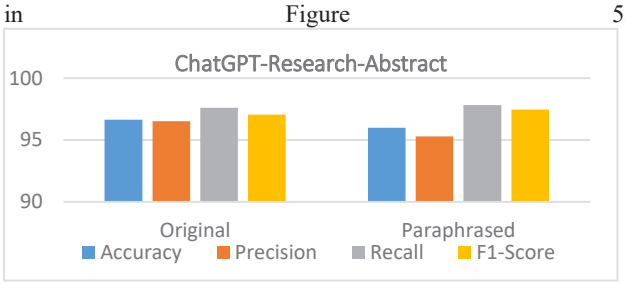
(d)

Fig. 4. Evaluation metrics of the proposed framework on ChatGPT Research Abstracts and ElectAI datasets for different text lengths

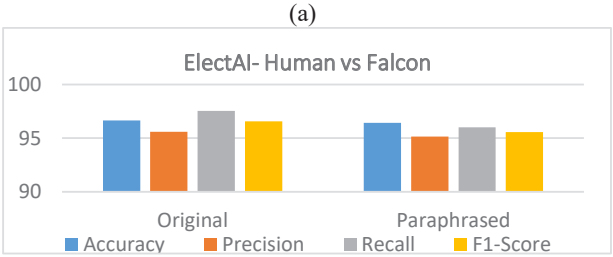
Figure 4 illustrates the evaluation of the proposed framework under varying text lengths on the ChatGPT Research Abstracts and ElectAI datasets supports the underlying robustness and stability of the framework, regardless of the text input length. On the ChatGPT Research Abstracts dataset, the model achieves strong performance with both short (≤ 50 words) and longer (> 50 words) texts. Notably, accuracy increased from 96.20% with short texts to 97.44% for longer texts. This trend continues throughout the ElectAI dataset across all generation models. In the Human vs Falcon classification, the model achieves strong accuracy for both short (96.42%) and long (96.67%) sample versions, with 100% recall being provided for the longer text. Additionally, performance improved among Mistral and LlaMA classifications, with both short and long text versions producing very high accuracy, precision, recall and F1-score measures. Across all text lengths, the results remained above 98% for almost every metric. This indicates that the proposed method is highly effective regardless of text size, consistently achieving strong precision and recall while generalizing well across different AI models and text-length variations.

3) Evaluation of the Proposed Framework on the Two Datasets under Text Paraphrasing

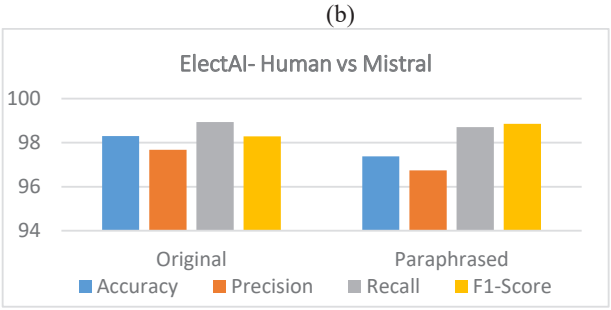
The proposed framework was evaluated on its ability to detect paraphrased LLM-generated texts that preserve semantic meaning while introducing syntactic variation. Paraphrased samples were generated using a pre-trained T5-base transformer model, which follows a text-to-text learning paradigm [36]. The model employs an encoder–decoder architecture, where the decoder creates a rephrased sentence based on the contextualized semantic representation produced by the encoder. The generated paraphrases introduce lexical and syntactic variation, including word substitutions, reordering, and structural reformulation, without changing the underlying meaning, T5's paraphrased outputs are intended to be semantically equivalent to the original text. This allows the evaluation of the robustness of the proposed framework under realistic paraphrasing-based text transformations. After the paraphrasing process, the proposed framework was applied to the modified dataset, and its performance was compared with that obtained on the original, non-paraphrased test set. The evaluation results under paraphrasing conditions are presented in Figure 5.



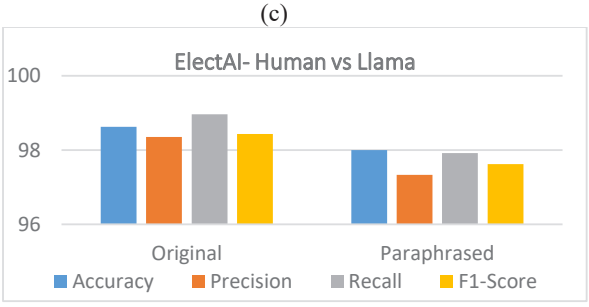
(a)



(b)



(c)



(d)

Fig. 5. Evaluation Metrics of the Proposed framework Between Original and Paraphrased Texts in the ChatGPT Research Abstracts and ElectAI Datasets.

Figure 5 illustrates the evaluation of the proposed framework under paraphrasing conditions and shows that the model remains highly effective when confronted with syntactically altered but semantically equivalent text. On the ChatGPT Research Abstracts dataset, performance declines only marginally, with accuracy decreasing from 96.63% on the original texts to 96% on paraphrased versions. Meanwhile, recall exhibits a slight increase, indicating sensitivity to AI-generated patterns despite syntactic variation. A similar trend is

observed on the ElectAI dataset. In the Human vs. Falcon category, performance remains stable, with only a minimal decrease in accuracy and F1-score after paraphrasing. Even for more advanced models such as Mistral and LLaMA, the framework continues to perform strongly, achieving accuracies of 97.38% and 98% on paraphrased texts, respectively, which are close to the results obtained on the original data. Although minor reductions in precision and F1-score are observed, overall detection performance remains robust. These results demonstrate that the proposed framework generalizes well to paraphrased LLM-generated content and effectively preserves deeper semantic and stylistic cues even after substantial rewriting.

4) Comparison of the Proposed Framework with Baseline Approaches

This section compares the proposed framework with representative baseline approaches. As shown in Table IV, the proposed framework consistently outperforms AuthentiGPT and SeqXGPT on both the ChatGPT Research Abstracts and ElectAI datasets.

TABLE IV
PERFORMANCE ANALYSIS OF THE PROPOSED FRAMEWORK IN COMPARISON WITH BASELINE APPROACHES

Dataset	Approach	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ChatGPT Research Abstracts	AuthentiGPT	88	84	87	85.5
	SeqXGPT	85.9	85.5	90.5	88
	Proposed	96.6	96.5	97.6	97
ElectAI (All)	AuthentiGPT	92	86.3	90.7	88.5
	SeqXGPT	86.4	82.3	93.3	87.5
	Proposed	97.8	97.2	98.5	98

The results in Table IV show that the proposed framework outperforms AuthentiGPT and SeqXGPT on both the ChatGPT Research Abstracts and ElectAI datasets. On the ChatGPT Research Abstracts dataset, the proposed framework achieves an accuracy of 96.6%, outperforming AuthentiGPT by 8.6 percentage points and SeqXGPT by 10.7 percentage points. Precision, recall, and F1-score all show comparable improvements. Notably, the proposed framework achieves a 97% F1-score, which is much higher than that of SeqXGPT (88%), and AuthentiGPT (85.5%). SeqXGPT's exhibits relatively high recall (90.5%), indicating sensitivity to AI-generated content; however, its lower precision and overall accuracy suggest a higher false-positive rate. On the other hand, the proposed framework effectively distinguishes between text produced by AI and text written by humans while maintaining high recall and precision (97.6%, 96.5%). A similar trend is observed on the ElectAI dataset. The proposed framework achieves an accuracy of 97.8%, outperforming AuthentiGPT (92%) and SeqXGPT (86.4%). It also records the highest precision (97.2%) and recall (98.5%), resulting in an F1-score of 98%. These results highlight the proposed framework's strong generalization capability across different domains and writing styles. While AuthentiGPT performs competitively on this dataset, its lower precision and F1-score suggest limitations in capturing nuanced stylistic and semantic differences. SeqXGPT again demonstrates relatively high recall (93.3%) but lower precision (82.3%), reinforcing the trade-off between sensitivity and specificity observed in feature-based baselines.

5) Ablation Study

To evaluate the contribution of each component in the proposed framework, an ablation study will be carried out by considering different feature types, sequences modeling architectures, and word embeddings on the ChatGPT Research Abstracts and ElectAI datasets. First, we evaluate handcrafted features syntactic and statistical stylometric indicators, to measure their individual contribution. Next, we evaluate semantic features extracted using CNN combined with sequence models such as RNN, LSTM, and BiLSTM, with GloVe and Word2Vec embeddings. We then examine the impact of feature fusion by combining handcrafted features with semantic features, where CNN and RNN are used with GloVe embeddings. Finally, the reference model combines handcrafted features with semantic features, which are learned with GloVe, CNN, and BiLSTM techniques, to serve as a reference model to evaluate all the individual metrics.

The results in Table V show that handcrafted syntactic and statistical cues are insufficient for reliable AI-generated text detection, as handcrafted features alone only achieve relatively low accuracy on both ChatGPT Research Abstracts (69.8%) and ElectAI (60.7%). Semantic-only configurations substantially improve performance, consistently exceeding 90% accuracy, highlighting the importance of deep semantic representations. The benefit of bidirectional contextual modeling is confirmed by the fact that using a standard RNN in place of BiLSTM causes a discernible drop in performance, while LSTM improves results but is still marginally less effective than the full proposed framework. Using Word2Vec embeddings further enhances semantic-only performance. The combined framework, integrating semantic and handcrafted features, achieves the highest accuracy and F1-score, demonstrating the complementary strengths of contextual and stylometric patterns.

E. Discussion

The experimental results confirmed the effectiveness and robustness of the proposed hybrid method for identifying LLM-generated. The most significant result from this study shows how important it is to combine different types of features (syntactic, statistical, and deep semantic) to achieve a reliable and generalized performance when detecting LLM-generated text. The addition of statistical and syntactical features accurately captures the surface structure and stylistic characteristics of LLM-generated text and the subtle differences between the writing styles of LLMs and humans. At the same time, using CNNs and BiLSTMs provides the model with a deeper understanding of the semantics of the text, enabling the model to detect more subtle differences in linguistic behaviour and the coherent use of word meaning. This framework shows strong performance (i.e., accuracy, precision, recall) as well as a low false-positive rate across both the scientific abstracts and political tweets datasets.

Generalization across different LLMs (GPT-3.5, LLaMA-2,

Mistral, Falcon) indicates robustness across varying datasets and styles of generation. The framework is highly consistent despite changing lengths of text, both less than and more than 50 tokens in length. Also, this framework maintains strong performance when classifying items that had been syntactically modified (paraphrased) using the T5 Model. For instance, paraphrasing the ChatGPT Research Abstracts dataset with T5 resulted in only a minor accuracy drop from 96.63% to 96%. The proposed framework consistently showed superior performance across all datasets and LLM categories When compared to the feature-based baseline AuthentiGPT and SeqXGPT. Therefore, the combination of multiple feature types is an effective way of increasing the sensitivity and reliability of an LLM text detector. Finally, the paraphrasing evaluation focuses on controlled, single-step neural rewriting to simulate common automated rewriting attacks. Future work will investigate more challenging cases, including multi-step paraphrasing, back-translation, and human-edited paraphrases, to further validate the framework's robustness.

TABLE V
ABLATION STUDY RESULTS OF THE PROPOSED FRAMEWORK

Dataset	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ChatGPT Research Abstracts	Handcrafted Features only	69.8	68.8	72.7	70.7
	Semantic Features only (Glove+ CNN+BiLSTM)	95.6	95	96.8	95.89
	Semantic Features only (Glove+ CNN+LSTM)	94.8	94.4	92.8	93.6
	Handcrafted+ Semantic (Glove+ CNN+RNN)	92.6	93	92.3	92.6
	Semantic Features only (Word2Vec+ CNN+BiLSTM)	94.6	93.9	96	95
	Proposed Framework	96.6	96.5	97.6	97
ElectAI (All)	Handcrafted Features only	60.7	62.1	59.2	60.6
	Semantic Features only (Glove+ CNN+BiLSTM)	96.2	96.5	97.6	97
	Semantic Features only (Glove+ CNN+LSTM)	94.6	95.7	96.6	96
	Handcrafted+ Semantic (Glove+ CNN+RNN)	93.7	92	92.8	93
	Semantic Features only (Word2Vec+ CNN+BiLSTM)	95.22	94.6	96.8	95.7
	Proposed Framework	97.8	97.2	98.5	98

V. CONCLUSION

This paper presents a hybrid detection framework that combines syntactic and statistical features with deep semantic representations derived from CNN and BiLSTM models. The proposed framework is effective in differentiating between human-written text and LLM-generated text across multiple domains and generative models by leveraging complementary feature types. Experiments on scientific abstracts and political tweet demonstrate high performance, with the framework achieving a maximum accuracy of 98.63%, an F1-score of 98.66%, and a minimum false positive rate (FPR) of 0.01. This framework also demonstrates strong robustness to variations in text length and strong resilience in detecting paraphrased LLM-generated content. These results indicate the advantage of fusing different feature types to improve LLM-generated text detection. Future work will involve extending the use of this framework with larger and more diverse multilingual datasets and evaluating its performance on data generated from more advanced LLMs. Additionally, this framework will be enhanced to improve adversarial robustness, such as real-world adversarial paraphrasing, and provide a more comprehensive comparative assessment with current detection approaches.

REFERENCES

[1] J. Achiam, S. Adler, S. Agarwal, and D. Almeida, "Gpt-4 technical report," *arXiv preprint; arXiv:2303.08774*, 2023, DOI: 10.48550/arXiv.2303.08774.

[2] A. Priyanshu, Z. Hong, and Y. Maurya, "AI Governance and Accountability: An Analysis of Anthropic's Claude," *arXiv preprint; arXiv:2407.01557*, 2024, DOI: 10.48550/arXiv.2407.01557.

[3] G. Team, R. Anil, J. Yu, R. Soricut, S. Borgeaud, J. B. Alayrac, and J. Schalkwyk, "Gemini: a family of highly capable multimodal models," *arXiv preprintar; arXiv:2312.11805*, 2023, DOI: 10.48550/arXiv.2312.11805.

[4] A. Giaretta and N. Dragoni, "Community targeted phishing, "in *Proc. 6th Int. Conf. Softw. Eng. Defence Appl.*, P. Ciancarini, M. Mazzara, A. Messina, A. Sillitti, and G. Succi, Eds. Cham, Switzerland: Springer, 2020, pp. 86–93, DOI: 10.1007/978-3-030-14687-0_8.

[5] K. Shu, S. Wang, D. Lee, and H. Liu, "Mining disinformation and fake news: Concepts, methods, and recent advancements, "in *Disinformation, Misinformation, and Fake News in social media. Cham*, Switzerland: Springer, pp.1–19, 2020, DOI: 10.48550/arXiv.2001.00623.

[6] N. Dehouche, "Plagiarism in the age of massive generative pre-trained transformers (GPT-3), "*Ethics Sci. Environ. Politics*, vol. 21, pp. 17–23, 2021, DOI: 10.3354/esep00195.

[7] R. Tang, and Y. N. Chuang, "The science of detecting llm-generated text." *Communications of the ACM* 6;, no. 4, pp. 50–59, 2024, DOI: 10.48550/arXiv.2303.07205.

[8] L. Floridi, and M. Chiriatti, "GPT-3: Its nature, scope, limits, and consequences." *Minds and Machines 30*; pp. 681–694, 2020, DOI: 10.1007/s11023-020-09548-1.

[9] J. Wu, L. S. Chao, S. Yang, Y. Yuan, and D. F. Wong, "A survey on LLM-generated text detection: Necessity, methods, and future directions." Computational Linguistics; pp. 1–65, 2025, DOI: 10.48550/arXiv.2310.14724.

[10] K. C. Fraser, H. Dawkins, and S. Kiritchenko, "Detecting ai-generated text: Factors influencing detectability with current methods." *Journal of Artificial Intelligence Research82*, pp. 2232–2278, 2025, DOI: 10.48550/arXiv.2406.15583.

[11] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding, "*arXiv preprint; arXiv:1810.04805*, 2018, DOI: 10.48550/arXiv.1810.04805.

[12] Y. Liu, M. Joshi, D. Chen, M. Ott, N. Goyal, J. Du, O. Levy, and M. Lewis, L. Zettlemoyer, and V. Stoyanov. "Roberta: A robustly optimized bert pretraining approach," *arXiv preprint; arXiv:1907.11692*, 2019, DOI: 10.48550/arXiv.1907.11692.

[13] J. Pennington, R. Socher, and C.D. Manning, "Glove: Global vectors for word representation," In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532–1543, 2014, DOI: 10.3115/v1/D14-1162.

[14] H. Zhou, "Research of text classification based on TF-IDF and CNN-LSTM", In *journal of physics: conference series*; vol. 2171, no. 1, p. 012 021. IOP Publishing, 2022, **DOI**: 10.1088/1742-6596/2171/1/012021.

[15] J. Kirchenbauer, J. Geiping, Y. Wen, J. Katz, I. Miers, and T. A. Goldstein, "A watermark for large language models", In *International Conference on Machine Learning*. PMLR, 2023, **DOI**: 10.48550/arXiv.2301.10226.

[16] K. C. Fraser, H. Dawkins, and S. Kiritchenko, "Detecting ai-generated text: Factors influencing detectability with current methods", *Journal of Artificial Intelligence Research* 82: 2233–2278, 2025, **DOI**: 10.48550/arXiv.2406.15583.

[17] R. Zhang, S. Hussain, P. Neekhara, and F. Koushanfar, "Remark-llm: A robust and efficient watermarking framework for generative large language models", *arXiv preprint arXiv:2310.12362*, 2024, **DOI**: 10.48550/arXiv.2310.12362.

[18] X. Yang, W. Cheng, Y. Wu, L. Petzold, W. Y. Wang, and H. Chen, "Dna-gpt: Divergent n-gram analysis for training-free detection of gpt-generated text," *arXiv preprint arXiv:2305.17359*. 2023, **DOI**: 10.48550/arXiv.2305.17359.

[19] S. Gehrmann, and A. M. Rush, "Gltr: Statistical detection and visualization of generated text", *arXiv preprint; arXiv:1906.04043*, 2019, **DOI**: 10.48550/arXiv.1906.04043.

[20] Z. Guo, and S. Yu, "Authentigpt: Detecting machine-generated text via black-box language models denoising", *arXiv preprint arXiv:2311.07700.2023*, **DOI**: 10.48550/arXiv.2311.07700.

[21] P. Wang, L. Li, B. Jiang, D. Zhang, K. Ren, and X. Qiu, "SeqXGPT: Sentence-level AI-generated text detection", *arXiv preprint; arXiv: 2310.08903*, 2023, **DOI**: 10.48550/arXiv.2310.08903.

[22] S. Alam, and N. Yao, "The impact of preprocessing steps on the accuracy of machine learning algorithms in sentiment analysis, "Computational and Mathematical Organization Theory, 2019, **DOI**: 10.1109/IISA62523.2024.10786699.

[23] C. Opar, "StyloAI: Distinguishing AI-generated content with stylometric analysis," In *International conference on artificial intelligence in education*, pp. 105–114. Cham: Springer Nature Switzerland, 2024, **DOI**: 10.48550/arXiv.2405.10129.

[24] N. T. Sivesind, "Chat GPT-Generated-Abstracts", Hugging Face, 2023.

[25] A. Dmonte, M. Zampieri, K. Lybarger, M. Albanese, and G. Coulter. "Classifying human-generated and ai-generated election claims in social media." *arXiv preprint arXiv:2404.16116,* 2024, **DOI**: 10.48550/arXiv.2404.16116.

[26] H. Touvron, L. Martin, K. Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov *et al*. "Llama 2: Open foundation and fine-tuned chat models." *arXiv preprint arXiv:2307.09288* (2023), **DOI**: 10.48550/arXiv.2307.09288.

[27] Q. J. Albert, S. Alexandre, M. Arthur, B. Chris, and S. C. Devendra. "Mistral 7b", *arXiv preprint arXiv:2310.06825*, 2023., **DOI**: 10.48550/arXiv.2310.06825.

[28] G. Penedo, Q. Malartic, D. Hesslow, R. Cojocaru, A. Cappelli, H. Alobeidli, B. Pannier, E. Almazrouei, and J. Launay. "The RefinedWeb dataset for Falcon LLM: outperforming curated corpora with web data, and web data only." *arXiv preprint arXiv:2306.01116*, 2023, **DOI**: 10.48550/arXiv.2306.01116.

[29] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, J. Vanderplas. "Scikit-learn: Machine learning in Python". The Journal of machine Learning research. 2011 Nov 1; 12:2825-30, **DOI**: 10.5555/1953048.2078195.

[30] S. Bird. "NLTK: the natural language toolkit". In *Proceedings of the COLING/ACL 2006 interactive presentation session*s 2006, pp. 69–72, **DOI**: 10.48550/arXiv.cs/0205028.

[31] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Corrado, S. Ghemawat. "Tensorflow: Large-scale machine learning on heterogeneous distributed systems". *arXiv preprint arXiv:1603.04467*. 2016, **DOI**: 10.48550/arXiv.1603.04467.

[32] M. Honnibal. "spaCy 2: Natural language understanding with Bloom embeddings", convolutional neural networks and incremental parsing. 2017.

[33] Textstat. https://textstat.org/.

[34] LanguageTool. https://pypi.org/project/language_tool_python/

[35] Pyspellchecker. https://pypi.org/project/pyspellchecker/

[36] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu. "Exploring the limits of transfer learning with a unified text-to-text transformer." *Journal of machine learning research 21*, no. 140, 1–67, 2020, **DOI**: 10.48550/arXiv.1910.10683.

Doaa Mostafa is a Ph.D. student in the computer science department at the faculty of computer and information science, Ain Shams University, Cairo, Egypt. She received her master's degree in computer science from the faculty of computer and information science, Ain Shams University in 2020. Her research interests include big data, artificial intelligence, natural language processing, and data mining.

Sally S. Ismail is a lecture of Computer Science at Ain Shams University, Cairo, Egypt. She holds a Ph.D. in Computer Science field from the University of Ain Shams. Her research interests include information retrieval, text summarization, and sentiment analysis.

Mostafa Aref is a professor of Computer Science at Ain Shams University, Cairo, Egypt. He holds a Ph.D. in engineering science in system Theory and engineering, June 1988, University of Toledo, Toledo, Ohio. He obtained his M.Sc. in computer science, 1983, University of Saskatchewan, Saskatoon, Sask. Canada. His research areas are natural language processing, knowledge representation, and ontology

Content Credentials: Trust Issues, Technical Solutions and Future Perspectives Using Encrypted Metadata in Image Processing

György Wersényi* and Victor Koech*

***Abstract*—Emerging technologies offer validation and authentication solutions in the field of audiovisual content creation. Visible or invisible watermarking, embedded metadata, and digital signatures can be used to maintain the validity and creditability of still images and video data. The Coalition for Content Provenance and Authenticity (C2PA) was established to create an open source framework and to provide technical solutions for image capture, processing, delivery, and verification. The leading market players in hardware and software development set the goal of applying encrypted metadata information to guarantee the authenticity of the data. Currently, only a few devices and applications are available and have been implemented based on this technology. This paper gives an introductory overview of the recent state, highlighting advantages, drawbacks, available implementations, and future perspectives on research directions.**

***Index Terms*—Content Credentials; Validation; Deepfake; Image Processing**

I. INTRODUCTION

In the age of rapid propagation of digital media, the truthfulness of the media has become increasingly important. Social networks have devolved journalistic responsibilities to the masses by making it possible for anyone with any motive to have unregulated access to audiences far and wide [1], [2]. This concern is also enhanced by the increased ease of access to photo-manipulation tools, some of which even include automation to lower the previously high technical bar [3]. Additionally, we have now entered the era of generative artificial intelligence (AI), which democratizes the ability to create synthetic content that has a convincing resemblance to real media. These developments have propagated and increased the complexity of disinformation and misinformation, which has caused severe challenges in determining the origin and authenticity of content [4], [5]. Malevolent people can take advantage of these tools to mislead, impersonate, and even conduct cyber-crimes which, in turn, erodes public confidence in systems such as governments, broadcast media and even personal communication [6]. This risk has resulted in an environment that requires verification of the origin and authenticity of any given digital content. A recent study on AI in the music industry revealed that 97% of people cannot tell the difference between fully AI-generated and human made music [7]. Furthermore, there was an overwhelming support for labeling of 100% AI-generated music, and more than half of the subjects felt uncomfortable by not being able to tell the difference.

* Széchenyi István University, Győr, Hungary (e-mail: wersenyi@sze.hu, koech.victor@sze.hu)

A. Trust and Validation

The rise of fake and AI-generated images poses major challenges to trust and validation in media, politics, and social networks [8]. Deepfakes and synthetic visuals can spread misinformation, manipulate public opinion, and undermine trust in authentic content. Such images are prevalent in journalism, advertising, and social media [9]. Technologies such as watermarking, blockchain-based provenance tracking, and encrypted metadata can help authenticate sources and restore trust in visual information [10], [11].

Key technological solutions to improve image authenticity and validation focus on provenance tracking, digital watermarking, cryptographic hashing, and AI-based forensic detection [12]–[16].

- Digital watermarking embeds invisible identifiers into images, allowing origin verification and tamper detection;
- Blockchain-based provenance systems record immutable metadata about image creation and modifications;
- Cryptographic signatures and hashing verify file integrity by comparing digital fingerprints; and
- AI forensics detect manipulations using pixel-level inconsistencies or generative adversarial network fingerprints.

B. Definition of Content Provenance

The foundation of the content provenance model is built on the understanding that there is a constant whack-a-mole going on between generative AI models and AI detection models, which is not sustainable. As generative AI models continue to improve, the detection of their outputs also becomes more complex, and the detection models have to play catch up [17].

DOI: 10.36244/ICJ.2026.1.7

The content provenance approach therefore tries to navigate around this problem by creating a parallel ecosystem of trust which shifts the burden of proof from the content consumer to the content producer. Content creators and publishers are the ones who are empowered to offer proof of their content history, rather than regular consumers who are skeptical and concerned about the trustworthiness of the content they encounter [5].

C. The Coalition for Content Provenance and Authenticity

According to the Coalition for Content Provenance and Authenticity explainer, C2PA is a conglomerate of major technology companies, media entities, and stakeholders spanning multiple industries with the sole objective of promoting an open and universal technical standard to build digital trust [5]. The coalition's main product is the C2PA technical specification, which outlines a framework for generating and embedding content credentials into digital media. The credentials serve as a secure electronic metadata package that is tied to the digital media. The CSI-Content Credentials report compares these credentials with nutrition labels that provide the consumer with details of the ingredients used and the origins of the products [18]. Inspection of these credentials is therefore meant to satisfy the curiosity of where the content originated, what tools were involved, and what manipulations were made over its entire history.

This paper provides a review of the C2PA framework and its applications. It presents the technical architecture, the evolution of its ecosystem, its security stance compared to the available technical specifications, the latest industry developments, and future perspectives in research. The evaluation will also highlight the limitations of the framework, the relationship with other alternative authentication technologies, and its implications on society and ethics. The paper is structured as follows. Section 2 introduces the C2PA framework, followed by basic technical specifications in Section 3. Section 4 deals with adoption and impact; Section 5 discusses limitations. Finally, privacy and ethical issues, as well as a comparison and outlook in development and research will be highlighted.

II. The Emergence and Evolution of C2PA

The C2PA coalition was officially established in the year 2021 as a Joint Development Foundation housed under the Linux Foundation. This happened as a result of the consolidation of two parallel initiatives, which had also been keen on solving the problem of digital provenance. This merging was a welcome change in the re-evaluation of a fragmented standards landscape, while it also helped bring together the expertise and resources of key industry players [19]. The parallel initiatives were The Content Authenticity Initiative (CAI) founded in 2019 by Adobe and Project Origin, also founded in 2019, by Microsoft and BBC. The former was mostly driven by creators with a desire to claim authorship of their work, while the latter was driven by the broadcast industry to counter disinformation and misinformation in the news ecosystem.

A. Governance of C2PA

The C2PA coalition is led by a committee of its bigger members, including Adobe and Microsoft who oversee the software, Arm and Intel who oversee the hardware and chip design, BBC who represent the media, and Truepic who specialize in the authentication technology. This group ensured that there would be good consideration for the entire life cycle of digital media from the point of capture to the editing and finally to its distribution [20]. On a more positive note, there has been a significant expansion of the steering committee, with more industries getting representation. Some of the newer members include early adopters like Sony and X, advertisers such as Publicis Groupe, and AI technology companies such as Google, Amazon, Meta, and OpenAI [21]. The involvement of big players from various industries highlights the coalition's deliberate strategy to expand the scope of the proposed solution. The problem of misinformation and disinformation is large and a solution by one entity would not be able to cover the vast landscape of the Internet [19]. This grouping therefore goes some way to bring about a network effect which can help bring about ubiquitous adoption which is open and interoperable.

B. The C2PA Charter

All work carried out by the C2PA working groups is governed by its official charter, which is a set of core principles framed as the constitution of the specification. The principles including interoperability, privacy, simplicity, global applicability, performance, prevention of abuse, and unbiased viewpoint are meant to ensure that the development is technically sound and ethical [5]. The principles are also designed to help with the decision-making of the working groups, so that all the builds undergo a set of checks through security reviews and harms modeling exercises before they are publicly implemented [22].

III. Technical Specifications and Architecture

A. Overview and Core Aspects of C2PA

To ensure that a digital asset's history is secure, tamper-evident, and verifiable, the C2PA framework layers on top of an architecture of data structures and cryptographic processes. According to the latest C2PA specification, version 2.2, the data model is based on hierarchy, by which the provenance information is segmented into specific components as follows [23]–[25].

1) Assertions: These are fundamentally the identifying units within the C2PA framework as they define a fact about a digital asset. They are all labeled using namespaced strings, such as c2pa.actions, which describe the actions that have been performed on the asset, or c2pa.ingredient that is linked to other assets that were used in the formation of the digital chain. These assertions are defined by C2PA or other entities.
2) Claim: This data structure ties all the assertions together to a single event in time. The claim is digitally signed

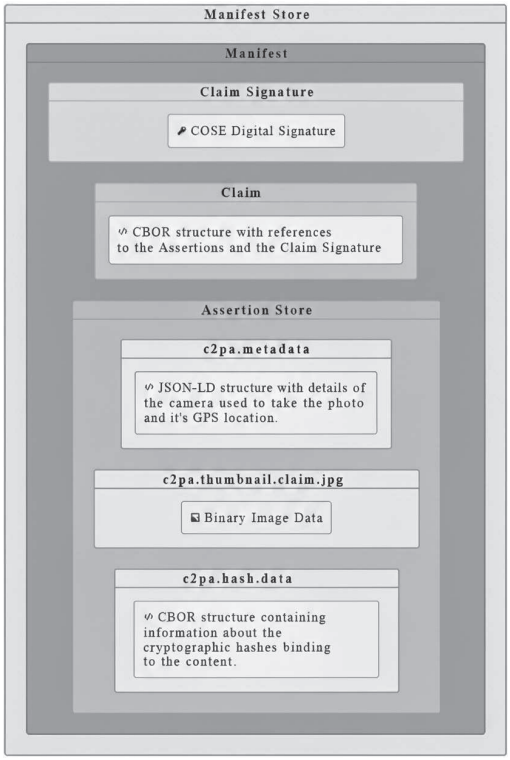


Fig. 1. C2PA Manifest of an image [24].

by the producer/modifier of the asset, cryptographically hard binding all the assertions to indicate their validity.

3) Manifest: This is the container of the verifiable asset provenance. It contains a single claim, the signature of the claim, and a set of assertions tied to the claim. This data package closely resembles the nutrition label of the asset. There can be multiple manifests in each asset that represent each tool that contributed to its lifecycle. Figure 1 shows a manifest of an image.
4) Manifest Store: This overarching container holds all the manifests associated with an asset. The store is a representation of the complete and cumulative provenance history of an asset and can be embedded in the asset file or hosted externally and linked to the asset.

B. Authenticity and Integrity

Establishing the trustworthiness of C2PA content credentials is based on cryptographic mechanisms that ensure that data can be verified to be authentic and have not been tampered with. These mechanisms are as follows.

1) Hashing: This is a hard-binding technique that creates a tamper-evident trace between a digital asset and its manifest. It involves generating a cryptographic hash that is tied to the bytes of the digital asset and is stored in an assertion contained in the manifest [25]. The hash is generated using any of the standard algorithms such as SHA-384 or SHA-256 and this generally results in a form of digital fingerprint of the asset which definitively determines the tamper status of the asset during validation [26].
2) Signing: The manifest containing the hash also contains a claim that is digitally signed by an actor that can either be a software application or a hardware device that creates it. This signing uses a public key infrastructure model in which the signature to the claim uses a private key that can be universally verified using the corresponding public key. Public keys are hosted by certificate authorities on their digital certificates (typically X.509 certificates) and show the creator and a tamper status since their creation [27].
3) Timestamping: For good record-keeping and chaining, details of when a manifest was signed are captured in the form of a time stamp. The signing process utilizes the services of a time stamp authority to tie the hash of the claim with a dated token. The token is bundled with the manifest and provides proof that the signature was created on the said date and time [26]. This can then be used to later verify the validity of the certificate according to the expiration or revocation status.

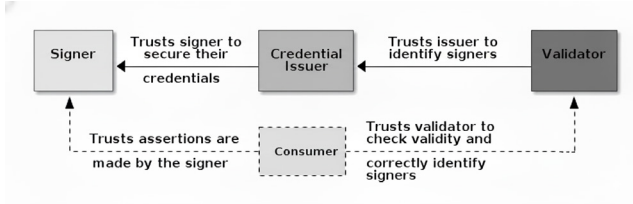


Fig. 2. The C2PA Trust Model [24].

C. The C2PA Trust Model

The framework aims to provide verifiable information about an asset's provenance so that the basis of trust is anchored in the identities of the signers. Devices and software that perform content credential validation do this by checking the integrity of the provenance data attached to the digital asset, so this is not just a check for 'realness' or 'fakeness'. There is a Trust List made up of certificate authorities whose certificates are known and acknowledged by the validating application or device. Therefore, the validation process tries to establish whether the certificate used originated from a certificate authority that is on the trust list. Depending on the results of the integrity checks, the manifest can be assigned a state such as valid, trusted, well-formed, or unknown [28]. To achieve the highest level of technical trust, an asset needs to have a well-designed manifest that meets the C2PA specification. The digital signature of the manifest should also be successfully validated as tamper-free, therefore valid, and finally the signature certificate for the manifest chain should be present on the trust list of the validator. Figure 2 shows the so-called trust model within C2PA.

Another aspect of the trust model is the binding of a manifest store with its associated asset either by embedding

the store in the asset file structure or storing the manifests externally in the cloud or in a distributed ledger [26]. Typically, in the embedding approach, JPEG Universal Metadata Box Format containers were used, ensuring that the provenance data moves with the asset as a self-contained file. In some other cases, such as when uploading files to social media apps whereby metadata can be stripped, C2PA advocates for using the external repository method where a soft binding is applied to relink assets to their stripped metadata. The soft binding identifier can be a digital watermark or a perceptual hash that visually identifies the asset or similar content [25].

C2PA uses cryptographic hashing and digital signatures to bind provenance to content in a tamper-evident way. Hashes (i.e., SHA-256) compute content digests over defined byte ranges; signatures over CBOR-serialized claims use X.509-based credentials via COSE (e.g., ECDSA, EdDSA, RSASSA-PSS). In the trust model, decisions rest on the identity of the signer (credential chain anchored in trusted roots) and validity of their certificate. Validators maintain trust lists after approval, check revocation, and assess whether a signing credential was valid at the time of signing. Identity assertions may include devices, applications, or pseudonyms. Information security and threat modeling outline ongoing threat modeling, encourage design against abuse, and emphasize key management, minimizing key reuse, revocation strategy, and securing claim generation systems. It also highlights the evaluation of harms, misuse and abuse as a continuous process that respects privacy, human rights, and evolving threats.

D. Privacy and Ethics

The foundation of C2PA is to ensure that transparency is brought to the forefront; however, this conflicts with the fundamental right to privacy. If the history of a digital asset is to be detailed and unalterable for accountability purposes, everything including who created it, where, when, and how should be included, but this will also bring about concerns of surveillance, misuse, and speech suppression [28]. These are concerns that the C2PA foundation is aware of and has taken some steps toward addressing. In the Harms Modeling documentation, C2PA outlines the possible harms that could occur, and the mechanisms by which creators can take control.

One of the key tools by which the C2PA tries to address concerns is by preaching that the whole project is voluntary and that no content should have credentials attached by default. This means that one must consent to the recording of provenance data, even if they are using a C2PA enabled tool [5]. To complement this, creators are also provided with a redaction feature which allows them or subsequent editors to subtract some assertions such as those containing sensitive information without invalidating the attached manifest. This subtraction leaves a record to ensure that transparency is preserved, but sensitive information is permanently removed [29]. Lastly, it is not a requirement for the manifest signature to identify the signer because pseudonyms or anonymous certificates can be used to protect vulnerable people like whistle-blowers and activists [30].

The Harms Modeling framework has also formalized a governance process in which potential social issues and their impacts are accounted for. The idea is to anticipate how technology might be used to negatively affect stakeholders and the world at large, by classifying intended users and their use cases, weighing the potential harms and creating countermeasures such as changes to specifications, design recommendations, and public interest campaigns [31]. These safeguards do not necessarily address all issues because as the adoption of C2PA increases, there is an ethical challenge that arises, whereby content lacking the defined identifier could be deemed suspicious in the future and this could penalize creators who have legitimate reasons not to adopt it [28].

E. C2PA versus Competing Technologies

There are several technologies that have been specifically built to address the issue of digital trust, but each of them has their shortcomings. When comparing C2PA with AI-based deepfake detection solutions, the most notable distinction is in the approach, which is proactive in the former and reactive in the latter. C2PA's approach is to establish the content's authenticity from the point of creation, which leaps ahead of the inherently delayed approach of trying to verify manipulation through AI [27], [32]. Although fake AI detection techniques might seem limited, especially considering their propensity to perform well with known manipulations but fail on new ones, it still serves as the best approach for the analysis of existing content created before C2PA [33].

When comparing C2PA with digital watermarking technologies, there is more of a synergetic relationship, whereby watermarks complement metadata packages. C2PA manifests can be easily removed from the digital asset in transmission while the digital watermarks remain attached to the pixels of the asset and therefore can survive a variety of transformations, including screen captures [34]. This has been the foundation of the Durable Content Credentials model, which combines the full external CPA manifest with a robust digital watermark bundled within the content itself. The resulting digital asset can be identified by a C2PA-compliant application even if the metadata had been removed, which means that the asset can have its provenance information restored [35].

C2PA also has a complementary relationship with perceptual hashing through the use of cryptographic functions to establish bindings. Hashes such as SHA-256 are known to be very profound, such that any change to the file, no matter how small, will result in a change in the hash value, which makes them ideal for bit-for-bit integrity but powerless for similarity comparisons. With perceptual hashes, the hash produced can be similar for content that is visually comparable despite changes to file size or minor alterations. They are therefore suitable for identifying copies or close duplicates of a digital asset shared on the Web [36]. C2PA has used perceptual hashing technology to establish soft bindings so that when manifests are detached, the hash can be used to find original or similar assets and reconnect them with provenance information [37].

Blockchain systems are also closely related to C2PA in that they are used to create immutable records, albeit with a different underlying architecture. For C2PA, cryptographic metadata can be embedded directly into the file for offline use, eliminating the need for a connection to a distributed network [38]. This means that the specification can utilize blockchain, but it is not necessary [26]. The advantage of using the decentralized distributed ledger of blockchain technology is that reliance on a single authority is removed, leading to extreme resistance to manipulation. Projects like the Numbers Protocol have combined these two technologies such that the C2PA provenance data are signed and the hash of the manifest is registered on a blockchain [39].

IV. DISCUSSION

Ensuring authenticity and trust in images and videos is critical to preserving the integrity of digital information. Manipulated or synthetic media can distort truth, fuel misinformation, and erode public confidence. Reliable provenance, cryptographic validation, and transparent metadata are essential to maintain accountability and verifiable trust in the dissemination of visual content. However, creating multi-metric evaluations combining robustness (survivability), security (tamper resistance, key compromise scenarios), utility (verification latency), and human-centric metrics (trust, comprehension) is a difficult task.

Providing open datasets and tools (scripts that apply common transforms, edits, platform-like re-encodings) are needed for others to reproduce results. Adversarial testbeds can simulate common distribution chains from the camera through the editor and the distribution network to the final result (screenshot/re-upload). Threat models must be defined explicitly to show what attackers can and cannot do (e.g., full access to platform storage vs. passive network attacker).

A C2PA manifest acts as a verifiable container of provenance data, enabling users and systems to confirm the authenticity and history of the modification of digital content. Furthermore, a C2PA claim is a trusted signed declaration embedded in the manifest that documents verifiable facts about the lifecycle or attributes of digital content. Schemas make provenance data understandable and consistent, while signatures make them trustworthy and tamper-proof. A schema is a descriptive data model that defines what data looks like (structure and validation); a signature is a cryptographic mechanism that verifies who created or altered data (authenticity and integrity). Together, they target a transparent, interoperable, and secure ecosystem.

A. Hardware Support

Currently, only a few vendors offer models with C2PA authentication (see Figure 3). These are available either out-of-the-box in hardware or via firmware/software updates. Some are partial or require licensing, and some vendors have announced intentions or are working on updates, but those features may not yet be available broadly.

Vendor	Camera Models / Details	Notes / Status
Leica	Leica M11-P, M11-D, SL3-S	Built-in C2PA provenance signatures.
Sony	A1 II, A1, A9 III, A7S III, A7 IV	Firmware-based C2PA; may need licensing.
Nikon	Z6 III	C2PA support via firmware rollout.
Fujifilm	X-T50, GFX100S II	Supports C2PA authenticity metadata.
Canon	EOS R1, EOS R5 Mark II	Native or firmware C2PA support.

Fig. 3. Vendors and models.

The movement officially started with Leica, who pioneered the M11-P camera in 2023 that featured built-in support for content credentials, and this was soon followed by another model, the SL3-S, in 2025 with the same capability [20]. This paved the way for other camera manufacturers such as Nikon, Canon, Panasonic, and Sony to join the initiative [27], [40]. Another significant milestone for hardware adoption was reached when Samsung, the largest smartphone manufacturer, announced that the Galaxy S25 lineup of phones would also adopt native C2PA support. This was shortly followed up by the Google Pixel lineup, also releasing with native support.

It is also possible that authenticity services may be available only in certain countries or regions, or only for certain user classes (i.e., press agencies) initially. Furthermore, even if the camera embeds the metadata, a software is needed in the workflow (editing, export, sharing) that preserves this metadata so the provenance chain remains intact. If metadata is stripped (by certain social media platforms or export tools), the C2PA signature might be lost.

The adoption of technology is evolving; many cameras require a specific firmware version, an optional licensed 'authenticity' service, or registration with the vendor cloud to load signing certificates. Furthermore, chip-level/phone implementations (i.e. Google Pixel 10 and Qualcomm platform integrations) are appearing that use content credentials at capture time when the phone OEM and camera stack implements it.

B. Software Support

A concise and up-to-date rundown of the leading software solutions and toolkits that implement content credentials for images can be assembled, whether they are primarily online (cloud/web), offline (desktop/mobile apps/local libraries), or hybrid services.

Primarily offline desktop apps (with optional online features) include Adobe Photoshop and Lightroom. Primarily online (cloud and APIs) solutions include Truepic (verification and authenticity of APIs), platform/Content delivery network integrations (Cloudflare), and platform verification features (YouTube, Google Search) [41]. Some vendors offer device-based hybrid solutions, i.e. Google Pixel or Sony and Leica, which embed manifests in-camera.

Developer toolkits and off-line libraries that are embedded into applications are available as open-source SDKs and reference tools (c2pa-js, Python examples, etc.). The CAI/C2PA community provides open-source SDKs, the Verify inspector tool, c2pa-js, Python/other examples, and reference implementations (GitHub) so developers can create or validate content

credentials in apps and pipelines. These are used to build server-side and client-side verification. Truepic is the best-known vendor, but other emerging platform vendors also plan to run private implementations. On the distribution side of the digital media, C2PA has also announced that some of the largest players have also begun integrating the framework into their platforms. This has been led by Meta (Instagram and Facebook), LinkedIn, and Tiktok [41].

The following services are currently available:

1) Adobe — Content Credentials (Photoshop, Lightroom, Web tools). Offline desktop apps can embed content credentials at export; online web tools for inspection and management (Adobe Content Authenticity web app, Inspect).
2) Truepic verify and authenticity platform. Its cloud and mobile SDKs provide capture and verification services, device attestation, and an authenticity platform for publishers, platforms, and enterprises.
3) Google Pixel and Photos for platform integrations is a hybrid device-level capture tool (Pixel/Android) that can embed content credentials at capture time (on-device) and online platform/display integrations (Google Photos, Search) to surface provenance.
4) Microsoft offers Project Origin and Azure integrations. It is an online platform and research initiative; developer documents show content credentials support for generated media.
5) Platform and Content delivery network adopters, such as Cloudflare, YouTube, and others are beginning to preserve and expose content credentials (example: Cloudflare added a “Preserve Content Credentials” option for hosted images; YouTube has experimented with C2PA-based labels).

C. Vulnerabilities and Adversarial Robustness of C2PA

The effectiveness of C2PA in achieving its objectives is challenged by several factors briefly mentioned. The main vulnerability is the fragility of embedded metadata, which can be easily removed or altered when uploaded to content delivery networks for optimization or privacy reasons [28]. A much simpler form of this vulnerability is the rampant laundering of content through screenshots and screen recording to create a new file for re-uploading [42]. Outside of these fundamental limitations, security researchers have also pointed out some sophisticated exploits such as provenance piggybacking, where an attacker takes an authentic credentialed asset and layers a manipulated or deep-faked element on top of it [43], and unprotected metadata manipulation, where the unprotected elements like EXIF data can be changed without invalidating the C2PA signature [44]. The RAND corporation has also pointed out that the framework’s threat model needs updating because it has been the same since version 1.0 from January 2022 with significant changes in the landscape driven by generative AI [45], [46]. This criticism highlights a significant privacy dilemma: the inherent push for transparency in C2PA conflicts with the need for safety/privacy for vulnerable creators. Another documented shortcoming of C2PA is that adding manifests to digital assets tends to increase file sizes, so assets that undergo multiple edits can become prohibitively large. This can result in bandwidth and storage problems for low-resource situations [47].

D. Outlook of C2PA

As the C2PA framework evolves and graduates from specification to general adoption, it will be critical that it is adopted by the International Organization for Standardization (ISO). Currently, there is some work to get this done in a fast track process published as ’ISO / DIS 22144, Authenticity of Information - Content Credentials’ [35]. If this happens, then it will carry more weight and increase adoption by governments and more corporations. So far, there has been some normal challenge to the technical details of the framework, but they are minor and do not dispute the core architecture [48].

The technical roadmap of the specification shows that it is focused on addressing several key challenges, such as solving metadata removal through vendor-agnostic watermarking, strengthening the threat model to counter AI attacks, addressing the issue of file size through compression or better manifest management, and achieving end-to-end provenance preservation [26], [46], [49]. There has also been a shift in the digital media landscape, with governments worldwide considering or enacting legislation that mandates transparency. In the United States, for example, Utah and California states have passed laws requiring verification of online content, while the United Nations has also made resolutions encouraging interoperability of authentication mechanisms. [50].

E. Research Directions

Due to the novelty of this topic, there is limited scientific research data available. Most of what is available is tech reports, preprints, repository entries, or non-reviewed publications. The most recent directions include cryptographic provenance, technical standard development, authenticity issues, and ethical frameworks [28], [51]–[53].

The most important field is trust enabling and calibration. Early user studies exist, but are limited. The goal is to experimentally measure how provenance displays affect user judgments for images and short videos across demographics and platforms. Controlled user experiments varying label wording, iconography, provenance depth should be designed in which accuracy in detecting manipulated media and changes in credibility/trust ratings can be tested. Different provenance user interfaces and labels can affect user trust, perception, and behavior (what users understand from “captured with camera and C2PA” vs. “AI-generated” labels). This issue is strongly related to trade-offs between provenance transparency, creator privacy, and identity binding. The support of selective disclosure, pseudonymous attestations, and legal/regulatory constraints is of paramount importance.

Emerging critiques warn against over-promising provenance [54]. Some of the vulnerabilities that have been highlighted, such as practical exploits and a weak threat model, show the

need to restrain enthusiasm until more research is conducted and solutions implemented. Personal data and details about the origin of the image (reliable data about the time and location of the capture, etc.), possible tracing of individuals based on metadata highlight privacy and data security issues. Designing cryptographic protocols that enable content to carry provenance claims without exposing sensitive identity fields, that is, zero-knowledge proofs to assert 'created by a verified source' while hiding identity, is essential to build trust.

Scalability and verification at Internet scale deals with the efficient verification across network edge nodes, browsers, social platforms, and truncation/screenshot scenarios (how to preserve provenance when content is transformed for distribution). For screenshots and re-upload, no concrete proposals have been made yet [55]. Studies have been carried out to determine how effective metadata manifests resist deletion and tampering vs. robust watermarks and ML-based verification [37]. Hybrid schemes that combine C2PA manifests with robust invisible watermarks are a major focus. There is a need to design an architecture that uses C2PA metadata and a perceptually robust watermarking scheme to survive common transforms (recompression, cropping, screenshotting) while preserving unforgeability will lead the way to increase trust.

Practical systems and prototype deployments are emerging regarding the embedding of continuous provenance for live streams, low-latency signing, and secure key handling [56], [57]. Building and measuring C2PA-compatible live streaming prototypes is the next step in development. Another focus could be on interoperability, which refers to the standards, protocols, technologies, and mechanisms that allow data to flow between diverse systems with minimal human intervention [58].

V. CONCLUSIONS

In this paper, we provided a brief overview of the current state of content provenance and authenticity regarding visual media, focusing on still images. The C2PA and the CAI association founded by leading market players paved the way for developments and technical solutions by formulation of an open, royalty-free technical standard that serves as a basis for the C2PA member's efforts against disinformation. Only a few experimental results have been available to date. The research directions and open questions were highlighted. Concerns about data privacy can undermine trust. Key and identity management at scale (who issues keys, revocation) is critical for platform trust. Provenance does not automatically stop misinformation; it can help, but a sociotechnical evaluation is necessary to avoid overpromising. The fine balance between the need for transparency and the right to privacy has introduced an ethical dilemma for the framework to figure out a solution.

The C2PA framework has great ambition and is serving a critical role in restoring trustworthiness to the digital ecosystem. Its success faces some hurdles that will depend on the coalition's resolve to improve the standard to meet the arising needs, the industry's appetite for privacy-preserving implementation, and the general acceptance of provenance by the public with its critical nuances. Future directions in research include the evaluation of human factors, but also technical issues such as cryptography, detection of fake images and malevolent intentions, real-time adaptation of the technology into live streams and the implementation on hardware (cameras) and software application level (social media platforms, search engines, generative and detective AI systems).

REFERENCES

[1] J. Golbeck, "Combining provenance with trust in social networks for semantic web content filtering," in *International Provenance and Annotation Workshop*. Springer, 2006, pp. 101–108.

[2] A. Bernstein and A. Gomila, "The truth in social media," *Topoi*, vol. 44, pp. 127–138, 3 2025. DOI: 10.1007/s11245-024-10039-6

[3] J. Jang, "Self presentation using photo-editing apps on social media," Master's Thesis, Seoul National University, Seoul, South Korea, February 2023. [Online]. Available: https://s-space.snu.ac.kr/handle/10371/194001

[4] E. Ferrara, "Genai against humanity: Nefarious applications of generative artificial intelligence and large language models," *Journal of Computational Social Science*, vol. 7, no. 1, pp. 549–569, 2024.

[5] Coalition for Content Provenance and Authenticity, "C2PA Explainer, v2.2," https://spec.c2pa.org/specifications/specifications/2.2/explainer/Explainer.html, 2025.

[6] S. Kreps and D. L. Kriner, "The potential impact of emerging technologies on democratic representation: Evidence from a field experiment," *New Media & Society*, vol. 26, pp. 6918–6937, 12 2024. DOI: 10.1177/14614448231160526

[7] Deezer, "Deezer/Ipsos Survey: 97% of people can't tell the difference between fully AI-generated and human made music," {https://newsroom-deezer.com/2025/11/deezer-ipsos-survey-ai-music/}, 2025.

[8] R. Chesney and D. Citron, "Deepfakes and the new disinformation war: The coming age of post-truth geopolitics," *Foreign Affairs*, vol. 98, no. 1, pp. 147–155, 2019.

[9] L. Verdoliva, "Media forensics and deepfakes: an overview," *IEEE journal of selected topics in signal processing*, vol. 14, no. 5, pp. 910–932, 2020.

[10] A. F. Qasim, F. Meziane, and R. Aspin, "Digital watermarking: Applicability for developing trust in medical imaging workflows state of the art review," *Computer Science Review*, vol. 27, pp. 45–60, 2018.

[11] A. Alexander, "Truth in the post-deepfake era: Can blockchain and watermarking restore digital trust?" Available at SSRN 5377851, 2025.

[12] M. Barni and F. Bartolini, Watermarking systems engineering: enabling digital assets security and other applications. Crc Press, 2004.

[13] B. Singh and G. Kasana, "A review of digital watermarking techniques: Current trends, challenges and opportunities," in *Web Intelligence*, vol. 22, no. 4. SAGE Publications Sage UK: London, England, 2024, pp. 523–553.

[14] X. Li, L. Wei, L. Wang, Y. Ma, C. Zhang, and M. Sohail, "A blockchain-based privacy-preserving authentication system for ensuring multimedia content integrity," *International journal of intelligent systems*, vol. 37, no. 5, pp. 3050–3071, 2022.

[15] Q.-u.-A. Mastoi, M. F. Memon, S. Jan, A. Jamil, M. Faique, Z. Ali, A. Lakhan, and T. A. Syed, "Enhancing deepfake content detection through blockchain technology." *International Journal of Advanced Computer Science & Applications*, vol. 16, no. 6, 2025.

[16] F. Marra, D. Gragnaniello, L. Verdoliva, and G. Poggi, "Do gans leave artificial fingerprints?" in *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*. IEEE, 2019. DOI: 10.1109/MIPR.2019.00103 pp. 506–511. [Online]. Available: https://ieeexplore.ieee.org/document/8698678

[17] K. Cheng, W. Li, N. Zhang, X. Liu, and H. Wu, "Principles and challenges of generative artificial intelligence detection," *British Journal of Anaesthesia*, vol. 133, no. 4, pp. 899–901, 2024.

[18] Canadian Centre for Cyber Security (CCCS) and NSA and UK NCSC and Australia ACSC, "Content Credentials: Strengthening Multimedia Integrity in the Generative AI Era," *Canadian Centre for Cyber Security, Tech. Rep.*, 2025, version 1.0. [Online]. Available: https://media.defense.gov/2025/Jan/29/2003634788/-1/-1/0/CSI-CONTENT-CREDENTIALS.PDF

[19] C2PA Founding, "C2PA Founding Press Release," https://c2pa.org/c2pa-founding-press-release/, 2 2021.

[20] Content Authenticity Initiative, "How it works," https://contentauthenticity.org/how-it-works, 2025.

[21] "C2PA - Announcements," https://c2pa.org/news/, 2021.

[22] Coalition for Content Provenance and Authenticity, "Guiding Principles," https://c2pa.org/principles/, 2021.

[23] CHESA, "Understanding c2pa: Enhancing digital content provenance and authenticity," https://chesa.com/, 2024.

[24] Coalition for Content Provenance and Authenticity (C2PA), "Content Credentials: C2PA Technical Specification, Version 2.1," C2PA, Tech. Rep. 2.1, 2025, accessed: 2025-10-10. [Online]. Available: https://spec.c2pa.org/specifications/specifications/2.1/specs/C2PA_Specification.html

[25] Coalition for Content Provenance and Authenticity, "Content Credentials: C2PA Technical Specification, v2.2," https://spec.c2pa.org/specifications/specifications/2.2/specs/C2PA_Specification.html, 2025.

[26] K. Rathi, S. S. Kumar, and A. N. Mandanna, "Insights into Coalition for Content Provenance and Authenticity (C2PA)," https://www.infosys.com/iki/techcompass/content-provenance-authenticity.html, 2 2024.

[27] M. Demey, "C2PA's Fight Against AI-Generated Deception," https://apryse.com/blog/ai-content-cp2a-authenticity, 11 2023.

[28] World Privacy Forum, "Privacy, Identity and Trust in C2PA: A Technical Review and Analysis of the C2PA Digital Media Provenance Framework," https://worldprivacyforum.org/posts/privacy-identity-and-trust-in-c2pa/, 2025.

[29] Coalition for Content Provenance and Authenticity, "C2PA Implementation Guidance, v2.2," https://spec.c2pa.org/specifications/specifications/ 2.2/guidance/Guidance.html, 2025.

[30] "C2PA Security Considerations, v1.0," https://spec.c2pa.org/specifications/specifications/1.0/security/SecurityConsiderations.html, 2022.

[31] "C2PA Harms Modelling," https://spec.c2pa.org/specifications/specifications/1.0/security/attachments/Harms_Modelling.pdf, 2021.

[32] M. Zorz, "The limits of AI-based deepfake detection – Help Net Security," https://www.helpnetsecurity.com/2024/11/22/ben-colman-reality-defender-deepfakes-detection/, 11 2024.

[33] Y. Xu, P. Terhörst, M. Pedersen, and K. Raja, "Analyzing fairness in deepfake detection with massively annotated databases," *IEEE Trans-actions on Technology and Society*, vol. 5, pp. 93–106, 2 2024. DOI: 10.1109/tts.2024.3365421

[34] A. Parsons, "Durable content credentials," https://contentauthenticity.org/blog/durable-content-credentials, 2024.

[35] U.S. Department of Defense, "Content Credentials: Strengthening Multimedia Integrity in the Generative AI Era," https://media.defense.gov/2025/Jan/29/2003634788/-1/-1/0/CSI-CONTENT-CREDENTIALS.PDF, 2025.

[36] L. Struppek, D. Hintersdorf, D. Neider, and K. Kersting, "Learning to break deep perceptual hashing: The use case neuralhash," in *ACM International Conference Proceeding Series. Association for Computing Machinery*, 6 2022., pp. 58–69. DOI: 10.1145/3531146.3533073

[37] J. Fairoze, G. Ortiz-Jimenez, M. Vecerik, S. Jha, and S. Gowal, "On the difficulty of constructing a robust and publicly-detectable watermark," *arXiv preprint arXiv:2502.04901*, 4 2025.

[38] Content Authenticity Initiative, "Open-source Tools for Content Authenticity and Provenance," https://opensource.contentauthenticity.org/docs/faqs/, 2025.

[39] Documentation for Numbers Protocol Team, "What is C2PA and why do we need it?" https://docs.numbersprotocol.io/introduction, 2025.

[40] J. Tse, "5,000 members: building momentum for a more trustworthy digital world," https://contentauthenticity.org/blog/5000-members-building-momentum-for-a-more-trustworthy_digital-world, 8 2025.

[41] Coalition for Content Provenance and Authenticity, "C2PA News,"https: //c2pa.org/news/, 2024.

[42] Online Brand Ambassadors, "C2pa: Certifying digital media's authenticity," https://www.onlinebrandambassadors.com/c2pa-certifying-digital-medias-authenticity/, 2024.

[43] SCW, "How c2pa can safeguard the truth from digital manipulation," https://www.scworld.com/perspective/how-c2pa-can-safeguard-the-truth-from-digital-manipulation, 2025.

[44] The Hacker Factor Blog, "C2pa from the attacker's perspective," https://www.hackerfactor.com/blog/index.php?/archives/1031-C2PA-from-the-Attackers-Perspective.html, 2024.

[45] RAND Corporation, "Overpromising on Digital Provenance and Security," https://www.rand.org/pubs/commentary/2025/06/overpromising-on-digital-provenance-and-security.html, 2025.

[46] A. Locker, C. Heitzenrater, and T. Helmus, "Overpromising on digital provenance and security," https://www.rand.org/pubs/commentary/2025/06/overpromising-on-digital-provenance-and-security.html, 6 2025.

[47] Infosys, "Insights into coalition for content provenance and authenticity (c2pa)," https://www.infosys.com/iki/techcompass/content-provenance-authenticity.html, 2024.

[48] "ISO/DIS 22144, Authenticity of Information - Content Credentials," 2025.

[49] G. Huszar, "Why Broadcasters Must Embrace C2PA for Content Trust," https://onediversified.com/insights/blog/c2pa, 2025.

[50] Coalition for Content Provenance and Authenticity, "C2PA NIST Response," https://downloads.regulations.gov/NIST-2024-0001-0030/attachment_1.pdf, 2024.

[51] P. Laskar, "Cryptographic Provenance and the Future of Media Authenticity: Technical Standards and Ethical Frameworks for Generative Content," *Journal of Computer Science and Technology Studies*, vol. 7, no. 6, pp. 967–972, 2025.

[52] E. Bureacă and I. Aciobănit, ei, "A Blockchain Blockchain-based Frame-work for Content Provenance and Authenticity," in *2024 16th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*. IEEE, 2024, pp. 1–5.

[53] C. Trattner, S. L. Forstner, A. Starke, and E. Knudsen, "C2PA Provenance Labels Increase Trust in News Platforms Across Western Countries," https://www.researchgate.net/publication/391419247 C2PA Provenance Labels Increase Trust in News Platforms Across Western Countries, 5 2025.

[54] A. R. Locker, C. Heitzenrater, and T. C. Helmus, "Overpromising on Digital Provenance and Security," https://www.rand.org/pubs/commentary/2025/06/overpromising-on-digital-provenance-and-security.html, RAND School of Public Policy, Tech. Rep. Note, 2025, accessed: 2025-10-10.

[55] J. Seynhaeve, "An introduction to C2PA — Certifying digital media's authenticity," https://www.onlinebrandambassadors.com/c2pa-certifying-digital-medias-authenticity/, 11 2024.

[56] M. Mesa-Simón, A. Escobar-Molero, B. Sáez-Mingorance, D. P. Morales, J. A. Álvarez-Bermejo, and F. J. Romero, "Enabling live video provenance and authenticity: A c2pa-based system with tpm-based security for livestreaming platforms," *Authorea Preprints*, 2025.

[57] S. Petrangeli, H. Wang, M. Fisher, D. Kozma, M. Mahamli, P. Blumenthal, and A. Parsons, "Integrating content authenticity with dash video streaming," in *Proceedings of the 15th ACM Multimedia Systems Conference*, 2024, pp. 492–498.

[58] J. C. Simmons and J. M. Winograd, "Interoperable provenance authentication of broadcast media using open standards-based metadata, watermarking and cryptography," *arXiv preprint arXiv:2405.12336*, 2024.

György Wersényi was born in 1975 in Győr, Hungary. He received his MSc degree in electrical engineering from the Technical University of Budapest in 1998 and PhD degree from the Brandenburg Technical University in Cottbus, Germany. Since 2002 he has been member of the Department of Telecommunications at the Széchenyi István University in Győr. From 2020 to 2022 he was the dean of Faculty of Mechanical Engineering, Informatics and Electrical Engineering, as well as the scientific president of the Digital Development Center at the university. Currently, he is a full professor, member of the European Acoustics Association (EAA) and the Audio Engineering Society (AES). His research focus is on acoustic measurements, virtual and augmented reality solutions, sonification, cognitive infocommunications, and assistive technologies.

Victor Koech was born in 1991 in Kenya. He received his MSc degree in 2020 in Computer Information Systems at the School of Science and Technology of Kenya Methodist University. He has been a postgraduate PhD candidate since 2025 at Széchenyi István University, Hungary.

Rate-Splitting Multiple Access for Satellite Short-Packet Communications: Finite Blocklength Modeling and Reliability Analysis

Sang-Quang Nguyen[(1)(*)], and Chi-Bao Le[(2)]

Abstract—**Short-packet transmission is becoming crucial for satellite services that cannot rely on long codewords to hit the required latency and reliability. This study investigates rate-splitting multiple access (RSMA) in that context and builds a finite-blocklength (FBL) model for a downlink satellite-terrestrial link affected by Shadowed-Rician fading. We obtain closed-form approximations for the block error rate (BLER) of both the common and private streams, explicitly incorporating imperfect successive interference cancellation (ipSIC) at the receivers. Compared with power-domain non-orthogonal multiple access (NOMA), RSMA exhibits more stable BLER the common stream helps dampen residual interference due to ipSIC and the short-packet effect–so RSMA generally needs less transmit power to attain the same error targets. Numerical results validate the analysis and demonstrate consistent RSMA advantages across a wide range of transmit powers, blocklengths, shadowing severities, and antenna configurations. The results suggest that RSMA is a very promising option for future satellite systems that need to provide reliable, low-latency, and short-packet communications in view of realistic SIC imperfections.**

Index Terms—**Rate-splitting multiple access, finite blocklength, short-packet communications, satellite-terrestrial systems, block error rate, imperfect successive interference cancellation, reliability analysis.**

I. INTRODUCTION

Satellite–terrestrial communication systems have attracted significant interest as future 6G networks begin to demand wider coverage, massive connectivity, and reliable links for devices scattered across remote or hard-to-reach regions [1]–[3]. Unlike traditional broadband satellite services that focus on large codewords and high throughput, many emerging applications are now dominated by short-packet transmission, where only a few hundred symbols are available to encode critical information. Examples include environmental monitoring sensors, safety-related signaling, UAV telemetry packets, and low-latency control messages for distributed IoT systems [4], [5]. These scenarios impose tight reliability and delay requirements, making it difficult to rely solely on classical Shannon-capacity results, which assume infinitely long blocklengths and often mask the performance degradation triggered by short packets [6]–[8]. As a result, finite-blocklength (FBL) analysis has become increasingly important for understanding the practical behavior of satellite links operating under strict latency constraints [9].

Another challenge that satellite systems must contend with is interference management. With the growth of spectrum sharing between terrestrial and non-terrestrial networks, serving multiple ground users simultaneously requires access schemes that can sustain reliability in the presence of strong co-channel interference. Non-orthogonal multiple access (NOMA) has been widely studied for this purpose, but its dependence on accurate successive interference cancellation (SIC) often leads to performance loss when SIC becomes imperfect—something that occurs frequently in Shadowed-Rician fading, user mobility, or channel estimation uncertainty [10]–[13]. Even small SIC errors may propagate and severely deteriorate reliability, particularly when packets are short and decoding margins are tight [14], [15]. In contrast, rate-splitting multiple access (RSMA) has emerged as a more flexible strategy that can blend partial interference decoding with partial interference treating-as-noise [16]. By dividing messages into a common stream and several private streams, RSMA can distribute interference more evenly across users and reduce the system's sensitivity to imperfect SIC. Several recent studies have shown that RSMA tends to provide more stable throughput and outage performance than NOMA in satellite or integrated satellite–terrestrial networks [17], especially when channel conditions fluctuate or users have highly unbalanced link qualities. However, most of these studies evaluate RSMA under the infinite blocklength assumption, where error probabilities naturally vanish at high SNR and decoding imperfections are less visible. This leaves an important open question unanswered: How does RSMA behave in short-packet satellite links where reliability, not capacity, becomes the primary metric?

But another layer of complexity arises from the Shadowed-Rician fading environment that characterizes many land-mobile satellite links. Driven by the degree of shadowing-from average to heavy-deep signal fluctuations make the block error rate swing wildly, even when the average signal-to-noise ratio is the same. These swings are even tougher to manage under FBL constraints, where tight margins around the coding gain come into play. Understanding the interplay of RSMA, FBL limits, imperfect SIC, and Shadowed-Rician fading is critical to shaping robust next-generation satellite-terrestrial systems. In this spirit, the present paper provides a detailed framework for the study of RSMA-enabled satellite short-packet communication under FBL constraints. The scaling of

[(1)] Sang-Quang Nguyen is with Posts and Telecommunications Institute of Technology, Ho Chi Minh City 70000, Vietnam. (e-mail: sangnq@ptit.edu.vn)
[(2)] Chi-Bao Le is with Transcosmos Vietnam, Ho Chi Minh City, Vietnam. (e-mail: bao.lc@trans-cosmos.com.vn)
[(*)] Corresponding author.

DOI: 10.36244/ICJ.2026.1.8

reliability with transmit power, codeword length, shadowing severity, and power-splitting choices is derived, as well as a comparison of RSMA against NOMA to uncover practical differences in robustness. These contributions shall serve as guidelines for the system designer to identify regimes where RSMA yields actual gains and where its common-message approach stabilizes error performance in realistic satellite environments. Different from existing RSMA finite-blocklength studies that primarily consider terrestrial or uplink systems, and RSMA-based satellite works that typically assume infinite blocklength, this paper focuses on downlink low Earth orbit (LEO) satellite communications under finite-blocklength constraints, Shadowed-Rician fading, and imperfect SIC.

A. Motivations and Contributions

The main contributions of this paper are outlined as follows:

- We develop a finite-blocklength reliability framework for downlink RSMA-enabled LEO satellite systems over Shadowed-Rician channels, which captures the decoding behavior of both common and private messages under short-packet constraints while explicitly accounting for imperfect SIC, a combination that has not been jointly investigated in existing works.
- Closed-form approximations for the BLER of RSMA streams are derived, explicitly incorporating the impact of imperfect SIC (ipSIC), which is a realistic challenge for practical receivers in satellite environments.
- A detailed comparison with NOMA is presented, showing when and why RSMA provides more reliable short-packet transmission-particularly at moderate SNRs, under severe shadowing, or when the blocklength becomes tight.
- Extensive numerical results validate the analysis and illustrate how transmit power, antenna configuration, blocklength, and shadowing severity jointly shape the reliability of RSMA-based satellite links.

Compared to existing RSMA studies under the finite-blocklength regime, which mainly focus on terrestrial or uplink scenarios, this work investigates downlink RSMA transmission in LEO satellite systems over Shadowed-Rician channels. Moreover, unlike prior satellite RSMA analyses that typically rely on infinite blocklength assumptions, we explicitly characterize finite-blocklength BLER performance while accounting for imperfect SIC effects on both common and private streams.

B. Organization & Notations

The remainder of this paper is structured as follows. Section II describes the system model, including the satellite-terrestrial channel characteristics and the RSMA transmission strategy. Section III develops the finite-blocklength analysis and derives the corresponding expressions for the achievable performance. Section IV reports and discusses the numerical results, highlighting the comparison between RSMA and conventional multiple-access schemes. Finally, Section V summarizes the main findings and outlines potential directions for future work.

Notations: Throughout this paper, we use several frequently used mathematical symbols. The operator $\mathbb{E}\{.\}$ denotes expectation; $J_i(\cdot)$ refers to the Bessel function of the first kind with order i; $\|.\|_F$ is the Frobenius norm; ${}_1F_1(\cdot;\cdot;\cdot)$ represents the confluent hypergeometric function of the first kind; $(\cdot)_t$ is the Pochhammer symbol; $\mathcal{B}(\cdot,\cdot)$ corresponds to the Beta function; $\Gamma(\cdot)$ is the Gamma function; $\gamma(\cdot,\cdot)$ indicates the lower incomplete Gamma function; the Gaussian $\mathcal{Q}$-function is written as $\mathcal{Q}(x) = \frac{1}{2\pi}\int_x^\infty e^{-t^2/2}\,dt$; and for any random variable X, its probability density and cumulative distribution functions are denoted by $f_X(\cdot)$ and $F_X(\cdot)$, respectively.

II. System Model and Terrestrial Channel Model

A. System Description

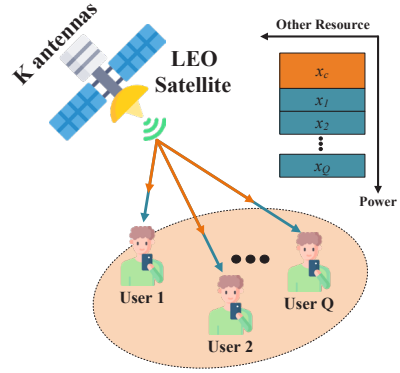


Fig. 1. An illustration of a downlink RSMA-enabled satellite-terrestrial communication system.

Fig. 1 sketches the downlink satellite-terrestrial network considered in this work, where rate-splitting multiple access (RSMA) is adopted. In this setup, a satellite S equipped with K antennas communicates simultaneously with Q ground users. The satellite operates in half-duplex mode and employs RSMA to handle interference while still aiming for better spectral efficiency. In practice, modern satellite platforms usually rely on multibeam architectures to extend coverage and capacity. For LEO systems, these beams are commonly produced by array-fed reflectors-mainly because they tend to be more efficient than direct radiating arrays. Since the beam pattern is fixed, the satellite avoids the need for heavy onboard signal processing.

The signal received at each user is shaped by the beamforming strategy, propagation environment, and power allocation used at the satellite. The channel gain from the satellite to user q is modeled as

$$\mathbf{h}_{U_q} = \mathcal{A}_q \mathbf{g}_{U_q}^{\dagger} \mathbf{w}_{U_q}, \tag{1}$$

where $\mathcal{A}_q$ includes both the antenna gain and free-space path-loss. The vector $\mathbf{g}_{U_q}$ represents the $K \times 1$ Shadowed-Rician fading channel between the satellite antennas and user q. The operator $(.)^{\dagger}$ denotes the conjugate transpose, and $\mathbf{w}_{U_q}$ is the corresponding $K \times 1$ beamforming vector. Following the maximum ratio transmission (MRT) rule, the beamformer is

selected as $\mathbf{w}_{U_q} = \frac{\|\mathbf{g}_{U_q}\|}{\|\mathbf{g}_{U_q}\|_F}$, where $\|.\|_F$ is the Frobenius norm. The factor $\mathcal{A}_q$ remains fixed and is written as

$$\mathcal{A}_q = \frac{c\sqrt{G_S G_R}}{4\pi f_c d_{S_q}}, \tag{2}$$

Here, c and f_c denoting the speed of light and the carrier frequency. The term d_{S_q} is the distance between the satellite and user q. The gains G_S and G_R correspond to the satellite beam gain and the user terminal gain, respectively. Specifically, $G_S = G_{\max}\left[\frac{J_1(w)}{2w} + 36\frac{J_3(w)}{w^3}\right]^2$, where $G_{\max}$ is the maximum beam gain and $w = \frac{2.07123\sin(\theta)}{\sin(\theta_{3\text{dB}})}$. Here, θ denotes the angular offset from the beam center, $\theta_{3\text{dB}}$ the 3 dB beamwidth, and $J_1(\cdot)$, $J_3(\cdot)$ are Bessel functions of order 1 and 3. In this model, different users are separated based on their channel quality obtained, for example, via pilot-assisted channel estimation. We further assume ideal CSI at the transmitter, and the results serve as a reference for future studies involving CSI imperfections [18].

B. The signal processing at transceivers

The satellite employs RSMA to serve all users concurrently. The transmitted waveform consists of two parts: a common message, intended to be decoded by every user, and a private message for each individual user. A power coefficient a_c is assigned to the common stream, whereas the remaining power is split among the private streams.

The transmitted signal is expressed as

$$x = \sqrt{P_S}\left(\sqrt{a_c}x_c + \sum\nolimits_{q=1}^{Q}\sqrt{a_q}x_q\right), \tag{3}$$

where P_S is the satellite transmit power, x_c is the common message with power $a_c P_S$, and x_q is the private message for user q with power $a_q P_S$. It holds that $a_c + \sum_{q=1}^{Q} a_q = 1$. The received signal at user q is then

$$\begin{aligned} y_{U_q} =& \mathbf{h}_{U_q}x + n_{U_q} \\ =& \underbrace{\mathbf{g}_{U_q}^{\dagger}\mathbf{w}_{U_q}\mathcal{A}_q\sqrt{a_c P_S}x_c}_{\text{Common Message}} + \underbrace{\mathbf{g}_{U_q}^{\dagger}\mathbf{w}_{U_q}\mathcal{A}_q\sqrt{a_q P_S}x_q}_{\text{Private Message}} \\ &+ \underbrace{\sum\nolimits_{j=1,j\neq q}^{Q}\mathbf{g}_{U_q}^{\dagger}\mathbf{w}_{U_q}\mathcal{A}_q\sqrt{a_j P_S}x_j}_{\text{Interference}} + \underbrace{n_{U_q}}_{\text{AWGN}}, \end{aligned} \tag{4}$$

where n_{U_q} is AWGN with zero mean and variance N_0. Each user decodes the received signal in two steps.

C. Decoding the common message

Users start by decoding the common message x_c while considering the private streams as interference. The SINR for decoding the common message at user q is

$$\begin{aligned} \bar{\gamma}_{c,q} =& \frac{a_c\mathcal{A}_q^2 P_S\left\|\mathbf{g}_{U_q}\right\|_F^2}{N_0 + \mathcal{A}_q^2 P_S\left(1-a_c\right)\left\|\mathbf{g}_{U_q}\right\|_F^2} \\ =& \frac{a_c\mathcal{C}_q}{1+\left(1-a_c\right)\mathcal{C}_q}, \end{aligned} \tag{5}$$

where $\varrho_S = \frac{P_S}{N_0}$ is the signal-to-noise ratio (SNR), $\mathcal{C}_q = \delta_q\left\|\mathbf{g}_{U_q}\right\|_F^2$ and $\delta_q = \varrho_S\mathcal{A}_q^2$.

D. Decoding the private message

After removing the common message, user q proceeds to decode its own private message. Interference now comes only from private messages of other users. The SINR for private decoding is

$$\bar{\gamma}_{p,q} = \frac{a_q\mathcal{C}_q}{1+\mathcal{C}_q\sum_{j=1,q\neq j}^{Q}a_j + \Xi_q a_c\mathcal{C}_q}, \tag{6}$$

where Ξ_q accounts for imperfect SIC (ipSIC), with $\Xi_q = 0$ corresponding to ideal cancellation [19].

E. Terrestrial Channel Model

To analyze performance metrics later on, we assume the fading coefficients are i.i.d. The PDF of the channel coefficient $g_{U_q}^{(k)}$ from the kth satellite antenna to user q is

$$f_{\left|g_{U_q}^{(k)}\right|^2}(x) = \alpha_q e^{-\beta_q x}{}_1F_1\left(m_q;1;\varpi_q x\right)\ ,\ x\geq 0\ , \tag{7}$$

where α_q, β_q, and ϖ_q are defined as in the original expression, with Ω_q, $2b_q$, and m_q denoting the LOS power, multipath power, and fading severity, respectively. The term ${}_1F_1(\cdot)$ is the confluent hypergeometric function [20, Eq. (9.210.1)].

Assuming m_q takes integer values, the PDF simplifies to

$$f_{\left|g_{U_q}^{(k)}\right|^2}(x) = \alpha_q e^{-(\beta_q-\varpi_q)x}\sum_{t=0}^{m_q-1}\zeta_q(t)\,x^t\ ,\ x\geq 0\ , \tag{8}$$

in which $\zeta_q(t)$ defined in terms of the Pochhammer symbol. Using [21], the PDF of $\mathcal{C}_q$ under i.i.d. Shadowed-Rician fading is

$$f_{\mathcal{C}_q}(x) = \sum_{j_1=0}^{m_q-1}\cdots\sum_{j_K=0}^{m_q-1}\frac{\Lambda_q(K)}{\delta_q^{\Delta_q}}x^{\Delta_q-1}e^{-\left(\frac{\psi_q}{\delta_q}\right)x}, \tag{9}$$

where

$$\Lambda_q(K) = \alpha_q^K\prod\nolimits_{l=1}^{K}\zeta_q(j_l)\prod\nolimits_{u=1}^{K-1}\mathcal{B}\left(\sum\nolimits_{p=1}^{u}j_p+u, j_{u+1}+1\right), \tag{10}$$

and

$$\Delta_q = \sum\nolimits_{l=1}^{K}j_l + K,\quad \psi_q = \beta_q - \delta_q. \tag{11}$$

Applying [20, Eq. (3.351.1)], the CDF of $\mathcal{C}_q$ becomes

$$F_{\mathcal{C}_q}(x) = \sum_{j_1=0}^{m_q-1}\cdots\sum_{j_K=0}^{m_q-1}\frac{\Lambda_q(K)}{\psi_q^{\Delta_q}}\gamma\left(\Delta_q, \frac{\psi_q x}{\delta_q}\right), \tag{12}$$

which can be further simplified via [20, Eq. (8.352.6)] to

$$F_{\mathcal{C}_q}(x) = 1 - \sum_{j_1=0}^{m_q-1}\cdots\sum_{j_K=0}^{m_q-1}\sum_{p=0}^{\Delta_q-1}\frac{\Lambda_q(K)\,\Gamma(\Delta_q)}{p!\psi_q^{\Delta_q-p}\delta_q^p}e^{-\frac{\psi_q x}{\delta_q}}x^p. \tag{13}$$

Remark 1: This characterization highlights how the channel statistics influence system behavior under the Shadowed-Rician fading considered in this work.

III. FINITE-BLOCKLENGTH ANALYSIS

In this section, we analyze the reliability performance of the RSMA-enabled satellite-terrestrial system when short-packet transmission is adopted. Unlike the infinite blocklength (IBL) regime, where decoding errors can be made arbitrarily small, finite-blocklength (FBL) theory reveals that the block error rate (BLER) remains strictly positive even when the coding rate lies below Shannon capacity. Therefore, the performance of the satellite link must be re-evaluated under the FBL regime to support delay-critical satellite applications such as emergency response, low-latency sensing, and real-time IoT services.

Throughout this section, the SINRs used for decoding the common and private RSMA messages follow (5) and (6), and the distribution of the effective channel gain $\mathcal{A}_q$ follows the Shadowed-Rician model in (9)-(13) in Section II.

A. Reliability Analysis

The decoding reliability at user U_q for both the common message and the private message can be approximated using the finite-blocklength framework. For moderately large blocklengths (typically $\mathcal{L}_{a,q} > 100$), the instantaneous BLER can be written as

$$\varepsilon_{a,q} \simeq \mathcal{Q}\left(\frac{\mathcal{C}(\bar{\gamma}_{a,q}) - \mathcal{R}_{a,q}}{\sqrt{\mathcal{V}(\bar{\gamma}_{a,q})\,\mathcal{L}_{a,q}^{-1}}}\right), \qquad a \in \{c, p\}, \tag{14}$$

where $a = c$ corresponds to the common message and $a = p$ refers to the private message. Here, $\mathcal{R}_{a,q} = \mathcal{T}_{a,q}/\mathcal{L}_{a,q}$ denotes the coding rate, with $\mathcal{L}_{a,q}$ being the blocklength and $\mathcal{T}_{a,q}$ the number of transmitted information bits. The term $\mathcal{C}(\bar{\gamma}_{a,q}) = \log_2(1+\bar{\gamma}_{a,q})$ represents the Shannon capacity associated with the received SINR $\bar{\gamma}_{a,q}$, while the channel dispersion is given by $\mathcal{V}(\bar{\gamma}_{a,q}) = \left[1 - (1+\bar{\gamma}_{a,q})^{-2}\right](\log_2 e)^2$. The function $\mathcal{Q}(\cdot)$ stands for the Gaussian Q-function.

From (14), the average BLER at user U_q for decoding both the common message and the private message can be expressed as

$$\tilde{\varepsilon}_{a,q} = \mathbb{E}\left\{\varepsilon_{a,q}\right\}. \tag{15}$$

The total average BLER at user U_q, denoted by $\tilde{\varepsilon}_q$, is obtained by averaging the BLER contributions from both the common and private messages. Accordingly, based on (15), we can write

$$\tilde{\varepsilon}_q = (\tilde{\varepsilon}_{c,q} + \tilde{\varepsilon}_{p,q})/2. \tag{16}$$

The closed-form formulation of the average BLER at each user is obtained in the following proposition.

Proposition 1: The closed-form approximate expression for the total average BLER at user U_q, denoted by $\tilde{\varepsilon}_q$, can be written as

$$\begin{aligned}\tilde{\varepsilon}_q \approx \frac{1}{2}\Bigg\langle &\Bigg[1 - \sum_{t=1}^{\mathbb{T}} \frac{1}{\mathbb{T}} \sum_{j_1=0}^{m_q-1} \cdots \sum_{j_K=0}^{m_q-1} \sum_{p=0}^{\Delta_q-1} \frac{\Lambda_q(K)\,\Gamma(\Delta_q)}{p!\psi_q^{\Delta_q-p}\delta_q^p} \\ &\times e^{-\frac{\psi_q}{\delta_q}\left(\frac{\Phi_{c,q}}{a_c - \Phi_{c,q}(1-a_c)}\right)} \left(\frac{\Phi_{c,q}}{a_c - \Phi_{c,q}(1-a_c)}\right)^p\Bigg] \\ &+ \Bigg[1 - \sum_{t=1}^{\mathbb{T}} \frac{1}{\mathbb{T}} \sum_{j_1=0}^{m_q-1} \cdots \sum_{j_K=0}^{m_q-1} \sum_{p=0}^{\Delta_q-1} \frac{\Lambda_q(K)\,\Gamma(\Delta_q)}{p!\psi_q^{\Delta_q-p}\delta_q^p} \\ &\times e^{-\frac{\psi_q}{\delta_q}\left(\frac{\Phi_{p,q}}{a_q - \Phi_{p,q}\left(\sum_{j=1,q\neq j}^{Q} a_j + \Xi_q a_c\right)}\right)} \\ &\times \left(\frac{\Phi_{p,q}}{a_q - \Phi_{p,q}\left(\sum_{j=1,q\neq j}^{Q} a_j + \Xi_q a_c\right)}\right)^p\Bigg]\Bigg\rangle.\end{aligned} \tag{17}$$

Remark 2: Under imperfect SIC, the average BLER does not vanish at high SNR and instead converges to an error floor dominated by residual interference. The impact of finite blocklength enters through the channel dispersion term, which becomes more pronounced in the short-packet regime. In this setting, RSMA benefits from the common stream, which provides an additional decoding layer and reduces sensitivity to residual interference compared to NOMA. Consequently, RSMA exhibits more stable reliability, particularly at moderate SNRs where finite-blocklength effects and imperfect SIC jointly limit performance.

Proof: Substituting $\varepsilon_{a,q}$ from (14) to (15), the average BLER at U_q for decoding both the common message and the private message can be expressed as

$$\tilde{\varepsilon}_{a,q} \simeq \mathbb{E}\left\{\mathcal{Q}\left(\frac{\mathcal{C}(\bar{\gamma}_{a,q}) - \mathcal{R}_{a,q}}{\sqrt{\mathcal{V}(\bar{\gamma}_{a,q})\,\mathcal{L}_{a,q}^{-1}}}\right)\right\}. \tag{18}$$

Obtaining a closed-form expression for (18) is, unfortunately, analytically intractable. For this reason, and following the approach in [22], we approximate $\tilde{\varepsilon}_{a,q}$ by

$$\tilde{\varepsilon}_{a,q} \approx \int_0^\infty \Upsilon_{a,q}(x)\, f_{\bar{\gamma}_{a,q}}(x)\, dx, \tag{19}$$

where $f_{\bar{\gamma}_{a,q}}(\cdot)$ denotes the PDF of the received SINR $\bar{\gamma}_{a,q}$. The function $\Upsilon_{a,q}(\cdot)$ serves as a tractable approximation of the term $\mathcal{Q}\left(\frac{\mathcal{C}(\bar{\gamma}_{a,q}) - \mathcal{R}_{a,q}}{\sqrt{\mathcal{V}(\bar{\gamma}_{a,q})\,\mathcal{L}_{a,q}^{-1}}}\right)$ and might be expressed as

$$\Upsilon_{a,q}(\bar{\gamma}_{a,q}) = \begin{cases} 1, & \bar{\gamma}_{a,q} \le \phi_{a,q}, \\ 0.5 - \mathcal{G}_{a,q}\sqrt{\mathcal{L}_{a,q}}\,(\bar{\gamma}_{a,q} - \chi_{a,q}), & \phi_{a,q} < \bar{\gamma}_{a,q} < \varsigma_{a,q}, \\ 0, & \bar{\gamma}_{a,q} \ge \varsigma_{a,q}, \end{cases} \tag{20}$$

where $\mathcal{G}_{a,q} = \left(2\pi\sqrt{2^{2\mathcal{R}_{a,q}} - 1}\right)^{-1}$, $\chi_{a,q} = 2^{\mathcal{R}_{a,q}} - 1$, $\phi_{a,q} = \chi_{a,q} - \left(2\mathcal{G}_{a,q}\sqrt{\mathcal{L}_{a,q}}\right)^{-1}$, and $\varsigma_{a,q} = \chi_{a,q} + \left(2\mathcal{G}_{a,q}\sqrt{\mathcal{L}_{a,q}}\right)^{-1}$. Next, by following the approach in [19] and applying integration by parts, the expression in (19) can be simplified as

$$\tilde{\varepsilon}_{a,q} \approx \mathcal{G}_{a,q}\sqrt{\mathcal{L}_{a,q}} \int_{\phi_{a,q}}^{\varsigma_{a,q}} F_{\bar{\gamma}_{a,q}}(x) dx. \tag{21}$$

Since the gap between $\phi_{a,q}$ and $\varsigma_{a,q}$ in (19) is relatively small [23], $\tilde{\varepsilon}_{a,q}$ can be further simplified by

$$\tilde{\varepsilon}_{a,q} \approx \begin{cases} \sum_{t=1}^{\mathbb{T}} \frac{1}{\mathbb{T}} F_{\mathcal{C}_q}\left(\frac{\Phi_{a,q}}{a_c - \Phi_{a,q}(1-a_c)}\right), & a \in c, \\ \sum_{t=1}^{\mathbb{T}} \frac{1}{\mathbb{T}} F_{\mathcal{C}_q}\left(\frac{\Phi_{a,q}}{a_q - \Phi_{a,q}\left(\sum_{j=1, q\neq j}^{Q} a_j + \Xi_q a_c\right)}\right), & a \in p, \end{cases} \tag{22}$$

Here, $\mathbb{T}$ implies the complexity accuracy trade-off parameter and $\Phi_{a,q} = \phi_{a,q} + (2t-1)(\varsigma_{a,q} - \phi_{a,q})/2\mathbb{T}$.

Using (13) and (22), we can derive the close-form of $\tilde{\varepsilon}_{a,q}$ as

$$\tilde{\varepsilon}_{a,q} \approx 1 - \sum_{t=1}^{\mathbb{T}} \frac{1}{\mathbb{T}} \sum_{j_1=0}^{m_q-1} \cdots \sum_{j_K=0}^{m_q-1} \sum_{p=0}^{\Delta_q-1} \frac{\Lambda_q(K)\,\Gamma(\Delta_q)}{p!\psi_q^{\Delta_q-p}\delta_q^p} \times e^{-\frac{\psi_q \xi_{a,q}}{\delta_q}} \xi_{a,q}^p, \tag{23}$$

where $\xi_{a,q} = \frac{\Phi_{a,q}}{a_c - \Phi_{a,q}(1-a_c)}$ for the common part ($a = c$), whereas for the private part ($a = p$), it is given by $\xi_{a,q} = \frac{\Phi_{a,q}}{a_q - \Phi_{a,q}\left(\sum_{j=1, j\neq q}^{Q} a_j + \Xi_q\right) a_c}$.

We derive the closed-form expression of the average BLER at each user in (17) by substituting (23) in (16).

The proof is completed. ■

Reliability at U_q is the chance that the packet is accurately identified, given as [24, Eq. (27)]

$$\nu_q = (1 - \tilde{\varepsilon}_q) \times 100\%. \tag{24}$$

Furthermore, the throughput for each user may be assessed as

$$\tau_q = (1 - \tilde{\varepsilon}_{c,q})\,\mathcal{R}_{c,q} + (1 - \tilde{\varepsilon}_{p,q})\,\mathcal{R}_{p,q}. \tag{25}$$

IV. Results and Discussions

In this segment, we conduct a series of numerical simulations to check how well the analytical expressions capture the behavior of the considered RSMA-enabled satellite-terrestrial system. The Shadowed-Rician fading parameters are chosen according to two typical shadowing conditions often used in satellite literature [25]: an average-shadowing case with $(b_q, m_q, \Omega_q) = (0.251, 5, 0.279)$ and a heavier shadowing scenario characterized by $(0.063, 1, 0.0007)$. To keep the setup simple, the system involves two terrestrial users ($Q = 2$), $\mathcal{L} = \mathcal{L}_{c,q} = \mathcal{L}_{p,q} = 200$, $\mathcal{T} = \mathcal{T}_{c,q} = \mathcal{T}_{p,q} = 80$ bits, both assumed to experience the same SIC imperfection level, i.e., $\Xi = \Xi_1 = \Xi_2$ and $\mathbb{T} = 5$. The equivalent noise power at each user terminal is modeled as $N_0 = \kappa B_w T$, where $\kappa = 1.38 \times 10^{-23}$ denotes the Boltzmann constant, $B_w = 50$ MHz is the transmission bandwidth, and the noise temperature is set to $T = 290$ K, a value commonly used for LEO downlink links [26]. The satellite operates with $K = 2$ antennas and radiates from a LEO platform toward a ground terminal located about $d_{S_q} = 1000$ km away. The carrier frequency is $f_c = 1.55$ GHz, and the propagation constants follow the typical free–space path-loss model. For the antenna characteristics, we adopt a receive antenna gain of $G_R = 5$ dBi and a maximum satellite beam gain of $G_{\max} = 25$ dBi, with an angular separation of $\theta = 0.8°$ between the users and a 3-dB beamwidth of $\theta_{3\text{dB}} = 0.4°$. To evaluate the system performance under these settings, each simulation curve is averaged over 10^6 Monte Carlo realizations. Regarding the power distribution, the common-message power ratio is fixed at $a_c = 0.4$, and the remaining power is proportionally divided among the users according to $a_1 = 0.4(1 - a_c)$ and $a_2 = 0.6(1 - a_c)$, reflecting a mild imbalance in private-message allocation. For NOMA, a fixed power allocation and the conventional SIC decoding order are assumed, which are commonly adopted for baseline comparisons.

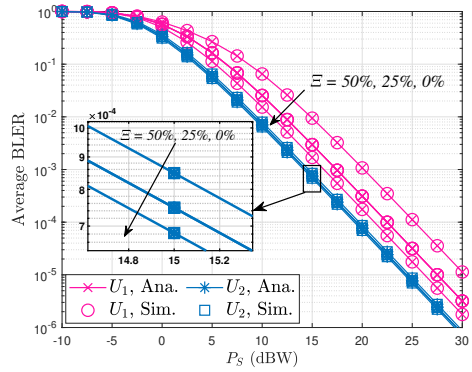


Fig. 2. Average BLER versus transmit power for different levels of SIC imperfection, with $K = 2$.

In Figure 2, we examine how the average BLER varies with transmit power under different levels of SIC imperfection. As expected, all curves improve gradually as the SNR increases, but the presence of residual interference still leaves a visible gap between the ideal and imperfect cases. At low SNR, the average BLER of all curves remains relatively high and close together, suggesting that noise dominates in this region. When the SNR approaches the medium range, the impact of the errors in the SIC starts becoming significant, and a small level of imprecision causes the average BLER convergence speed to reduce. For the high SNR range, it appears that the system with the highest level of imprecision for the SIC will eventually converge to the error floor, which suggests that the presence of the interference will affect the reliability of the RSMA, even when it operates at higher powers.

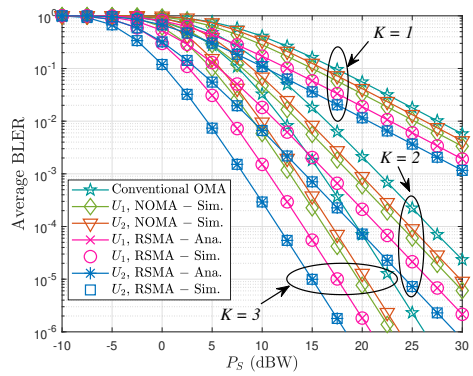


Fig. 3. Average BLER versus transmit power with increasing number of antennas at the satellite, for $\Xi = 10\%$.

Figure 3 illustrates the impact of the number of satellite antennas on the BLER performance of RSMA, NOMA, and conventional OMA. As the antenna number increases, RSMA tends to achieve a more pronounced BLER reduction, benefiting more effectively from the additional array gain. In the low-SNR region, the performance of RSMA and NOMA remains close, since noise largely dominates the system behavior. Once the SNR exceeds a certain threshold, however, a clear performance separation emerges, with RSMA outperforming both NOMA and OMA. It is also apparent that $K = 3$, the RSMA curve drops more rapidly than those of NOMA and conventional OMA, which highlights the effectiveness of multi-antenna processing combined with rate-splitting. At higher SNR values, this performance advantage is largely preserved, while the improvement of NOMA gradually saturates due to residual interference.

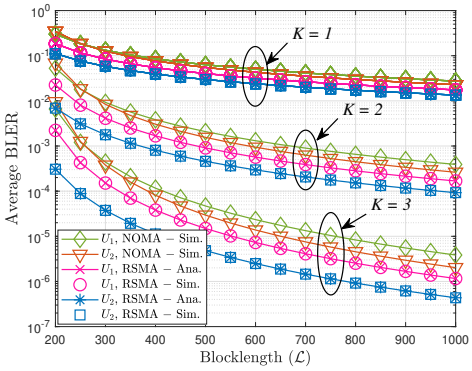


Fig. 4. Average BLER versus blocklength with increasing number of antennas at the satellite, for $\Xi = 20\%$ and $P_S = 10$ dBW.

Figure 4 examines the impact of blocklength on the BLER performance of RSMA and NOMA. As expected, increasing the blocklength improves reliability for both schemes due to reduced channel dispersion. However, RSMA exhibits a faster BLER decay, particularly in the short and moderate blocklength regimes, since the common message provides an additional decoding path for users experiencing unfavorable channel conditions. In contrast, NOMA degrades more rapidly when packets are short, as its performance relies heavily on reliable SIC, which becomes less effective under FBL. As the blocklength grows large, the performance gap gradually narrows and eventually saturates, indicating diminishing FBL effects. Overall, the results suggest that RSMA achieves higher reliability and better energy efficiency than NOMA in short-packet satellite transmissions.

Figure 5 illustrates the BLER behavior of the satellite link under average shadowing (AS) and heavy shadowing (HS) conditions. Under AS, where signal blockage is moderate, the channel remains relatively stable and BLER decreases rapidly with increasing transmit power. In contrast, HS introduces severe signal obstruction and deep fading, resulting in persistently higher BLER and a much slower improvement even with power boosting. These results indicate that while RSMA remains robust under AS conditions, severe shadowing imposes a fundamental limitation that cannot be fully mitigated by transmit power alone.

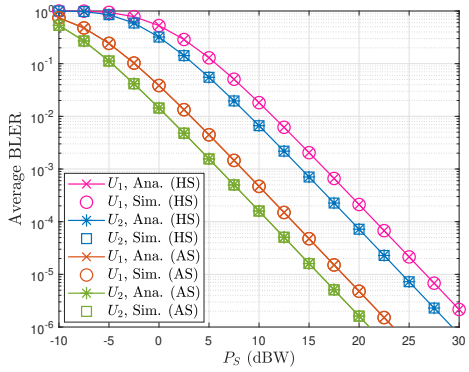


Fig. 5. Average BLER versus transmit power under various shadow fading, with $\Xi = 10\%$ and $K = 2$.

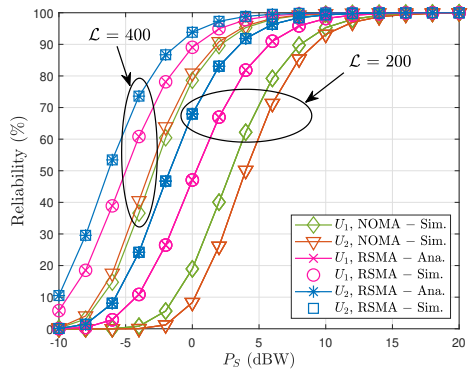


Fig. 6. Reliability versus transmit power for different blocklength values, with $\Xi = 10\%$ and $K = 2$.

Figure 6 presents the reliability performance versus transmit power for different blocklengths, comparing RSMA and NOMA. For all blocklength values, increasing power enhances reliability; however, RSMA consistently achieves a target BLER at lower SNR than NOMA. This advantage becomes more pronounced for short packets, where NOMA suffers from residual interference and imperfect SIC, leading to a noticeably higher BLER in the mid-SNR region. As the blocklength increases, both schemes benefit, yet RSMA maintains a clear edge owing to its common stream, which provides additional decoding robustness. These observations confirm that RSMA is more suitable for reliability-critical short-packet satellite links.

Finally, Figure 7 compares the throughput performance of RSMA and NOMA under different satellite antenna configurations. Increasing the transmit power improves throughput for both schemes; however, RSMA consistently achieves higher throughput due to its superior decoding reliability. Moreover, the use of additional antennas enhances the SINR of both common and private streams in RSMA, allowing it to better exploit spatial diversity. At moderate SNR, RSMA exhibits a steeper throughput increase, whereas NOMA remains constrained by

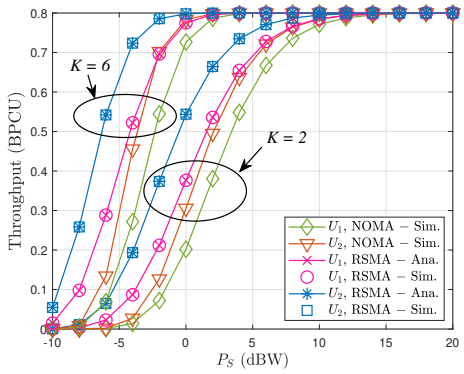


Fig. 7. Throughput versus transmit power with increasing number of antennas at the satellite, for $\Xi = 10\%$ and $K = \{2, 6\}$.

residual interference and BLER effects. Even at high SNR, the throughput gap persists, highlighting that RSMA more effectively converts reliability gains into throughput benefits under short-packet conditions.

It is worth noting that the analytical and simulation results presented in this section are obtained under the assumption of perfect CSI at the receiver. In practical LEO satellite systems, channel estimation errors may arise due to mobility, Doppler effects, and limited pilot resources, which can degrade the reliability performance to some extent. Nevertheless, the results reported here serve as a useful benchmark for evaluating the fundamental behavior of RSMA under finite-blocklength transmission. Incorporating CSI imperfections into the proposed framework is an interesting extension and will be considered in future work.

V. Conclusion

This paper explored how RSMA behaves in satellite short-packet communications when finite-blocklength and realistic channel conditions are taken into account. By building a tractable FBL model over Shadowed-Rician fading, we derived BLER expressions for both common and private streams, including the effect of imperfect SIC–something that often becomes a real bottleneck in practical receivers. The analysis and simulations align well and point to a consistent trend: RSMA tends to hold its reliability better than NOMA, especially when the packets are short or the shadowing becomes heavy. The common stream, although simple in structure, seems to play a meaningful role in stabilizing the error performance when the SNR is moderate and when ipSIC degrades the private layers. Numerical results further show that RSMA normally requires less transmit power to reach comparable BLER targets and scales more gracefully with blocklength and antenna count. While this study focused on a basic two-user setting, the insights suggest that RSMA is a promising direction for future satellite systems that must deliver reliable, low-latency short-packet links. Extending the model to multi-beam satellites, multi-user scenarios, or imperfect CSI would be a natural next step for future work.

References

[1] X. Liu, K.-Y. Lam, F. Li, J. Zhao, L. Wang, and T. S. Durrani, "Spectrum sharing for 6g integrated satellite-terrestrial communication networks based on noma and cr," *IEEE Network*, vol. 35, no. 4, pp. 28–34, 2021. DOI: 10.1109/MNET.011.2100021.

[2] X. Fang, W. Feng, T. Wei, Y. Chen, N. Ge, and C.-X. Wang, "5g embraces satellites for 6g ubiquitous iot: Basic models for integrated satellite terrestrial networks," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 14 399–14 417, 2021. DOI: 10.1109/JIOT.2021.3068596.

[3] Q. Wang, X. Chen, and Q. Qi, "Energy-efficient design of satellite-terrestrial computing in 6g wireless networks," *IEEE Transactions on Communications*, vol. 72, no. 3, pp. 1759–1772, 2024. DOI: 10.1109/TCOMM.2023.3334813.

[4] N. Abbas, A. Mrad, A. Ghazleh, and S. Sharafeddine, "Uav-based relay system for iot networks with strict reliability and latency requirements," *IEEE Networking Letters*, vol. 3, no. 3, pp. 110–113, 2021. DOI: 10.1109/LNET.2021.3077869.

[5] E. E. Haber, H. A. Alameddine, C. Assi, and S. Sharafeddine, "Uav-aided ultra-reliable low-latency computation offloading in future iot networks," *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6838–6851, 2021. DOI: 10.1109/TCOMM.2021.3096559.

[6] C. Feng, H.-M. Wang, and H. V. Poor, "Reliable and secure short-packet communications," *IEEE Transactions on Wireless Communications*, vol. 21, no. 3, pp. 1913–1926, 2022. DOI: 10.1109/TWC.2021.3108042.

[7] J. Cao, X. Zhu, S. Sun, Z. Wei, Y. Jiang, J. Wang, and V. K. Lau, "Toward industrial metaverse: Age of information, latency and reliability of short-packet transmission in 6g," *IEEE Wireless Communications*, vol. 30, no. 2, pp. 40–47, 2023. DOI: 10.1109/MWC.2001.2200396.

[8] X. Gao, L. Shi, Y. Ye, G. Zheng, and G. Lu, "Reliability-oriented resource allocation in short-packet backscatter communications," *IEEE Transactions on Vehicular Technology*, vol. 74, no. 2, pp. 3468–3473, 2025. DOI: 10.1109/TVT.2024.3472093.

[9] J. Xu, O. Dizdar, and B. Clerckx, "Rate-splitting multiple access for short-packet uplink communications: A finite blocklength analysis," *IEEE Communications Letters*, vol. 27, no. 2, pp. 517–521, 2023. DOI: 10.1109/LCOMM.2022.3226817.

[10] H. Shuai, K. Guo, K. An, Y. Huang, and S. Zhu, "Transmit antenna selection in noma-based integrated satellite-hap-terrestrial networks with imperfect csi and sic," *IEEE Wireless Communications Letters*, vol. 11, no. 8, pp. 1565–1569, 2022. DOI: 10.1109/LWC.2022.3165710.

[11] G. Xu, Z. Zhao, Z. Song, Q. Zhang, and B. Ai, "Symbol error analysis for integrated satellite- terrestrial relay networks with non-orthogonal multiple access under hardware impairments," *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 12 980–12 994, 2024. DOI: 10.1109/TWC.2024.3397847.

[12] S. Mondal, K. Singh, D. W. Kwan Ng, C.-P. Li, and Z. Ding, "Outage performance of ris-aided noma isac network for leo satellite system," in *2025 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2025, pp. 1556–1561. DOI: 10.1109/ICCWorkshops67674.2025.11162359.

[13] Z. Belső and L. Pap, "Effect of the imperfect channel estimation on achievable noma rate," vol. 17, no. 1, pp. 2–10, 2023. DOI: 10.21203/rs.3.rs-3310816/v1.

[14] H. Zeng, R. Zhang, X. Zhu, J. Cao, Y. Jiang, F.-C. Zheng, and P. Dai, "Frame structure and resource optimization for hybrid long- and short-packet noma-based data collection in iiot with imperfect sic," *IEEE Internet of Things Journal*, vol. 11, no. 23, pp. 37 799–37 812, 2024. DOI: 10.1109/JIOT.2024.3444462.

[15] S. Kumar, B. Kumbhani, and S. Darshi, "Performance analysis of star-ris aided short-packet noma network under imperfect sic and csi," *IEEE Systems Journal*, vol. 19, no. 2, pp. 600–611, 2025. DOI: 10.1109/JSYST.2025.3562732.

[16] V. S. Nguyen, A. Le-Thi, V. D. Thuan, C.-B. Le, T. H. Nguyen, and S.-Q. Nguyen, "Analysis of ergodic sum rate in rsma with perfect and imperfect sic: A multiple-antenna selection approach for optimizing uav positioning," *Physical Communication*, vol. 72, p. 102 741, 2025. DOI: 10.1016/j.phycom.2025.102741.

[17] C. Hu, H. Yang, Z. Zhou, B. Li, X. Jiang, N. Zhao, and A. Nal- lanathan, "Outage analysis of rsma enabled integrated satellite-terrestrial networks," *IEEE Transactions on Vehicular Technology*, vol. 74, no. 6, pp. 10 023–10 028, 2025. DOI: 10.1109/TVT.2025.3539758.

[18] C. K. Singh, P. K. Upadhyay, J. Lehtomäki, and M. Juntti, "Performance analysis with deep learning assay for cooperative uav-borne irs noma networks under non-ideal system imperfections," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 1, pp. 1065–1083, 2024. DOI: 10.1109/TVT.2023.3309619.

[19] S. K. Singh, K. Agrawal, K. Singh, Y.-M. Chen, and C.-P. Li, "Performance analysis and optimization of rsma enabled uav-aided ibl and fbl communication with imperfect sic and csi," *IEEE Transactions on Wireless Communications*, vol. 22, no. 6, pp. 3714–3732, 2023. DOI: 10.1109/TWC.2022.3220785.

[20] I. S. Gradshteyn and I. M. Ryzhik, Table of integrals, series, and products. *Academic press*, 2014. DOI: 10.1016/C2010-0-648395.

[21] V. Bankey, P. K. Upadhyay, D. B. Da Costa, P. S. Bithas, A. G. Kanatas, and U. S. Dias, "Performance analysis of multi-antenna multiuser hybrid satellite-terrestrial relay systems for mobile services delivery," *IEEE Access*, vol. 6, pp. 24 729–24 745, 2018. DOI: 10.1109/ACCESS.2018.2830801.

[22] C. D. Ho, T.-V. Nguyen, T. Huynh-The, T.-T. Nguyen, D. B. da Costa, and B. An, "Short-packet communications in wireless-powered cognitive iot networks: Performance analysis and deep learning evaluation," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2894–2899, 2021. DOI: 10.1109/TVT.2021.3061157.

[23] T.-H. Vu, T.-V. Nguyen, Q.-V. Pham, D. Benevides da Costa, and S. Kim, "Star-ris-enabled short-packet noma systems," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 10, pp. 13 764–13 769, 2023. DOI: 10.1109/TVT.2023.3278737.

[24] G. N. Tran and S. Kim, "Performance analysis of short packets in noma vlc systems," *IEEE Access*, vol. 10, pp. 6505–6517, 2022. DOI: 10.1109/ACCESS.2022.3141865.

[25] N. I. Miridakis, D. D. Vergados, and A. Michalas, "Dual-hop communication over a satellite relay and shadowed rician channels," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 9, pp. 4031–4040, 2015. DOI: 10.1109/TVT.2014.2361832.

[26] A. Abdi, W. Lau, M.-S. Alouini, and M. Kaveh, "A new simple model for land mobile satellite channels: first- and second-order statistics," *IEEE Transactions on Wireless Communications*, vol. 2, no. 3, pp. 519–528, 2003. DOI: 10.1109/TWC.2003.811182.

Sang-Quang Nguyen received the B.E. degree in Electrical Engineering from Ho Chi Minh City University of Transport, Vietnam, in 2010, the M.E. degree in Telecommunications Engineering from Ho Chi Minh City University of Technology, Vietnam, in 2013, and the Ph.D. degree in Electrical Engineering from the University of Ulsan, South Korea, in 2017. From 2017 to 2021, he was a Lecturer at Duy Tan University, Vietnam. Since May 2021, he has been a Lecturer at Ho Chi Minh City University of Transport, Vietnam. In September 2024, he joined the Post and Telecommunications Institute of Technology, Ho Chi Minh City, as a Lecturer. He also served as a Research Fellow at Queen's University Belfast, United Kingdom, where he contributed to advancements in wireless communications. Dr. Sang's research interests include cooperative communications, cognitive radio networks, physical layer security, non-orthogonal multiple access (NOMA), short-packet communications, and backscatter communications. His work primarily focuses on secure and energy-efficient communication solutions for next-generation wireless networks. Dr. Sang can be contacted via email at sangnq@ptit.edu.vn.

Chi-Bao LE was born in Binh Thuan province, Viet Nam. He has worked closely with Dr. Thuan at Wireless Communications and Signal Processing Research Group at Industrial University of Ho Chi Minh City, Vietnam. He is currently pursue Master degree in field of wireless communications. His research interest includes electronic design, signal processing in wireless communications network, non- orthogonal multiple access, physical layer security.

Evaluating Data Transmission Performance in 5G mmWave Networks using Multi-Layer Transmission and MIMO Technology

John Baghous[1], and Mohamed Khaled Chahine[2]

***Abstract*—The transition from 4G to 5G networks was necessitated by fundamental limitations in spectral efficiency and data capacity inherent to the existing framework. Fifth-generation (5G) systems address these constraints by capitalizing on key enabling technologies, such as mmWave spectrum, multiple-input multiple- output (MIMO), massive MIMO (mMIMO), beamforming (BF) or Precoding. This paper investigates a multi-layer transmission scheme employing MIMO with Precoding (SVD) as a cooperative technique to enhance downlink (DL) data transmission performance in an enhanced mobile broadband (eMBB) scenario. The study operates within the standard 5G mmWave frequency band (FR2 at 40 GHz). We differentiate between key performance metrics: the user-experienced data rate (or throughput) Measured at the Receiver (Rx) and the peak theoretical data rate (or Bit Rate) Measured at the Transmitter (Tx). Simulation results, conducted using MATLAB, demonstrate that the proposed approach significantly improves both the achievable throughput and spectral efficiency within a fixed bandwidth. Throughput is evaluated in absolute terms (Mbps) and as a normalized percentage of the peak theoretical data rate (Bit Rate). The core of this study examines the impact of the number of spatial data streamng layers on a 5G-NR system performance. While increasing the transmission layers enhances the potential peak data rate at the transmitter, it concurrently elevates the bit error rate (BER) at the receiver, ultimately degrading the net throughput. This underscores the necessity for advanced receiver-side technologies, such as MIMO processing, to counteract high-path loss and other impairments prevalent at mmWave frequencies. The results confirm that augmenting the number of antennas in the MIMO configuration effectively mitigates this limitation. It improves the overall throughput and reduces the received BER by enhancing spatial diversity and signal recovery capabilities.**

***Index Terms*—mmWave, 5G, MIMO, massive MIMO, eMBB, multi-layer transmission scheme, CDL channel model, Data transmission, Bit Rate, Throughput, Spectrum efficiency.**

I. INTRODUCTION

5G technology, the fifth generation of wireless communication systems, represents a significant enhancement to global communication infrastructure, offering substantially higher data rates, lower latency, and greater network capacity [1,2].

[1] Department of Electronics and Communications Engineering, Faculty of Mechanical and Electrical Engineering, DU, Syria (DU e-mail: john.baghous@damascusuniversity.edu.sy) (Private e-mail: m.e.john.baghous@gmail.com)

[2] Department of Electronics and Communications Engineering, Faculty of Mechanical and Electrical Engineering, DU, Syria (e-mail: mk.chahine@damascusuniversity.edu.sy)

DOI: 10.36244/ICJ.2026.1.9

These systems enable a wide range of new applications across diverse fields such as manufacturing, healthcare, and intelligent transportation. This expansion, driven by the proliferation of data-intensive applications, has led to an exponential increase in user-generated data. Consequently, both academia and industry face the ongoing challenge of developing advanced techniques and technologies to meet rising demands and enhance the quality of service (QoS) for end-users [3].

As the number of connected devices continues to rise, global mobile data traffic is projected to grow by several orders of magnitude, with some estimates suggesting an increase of up to 20,000 times current volumes by 2030 [4]. This trend is already observable, with many countries reporting substantial annual growth in total data traffic [5].

This study investigates the prospective use cases for data transmission within the 37-40 GHz frequency band (n260 band as 3GPP TS 38.101-2), a spectrum specified in the 5G NR FR2 standard by 3GPP Release 15 and beyond. The simulation platform was fully developed and validated in compliance with the relevant 3GPP specifications to ensure standardization fidelity.

The remainder of this paper is organized as follows: Section II presents the adopted methodology and a concise overview of the key theoretical foundations. Section III details the practical system implementation and defines the performance metrics. Section IV presents and analyzes the simulation results. Finally, Section V provides a comprehensive discussion and concludes the paper.

II. Methods and Experiments

Given the infeasibility of conducting physical experiments and measurements, this study employed mathematical modeling, programming, and simulation as its primary methodological tools. This computational approach was deemed appropriate and sufficient for the defined research objectives.

The validity and accuracy of the implemented simulation platform were rigorously verified through a multi-faceted comparison. This included aligning the simulation outputs

with established mathematical models, benchmarking against performance tables and specifications outlined in relevant 3GPP technical reports and standards, and evaluating the consistency of curve behaviors and numerical trends with findings from prior research. All simulations were executed on a standard computing workstation using industry-standard software.

III. 5G System and Millimeter Wave Technology

South Korea pioneered the commercial deployment of 5G networks in March 2019. Since that initial launch, 5G technology has undergone substantial evolution and enhancement. The 5G system architecture is fundamentally designed to support three primary service categories, as illustrated in Figure 1 [6]:

Enhanced Mobile Broadband (eMBB), Ultra-Reliable Low-Latency Communications (URLLC), and Massive Machine-Type Communications (mMTC).

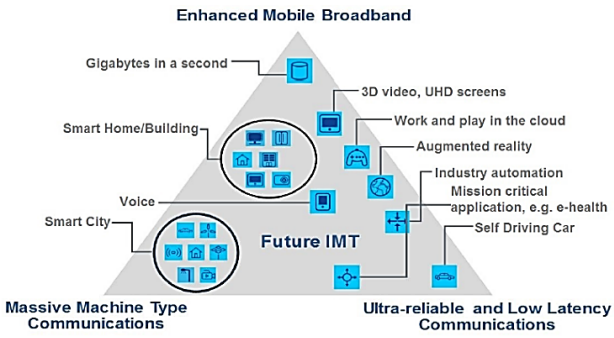


Figure 1. 5G-NR system use cases according to IMT-2020 [6].

The 3GPP Release 18 specification defines target peak data rates exceeding 10 gigabits per second (Gb/s) per cell for the Enhanced Mobile Broadband (eMBB) use case [7]. While eMBB primarily targets high data rates, it also aims to deliver a low-latency user experience, with an objective of 1 ms latency for users in motion. To achieve these ambitious targets, 5G systems leverage a suite of advanced technologies, including mmWave spectrum, small cell densification, Massive MIMO, and beamforming [8, 9].

Operating at higher frequencies than their 4G counterparts, 5G networks face inherent challenges related to signal propagation, such increased path loss and susceptibility to blockage. These challenges are mitigated through the application of sophisticated antenna technologies and advanced signal processing techniques to ensure robust and reliable network performance [10].

In the context of 5G, mmWave generally refers to the frequency range between 24 GHz and 100 GHz. This spectrum offers immense bandwidth potential, enabling exceptionally high data rates, particularly when combined with modern channel coding schemes.

5G deployments utilize a heterogeneous spectrum strategy, operating across licensed, unlicensed, and shared bands within two primary frequency ranges. Frequency Range 1 (FR1 or "sub-6 GHz") spans from 410 MHz to 7.125 GHz, encompassing low- and mid-band frequencies. Frequency Range 2 (FR2) spans from 24.25 GHz to 52.6 GHz, occupying the lower portion of the mmWave spectrum. While "mmWave" is a broader term often describing frequencies from approximately 24 GHz to 100 GHz and beyond, FR2 represents the specific, standardized 5G allocation within this range, as illustrated in Figure 2 [11].

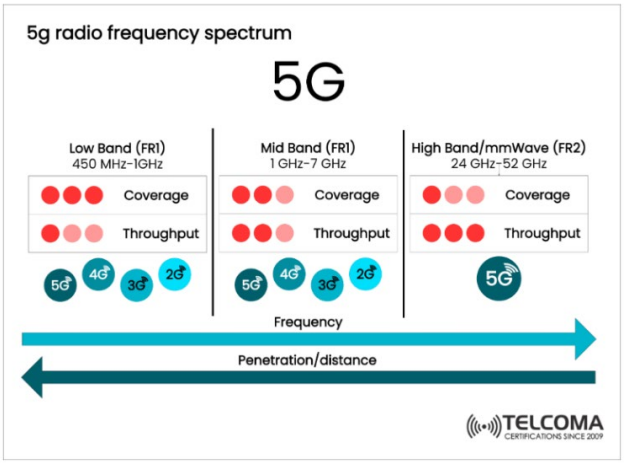


Figure 2. 5G radio frequency spectrum [11].

IV. CDL Channel Model

The Clustered Delay Line (CDL) model, standardized by the 3GPP, serves as a fundamental reference channel model for designing and evaluating 5G communication systems [12].

It is a statistical model well-suited for characterizing environments with clustered multipath propagation, where multiple signal paths arrive at the receiver in groups (clusters) with similar delay, angle, and power characteristics [13].

This makes it particularly applicable for system design in the mmWave frequency bands and for technologies like Massive MIMO. Defined for the frequency range from 0.5 GHz to 100 GHz, the CDL models are broadly categorized into two primary propagation conditions: Non-Line-of-Sight (NLOS) and Line-of-Sight (LOS). Each category is further divided into five subtypes, denoted as CDL-A, CDL-B, CDL-C, CDL-D, and CDL-E, which offer varying degrees of delay spread and K-factor to model different channel realism and severity [12, 14].

The CDL-E channel model is standardized for pure Line-of-Sight (LOS) propagation environments, representing an ideal strong direct path with minimal dependence on scattered multipath components. It is characterized by the smallest delay spread among all CDL models, reflecting very low latency variation and dominant Rician fading behavior.

This profile models a realistic LOS-dominant condition, featuring one extremely strong primary path (the direct ray) followed by very weak, short-delay multipath echoes. It is well-suited for simulating communication links in Rural Macrocell (RMa), highway, and Macrocell (UMa) scenarios, where the transmitter and receiver have a clear, unobstructed view (LOS) [12, 15]. The key parameters defining the CDL-E model are summarized in Table 1.

Table I. Parameters of the CDL-C model.[12]

Cluster #	Cluster PAS	Normalized Delay	Power in [dB]	AOD in [°]	AOA in [°]	ZOD in [°]	ZOA in [°]
1	Specular (LOS path)	0.000	-0.03	0	-180	99.6	80.4
	Laplacian	0.000	-22.03	0	-180	99.6	80.4
2	Laplacian	0.5133	-15.8	57.5	18.2	104.2	80.4
3	Laplacian	0.5440	-18.1	57.5	18.2	104.2	80.4
4	Laplacian	0.5630	-19.8	57.5	18.2	104.2	80.4
5	Laplacian	0.5440	-22.9	-20.1	101.8	99.4	80.8
6	Laplacian	0.7112	-22.4	16.2	112.9	100.8	86.3
7	Laplacian	1.9092	-18.6	9.3	-155.5	98.8	82.7
8	Laplacian	1.9293	-20.8	9.3	-155.5	98.8	82.7
9	Laplacian	1.9589	-22.6	9.3	-155.5	98.8	82.7
10	Laplacian	2.6426	-22.3	19	-143.3	100.8	82.9
11	Laplacian	3.7136	-25.6	32.7	-94.7	96.4	88
12	Laplacian	5.4524	-20.2	0.5	147	98.9	81
13	Laplacian	12.0034	-29.8	55.9	-36.2	95.6	88.6
14	Laplacian	20.6419	-29.2	57.6	-26	104.6	78.3

Per-Cluster Parameters					
Parameter	c_{ASD} in [°]	c_{ASA} in [°]	c_{ZSD} in [°]	c_{ZSA} in [°]	XPR in [dB]
Value	5	11	3	7	8

Furthermore, the CDL propagation model is fundamentally structured around discrete clusters. Each cluster comprises multiple rays that share a common propagation delay but exhibit distinct spatial characteristics, specifically through variations in their Azimuth Angle of Departure (AoD) and Azimuth Angle of Arrival (AoA) [12, 16].

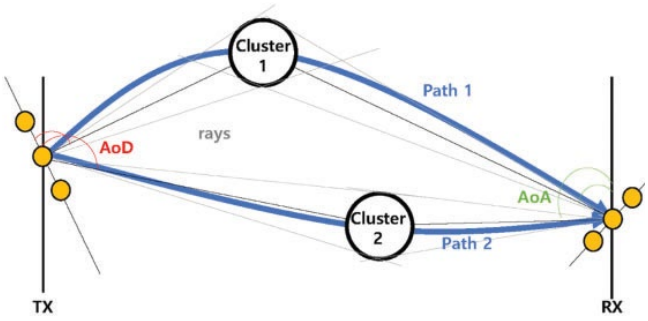


Figure 3. Concept of the CDL model [16].

V. MIMO Technology

Multiple-Input Multiple-Output (MIMO) technology enhances wireless communication by employing multiple antennas at both the transmitter and receiver. A cornerstone of modern 4G and 5G networks, MIMO enables the simultaneous transmission and reception of multiple data streams. This spatial multiplexing capability significantly increases spectral efficiency, leading to higher data throughput and improved link reliability without requiring additional bandwidth. Consequently, MIMO is a highly flexible technology for augmenting network capacity and peak data rates.

Conventional MIMO implementations typically involve a moderate number of antennas (e.g., fewer than 10) at the base station (BS) and a limited number (e.g., two or four) at the user equipment (UE), constrained by factors like physical size and hardware complexity [17, 18].

A simplified block diagram of a MIMO system is presented in Figure 4. Massive MIMO, an advanced evolution of MIMO, scales this concept by utilizing very large antenna arrays (often comprising dozens to hundreds of elements) at the base station. By leveraging this extensive spatial dimension, Massive MIMO can serve multiple users concurrently with highly focused signal beams, dramatically improving spectral efficiency, energy efficiency, and overall network capacity [18, 19].

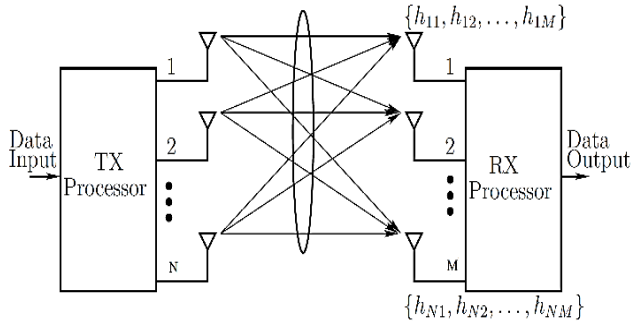


Figure 4. MIMO system block diagram. [20]

The evolution from conventional MIMO to Massive MIMO has been propelled by the escalating demand for higher data rates and the exponential growth in the number of connected devices [21].

In MIMO systems, data transmission is primarily enabled through two core techniques: spatial diversity and spatial multiplexing. Spatial diversity enhances link reliability by transmitting replicas of the same data stream across multiple independent paths. The receiver combines these signals to mitigate fading and improve the probability of correct detection.

Conversely, spatial multiplexing increases the peak data rate by splitting the data into multiple independent substreams transmitted simultaneously over different spatial channels. While this maximizes spectral efficiency, it can compromise link robustness compared to diversity schemes, creating a fundamental trade-off [23]. The general input-output relationship for a narrowband MIMO system is expressed by Equation (1) [23].

$$Y = H.X + N \ldots (1)$$

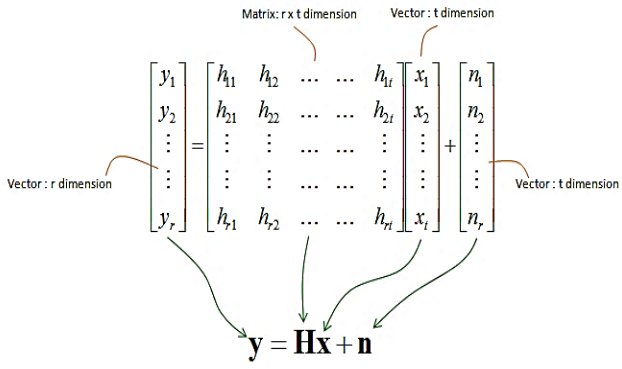


Figure 5. General MIMO equation. [20]

Where: Y is an N × 1 or (r × 1) received signal vector, H is an N × M or (r × t) channel matrix, X is an M × 1 or (t × 1) transmitted signal vector, and N represents the additive white Gaussian noise (AWGN) and it is (r × 1). This is based on the two Figures 4 and 5.

VI. Key Performance Indicators

Given the focus on the enhanced Mobile Broadband (eMBB) use case and data transmission in the mmWave frequency band, suitable key performance metrics were selected for this analysis:

1. Bit Rate:

Bit rate equation of the 5G system shown in equation (2) [24]

$$\text{Data Rate (Mbps)} = 10^{-6} \sum_{j=1}^{J} \left(v_{Layers}^{(j)} \cdot Q_m^{(j)} \cdot f^{(j)} \cdot R_{max} \cdot \frac{N_{PRB}^{BW(j),u} \cdot 12}{T_s^u} \cdot \left(1 - OH^{(j)}\right) \right) \ldots (2)$$

where J is the number of aggregated component carriers in a band or band combination; Rmax=948/1024; $v_{layers}^{(j)}$ is the maximum number of layers; $Q_m^{(j)}$ is the maximum modulation order and takes the following values (2 for QPSK, 4 for 16-QAM, 6 for 64-QAM, 8 for 256-QAM); $f^{(j)}$ is the scaling factor, the scaling factor can take the values 1, 0.8, 0.75, and 0.4. μ is the numerology (as defined in 3GPP TS 38.211) and can takes values from 0 to 5. T_S^{μ} is the average OFDM symbol duration in a subframe for numerology. $N_{PRB}^{BW(j),\mu}$ is the maximum RB allocation in bandwidth. $BW^{(j)}$ with numerology μ where $BW^{(j)}$ is the UE supported maximum bandwidth. $OH^{(j)}$ is the overhead and takes the following values: FR1 frequency range: DL: 0.14; UL: 0.08 and FR2 frequency range: DL: 0.18; UL: 0.1

2. Throughput:

Throughput can be calculated using equation (3) [22].

$$Throughput\ \left(\frac{b}{s}\right) = BW\ (Hz) \times SE\left(\frac{\frac{b}{s}}{Hz}\right) \quad (3)$$

Where BW is the bandwidth, and SE is the spectral efficiency. Or equation (2) after incorporating the effects of the Radio channel (CDL-E model here). Therefore, enhancing the achievable data rate necessitates an increase in either the channel bandwidth or the spectral efficiency. Given that the radio spectrum is a finite and often congested resource, significant bandwidth expansion is seldom feasible. Consequently, the primary focus for performance improvement shifts to maximizing spectral efficiency.

A canonical method for achieving this is the deployment of Multiple-Input Multiple-Output (MIMO) systems, which utilize multiple antennas at both link ends. MIMO technology enhances spectral efficiency by exploiting spatial diversity and multiplexing, thereby enabling high-speed data transmission even under challenging channel conditions [22].

3. Spectral efficiency:

The spectral efficiency relationship is given by equation (4) [24,25].

$$SE_{5G}\left(\frac{bps}{Hz}\right) = \frac{5G\ Throughput\ or\ R\ (bps)}{Channel\ BW(Hz)} \quad \ldots \quad (4)$$

Where, R represents the bit rate, and BW represents the bandwidth. It should be noted that bandwidth represents the maximum bandwidth supported by the User Equipment (UE) in a given frequency band or multiple bands. The Resource Elements (REs) are grouped into Physical Resource Blocks (PRBs), where each PRB consists of 12 subcarriers with a specific Subcarrier Spacing (SCS). Therefore, the actual bandwidth—excluding guard bands—can be calculated according to the new 5G radio standard using equation (5) [26].

$$BW(MHz) = NPRB \times SCS \times 12 \times 10^{-3} \quad \ldots \quad (5)$$

VII. Improving Network Throughput

Enhancing throughput in wireless communication systems presents a significant engineering challenge. The network throughput, often defined as the total data rate successfully delivered over a given coverage area, can be modeled by the general expression in Equation (6) [27].

$$Area\ throughput\left(\frac{\frac{b}{s}}{km^2}\right) = BW(Hz) \times CD\left(\frac{cells}{km^2}\right) \times SE\left(\frac{\frac{b}{s}/Hz}{Cell}\right) \quad \ldots \quad (6)$$

where BW is the allocated channel bandwidth, CD represents the network cell density (e.g., base stations per km²), and SE is the average spectral efficiency (in bits/s/Hz). A fundamental physical-layer challenge is to reliably and

uniformly enhance the aggregate wireless throughput across a target coverage area [27]. As indicated by Equation (5) and established in prior studies [28, 29], achieving higher throughput relies on three principal levers:

1) **Bandwidth Expansion:** Allocating additional spectrum to 5G services.
2) **Network Densification:** Condensing the network topology by deploying more cells and access points.
3) **Spectral Efficiency Enhancement:** Utilizing technologies like Multiple-Input Multiple-Output (MIMO) to improve the data transmission efficiency per cell within a fixed bandwidth.

Accordingly, the objective of this research is to evaluate the end-to-end performance of 5G mmWave communication systems under varied scenarios and parameter configurations. The analysis is based on three key performance metrics—spectral efficiency, transmitted bit rate, and achievable throughput—for a given channel bandwidth.

VIII. Implementation and Results

This paper investigates the application of multi-layer transmission (spatial multiplexing) in 5G mmWave communication systems operating at 40 GHz, corresponding to the 3GPP FR2 band n260. The primary objective is to demonstrate a significant enhancement in system performance—specifically in data transmission quality—by improving key performance indicators (KPIs) for the enhanced Mobile Broadband (eMBB) use case, such as spectral efficiency and throughput, under realistic Line-of-Sight (LOS) propagation conditions modeled using the 3GPP CDL-E channel model.

The proposed approach targets performance gains at both the transmitter, through layered data streams, and the receiver, leveraging Multiple-Input Multiple-Output (MIMO) signal processing with practically realizable antenna configurations for user equipment. The analysis focuses on the downlink (DL) transmission scenario.

The primary contribution of this work is twofold. First, it demonstrates the efficacy of the multi-layer transmission (spatial multiplexing) concept in significantly improving the achievable bit rate at the transmitter (Tx). We quantify this performance gain and analyze its scalability with an increasing number of layers.

Second, the study reveals a critical limitation: these transmitter-side gains do not directly translate to proportional improvements at the receiver (Rx) without additional signal processing support. To address this, we employ Multiple-Input Multiple-Output (MIMO) precoding techniques using Singular Value Decomposition (SVD) to recover the spatial streams and mitigate inter-layer interference. The use of multiple codewords, as defined in 3GPP specifications, provides greater flexibility in controlling the Modulation and Coding Scheme (MCS) for each codeword independently.

The system was implemented, and simulated using MATLAB R2024a. Figure 6 illustrates the block diagram of the implemented 5G NR downlink transmission system.

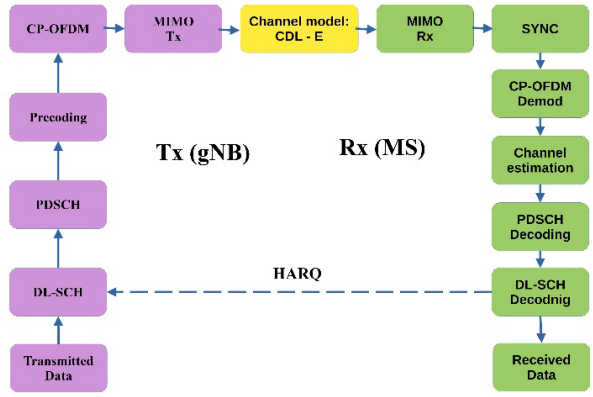


Figure 6. The block diagram of the implemented system.

The architecture supports multi-layer spatial multiplexing with independent control over the number of transmission layers and the antenna array dimensions at both the transmitter (gNB) and receiver (UE). The system employs Singular Value Decomposition (SVD)-based precoding at the transmitter to separate the spatial streams, which effectively performs digital beamforming, and MIMO processing at the receiver to recover the transmitted data. A key feature is the use of multiple codewords, each associated with an independent Modulation and Coding Scheme (MCS), providing flexible link adaptation.

The Demodulation Reference Signal (DMRS) symbols are allocated within each time slot according to 3GPP specifications. which according to the 3GPP TS 138 211 specification, the following conditions must be satisfied: the number of transmit and receive antennas must be at least equal to the number of spatial layers $(N_{TX}, N_{RX}) \geq N_{\text{layers}}$, and the DMRS length is configured as a function of the layer count: DMRS length = 1 for $N_{\text{layers}} \leq 4$, and DMRS length = 2 for $4 < N_{\text{layers}} \leq 8$.

In the results section, we present the effect of varying the number of layers and the number of antennas on the data transmission performance of the system. The primary simulation parameters are listed in Table II. For the scope of this analysis, we apply a uniform MCS across all layer.

Table II. Simulation parameters

Parameter	Value
Number of frames	1
SNR (dB)	-2 to 8
PRBs	135
SCS (kHz)	60
CP	Normal
Mapping Type	A
1st Modulation Order	16QAM
1st Code Rate	0.5
2nd Modulation Order	16QAM
2nd Code Rate	0.5
MIMO (Tx)	16, 32
MIMO (Rx)	4, 8
Channel model (Delay Profile)	CDL-E
Delay Spread (μsec)	300
Frequency (GHz)	40
Num of HARQ Processes	16
Waveform Type	CP-OFDM
Channel BW (MHz)	100

Three distinct MIMO system configurations, denoted by their antenna dimensions $N_t \times N_r$, are evaluated in this study. These scenarios are designed to isolate the impact of spatial multiplexing and antenna scaling. The configurations are as follows:

Scenario 1 (Baseline): A single codeword is transmitted using a 16 × 4 MIMO configuration. The number of spatial layers is varied from 1 to 4, corresponding to the maximum rank supported by this antenna configuration.

Scenario 2 (Multi-Layer with Fixed Antennas): Two independent codewords are transmitted concurrently using the same 16 × 8 MIMO configuration as Scenario 1. The number of spatial layers is varied from 5 to 8.

Scenario 3 (Multi-Layer with Antenna Scaling): Two independent codewords are transmitted using an expanded MIMO configuration of 32 × 8 antennas. The number of spatial layers is again varied from 5 to 8.

This progression allows for a comparative analysis of performance gains attributable to spatial multiplexing (Scenario 2) versus the combined effect of multiplexing and increased spatial degrees of freedom (Scenario 3).

The choice of 4 and 8 receive antennas at the UE reflects practical hardware constraints in commercial mobile devices, where physical size, power consumption, and thermal limitations restrict the number of antenna elements that can be integrated. These configurations are sufficient to support up to 4 spatial layers for the single-codeword scenario and up to 8 spatial layers for the dual-codeword scenario, respectively, which represent the maximum numbers considered in this study.

Case 1: 16 × 4 MIMO with a Single Codeword

The throughput and spectral efficiency (SE) performance for a 16 × 4 MIMO configuration employing a single codeword (mapped to 1 to 4 spatial layers) are presented in Figures 7, 8, and 9, respectively.

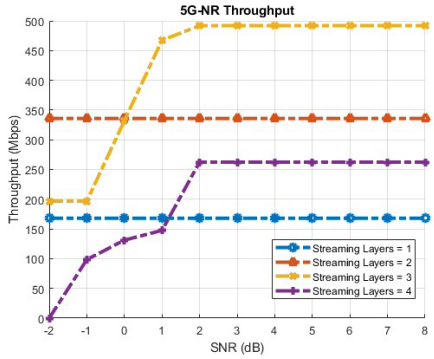


Figure 7. Throughput of the 5G-NR system using a single codeword with 16 × 4 MIMO.

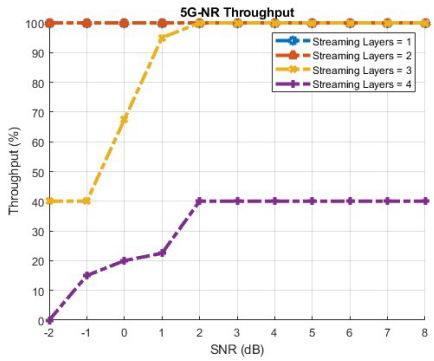


Figure 8. Normalized throughput (% of peak theoretical rate) of the 5G-NR system using a single codeword with 16 × 4 MIMO.

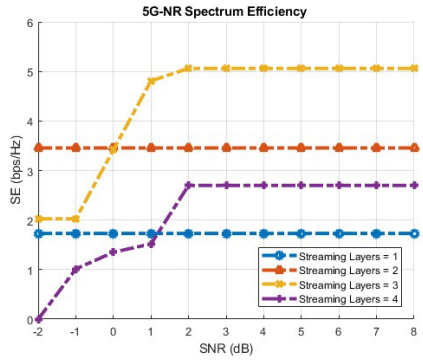


Figure 9. Spectral efficiency of the 5G-NR system using a single codeword with 16 × 4 MIMO.

Case 2: 16 × 8 MIMO with Dual Codewords

The corresponding throughput and spectral efficiency (SE) results for a 16 × 8 MIMO system utilizing two independent

codewords (mapped to 5 to 8 spatial layers) are presented in Figures 10, 11, and 12, respectively.

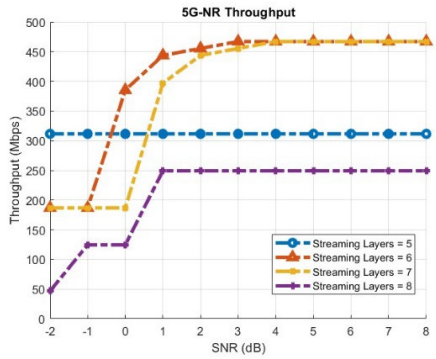


Figure 10. Throughput of the 5G-NR system using dual codewords with 16 × 8 MIMO.

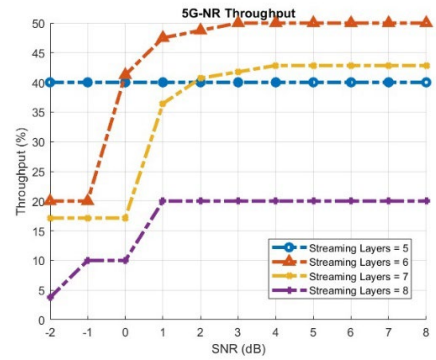


Figure 11. Normalized throughput (% of peak theoretical rate) of the 5G-NR system using dual codewords with 16 × 8 MIMO.

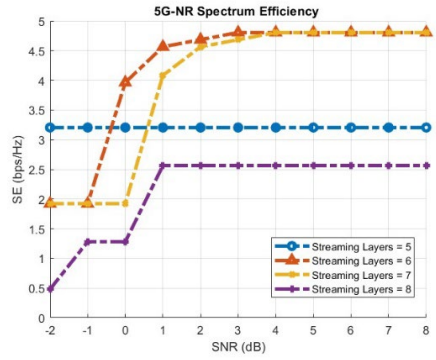


Figure 12. Spectral efficiency of the 5G-NR system using dual codewords with 16 × 8 MIMO.

The increase to 8 receive antennas is necessary to support up to 8 spatial layers, as the condition $N_{Rx} \geq N_{layers}$ must be satisfied. While this represents an optimistic upper bound for current UE hardware, it allows us to evaluate the maximum potential of spatial multiplexing in the dual-codeword scenario.

Case 3: 32 × 8 MIMO with Dual Codewords

Figures 13, 14, and 15 present the throughput and spectral efficiency (SE) for an expanded 32 × 8 MIMO configuration, maintaining the use of two independent codewords (mapped to 5 to 8 spatial layers).

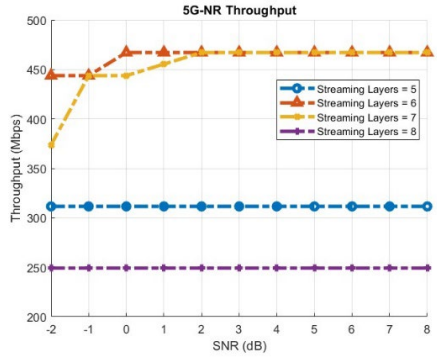


Figure 13. Throughput of the 5G-NR system using dual codewords with 32 × 8 MIMO.

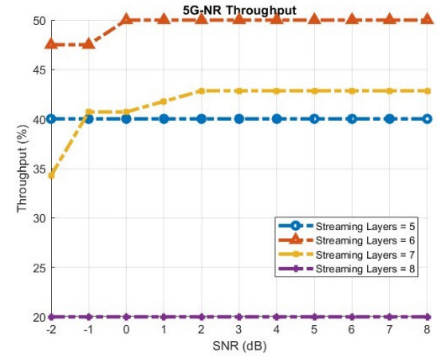


Figure 14. Normalized throughput (% of peak theoretical rate) of the 5G-NR system using dual codewords with 32 × 8 MIMO.

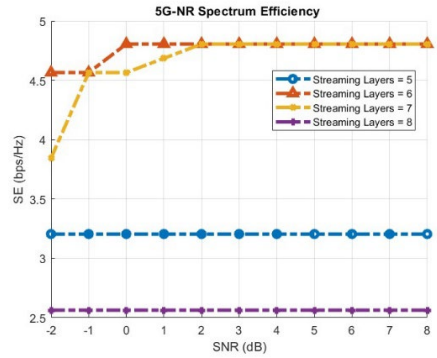


Figure 15. Spectral efficiency of the 5G-NR system using dual codewords with 32 × 8 MIMO.

Finally, Figures 16, 17, and 18 illustrate the variation of three key metrics—peak throughput, peak spectral efficiency, and user-experienced data rate (as a percentage of the maximum)—as a function of the number of spatial (streaming) layers.

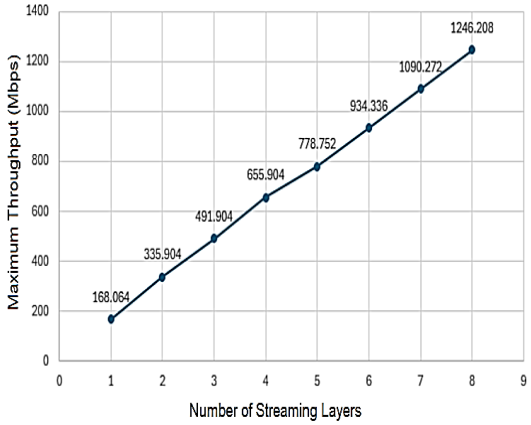


Figure 16. Bit Rate versus the number of spatial layers.

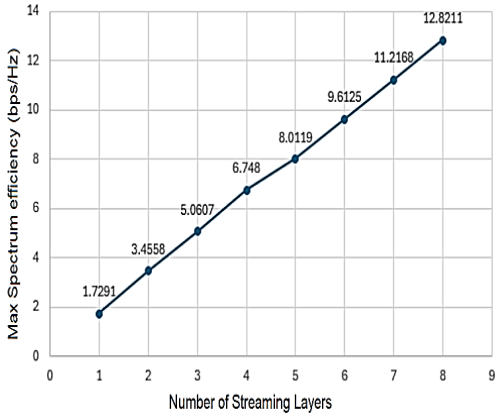


Figure 17. Maximum spectral efficiency versus the number of spatial layers.

IV. Discussion

This section provides a comprehensive analysis of the simulation results presented in Figures 7–15. The discussion is structured according to the three investigated scenarios, each defined by a specific MIMO configuration and range of spatial layers. All simulations were conducted at 40 GHz using the 3GPP CDL-E channel model, which represents a realistic Line-of-Sight (LOS) propagation environment with rich angular dispersion.

Scenario A (Tx = 16, Rx = 4, Layers 1–4), Figures 7, 8, and 9 illustrate the normalized throughput (%), absolute throughput (Mbps), and spectral efficiency (bps/Hz) for a single codeword transmitted over 1 to 4 spatial layers.

Low SNR regime (SNR < 0 dB): At very low Signal-to-Noise Ratio (SNR), thermal noise dominates the link. Consequently, configurations with fewer layers (e.g., 1 or 2 layers) achieve higher normalized throughput than those with 3 or 4 layers. This is because the channel lacks sufficient capacity to support reliable demultiplexing of multiple spatial streams under noisy conditions.

High SNR regime (SNR ≥ 2 dB): As SNR increases, the advantage of spatial multiplexing becomes evident. For example, at SNR = 2 dB, the 3-layer configuration reaches 100% normalized throughput, while the 4-layer configuration reaches 40%. In absolute terms, the 4-layer setup delivers ~260 Mbps compared to only ~170 Mbps for a single layer. This demonstrates that increasing the number of spatial layers significantly enhances spectral efficiency when channel conditions are favorable.

Practical insight: With 4 receive antennas at the User Equipment (UE), supporting up to 4 layers represents a realistic upper bound for commercial devices. The results confirm that this configuration can fully exploit the available spatial degrees of freedom at moderate to high SNR.

Scenario B (Tx = 16, Rx = 8, Layers 5–8), Figures 11, 12, and 13 show the performance when transmitting two independent codewords mapped to 5 to 8 spatial layers, while keeping the number of transmit antennas at 16 and increasing receive antennas to 8.

Limited effective rank: Despite increasing the number of receive antennas to 8, the system struggles to support more than 4 effective layers. At SNR = 8 dB, the normalized throughput for 5 layers is only 40%, dropping to 20% for 8 layers. This indicates that with only 16 transmit antennas and the CDL-E channel model, the effective rank of the channel matrix is limited to approximately 2.

Non-linear behavior (Layers 6 and 7 outperforming Layer 5): A notable observation in Figures 11 and 12 is that at moderate SNR values (e.g., 0–4 dB), the throughput for 6 and 7 layers exceeds that of 5 layers. This counterintuitive behavior arises from the Transport Block Size (TBS) quantization and Code Block Segmentation procedures defined in 3GPP NR. Specifically:

Layer 5 falls into an unfavorable quantization region, leading to a lower effective code rate and higher overhead.

Layers 6 and 7 achieve better alignment with the code block size limits, resulting in larger TBS and more efficient utilization of the available resources.

This finding underscores that performance is not a simple linear function of the number of layers; rather, it is influenced by higher-layer procedural effects.

Inter-layer interference: As the number of layers increases beyond the effective rank, inter-layer interference (ILI) becomes severe. The receiver, even with 8 antennas, cannot fully suppress this interference, leading to degraded throughput despite the higher theoretical peak rate.

Scenario C (Tx = 32, Rx = 8, Layers 5–8), Figures 13, 14, and 15 present the results for an expanded configuration with 32 transmit antennas at the gNB while maintaining 8 receive antennas at the UE. This scenario isolates the impact of increasing spatial degrees of freedom at the transmitter.

Improved array gain: Compared to Scenario B, the 32-transmit configuration exhibits a clear shift of the throughput curves toward lower SNR. For example, at SNR = 0 dB, the normalized throughput for 8 layers increases from 10% (Scenario B) to 20% (Scenario C). This improvement is attributed to the higher array gain provided by the larger antenna array, which enhances the received signal power without increasing transmit power.

Enhanced support for higher layers: Increasing the number of transmit antennas to 32 raises the statistical probability of a higher channel rank. At SNR = -2 dB, the normalized throughput for 8 layers reaches 20% in Scenario C, compared to only 4% in Scenario B. This demonstrates that massive MIMO at the base station is a key enabler for high-order spatial multiplexing in mmWave bands.

Remaining bottleneck: Even with 32 transmit antennas, the performance for 8 layers saturates at 20% (not 100%). This limitation is due to the receive antenna count (8) , which now becomes the limiting factor. With only 8 receive antennas, the system cannot fully resolve 8 independent spatial streams under realistic CDL-E channel conditions, which impose spatial correlation.

The comparison clearly shows that scaling the number of transmit antennas at the gNB significantly improves both the achievable throughput and the robustness against inter-layer interference, even when the number of receive antennas remains constant.

Figures 16 and 17 illustrate the peak theoretical bit rate and maximum spectral efficiency as functions of the number of spatial layers, respectively. Both metrics exhibit an approximately linear relationship with the number of layers, as expected from the fundamental MIMO capacity formula. For instance, increasing the number of layers from 1 to 8 raises the peak bit rate from 168 Mbps to 1246 Mbps, and the spectral efficiency from 1.73 bps/Hz to 12.82 bps/Hz.

However, these values represent an upper-bound benchmark achievable only under ideal channel conditions with no noise, no interference, and perfect receiver processing. In practice, as demonstrated in Figures A1–C3, the actual user throughput is significantly lower, particularly for higher layer counts and lower SNR values.

The gap between the peak theoretical rate and the achieved throughput highlights the impact of inter-layer interference, channel rank limitations under the CDL-E model, and the finite capabilities of the MIMO receiver.

The non-uniform incremental gains observed between successive layer counts (e.g., the marginal increase from 4 to 5 layers is 131 Mbps, while from 7 to 8 layers it is 219 Mbps) are attributed to the Transport Block Size (TBS) quantization and code block segmentation procedures inherent to the 3GPP NR standard.

These procedural effects introduce rounding and alignment constraints that cause the peak rate to deviate slightly from perfect linearity. Overall, Figures 16 and 17 confirm that spatial multiplexing is a powerful enabler for high-data-rate transmission in 5G mmWave systems, while also underscoring the need for advanced receiver processing and sufficient antenna resources to approach the theoretical limits in practical deployments.

X. Conclusion and Future Work

This paper investigated the performance of multi-layer spatial multiplexing in a 5G NR downlink system operating at 40 GHz (3GPP FR2 band n260) using the CDL-E channel model for realistic Line-of-Sight propagation. The study focused on the impact of the number of spatial layers and MIMO antenna configurations on key performance indicators, namely throughput and spectral efficiency, under practically feasible UE antenna constraints (4 and 8 receive antennas).

The experimental findings yield the following scientific insights: Non-linear relationship between layers and throughput: Increasing the number of spatial layers does not guarantee a proportional increase in user throughput.

Performance is governed by the effective channel rank, the TBS quantization effects inherent to 3GPP NR, and the balance between transmit and receive antennas. This was particularly evident in Scenario B, where layers 6 and 7 outperformed layer 5 due to more favorable code block segmentation.

Massive MIMO at the gNB is essential for mmWave: The transition from 16 to 32 transmit antennas (Scenario C) resulted in a substantial improvement in both array gain and spatial multiplexing capability. At SNR = -2 dB, the normalized throughput for 8 layers increased from 4% to 20%, confirming that massive MIMO at the base station is a critical enabler for high-order spatial multiplexing in the FR2 band.

UE antenna count remains a practical bottleneck: While increasing transmit antennas improves performance, the receive antenna count at the UE (4 or 8) ultimately limits the maximum achievable rank. Even with 32 transmit antennas, the 8-layer configuration could not reach 100% normalized throughput, indicating that further gains require more advanced UE antenna designs or collaborative MIMO schemes. Trade-off between layers and SNR: Systems with fewer spatial layers (e.g., 1–4 layers) achieve near-100% throughput efficiency at lower SNR values and require fewer receive antennas.

This makes them suitable for cell-edge users or devices with tight form-factor constraints. Conversely, higher-order multiplexing (5–8 layers) delivers superior peak throughput but demands higher SNR and more sophisticated interference

mitigation. Flexibility through multi-codeword transmission: The use of multiple independent codewords, each with its own Modulation and Coding Scheme (MCS), provides valuable adaptability.

This allows the system to optimize link performance based on real-time channel conditions, user requirements, and quality-of-service (QoS) targets. In summary, this work demonstrates that practical, commercially relevant UE antenna configurations (4–8 Rx) can support up to 8 spatial layers when combined with massive MIMO at the gNB (32 Tx) , provided that the channel conditions (LOS, CDL-E) are favorable.

The results also highlight the importance of considering higher-layer procedural effects (TBS quantization, code block segmentation) when evaluating MIMO performance, as these can lead to non-intuitive behaviors such as higher layers outperforming lower ones.

Future research directions include:

- Investigating hybrid beamforming architectures that better reflect practical mmWave implementations.
- Exploring higher receive antenna counts (e.g., 16 or 32) at the UE, possibly through collaborative or distributed MIMO approaches.
- Analyzing the impact of imperfect channel estimation and realistic HARQ processes on multi-layer performance.
- Extending the analysis to multi-user MIMO (MU-MIMO) scenarios with inter-user interference.

REFERENCES

[1] Ashraf, A., Gunawan, T. S., Kartiwi, M., Nur, L. O., Nugroho, B. S., & Astuti, R. P. (2024). Advancements and challenges in scalable modular antenna arrays for 5G massive MIMO networks. *IEEE Access*, 12, 57 895–57 916.

[2] Nkrumah, M. (2023). The Impact of 5G Technology on Communication Infrastructure. *Journal of Communication*, 4(1), 43–55. DOI: 10.47941/jcomm.1655

[3] Alrubaye, J. S., Abdkhaleq, M. H. G., & Alaidi, A. H. (2024). A Comprehensive Review of Routing in 4G/5G Cellular Networks: Challenges, Trends, and Future Directions. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, 16(4), 215–230. DOI: 10.29304/jqcsm.2024.16.41784

[4] Ashraf, S., Sheikh, J. A., Ashraf, A., & Rasool, U. (2024). 5G Millimeter Wave Technology: An Overview. In book: *Intelligent Signal Processing and RF Energy Harvesting for State of Art 5G and B5G Networks*. Springer, 97–112. DOI: 10.1007/978-981-99-8771-9_6

[5] Földes, G. (2023). Techno-Economic Analysis on Mobile Network Sharing Contribution to Social Welfare at 4G–5G Area in Hungary. *Infocommunications Journal*, 15(1), 87–97, DOI: 10.36244/icj.2023.1.9

[6] Shafi, M., et al., (2017). 5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice. *IEEE Journal on Selected Areas in Communications*, 35(6), 1201–1221. DOI: 10.1109/jsac.2017.2692307

[7] Makara, Á. L., & Csurgai-Horváth, L. (2021). Improved Model for Indoor Propagation Loss in the 5G FR2 Frequency Band. *Infocommunications Journa*l, 13(1), 2–10. DOI: 10.36244/icj.2021.1.1

[8] Zreikat, A. I., & Mathew, S. (2024). Performance evaluation and analysis of urban-suburban 5g cellular networks. *Computers*, 13(4), 108.

[9] Marinova, S., & Leon-Garcia, A. (2024). Intelligent O-RAN Beyond 5G: Architecture, Use cases, Challenges, and Opportunities. *IEEE Access*, 12, 27 088–27 114 DOI: 10.1109/access.2024.3367289.

[10] Islam, S., Abdulsalam, A. Z., Kumar, B. A., Hasan, M. K., Kolandaisamy, R., & Safie, N. (2024). Mobile Networks Toward 5G/6G: Network Architecture, Opportunities and Challenges in Smart City. *IEEE Open Journal of the Communications Society,* 6, 3082–3093. DOI: 10.1109/ojcoms.2024.3419791

[11] https://telcomaglobal.com/p/5g-frequency-bands

[12] ETSI. "TS 138 901 V17.0.0 (2022-04) – 5G; Study on channel model for frequencies from 0.5 to 100 GHz (3GPP TR 38.901 Release 17).

[13] Zhu, Q. et al. (2019). 3GPP TR 38.901 Channel Model. Chapter in Wiley 5G Ref: The Essential 5G Reference Online. *John Wiley & Sons*. DOI: 10.1002/9781119471509.w5gref048

[14] Zhang, Y., Sun, J., Gui, G., Gacanin, H., & Adachi, F. (2022). A Novel Channel Identification Architecture for mmWave Systems Based on Eigen Features. *IEEE 14th International Conference on Wireless Communications and Signal Processing (WCSP)*, 550–555. DOI: 10.1109/wcsp55476.2022.10039221

[15] Kumari, N., Rajput, R., & Sharma, S. (2023). CDL Channel Model: Revolutionizing Wireless Communication. *International Journal of Innovative Science and Research Technology,* 8(7), 1937–1949.

[16] Lee, K.H. (2022). LCF: A Deep Learning-Based Lightweight CSI Feedback Scheme for MIMO Networks. *Computers, Materials & Continua*, 71(3), 5561–5580. DOI: 10.32604/cmc.2022.024562

[17] Bhatti, U.S. (2025). Performance Enhancement of 5G Networks Remodeling Power Domain Scheme through NOMA-MIMO Technologies Integration. *International Journal of Advanced Engineering, Management and Science,* 11(4), 280–300. DOI: 10.22161/ijaems.114.28

[18] De Figueiredo, F. A. P. (2022). An overview of massive MIMO for 5G and 6G. *IEEE Latin America Transactions*, 20(6), 931–940.

[19] Jwair, M. H., & Elwi, T. A. (2023). Meta Surface Antenna Circuitry for 5G Communication Networks. *Infocommunications Journal,* 15(2), 2–7. DOI: 10.36244/icj.2023.2.1

[20] Halak, B., *et al*. (2018). Hardware Efficient Architecture for Element-Based Lattice Reduction Aided K-Best Detector for MIMO Systems. *Journal of Sensor and Actuator Networks*, 7(22), 1–16. DOI: 10.3390/jsan7020022

[21] Lu, L., Li, G. Y., Swindlehurst, A. L., Ashikhmin, A., & Zhang, R. (2014). An Overview of Massive MIMO: Benefits and Challenges. *IEEE Journal of Selected Topics in Signal Processing*, 8(5), 742–758. DOI: 10.1109/jstsp.2014.2317671

[22] Borges, D., Montezuma, P., Dinis, R., & Beko, M. (2021). Massive mimo techniques for 5g and beyond—opportunities and challenges. *Electronics*, 10(14), 1667.

[23] Baghous, J. (2021). 5G System Throughput Performance Evaluation using Massive-MIMO Technology with Cluster Delay Line Channel Model and Non-Line of Sight Scenarios. *Infocommunications Journal*, 13(2), 40–45. DOI: 10.36244/icj.2021.2.6

[24] 3GPP TS 38.306 v15.7.0, (2019-09). 5G-NR-User Equipment (UE) Radio Access Capabilities. (Release 15).

[25] Barri, E., Bouras, C., Kokkinos, V., & Koukouvela, A. (2021, November). A Mechanism for Improving the Spectral Efficiency in mu-MIMO for 5G and Beyond Networks. In *Proceedings of the 19th ACM International Symposium on Mobility Management and Wireless Access* (pp. 11–16).

[26] Furht, B. (2012). Encyclopedia of Wireless and Mobile Communications. *CRC Press*. DOI: 10.1201/noe1420043266

[27] 3GPP TS 38.101-1 version 15.3.0, (2018-10). 5G-NR-User Equipment (UE) Radio Transmission and Reception; Part 3: Range 1 and Range 2. (Release 15).

[28] Marzetta, T. L. (2015). Massive MIMO: an Introduction. *Bell Labs Technical Journal*, 20, 11–22. DOI: 10.15325/bltj.2015.2407793

[29] Xiang, W., Zheng, K., and Shen, X. (Eds). (2017). *5G Mobile Communications*. Springer. DOI: 10.1007/978-3-319-34208-5

John Baghous received his BE in Electronics and Communications Engineering (2017) from Damascus University (DU), Syria, and his Master in Advanced Communications Engineering (2021) from DU. His is currently a PhD student at the department of Electronics and Communications Engineering, Faculty of Mechanical and Electrical Engineering, DU, Syria. His research areas include cellular communications, wireless communication, wireless sensor networks, and D2D communications.

Mohamed Khaled Chahine received his BE in Electronics (1985) from INPG Grenoble France, Postgraduate MEng in Electronics and Optoelectronics (1990), and PhD in Satellite Communications (1994) from University of Pierre and Marie Curie Paris France. From 1994 to 2004, he was working at HIAST Damascus Syria as researcher and teacher. From 2004 to 2006 he was Director of COMSAT-COMSTECH-ITC Syria. Since 2006, he is an Associate Professor at the Department of Electronics and Communication Engineering at the Faculty of Mechanical and Electrical Engineering at Damascus University, Syria. His research areas include satellite communications, mobile communications, communication networks, optical communications, electronics engineering, wireless sensor networks, vehicular communications, software-defined networks, software-defined radios, quality control, ground penetrating radar, and audio and video streaming.

A Review of Security Challenges and Intrusion Detection Mechanisms to Mitigate Sub-Optimization Attacks in RPL-Based 6LoWPAN IoT Networks

Angel D[1]*, and Dr. Robin Rohit Vincent[2]

***Abstract*—The Internet of Things (IoT) has transformed device connectivity with the smooth interfacing for real-time data exchange across multiple applications, from smart homes to industrial automation. Nonetheless, as networks under IoT, especially those using the routing protocol for low-power and lossy networks (RPL), continue in their expansion, the security penetration becomes much more evident. One of the major security constraints is sub-optimization attacks-they negatively affect network performance, scalability, and data integrity. These attacks impede the very efficiency of the IoT systems, thereby making it so challenging for the systems to be secured and maintained successfully. Traditional IDS and cryptographic solutions are seldom fit-for-purpose in dynamic IoT environments, which opens up the need for the ability to provide scalable and energy-aware security solutions. This review investigates and surveys existing IDS, cryptographic solutions, and machine learning techniques targeting and working against such threats. It puts forth an integrated solution where an adaptive IDS is combined with scalable, energy-efficient, real-time anomaly detection to make IoT networks more resilient to sub-optimization attacks. According to this study, dynamic, context-aware safety measures are essential, as they are capable of addressing the new challenges arising from IoT environments.**

***Index Terms*—Internet of Things, Routing Protocol for Low-Power and Lossy Networks, Intrusion Detection Systems, sub-optimization attacks, and security mechanisms.**

I. INTRODUCTION

IoT has changed how systems and devices are capable of speaking to one another seamlessly, whether in an application for a smart home or healthcare facility, in industrial automation and transport sectors, while a lot of IoT networks are also expanding with connected devices in vast numbers to the extent of billions in exchange for exchanging large data in real time [1]. However, this growth raises major security risks, as weak protection of sensitive sensor data leads to breaches with economic and safety consequences [2].

[1]* PhD scholar, Presidency School of Computer Science and Engineering Presidency University, Bengaluru, Rajanukunte, Yelahanka, Bangalore North, Karnataka. (e-mail: ANGEL.20233CSC0005@presidencyuniversity.in, Angelveena.praveena@gmail.com)

[2] Professor and Head, Nvidia - CoE
Presidency School of Computer Science and Engineering, Presidency University, Bengaluru, Rajanukunte, Yelahanka, Bangalore North, Karnataka. (e-mail: robinrohit.vincent@presidencyuniversity.in, robinrohit@gmail.com)

DOI: 10.36244/ICJ.2026.1.10

Security is difficult due to limited resources, dynamic topologies, and diverse protocols [3]. IoT networks face threats such as data manipulation, DoS, and unauthorized access, endangering critical sectors like energy, healthcare, and transport. To address this, research focuses on cryptography, authentication, IDS, and AI-driven threat detection [4]. The review emphasizes scalability, adaptability, and resource issues in IoT security. Its aims are to:

- Examine threats in RPL-based 6LoWPAN, with emphasis on sub-optimization attacks.
- Investigate IDS frameworks and their weaknesses/strengths.
- Evaluate impacts of attacks on performance, efficiency, and scalability.
- Explore solutions such as dynamic/compressive IDS.
- Recognize research gaps in IDS optimization and scalability.
- Provide recommendations to enhance RPL-based 6LoWPAN security frameworks.

II. IOT AND RPL FUNDAMENTALS

In an IoT system (illustrated in Figure 1) four distinct components exist: sensors/devices, connectivity, data processing, and user interface. Sensors, which can be simple (e.g., temperature) or sophisticated (e.g., video camera), obtain real-time data from the environment. In the case of connectivity, the sensor should transmit the gathered data to the cloud using a transmission method, which includes Bluetooth, Wi-Fi, WAN, satellite, or mobile, and is retained in the cloud for further processing, and is the key for IoT implementation and is most often overlooked. Following this, the data is processed, which can be something simple like monitoring a temperature or sophisticated like computer vision or object detection. The end user then can interact with the IoT system using the user interface of notifications, alarms or live monitoring via a web server [5].

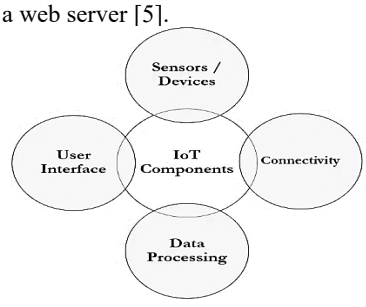


Figure 1: Components of an IoT System

A. IoT four-layer architecture

The four-layer IoT architecture consists of separate layers, with the application layer acting as a bridge between the network and devices, defining all IoT applications based on sensor data, facilitating user engagement and unique functions [6]. The perception layer sends data to the data processing layer, ensuring data safety and originates from real users. The network layer connects devices and facilitates data flow from sensors. The sensor layer detects and collects data from IoT devices, controlling their operating mechanisms for precise data gathering. These layers form a strong architecture for IoT systems, ensuring data safe and effective transmission from sensors to apps.

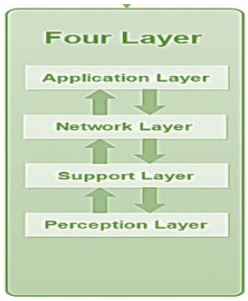


Figure 2: IoT Four-Layer Architecture

B. RPL - Network Layer Routing Protocol

RPL is a distance-vector routing protocol designed to accommodate a variety of data link layer protocols and to provide routing solutions for networks with restricted resources. RPL creates a Destination-Oriented Directed Acyclic Graph (DODAG) to ensure that each leaf node has a single path to the Root, which serves as the center for all traffic routing. The DODAG structure is established by nodes first promoting themselves as the Root by broadcasting DODAG Information Objects (DIOs), which spread throughout the network. To enable their parent nodes to make routing decisions for the destination, nodes send a Destination Advertisement Object (DAO) to them. Once the data transfer object (DTO) receives the DAO Acknowledgment (DAO-ACK), it is up to it to undertake the process to join the network. There are two modes of operation for RPL nodes: stateful and stateless. The most prevalent stateless nodes simply monitor their parent nodes, whereas the Root is fully aware of the DODAG. Stateful nodes, on the other hand, keep track of data about their parents and offspring, allowing for effective communication inside sub-trees without going via the Root. RPL is the perfect protocol for low-power and lossy situations because of its hierarchical structure and adaptable node operation.

RPL is a proactive distance-vector protocol developed by IETF for resource-constrained environments such as IoT systems running 6LoWPAN. RPL organizes nodes into a DODAG, which is rooted at a central node generally termed the gateway, acting as the destination for all traffic.

i. Key Concepts

- Rank: Indicates distance from root; exploitable for rank attacks.
- Version Number: Identifies DODAG; misuse can cause reconvergence and energy drain.
- Objective Function (OF): Guides rank/parent selection (e.g., OF0, MRHOF), affecting security and performance.

ii. RPL Control Messages

- DIO: Advertises DODAG parameters (version, rank, OF).
- DIS: Requests for DIOs for route discovery.
- DAO: Sent upstream to publish downward routes (storing mode).
- DAO-ACK: Acknowledges receipt of DAO and route advertisement.

iii. Modes of Operation

- Storing Mode: The intermediate nodes store routing tables.
- Non-Storing Mode: The root alone stores full routing information and uses source routing.

Even though optimal for low-power, lossy networks, RPL's management of rank, version numbers, and control messages provides weaknesses that facilitate sub-optimization attacks, reducing performance without stopping communication [7,8,9].

III. SUB-OPTIMIZATION ATTACKS ON RPL

Sub-optimization attacks refer to the routing layer attacks in RPL-based 6LoWPAN networks that use the weaknesses in design in the RPL protocol's objective functions and control mechanisms in order to deliberately degrade routing performance but not completely disrupt it. The attacks may involve manipulating metrics like rank values, suppressing DIO messages, triggering excessive DIS messages, or influencing parent selection so as to cause inefficient routing paths. Unlike conventional threats such as sinkhole or blackhole attacks that primarily capture or drop packets, sub-optimization attacks are usually covert and result, rather gradually, in higher latencies, energy burns, and congestion. Their subtlety makes it harder for an ordinary IDS to detect. Hence, emphasis must be put on security mechanisms that exploit deviations in performance rather than clear-cut malicious behaviors.

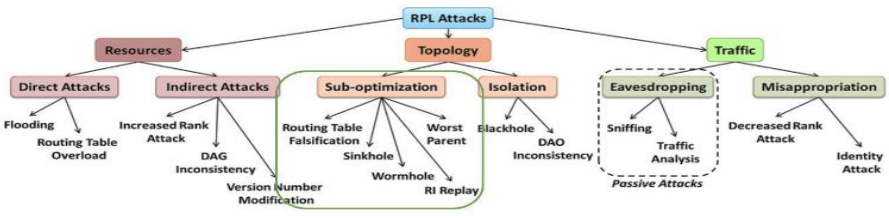


Figure 3: Classification of attacks in RPL Network

A. Routing Table Falsification

Several IoT IDS and secure routing approaches have been proposed. A hybrid SVM–Decision Tree IDS [10] achieved high accuracy with low false positives, while an ANN-based protocol [11] enhanced QoS. DETONAR [12] combined anomaly and signature detection without extra RPL overhead. A secure RPL protocol [13] mitigated DIS flooding with improved resilience, and [14] highlighted the negative impact of Hello Flood, Version Number, and Rank Reduction attacks on DODAG stability, efficiency, and scalability.

B. Sinkhole Attacks

Several solutions address sinkhole and selective forwarding in 6LoWPAN/RPL networks. A power-efficient IDS [15] achieved high detection accuracy with lightweight implementation. UVM [16] detects sinkholes via rank, power, and DIO metrics using equal voting. RFTrust [17] integrates Random Forest with trust models for reliable routing. CLS-RPL [18] uses cross-layer and overhearing mechanisms, while a multidirectional trust model [19] applies fuzzy and subjective logic with entropy-based trust weights to reduce false positives and delays.

TABLE I
COMPARISON OF SINKHOLE ATTACK DETECTION METHODS IN IOT NETWORKS

Paper	Objective	Techniques Used	Advantages	Disadvantages
[15]	Lightweight IDS for sinkhole attacks	Lightweight & Roving IDS	High performance, low energy	Limited to 6LoWPAN
[16]	Detect sinkhole nodes via behavior	UVM, Behavioral Indicators	Simple detection, effective features	Equal weight may misjudge importance
[17]	Trust-aware detection with ML	RF, SL, Trust Routing	Combines ML & trust for reliability	RF may be resource-heavy
[18]	Cross-layer sinkhole prevention	CLS-RPL, Overhearing, Secured-RPL	Stronger detection via layer integration	Overhearing adds energy cost
[19]	Multidirectional trust detection	FLS, SL, Trust Weight Adjustment	Low delay, fewer false positives	Higher complexity with adaptability

B. Wormhole Attacks

Wormhole attacks in RPL-based IoT networks create covert tunnels that disrupt routing, causing loss, delay, and integrity issues, and are harder to detect than sinkhole or blackhole attacks. Detection methods include trust models (e.g., SLF-RPL), ML frameworks using routing metrics, and traffic feature analysis. Recent solutions like MC-MLGBM [20], SLF-RPL [21], and hybrid IDSs [22] improve detection with low overhead, though wormholes remain less studied. Future work should emphasize hybrid ML–trust models for adaptive, lightweight, and scalable protection.

TABLE II
COMPARISON OF WORMHOLE ATTACK DETECTION METHODS IN RPL-BASED IOT NETWORKS

Paper	Objective	Techniques Used	Advantages	Disadvantages
[20]	To detect rank and wormhole attacks in RPL-based IoT networks	Machine Learning (MC-MLGBM Model)	Lightweight, multi-class classification, high accuracy	May require large datasets, computational overhead for training
[21]	To detect wormhole attacks using a trust model	Subjective Logic-based Trust Model (SLF-RPL)	Energy-efficient, adaptive to dynamic IoT environments	Potential delays in detecting malicious nodes, trust model sensitivity
[22]	Energy-efficient wormhole attack detection	Signal Strength, Hop Count-based Detection	Low energy overhead, suitable for constrained environments	May have detection delays in dynamic topologies

C. Other attacks

To limit IPv6 spoofing in 6LoWPAN, short-lived node addresses were suggested [23], minimizing disruption by updating addresses constantly.

An IDS based on PSO was designed [24], utilizing real-time Cooja IoT simulator information. PSO trained ML algorithms for enhanced accuracy in detecting routing threats.

A reinforcement learning (RL) agent [25] efficiently identified and prevented rank attacks in software-defined low-power IoT networks, with low latency, minimized duty cycles, and increased packet delivery ratios. An IDS model for Ping of Death attacks [26] utilized integer optimization to reduce false alarms and missed detections.

In the case of RPL routing attacks, SRPL-RP [27] resisted rank and version number tampering by ranking strategy comparison and blacklist table maintenance, supporting multiconfiguration and multiprotocol topologies.

Replay-based DoS attacks (copycat) [28] resulted in delivery degradation, energy consumption, and delay. CoSec-

RPL [30] employed Outlier Detection to counter non-spoofed copycat attacks better than baseline RPL.

To protect against DIS flooding, Secure-RPL [29] minimized control overhead and energy usage in both static and dynamic environments. In the case of version number flooding, [31] identified weaknesses in resource-constrained RPL, but only limited experiments were conducted.

TABLE III
COMPARISON OF IOT NETWORK SECURITY PAPERS

Paper	Objective	Techniques Used	Advantages	Disadvantages
[25]	Prevent rank attacks using SDN & RL	RL for route optimization, SDN for QoS	Cost-efficient, improved forwarding & latency	Higher complexity due to RL-SDN integration
[26]	Prevent Ping of Death via oversized packet filtering	Integer optimization, packet filtering	Low false alarms, reduced missed detections	Limited to DoS packet-size attacks
[27]	Detect & isolate rank/version number attacks in RPL	Secure RPL protocol	Supports multiple topologies	Focused only on rank/version threats
[28]	Study & mitigate copycat (replay) DoS attacks	Experimental analysis (delivery ratio, delay, power)	Detailed impact analysis of copycat attacks	Narrow focus, not generalizable
[29]	Mitigate DIS flooding in 6LoWPAN	Secure-RPL detection method	Reduces overhead & power drain	Attack-specific, not broadly applicable
[30]	Detect replay-based attacks with anomaly methods	Outlier Detection, CoSec-RPL IDS	Reduces replay attack effects	May miss varied replay/DoS attacks
[31]	Analyze flooding via version manipulation in RPL	Experimental performance study	Shows vulnerabilities in RPL versioning	Limited topologies, results not generalizable

IV. CURRENT SECURITY APPROACHES

A. Intrusion Detection Systems (IDS)

RPL, though widely adopted in IoT, remains vulnerable due to open environments and node constraints [32],[33]. Several IDS solutions have been proposed: a hybrid IDS with incremental ML for detecting DIO Suppression, Worst Parent, and Rank attacks [34]; an energy-aware cooperative IDS for host/edge devices [35]; and a heterogeneous IDS for dynamic 6LoWPAN environments [36]. Other works include KNN-based IPv6-IDS for WSNs [37], a Zigbee hybrid rule/ML IDS [38], and optimized RPL IDSs targeting routing and DoS attacks [39]. ASSET [40], a softwarized IDS, further detects 13 attack types with adaptive monitoring.

TABLE IV
COMPARISON OF IDS FOR IOT NETWORKS

Paper	Objective	Techniques Used	Advantages	Disadvantages
[36]	Adaptive IDS for RPL attacks in dynamic IoT.	Hybrid IDS, mobility- & attack-aware.	Works well in dynamic/heterogeneous settings.	High resource use, complex deployment.
[37]	Secure framework & IDS for IPv6 WSNs.	K-NN, profile-based detection.	Fast detection, easy integration.	Not scalable; K-NN is costly.
[38]	Hybrid IDS for Zigbee IoT.	ML anomaly detection + rule-based IDS.	Detects known & unknown attacks.	Rule creation hard; complex upkeep.
[39]	IDS for RPL IoT (routing & DoS attacks).	Multi-technique detection.	Detects many attacks (version, blackhole, grayhole, flooding).	Needs tuning for different networks.
[40]	Softwarized IDS (ASSET) for RPL IoT.	Multi-mechanism, configurable IDS.	Flexible, expandable, good trade-offs.	Complex setup; may add overhead.

B. Trust-based models

The rapid growth of IoT has amplified security risks, leading to increasing focus on trust-based IDS models for RPL networks. TIDSRPL [43] forwards node trust evaluations to the root, improving efficiency and outperforming MRHOF-RPL against Sybil, Sinkhole, and Selective Forwarding attacks. DSTIDS [44] strengthens sinkhole resistance and maintains QoS under attack through direct neighbor reputation. A behavioral trust model [45] mitigates version spoofing and hello flooding, analyzing both localized and global energy impacts. SMTrust [46] leverages mobility-aware trust to enhance resilience against rank and blackhole threats. SRF-IoT [47] integrates trust with an external IDS to isolate attackers in Contiki-NG. More advanced, a hybrid deep learning approach [48] combining RNN and stacked LSTM autoencoders predicts routing behaviors and mitigates blackhole, DIS flooding, rank, and version number attacks, showcasing the synergy between deep learning and trust mechanisms.

TABLE V
COMPARISON OF TRUST-BASED MODELS FOR RPL NETWORK SECURITY

Paper	Objective	Techniques Used	Advantages	Disadvantages
[43]	TIDSRPL: trust-based IDS to spot malicious nodes.	Trust eval, root-node offload; detects Sybil, Sinkhole, Selective Forwarding.	Efficient trust use; multi-attack detection.	Heavy root-node reliance, bottleneck risk.
[44]	DSTIDS to block sinkhole attacks in RPL IoT.	Trust-based, tested on Contiki 3.0 & Cooja.	Prevents sinkhole, improves QoS.	Limited to sinkhole only.
[45]	Trust-based RPL to curb energy fatigue (version/hello flooding).	Trust vs. version tampering & hello flooding.	Reduces energy drain, local mitigation.	Handles only two attacks, limited scope.
[46]	SMTrust for secure IoT routing, Rank & Blackhole prevention.	Security-Mobility-Trust model, trust metrics.	Better security, works for mobile/static nodes.	Needs parameter tuning for diverse nets.
[47]	SRF-IoT IDS for Rank & Blackhole detection.	Trust + IDS framework, external IDS.	Isolates malicious nodes well.	Extra IDS integration adds overhead.
[48]	Hybrid DL trust model for anomaly detection.	LSTM seq2seq autoencoder + trust.	Detects multiple RPL attacks via DL.	High complexity, tough for low-power IoT.

C. Anomaly-based detection techniques

In RPL networks, attackers can exploit the lack of parent monitoring by advertising falsely low rank values, drawing excessive traffic and enabling Rank Attacks such as sinkhole or selective forwarding (RA1) and topological instability (RA2) that degrade network performance [49]. To address these, SARPL applies statistical anomaly detection to identify and counter RA1 and RA2 effectively. Another approach, GAIDS [50], introduces an anomaly-based IDS leveraging stochastic games for attack detection and evolutionary games to confirm malicious intent. Since RPL's constraints may lead to false positives by misclassifying honest nodes, GAIDS employs an adaptive game-theoretic framework and clustered network design to improve verification accuracy, reliability, and robustness.

D. Cryptographic solutions

In [51], a lightweight Authentication and Key Exchange (AKE) framework was proposed for 6LoWPAN, using hashing and authenticated encryption to establish keys securely between sensor nodes and a server with low cost, avoiding IP security protocols. In [52], SRUA-IoT, a resource-efficient remote user authentication scheme, applied symmetric encryption, XOR, and hashing to generate secure session credentials, resisting multiple threats through formal and informal analysis. In [53], DSHRPL integrated encryption with node rating to secure RPL in 6LoWPAN through four stages: building trusted RPL, detecting sinkhole attacks, quarantining malicious nodes, and encrypting transmissions.

E. IDS using optimization techniques

In [54], a mobility management framework based on firefly algorithm (mRPL+firefly optimizer) was presented to improve RPL routing protocol in 6LoWPAN networks, realizing improvements in Packet Delivery Ratio (PDR), hop count, end-to-end latency, and energy consumption over available systems. To meet data dissemination overhead in IoT-enabled systems, [55] proposed Tabu RPL, which incorporates Tabu Search Routing (TSR), an adaptive routing that constantly optimizes data distribution according to network conditions and device capabilities, efficiently balancing routing options for enhanced efficiency. In the same vein, [56] introduced an equilibrium optimizer-based RPL (EO-RPL) protocol specifically designed for smart city structural health monitoring, where the EO algorithm optimizes parent node selection by weighing various routing metrics together, breaking the constraint of conventional RPL methods depending on single or composite rigid metrics.

F. Recent Studies on Sub-Optimization Attacks in RPL-Based IoT Networks (2025)

Rank attacks on RPL were first deployed in the Cooja Simulator in order to assess their impact. Subsequently, a new trust-based mitigation strategy was put forward [57] that catered to the requirements of IoT systems with limited resources. Meanwhile, an MLP trained on simulated data was put to use in a combined deep learning and machine learning framework [58] so as to improve the detection and classification of 10 types of RPL routing attacks. To effectively counter rank attacks, a lightweight ensemble IDS was introduced in [59], integrating SVM and XGBoost. This system, feature selection of which was carried out using RFE and Mutual Information, was evaluated under static and dynamic conditions. While the attacks in were addressed, attention in [60] was given to the version number attacks, conducting an in-depth study of the attacks and existing detection and prevention methods, and outlining the burning challenges in the research of RPL security. Edge-layer scrutiny of traffic patterns for the detection of Clone ID attacks was enhanced with the proposal of [61] through a DNN-based approach, bolstering the security, robustness, and efficiency of RPL IoT networks.

G. Comparison of Intrusion Detection Techniques

The earlier section presented various IDS and security mechanisms for RPL-based 6LoWPAN IoT networks;

however, for a practical application perspective, it is important to have a critical comparison. Table 6, given below, contrasts the prominent techniques according to detection accuracy, energy efficiency, scalability, and readiness for real-world deployment. Such an analysis facilitates identifying the solutions best suited for the context of an IoT setup with varied constraints.

TABLE VI
Comparative Analysis of IDS Techniques for RPL-Based 6LoWPAN IoT Networks

Technique / Model	Attack Type(s) Detected	Detection Accuracy	Energy Efficiency	Scalability	Real-World Readiness
Adaptive Hybrid IDS [36]	DIO Suppression, Worst Parent, Rank	High (90–98%)	Moderate	Moderate	Tested in dynamic data settings
RFTrust [17]	Sinkhole	High	Low–Moderate	High	Lightweight, trust-aware
DETONAR [12]	Routing Attacks (DIS, DAO, etc.)	Medium–High	High	Moderate	Packet-sniffing-based, tested
LSTM Autoencoder + RNN [48]	Blackhole, DIS Flooding, Rank, Version	Very High	Low	Low (high overhead)	High accuracy but resource-heavy
KNN-based IDS [37]	General routing anomalies	Moderate	Low	Low	Simple but not scalable
SM Trust Framework [46]	Rank, Blackhole (with node mobility)	High	Moderate	High	Suitable for mobile/static IoT
GAIDS (Game-Theoretic Anomaly IDS) [50]	Rank-based statistical anomalies	Moderate–High	Moderate	Moderate	Realistic with clustered topology
PSO-Optimized IDS [52]	Multiple attacks (ML-based)	High	Moderate	High	Efficient but setup-sensitive

The study reveals that intrusion detection techniques for RPL-based 6LoWPAN networks vary significantly in terms of accuracy, resource consumption, scalability, and practical applicability. Deep learning models, like LSTM autoencoders, have the best detection rates but are not suitable for resource-constrained IoT settings. Trust-based models like RFTrust and SMTrust offer a balanced trade-off, with high detection accuracy and good scalability. PSO-based IDSs compromise adaptability and efficiency, while KNN-based methods lack scalability and real-time applicability. The study emphasizes that no IDS solution is universally best, but its selection depends on deployment criteria.

H. Comparative Analysis of Detection Techniques

Detection techniques currently vary in terms of cost, scalability, and deployment readiness. A hybrid IDS model balances energy efficiency and scalability, ensuring reliable detection of diverse attack types. Deep learning-based architectures show better detection accuracy but have high computational costs. Rule-based and lightweight IDS approaches are practical but lack resilience against adaptive or sophisticated attack patterns. There is a need for an integrated framework that combines these approaches while keeping computational and energy costs low. The framework should adapt to IoT traffic drift and be scalable and fault-resilient for heterogeneous and large-scale environments.

I. Critical Evaluation of Existing Studies

Despite studies presenting effective measures against sub-optimization attacks in RPL-based 6LoWPAN networks, most do not address their limitations in real-world scenarios. Techniques like deep learning-based IDS have limitations in scalability and energy consumption, while lightweight alternatives can improve energy consumption but compromise detection rates. Trust-based and rule-based mechanisms offer balanced tradeoffs but are still limited in their ability to detect evolving or hybrid attacks. There is a lack of an honest comparative evaluation considering detection accuracy, energy efficiency, scalability, and deployability for RPL-based IoT networks.

V. TRENDS AND CHALLENGES

Extending This section explores emerging trends and challenges in IoT security for RPL-based 6LoWPAN networks. Key trends include the use of machine learning and artificial intelligence in IDS for real-time threat detection, decentralized security systems like trust-based and blockchain, and energy efficiency. Context-based security solutions are needed to support network morphisms due to network size and complexity. Privacy issues remain a challenge, requiring scalable, adaptive, and energy-efficient solutions to safeguard sensitive information while adhering to privacy regulations.

i) AI and machine learning for dynamic threat detection

ML/DL development is limited by the resources of the devices, but deployment at RPL, fog/edge, and cloud layers facilitates AI-based security. A hybrid IDS [63] addressed Flooding, Black Hole, DODAG Version Number, and Reduced Rank attacks employing ROUT-4-2023 and compared ML and DL through confusion matrices and processing time. To improve RPL security and QoS, [64] presented a mixed solution with varying OF, random forest classification, and RL-based adaptive routing against rank, sinkhole, and wormhole attacks. A hybrid DL IDS [65] integrating semi-supervised and supervised learning with IoTR-DS dataset was able to detect

DIS, Rank, and Wormhole attacks more accurately than current models. A GRU-based DL model [66] performed better than SVM and logistic regression in detecting HF attacks, enhancing efficiency, accuracy, and energy savings.

ii) Lightweight solutions for constrained devices

In [67], a distributed and collaborative RPL-based security mechanism (CDRPL) was proposed to detect and counteract Version Number (VN) attacks in order to efficiently identify and achieve fast topology convergence. To address flooding-based threats, [68] exhibited a Destination Information Object Flooding (DIOF) attack model in Cooja and suggested an easy-to-implement countermeasures. An energy-efficient ML-based trust-based IDS [69] for Rank, Sybil, and Wormhole attacks was proposed with dynamic trust estimation and low memory consumption. The Hatchet Man attack in RPL was investigated in [70], where low-complexity game-theoretic defense mechanism mitigated denial-of-service effects in 6LoWPAN IoT networks.

VI. KEY CHALLENGES

The key challenges of the review are listed as,

- **Resource Constraints:** Limited power, memory, and computing capacity demand lightweight solutions.
- **Dynamic & Scalable Networks:** Constantly changing and expanding devices make stable security difficult.
- **Energy-Efficient Security:** Battery limits require strong yet low-energy mechanisms.
- **False Positives & Latency:** Diverse traffic causes false alarms; real-time detection must balance resource limits.
- **Advanced Attacks:** Sinkhole, wormhole, and rank attacks exploit protocol flaws, needing multi-layered defenses.
- **Dataset Limitations:** Lack of large, realistic datasets hinders ML/AI-based IDS accuracy.
- **Interoperability:** Security must span heterogeneous IoT protocols (RPL, Zigbee, LoRaWAN, 6LoWPAN).
- **Privacy & Data Protection:** Safeguarding sensitive IoT data against misuse is challenging.
- **Concept Drift:** Changing traffic/attack patterns require adaptive models.
- **Centralized vs. Decentralized Trade-off:** Centralized models risk single-point failures; decentralized ones add overhead.
- **Real-time Adaptation:** Rapidly evolving threats reduce detection accuracy if systems cannot adapt.

VII. RESEARCH GAPS

Although there has been significant work toward resolving RPL-based 6LoWPAN security, numerous important research gaps remain [32]– [48]. IDS solutions today are not advanced or scalable enough to accommodate future threats like wormhole and rank manipulation in dynamic IoT networks. Lightweight methods that support security in conjunction with energy and computational constraints are still insufficient, while the majority of frameworks are based on centralized architectures with single points of failure and poor resilience. There is also the requirement for topology-aware and traffic-aware security systems that would evolve with topology and traffic changes, and privacy-aware methods in conformance with regulatory needs. Additionally, the incorporation of emerging technologies like machine learning and blockchain into IoT security solutions presents the potential to advance detection, data integrity, and scalability but remains untapped.

A significant challenge is the unavailability of realistic IoT attack datasets. Available datasets tend to be either synthetic or simulated, and this confines the strength of IDS models upon deployment in realistic settings. Much research has concentrated on classic attacks such as sinkhole and rank manipulation, with little investigation into hybrid, multi-layered, or adaptive threats. The majority of IDS solutions are static and do not learn adaptability against concept drift in IoT traffic, leading to greater false positives and lower accuracy. Scalability is also not addressed, as most methods are tested only on controlled or small networks and not on large, heterogeneous deployments.

Privacy is also an open gap, as most solutions focus on protecting communications and intrusion detection but fail to anonymize data and ensure secure handling of data. Most IoT security models also are not real-time adaptive and cannot automatically respond dynamically to changing threats without human action. Excessive use of single-layer defenses omits cross-layer vulnerabilities almost completely. To address these challenges, future work should emphasize scalable, lightweight, multi-layered, and adaptive security solutions that incorporate AI, blockchain, and context-awareness to provide strong and resilient IoT security.

VIII. PROPOSED ROADMAP FOR FUTURE RESEARCH

This article suggests a hybrid evaluation matrix as a method for systematically comparing intrusion detection and protection mechanisms across four key axes:

- Detection Performance – identifying different/novel attack types as they evolve while minimizing false positives.
- Resource Efficiency – ensuring low utilization of resources such as CPU or memory space suitable for constrained IoT devices.
- Scalability – potential for supporting large scale, dynamic, heterogeneous IoT networks depending on the size of the IoT ecosystem.
- Deployment Maturity – practical integration within the real world IoT ecosystem beyond simulation environments.

Furthermore, it suggests a multi-layer integrated solution integrating anomaly detection (e.g. ML/DL), trust management, and lightweight cryptography, where old adversaries are still relevant in the evolving threats space (i.e. concept drift) while maintaining energy consumption for resilient and feasible IoT solution designs. A single ecosystem that can also reduce the number of false alarms.

IX. CONCLUSION AND FUTURE DIRECTIONS

Though research in the field of IoT security has made progress in addressing vulnerabilities, there are still serious limitations. Many currently existing solutions are resource-heavy thus are not a feasible solution for low resource IoT devices. Moreover, most intrusion detection systems (IDS) are static, meaning their modifications are limited, which makes it unsuitable to address the dynamic and evolving nature of IoT

environments. The privacy concerns you can attack with a scalable prototype that has been tested in the real world are still serious issues. While there has been much done about countering known attacks, such as sinkhole, rank and wormhole attacks, there are sophisticated and emerging threats that still create challenges to IoT Security. To address these gaps, there is posting future research on efficient and adaptive frameworks, with promising avenues including AI/ML driven IDS capable of detecting and responding in real-time, blockchain integration for data integrity and decentralized trust, and post quantum cryptography can be utilized to mitigate the risks associated with quantum computing's threats to secure IoT. Another key area of interest is using decentralized security models, eliminating a single point of failure, to encourage resilience in a diverse IoT ecosystem. Furthermore, the future of IoT security ultimately resides in a state of development that is scalable, lightweight, and adaptive, capable of securing heterogeneous and large-scale networks against evolving and complex threats.

REFERENCES

[1] A. Seyfollahi, M. Mainuddin, T. Taami, and A. Ghaffari, "RM-RPL: reliable mobility management framework for RPL-based IoT systems,"*Clust. Comput.*, vol. 27, no. 4, pp. 4449–4468, 2024, DOI: 10.1007/s10586-023-04199-0.

[2] S. Gonen, "A methodical examination of single and multi-attacker flood attacks using RPL-based approaches," *Comput. Ind. Eng.*, vol. 194, p. 110 356, 2024, DOI: 10.1016/j.cie.2024.110356.

[3] S. Sahraoui and N. Henni, "SAMP-RPL: secure and adaptive multipath RPL for enhanced security and reliability in heterogeneous IoT-connected low power and lossy networks," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 1, pp. 409–429, 2023, DOI: 10.1007/s12652-021-03303-9.

[4] B. Isong, O. Kgote, and A. Abu-Mahfouz, "Insights into Modern Intrusion Detection Strategies for Internet of Things Ecosystems," *Electron. Switz.*, vol. 13, no. 12, 2024, DOI: 10.3390/electronics13122370.

[5] M. Kamal, I. Rashid, W. Iqbal, M. H. Siddiqui, S. Khan, and I. Ahmad, "Privacy and security federated reference architecture for Internet of Things," *Front. Inf. Technol. Electron. Eng.*, vol. 24, no. 4, pp. 481–508, 2023, DOI: 10.1631/FITEE.2200368.

[6] M. Ghaleb and F. Azzedin, "Towards Scalable and Efficient Architecture for Modeling Trust in IoT Environments," *Sensors*, vol. 21, Apr. 2021, DOI: 10.3390/s21092986.

[7] M. Asim, T. Baker, J. Nisar, and N. Tariq, "CTrust-RPL: A control layer-based trust mechanism for supporting secure routing in routing protocol for low power and lossy networks-based Internet of Things applications," *Trans. Emerg. Telecommun. Technol.*, vol. 32, Mar. 2021, DOI: 10.1002/ett.4224.

[8] A. Idrees and A. Witwit, "Energy-efficient Load-balanced RPL routing protocol for Internet of Things (IoTs) Networks," *Int. J. Internet Technol. Secur. Trans.*, vol. 11, pp. 286–306, Apr. 2021, DOI: 10.1504/IJITST.2020.10030144.

[9] T. Winter et al., RFC 6550: RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks. USA: RFC Editor, 2012.

[10] A. Alazab, A. Khraisat, S. Singh, S. Bevinakoppa, and O. Mahdi, "Routing Attacks Detection in 6LoWPAN-Based Internet of Things," *Electronics*, vol. 12, p. 1320, Mar. 2023, DOI: 10.3390/electronics12061320.

[11] J. Lu, D. Li, P. Wang, F. Zheng, and M. Wang, "Security-Aware Routing Protocol Based on Artificial Neural Network Algorithm and 6LoWPAN in the Internet of Things," *Wirel. Commun. Mob. Comput.*, vol. 2022, pp. 1–8, Jan. 2022, DOI: 10.1155/2022/8374473.

[12] A. Agiollo, M. Conti, P. Kaliyar, T.-N. Lin, and L. Pajola, "DETONAR: Detection of Routing Attacks in RPL-Based IoT," *IEEE Trans. Netw. Serv. Manag.*, vol. 18, no. 2, pp.1178–1190, 2021, DOI: 10.1109/TNSM.2021.3075496.

[13] E. V Abhinaya and B. Sudhakar, "A secure routing protocol for low power and lossy networks based 6LoWPAN networks to mitigate DIS flooding attacks," *J. Ambient Intell. Humaniz. Comput.*, 2021, DOI: 10.1007/s12652-020-02804-3.

[14] S. Sharma and V. K. Verma, "Security explorations for routing attacks in low power networks on internet of things," *J. Supercomput.*, vol. 77, no. 5, pp. 4778–4812, 2021, DOI: 10.1007/s11227-020-03471-z.

[15] P. Bhale, S. Dey, S. Biswas, and S. Nandi, "Energy Efficient Approach to Detect Sinkhole Attack Using Roving IDS in 6LoWPAN Network BT – Innovations for Community Services," S. S. Rautaray, G. Eichler, C. Erfurth, and G. Fahrnberger, Eds., Cham: Springer International Publishing, 2020, pp. 187–207.

[16] S. Al-Sarawi, M. Anbar, B. A. Alabsi, M. A. Aladaileh, and S. D. Ahmed Rihan, "Unweighted Voting Method to Detect Sinkhole Attack in RPL-Based Internet of Things Networks," *Comput. Mater. Contin.*, vol. 77, no. 1, pp. 491–515, 2023, DOI: 10.32604/cmc.2023.041108.

[17] K. Prathapchandran and T. Janani, "A trust aware security mechanism to detect sinkhole attack in RPL-based IoT environment using random forest – RFTRUST," *Comput. Netw.*, vol. 198, p. 108 413, 2021, DOI: 10.1016/j.comnet.2021.108413.

[18] A. Jamil, M. Ali, and M. Tharwat, "Sinkhole Attack Detection and Avoidance Mechanism for RPL in Wireless Sensor Networks," *Ann. Emerg. Technol. Comput.*, vol. 5, pp. 94–101, Mar. 2021, DOI: 10.33166/AETiC.2021.05.011.

[19] S. Khoeurt, C. So-In, P. Musikawan, and P. Aimtongkham, "Multidirectional Trust-Based Security Mechanisms for Sinkhole Attack Detection in the RPL Routing Protocol for Internet of Things," *J. Wirel. Mob. Netw.*, vol. 14, pp. 48–76, Sep. 2023, DOI: 10.58346/JOWUA.2023.I3.005.

[20] F. Zahra, N. Z. Jhanjhi, S. N. Brohi, N. A. Khan, M. Masud, and M. A. AlZain, "Rank and Wormhole Attack Detection Model for RPL-Based Internet of Things Using Machine Learning," *Sensors*, vol. 22, no. 18, 2022, DOI: 10.3390/s22186765.

[21] S. Javed *et al.*, "A Subjective Logical Framework-Based Trust Model for Wormhole Attack Detection and Mitigation in Low-Power and Lossy (RPL) IoT-Networks," *Information*, Aug. 2023, DOI: 10.3390/info14090478.

[22] S. A. Bhosale and S. S. Sonavane, "Wormhole Attack Detection System for IoT Network: A Hybrid Approach," *Wirel. Pers. Commun.*, vol. 124, no. 2, pp. 1081–1108, 2022, DOI: 10.1007/s11277-021-09395-y.

[23] M. Mavani and K. Asawa, "Resilient against spoofing in 6LoWPAN networks by temporary-private IPv6 addresses," *Peer--Peer Netw. Appl.*, vol. 13, no. 1, pp. 333–347, Jan. 2020, DOI: 10.1007/s12083-019-00792-6.

[24] M. Belaissaoui and Y. Maleh, "Machine Learning techniques optimized by Practical Swarm optimization for Intrusions Detection in IoT".

[25] C. Miranda, G. Kaddoum, A. Boukhtouta, T. Madi, and H. A. Alameddine, "Intrusion Prevention Scheme Against Rank Attacks for Software-Defined Low Power IoT Networks," *IEEE Access,* vol. 10, pp. 129 970–129 984, 2022, DOI: 10.1109/ACCESS.2022.3228170.

[26] A. Abdollahi and M. Fathi, "An Intrusion Detection System on Ping of Death Attacks in IoT Networks," *Wirel. Pers. Commun.*, vol. 112, no. 4, pp. 2057–2070, 2020, DOI: 10.1007/s11277-020-07139-y.

[27] Z. Almusaylim, N. Jhanjhi, and A. Alhumam, "Detection and Mitigation of RPL Rank and Version Number Attacks in the Internet of Things: SRPL-RP," *Sensors*, vol. 20, p. 5997, Oct. 2020, DOI: 10.3390/s20215997.

[28] A. Verma and V. Ranga, "The impact of copycat attack on RPL based 6LoWPAN networks in Internet of Things," *Computing*, vol. 103, no. 7, pp. 1479–1500, 2021, DOI: 10.1007/s00607-020-00862-1.

[29] A. Verma and V. Ranga, "Mitigation of DIS flooding attacks in RPL-based 6LoWPAN networks," *Trans. Emerg. Telecommun. Technol.*, pp. 1–25, Feb. 2020, https://doi.org/10.1002/ett.3802.

[30] A. Verma and V. Ranga, "CoSec-RPL: detection of copycat attacks in RPL based 6LoWPANs using outlier analysis," *Telecommun. Syst.*, vol. 75, no. 1, pp. 43–61, 2020, DOI: 10.1007/s11235-020-00674-w.

[31] M. Rouissat, B. Mohammed, and B. Sid Ahmed Hichame, "A potential flooding version number attack against RPL based IOT networks," *J. Electr. Eng.*, vol. 73, 2022, pp. 67–275, Sep. 2022, DOI: 10.2478/jee-2022-0035.

[32] G. Simoglou, G. Violettas, S. Petridou, and L. Mamatas, "Intrusion detection systems for RPL security: A comparative analysis," *Comput. Secur.*, vol. 104, p. 102 219, May 2021, DOI: 10.1016/j.cose.2021.102219.

[33] A. Verma and V. Ranga, "Security of RPL Based 6LoWPAN Networks in the Internet of Things: A Review," *IEEE Sens. J.*, vol. 20, no. 11, pp. 5666–5690, Jun. 2020, DOI: 10.1109/JSEN.2020.2973677.

[34] A. M. Pasikhan, J. A. Clark, and P. Gope, "Incremental hybrid intrusion detection for 6LoWPAN," *Comput. Secur.*, vol. 135, p. 103 447, 2023, DOI: 10.1016/j.cose.2023.103447.

[35] J. Arshad, M. A. Azad, M. M. Abdeltaif, and K. Salah, "An intrusion detection framework for energy constrained IoT devices," *Mech. Syst. Signal Process.*, vol. 136, p. 106 436, 2020, DOI: 10.1016/j.ymssp.2019.106436.

[36] A. Pasikhani, J. Clark, and P. Gope, Adaptive Hybrid Heterogeneous IDS for 6LoWPAN. 2022. DOI: 10.48550/arXiv.2205.09170.

[37] M. Wei, C. Rong, E. Liang, and Y. Zhuang, "An intrusion detection mechanism for IPv6-based wireless sensor networks," *Int. J. Distrib. Sens. Netw.*, vol. 18, p. 155013292210779, Mar. 2022, DOI: 10.1177/15501329221077922.

[38] F. Sadikin, T. van Deursen, and S. Kumar, "A ZigBee Intrusion Detection System for IoT using Secure and Efficient Data Collection," *Internet Things*, vol. 12, p. 100 306, 2020, DOI: 10.1016/j.iot.2020.100306.

[39] E. Garcia Ribera, B. Martinez Alvarez, C. Samuel, P. P. Ioulianou, and V. G. Vassilakis, "An Intrusion Detection System for RPL-Based IoT Networks," *Electron. Switz.*, vol. 11, no. 23, 2022, DOI: 10.3390/electronics11234041.

[40] G. Violettas, G. Simoglou, S. Petridou, and L. Mamatas, "A Softwarized Intrusion Detection System for the RPL-based Internet of Things networks," *Future Gener. Comput. Syst.*, vol. 125, pp. 698–714, 2021, DOI: 10.1016/j.future.2021.07.013.

[41] K. Avila, D. Jabba, and J. Gomez, "Security Aspects for Rpl-Based Protocols: A Systematic Review in IoT," *Appl. Sci.*, vol. 10, no. 18, Art. no. 18, Jan. 2020, DOI: 10.3390/app10186472.

[42] A. Kamble, V. S. Malemath, and D. Patil, "Security attacks and secure routing protocols in RPL-based Internet of Things: Survey," in *2017 International Conference on Emerging Trends & Innovation in ICT (ICEI)*, Feb. 2017, pp. 33–39. DOI: 10.1109/ETIICT.2017.7977006.

[43] S. Remya, M. J. Pillai, C. Arjun, S. Ramasubbareddy, and Y. Cho, "Enhancing Security in LLNs Using a Hybrid Trust-Based Intrusion Detection System for RPL," *IEEE Access*, vol. 12, pp. 58 836–58 850, 2024, DOI: 10.1109/ACCESS.2024.3391918.

[44] B. Patel and P. Shah, "Direct Neighbour Sink Reputed Trust Based Intrusion Detection System to Mitigate Sinkhole Attack in RPL for IoT Networks," *J. Eng. Sci. Technol. Rev.*, vol. 14, pp. 35–38, Feb. 2021, DOI: 10.25103/jestr.141.03.

[45] F. Azzedin, "Mitigating Denial of Service Attacks in RPL-Based IoT Environments: Trust-Based Approach," *IEEE Access*, vol. PP, p. 1, Jan. 2023, DOI: 10.1109/ACCESS.2023.3331030.

[46] S. M. Muzammal, R. K. Murugesan, N. Jhanjhi, M. Humayun, A. Osman, and A. Abdelmaboud, "A Trust-Based Model for Secure Routing against RPL Attacks in Internet of Things," *Sensors*, vol. 22, p. 7052, Sep. 2022, DOI: 10.3390/s22187052.

[47] P. P. Ioulianou, V. G. Vassilakis, and S. F. Shahandashti, "A Trust-Based Intrusion Detection System for RPL Networks: Detecting a Combination of Rank and Blackhole Attacks," *J. Cybersecurity Priv.*, vol. 2, no. 1, pp. 124–153, 2022, DOI: 10.3390/jcp2010009.

[48] K. Ahmadi and R. Javidan, "A Trust Based Anomaly Detection Scheme Using a Hybrid Deep Learning Model for IoT Routing Attacks Mitigation," *IET Inf. Secur.*, vol. 2024, no. 1, p. 4 449 798, Jan. 2024, DOI: 10.1049/2024/4449798.

[49] M. A. Alqarni and S. H. Chauhdary, "A Security Scheme for Statistical Anomaly Detection and the Mitigation of Rank Attacks in RPL Networks (IoT Environment)," *Eng. Technol. Appl. Sci. Res.*, vol. 13, no. 6, pp. 12 409–12 414, 2023, DOI: 10.48084/etasr.6433.

[50] D. B. Gothawal and S. V Nagaraj, "Anomaly-Based Intrusion Detection System in RPL by Applying Stochastic and Evolutionary Game Models over IoT Environment," *Wirel. Pers. Commun.,* vol. 110, no. 3, pp. 1323–1344, 2020, DOI: 10.1007/s11277-019-06789-x.

[51] M. Tanveer, G. Abbas, Z. Abbas, M. Waqas, F. Muhammad, and S. Kim, "S6AE: Securing 6LoWPAN using Authenticated Encryption Scheme," *Sensors*, vol. 20, May 2020, DOI: 10.3390/s20092707.

[52] G. Abbas, M. Tanveer, Z. Abbas, M. Waqas, T. Baker, and D. OBE, "A secure remote user authentication scheme for 6LoWPAN-based Internet of Things," *PLOS ONE*, vol. 16, p. e0258279, Nov. 2021, DOI: 10.1371/journal.pone.0258279.

[53] M. Zaminkar, F. Sarkohaki, and R. Fotohi, "A method based on encryption and node rating for securing the RPL protocol communications in the IoT ecosystem," *Int. J. Commun. Syst.*, vol. 34, Nov. 2020, DOI: 10.1002/dac.4693.

[54] K. Manikannan and V. Nagarajan, "Optimized mobility management for RPL/6LoWPAN based IoT network architecture using the firefly algorithm," *Microprocess. Microsyst.*, vol. 77, p. 103 193, 2020, DOI: 10.1016/j.micpro.2020.103193.

[55] V. K. Prajapati, T. P. Sharma, and L. K. Awasthi, "Data Dissemination Framework for Optimizing Overhead in IoT-Enabled Systems Using Tabu-RPL," *SN Comput. Sci.*, vol. 5, no. 4, p. 343, 2024, DOI: 10.1007/s42979-024-02694-8.

[56] K. A. Darabkh, H. H. AlAdwan, M. Al-Akhras, F. Jubair, and S. Rahamneh, "A revolutionary RPL-based IoT routing protocol for monitoring building structural health in smart city domain utilizing equilibrium optimizer algorithm," *Soft Comput.*, vol. 28, no. 17, pp. 10 099–10 138, 2024, DOI: 10.1007/s00500-024-09677-0.

[57] M. Yadav and R. Kaur, "Implementation of Rank Attack and Its Mitigation in RPL-Based IoT Networks:," in *Proceedings of the 10th International Conference on Internet of Things, Big Data and Security, Porto, Portugal: SCITEPRESS - Science and Technology Publications*, 2025, pp. 215–222. DOI: 10.5220/0013205100003944.

[58] A. Krari, A. Hajami, A. Toubi, and M. A. Said, "Securing IoT Networks: Multi-Attack Detection of RPL Routing Threats Using Deep Learning," *J. Comput. Sci.*, vol. 21, no. 4, pp. 836–850, Mar. 2025, DOI: 10.3844/jcssp.2025.836.850.

[59] S. Kalyani and D. Vydeki, "A Resource-Efficient Ensemble Learning Framework for Detecting Rank Attacks in RPL-Based IoT Networks," *J. Econ. Technol.*, Jul. 2025, DOI: 10.1016/j.ject.2025.06.003.

[60] M. Boudouaia, V. Tournois, S. Ouchani, and A. Abuarqoub, "Version Number Attacks in RPL based IoT Networks State of the Art and Future directions," in *Proceedings of the 8th International Conference on Future Networks & Distributed Systems, in ICFNDS '24*. New York, NY, USA: Association for Computing Machinery, Jul. 2025, pp. 276–282. DOI: 10.1145/3726122.3726163.

[61] F. Al-Quayed, S. R. Awan, N. Tariq, M. Humayun, T. S. Alnusairi, and T. Rehman, "CID-RPL: Clone ID Attack Detection Using Deep Neural Network for RPL-Based IoT Networks," *IET Commun.*, vol. 19, no. 1, p. e70067, 2025, DOI: 10.1049/cmu2.70067.

[62] M. Belaissaoui and M. Yassine, "Machine Learning techniques optimized by Practical Swarm optimization for Intrusions Detection in IoT. In *Journal of Information Assurance and Security*. ISSN 1554-1010," vol. 16, pp. 105–116, Jul. 2021.

[63] U. Shahid, M. Z. Hussain, M. Z. Hasan, A. Haider, J. Ali, and J. Altaf, "Hybrid Intrusion Detection System for RPL IoT Networks Using Machine Learning and Deep Learning," *IEEE Access*, vol. PP, p. 1, Jan. 2024, DOI: 10.1109/ACCESS.2024.3442529.

[64] A. Wakili, S. Bakkali, and A. E. H. Alaoui, "Machine learning for QoS and security enhancement of RPL in IoT-Enabled wireless sensors," *Sens. Int.*, vol. 5, p. 100 289, 2024, DOI: 10.1016/j.sintl.2024.100289.

[65] N. W. Khan et al., "A hybrid deep learning-based intrusion detection system for IoT networks," *Math. Biosci. Eng.*, vol. 20, no. 8, pp. 13 491–13 520, 2023, DOI: 10.3934/mbe.2023602.

[66] S. Çakır, S. Toklu, and N. Yalçın, "RPL Attack Detection and Prevention in the Internet of Things Networks Using a GRU Based Deep Learning," *IEEE Access*, vol. 8, pp. 183 678–183 689, Oct. 2020, DOI: 10.1109/ACCESS.2020.3029191.

[67] I. S. Alsukayti and A. Singh, "A Lightweight Scheme for Mitigating RPL Version Number Attacks in IoT Networks," *IEEE Access*, vol. 10, pp. 111 115–111 133, 2022, DOI: 10.1109/ACCESS.2022.3215460.

[68] M. Rouissat, B. Mohammed, I. Alsukayti, and A. Mokaddem, "A Lightweight Mitigation Approach against a New Inundation Attack in RPL-Based IoT Networks," *Appl. Sci.*, vol. 13, Sep. 2023, DOI: 10.3390/app131810366.

[69] A. Burange and V. Deshmukh, "Securing IoT Attacks: A Machine Learning Approach for Developing Lightweight Trust-Based Intrusion Detection System," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, pp. 14–22, Sep. 2023, DOI: 10.17762/ijritcc.v11i7.7788.

[70] G. Sharma, J. Grover, and A. Verma, "A Lightweight Security Solution for Mitigation of Hatchetman Attack in RPL-based 6LoWPAN," in *2023 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, 2023, pp. 750–755. DOI: 10.1109/ANTS59832.2023.10469481.

Angel D is an experienced Information Technology Senior Lecturer with over 12 years in academia and industry, specializing in curriculum development, innovative teaching methodologies, and research. She is currently pursuing a Ph.D. in Computer Science Engineering at Presidency University and holds a master's degree in computer application. Her research interests include Machine Learning, AI, IoT, and Data Analytics. Angel has also contributed to international conferences, authored academic articles, and led workshops on emerging IT trends. She also held senior roles in the IT industry, as Senior Programmer Analyst at Cognizant and Infosys, specialized in BI applications and data reporting, working with clients such as Royal Bank of Scotland, AstraZeneca, Philips, and British Gas. She been recognized with several awards from the Ministry of Higher Education, Oman, and had the privilege of speaking at various academic institutions on the importance of strengthening soft skills.

Robin Rohit Vincent is a distinguished academician and researcher in the field of Computer Science and Engineering. He currently serves as Professor and Head of the NVIDIA Centre of Excellence (CoE) at Presidency School of Computer Science and Engineering, Presidency University, Bengaluru. He holds B.E., M.E., and Ph.D. degrees, along with a Postdoctoral Fellowship (PDF) from the UK, reflecting his strong academic and research background.

With extensive experience in teaching, research, and academic leadership, Dr. Vincent has contributed significantly to emerging areas of technology, mentoring students and guiding innovative projects. His work focuses on advancing knowledge in cutting-edge domains and fostering industry– academia collaboration through initiatives like the NVIDIA CoE.

Path Planning Transformer Supervised by Improved RRT* with Reduced Random Map Size for Mobile Robots

Aphilak Lonklang and János Botzheim

***Abstract*—The Improved Rapidly-exploring Random Tree with Reduced Random Map Size (IRRT*-RRMS) algorithm was previ- ously developed to find collision-free paths for mobile robot path planning. Given the excellent performance of Transformer Neural Networks with sequential data, we propose an encoder-decoder transformer model combining a Vision Transformer (ViT) as the encoder and a time-series forecasting module as the decoder to learn the path planning algorithm. The novelty of this paper lies in developing a model supervised by a dataset generated from the IRRT*-RRMS algorithm and using this trained model for the path planning task. The trained model efficiently predicts intermediate points between the desired starting and goal points. The performance was validated on a real robot, demonstrating that the trained model required less computation time compared to the IRRT*-RRMS algorithm.**

***Index Terms*—Path Planning Transformers, Mobile Robots, RRT*.**

I. INTRODUCTION

Material distribution is a key operation in modern manufacturing and logistics, where autonomous mobile robots (AMRs) and automated guided vehicles (AGVs) are deployed to improve efficiency, reduce human effort, and optimize resource use [1]. Path-planning algorithms enable these robots to navigate safely through complex environments involving static and dynamic obstacles. Classical methods such as A* and Dijkstra [2], [3] generate collision-free paths but rely on grid-based maps, which can be computationally and memory-intensive in large-scale environments [4], [5].

Sampling-based algorithms like Rapidly-exploring Random Trees (RRT) and RRT* [6], [7] address scalability by incrementally building space-filling trees. However, they often suffer from slow convergence and require numerous random samples. Informed RRT* [8] improves efficiency by focusing sampling in promising regions. Despite improvements, these methods remain costly in dynamic or large environments.

RRT, introduced by LaValle et al. [9], is widely used due to its simplicity and effectiveness in high-dimensional spaces [10]. It explores the environment via random sampling and builds a tree from the start to the goal. RRT* enhances this by optimizing path length [11], [12], but it still suffers from redundant sampling and inconsistent path quality [13].

Department of Artificial Intelligence, Faculty of Informatics, ELTE Eötvös Loránd University, Budapest, Hungary (E-mail: {aphilak, botzheim}@inf.elte.hu)

To improve efficiency, the Improved RRT* with Reduced Random Map Size (IRRT*-RRMS) [14] was proposed. It removes nodes in obstacle regions from the sampling space and excludes reused nodes in future iterations, preventing redundancy. This modification accelerates tree exploration and convergence. IRRT*-RRMS is especially suitable for dynamic environments with unknown static obstacles [15].

As environments become complex, traditional planners encounter limitations in speed and generalization [16], [17]. To overcome this, machine learning approaches—particularly neural networks—have been proposed [18]–[20]. CNN-based planners can encode spatial data but are often constrained by fixed-size maps. Recently, transformer-based models like the Motion Planning Transformer (MPT) [4] have shown promising results by learning to predict feasible paths in complex, high-dimensional spaces.

Transformers [21], originally designed for NLP, leverage attention mechanisms to model long-range dependencies. They have been successfully applied to tasks like image classification [22], visual inspection [23], and motion planning [24]. Variants have also shown success in time-series forecasting [25], [26] across diverse domains including traffic [27] and agriculture [28].

Although transformer-based planners like MPT offer enhanced flexibility and planning efficiency, they still face challenges such as generalization and dependency on large training datasets. To address this, we propose the Path Planning Transformer (PPT), which leverages IRRT*-RRMS as a data generator. This allows for high-quality, structured datasets with fewer noisy waypoints. Furthermore, post-processing techniques like Bacterial Mutation and Node Deletion are integrated to refine the predicted paths.

Classical sampling-based planners such as RRT* and the proposed IRRT*-RRMS algorithm are deterministic methods that explicitly explore the configuration space to guarantee collision-free solutions when feasible paths exist. However, such algorithms often require extensive sampling and iterative computation, which may lead to high planning times in complex environments. In contrast, learning-based approaches such as transformer neural networks operate as stochastic predictive models that approximate feasible paths based on patterns learned from training data. While these models can generate solutions significantly faster during inference, they may introduce uncertainty and occasionally produce infea-

DOI: 10.36244/ICJ.2026.1.11

sible paths. To leverage the advantages of both paradigms, the proposed framework uses the deterministic IRRT*-RRMS algorithm to generate high-quality training data, while the Path Planning Transformer learns to predict intermediate waypoints efficiently. Post-processing techniques such as Bacterial Mutation and Node Deletion are then applied to improve path feasibility and robustness.

In summary, our approach combines the strengths of classical and learning-based methods to deliver a robust, real-time path-planning framework for mobile robots, improving accuracy, generalization, and computational efficiency.

II. Problem Definition

In mobile robot navigation systems, a global path planner is typically used to compute an initial collision-free path between the starting location and the goal position using a global occupancy map. Classical sampling-based algorithms such as RRT* provide reliable solutions but often require significant computational effort due to repeated random sampling and iterative optimization. The objective of this work is to investigate whether a transformer-based neural network can learn the path generation behavior of the IRRT*-RRMS algorithm and generate feasible global paths more efficiently. The proposed Path Planning Transformer therefore focuses on accelerating the global path planning stage by predicting intermediate waypoints from map representations and start–goal configurations.

III. Improved RRT* with Reduced Random Map Size

A. Environment and Mapping

The environment is modeled as an $M \times N$ matrix, where each cell contains a unique index and its 2D Cartesian coordinates. As shown in Fig. 1, white cells indicate feasible regions, while red cells denote obstacles.

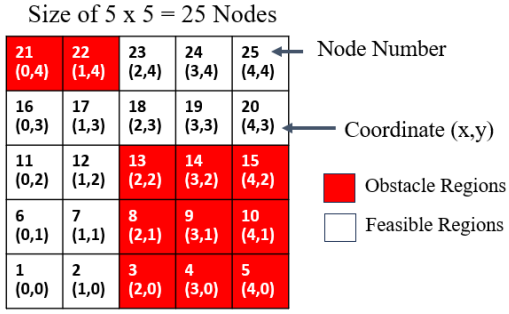


Fig. 1: Environment Mapping

B. Traditional RRT* and Limitations

Traditional RRT* incrementally builds a tree by sampling random nodes (q_{rand}), steering toward them from the nearest node (q_{near}), and optimizing via nearby nodes (q_{min}) within radius R [9], [11]. However, it retains unusable nodes (in red) [13], increasing computational cost in complex environments (Fig. 2).

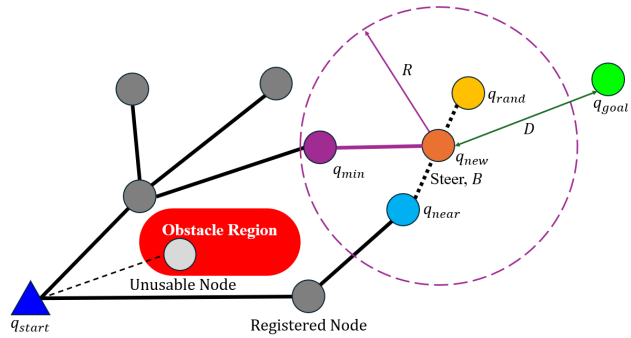


Fig. 2: Traditional RRT* Illustration

C. Feasible Region Mapping

To overcome this, the global map is flattened into a 1D vector ($randMap$), and obstacle nodes are removed [14]. Fig. 3 illustrates this pruning, which retains only feasible nodes for sampling.

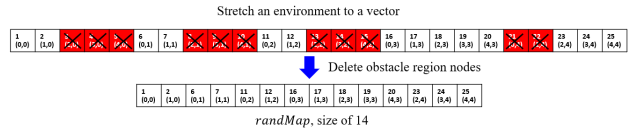
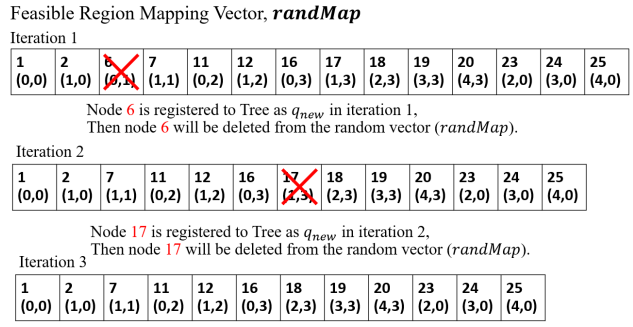


Fig. 3: Feasible Region Mapping

D. IRRT* with Reduced Random Map Size

IRRT*-RRMS [15], [29] further improves sampling by removing each used node from $randMap$ after registration, ensuring no repetitions and more efficient exploration. This is shown in Fig. 4, where node "6" is removed after iteration.



"Fig. 4: Reduced Random Map Size Technique"

The pseudocode in Alg. 1 summarizes the full algorithm, including pre-processing and tree construction steps.

E. Post-processing: Bacterial Mutation and Node Deletion

To refine paths, we apply two post-processing steps:

- **Bacterial Mutation:** introduces small variations to improve smoothness and collision avoidance.

Algorithm 1 Improved RRT* with Reduced Random Map Size Algorithm

$Map =$ **ReadMap** from file (.bmp)
$randMap =$ **StretchMap** from matrix to row vector
for $i <$ Length($randMap$) **do**
  if $randMap(i)$ is an obstacle region **then**
    Delete $randMap(i)$ from $randMap$ vector
  end if
end for
Initialize q_{start} and q_{goal}
for $i < MaxIteration$ **do**
  $q_{rand} \leftarrow$ random node $randMap$
  $q_{near} \leftarrow$ find nearest node from Tree
  if obstacle free between q_{near} and q_{new} **then**
    $q_{new} \leftarrow$ steer from q_{near}
    Find minimum cost from q_{min} and q_{new} in radius of R
    Add q_{new} to Tree
    Remove q_{new} from $randMap$
  end if
  if distance between q_{new} and $q_{goal} \leq D$ **then**
    Stop iteration
  end if
end for
Return Tree and Define Path as a Bacterium
for $i <$ size of bacterium **do**
  Bacterial Mutation
end for
Return Fine-tuned Bacterium
for $i <$ size of bacterium **do**
  Node Deletion
end for
Return Final Path
End

- **Node Deletion:** removes redundant nodes to reduce path length and complexity.

The cost function (Eq. (1)) evaluates path fitness using length L, Penalty for collisions, Turns, and smoothness factor S.

$$Cost(path) = L(path) + Penalty \cdot Coll(path) + S \cdot Turn(path) \quad (1)$$

Parameter settings are summarized in Table I. These were tuned for a 480×480 map with max iterations of 600, steering distance $B = 45$, and optimized radius $R = 80$.

F. Application and Integration

We validated IRRT*-RRMS using a real robot in unknown static obstacle scenarios via ROS and MATLAB. Fig. 5 shows the result: the magenta path is the raw IRRT*-RRMS output, and the red path is the final result after post-processing. The path is smooth, collision-free, and deployable in real-time scenarios.

TABLE I
PARAMETER SETTINGS

Algorithm	Parameter	Value
RRT*	Max Iteration	600
	Map Size	480 × 480
	Steering Distance (B)	45
	Radius (R)	80
Bacterial Mutation	N_{clone}	20
	Generations	Size of Bacterium
	$Penalty$	10,000
	S	0.1
Node Deletion	Iteration	Size of Bacterium

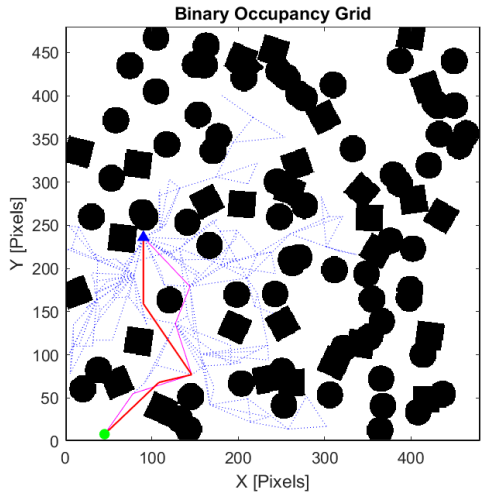


Fig. 5: Final path result after IRRT*-RRMS and post-processing from the start point (blue triangle) to the goal point (green circle)

In summary, IRRT*-RRMS reduces computational complexity by limiting the sampling space while maintaining high-quality path outputs. These paths are used to train the Path Planning Transformer, enabling learning from structured, optimized examples.

IV. TRANSFORMER NEURAL NETWORKS FOR MOBILE ROBOT PATH PLANNING

A. Transformer Architectures and Variants

The Transformer Neural Network, introduced by Vaswani et al. [21], revolutionized sequence modeling through its self-attention mechanism, replacing RNNs and CNNs in many NLP tasks. Its encoder-decoder structure allows efficient parallelization and long-range dependency modeling, forming the backbone of models like BERT and GPT.

Due to its scalability and modular design, the Transformer has been widely adopted beyond NLP, including vision and time-series domains. The Vision Transformer (ViT) [30] treats images as sequences of patches, applying attention mechanisms to capture spatial features, achieving competitive results in classification and segmentation. Similarly, Time Series Transformer variants extend this architecture to model temporal dependencies for forecasting applications, although a canonical form has not yet been standardized.

B. Motion Planning Transformer (MPT)

Transformers have emerged as a powerful tool in robotics, enabling efficient sensor data processing and spatial understanding for tasks such as navigation and obstacle avoidance. The Motion Planning Transformer (MPT) [24] introduces a data-driven approach to reduce the path planning search space by predicting feasible regions from environmental data. This reduces planning time and improves generalization across diverse environments.

MPT combines Transformer-based predictions with sampling-based motion planning (SMP), enhancing planning speed and accuracy. As shown in Fig. 6, the model outputs "Region Proposals" (green areas) for planning between start (blue) and goal (green) points. The resulting path (red) is computed based on these feasible regions.

In contrast, our approach uses IRRT*-RRMS to generate training data and feasible paths. The Transformer model can generate solutions directly, with optional refinement through post-processing algorithms.

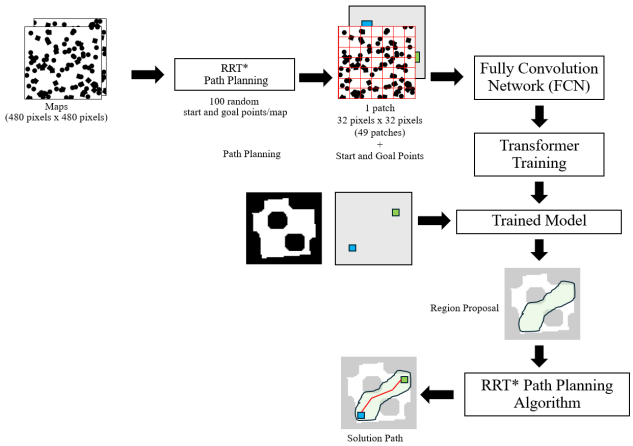


Fig. 6: Motion Planning Transformer (MPT)

V. Proposed Algorithm: Path Planning Transformers

This work proposes Path Planning Transformers, a transformer-based model for predicting collision-free paths in mobile robot navigation (Fig. 7). The pipeline begins with dataset generation using IRRT*-RRMS. Two datasets were created: a point robot dataset with 100 maps and 10,000 paths, and a real-size robot dataset with 30 maps and 3,000 paths. Each map is 480×480 pixels, and paths are stored as CSV files with start, goal, and intermediate coordinates.

A. Dataset Generation

We used random forest maps from the MPT paper. The IRRT*-RRMS algorithm, with post-processing, was used to generate paths. For real-robot datasets, obstacle areas were expanded by half the robot size. Maps were resized to 224×224 pixels for ViT compatibility, and 100 paths were generated per map. Fig. 8 shows dataset examples and CSV files with start (x_s, y_s), goal (x_g, y_g), and intermediate points (x_i, y_i).

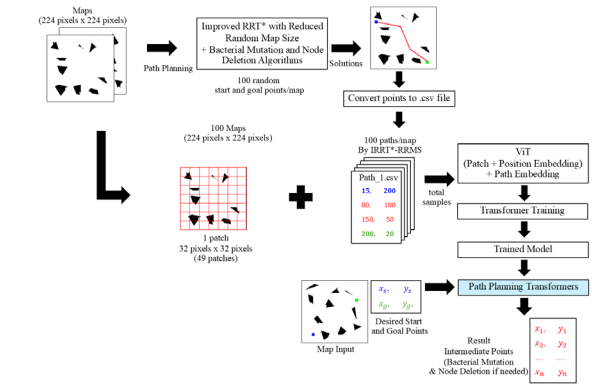


Fig. 7: Path Planning Transformers (PPT)

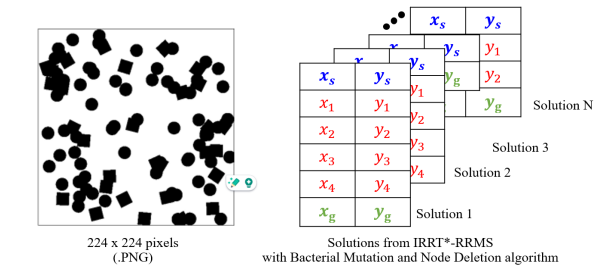


Fig. 8: Dataset Illustration

B. Dataset Encoding

We followed the ViT approach [30] to tokenize map patches (32×32) into 49 tokens, flattened to 1D embeddings. Each path is limited to 32 coordinates, including start (type-0), goal (type-1), intermediate (type-2), stop (type-4), and padding (type-3). Stop is marked as (0,0); padding as (-1,-1). Fig. 9 shows reordering for transformer input: $(0, 1, 2, 4, 3, \ldots)$.

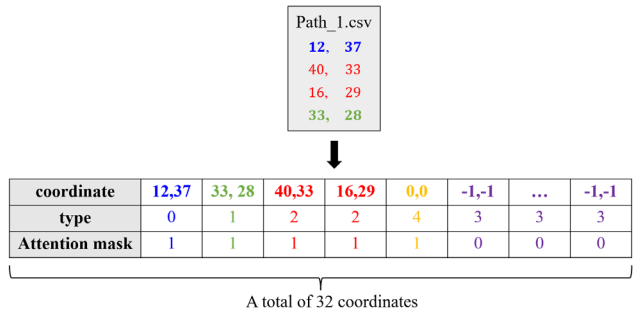


Fig. 9: Path Embedding and Reordering

C. Decoder and Loss Function

The decoder is based on the Time Series Transformer [31]. We used 3 decoder layers, a feature size of 768, and predicted 32 waypoints. The loss function (Eq. (2)) combines coordinate MSE and type classification cross-entropy.

$$L = \frac{1}{S}\sum_{j=1}^{S}\left(\frac{1}{n}\sum_{i=1}^{n}(x_{ij}-\hat{x}_{ij})^2 + (y_{ij}-\hat{y}_{ij})^2 - t_{ij}\log p_{ij}\right) \tag{2}$$

where

n = Number of intermediate points in each sample
S = Total number of samples
x = Prediction x-coordinate of each point
y = Prediction y-coordinate of each point
$\hat{x}$ = Ground truth x-coordinate of each point
$\hat{y}$ = Ground truth y-coordinate of each point
t = One hot encoded of the truth value of coordinate type
p = Soft-Max probability for the type of each point

D. Training

Model training was performed using PyTorch on an Intel i9 laptop with an RTX 3080 GPU. Learning started at $5e-5$ and decayed to $1e-5$ as validation loss converged. Each dataset was trained with the Adam optimizer, batch size 100.

Point Robot Model: Trained on 10,000 samples for 156 epochs. The best model was saved at epoch 105 with loss 1,571 (Fig. 10).

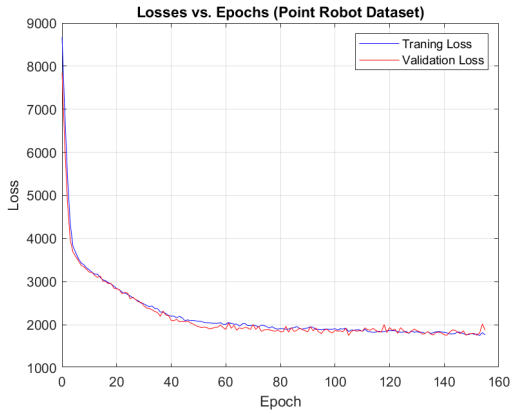


Fig. 10: Validation Loss (Point Robot Dataset)

Real-size Robot Model: Trained on 3,000 samples for 450 epochs. Optimal validation loss: 1,371 at epoch 353 (Fig. 11).

Figures 10 and 11 illustrate the training and validation loss curves for the transformer model trained on the point robot dataset and the real-size robot dataset, respectively. In both cases, the loss decreases rapidly during the initial training stage and gradually stabilizes as the training progresses. The validation loss closely follows the training loss throughout the training process, indicating stable convergence and suggesting that the model does not suffer from significant overfitting.

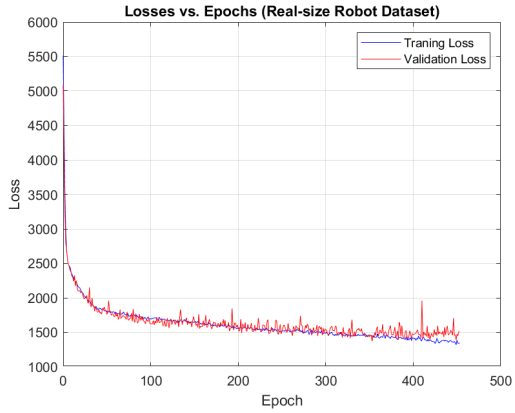
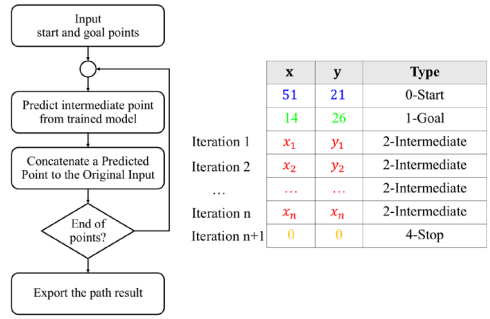


Fig. 11: Validation Loss (Real-size Robot Dataset)

These results demonstrate that the transformer model successfully learns the relationship between the map representation and the corresponding waypoint sequences generated by the IRRT*-RRMS algorithm.

E. Prediction Flow

The prediction processes are depicted in Fig. 12. The iterations start by determining the starting and goal points. The model predicts the result and concatenates the first coordinate to the original series. The iterations continue until the predicted coordinate is the padding or stop point; then, the iterations stop, and the solution is exported as a series of coordinates. The path results are checked for collision with the environment. If the path is collision-free, the solution can be used. On the other hand, the path results are fine-tuned with post-processing algorithms, which were mentioned in subsection post-processing algorithms. The solution path with the post-processing algorithms is called a Path Planning Transformer with Bacterial Mutation and Node Deletion algorithms (PPT-BM-ND).



	x	y	Type
	51	21	0-Start
	14	26	1-Goal
Iteration 1	x_1	y_1	2-Intermediate
Iteration 2	x_2	y_2	2-Intermediate
…	…	…	2-Intermediate
Iteration n	x_n	x_n	2-Intermediate
Iteration n+1	0	0	4-Stop

Fig. 12: Prediction Process of PPT

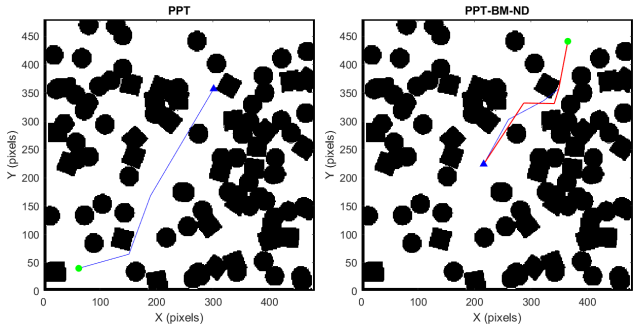
VI. RESULTS

In the following result figures, a blue triangle marker denotes the starting point, while a green circle marker represents

the goal point. The blue line indicates the path generated by the Proposed Path Transformer model. For post-processing enhancement methods, the red line illustrates the final trajectory produced by the PPT-BM-ND approach, which incorporates bacterial mutation and node deletion algorithms.

A. Point Robot Case

Fig. 13 compares path planning results using PPT and PPT-BM-ND. PPT alone (blue) occasionally results in collisions, while PPT-BM-ND (red) corrects these. As seen in Table II, PPT-BM-ND achieves 100% collision-free paths with minimal increase in path length and substantially lower computational time than IRRT*-RRMS. The average computation time for 100 samples is 0.614 s. Although maps are from training, start and goal points were randomly chosen.



(a) PPT Only (blue) (b) PPT-BM-ND (red) vs. PPT (blue)

Fig. 13: Point Robot Path Planning Results from Start (blue triangle) to Goal (green circle) Points

TABLE II
NUMERICAL RESULTS: POINT ROBOT (Fig. 13b)

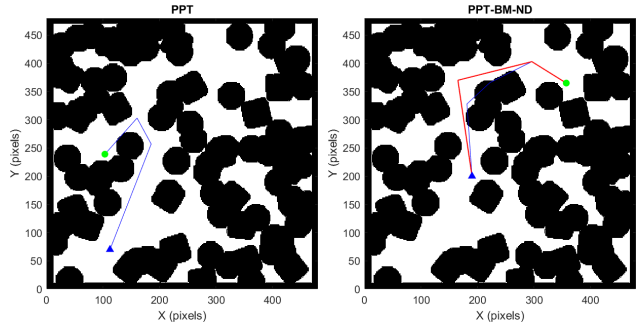
Method	Path Length (px)	Time (s)	Collision-Free?
IRRT*-RRMS	295	26.45	Yes
PPT	280	0.09	No
PPT-BM-ND	296	2.09	Yes

B. Real-size Robot Case

Fig. 14 shows results for real-size robots. PPT-BM-ND achieves 100% success, improving upon PPT's 75% and maintaining a shorter computation time than IRRT*-RRMS. The average inference time was 0.520 s (Table III).

TABLE III
NUMERICAL RESULTS: REAL-SIZE ROBOT (Fig. 14b)

Method	Path Length (px)	Time (s)	Collision-Free?
IRRT*-RRMS	367	17.25	Yes
PPT	338	0.88	No
PPT-BM-ND	379	2.70	Yes

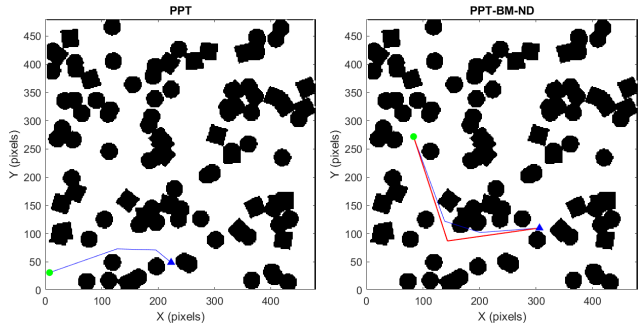


(a) PPT Only (blue) (b) PPT-BM-ND (red) vs. PPT (blue)

Fig. 14: Real-size Robot Results from Start (blue triangle) to Goal (green circle) Points

C. Unseen Map (Point Robot)

We tested generalization using unseen maps. As Fig. 15 shows, PPT produced valid paths in 40% of cases, while PPT-BM-ND achieved 100%. Table IV shows comparable path lengths but faster computation than IRRT*-RRMS. Average inference time: 1.22 s.



(a) PPT Only (blue) (b) PPT-BM-ND (red) vs. PPT (blue)

Fig. 15: Unseen Map Test: Point Robot from Start (blue triangle) to Goal (green circle) Points

TABLE IV
UNSEEN MAP RESULTS (Fig. 15b)

Method	Path Length (px)	Time (s)	Collision-Free?
IRRT*-RRMS	321	13.48	Yes
PPT	332	0.05	No
PPT-BM-ND	331	2.06	Yes

D. Real-World Implementation

A real environment was built to evaluate PPT on unseen data using the "Stephen" robot (Fig. 16). The test area (5.2 m. × 4.8 m.) was mapped using ROS and LiDAR (Fig. 17). The robot first explored the environment to construct an occupancy grid map representing obstacle regions and free space. This map was then used as the global map input for the path planning

algorithm. The real-world experiment primarily serves as a proof-of-concept validation demonstrating that the proposed transformer-based planning approach can be deployed on a physical robot platform. Obstacle regions were added to ensure safety. Fig. 18 shows that PPT only produced a non-usable path (blue), while PPT-BM-ND (red) generated a collision-free solution suitable for execution. A demonstration video is available: (Video Link).

To further evaluate the robustness of the proposed approach, an additional real-world navigation experiment was conducted. In this experiment, different start and goal positions were selected within the environment to test the ability of the Path Planning Transformer to generate feasible paths under varying navigation conditions.

As shown in Fig. 19, the transformer model predicts intermediate waypoints between the start and goal locations. The predicted path is then refined using the Bacterial Mutation and Node Deletion post-processing algorithms to ensure collision-free navigation.

The results demonstrate that the proposed framework successfully generates feasible paths in different navigation scenarios within the same real-world environment. The refined path avoids obstacle regions and produces smooth trajectories that can be executed by the robot.

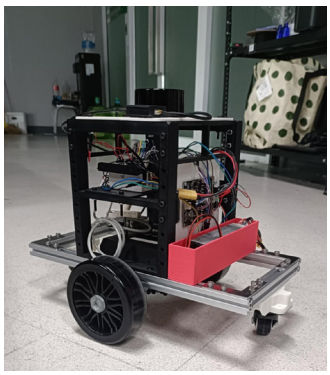


Fig. 16: Stephen Robot

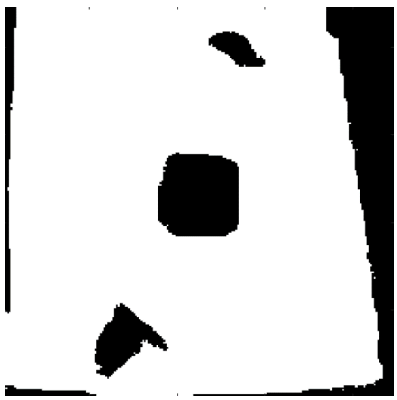


Fig. 17: Scanned Environment (SLAM)

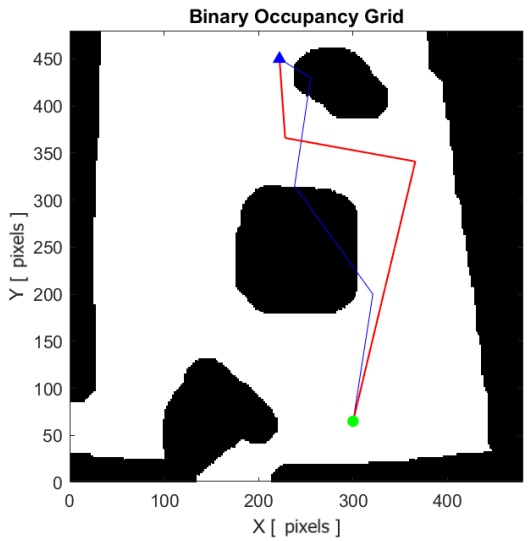


Fig. 18: Case I: Path Planning in Real Environment, The blue line shows the path predicted by the Path Planning Transformer, while the red line represents the refined trajectory obtained after applying the Bacterial Mutation and Node Deletion post-processing algorithms.

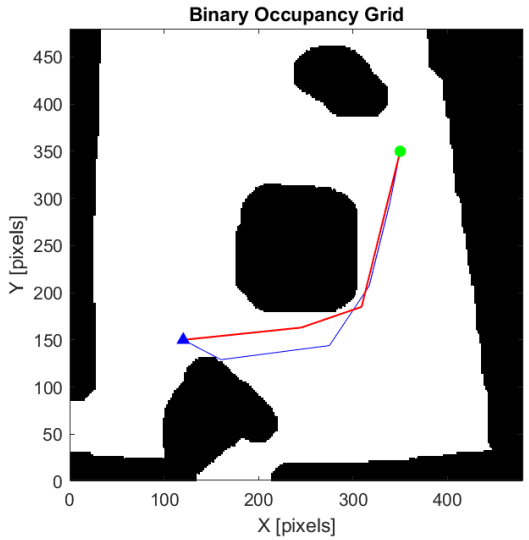


Fig. 19: Case II: Path Planning in Real Environment.

VII. DISCUSSION

The PPT-BM-ND approach consistently generated accurate, collision-free paths with fewer vertices and shorter inference times than IRRT*-RRMS and MPT. Table V summarizes the performance across algorithms. PPT-BM-ND achieved 100% collision-free path success rate in all tested cases, including unseen maps and real-world deployment. Although randomly generated maps were used during dataset generation to create diverse navigation scenarios, future work will investigate structured environments such as factory layouts with corridors and workstations.

TABLE V
PERFORMANCE COMPARISON: MPT vs. PPT VARIANTS

Method	Collision-free Path Success Rate (%)	Vertices Count
IRRT*-RRMS	100	0–10
MPT	100	60
PPT (Point Robot)	70	0–10
PPT (Real-size Robot)	75	0–10
PPT (Unseen Map)	40	0–10
PPT-BM-ND	100	0–10

A. Comparison of Deterministic and Stochastic Approaches

The IRRT*-RRMS algorithm operates as a deterministic sampling-based planner that guarantees convergence to an optimal solution given sufficient sampling. However, its iterative nature can be computationally intensive in complex or dynamic environments. In contrast, the proposed Path Planning Transformer is a stochastic model that learns the underlying patterns of collision-free paths from the IRRT*-RRMS dataset.

The primary benefits of this hybrid approach include:

- Computational Efficiency: Once trained, the PPT predicts paths in a single forward pass, significantly reducing the computation time compared to the iterative sampling required by IRRT*-RRMS.
- Knowledge Distillation: The model leverages high-quality, structured datasets generated by the improved algorithm to achieve better generalization than models trained on raw RRT data.

The primary drawback is the inherent uncertainty (stochasticity) of neural network predictions, which may occasionally result in sub-optimal or infeasible paths. To address this, we integrated Bacterial Mutation and Node Deletion as post-processing steps to refine these stochastic outputs into reliable, collision-free trajectories.

Although the proposed framework demonstrates promising performance in the evaluated scenarios, additional robustness tests such as noisy map inputs, partial observability, and dynamic obstacle environments could further strengthen the evaluation. These conditions are common in real-world robotic systems where sensor measurements may be uncertain or incomplete. Investigating the performance of transformer-based path planning models under such conditions is an important direction for future research.

VIII. CONCLUSION

In this paper, a Path Planning Transformer framework supervised by the Improved Rapidly-exploring Random Tree with Reduced Random Map Size algorithm was proposed for mobile robot navigation. The deterministic IRRT*-RRMS algorithm was first used to generate high-quality training datasets, which enabled the transformer model to learn structured path planning patterns from optimized solutions. The trained model was then applied to predict intermediate waypoints between the starting and goal positions in unknown environments.

Experimental results demonstrate that the proposed transformer-based model significantly reduces the computation time required for path planning compared with the original IRRT*-RRMS algorithm. While the PPT model alone occasionally produces infeasible paths due to the stochastic nature of neural network predictions, the integration of Bacterial Mutation and Node Deletion post-processing algorithms (PPT-BM-ND) effectively improves the reliability of the predicted trajectories. The results show that the proposed framework achieves a 100% collision-free success rate while maintaining substantially lower computation time than traditional sampling-based planning methods.

The proposed approach highlights the potential of combining deterministic planning algorithms with learning-based models, where classical planners provide reliable training data and transformer architectures enable fast inference during deployment. The experimental validation on both simulated environments and a real robot platform demonstrates the practical applicability of the method.

Future work will focus on extending the proposed framework to more complex navigation scenarios, including dynamic environments, multi-robot coordination, and larger-scale environments, as well as investigating improved learning strategies to further enhance the generalization capability of transformer-based path planning models.

REFERENCES

[1] M. Balogh and A. Vidács, "Optimizing camera stream transport in cloud-based industrial robotic systems," *Infocommunications Journal*, vol. XIV, no. 1, pp. 36–42, Mar. 2022. DOI: 10.36244/ICJ.2022.1.5. [Online]. Available: https://doi.org/10.36244/ICJ.2022.1.5.

[2] L. Liu, X. Wang, X. Yang, H. Liu, J. Li, and P. Wang, "Path planning techniques for mobile robots: Review and prospect," *Expert Systems with Applications*, vol. 227, p. 120 254, 2023, ISSN: 0957-4174. DOI: 10.1016/j.eswa.2023.120254. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095741742300756X.

[3] S. Venu and M. Gurusamy, "A comprehensive review of path planning algorithms for autonomous navigation," *Results in Engineering*, vol. 28, p. 107 750, 2025, ISSN: 2590-1230. DOI: 10.1016/j.rineng.2025.107750. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2590123025038034.

[4] J. J. Johnson, U. S. Kalra, A. Bhatia, L. Li, A. H. Qureshi, and M. C. Yip, Motion planning transformers: A motion planning framework for mobile robots, 2022. *arXiv: 2106.02791* [cs.RO].

[5] J. Tang, Z. Mao, and H. Ma, "Large-scale multi-robot coverage path planning on grids with path deconfliction," *IEEE Transactions on Robotics*, vol. 41, pp. 3348–3367, 2025, ISSN: 1941-0468. DOI: 10.1109/tro.2025.3567476. [Online]. Available: http://dx.doi.org/10.1109/TRO.2025.3567476.

[6] F. Tao, Z. Ding, Z. Fu, M. Li, and B. Ji, "Efficient path planning for autonomous vehicles based on RRT* with variable probability strategy and artificial potential field approach," *Scientific Reports*, vol. 14, no. 1, p. 24 698, 2024, ISSN: 2045-2322. DOI: 10.1038/s41598-024-76299-9. [Online]. Available: https://doi.org/10.1038/s41598-024-76299-9.

[7] C. Cui, Z. Wang, J. Sui, Y. Zhang, and C. Guo, "An improved RRT behavioral planning method for robots based on PTM algorithm," *Scientific Reports*, vol. 14, no. 1, p. 21 776, 2024, ISSN: 2045-2322. DOI: 10.1038/s41598-024-72616-4. [Online]. Available: https://doi.org/10.1038/s41598-024-72616-4.

[8] J. D. Gammell, S. S. Srinivasa, and T. D. Barfoot, "Informed RRT*: Optimal sampling-based path planning focused via direct sampling of an admissible ellipsoidal heuristic," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE*, Sep. 2014, pp. 2997–3004. DOI: 10.1109/iros.2014.6942976. [Online]. Available: http://dx.doi.org/10.1109/IROS.2014.6942976.

[9] S. M. LaValle, "Rapidly-exploring random trees: A new tool for path planning," *The annual research report*, 1998. [Online]. Available: https://api.semanticscholar.org/CorpusID:14744621.

[10] J. Wang, J. Li, Y. Song, Y. Tuo, and C. Liu, "FC-RRT*: A modified RRT* with rapid convergence in complex environments," *Journal of Computational Science*, vol. 77, p. 102 239, 2024, ISSN: 1877-7503. DOI: 10.1016/j.jocs.2024.102239. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877750324000322.

[11] S. Karaman and E. Frazzoli, "Incremental sampling-based algorithms for optimal motion planning," *CoRR*, vol. abs/1005.0416, 2010. *arXiv: 1005.0416*. [Online]. Available: http://arxiv.org/abs/1005.0416.

[12] O. Arslan, K. Berntorp, and P. Tsiotras, "Sampling-based algorithms for optimal motion planning using closed-loop prediction," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 4991–4996. DOI: 10.1109/ICRA.2017.7989581.

[13] J. Ding, Y. Zhou, X. Huang, K. Song, S. Lu, and L. Wang, "An improved RRT* algorithm for robot path planning based on path expansion heuristic sampling," *Journal of Computational Science*, vol. 67, p. 101 937, 2023, ISSN: 1877-7503. DOI: 10.1016/j.jocs.2022.101937. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877750322002964.

[14] A. Lonklang and J. Botzheim, "Improved rapidly exploring random tree with bacterial mutation and node deletion for offline path planning of mobile robot," *Electronics*, vol. 11, no. 9, 2022, ISSN: 2079-9292. DOI: 10.3390/electronics11091459. [Online]. Available: https://www.mdpi.com/2079-9292/11/9/1459.

[15] A. Lonklang and J. Botzheim, "Mobile robot path planning for unknown static obstacle avoidance by improved RRT* algorithm," in 2024 10th International Conference on Automation, *Robotics and Applications (ICARA)*, 2024, pp. 155–159. DOI: 10.1109/ICARA60736.2024.10553042.

[16] Y. Pan, Y. Tao, W. Lu, G. Li, and J. Cong, "Dynamic path planning of vehicles based on the adaptive potential field and hierarchical replacement immune algorithm," *Arabian Journal for Science and Engineering*, 2024, ISSN: 2191-4281. DOI: 10.1007/s13369-023-08541-x. [Online]. Available: https://doi.org/10.1007/s13369-023-08541-x.

[17] F. Hamad, H. N. Fakhouri, F. Alzghoul, and J. Zraqou, "Development and design of object avoider robot and object, path follower robot based on artificial intelligence," *Arabian Journal for Science and Engineering*, 2024, ISSN: 2191-4281. DOI: 10.1007/s13369-024-09365-z. [Online]. Available: https://doi.org/10.1007/s13369-024-09365-z.

[18] J. Yu, Y. Su, and Y. Liao, "The path planning of mobile robot by neural networks and hierarchical reinforcement learning," *Frontiers in Neurorobotics*, vol. 14, 2020, ISSN: 1662-5218. DOI: 10.3389/fnbot.2020.00063. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnbot.2020.00063.

[19] H. Bharadwaj and V. K. E, "Comparative study of neural networks in path planning for catering robots," *Procedia Computer Science*, vol. 133, pp. 417–423, 2018, International Conference on Robotics and Smart Manufacturing (RoSMa2018), ISSN: 1877-0509. DOI: 10.1016/j.procs.2018.07.051. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050918309967.

[20] R. Yonetani, T. Taniai, M. Barekatain, M. Nishimura, and A. Kanezaki, "Path planning using neural A* search," *CoRR*, vol. abs/2009.07476, 2020. *arXiv:2009.07476*. [Online]. Available: https://arxiv.org/abs/2009.07476.

[21] A. Vaswani, N. Shazeer, N. Parmar, *et al*., "Attention is all you need," *CoRR*, vol. abs/1706.03762, 2017. *arXiv: 1706.03762*. [Online]. Available: http://arxiv.org/abs/1706.03762.

[22] Y. Bazi, L. Bashmal, M. M. A. Rahhal, R. A. Dayil, and N. A. Ajlan, "Vision transformers for remote sensing image classification," *Remote Sensing*, vol. 13, no. 3, 2021, ISSN: 2072-4292. DOI: 10.3390/rs13030516. [Online]. Available: https://www.mdpi.com/2072-4292/13/3/516.

[23] N. Hütten, R. Meyes, and T. Meisen, "Vision trans-former in industrial visual inspection," *Applied Sciences*, vol. 12, no. 23, 2022, ISSN: 2076-3417. DOI: 10.3390/app122311981. [Online]. Available: https://www.mdpi.com/2076-3417/12/23/11981.

[24] J. J. Johnson, L. Li, A. H. Qureshi, and M. C. Yip, "Motion planning transformers: One model to plan them all," *CoRR*, vol. abs/2106.02791, 2021. *arXiv: 2106.02791*. [Online]. Available: https://arxiv.org/abs/2106.02791.

[25] S. Jung, K. Kim, H. Kwak, and Y. Park, "A worrying analysis of probabilistic time-series models for sales forecasting," *CoRR*, vol. abs/2011.10715, 2020. *arXiv: 2011.10715*. [Online]. Available: https://arxiv.org/abs/2011.10715.

[26] Q. Wen, T. Zhou, C. Zhang, *et al*., Transformers in time series: A survey, 2023. *arXiv: 2202.07125* [cs.LG].

[27] V. Flunkert, D. Salinas, and J. Gasthaus, "DeepAR: Probabilistic forecasting with autoregressive recurrent networks," *CoRR*, vol. abs/1704.04110, 2017. *arXiv: 1704.04110*. [Online]. Available: http://arxiv.org/abs/1704.04110.

[28] L. Bi, O. Wally, G. Hu, A. U. Tenuta, Y. R. Kandel, and D. S. Mueller, "A transformer-based approach for early prediction of soybean yield using time-series images," *Front Plant Sci*, vol. 14, p. 1 173 036, 2023, ISSN: 1664-462X (Print) 1664-462x. DOI: 10.3389/fpls.2023.1173036.

[29] A. Lonklang and J. Botzheim, "A rapidly-exploring random tree algorithm with reduced random map size," in *2023 9th International Conference on Automation, Robotics and Applications (ICARA)*, 2023, pp. 356–361. DOI: 10.1109/ICARA56516.2023.10125934.

[30] A. Dosovitskiy, L. Beyer, A. Kolesnikov, *et al*., "An image is worth 16x16 words: Transformers for image recognition at scale," *CoRR*, vol. abs/2010.11929, 2020. *arXiv: 2010.11929*. [Online]. Available: https://arxiv.org/abs/2010.11929.

[31] K. Rasul, A. Ashok, A. R. Williams, et al., Lag-Llama: Towards foundation models for probabilistic time series forecasting, 2024. *arXiv: 2310.08278* [cs.LG]. [On-line]. Available: https://arxiv.org/abs/2310.08278.

Aphilak Lonklang's degrees earned: Bachelor's degree in Aeronautical Engineering (First Class Honor B.Eng) and Master's degree in Mechanical Engineering (M.Eng) from the Suranaree University of Technology (SUT), Nakhon Ratchasima, Thailand. Experiences: Research Assistant in Industrial Automation Research Unit for three years in SUT (2013-2016). Lecturer at the School of Mechanical Engineering, Institute of Engineering, SUT (August 2016 - present). Award in Teaching Experience: Associate Fellow of the Higher Education Association from UKPSF (December 1st 2021). Research Interests: Industrial Automation, Industrial Robotics, Computational Intelligence.

János Botzheim received the M.Sc. and Ph.D. degrees in computer engineering from the Budapest University of Technology and Economics, in 2001 and 2008, respectively. He is an associate professor and the Head of Department of Artificial Intelligence at Eötvös Loránd University (ELTE), Faculty of Informatics, Budapest, Hungary. He had several international visiting fellowships. For six years he was an associate professor at Tokyo Metropolitan University. He is a member of several scientific societies such as IEEE, Hungarian Academy of Engineering, John von Neumann Computer Science Society, and Hungarian Fuzzy Association. His research interests are computational intelligence and cognitive robotics.

General Chair
Abderrahim Benslimane
Avignon University, France

Steering Committee Chair
Samuel Pierre
Polytechnic of Montreal, Canada

Technical Program Co-Chairs
Wessam Ajib
Université du Québec à Montréal, Canada
Chiara Boldrini
IIT-CNR Istituto di Informatica e Telematica, Italy
Rasheed Hussain
University of Bristol, UK
Xiaohui Liang
University of Massachusetts Boston, USA
Chrysa Papagianni
University of Amsterdam, Netherlands

Workshop co-Chairs
Panagiotis Papadimitratos
KTH Royal Institute of Technology, Sweden
Cheng Li
Simon Fraser University, Canada

Publication co-Chair
Saadi Boudjit
University of Rouen Normandy, France

Publicity co-Chairs
Ranwa Al Mallah
Polytechnique Montréal, Canada
Bintao Hu
Xi'an Jiaotong-Liverpool University, China
Feng Ye
University of Wisconsin-Madison, USA

Student Travel Grant Chair
Jennifer Simonjan
Technology Innovation Institute, UAE

Short Papers, Demos and Posters Symposium Co-Chairs
Marica Amadeo
University Mediterranea of Reggio Calabria, Italy
Valeria Loscri
INRIA, France

Webmaster Chair
Ngoran Magnuss Dufe
Avignon University, France

Local Organization Chair
Abderrahim Benslimane
Avignon University, France

The WiMob conference is an international forum for the exchange of experience and knowledge among researchers and developers concerned with wireless and mobile technology. For twenty-two years, the International WiMob conference has provided unique opportunities for researchers to interact, share new results, show live demonstrations, and discuss emerging directions in next-generation networks, 5G/6G, IoT and Artificial Intelligence based Communication and Networking.

WiMob 2026 will take place on 14-16 October 2026, at the Avignon Grand Hotel, Avignon, France. Avignon is the very charming and historical city in south of France, the Provence. It is a vibrant city, seat of the popes and city of the art..

WiMob 2026 is soliciting high-quality technical papers addressing research challenges in the areas of Wireless Communication, Next Generation Mobile Networking, Mobility and Nomadicity, Ubiquitous Computing, Services and Applications, Green and sustainable communications and network computing, Security on Wireless and mobile Networks and Artificial Intelligence based Communications and Networking. Papers should present original work validated via analysis, simulation or experimentation. Practical experiences and testbed trials also are welcome.

IEEE WiMob 2026 will host FIVE parallel symposia, including but not limited to the following topics:

- Wireless Communications (WC)
- Next Generation Mobile Networking (NGMN)
- Artificial Intelligence based Communications & Networking (AICN)
- Blockchain and Cryptocurrency in Mobile Networks (BCMN)
- Trust, Security and Privacy in Wireless and Mobile Networks (SWMN)

IMPORTANT DATES

Papers Submission: June 1, 2026
Notification of Acceptance: July 30, 2026
Camera Ready & Author Registration: September 1, 2026
Registrations for Authors: September 1, 2026

INSTRUCTIONS FOR PAPER SUBMISSION

Authors are required to submit fully formatted, original papers (PDF), with graphs, images, and other special areas arranged as intended for the final publication.

Papers should be written in English conforming to the IEEE standard conference format (8.5" x 11" - US letter, Two-Column). The initial submission for review will be limited to 8 pages. The final manuscript for publication will be limited to 8 IEEE pages. Additional charges may apply for additional pages.

Conference content will be submitted for inclusion into IEEE Xplore as well as other Abstracting and Indexing (A&I) databases.

Each accepted paper must be presented at the conference by one of the co-authors or a third party.

Only timely submissions through EDAS here will be accepted.

For more details please visit the WiMob 2026 official website (www.wimob.org/wimob2026).

ACCEPTED PAPERS

All accepted papers will be published in the conference proceedings and will be accessible via the IEEE Xplore Digital Library.

IEEE IEEE ComSoc IEEE Communications Society WiMOB

22nd International Conference on Network and Service Management

Evolving Network and Service Management: From Automation to Agentic Intelligence

Alcalá de Henares (Madrid), Spain // 26 - 30 October, 2026

CALL FOR PAPERS

The 22nd International conference on Network and Service Management (CNSM) is inviting authors to submit original contributions to network and service management research. CNSM is a selective single-track conference that covers all aspects of network and service management, pervasive systems, enterprises, and cloud computing environments. In particular, CNSM 2026 will focus on **Evolving Network and Service Management: From Automation to Agentic Intelligence**.

Papers accepted and presented at CNSM 2026 will be submitted for inclusion into IEEE Xplore, subject to meeting IEEE Xplore's scope and quality requirements. Authors of selected papers accepted for publication in the CNSM 2026 proceedings will be invited to submit an extended version of their papers to the IEEE Transactions on Network and Service Management journal.

Topics of Interest (but not limited to)

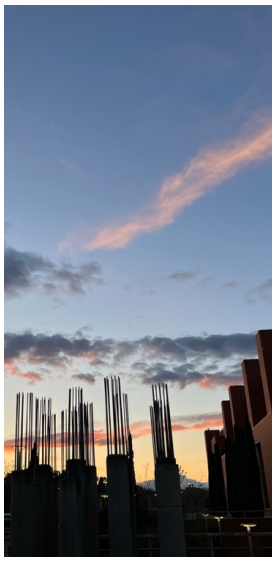
Technologies
- Communication Protocols
- Middleware
- Overlay Networks
- Peer-to-Peer Networks
- Technologies for Computing Continuum
- 5G/6G Networks
- Federated and Distributed Learning
- Generative AI and Large Language Models
- Green and Sustainable Networking
- Information Visualization
- Software-Defined Networking
- Monitoring and Measurements
- Multi-Access Edge Computing
- Network Function Virtualization
- Orchestration
- Operations and Business Support Systems
- Control and Data Plane Programmability
- Distributed Ledger Technology
- Digital Twins for Networks and Services
- Reinforcement Learning
- Secure and Dependable Networking

Service Management
- Multimedia Services
- Content Delivery Services
- Cloud/Edge Computing Services
- Data Services
- Internet Connectivity and Internet Access Services
- Internet of Things Services
- Security Services
- Context-Aware Services
- Information Technology Services
- Service Assurance

Functional Areas
- Fault Management
- Configuration Management
- Accounting Management
- Performance Management
- Security Management

Management Paradigms
- Centralized Management
- Hierarchical Management
- Distributed Management
- Federated Management
- Autonomic and Cognitive Management
- Policy- and Intent-Based Management
- Model-Driven Management
- Pro-active Management
- Energy-aware Management
- QoE-Centric Management

Methods
- Artificial Intelligence and Machine Learning
- Mathematical Logic and Automated Reasoning
- Optimization Theories
- Control Theory
- Probability Theory, Stochastic Processes, and Queuing Theory
- Artificial Intelligence and Machine Learning
- Evolutionary Algorithms
- Economic Theory and Game Theory
- Monitoring and Measurements
- Data Mining and (Big) Data Analysis
- Computer Simulation Experiments
- Testbed Experimentation and Field Trials
- Software Engineering Methodologies



Important Dates

Paper Submission:
15 June 2026
Rebuttal Period:
19-21 August 2026
Acceptance Notification:
26 August 2026
Camera Ready due:
14 September 2026

Technical Program Co-Chairs

Molka Gharbaoui, Scuola Superiore Sant'Anna, Italy

Mohamed Faten Zhani, King Fahd University of Petroleum and Minerals, Saudi Arabia

General Co-Chairs

Elisa Rojas, Universidad de Alcalá, Spain

Jaime Galán-Jiménez, University of Extremadura, Spain

Paper Submission

Authors are invited to submit original contributions that have not been published or submitted for publication elsewhere. Papers should be prepared using the IEEE 2-column conference style and are limited to 9 pages including references (full papers) or 5 pages including references (short papers). Papers must be submitted electronically in PDF format through EDAS at https://edas.info/N35030

Papers exceeding page limits, multiple submissions, and self-plagiarized papers will be rejected without further review. All other papers will get a thorough single-blind review process, followed by a rebuttal phase.

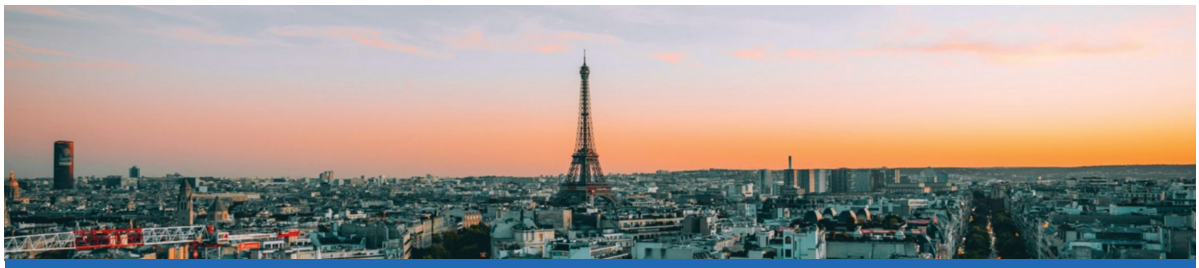
For further information, please check http://www.cnsm-conf.org/2026/

International Conference on Interconnected AI and NETworks

AIxNET 2026

November 23-25, 2026

Paris, France



CALL FOR PAPERS

Networks are entering an era where both classical ML and emerging generative and agentic AI are transforming end-to-end networking—from intent capture to closed-loop control across RAN, Core, transport, and edge/cloud. AIxNET welcomes contributions that advance algorithms, architectures, protocols, evaluations, and safeguards for trustworthy, explainable, and safe-to-operate AI-driven networking. We particularly encourage rigorous comparative studies across control layers (SMO/intent vs near-RT vs lower-layer control), and the release of open datasets and artifacts to help the community build together.

AIxNET is intending to build a stimulating, open, dynamic, and friendly forum to co-create the future and spark collaborations across teams. The conference will be a unique opportunity to gather academic and industry research on this crucial topic for 2030 networks. Expect interactive sessions, demos, and time for discussion.

Main Topics of Interest include (but are not limited to)

1. Agentic AI: from Human Intent to Action Autonomy
- Networked "xLM" challenges: Intent capture/parsing/policy synthesis at SMO and service layers, use of Large, Small or Machine Language Models (LLM, SLM, MLM)
- Hierarchical/heterogeneous agents spanning non-RT and near-RT control (e.g., O-RAN RIC), Core CNFs, and edge resources
- Agentic 6G functions
- Interconnection and collaboration between AI agents
- Tool and protocols for network-facing agents (e.g., MCP-enabled clients/servers), conflict resolution, safe rollbacks

2. New paradigms for networking: from Classical ML to xLM-based Control at Scale
- Supervised/unsupervised/self-supervised learning for prediction, anomaly detection, resource allocation, QoE optimization
- ML and LLM techniques for scheduling, slicing, mobility, energy saving; cross-domain orchestration across RAN/Core/transport for B5G and 6G
- Programmable data planes (P4/eBPF) and SDN control plane with ML-in-the-loop; NWDAF-enabled analytics
- Challenges for access networks and edge networking, use of alternative models, SLM, TRM
- Architecture and framework for agentic AI networking
- Data collection and labeling

3. Machine Learning and Artificial Intelligence for the Physical Layer
- AI/ML for PHY layer optimization: new air interfaces, waveforms, modulation and coding techniques
- AI/ML-augmented next generation multiple access techniques (SDMA, NOMA, RSMA)
- Physical layer AIML techniques for Massive MIMO; Cell-free and distributed massive MIMO; Massive, Ultra-Massive, extreme, and fluid MIMO
- Integrated sensing and communication (ISAC)
- AI/ML-based waveform design techniques tailored for emerging multi-antenna solutions, including different RIS architectures, XL-MIMO, pinching antennas, moveable antennas, fluid antennas, etc.
- ML for Terahertz and Millimeter Wave communications

4. AI/ML for wireless/optical/satellite networks
- Open, Programmable and AI-Native Radio Access Network and nodes
- Intelligent radio resource management and spectrum allocation
- Self-organizing networks (SON) and autonomous network management
- AI-native air interface design for 6G
- Autonomous satellite network management and orchestration
- AI for LEO/MEO/GEO constellation optimization; Integrated terrestrial–satellite network intelligence
- AI and ML for optical systems and networks; Optical network control and management; elastic optical networks and software-defined optics; Digital twins for optical network planning and optimization

5. Comparative Designs Across Layers: SMO/Intent vs Near-RT vs Lower-Layer Control
- Side-by-side evaluations of top-down (intent-driven) vs bottom-up (local) autonomy
- Responsibility split across SMO policies, RIC xApps/rApps, Core functions, device/edge controllers
- Stability, latency and safety; arbitration under competing objectives (QoE, energy, cost, SLAs)
- Cross-layer observability, auditability, and explainability methodologies

6. Explainability and trustworthiness: Bias and Functional Safety
- Human in the loop supervision and autonomy levels for safe operations
- Explainability for operator oversight (pre/post methods, rationales, provenance, accountability logs)
- Security and governance for AI-operated changes (access control, authorization, verification, compliance-by-design)
- Possible Bias sources and mitigation (data, prompts, tools, policies); fairness in resource allocation and service admission
- Trust, safety and ethical considerations in generative and agentic AI networking

7. Explainability and trustworthiness: Bias and Functional Safety
- Human in the loop supervision and autonomy levels for safe operations
- Explainability for operator oversight (pre/post methods, rationales, provenance, accountability logs)
- Security and governance for AI-operated changes (access control, authorization, verification, compliance-by-design)
- Possible Bias sources and mitigation (data, prompts, tools, policies); fairness in resource allocation and service admission
- Trust, safety and ethical considerations in generative and agentic AI networking

8. Evaluation, Benchmarks, Open Datasets, and experimentations
- Public datasets/benchmarks for RAN/Core/transport/edge; simulated vs real testbeds
- Evaluation methodology and built of meaningful KPIs (e.g., relying on MTTR, SLO, energy–QoE trade-offs…)
- Network performance metric in generative and agentic AI communication systems
- Digital twins, experimentation platforms, and testbeds for generative and agentic AI networking
- Reproducible pipelines, artifact sharing, and insightful negative results, robustness to drift
- Sustainability and cost modeling (e.g., compute budgets, edge vs cloud placement)

https://aixnet.dnac.org/

IMPORTANT DATES:

Paper Submission Due:
June 20, 2026

Notification of Acceptance:
September 15, 2026

Camera-Ready Papers due:
September 30, 2025

ORGANIZING COMMITTEE:

GENERAL CO-CHAIRS:

Stefano Secci
Cnam, France

Emmanuel Bertin
Orange Innovation, France

TPC CO-CHAIRS:

Sahar Hoteit
Université Paris-Saclay, France

Chiara Contoli
University of Urbino, Italy

Guidelines for our Authors

Format of the manuscripts

Original manuscripts and final versions of papers should be submitted in IEEE format according to the formatting instructions available on

https://journals.ieeeauthorcenter.ieee.org/
Then click: "IEEE Author Tools for Journals"
- "Article Templates"
- "Templates for Transactions".

Length of the manuscripts

The length of papers in the aforementioned format should be 6-8 journal pages.
Wherever appropriate, include 1-2 figures or tables per journal page.

Paper structure

Papers should follow the standard structure, consisting of *Introduction* (the part of paper numbered by "1"), and *Conclusion* (the last numbered part) and several *Sections* in between.
The Introduction should introduce the topic, tell why the subject of the paper is important, summarize the state of the art with references to existing works and underline the main innovative results of the paper. The Introduction should conclude with outlining the structure of the paper.

Accompanying parts

Papers should be accompanied by an *Abstract* and a few *Index Terms (Keywords)*. For the final version of accepted papers, please send the short cvs and *photos* of the authors as well.

Authors

In the title of the paper, authors are listed in the order given in the submitted manuscript. Their full affiliations and e-mail addresses will be given in a footnote on the first page as shown in the template. No degrees or other titles of the authors are given. Memberships of IEEE, HTE and other professional societies will be indicated so please supply this information. When submitting the manuscript, one of the authors should be indicated as corresponding author providing his/her postal address, fax number and telephone number for eventual correspondence and communication with the Editorial Board.

References

References should be listed at the end of the paper in the IEEE format, see below:

a) Last name of author or authors and first name or initials, or name of organization
b) Title of article in quotation marks
c) Title of periodical in full and set in italics
d) Volume, number, and, if available, part
e) First and last pages of article
f) Date of issue
g) Document Object Identifier (DOI)

[11] Boggs, S.A. and Fujimoto, N., "Techniques and instrumentation for measurement of transients in gas-insulated switchgear," IEEE Transactions on Electrical Installation, vol. ET-19, no. 2, pp.87–92, April 1984. DOI: 10.1109/TEI.1984.298778

Format of a book reference:

[26] Peck, R.B., Hanson, W.E., and Thornburn, T.H., Foundation Engineering, 2nd ed. New York: McGraw-Hill, 1972, pp.230–292.

All references should be referred by the corresponding numbers in the text.

Figures

Figures should be black-and-white, clear, and drawn by the authors. Do not use figures or pictures downloaded from the Internet. Figures and pictures should be submitted also as separate files. Captions are obligatory. Within the text, references should be made by figure numbers, e.g. "see Fig. 2."
When using figures from other printed materials, exact references and note on copyright should be included. Obtaining the copyright is the responsibility of authors.

Contact address

Authors are requested to submit their papers electronically via the following portal address:

https://www.ojs.hte.hu/infocommunications_journal/about/submissions

If you have any question about the journal or the submission process, please do not hesitate to contact us via e-mail:

Editor-in-Chief: Pál Varga – pvarga@tmit.bme.hu

Associate Editor-in-Chief:

József Bíró – biro@tmit.bme.hu

László Bacsárdi – bacsardi@hit.bme.hu

Special Issue
of the **Infocommunication Journal**

Call for Papers

Data Science and Information Technology

The topics include, but are not limited to:

Artificial Intelligence and its Applications • Stochastic Models in Data Science • Automata, Logic, and Models of Computation • Predictive Analytics • Cybersecurity • Temporal Data Science with Applications • Autonomous Vehicles and Embedded Systems • Virtual Reality and Its Applications • Software Technology • Internet of Things, Operational Technology, and Network Intelligence • AI Methods in Experimental Particle Physics

This special issue collects the latest results emerging on the field of Data Science and Information Technology.

Special Issue Editors:

Dr. András Hajdu
University of Debrecen, Hungary
hajdu.andras@inf.unideb.hu

Dr. Balázs Harangi
University of Debrecen, Hungary
harangi.balazs@inf.unideb.hu

TPC members:

Dr. Do Van Tien
Budapest University of Technology and Economics, Hungary

Dr. Tamás Márton Bérczes; **Dr. Zoltán Gál**; **Andrea Pintér-Huszti**; **Dr. Attila László Gilányi**; **Dr. Sándor Baran**; **Dr. Imre Varga**
University of Debrecen, Hungary

Prof. Péter Baranyi
CIAS at the Corvinus University of Budapest

Important date:

Submission deadline: **30 November, 2026**
Expected publication: **March, 2027**

Regarding manuscript submission information, please visit:
https://www.infocommunications.hu/for-our-authors

SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS

Who we are

Founded in 1949, the Scientific Association for Infocommunications (formerly known as Scientific Society for Telecommunications) is a voluntary and autonomous professional society of engineers and economists, researchers and businessmen, managers and educational, regulatory and other professionals working in the fields of telecommunications, broadcasting, electronics, information and media technologies in Hungary.

Besides its 1000 individual members, the Scientific Association for Infocommunications (in Hungarian: Híradástechnikai és Informatikai Tudományos Egyesület, HTE) has more than 60 corporate members as well. Among them there are large companies and small-and-medium enterprises with industrial, trade, service-providing, research and development activities, as well as educational institutions and research centers.

HTE is a Sister Society of the Institute of Electrical and Electronics Engineers, Inc. (IEEE) and the IEEE Communications Society.

What we do

HTE has a broad range of activities that aim to promote the convergence of information and communication technologies and the deployment of synergic applications and services, to broaden the knowledge and skills of our members, to facilitate the exchange of ideas and experiences, as well as to integrate and harmonize the professional opinions and standpoints derived from various group interests and market dynamics.

To achieve these goals, we…

- contribute to the analysis of technical, economic, and social questions related to our field of competence, and forward the synthesized opinion of our experts to scientific, legislative, industrial and educational organizations and institutions;
- follow the national and international trends and results related to our field of competence, foster the professional and business relations between foreign and Hungarian companies and institutes;
- organize an extensive range of lectures, seminars, debates, conferences, exhibitions, company presentations, and club events in order to transfer and deploy scientific, technical and economic knowledge and skills;
- promote professional secondary and higher education and take active part in the development of professional education, teaching and training;
- establish and maintain relations with other domestic and foreign fellow associations, IEEE sister societies;
- award prizes for outstanding scientific, educational, managerial, commercial and/or societal activities and achievements in the fields of infocommunication.

Contact information

President: **FERENC VÁGUJHELYI** • *elnok@hte.hu*
Secretary-General: **GÁBOR KOLLÁTH** • *kollath.gabor@hte.hu*
Operations Director: **PÉTER NAGY** • *nagy.peter@hte.hu*

Address: H-1051 Budapest, Bajcsy-Zsilinszky str. 12, HUNGARY, Room: 502
Phone: +36 1 353 1027
E-mail: *info@hte.hu*, Web: *www.hte.hu*