Infocommunications Journal

A PUBLICATION OF THE SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS (HTE)

December 2025 Volume XVII Number 4 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)

The publication was produced with the support of the Hungarian Academy of Sciences and the NMHH

nmhh National Media and Infocommunications Authority | Hungary

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 December 2025 issue

Margit Antal, Hsin-Ying Liang,
Chuan-Bi Lin, Yasmeen Taha Yaeeen,
Mohd Fadzli Mohd Salleh, Taha A. Elwi,
Ádám Kiss, László Schäffer,
Yuewen Xu, Qian Hu, Zhongqiang Luo,
Wenshi Xiao, T. Padmavathi,
Kusma Kumari Cheepurupalli, R. Madhu,
Sunil Kumar Shah, Raghvendra Sharma,
Neeraj Shukla, Márton L. Márton,
László Bacsárdi

Evaluation of Embedding Models for Hungarian Question-Answer Retrieval on Domain-Specific and Public Benchmarks

Margit Antal

Abstract—**Embedding models have become a fundamental component of modern natural language processing, yet their performance in morphologically rich, low-resource languages such as Hungarian remains underexplored. In this paper, we present a systematic evaluation of state-of-the-art embedding models for Hungarian question–answer retrieval. We construct two complementary evaluation datasets: (i) a domain-specific corpus collected from company documentation, preprocessed into topical chunks with human-verified question–answer pairs and (ii) the publicly available HuRTE benchmark. Using Chroma as the vector database, we compare eight multilingual and cross-lingual embedding models alongside keyword-based search baseline.**

Performance is measured using Mean Reciprocal Rank (MRR) and Recall@k. Results show substantial variation across mod- els and datasets, with notable differences between domain-specific and general-purpose retrieval tasks. BGE-M3 and XLM-ROBERTA achieved the highest accuracy (MRR: 0.90) on the Clearservice dataset, while GEMINI demonstrated superior performance on HuRTE (MRR: 0.99). We complement the evaluation with comprehensive error analysis, highlighting challenges posed by Hungarian domain-specific terminology, synonyms, and overlapping topics, and discuss trade-offs in efficiency through index build time and query latency measurements. Our findings provide a comparative study of embedding-based retrieval in Hungarian, offering practical guidance for downstream applications and setting a foundation for future research in Hungarian representation learning. The dataset and the corresponding evaluation code are publicly accessible at https://github.com/margitantal68/hungarian-embeddings.

Index Terms—**Hungarian language, embedding models, question-answer retrieval, vector similarity search.**

I. INTRODUCTION

Embedding models have become fundamental to modern natural language processing (NLP), providing dense vector representations that encode semantic relationships between words, phrases, and documents. The evolution of embedding techniques has progressed from early feedforward neural networks to static models like Word2Vec [1] and FastText [2], and subsequently to dynamic, contextualized embeddings derived from Transformer-based architectures including BERT [3], GPT [4], and T5 [5]. These advances have significantly improved performance across numerous NLP tasks, particularly in high-resource languages such as English.

Margit Antal is with Faculty of Technical and Human Sciences, Sapientia Hungarian University of Transylvania, Târgu-Mures, Romania. (E-mail: manyi@ms.sapientia.ro)

DOI: 10.36244/ICJ.2025.4.1

However, the effectiveness of modern embedding models in morphologically rich, low-resource languages remains insufficiently explored. Hungarian exemplifies these challenges due to its agglutinative morphology, where words can contain multiple morphemes that substantially alter meaning and grammatical function. This morphological complexity often results in performance degradation compared to English [6].

The importance of robust embeddings extends beyond traditional NLP tasks to modern applications such as Retrieval-Augmented Generation (RAG) systems, where embeddings enable efficient knowledge retrieval from large databases by capturing semantic similarity beyond surface-level keyword matching. In these systems, embedding quality directly impacts retrieval accuracy and, consequently, the overall system performance. Despite this critical role, systematic evaluation of state-of-the-art embedding models for Hungarian remains limited.

Previous work on Hungarian embeddings has been sparse and focused primarily on static representations. Gedeon [7] presented the most comprehensive evaluation to date, but concentrated exclusively on static word embeddings, leaving modern contextualized models largely unexplored. To the best of our knowledge, only a single study [8] to date has systematically evaluated embedding models for Hungarian texts, focusing exclusively on the legal domain. However, no comprehensive assessment has yet been conducted for other types of Hungarian texts.

This paper addresses these research gaps by presenting the first comprehensive evaluation of state-of-the-art embedding models for Hungarian question-answer retrieval. Our primary contributions are threefold: (1) we provide a systematic comparison of modern embedding models on Hungarian retrieval tasks, (2) we establish evaluation benchmarks using both domain-specific and general-purpose datasets, and (3) we offer practical guidance for selecting appropriate models for Hungarian NLP applications.

To achieve these objectives, we construct two complementary evaluation datasets. The first comprises domain-specific data extracted from technical documentation, preprocessed into semantically coherent chunks with human-annotated question-answer pairs. The second utilizes the publicly available HuRTE benchmark [9], providing standardized evaluation conditions. We employ Chorma [10] for efficient vector storage and similarity search.

Our evaluation examines eight diverse embedding models, including multilingual transformers (BGE-M3, E5-BASE,

XLMROBERTA, NOMIC), language-specific models (HUBERT), and commercial API solutions (OpenAI, Google). Performance is assessed using established information retrieval metrics: Mean Reciprocal Rank (MRR) and Recall@k. Beyond quantitative analysis, we conduct detailed error analysis to identify failure patterns related to Hungarian morphology, compound word processing, and domain-specific terminology. Additionally, we analyze practical considerations including inference latency, computational requirements, and cost-effectiveness to provide comprehensive guidance for practitioners.

The remainder of this paper is organized as follows: Section 2 reviews related work in multilingual embeddings and Hungarian NLP; Section 3 details our experimental methodology and datasets; Section 4 presents quantitative results and comparative analysis; Section 5 discusses error patterns and morphological challenges; Section 6 discusses the results, highlighting practical trade-offs and deployment considerations, while Section 7 concludes by outlining the implications for future Hungarian NLP research.

II. Related Work

Teaching machines to comprehend human language is a fundamental step in developing intelligent systems, a task often facilitated by word embeddings. These dense vector representations map similar words to similar vectors and are capable of capturing complex semantic relationships. The field has evolved from early feedforward neural networks for language modeling to highly effective static models like Word2Vec and FastText. However, a key limitation of these static embeddings is their inability to capture context-dependent meanings, leading to the development of dynamic, contextualized word embeddings from Transformer-based models such as BERT, GPT, and T5.

For the Hungarian language, which is considered an underrepresented language due to its complex morphology and agglutinative nature, high-quality embedding models are insufficiently evaluated. Few empirical measurements exist to assess embedding model performance specifically for Hungarian, making it difficult for developers of Hungarian Q&A systems to determine which models are best suited for their applications.

Despite the success of dynamic models, static word embeddings remain relevant for various applications due to their lower computational requirements. Research has shown a significant performance drop in Hungarian word analogy tasks compared to English, attributed to the language's high morphological variation and less stable semantic representations.

Several studies have focused on evaluating word embeddings in Hungarian. Gedeon [7] provides a comprehensive analysis of various static word embeddings, including traditional models like Word2Vec and FastText, as well as static embeddings derived from BERT-based models using different extraction methods. For intrinsic evaluation using a word analogy task (measuring embedding quality without using them in a real application), FastText demonstrated superior performance, achieving high accuracy and Mean Reciprocal Rank (MRR) scores. Among the BERT-based models, the X2Static method for extracting static embeddings showed superior performance compared to decontextualized and aggregate methods, approaching the effectiveness of traditional static embeddings. This method leverages contextual information from a teacher model to generate static embeddings, and a Turkish study similarly found X2Static to be the most effective for extracting static embeddings from BERT-based models. For extrinsic evaluation (test embeddings in a real application, such as NER or POS tagging), Gedeon utilized a bidirectional LSTM model for Named Entity Recognition (NER) and Part-of-Speech (POS) tagging tasks. The results indicated that embeddings derived from dynamic models, particularly those extracted using the X2Static method, outperformed purely static embeddings. ELMo embeddings achieved the highest accuracy in both NER and POS tagging, highlighting the benefits of contextualized representations even when used in a static form. ELMo generates contextualized word embeddings using a bidirectional LSTM language model, capturing polysemy and context-dependent meanings.

BERT (Bidirectional Encoder Representations from Transformers) [3] and its derivatives have become central to modern NLP. For Hungarian, huBERT is a state-of-the-art Hungarian cased BERT-base model trained on the Webcorpus 2.0. It has been shown to outperform multilingual BERT models in tasks such as morphological probing, POS tagging, and NER. Nemeskey [11] introduced the huBERT family, which achieved state-of-the-art performance in NER and NP chunking for Hungarian. Another significant Hungarian BERT model is PULI BERT-Large [12], a BERT large model with 345 million parameters. XLM-RoBERTa (XLM-R) is a transformer-based multilingual masked language model that also includes Hungarian data.

The sentence transformers method has gained popularity for creating semantically meaningful sentence embeddings that enable comparison using cosine similarity. This approach, sometimes extended to multilingual models using knowledge distillation, involves a teacher model generating desired sentence embeddings in one language, which a student model then replicates across multiple languages using parallel sentences. Hatvani and Yang [13] addressed the lack of high-quality embedding models for Hungarian in RAG systems. They developed three encoder-only language models: xml_roberta_sentence_hu, hubert_sentence_hu, and minilm_sentence_hu. These models, trained using a distillation method with paraphrase-distilroberta-base-v2 as the teacher model and FLORES-200 and OpenSubtitles corpora, demonstrated substantial improvements in semantic similarity tasks. The hubert_sentence_hu model achieved the highest accuracy and F1-Score on a custom news article test corpus.

Beyond general NLP tasks, Hungarian embedding models have been applied and evaluated in specific domains. Osváth et al. [14] used BERT topic modeling with huBERT and HIL-SBERT embeddings to analyze patient narratives from a Hungarian online forum, identifying major topics and using a fine-tuned BERT model for sentiment analysis. Their findings highlighted dominantly negative sentiments in patient experiences and comments.

Yang and Váradi [15] explored developing deep neural network language models for Hungarian with low computational and data resources. They pre-trained and fine-tuned five transformer models: ELECTRA, ELECTRIC, RoBERTa (small), BART (base), and GPT-2 on various NLP tasks, including sentence-level sentiment analysis, NER, noun phrase chunking, and text summarization. While these experimental models generally did not surpass the state-of-the-art huBERT model in classification tasks, they achieved competitive results with fewer parameters and resources. Notably, their BART model achieved a significantly higher F-score in abstractive summarization compared to huBERT-based tools, and the models offered advantages in terms of smaller carbon footprint and mobile application suitability.

Tóth et al. [16] developed LMEZZ, a learning application to help students with Hungarian sentence analysis based on school grammar rules, utilizing transformer-based BERT models (huBERT and PULI BERT-Large) for improved reliability over convolutional neural network-based SpaCy models.

A recent study [8] presents a semantic search system developed to efficiently identify Hungarian court decisions with similar factual backgrounds. Its primary objective is to retrieve relevant legal precedents by matching court rulings based on semantic similarity, using factual case summaries as queries. The research evaluated twelve embedding models on a corpus of 1,172 Hungarian court decisions. Given that legal documents are typically lengthy—often exceeding the context window of most transformer-based architectures—the authors examined seven different strategies for handling long texts, including simple chunking, striding (overlapping chunks), and Last Chunk Scaling (LCS), which mitigates the overrepresentation of small final segments in the averaged embedding vector. Model performance was assessed using the Mean Reciprocal Rank (MRR) metric. The study found that the Cohere embed-multilingual-v3.0 model achieved the best results, reaching an MRR of 0.95. Notably, this demonstrates that a well-optimized 512-token model can outperform several models with substantially larger context windows (up to 8192 tokens). The authors also evaluated models pretrained specifically for the Hungarian language, including the base huBERT model [11] without fine-tuning, as well as two adapted variants: the sbert_hubert model [14], fine-tuned for sentence-level semantic similarity, and the danieleff model, fine-tuned for question–answering (Q&A) tasks. The danieleff model was trained on 170 question–answer pairs derived from sections (1,000–5,000 characters) of university academic regulations. Among the Hungarian-language models, danieleff achieved the highest performance.

Modern multilingual embedding models have made significant progress in understanding multiple languages simultaneously. Current state-of-the-art models can work with more than 100 languages and perform well on standard evaluation benchmarks [17]. Transformer-based models, especially BERT, XLM-R, and XLM-RoBERTa, have become the most widely used approaches for this task. Several key innovations have improved these models. First, researchers developed methods to adapt models trained on one language to work with others by creating specialized word representations [18]. Second, they created systems that can understand sentences across languages by sharing vocabulary encoding methods, enabling models to work on new languages without additional training [19]. Third, they combined different types of embeddings with improved alignment techniques to better match meanings across languages [20]. These advances have led to efficient multilingual systems that can process long texts (up to 8192 tokens) while maintaining good performance across different language tasks [17]. However, these models have not been thoroughly tested on morphologically complex languages like Hungarian. Most evaluations focus on widely-used languages, which may not reveal the challenges that arise with Hungarian's complex word structure and limited available training data.

III. METHODS

A. Embeddings

Text embeddings are numerical representations of words, sentences, or documents in a continuous vector space. They capture semantic meaning, so texts with similar meanings end up close together in that space, even if they use different wording. In RAG systems, embeddings are crucial because they enable efficient retrieval of relevant knowledge from large databases. Instead of relying only on keyword matching, embeddings allow the system to understand context and intent, leading to more accurate and meaningful results.

Multilingual embeddings extend this capability across languages, mapping semantically similar texts in different languages to nearby positions in the same vector space. This makes it possible for a RAG system to retrieve knowledge in one language and use it to answer questions in another, breaking down language barriers and improving accessibility. In practice, high-quality multilingual embeddings are essential for building global, cross-lingual RAG applications that can serve diverse users and knowledge sources.

Embedder models were utilized in three ways: commercial models accessed via their APIs, and open-source models run either through a local Ollama server or via SentenceTransformers [21], a Python library for generating dense vector representations (embeddings) of sentences, paragraphs, and documents.

In this paper we employed the following embedder models:

- BGE-M3 – bge-m3 [22]: This model offers robust embeddings for multilingual and general-purpose semantic tasks, with emphasis on large-scale retrieval and clustering.
- E5-BASE – intfloat/multilingual-e5-base [23]: The Multilingual E5-Base model is a transformer-based text embedder that generates semantically rich, language-agnostic sentence embeddings across over 100 languages, enabling effective multilingual retrieval, clustering, and semantic similarity tasks.
- GEMINI – gemini-embedding-001 [24]: Text embeddings were generated using the *Gemini-embedding-001* model, which produces 768-dimensional vectors. To ensure consistency with other embedding models in our evaluation, the input-type parameter was not specified.

TABLE I
COMPARISON OF POPULAR EMBEDDING MODELS BY USAGE MODE, DIMENSION, SEQUENCE LENGTH, DOMAIN, AND SIZE.

Model Name	Usage	Dimension	Sequence Length (#tokens)	Domain	Model Size
BGE-M3	SentenceTransformer - local	1024	8192	Multilingual	≈ 560M
E5-BASE	SentenceTransformer - local	768	512	Multilingual	≈ 278M
GEMINI	Google API	768	2048	General	Undisclosed
HUBERT	SentenceTransformer - local	768	512	Hungarian	≈ 110M
NOMIC	Ollama - local	768	2048	General	≈ 137M
OPENAI-3SMALL	OpenAI API	1536	8192	General	Undisclosed
OPENAI-ADA	OpenAI API	1536	8192	General	Undisclosed
XLMROBERTA	SentenceTransformer - local	768	128	Multilingual	≈ 270M

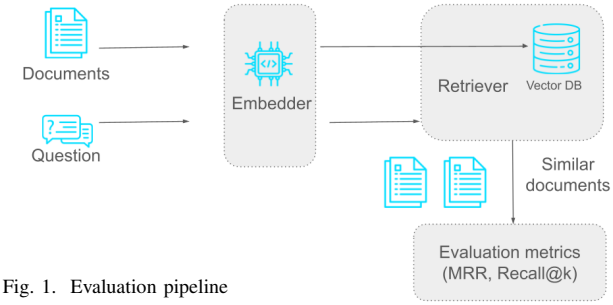


Fig. 1. Evaluation pipeline

The embeddings were obtained through the Gemini API's v1beta endpoint.

- HUBERT – danieleff/hubert-base-cc-sentence-transformer [8]: This model model was fine-tuned on 170 Hungarian question–answer pairs derived from sections of university academic regulations ranging from 1,000 to 5,000 characters in length.
- NOMIC – nomic-embed-text-v1 [25]: Designed for general-purpose embeddings, NOMIC excels in large-scale retrieval, clustering, and semantic search tasks with high efficiency.
- OPENAI-3SMALL – text-embedding-3-small: A lightweight and cost-effective embedding model from OpenAI's API v1, optimized for production-scale semantic tasks where efficiency and performance must be balanced.
- OPENAI-ADA – text-embedding-ada-002: A versatile embedding model available through OpenAI's API v1, widely adopted for applications such as semantic search, clustering, classification, and recommendation systems, supporting a broad range of text inputs.
- XLMROBERTA – paraphrase-xlm-r-multilingual-v1 [26], [27]: Tis model is a multilingual Sentence-Transformer based on XLM-RoBERTa, designed to produce high-quality, language-agnostic sentence embeddings for over 50 languages, optimized for tasks like semantic similarity and multilingual retrieval.

Table I presents the key characteristics of the embedding models.

B. Evaluation pipeline

The evaluation pipeline is illustrated in Fig. 1. During the ingestion stage, documents are transformed into vector representations using an embedding model and subsequently stored in a vector database, specifically Chroma in our implementation. Following ingestion, the evaluation of a question dataset proceeds through three steps: (1) vectorization of the input questions, (2) retrieval of the top-k most similar documents for each question, and (3) computation of retriever performance metrics.

As a baseline for comparison with semantic search, we incorporated a keyword-based retrieval method. Specifically, we employed the BM25 [28] algorithm to retrieve the top-k most relevant documents.

C. Datasets

1) Clearservice: The Clearservice dataset[1] is a custom-made dataset created from the data of the company of the same name. The dataset consists of two parts: (1) A file called topics.txt, which groups the data into topics and serves as the search space. (2) A set of questions in cs_qa.csv, containing 50 questions. Each question is associated with a specific topic and can be answered based on it. For each question, the corresponding topic and a reference answer are provided. The reference answer, however, is not used in this study.

2) HuRTE: The HuRTE dataset[2] is the Hungarian adaptation of the Recognizing Textual Entailment (RTE) corpora originally included in the GLUE benchmark. It forms part of the Hungarian Language Understanding Evaluation Benchmark Kit (HuLU) [29] [9] and was created through translation and re-annotation of the English RTE instances. The dataset consists of 4,504 examples, each comprising a premise — sometimes a multi-sentence passage — and a single-sentence hypothesis, with the task being to determine whether the premise entails the hypothesis. This is framed as a binary classification problem, where labels indicate entailment ("1") or non-entailment ("0"). The corpus is divided into training (2,132 instances), validation (243 instances), and test splits; however, test labels are not provided. The data is distributed in JSON format, with each entry containing an identifier, a premise, a hypothesis, and the corresponding label.

We measure the quality of retrieval using two types of evaluations. The *HuRTE-Positive* evaluation is performed using only the positive examples (label 1) both in the index and in the question evaluation. In this setting, the training set contains 1,092 positive examples, and the validation set contains 135. The hypothesis sentences are searched for within the premise texts, and retrieval quality is assessed accordingly.

The *HuRTE-All* evaluation is performed using all examples in the index, while still evaluating the questions using only

[1]https://github.com/margitantal68/hungarian-embeddings/tree/master/data/clearservice
[2]https://github.com/nytud/HuRTE

the positive examples. This allows us to assess retrieval performance in a more realistic setting, where irrelevant data is present in the index, but only the positives matter for evaluation.

D. Metrics

We evaluated retrieval performance using Mean Reciprocal Rank (MRR) and Recall. MRR measures the rank position of the first relevant document, averaged across all queries. For each (query, document) pair, documents were retrieved using semantic search in Chroma, and the rank of the corresponding ground-truth context was recorded. Recall@1 and Recall@3 capture the proportion of queries for which the correct context appears within the top one or top three retrieved results, respectively.

IV. RESULTS

All measurements were conducted on a MacBook Pro equipped with an Apple M1 Pro processor and 32 GB of unified memory, running macOS Sequoia version 15.7.1. The experiments involving the Nomic embedder utilized the Ollama runtime (version 0.12.6).

TABLE II
EMBEDDING MODELS EVALUATION ON CLEARSERVICE DATASET.

Embedder	MRR	Recall@1	Recall@3
BGE-M3	0.90	0.86	0.96
E5-BASE	0.79	0.70	0.92
GEMINI	0.87	0.78	0.98
HUBERT	0.78	0.74	0.84
NOMIC	0.71	0.64	0.80
OPENAI-3SMALL	0.80	0.70	0.94
OPENAI-ADA	0.80	0.72	0.90
XLMROBERTA	0.90	0.86	0.96
BM25	0.77	0.68	0.80

TABLE III
COMPARISON OF MODELS ON HURTE DATASET *HuRTE-Positive* EVALUATION.

Model	MRR		Recall@1		Recall@3	
	Val	Train	Val	Train	Val	Train
BGE-M3	0.98	0.89	0.96	0.82	1.00	0.97
E5-BASE	0.93	0.84	0.90	0.77	0.97	0.92
GEMINI	0.99	0.91	0.97	0.85	1.00	0.98
HUBERT	0.82	0.63	0.77	0.53	0.88	0.74
NOMIC	0.90	0.72	0.85	0.65	0.95	0.80
OPENAI-3SMALL	0.94	0.85	0.92	0.78	0.97	0.92
OPENAI-ADA	0.94	0.84	0.91	0.78	0.98	0.92
XLMROBERTA	0.94	0.82	0.91	0.75	0.98	0.91
BM25	0.82	0.72	0.78	0.64	0.84	0.79

We applied our evaluation pipeline to both the Clearservice and HuRTE datasets. For HuRTE, we conducted two types of evaluations: *HuRTE-Positive*, which uses only the positive examples in both the index and the evaluation, and *HuRTE-All*, which uses all examples in the index while evaluating only the positive questions. Each type of evaluation was performed separately on the validation set (243 samples, 135 positives) and the training set (2132 samples, 1092 positives), allowing us to analyze how performance generalizes from a smaller dataset to a larger one of the same type.

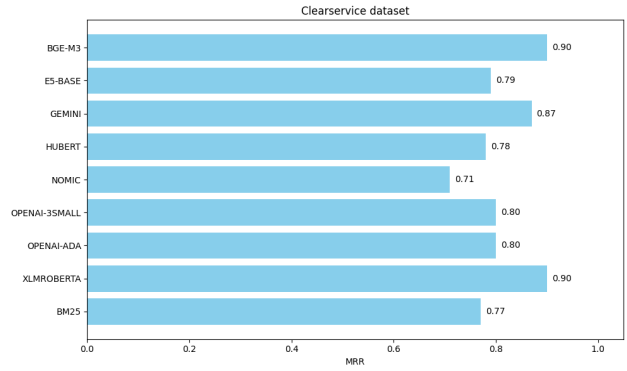


Fig. 2. Models' performance on the Clearservice dataset using the MRR metric.

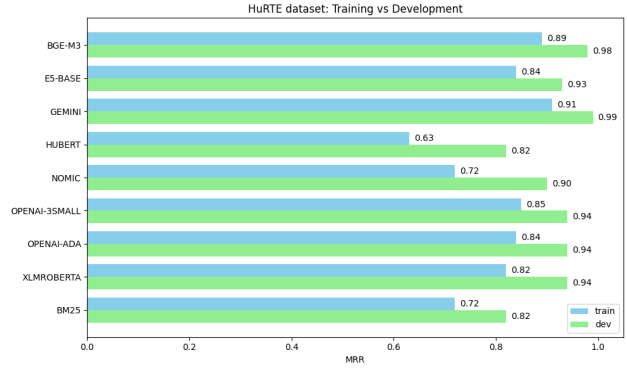


Fig. 3. Model performance on the HuRTE dataset (*HuRTE-Positive* evaluation) using the MRR metric.

The evaluation protocol is described in III-B. The results are summarized in the following tables: Table II presents the outcomes for the Clearservice dataset, while Table III reports the results for the HuRTE-Positive dataset.

Among the available metrics, MRR (Mean Reciprocal Rank) was chosen for visual representation, as it reflects both the position and relevance of the first correct result, providing a more informative measure of retrieval effectiveness than Recall@1 or Recall@3. Figs. 2, 3, and 4 show visual representations of the results.

A. Analysis of Model Performance

Two types of time measurements were performed: index build time, representing the time required to create the model index, and average query latency, indicating the response time per query. The results are shown in Fig. 5.

HUBERT and XLMROBERTA achieved the best overall time performance, with both low build times and minimal latency. E5-BASE also performed efficiently across both metrics. GEMINI, OPENAI-3SMALL, and OPENAI-ADA exhibited notably higher query latencies despite moderate build times, likely due to API communication overhead. NOMIC, run locally via Ollama, and BGE-M3 had longer index build times but maintained low query latency. Overall, HUBERT

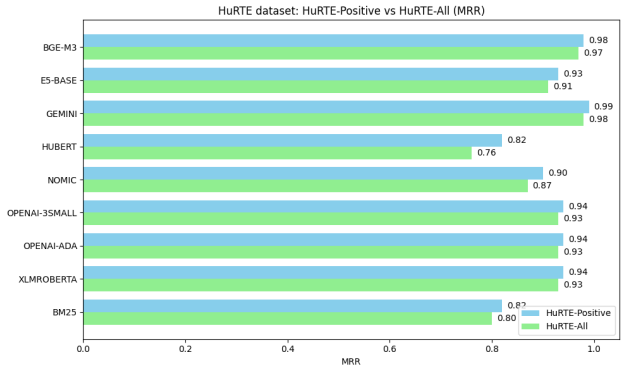


Fig. 4. Model performance on the HuRTE dataset (validation subset), comparing *HuRTE-Positive* and *HuRTE-All* evaluations using the MRR metric.

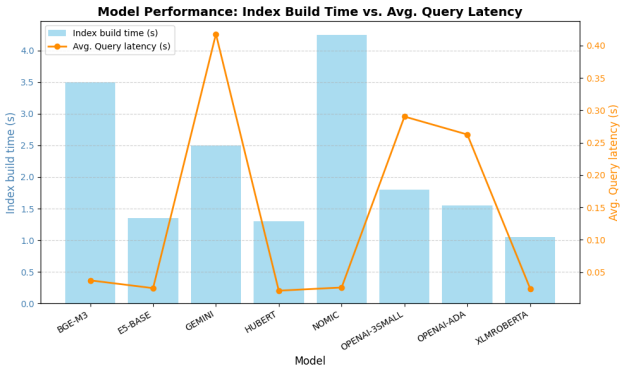


Fig. 5. Index build time vs. average query latency evaluated on the HuRTE dataset, validation subset, (*HuRTE-Positive* evaluation).

and XLMROBERTA demonstrated the best balance between index setup efficiency and query responsiveness.

We evaluated both the retrieval quality and efficiency of embedding models, which are critical for performance in RAG systems. By comparing MRR (ranking quality) against query latency and index build time, this reveals how effectively each model balances accuracy with speed. Models positioned toward the top-left of the diagram achieve the best trade-off, offering high-quality retrieval with minimal response time. The results are shown in Fig. 6.

V. Error Analysis

We conducted a comprehensive error analysis on the Clearservice dataset. The questions and their corresponding error rates are presented in Appendix A.

Our analysis revealed that some errors are systematic—where seven out of eight embedding models failed—while others are occasional. The failures can be broadly categorized into three groups: (i) synonyms and paraphrases - The embeddings failed to recognize equivalence between different phrasings of the same concept. (ii) Overlapping topics - Relevant information appears across multiple sections, leading to confusion between semantically related topics. (iii) Domain-specific terminology - Specialized vocabulary (e.g., HR or legal terms) was not consistently captured by the embeddings.

In the following, we analyze the questions with the highest error rates: Q23 and Q9.

Q23: *Milyen elvárás van a munkavégzéssel kapcsolatban?*(What are the expectations regarding the work?) - error_rate = 0.875. This question includes the terms elvárás (expectation) and munkavégzés (work). The correct match is Topic 6 – *Munkavégzés* (Work), but Topic 9 – *Elvárások és Dokumentáció* (Expectations and Documentation) is also semantically close, even though it refers to application requirements rather than work performance. This semantic proximity likely caused confusion among the embedding models.

Q9: *Mi jár vasárnapi és ünnepnapi munkára?* (What is provided for Sunday and holiday work?) — error_rate = 0.625. The ground-truth topic is Topic 2 – *Fizetés* (Salary). The question is framed in terms of benefits (mi jár), while the relevant text specifies compensation percentages. Embeddings may incorrectly associate it with Topic 10 – *Szabadság és Hazautazás* (Vacation and Travel Home) due to the lexical overlap with ünnep (holiday).

To further investigate these confusions, we computed the Recall@3 confusion matrix, which measures how often embedding models retrieved the correct topic among their top three results. For each question, the ground-truth topic was compared against the top three retrieved topics, and the results were aggregated into a matrix with true topics as rows and retrieved topics as columns. Off-diagonal entries highlight frequent mismatches between semantically related topics.

By visualizing this matrix as a heatmap, we identified which topics are most frequently confused, revealing systematic weaknesses such as overlapping categories, synonym mismatches, and domain-specific ambiguities. Confusion matrices for all embedding models are shown in Appendix A.

VI. Discussion

Our evaluation of eight embedding models and BM25 across two Hungarian-language datasets reveals important insights into retrieval performance for domain-specific applications.

BGE-M3 and XLMROBERTA emerged as top performers on the Clearservice dataset, both achieving an MRR of 0.90 and Recall@1 of 0.86. GEMINI demonstrated the strongest performance on HuRTE-Positive (MRR: 0.99 validation, 0.91 training), followed closely by BGE-M3 (MRR: 0.98 validation, 0.89 training). The consistent performance gap between validation and training sets suggests that model behavior generalizes well from smaller to larger datasets of similar characteristics.

The traditional BM25 baseline achieved competitive results (MRR: 0.77 on Clearservice), outperforming NOMIC and matching HUBERT on certain metrics, demonstrating that lexical matching remains valuable for Hungarian text retrieval. However, neural embedding models consistently surpassed BM25, particularly on Recall@3 metrics.

Efficiency analysis revealed critical trade-offs between accuracy and speed. HUBERT and XLMROBERTA offered the best balance, with low index build times and minimal query latency. While GEMINI achieved superior retrieval quality,

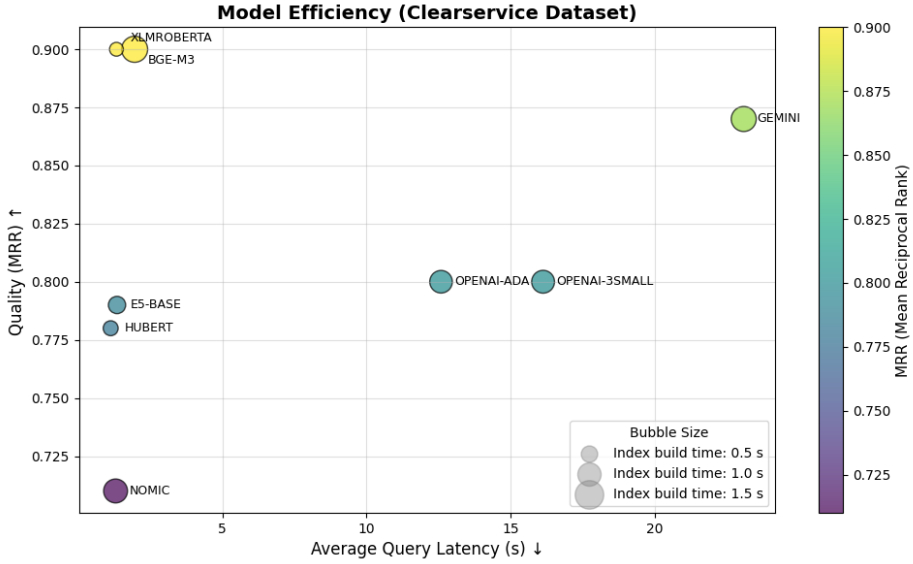


Fig. 6. Model efficiency on the Clearservice dataset using the MRR metric.

it exhibited significantly higher query latency due to API communication overhead, making it less suitable for real-time applications. BGE-M3, despite longer index build times, maintained competitive query latency while delivering top-tier accuracy.

The error analysis on Clearservice exposed three primary failure modes: (i) synonyms and paraphrases, where embeddings failed to recognize semantic equivalence; (ii) overlapping topics, particularly when relevant information spans multiple sections; and (iii) domain-specific terminology, especially HR and legal vocabulary.

Systematic errors, where seven of eight models failed, highlight fundamental limitations in capturing Hungarian domain-specific semantics. The confusion between Topics 6 and 9 (Q23) and between Topics 2 and 10 (Q9) demonstrates that lexical overlap and semantic proximity can mislead even state-of-the-art embeddings. The Recall@3 confusion matrices further confirm these patterns, revealing consistent misclassifications between semantically related topics.

VII. Conclusions

In this paper, we conducted a comprehensive analysis of embedding models for Hungarian texts. Eight embedding models were evaluated against a baseline lexical search in the context of an information retrieval task for a Q&A system.

Our findings demonstrate that BGE-M3 and XLM-ROBERTA offer the best overall performance for Hungarian text retrieval, balancing high accuracy (MRR: 0.90) with operational efficiency. While GEMINI achieves superior accuracy, it comes at the cost of increased latency, making the choice between these models dependent on specific application requirements.

For production RAG systems, the trade-off between accuracy and speed is critical. HUBERT and XLMROBERTA provide optimal latency profiles, while BGE-M3 offers a strong middle ground for applications that can tolerate longer index build times in exchange for improved retrieval quality. This efficiency analysis is particularly valuable for practitioners deploying real-time information retrieval systems.

The systematic errors observed across models indicate that domain-specific Hungarian terminology and subtle semantic distinctions remain challenging for current embedding approaches. These persistent challenges suggest that future work should focus on fine-tuning strategies that explicitly incorporate domain knowledge and synonym relationships to better capture the nuances of specialized vocabulary.

While the evaluation provides valuable insights into Hungarian embedding-based retrieval, the limited size of the Clearservice dataset and the entailment-only focus of the HuRTE subset constrain the generalizability of the results. Future work will address these limitations by expanding the domain-specific dataset and incorporating more diverse and balanced retrieval benchmarks to ensure broader applicability.

Declarations

Declaration of generative AI and AI-assisted technologies in the writing process: During the preparation of this work the author used ChatGPT in order to refine language. After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

References

[1] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *Proceedings of Workshop at ICLR*, vol. 2013, 01 2013.

[2] A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov, "Bag of tricks for efficient text classification," in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, M. Lapata, P. Blunsom, and A. Koller, Eds. Valencia, Spain: Association for Computational Linguistics, Apr. 2017, pp. 427–431. [Online]. Available: https://aclanthology.org/E17-2068/

[3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, J. Burstein, C. Doran, and T. Solorio, Eds. Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. [Online]. Available: https://aclanthology.org/N19-1423

[4] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving language understanding by generative pre-training," OpenAI, San Francisco, CA, USA, Tech. Rep., 2018, openAI Technical Report. [Online]. Available: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf

[5] 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," 2019. [Online]. Available: https://arxiv.org/abs/1910.10683

[6] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," 2017. [Online]. Available: https://arxiv.org/abs/1607.04606

[7] M. Gedeon, "A comparative analysis of static word embeddings for hungarian," *Infocommunications Journal*, vol. XVII, pp. 28–34, 06 2025. [Online]. Available: https://doi.org/10.36244/ICJ.2025.2.4

[8] G. M. Csányi, D. Lakatos, J. P. Vadász, D. Nagy, and I. Üveges, "A kontextusablakon kihajolni nem veszélyes: jogi szövegek hatékony szemantikus keresése," in *Proceedings of the XXI. Magyar Számítógépes Nyelvészeti Konferencia*, Szeged, Hungary, February 6–7 2025, pp. 3–17, mONTANA Knowledge Management Ltd.; National University of Public Service.

[9] N. Ligeti-Nagy, G. Ferenczi, E. Héja, L. J. Laki, N. Vadász, Z. G. Yang, and T. Váradi, "HuLU: Hungarian language understanding benchmark kit," in *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, and N. Xue, Eds. Torino, Italia: ELRA and ICCL, May 2024, pp. 8360–8371. [Online]. Available: https://aclanthology.org/2024.lrec-main.733

[10] ChromaDB Team, "Chroma: Open-source embedding and vector database," https://github.com/chroma-core/chroma, 2023, version 1.3.0. Available at https://www.trychroma.com. Licensed under Apache 2.0.

[11] D. M. Nemeskey, "Introducing hubert," in *XVII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY2021)*, 2021, pp. 3–14.

[12] Z. G. Yang, R. Dodé, G. Ferenczi, E. Héja, K. Jelencsik-Mátyus, Kőrös, L. J. Laki, N. Ligeti-Nagy, N. Vadász, and T. Váradi, "Jönnek a nagyok! bert-large, gpt-2 és gpt-3 nyelvmodellek magyar nyelvre," in *XIX. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2023)*. Szegedi Tudományegyetem, Informatikai Intézet, 2023, pp. 247–262.

[13] P. Hatvani and Z. Yang, "Training embedding models for hungarian," in *Proceedings of the 3rd Conference on Information Technology and Data Science (CITDS)*, 08 2024, pp. 1–6.

[14] M. Osváth, Z. Yang, and K. Kósa, "Analyzing narratives of patient experiences: A bert topic modeling approach," *Acta Polytechnica Hungarica*, vol. 20, pp. 153–171, 01 2023.

[15] Z. Yang and T. Váradi, "Training experimental language models with low resources, for the hungarian language," *Acta Polytechnica Hungarica*, vol. 20, pp. 169–188, 01 2023.

[16] N. Tóth, B. Oszkó, and Z. Yang, "Hungarian sentence analysis learning application with transformer models," *Acta Cybernetica*, vol. 27, pp. 83–91, 03 2025.

[17] J. Chen, S. Xiao, P. Zhang, K. Luo, D. Lian, and Z. Liu, "Bge m3-embedding: Multi-lingual, multi-functionality, multi-granularity text embeddings through self-knowledge distillation," 2024. [Online]. Available: https://arxiv.org/abs/2402.03216

[18] M. Artetxe, S. Ruder, and D. Yogatama, "On the cross-lingual transferability of monolingual representations," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, D. Jurafsky, J. Chai, N. Schluter, and J. Tetreault, Eds. Online: Association for Computational Linguistics, Jul. 2020, pp. 4623–4637. [Online]. Available: https://aclanthology.org/2020.acl-main.421/

[19] M. Artetxe and H. Schwenk, "Massively multilingual sentence embeddings for zero-shot cross-lingual transfer and beyond," *Transactions of the Association for Computational Linguistics*, vol. 7, pp. 597–610, 09 2019. [Online]. Available: DOI: 10.1162/tacl_a_00288

[20] T. Schuster, O. Ram, R. Barzilay, and A. Globerson, "Cross-lingual alignment of contextual word embeddings, with applications to zero-shot dependency parsing," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, J. Burstein, C. Doran, and T. Solorio, Eds. Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 1599–1613. [Online]. Available: https://aclanthology.org/N19-1162/

[21] N. Reimers and I. Gurevych, "Sentence-bert: Sentence embeddings using siamese bert-networks," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. *Association for Computational Linguistics*, 11 2019. [Online]. Available: http://arxiv.org/abs/1908.10084

[22] J. Chen, S. Xiao, P. Zhang, K. Luo, D. Lian, and Z. Liu, "M3-embedding: Multi-linguality, multi-functionality, multi-granularity text embeddings through self-knowledge distillation," in *Findings of the Association for Computational Linguistics: ACL 2024*, L.-W. Ku, A. Martins, and V. Srikumar, Eds. Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 2318–2335. [Online]. Available: https://aclanthology.org/2024.findings-acl.137/

[23] L. Wang, N. Yang, X. Huang, L. Yang, R. Majumder, and F. Wei, "Multilingual e5 text embeddings: A technical report," arXiv preprint arXiv:2402.05672, 2024, model "multilingual-e5-base" released on Hugging Face, intfloat/multilingual-e5-base.

[24] J. Lee, F. Chen, S. Dua, D. Cer, M. Shanbhogue, I. Naim, G. H. Ábrego, Z. Li, K. Chen, H. S. Vera, X. Ren, S. Zhang, D. Salz, M. Boratko, J. Han, B. Chen, S. Huang, V. Rao, P. Suganthan, F. Han, A. Doumanoglou, N. Gupta, F. Moiseev, C. Yip, A. Jain, S. Baumgartner, S. Shahi, F. P. Gomez, S. Mariserla, M. Choi, P. Shah, S. Goenka, K. Chen, Y. Xia, K. Chen, S. M. K. Duddu, Y. Chen, T. Walker, W. Zhou, R. Ghiya, Z. Gleicher, K. Gill, Z. Dong, M. Seyedhosseini, Y. Sung, R. Hoffmann, and T. Duerig, "Gemini embedding: Generalizable embeddings from gemini," 2025. [Online]. Available: https://arxiv.org/abs/2503.07891

[25] Z. Nussbaum, J. X. Morris, B. Duderstadt, and A. Mulyar, "Nomic embed: Training a reproducible long context text embedder," 2024.

[26] N. Reimers and I. Gurevych, "Sentence-bert: Sentence embeddings using siamese bert-networks," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11 2019. [Online]. Available: http://arxiv.org/abs/1908.10084

[27] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, and V. Stoyanov, "Unsupervised cross-lingual representation learning at scale," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 2020. [Online]. Available: https://arxiv.org/abs/1911.02116

[28] S. Robertson and H. Zaragoza, "The probabilistic relevance framework: Bm25 and beyond," *Found. Trends Inf. Retr.*, vol. 3, no. 4, p. 333–389, Apr. 2009. [Online]. Available: DOI: 10.1561/1500000019

[29] N. Ligeti-Nagy, G. Ferenczi, E. Héja, K. Jelencsik-Mátyus, L. J. Laki, N. Vadász, Z. G. Yang, and T. Váradi, "Hulu: magyar nyelvű benchmark adatbázis kiépítése a neurális nyelvmodellek kiértékelése céljából," in *XVIII. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem, Informatikai Intézet, 2022, pp. 431–446.

Margit Antal received her M.Sc. degree in Computer Science from Babeş–Bolyai University in 1991 and her Ph.D. in Electronics and Telecommunications from the Technical University of Cluj-Napoca in 2006. Since 2016, she has been working as an Associate Professor in the Mathematics–Informatics Department at the Sapientia Hungarian University of Transylvania. Her main research interests include biometrics, pattern recognition, and representation learning. She has published several research papers in international and national journals. She is an editor of the Acta Universitatis Sapientiae, Informatica journal.

TABLE IV
RETRIEVAL ERROR RATES FOR THE CLEARSERVICE DATASET

Index	Question	Error Rate
23	*Milyen elvárás van a munkavégzéssel kapcsolatban?*	0.875
9	*Mi jár vasárnapi és ünnepnapi munkára?*	0.625
5	*Milyen szállodákban biztosít munkát a cég?*	0.250
15	*Kiutazás előtt mit kell teljesíteni?*	0.250
26	*Hogyan biztosítják a munkaegyenlőséget?*	0.250
39	*Milyen személyazonosító okmány szükséges?*	0.250
48	*Hol van a munkavállaló hivatalosan bejelentve?*	0.250
2	*Hány magyar munkavállaló dolgozik jelenleg a cégben?*	0.125
10	*A magyar alapbér előleg?*	0.125
12	*Milyen típusú lakásokban szállásolják el a dolgozókat?*	0.125
17	*Ki biztosítja a nyelvoktatást?*	0.125
19	*Milyen pozíciók érhetők el?*	0.125
24	*Milyen egészségügyi állapot kizáró ok?*	0.125
25	*Ki állapítja meg az egészségügyi alkalmasságot?*	0.125
28	*Mennyi ideig tart a tréning?*	0.125
30	*Ki fizeti a tréninget?*	0.125
35	*Van lehetőség hétvégén hazautazni?*	0.125
38	*Milyen munkaviszony szükséges az elmúlt egy évben?*	0.125
40	*Milyen erkölcsi feltétel van?*	0.125
41	*Hogyan kell felmondani a meglévő munkahelyen?*	0.125

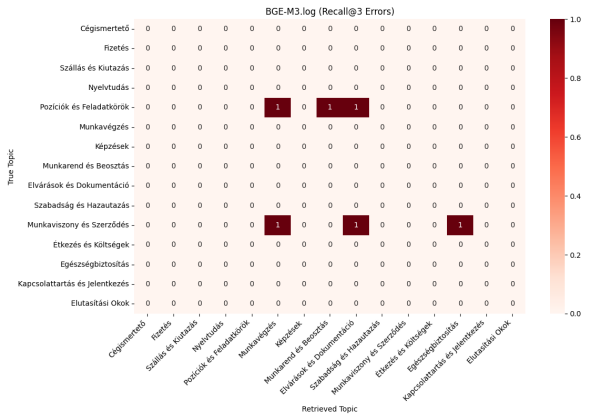


Fig. 7. Confusion matrix for the BGE-M3 model

APPENDIX

Fig. 8. Confusion matrix for the E5-BASE model

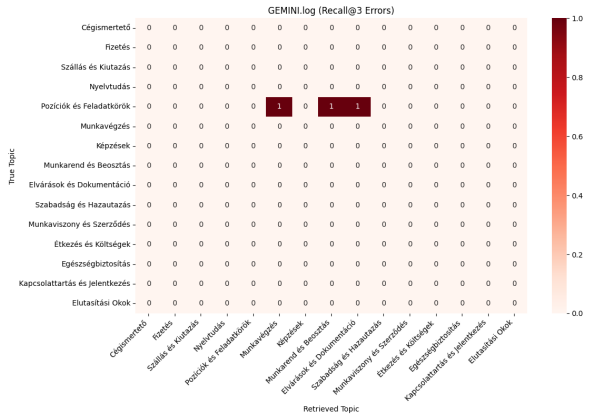


Fig. 9. Confusion matrix for the GEMINI model

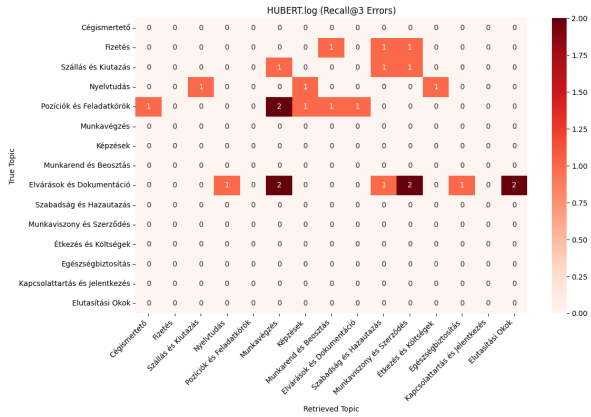


Fig. 10. Confusion matrix for the HUBERT model

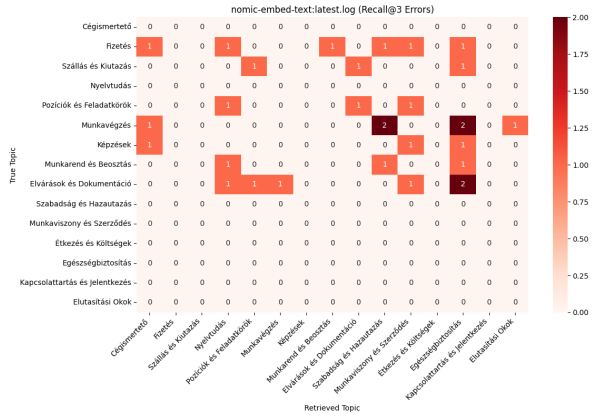


Fig. 11. Confusion matrix for the NOMIC model

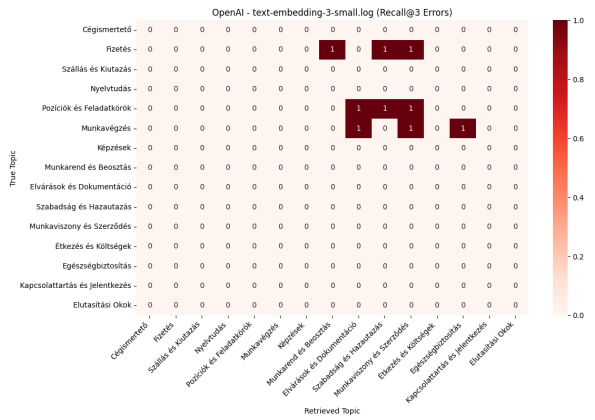


Fig. 12. Confusion matrix for the OPENAI-3SMALL model

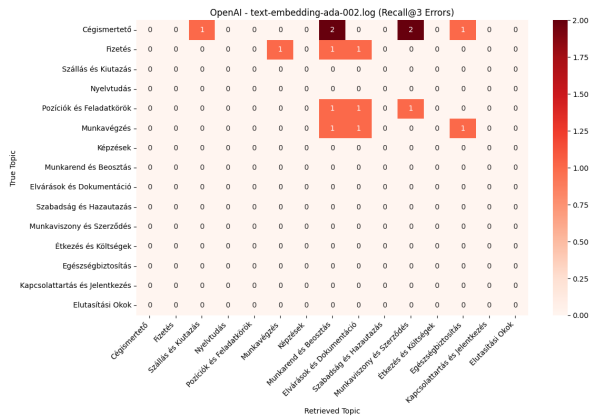


Fig. 13. Confusion matrix for the OPENAI-ADA model

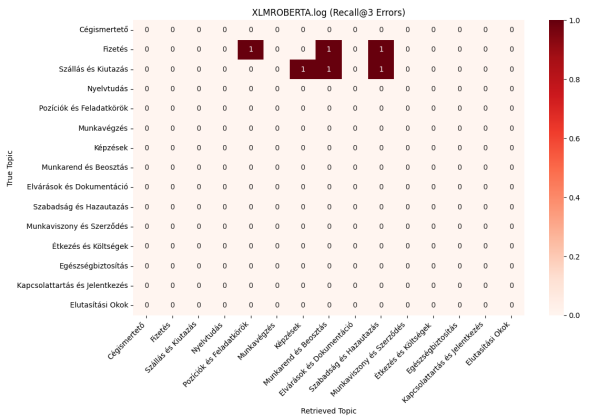


Fig. 14. Confusion matrix for the XLMROBERTA model

PAPR Reduction in OTSM Systems: A Comparative Analysis of SLM Techniques with Novel Phase Matrix Designs

Hsin-Ying Liang and Chuan-Bi Lin

Abstract—**Orthogonal Time Sequency Multiplexing (OTSM) represents a pivotal advancement in wireless communication technology. Nevertheless, its high Peak-to-Average Power Ratio (PAPR) imposes significant constraints on its practical applications and future development. The definition of PAPR refers to the ratio of the maximum instantaneous power to the average power of a signal, and it is commonly used to assess the performance of high-power amplifiers. When PAPR values are excessively high, they reduce the efficiency of high-power amplifiers and increase the complexity of the transmission system. To mitigate this challenge, this paper explores and evaluates the efficacy of the Selective Mapping (SLM) technique for enhancing PAPR performance in OTSM systems. Leveraging the unique two-dimensional data structure inherent to OTSM, a specialized SLM approach is introduced in this paper. The proposed SLM method incorporates a Phase Generation Mechanism (PGM) that utilizes a pre-constructed perturbation phase matrix. This matrix undergoes cyclic shifts to produce multiple perturbation phase matrices. To assess the effectiveness of the proposed SLM technique, this paper investigates three distinct perturbation phase matrix generation mechanisms: Zadoff-Chu Transform (ZCT) matrices, Discrete Cosine Transform (DCT) matrices, and Randomly Generated Phase (RGP) matrices. Additionally, for evaluating PAPR performance improvement, the Complementary Cumulative Distribution Function (CCDF) is used, a statistical method that estimates the probability of high PAPR occurrences. Simulation results indicate that the RGP-based phase generation mechanism consistently outperforms the other methods in achieving significant PAPR reduction.**

Index Terms—**OTSM, PAPR, ZCT matrices, DCT matrices.**

I. INTRODUCTION

Addressing the stringent demands of 5G and future wireless communication for high throughput and enhanced spectral efficiency necessitates advanced physical layer technologies[1][2]. Filter Bank Multi-Carrier (FBMC) modulation, a key 5G enabler, improves spectral efficiency and bandwidth conservation over Orthogonal Frequency Division Multiplexing (OFDM), especially in SUI-6 multipath fading channels, where it can outperform OFDM by up to 20%[1]. Concurrently, Sparse Code Multiple Access (SCMA)[2], a non-orthogonal multiple access (NOMA) technique, augments multi-user processing capability. SCMA codebook design utilizing chaotic interleaving based on Arnold's Cat map can reduce computational complexity by up to 32% for $M = 16$ codewords while maintaining performance[2]. For high-mobility wireless communication environments demanding superior transmission stability and reduced system complexity, Orthogonal Time Sequency Multiplexing (OTSM) is introduced to fulfill these specialized requirements[3][4][5].

Orthogonal Time Sequency Multiplexing (OTSM) is a novel single-carrier modulation technology designed for high-mobility wireless communication environments. OTSM not only achieves comparable performance to Orthogonal Time Frequency Space (OTFS) modulation but also significantly reduces the system's implementation complexity [3][4][5]. The core concept of OTSM is multiplexing in the delay-sequency domain, utilizing Walsh-Hadamard Transform (WHT) to convert signals into the delay-time domain, and finally transmitting and receiving signals in the time domain. This modulation technique demonstrates exceptional transmission stability in high-mobility channels while maintaining system implementation simplicity.

At the transmitter, OTSM arranges information symbols into a two-dimensional delay-sequency domain matrix. Each row undergoes WHT to transform it into the delay-time domain, and the transformed matrix rows are sequentially transmitted. At the receiver, the received time-domain signals are reconstructed into a delay-time domain matrix, and each row is processed by WHT to recover the original delay-sequency domain symbols. By leveraging the simplicity of WHT, which only requires addition and subtraction operations, OTSM achieves significant simplification in modulation and demodulation compared to OTFS, which relies on Fast Fourier Transform (FFT). In addition to reduced computational complexity, OTSM also exhibits similar performance to OTFS in high-mobility scenarios and outperforms conventional Orthogonal Frequency Division Multiplexing (OFDM) [3][4][5]. OTSM can be seamlessly integrated with existing OFDM systems, where Inverse WHT (IWHT) and FFT are used to generate time-frequency signals for OFDM transmission. Compared to OTFS, OTSM can be seen as a simplified version, primarily replacing FFT with WHT for domain transformations. Notably, both techniques effectively separate delay and Doppler effects at the receiver, granting them significant advantages in high-mobility channels. Moreover, WHT's low complexity makes OTSM particularly suitable for future high-mobility communication systems sensitive to computational load, showcasing substantial application potential.

Peak-to-Average Power Ratio (PAPR) is a critical performance metric in wireless communication systems[6][7][8][9].

H. -Y. Liang was with the Department of Information and Communication Engineering, Chaoyang University of Technology, Wufeng District, Taichung, 413310 Taiwan, R.O.C. (e-mail: hyliang@gm.cyut.edu.tw).

H. -Y. Liang and C. -B. Lin are with Chaoyang University of Technology.

DOI: 10.36244/ICJ.2025.4.2

It is defined as the ratio of the maximum instantaneous power of a signal to its average power. Signals with high PAPR are prone to nonlinear distortions when passed through high-power amplifiers (HPAs), thereby affecting energy efficiency and system performance. Consequently, effective PAPR reduction is essential to improving system robustness and influences the design, cost, and efficiency of HPAs. While both OTSM and OTFS are considered promising candidate technologies for future wireless systems, one shared drawback is their inherently high PAPR. In general, the PAPR characteristics of OTSM systems are closely related to the arrangement of data symbols within the delay-sequency grid and the properties of the WHT. In contrast, the PAPR characteristics of OTFS systems are primarily influenced by how data symbols in the delay-Doppler domain are expanded into the time-frequency domain via unitary transformations, typically involving FFT operations. Since traditional PAPR reduction techniques are designed specifically for OFDM-based systems, they are not directly applicable to OTFS or OTSM due to fundamental differences in modulation and signal structure. This limitation has prompted growing interest in adapting or re-designing PAPR reduction schemes tailored specifically for OTFS and OTSM systems. In addition to PAPR reduction, minimizing bit error rate (BER) and computational complexity (CC) are also active areas of research within the OTFS and OTSM domains. Nevertheless, the primary focus of the present paper is to investigate a PAPR reduction technique specifically designed for OTSM systems. This work does not aim to directly optimize BER or computational complexity. Instead, it proposes a method for enhancing PAPR performance in OTSM systems, providing a foundation for future research that may explore integrated solutions capable of jointly improving PAPR and BER performance while maintaining low computational complexity. PAPR is one of the primary drawbacks of OTSM and is a critical factor affecting system performance and efficiency[7][8][9]. High PAPR primarily results in nonlinear distortion, decreased transmission efficiency, and elevated power consumption, all of which significantly constrain the practical implementation of OTSM. In addressing PAPR reduction, Al Ahsan et al. [7] proposed an Adapted Tone Reservation (A-TR) method and analyzed its performance in the delay domain (A-TR-DD), frequency domain (A-TR-SD), and delay-frequency domain (A-TR-BD). The study indicated that A-TR techniques effectively reduce PAPR while maintaining good BER performance. However, the authors focused solely on A-TR methods without considering alternative PAPR reduction techniques. To address this, Neelam and Sahu [8] proposed an SIP-based method that reduces pilot data interference (PDI) and inter-block interference (IBI) through effective channel estimation and data detection techniques. They also analyzed the linear relationship between PAPR and the number of frequency-domain resource blocks, but the design and parameter optimization of the SIP technique remain unexplored. Additionally, Doosti-Aref et al. [9] introduced PSeIM-OTSM, which leverages index modulation to further reduce PAPR and BER, improving energy efficiency. However, its adaptability to more complex scenarios, such as MIMO systems or multi-user environments, has yet to be studied.

Based on the aforementioned literature, OTSM remains an active field of research, with numerous new discoveries being published regularly. As indicated in the aforementioned literature, research aimed at simultaneously optimizing both the PAPR and BER performance of OTSM systems presents significant challenges. Optimizing various performance metrics of the OTSM system may inadvertently lead to an increase in circuit complexity. Furthermore, a holistic solution that addresses the optimization of PAPR, BER, and circuit complexity concurrently remains an open problem. For example, the traditional Selective Mapping (SLM) technique has primarily focused on optimizing the PAPR performance of OFDM systems. While it has proven effective in improving the PAPR of OFDM systems, the identification of an optimal phase generation mechanism for PAPR reduction continues to be a popular research topic. However, unlike OFDM systems, the development of an SLM technique to optimize PAPR in OTSM systems remains an unresolved issue, primarily due to the structural differences between OFDM and OTSM. For instance, OFDM is a multi-carrier technique, while OTSM is a single-carrier technique. Additionally, the appropriate phase generation mechanism for applying the SLM technique to OTSM systems is also still to be determined. This paper proposes a modified SLM[10] technique for OTSM systems, utilizing Zadoff-Chu Transform (ZCT)[11] matrices, Discrete Cosine Transform (DCT)[11] matrices and Randomly Generated Phase (RGP) matrices to construct phase perturbation matrices. By applying cyclic shifts, multiple distinct phase perturbation matrices are generated. The SLM technique uses these matrices to create multiple candidate signals and selects the signal with the lowest PAPR for transmission. Simulation results indicate that the proposed method achieves optimal PAPR reduction performance using RGP matrices, outperforming ZCT or DCT matrices.

This work builds upon the precoding techniques proposed in Reference [11] for application in OTFS systems. In particular, this paper investigates how the two precoding techniques introduced in [11], namely ZCT and DCT, can be integrated with the conventional SLM technique to address the PAPR issue in OTSM systems. Unlike OTFS, which employs a multicarrier modulation scheme, OTSM is based on single-carrier modulation, resulting in notable differences in system architecture. Given that OTSM systems also experience the high PAPR problem, this paper proposes a PAPR reduction approach tailored to the specific characteristics of OTSM. The effectiveness of the proposed method in improving PAPR performance is analyzed and evaluated through simulations. The main contributions of this paper are summarized as follows:

- Modified Phase Perturbation Mechanism in the SLM Technique
  The conventional SLM technique employs a phase perturbation mechanism that generates U phase sequences of the same size as the input data, where each sequence is multiplied element-wise with the input to produce U candidate signals. For example, if the input data size in OTSM is $M \times N$, the phase perturbation sequences are

also of size $M \times N$. The proposed method introduces two modifications:

  - Precoding-Based Phase Perturbation: Rather than using element-wise multiplication, the proposed method applies an N×N precoding matrix to the input data via matrix multiplication, achieving phase perturbation through a structural transformation.
  - Circular Shift Generation of Phase Sequences: Unlike the traditional method that requires generating U distinct phase sequences, the proposed method creates only one perturbation matrix and obtains the remaining $U-1$ sequences via circular shifting.

  These modifications aim to reduce the computational complexity associated with phase sequence generation and introduce a structurally efficient mechanism using precoding and circular shifting.

- Phase Sequence Generation Mechanisms and Their Impact on PAPR
  The design of phase perturbation sequences is a critical area of SLM-related research due to its influence on PAPR reduction performance. This paper investigates three phase sequence generation mechanisms—ZCT, DCT, and RGP—within the framework of the proposed method. Their effectiveness is evaluated through simulations tailored to the OTSM system. The results indicate that the proposed method can achieve PAPR reduction under the evaluated conditions, while maintaining transmission performance comparable to that of the original OTSM system.
- Design of an SLM Technique Adapted for OTSM Systems
  Since the OTSM system already incorporates the Walsh–Hadamard Transform (WHT), as described in Reference [11], this paper presents an enhanced SLM technique that combines both SLM and precoding. The aim is to preserve the respective advantages of both methods.

Furthermore, based on the structure of the OTSM system, three types of phase perturbation matrices are analyzed within the proposed framework, and simulation results are provided to assess their applicability and performance in the OTSM context.

The chapter organization of this paper is structured as follows. Section II presents the mathematical definitions of OTSM signals and PAPR. Section III introduces the proposed method for improving the conventional SLM technique and its application in OTSM systems. Additionally, the definitions of the ZCT and DCT matrices are provided to demonstrate how the proposed method constructs an enhanced Phase Generation Mechanism (PGM) for the SLM technique. Sections IV and V focus on the discussion of simulation results and the conclusion, respectively.

II. Definition of OTSM Signals and PAPR

OTSM is a wireless communication signal modulation technique based on the Walsh-Hadamard Transform (WHT). This technique primarily utilizes WHT for signal processing in the delay-sequency domain and combines it with pulse-shaping filters to generate time-domain signals. The mathematical representation and generation method of OTSM signals are introduced as follows. OTSM signals are computed through the following steps. First, the information symbols to be transmitted are mapped onto a two-dimensional Delay-Sequency (DS) domain matrix $\mathbf{X}_{DS}$. Subsequently, each row of $\mathbf{X}_{DS}$ undergoes an N-point Walsh-Hadamard Transform, transforming the data into the delay-time domain to form a new matrix $\tilde{\mathbf{s}}$. This process is critical for OTSM signal processing as it achieves the conversion from the delay-sequency domain to the delay-time domain. The mathematical expression for any OTSM signal in the delay-time domain can be described as:

$$\tilde{s}[m+nM] = \sum_{k=0}^{N-1} X_{DS}[m,k] W_N[k,n], \tag{1}$$

where $X_{DS}[m,k]$ is the information symbol matrix in the delay-sequency domain, and $W_N[k,n]$ represents the normalized N-point Walsh-Hadamard Transform matrix. The index m corresponds to the delay dimension, ranging from 0 to $M-1$, and n corresponds to the sequency dimension, ranging from 0 to $N-1$, where M and N are the total numbers of delays and sequences, respectively. In the delay-sequency domain, zero-symbol vectors are typically inserted into the data matrix to mitigate inter-block interference (IBI) caused by channel delay spread. Channel delay spread leads to signal leakage between adjacent blocks, resulting in interference. To address this issue, OTSM technology introduces zero-symbol vectors into the data matrix, analogous to inserting guard bands in the time domain to prevent signal overlap. In OTSM systems, zero-symbol vectors are generally placed in the last few rows (z_l) of the data matrix, with the number of rows being equal to or greater than the channel delay spread index. This approach reduces the computational complexity at the receiver and enhances system performance in high-mobility environments. However, the inclusion of zero symbols reduces spectral efficiency as these symbols occupy transmission resources. Consequently, zero-padding offers a balanced solution between performance improvement and resource utilization, making it suitable for various application scenarios.

Next, OTSM applies pulse shaping to the time-domain vector $\tilde{\mathbf{s}}$ by using a pulse-shaping filter $g(t)$. This process converts the signal $\tilde{\mathbf{s}}$ into a form suitable for wireless transmission, generating the final time-domain signal $s(t)$. The representation of the OTSM signal in the time domain can be described as:

$$s(t) = \sum_{0 \le n < N, 0 \le m < M} \tilde{s}[m+nM] g(t-nT), \tag{2}$$

where $s(t)$ is the time-domain signal, $\tilde{s}[m+nM]$ represents the signal in the delay-time domain, $g(t)$ is the pulse-shaping filter, and T denotes the symbol period.

The primary advantage of OTSM lies in its lower computational complexity. Specifically, compared to traditional OFDM, OTSM employs WHT for modulation and demodulation, whereas OFDM uses the Fast Fourier Transform (FFT). As FFT requires extensive adders and multipliers, while WHT involves only addition and subtraction operations, OTSM reduces computational complexity and implementation costs effectively. Despite its advantages, OTSM signals exhibit the drawback of a high PAPR. PAPR is an important metric for evaluating system performance and design. It is defined as the ratio of the peak power to the average power of a signal. A high PAPR value indicates significant peak power, which

can have a profound impact on OTSM system performance. The calculation method and implications of PAPR for OTSM systems are discussed below.

The PAPR of OTSM signals can be calculated as follows. First, based on the time-domain representation of OTSM signals, as shown in Equation (2), the signal $s(t)$ is sampled at an appropriate time to obtain a discrete time-domain sample sequence $s[u]$, where $u = m+nM$ and $u = 0, 1, \ldots, MN-1$. After obtaining the time-domain sample sequence $s[u]$, the maximum absolute value among all samples is identified, and its squared value is taken as the signal's peak power. The average power of the signal is computed by summing the squared values of all samples and dividing by the total number of samples. Finally, the PAPR of the OTSM signal is determined by dividing the peak power by the average power and converting it to decibel (dB) units, as expressed by:

$$\mathrm{PAPR}(s[u]) = 10log_{10}\left(\frac{\max\limits_{u}|s[u]|^2}{P_{avg}}\right) \tag{3}$$

where $\max\limits_{u}|s[u]|^2$ represents the peak power of the OTSM signal, and P_{avg} is the average power.

The Complementary Cumulative Distribution Function (CCDF) is a statistical tool used to quantify the probability that a random variable exceeds a specified value. More precisely, the CCDF represents the likelihood that a random variable is greater than or equal to a particular threshold. In communication systems, the CCDF is widely applied, especially in analyzing the PAPR of signals. Mathematically, the CCDF is defined as follows: for a random variable V and a given threshold v, the CCDF is expressed as

$$\mathrm{CCDF}(v) = P(V > v) \tag{4}$$

where $P(V > v)$ denotes the probability that the random variable V exceeds v. In communication systems, the CCDF is widely employed to evaluate the PAPR of signals. It facilitates an effective comparison of the performance of various PAPR reduction techniques.

III. PROPOSED METHOD

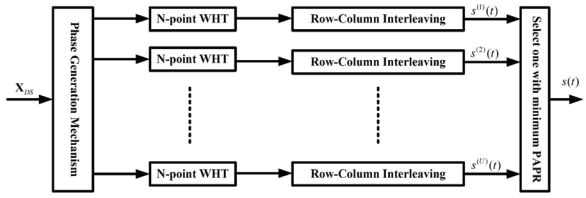


Fig. 1. Block Diagram of the Proposed Method.

Figure 1 illustrates the block diagram of the proposed method when applied to the OTSM system. In this figure, the primary focus of this paper is on the application of the SLM technique to the OTSM system, where phase perturbation is applied to the two-dimensional matrix of data. The phase generation mechanism is central to the main research topic of this paper. Additionally, this section will provide a brief review of the SLM technique's principles, followed by a discussion of the unresolved research challenges associated with phase generation mechanisms. Subsequently, the paper will present a low-complexity phase generation mechanism and use three commonly employed phase perturbation matrices for PAPR reduction to construct an enhanced phase generation method. In the next section, the PAPR performance of three low-complexity phase perturbation schemes will be evaluated.

High PAPR signals necessitate the use of power amplifiers with high linearity to avoid signal distortion. However, high-linearity amplifiers typically have lower efficiency, leading to increased power consumption and reduced system performance. Addressing high PAPR signals often requires more expensive linear power amplifiers, directly increasing system hardware costs, particularly in large-scale deployments. Conversely, using nonlinear power amplifiers to amplify high PAPR signals can result in nonlinear distortion, degrading signal quality. This not only affects system reliability but may also increase the bit error rate, reducing communication performance. Furthermore, high PAPR signals have a larger dynamic range, which implies a higher resolution requirement for Analog-to-Digital Converters (ADCs). If the ADC resolution is insufficient, signal distortion or information loss may occur, adversely affecting overall system performance. In summary, high PAPR introduces a range of negative effects, including low amplifier efficiency, increased system costs, degraded signal quality, and higher ADC resolution requirements. Therefore, developing effective PAPR suppression techniques is crucial for enhancing OTSM system performance and reducing costs.SLM is an effective approach for mitigating the PAPR in communication systems and is categorized under multiple signal representation (MSR) techniques. MSR techniques operate by generating multiple candidate signals through phase manipulation of the transmitted data, followed by the selection of the candidate signal with the lowest PAPR for transmission. In alignment with this principle, SLM generates U candidate signals by multiplying the transmitted data with U predesigned phase perturbation sequences. These modified signals are then processed further, often through OFDM modulation, to produce the final set of candidate signals. Ultimately, the signal exhibiting the minimum PAPR is chosen for transmission.

It is important to note that the modified SLM technique proposed in this paper primarily focuses on reducing the PAPR of OTSM systems, while attempting to introduce a phase generation mechanism that offers both low computational complexity and improved PAPR performance. The main goal is not to enhance the error correction capability of the transmitted signal. Furthermore, for the conventional SLM technique used in OFDM systems, the size of the phase perturbation sequence is consistent with the size of the input data and is represented as a row vector. In contrast, for the traditional SLM technique applied to OTSM systems, the size of the phase perturbation sequence must also match the input data size, but it will be represented as a two-dimensional matrix. Therefore, the design of a phase generation mechanism to create phase perturbation sequences that optimize PAPR performance will differ from the OFDM system case, making this an entirely new and important research topic. Reference [11] demonstrates the use of precoding techniques to reduce the PAPR of OTFS systems, with simulation results showing that the application of a precoding matrix effectively improves the PAPR performance of OTFS systems. Based on this, this paper proposes an enhanced SLM technique that combines

precoding and SLM techniques. Specifically, the precoding matrix is employed as the phase perturbation matrix to apply phase disturbance to the transmitted data, and cyclic shifts are used to generate multiple phase perturbation matrices. However, it should be noted that while the precoding matrix can effectively improve the PAPR performance of OTFS systems, it does not necessarily yield the same improvement for OTSM systems.

In this paper, the authors have adopted a precoding matrix from reference [11] and modified it to serve as the phase generation mechanism for the SLM technique. The performance of the proposed method in reducing PAPR is specifically demonstrated in the simulation results section, which aims to evaluate whether the precoding matrix, originally used for OTFS systems, can serve as a viable phase generation mechanism to improve the PAPR performance of OTSM systems. To compare whether the precoding matrix described in reference [11] can improve the PAPR performance of OTSM systems in the same manner as it does for OTFS systems, the ZCT matrix and DCT matrix described in reference [11] will be introduced as follows. Given the limited exploration of SLM in the context of OTSM systems, this paper proposes an advanced SLM approach specifically adapted to OTSM. Conventional SLM methods typically employ randomly generated phase perturbation sequences, which, while effective, tend to increase system circuit complexity. Consequently, the development of low-complexity PGMs capable of maintaining or improving PAPR reduction performance remains a significant area of investigation. To address this issue, this paper evaluates conventional random PGMs and introduces two alternative mechanisms utilizing the ZCT matrix and the DCT matrix. These mechanisms are assessed for their ability to enhance PAPR performance while minimizing complexity. The definitions of the ZCT and DCT matrices are provided below[11].

1) **ZCT Matrix**
A Zadoff-Chu sequence of length K is expressed as:

$$z[k] = \begin{cases} \exp\left(j\frac{2\pi r}{K}\left(\frac{k^2}{2}+qk\right)\right), & \text{if } K \text{ is even,} \\ \exp\left(j\frac{2\pi r}{K}\left(\frac{k(k+1)}{2}+qk\right)\right), & \text{if } K \text{ is odd.} \end{cases}$$

From this sequence, a ZCT matrix of size $M \times M$ is constructed as:

$$z_M[m,l] = z[m+lK], \quad m,l = 0,1,\ldots,M-1.$$

For example, with $M = 4$, $q = 7$, and $r = 1$, the ZCT matrix is given by:

$$\begin{bmatrix} 1 & i & -1 & -i \\ -0.98+0.20i & 0.98-0.20i & -0.98+0.20i & 0.98-0.20i \\ 1 & -i & -1 & i \\ -0.83-0.56i & -0.83-0.56i & -0.83-0.56i & -0.83-0.56i \end{bmatrix}$$

2) **DCT Matrix**
A Discrete Cosine Transform matrix C_M of size $M \times M$ is defined as:

$$C_M[m,l] = \begin{cases} \sqrt{\frac{1}{M}}, & m = 0,\ 0 \le l \le M-1, \\ \sqrt{\frac{2}{M}}\cos\left(\frac{(2l+1)m\pi}{2M}\right), & 1 \le m \le M-1,\ 0 \le l \le M-1. \end{cases}$$

For $M = 4$, the DCT matrix is represented as:

$$\begin{bmatrix} 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.6533 & 0.2706 & -0.2706 & -0.6533 \\ 0.5000 & -0.5000 & -0.5000 & 0.5000 \\ 0.2706 & -0.6533 & 0.6533 & -0.2706 \end{bmatrix}$$

The proposed method modifies the phase of the data matrix $\mathbf{X}_{\text{DS}}$ in the delay-sequency domain by multiplying it with a phase perturbation matrix $\mathbf{P}^{(u)}$, resulting in U candidate signals as:

$$s^{(u)}(t) = \sum_{\substack{0\le k<N \\ 0\le n<N \\ 0\le m<M}} \mathbf{P}^{(u)} X_{\text{DS}}[m,k] W_N[k,n] g(t-nT), \tag{5}$$

where $u = 1,\ldots,U$ and $\mathbf{P}^{(u)}$ denotes an $M \times M$ phase perturbation matrix. To reduce circuit complexity, these phase perturbation matrices are generated using the ZCT and DCT matrices, which are then modified through cyclic right shifts. Furthermore, as shown in Equation (1), the generation of OTSM signals involves the superposition of multiple modulation symbols. When several of these symbols exhibit phase alignment, constructive interference may occur, resulting in signal peaks with significantly elevated amplitudes. This phenomenon directly contributes to the high PAPR characteristic of OTSM signals. To address this, a phase perturbation matrix is introduced, whose primary function is to impose deliberate phase variations on the transmitted symbols. This phase diversity disrupts potential phase alignment, thereby reducing the probability of peak formation and lowering the resulting PAPR. To illustrate the construction method of such a phase perturbation matrix, consider the use of a 4×4 DCT matrix. Based on this reference matrix, four phase perturbation matrices $\mathbf{P}^{(u)}$, where $u = 0,1,2,3$, are derived by applying successive circular shifts along the columns. These matrices are expressed as follows:

$$\mathbf{P}^{(0)} = \begin{bmatrix} 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ -0.6533 & 0.6533 & 0.2706 & -0.2706 \\ -0.5000 & -0.5000 & 0.5000 & 0.5000 \\ -0.2706 & 0.2706 & -0.6533 & 0.6533 \end{bmatrix},$$

$$\mathbf{P}^{(1)} = \begin{bmatrix} 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ -0.2706 & -0.6533 & 0.2706 & 0.6533 \\ 0.5000 & 0.5000 & -0.5000 & -0.5000 \\ 0.6533 & -0.2706 & -0.6533 & 0.2706 \end{bmatrix},$$

$$\mathbf{P}^{(2)} = \begin{bmatrix} 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.2706 & -0.2706 & -0.6533 & 0.6533 \\ -0.5000 & 0.5000 & -0.5000 & 0.5000 \\ -0.6533 & -0.6533 & 0.2706 & 0.2706 \end{bmatrix},$$

$$\mathbf{P}^{(3)} = \begin{bmatrix} 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.6533 & 0.2706 & -0.6533 & -0.2706 \\ 0.5000 & -0.5000 & -0.5000 & 0.5000 \\ 0.2706 & -0.6533 & 0.6533 & -0.2706 \end{bmatrix}.$$

The above method demonstrates that multiple distinct phase perturbation matrices can be generated without the need for additional arithmetic operations such as multiplication or addition. These matrices can be employed to modulate the transmission sequence, generating a set of candidate signals. Subsequently, the transmitter selects the candidate signal with the lowest PAPR for actual transmission. More specifically, this approach generates U distinct phase perturbation matrices, where U is at most M. Furthermore, unlike the traditional SLM technique, which requires the generation of U phase perturbation matrices in the OTSM system, the proposed method only requires one phase perturbation matrix. This simplification reduces the hardware complexity required for the phase generation mechanism in circuit design. Additionally, the proposed method uses $U-1$ cyclic shifts to generate $U-1$ phase perturbation matrices. Since cyclic shifts do not require adders or multipliers, this further alleviates the hardware implementation complexity. In summary, the proposed method is an enhanced SLM technique based on the OTSM system architecture, offering suboptimal PAPR improvement performance with low circuit complexity for PAPR reduction. On the receiver side, the proposed method follows the same

procedure as the conventional SLM technique. Since both the transmitter and receiver have prior knowledge of the U phase perturbation matrices, the amount of side information required remains $\log_2(U)$ bits, as in conventional SLM. No additional bits are needed to indicate which of the U candidate signals is selected for transmission. Thus, after signal processing at the OTSM receiver, the proposed method can be recovered using the standard SLM demodulation procedure.

IV. SIMULATION RESULTS

This paper proposes a modified SLM technique to improve the PAPR performance of OTSM systems, achieved by combining traditional SLM with precoding techniques. The proposed method simplifies the complexity of phase perturbation generation in traditional SLM techniques by employing cyclic shifting, and optimizes PAPR performance using precoding techniques. In this section, the paper compares the PAPR performance of the modified SLM technique with that of the conventional SLM technique to assess whether the proposed method can retain the PAPR improvement benefits of the traditional SLM technique. To assess the PAPR reduction performance of the proposed method in OTSM systems, simulations are conducted using the following parameters. A total of $10,000$ OTSM signals are generated randomly, employing digital modulation schemes of 4-QAM and 16-QAM. The total number of delays (M) and sequences (N) are configured to 64 and 256, respectively. For the number of candidate signals, values of $\{4, 16\}$ are utilized. In the conventional SLM approach, two randomly generated phase variations $\{-1, 1\}$ are applied. For the proposed method, the ZCT matrix parameters are set as $q = 7$ and $r = 1$. Additionally, the number of zero symbols inserted into the last row of the data matrix (z_l) is configured to 3. All simulation outcomes in this section are illustrated using CCDF curves. The CCDF curves for various methods are compared and analyzed to provide a thorough evaluation of their relative PAPR reduction performance.

Figure 2 and Figure 3 compare the performance of three PAPR reduction techniques in a 4QAM-modulated OTSM system. These techniques include the traditional SLM technique, the proposed method based on the RGP matrix, the proposed method based on the DCT matrix, and the proposed method based on the ZCT matrix. As shown in Figure 2 and Figure 3, the RGP-based method demonstrates superior PAPR reduction performance compared to the ZCT-based method and DCT-based method. Additionally, the PAPR reduction performance of the RGP-based method improves significantly with an increasing number of candidate signals. To analyze the impact of the total delay number (M) and the total sequency number (N) on the PAPR reduction performance of the proposed methods, Figures 4 and 5 present simulation results under the condition of four candidate signals, while Figure 6 shows results for 16 candidate signals. These results are obtained by varying the total sequency number (or total delay number) while fixing the total delay number (or total sequency number) at 64. In Figure 4, when $M = 64$ and $N = 128$ or $N = 256$, both the DCT-based and ZCT-based methods exhibit a slight degradation in PAPR reduction performance as the total sequency number increases. Similarly, in Figure 5, when $N = 64$ and $M = 128$ or $M = 256$, the PAPR reduction performance of both methods slightly decreases with an increase in the total delay number. Nevertheless, the results presented in Figures 4, 5, and 6 demonstrate that, regardless of variations in M or N, the RGP-based method consistently outperforms both the ZCT-based and DCT-based methods in terms of PAPR reduction. It is noteworthy that when M is smaller than N, the proposed method yields improved PAPR performance, whereas, when M exceeds N, the improvement in PAPR performance decreases. To further investigate whether the proposed method's PAPR reduction performance is influenced by variations in the T-value of T-ary QAM modulation, this paper simulated and analyzed the performance of the proposed method in a 16-QAM-modulated OTSM system. Figure 7 and 8 presents a comparison of three PAPR reduction techniques in the 16-QAM-modulated OTSM system. Consistent with the results shown in Figure 2 and Figure 3, the RGP-based method demonstrates superior PAPR reduction performance in the 16-QAM system compared to the ZCT-based method and the DCT-based method. The overall simulation results confirm that the RGP-based method consistently outperforms both the ZCT-based method and the DCT-based method in terms of PAPR reduction performance in both 4-QAM and 16-QAM-modulated OTSM systems. Figure 9 compares the bit error rate (BER) curves of the proposed method and the OTSM system over an additive white Gaussian noise (AWGN) channel. The results in Figure 9 show that when $M \leq N$, the RGP-based proposed method performs slightly worse than the OTSM system in the AWGN channel; however, when $M > N$, the RGP-based proposed method outperforms the OTSM system. Therefore, from the perspective of both PAPR performance improvement and channel transmission performance, the RGP-based proposed method effectively combines the advantages of both SLM and precoding techniques in improving PAPR. Additionally, the integration of precoding techniques helps reduce the complexity of the phase perturbation mechanism in traditional SLM techniques and provides a better phase perturbation matrix for PAPR improvement.

V. CONCLUSION

To mitigate the issue of high PAPR in OTSM systems and enhance overall transmission performance, this paper introduces an improved Selected Mapping (SLM) technique specifically designed for OTSM systems. The proposed approach constructs the phase generation matrix (PGM) using ZCT, DCT, and RGP matrices, in conjunction with cyclic shifts, to generate multiple candidate signals for selecting the optimal transmission signal. Simulation results show that the PGM based on the RGP matrix outperforms the ZCT-based and DCT-based mechanisms, as well as conventional SLM techniques, in terms of PAPR reduction. Moreover, the proposed method, utilizing the RGP-based PGM, consistently delivers significant PAPR reduction in OTSM systems, regardless of whether 4-QAM or 16-QAM modulation schemes are employed. The findings of this paper may serve as a reference

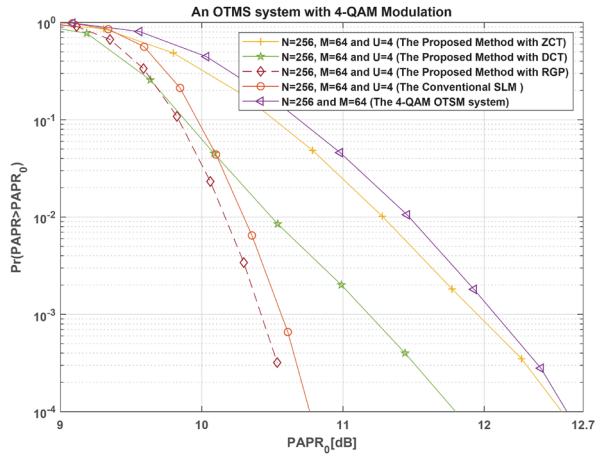


Fig. 2. Performance Comparison of PAPR Reduction Techniques in 4QAM-modulated OTSM Systems with $M = 64$ and $N = 256$.

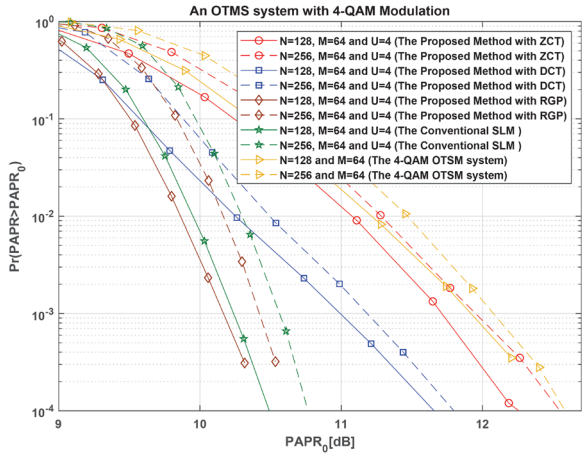


Fig. 4. PAPR Reduction Performance Comparison of the Proposed Method in 4QAM-modulated OTSM systems with $M = 64$ and $N = 128$ or 256.

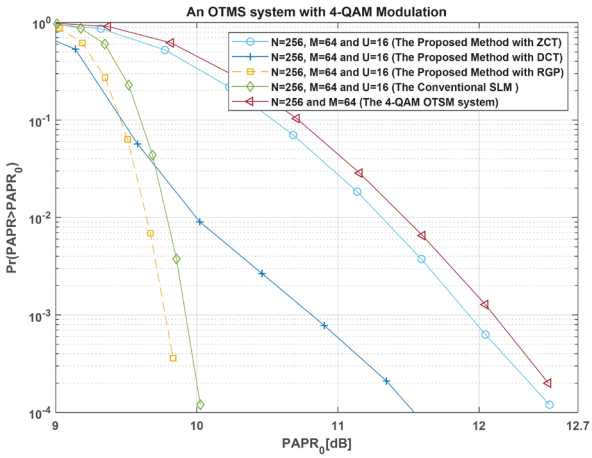


Fig. 3. Performance Comparison of PAPR Reduction Techniques in 4QAM-modulated OTSM Systems with $M = 64$ and $N = 256$.

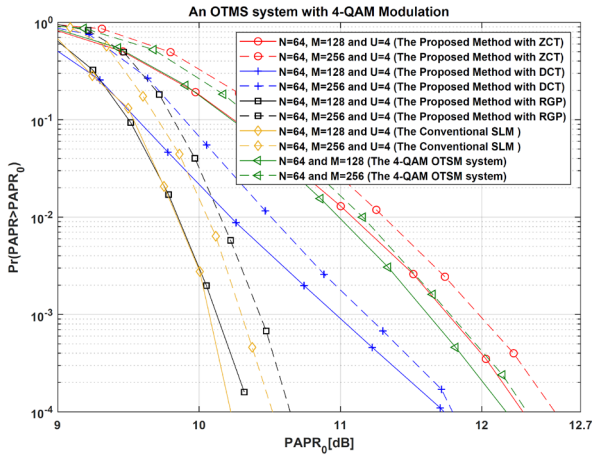


Fig. 5. PAPR Reduction Performance Comparison of the Proposed Method in 4QAM-modulated OTSM systems with $N = 64$ and $M = 128$ or 256.

for improving OTSM system performance in areas such as channel efficiency and circuit design. These insights can also contribute to the optimization of key issues in OTSM systems, including channel performance, circuit implementation, and PAPR reduction.

REFERENCES

[1] S. S. Mohamed and H. N. Abdullah, "Interleaving based SCMA Codebook Design Using Arnold's Cat Chaotic Map", *Iraqi Journal of ICT*, vol. 6, no. 2, pp. 42–57, Aug. 2023. **DOI**: 10.31987/ijict.6.2.225.

[2] A. A. Abdulhussein and H. N. Abdullah, "Performance evaluation of filter bank multi-carrier modulation in multipath fading channels", *Iraqi Journal of ICT*, vol. 1, no. 1, pp. 1–10, Dec. 2021. **DOI**: 10.31987/ijict.1.1.181.

[3] T. Thaj and E. Viterbo, "Orthogonal time sequency multiplexing modulation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, pp. 1–7, Mar. 2021. **DOI**: 10.1109/WCNC49053.2021.9417451.

[4] T. Thaj, E. Viterbo, and Y. Hong, "Orthogonal time sequency multiplexing modulation: Analysis and low-complexity receiver design," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 7842–7855, Dec. 2021. **DOI**: 10.1109/TWC.2021.3088479.

[5] H. Wang, Q. Chen, X. Wang, W. Du, X. Li and A. Nallanathan, "Adaptive Block Sparse Backtracking-Based Channel Estimation for Massive MIMO-OTFS Systems," *IEEE Internet of Things Journal*, vol. 12, no. 1, pp. 673–682, Jan. 2025. **DOI**: 10.1109/JIOT.2024.3466911.

[6] A. Kumar and A. Nanthaamornphon, "Reducing the peak to average power ratio in optical NOMA waveform using Airy-special function based PTS algorithm," *Infocommunications Journal*, vol. XVII, no. 1, pp. 11–18, Mar. 2025. **DOI**: 10.36244/ICJ.2025.1.2.

[7] R. Al Ahsan, F. M. Ghannouchi, and A. O. Fapojuwo, "Analysis of adapted tone reservation PAPR reduction techniques in OTSM system," in *Proc. 2023 33rd Int. Telecommun. Netw. Appl. Conf. (ITNAC)*, Melbourne, Australia, pp. 1–6, Nov. 2023. **DOI**: 10.1109/ITNAC59571.2023.10368554.

[8] S. G. Neelam and P. R. Sahu, "Iterative channel estimation and data detection of OTSM with superimposed pilot scheme and PAPR analysis," *IEEE Commun. Lett.*, vol. 27, no. 8, pp. 2001–2005, Aug. 2023. **DOI**: 10.1109/LCOMM.2023.3281575.

[9] A. Doosti-Aref, C. Masouros, E. Basar, and H. Arslan, "Pairwise sequency index modulation with OTSM for green and robust single-carrier communications," *IEEE Commun. Lett.*, vol. 28, no. 4, pp. 1164–1168, Apr. 2024. **DOI**: 10.1109/LWC.2024.3360256.

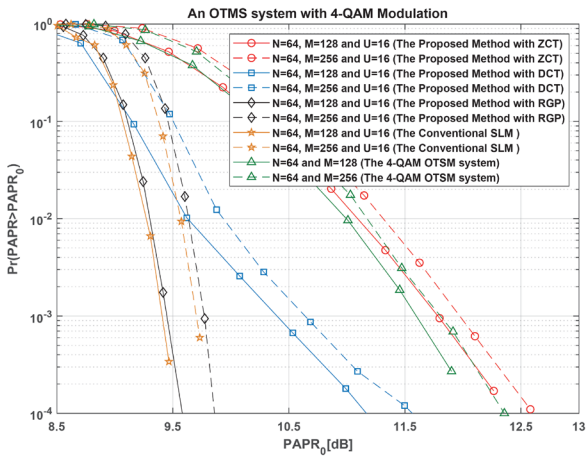


Fig. 6. PAPR Reduction Performance Comparison of the Proposed Method in 4QAM-modulated OTSM systems with $N = 64$ and $M = 128$ or 256.

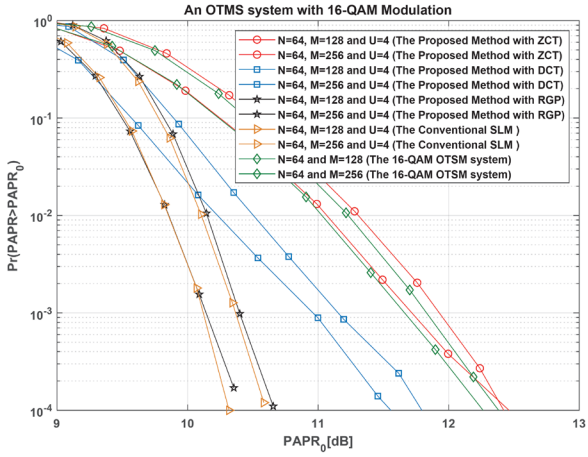


Fig. 8. PAPR Reduction Performance Comparison of the Proposed Method in 16QAM-modulated OTSM systems with $N = 64$ and $M = 128$ or 256.

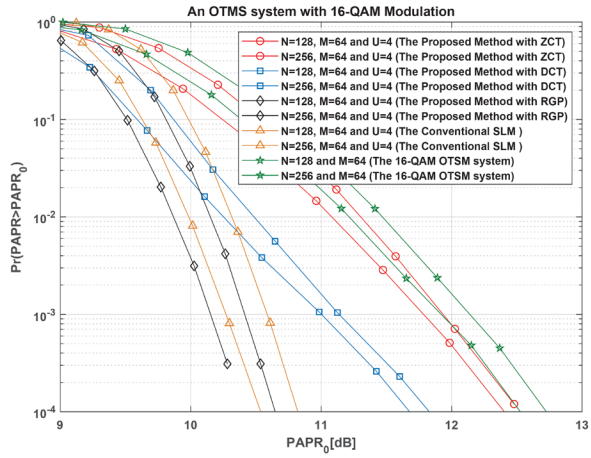


Fig. 7. PAPR Reduction Performance Comparison of the Proposed Method in 16QAM-modulated OTSM systems with $M = 64$ and $N = 128$ or 256.

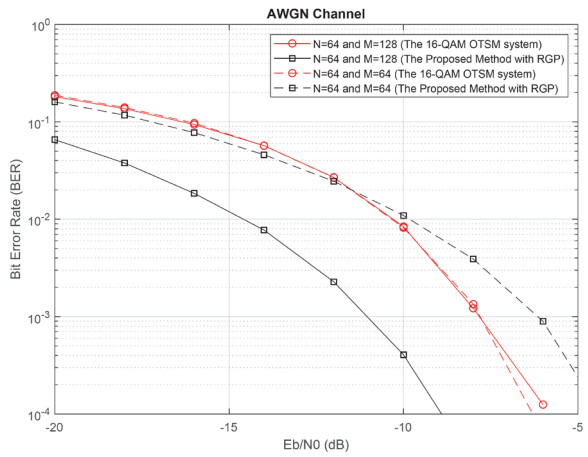


Fig. 9. BER Performance of OTSM Systems with and without the Proposed Method under AWGN

[10] T. Jiang, and Y. Wu, "An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM Signals," *IEEE Transactions on Broadcasting*, vol. 54, no. 2, pp. 257–268, June 2008. DOI: 10.1109/TBC.2008.915770.

[11] Y. I. Tek and E. Basar, "PAPR Reduction Precoding for Orthogonal Time Frequency Space Modulation," *2023 46th International Conference on Telecommunications and Signal Processing (TSP)*, Prague, Czech, July 2023. DOI: 10.1109/TSP59544.2023.10197718.

Hsin-Ying Liang received the B.S. degree from National Taipei University of Technology, Taiwan, in 2001, the M.S. degree from National Chi Nan University, Taiwan, in 2003, and the Ph.D. degree from National Chung Hsing University, Taiwan, in 2008, all in electrical engineering. Since 2008, he has been with the Department of Information and Communication Engineering, Chaoyang University of Technology, Taiwan, where he is currently a Professor. His research interests in-clude error-correcting codes and OFDM-based communication systems.

Chuan-Bi Lin received his B.S. degree from Tatung University, Taiwan, in 1994, his M.S. degree from the University of Bridgeport, USA, in 2001, and his Ph.D. degree in Electrical and Computer Engineering from the New Jersey Institute of Technology, Newark, USA, in 2008. He served as a postdoctoral researcher in the Department of Electrical and Computer Engineering at the New Jersey Institute of Technology from 2008 to 2009. He joined Chaoyang University of Technology in 2009 as an assistant professor in the Department of Information and Communication Engineering and was promoted to associate professor on August, 2023. His research interests include computer networks, switching and routing, Internet of Things (IoT) applications, image recognition, and artificial intelligence.

Design of Enhanced Dual Band Rectenna with Binary Coding Technique of Genetic Algorithm

Yasmeen Taha Yaseeen[1], Mohd Fadzli Mohd Salleh[2], and Taha A. Elwi[3,*]

***Abstract*—In this paper, an artificial intelligence-based rectenna design is proposed for Wi-Fi applications. The rectenna design is optimized using a Genetic Algorithm (GA) integrated with a Binary Coding (BC) scheme. The proposed rectenna is configured to operate at 2.45 GHz and 5.8 GHz with a maximum size of 27×30×10 mm3. The performance of the optimized rectenna has been characterized in terms of S-parameters, bandwidth, radiation patterns, and gain. For this, a dual-bandwidth patch is designed to suit the applications of 2.45 GHz and 5.8 GHz. The measured radiation patterns and S11 spectra are evaluated to obtain a peak radiation efficiency of 52% and 56% to realize a gain of 6.2 dB at 2.45 GHz and 7.12 dB at 5.8 GHz, respectively. The proposed rectenna is integrated with an RF–DC rectifier (RFD102A module) to evaluate the harvested power in terms of DC output voltage under outdoor conditions. The maximum obtained harvested DC output voltage is found to be 2.27V and 2.3V at 2.45GHz and 5.8GHz, respectively. Finally, the obtained measurements are compared to the simulated results to realize good agreements between them.**

***Index Terms*—Binary coding, energy harvesting, genetic algorithm, microstrip rectenna, rectifier, Wi-Fi, sub-6GHz.**

I. INTRODUCTION

THIS growing demand for wireless technologies has led to an environment increasingly surrounded by microwave sources [1]. In addition to mobile communication systems, everyday life extensively utilizes wireless applications like Bluetooth, Wi-Fi, WLAN, and the Internet of Things (IoT) [2]. These wireless systems operate within standardized frequency bands [1]. Consequently, both indoor and outdoor environments contain abundant microwave energy across the frequency spectrum, which can be harvested using single- or multiband rectennas [3]. Recently, many designs have been reported for RF energy harvesting systems [4], which can be classified into three categories: single-band, multi-band, and broadband. These systems use different rectennas to improve gain and rectification efficiency by connecting impedance-matching circuits using GA techniques. Such research studies provided various optimization algorithms, including GA and particle swarm optimization, to enhance antenna design performance for multiple applications in modern wireless communication networks. In [5], a microstrip patch antenna array was designed using defective ground structure with the aid of GA to realize a dual-band of operation at 2.45 GHz and 5.8 GHz for wireless power transfer applications. The design of an antenna array was optimized to minimize side lobe effects by utilizing a microstrip patch antenna element that offers improved bandwidth, directivity, and efficiency at multiple resonant frequencies [6]. A study was developed in [7] with the aid of web-tool-based GAs to integrate seven empirical propagation loss models to optimize antenna performance and improve wireless coverage and network capacity. A report was published in [8] to optimize a 28 GHz microstrip antenna with GA in terms of width, microstrip line width, and dielectric permittivity to achieve remarkable performance. A tri-band miniaturized rectangular patch antenna based on a defected ground structure optimized using GA to achieve a size reduction of about 82% smaller than a conventional single-band structure that covers the frequency bandwidths between 3.2 GHz and 3.5 GHz, 5.5 GHz and 5.9 GHz, and 6.3 GHz and 7.1 GHz with gains of 0.7 dBi, 1.76 dBi, and 2.93 dBi, respectively [9]. The proposed antenna in [10] was optimized from a single microstrip patch with a binary-coded genetic algorithm (BCGA) to achieve triple-band operation at 28 GHz, 40 GHz, and 47 GHz for mm-wave applications; this antenna showed a gain of 7.7 dBi, 12.1 dBi, and 8.2 dBi, respectively. The published work in [11] proposed a method to enhance miniaturized microstrip antenna performance using GA to realize bandwidths with sub-6 GHz. An antenna patch design was created using GA, featuring a miniaturized size that operates within the frequency band of 1.8 GHz to 3 GHz, resulting in an array of 9×9 elements [12]. A model that utilizes a high-frequency electromagnetic simulator was proposed in [13] to analyze binary mixtures, employing GA to operate between 8.2 GHz and 12.4 GHz. A comprehensive study on GA for optimizing electromagnetic problems, including complex issues, was presented in [14]. A broadband triple-band frequency patch antenna for WLAN applications was designed with the aid of GA optimization [15]. The study referenced in [16] introduced a compact, cost-effective microstrip antenna for a V2V communication system by implementing a defected ground structure to reduce antenna losses through GA optimization. In another context, the application of GA in antenna design has revolutionized the field by offering an efficient solution to the optimization of complex, multi-parameter problems. One of the most significant benefits of using BCGA in antenna design is the simplification of the encoding process [17]. The application of BCGA in antenna design has been extensively studied, with numerous successful implementations. For example, the effectiveness of GAs in optimizing linear array antennas was demonstrated in [15]. By encoding the array elements' positions and excitation amplitudes as binary strings, they were able to optimize the array radiation pattern, achieving significant improvements in

[1] School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Nibong Tebal, Malaysia, (e-mail: yasmeen@student.usm.my).

[2] School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Seri Ampangan, 14300 Nibong Tebal, Pulau Pinang, Malaysia (e-mail: fadzlisalleh@usm.my).

[3] Department of Automation and Artificial Intelligence Engineering, College of Information Engineering at Al-Nahrain University, Baghdad, Iraq

[*] Corresponding author: Taha A. Elwi, (e-mail: taelwi82@nahrainuniv.edu.iq).

DOI: 10.36244/ICJ.2025.4.3

directivity and sidelobe levels. Similarly, in [18], GA was used to optimize microstrip patch antennas, focusing on parameters such as the patch dimensions and feed position to enhance bandwidth and radiation patterns. In phased antenna arrays, BCGA has proven particularly useful, as in [17], which employed GAs to optimize the design of phased arrays, targeting improvements in sidelobe levels and directivity. In [19], a design was explored using GAs to optimize the antenna multiband performance by encoding the antenna geometry and metamaterial properties as binary strings. In [20], a GA was used to optimize wire antennas, focusing on the adaptability of BC to manage complex design challenges. In this paper, an enhanced rectenna design with distinguished performances is introduced for RF energy harvesting at Wi-Fi frequency bands. It is built to achieve excellent conversion efficiency when introduced to an RF energy harvester. Suitable scaling and slight tuning can make the energy harvester system applicable for a wide range of low-power applications, including those operating within the sub-6 GHz band applications and Wi-Fi bands.

II. RECTENNA DESIGN AND GEOMETRICAL DETAILS

The basic rectenna configuration is based on a conventional printed monopole design with coplanar waveguide (CPW) 50Ω feed, as seen in Figure 1(a). Next, the rectenna back panel is covered with a fractal based on Minkowski geometry, as shown in Figure 1(b). The rectenna is printed on an FR4 substrate of 1.6 mm thickness. The rectenna is mounted below another FR4 substrate of the same dimensions. The back panel of the second substrate, see Figure 1(c), is covered with a copper layer of a square aperture, while the other side of the substrate is a rectangular patch designed with a BC scheme, as seen in Figure 1(d). The first layer enables the 2.45 GHz resonant mode, which is achieved through the fractal-based geometry. The second mode is realized from the second rectenna patch based on the BC scheme, as will be shown later.

Now, to obtain the proposed rectenna performance, a full wave analysis is conducted to realize the optimal design by using a parametric study. However, the optimization issue that is defined in this paper is resolved using a BC scheme. This geometry is designed using a straightforward procedure to create the proposed rectenna. The main design challenge lies in optimizing the second-layer BC scheme used to form the patch shown in Figure 1(d). The upper and the lower layers are coupled to realize a 3D-printed geometry with a separation distance of 10 mm. For this, GA is used to find the best rectenna bandwidth with the optimal gain at Wi-Fi bands. The binary-coded patch area is divided into small cells, and the GA uses a code of 0s and 1s to define the conductor regions of the patch. In a binary genetic algorithm (BGA) procedure, the important genetic operators are selection, crossover, and mutation [20].

The selection operator chooses two parent chromosomes from the population at random. The crossover operation mixes two parents' chromosomes to create a new child chromosome [21]; then a mutation introduces changes to the generated chromosome with a certain probability [22].

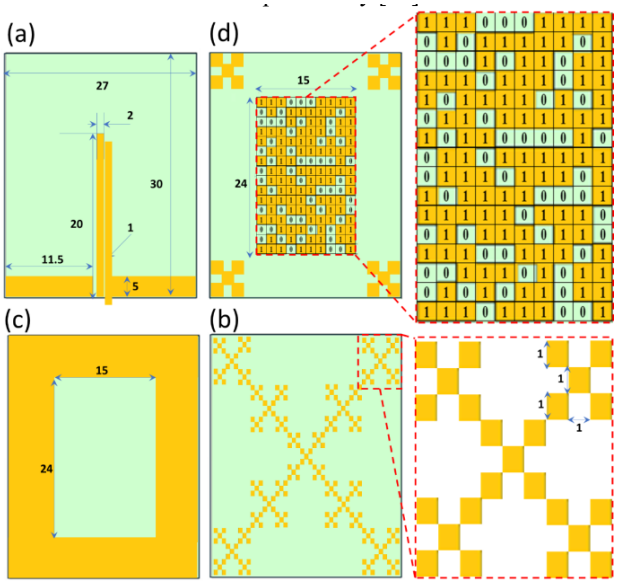


Figure 1: Rectenna design: (a) The printed monopole, (b) the Minkowski geometry, (c) rectangular patch, and (d) BC patch.

The initial rectenna design based on BC is performed at 2.45 GHz and 5.8 GHz on FR4 substrate. The conducted BGA simulations are used to access the optimum patch pattern by dividing it into 160-unit cells (10×16) with a size of 1.5×1.5 mm². A single unit cell is treated as a gene and is encoded with a binary value of 1 if it is a metal pixel and 0 if it is not a metal pixel. The overlaps between adjacent unit cells are minimized to maintain electrical continuity during fabrication. Since the material of each unit cell is directly mapped into a binary value, the patch geometry is arranged as a binary string involving a series of zeros and ones. Each binary string will lead to an arbitrary structure for the radiating patch of the rectenna. The performance of the rectenna can be improved by varying the values in the binary string. Thus, the desired performance can be achieved without increasing the physical size of the rectenna by controlling the patch surface current. The optimal binary string for the patch rectenna is investigated by a number of iterations. However, to preserve the continuity of the structure, zeros are replaced with ones if they are surrounded by ones, and vice versa. This method is applied to the patch rectenna by using a fitness function that is meant to minimize the reflection coefficient and increase the bandwidth of the rectenna. The fitness function is considered to be the rectenna gain bandwidth product (G*BW) and gain as given in equations (1) and (2). If the gain-bandwidth product exceeds 50%, the fitness function is assigned a value of 1; otherwise, it is set to 0. The multiplication is evaluated as a fitness function.

$$fitness(F) = \frac{1}{N}\sum_{i=1}^{N} G * BW \qquad (1)$$

where

$$G * BW(\%) = \begin{cases} \geq 50\% \; than \; set \; 1 \\ < 50\% \; than \; set \; 0 \end{cases}. \qquad (2)$$

N is the number of sampling frequencies in a given band. The coefficient G*BW should be maximum at the resonance frequency. GA is one of the EM optimization techniques integrated with electromagnetic software packages such as CST MWS and HFSS. The simulation, along with the BGA script, is

fetched to the CST MWS macro command code window with Visual Basic scripts based on the listed parameters in Figure 1(d). The code iterations are executed after reaching the desired value of the fitness function with an error of 2% or reaching the best fitness function value that shows no significant change for 20 generations.

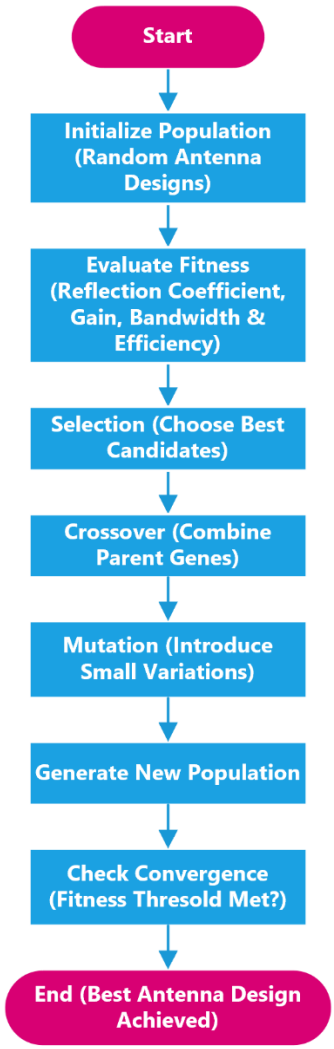


Figure 2. Flowchart of BGA optimization process.

The computing performance is related to the full-wave electromagnetic simulation time for the structure; therefore, the proposed BGA algorithm provides the most robust solution among potential solutions. The optimal patch geometry was achieved after evaluating 120 chromosome genes. The solution space consists of 2160 solutions in comparison to the traditional full-wave analysis techniques, which require a substantial amount of time for analysis. However, the proposed solution based on the proposed BGA algorithm takes only a few attempts to find the solution. The optimized patch exhibits S11 below -10dB at 2.45GHz and 5.8GHz with gains of 6.2dBi and 7.12dBi, respectively. The optimized patch structure is presented in Figure 1 with all geometrical details. For further details, the proposed method is represented as flowchart as shown in Figure 2. The considered algorithm iteratively refines the rectenna geometry to maximize efficiency while minimizing computational overhead. The step-by-step workflow of the GA-based rectenna optimization process used in this study.

III. RECTENNA DESIGN AND GEOMETRICAL DETAILS

The performance of the optimized rectenna has been characterized in terms of S11, radiation pattern, and gain; these characteristics are shown in Figure 3. The simulations are performed and compared to the experimental measurements for validations. The rectenna is fabricated using chemical etching and measured inside an RF chamber, as seen in Figure 3. By considering the operating frequencies of the rectenna, the reflection coefficient is measured by the Agilent E5071C Vector Network Analyzer (VNA).

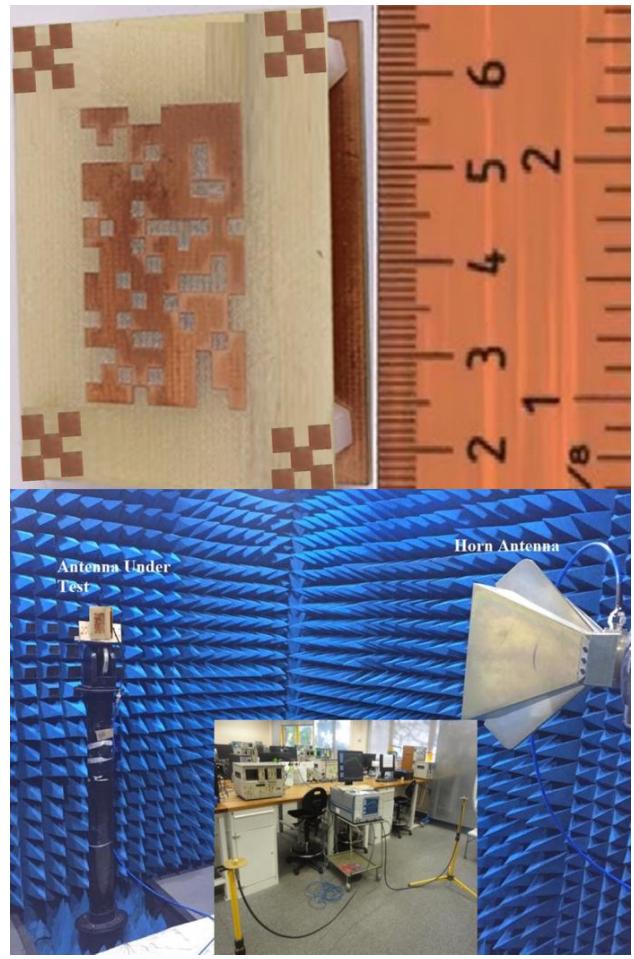


Figure 3: The fabricated rectenna prototype.

The experimental validation of the simulated results is carried out from a software package of CST MWS as given in

Figure 4. The measured and simulated S_{11} spectra of the proposed rectenna is shown in Figure 3(a). The obtained results show that the rectenna bandwidth found between 2.25GHz and 2.64GHz with S_{11} below -10dB. The measured bandwidth of the proposed rectenna, see Figure 4(a), at the second band is found from 5.38GHz to 5.86GHz with S_{11} lower than -10dB. Figure 4(b) shows the measured and simulated radiation patterns at 2.45GHz and 5.8GHz for azimuth and zenith for both co-polarization and cross-polarization. The measured radiation patterns are presented in Figure 4(c) for both azimuth and elevation planes, showing co- and cross-polarization components. The simulated results show excellent agreements with the obtained from measurements. Minor discrepancies between the simulated and measured results are attributed to fabrication tolerances.

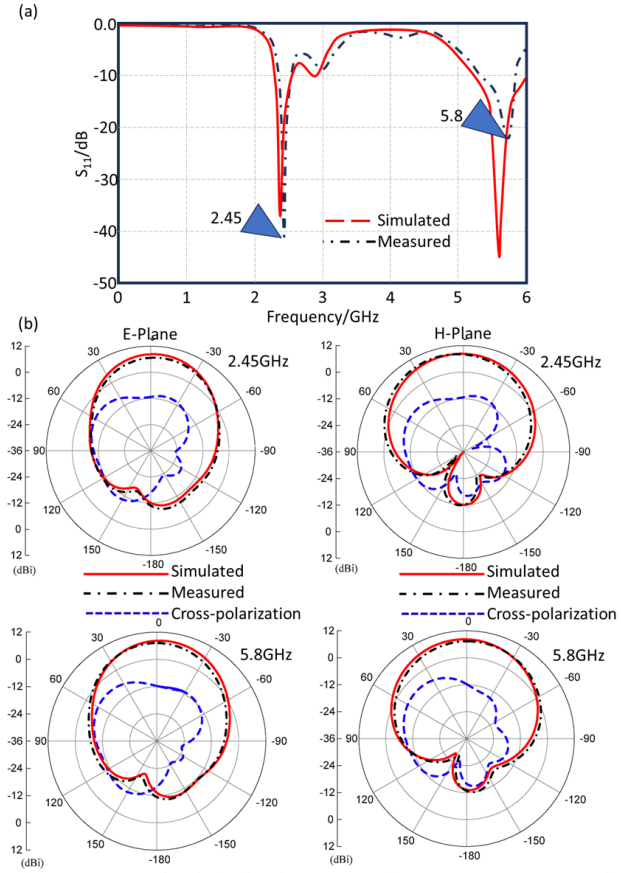


Figure 4. Measured and simulated rectenna performances (a) S11 spectra (b) Radiation patterns.

Figure 5(a) shows the surface current distributions for the proposed rectenna at 2.45 GHz and 5.8 GHz on the second patch. From the obtained distributions, the aperture slots are found to realize significant effects on the surface current distributions. Consequently, such observation shows a significant effect on the rectenna gain to realize a broadside directional radiation pattern as seen in the 3D far field in Figure 5(b). The rectenna gain is improved 6.2 dBi and 7.12 dBi at 2.45 GHz and 5.8 GHz, respectively.

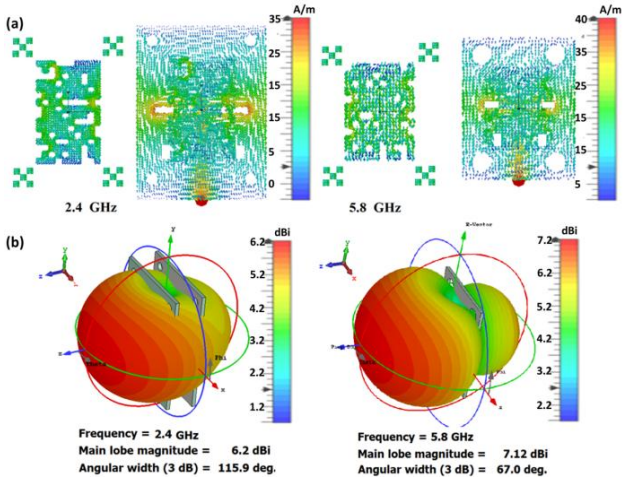


Figure 5. Simulated results for the proposed rectenna at 2.45GHz and 5.8GHz: (a) Surface current distributions and (b) 3D radiation patterns.

IV. RF ENERGY HARVESTING MEASUREMENTS

Figure 6 shows the experimental setup for testing how well the RF–DC energy harvesting works. The optimized dual-band rectenna is connected to the RFD102A-TB RF harvester module by a SMA coaxial cable. The tests took place in a semi-anechoic RF chamber to reduce outside noise and reflections. The RF source was a signal generator (Agilent E8257D PSG), and the power amplifier (Mini-Circuits ZHL-42W) was used to control and boost the RF power sent at the right frequency bands. The rectenna was placed 1.5 meters away from the transmitting antenna (a standard gain horn antenna with 10 dBi gain) and was set up so that both antennas would be as direct as possible. The transmitting antenna received continuous-wave (CW) signals at 2.45 GHz and 5.8 GHz. We looked at three different RF power levels at the rectenna terminals: 0 dBm, 8 dBm, and 16 dBm. To make sure that the calibration was consistent, we used a power meter (Keysight N1914A) to measure the corresponding incident power densities.

A high-impedance digital voltmeter connected to the RFD102A-TB output terminals was used to measure the rectified DC output voltage. The harvested voltage was recorded over several measurements for each input power level to make sure the results were consistent and averaged for accuracy. We also used the Agilent E5071C Vector Network Analyzer (VNA) to get the measured S11 spectra for the rectenna–harvester system. This was done to make sure that the impedance was matched under the same test conditions. In these controlled tests, the highest DC output voltage that could be harvested was 2.27 V at 2.45 GHz and 2.3 V at 5.8 GHz when the RF power was 16 dBm. The RF-to-DC conversion efficiencies were 52% and 56%, respectively. This shows that the GA-optimized dual-band rectenna design works well for collecting energy from Wi-Fi signals below 6 GHz.

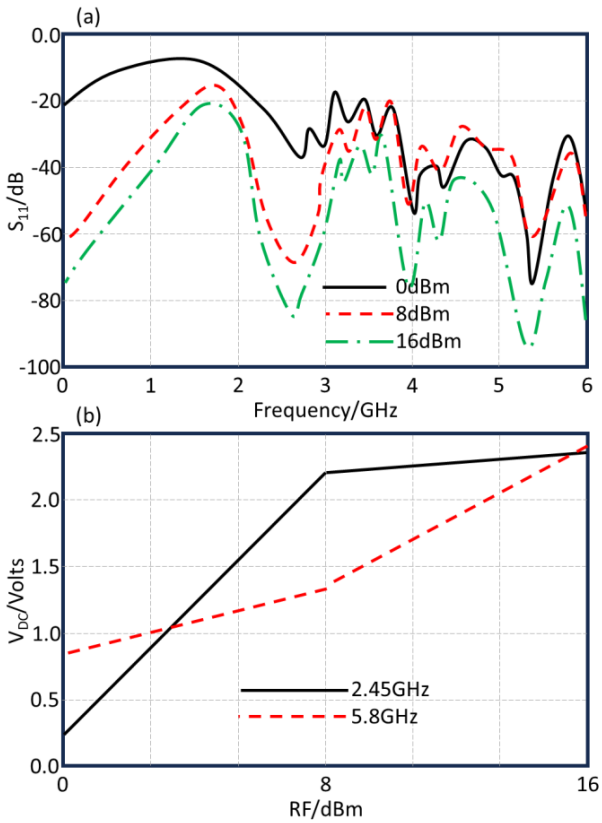


Figure 6: The evaluated results at the considered input RF energy: (a) S_{11} spectra and (b) Harvested DC voltage.

V. Results Comparison and Discussion

In Table 1, a comparison between the proposed work and other published designs is listed. The comparison between the proposed rectenna performance with their relatives is considered in terms of size, substrate, efficiency, frequency band, gain, and V_{DC}. It is observed that the proposed rectenna shows an excellent size reduction with an observable gain enhancement and harvested VDC. This realizes an advancement over other published designs using traditional techniques. The archived results in Table 1 indicate that the proposed rectenna achieves a significant gain enhancement at 2.45 GHz and 5.8 GHz by controlling the surface current on the rectenna patch, as illustrated in Figure 5(a). This gain enhancement is reflected in the energy harvesting that reached 2.3 volts, as seen in Figure 6(b).

TABLE I
Comparison of the RF energy harvesting systems with the proposed work.

Size/ mm²	Substrate	Efficiency/ %	Freq./ GHz	Gain/ dBi	V_{DC}/Volt	Reference
28×32	Taconic	50-80	2.4, 5.8	5-8	0.5-1.5	**[23]**
280×280	Taconic	42	2.45	11	3.4	**[24]**
60×70	Polymer	55	3.4	5.5	2	**[25]**
55×74	FR4	49-62	2.4, 5.8	2-4	0.4-1.2	**[26]**
38×52	Polymer	70-90	2-3	1.8-4	1.0-3.0	**[27]**
68×102	FR4	45	0.9	2.4	0.8	**[28]**
45×68	FR4	65-90	2.45, 5.8	3-5	0.7-1.5	**[29]**
27×30	FR4	52-56%	2.45, 5.8	6.2-7.12	2.3	**proposed work**

VI. Conclusion

The application of BCGA to antenna design represents a breakthrough for renewable energy research, opening new perspectives in RF energy harvesting. In the framework of this work, we have proposed a backup as RF energy harvesting for different applications, including wireless sensor networks. The use of hybrid RF energy harvesting systems prevents different realizations of low-energy devices that were previously not possible to be functionalized. The proposed design uses rectennas with enhanced characteristics that are designed with the aid of the BCGA scheme. AI-based designs of the rectennas are developed using the GA with the BC method to realize enhanced bandwidth and gain. The proposed rectenna shows a compact and low cost with enhanced performance to provide efficiency of 52% and 56% at 2.48 GHz and 5.8 GHz, respectively. The output DC power of the RF energy harvesting system is increased by using dual frequency bands. The proposed rectenna size is miniaturized to 27×30 mm² with a harvested real voltage of 2.3 volts. Finally, the measured and simulated results show excellent agreement, validating the proposed design methodology.

Acknowledgment: The authors would like to express their thanks to the International Applied and Theoretical Research Center, (IATRC), Baghdad, Iraq.

References

[1] M. S. Abood, W. Hua, B. S. Virdee, D. He, M. Fathy, A. A. Yusuf, O. Jamal, T. A. Elwi, M. Alibakhshikenari, L. Kouhalvandi, and A. Ahmad, "Improved 5G network slicing for enhanced QoS against attack in SDN environment using deep learning," *IET Communications*, 2023, ISSN: 1751-8636.

[2] M. S. Abood *et al*., "An LSTM-based network slicing classification future predictive framework for optimized resource allocation in C-V2X," *IEEE Access*, vol. 11, pp. 129 300–129 310, 2023, DOI: 10.1109/ACCESS.2023.3332225.

[3] R. K. Abdulsattar, S. M. Sadeq, T. A. Elwi, Z. A. A. Hassain, and M. Y. Muhsin, "Artificial neural network approach for estimation of moisture content in crude oil by using a microwave sensor," *Int. J. Microwave & Optical Technology*, vol. 18, no. 5, pp. 511–518, 2023.

[4] T. A. Elwi, N. M. Noori, and M. N. Majeed, "On the performance of adaptive intelligent wireless sensor nodes nanostructured array for IoT applications," *Int. J. Telecommunications & Emerging Technologies*, vol. 9, no. 2, 2023.

[5] A. M. Al-Saegh *et al*., "AI-based investigation and mitigation of rain effect on channel performance with aid of a novel 3D slot array antenna design for high throughput satellite system," *IEEE Access*, vol. 12, pp. 29 926–29 939, 2024, DOI: 10.1109/ACCESS.2024.3368829.

[6] W. En-Naghma, H. Halaq, and A. El Ougli, "Design and optimization using genetic algorithms of a dual-band microstrip antenna based on defective ground structure for wireless power transmission applications," Aug. 31, 2024.

[7] K. Karthika, K. Anusha, K. Kavitha, and D. M. Geetha, "Optimization algorithms for reconfigurable antenna design," 2023.

[8] A. M. Arun and M. R. E. Jebarani, "Efficient design and analysis of microstrip patch antenna array with real-coded genetic algorithm," 2024.

[9] Hassain, Z. A. A.; Farhan, M. J.; Elwi, T.A.; Mocanu, I.A. Design and Optimization of an Inductive-Stub-Coupled CSRR for Non-Invasive Glucose Sensing. *Sensors 2025*, 25, 7592. DOI: 10.3390/s25247592.

[10] Marwah H. Jwair, Taha A. Elwi, Salam K. Khamas, Aydin Farajidavar, Alyani Binti Ismail, "Circularly Shaped Metamaterial Fractal Reconfigurable Antenna for 5G Networks", *Iraqi Journal of Information and Communication Technology*: Vol. 6 No. 3 (2023): Iraqi Journal Of Information and Communication Technology.

[11] A. R. Hamad *et al*., "Rectenna design optimized by binary genetic algorithm for hybrid energy harvesting applications across 5G sub-6 GHz band," in *Radio Science,* vol. 60, no. 6, pp. 1–15, June 2025, DOI: 10.1029/2024RS008154.

[12] K. Abouhssous, L. Wakrim, A. Zugari, and A. Zakriti, "A three- band patch antenna using a defected ground structure optimized by a genetic algorithm for modern wireless mobile applications," *Jordanian J. Comput. Inf. Technol*., vol. 9, no. 1, pp. 11–20, 2023.

[13] A. Dejen, J. M. J. W. Jayasinghe, M. Ridwan, and J. Anguera, "Genetically engineered tri-band microstrip antenna with improved directivity for mm-wave wireless applications," 2021.

[14] Yahiea Al Naiemy, Aqeel N Abdulateef, Ahmed Rifaat Hamad, Mohammed Saadi Ismael, Balachandran Ruthramurthy, Taha A Elwi, Lajos Nagy, and Thomas Zwick "A Further Realization of Binary Genetic Algorithm to Design a Dual Frequency Band Rectenna for Energy Harvesting in 5G Networks", *DJES*, vol. 18, no. 2, pp. 203–214, Jun. 2025, DOI: 10.24237/djes.2024.18213.

[15] Amna S. Kamel, Ali S. Jalal, "Reconfigurable Monopole Antenna Design Based On Fractal Structure for 5G Applications", *Iraqi Journal of Information and Communication Technology*: Vol. 1 No. 1 (2021): Special Issue-Conference series: ARIE2021.

[16] N. Herscovici and M. F. Osorio, "Miniaturization of rectangular microstrip patches using genetic algorithms," *IEEE Antennas Wireless Propag. Lett*., vol. 1, pp. 94–97, 2002.

[17] A. Reddaf, F. Djerfaf, K. Ferroudji, M. Boudjerda, K. Hamdi-Chérif, and I. Bouchachi, "Modeling of electromagnetic behavior of composite thin layers using genetic algorithm," *Math. Comput. Simul*., vol. 167, pp. 281–295, 2020.

[18] J. M. Johnson and Y. Rahmat-Samii, "Genetic algorithms in engineering electromagnetics," *IEEE Trans. Antennas Propag*., vol. 39, no. 1, pp. 7–21, 1997.

[19] D. S. Weile and E. Michielssen, "Genetic algorithm optimization applied to electromagnetics: A review," *IEEE Trans. Antennas Propag*., vol. 45, pp. 343–353, 1997.

[20] J. M. J. W. Jayasinghe and D. N. Uduwawala, "A broadband triple-frequency patch antenna for WLAN applications using genetic algorithm optimization."

[21] V. Renuga Kanni and R. Brinda, "Design of high gain microstrip antenna for vehicle-to-vehicle communication using genetic algorithm," *Prog. Electromagn. Res. M*, vol. 81, pp. 167–179, 2019.

[22] M. Lamsalli, A. E. Hamichi, M. Boussouis, N. A. Touhami, and T. Elhamadi, "Genetic algorithm optimization for microstrip patch antenna miniaturization," *Prog. Electromagn. Res. Lett*., vol. 60, pp. 113–120, 2016.

[23] T. A. Elwi, "A systematic study on the metamaterial microstrip antenna design for self-powered wireless systems," in *Energy Harvesting in Wireless Sensor Networks and Internet of Things*, vol. 124, pp. 125–135, 2021.

[24] H. Almizan, M. H. Jwair, Y. Al Naiemy, Z. A. A. Hassain, L. Nagy, and T. A. Elwi, "Novel metasurface-based microstrip antenna design for gain enhancement RF harvesting," *Infocommunications Journal*, vol. 15, no. 1, pp. 2–8, 2023.

[25] M. N. N. Alaukally, T. A. Elwi, and D. C. Atilla, "Miniaturized flexible metamaterial antenna of circularly polarized high gain-bandwidth product for RF energy harvesting," *Int. J. Commun. Syst*., vol. 35, no. 3, Art. no. e5024, 2022, DOI: 10.1002/dac.5024.

[26] T. S. A. Al-Rawe, T. A. Elwi, and Ö. Ü. D. K. Türeli, "A dual-band high efficiency fractal rectenna for RF energy harvesting systems," in *Proc. 5th Int. Congr. Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Istanbul, Türkiye, 2023, pp. 1–4, DOI: 10.1109/HORA58378.2023.10156661.

[27] T. A. Elwi, "A further realization of a flexible metamaterial-based antenna on nickel oxide polymerized palm fiber substrates for RF energy harvesting," *Wireless Personal Communications*, vol. 115, pp. 1623–1634, 2020, DOI: 10.1007/s11277-020-07646-y.

[28] S. Ghosh, "Design and testing of RF energy harvesting module in GSM 900 band using circularly polarized antenna," in *Proc. IEEE Int. Conf. Research in Computational Intelligence and Communication Networks (ICRCICN)*, 2015, pp. 386–389.

[29] G. P. Ramesh and A. Rajan, "Microstrip antenna designs for RF energy harvesting," in *Proc. Int. Conf. Communication and Signal Processing (ICCSP)*, Melmaruvathur, India, 2014, pp. 1653–1657, DOI: 10.1109/ICCSP.2014.6950129.

Yasmeen Taha Yaseen was born in January 1991. She received the B.Sc. degree in Computer Engineering from Al-Mamoun University College, Baghdad, Iraq, in 2017. She obtained the M.Sc. degree in Computer and Electrical Engineering from Altınbaş University, Istanbul, Turkey, in 2020. She is currently pursuing the Ph.D. degree with the School of Electrical and Electronic Engineering, Universiti Sains Malaysia, since 2023. Her research interests include artificial intelligence, embedded systems, and smart technologies.

Mohdfe Fadzli Bin Mohd Salleh worked as a Software Engineer at the Research and Development Department (RnD), Motorola Penang, (1996-2001). Currently, he is an Associate Professor at the School of Electrical and Electronic Engineering, USM. His main research interests include source coding and signal processing for applications in telecommunications and wireless communication networks. He has supervised 12 Ph.D. degree students to graduation as principal supervisor.

Taha A. Elwi received his B.Sc. in Electrical Engineering Department (2003) (Highest Graduation Award), and Postgraduate M.Sc. in Laser and Optoelectronics Engineering Department (2005) (Highest Graduation Award) from Al-Nahrain University Baghdad, Iraq. From April 2005 to August 2007, he worked with Huawei Technologies Company, in Baghdad, Iraq. On January 2008, he joined the University of Arkansas at Little Rock and he obtained his Ph.D. in December 2011 in system engineering and Science. He is considered of Stanford University's top 2% scientists in 2022. His research areas include wearable and implantable antennas for biomedical wireless systems, smart antennas, WiFi deployment, electromagnetic wave scattering by complex objects, design, modeling, and testing of metamaterial structures for microwave applications, design and analysis of microstrip antennas for mobile radio systems, precipitation effects on terrestrial and satellite frequency re-use communication systems, effects of the complex media on electromagnetic propagation and GPS. His research is conducted to consider wireless sensor networks based on microwave terminals and laser optoelectronic devices. The nano-scale structures in the entire electromagnetic spectrum are a part of his research interest. Also, his work is extended to realize advancements in reconfigurable intelligent surfaces and control the channel performance. Nevertheless, the evaluation of modern physics phenomena in wireless communication networks including cognitive radio networks and squint effects is currently part of his research. His research interests include pattern recognition, signal and image processing, machine learning, deep learning, game theory, and medical image analysis-based artificial intelligence algorithms and classifications. He serves as an editor in many international journals and publishers like, MDPI, IEEE, Springer, and Elsevier. He is currently the head of the International Applied and Theoretical Research Center (IATRC), Baghdad Quarter, Iraq. Also, he has been a member of the Iraqi scientific research consultant since 2016. He is leading three collaborations around the world regarding biomedical applications using microwave technology. He is the supervisor of many funded projects and Ph.D. theses with corresponding of more than 150 published papers and holding 10 patents.

Significant performance improvement in polarization-diversity GMSK reception using atypical filters

Ádám Kiss, and László Schäffer

***Abstract*—This paper presents novel signal processing methods to enhance the reception performance of Gaussian Minimum Shift Keying (GMSK) signals from pico- and nanosatellites, emphasizing software-defined approaches over hardware upgrades. Atypical filtering techniques, including phase-domain median and FIR filtering, as well as polarization-diverse multi-channel methods, are explored and evaluated. Real-world experiments were conducted using coherently sampled dual-polarization channels from a ground station in Szeged, Hungary, receiving transmissions from the MRC-100 satellite. Various single- and multi-channel preprocessing strategies were benchmarked using packet decoding success and bit-error rates. Results show that non-linear phase filtering and blind source separation techniques, notably FastICA, significantly increase the number of correctly decoded packets – achieving up to a 16% improvement compared to conventional demodulation without preprocessing. This study demonstrates the utility and relative independence of these methods and highlights their potential for improving satellite data throughput with no hardware modification. These techniques are suitable for integration into existing ground stations to enhance data reception performance.**

***Index Terms*—CubeSat, GMSK, Independent Component Analysis, Median filter, PocketQube, Polarization-diversity.**

I. INTRODUCTION

SEVERAL approaches are available for transferring data from a space satellite. Yet, the most notable example is the use of radio signals, particularly sub-gigahertz frequencies, in educational and research satellites. One classic, widely used modulation is Gaussian Minimum Shift Keying (GMSK) [1] on the 70 cm radio amateur band and on the industrial satellite band on slightly lower frequencies. The typical effective radiated power is 100 mW, with an average path loss of around 150 dB. To ensure a reliable transmission, a robust error-correcting method and sufficient signal levels are required. If a specific satellite is already in orbit, the only way to increase the data rate is to build a better reception architecture. Although one possible improvement of the ground station is to enhance antenna dimensions and optimize filter characteristics, realizing these are not cost-efficient. The primary motivation of this research was to identify and analyze receiver algorithms that enhance nano- and pico-satellite reception without requiring an investment in a more expensive reception system. To the best of our knowledge, no published results are using such non-linear filters to improve GMSK reception. The Independent Component Analysis (ICA) algorithm was used for separating parallel transmissions [2], [3], [4]. However, utilizing it to select the desired signal from a noisy environment is poorly covered.

Department of Technical Informatics, University of Szeged, Hungary
Corresponding author: Ádám Kiss (e-mail: kissadam@inf.u-szeged.hu).

DOI: 10.36244/ICJ.2025.4.4

In this paper, following a brief introduction to the measurement setup, some non-linear methods are introduced in Section III to preprocess the radio signals, which only require a single channel as input. A separate section –numbered IV– presents methods operating on two coherently sampled, independent channels. As we aim to enlarge the size of downloaded data, we quantitatively compare the number of successfully decoded correct packets and qualitatively compare the bit-error rate of the same packet among methods. Every comparison is shown with the reference state-of-the-art GMSK demodulator. On amateur radio bands, it is common to conduct experiments like this to investigate radio channel performance. [5] Therefore, we have chosen a satellite operating on these bands.

II. MEASUREMENT SETUP

The experiments presented in this paper were conducted in the central area of Szeged, Hungary, in an environment where TV and FM broadcast transmitters, along with various cellular base stations, loaded the receiver input in addition to the useful signal. Generally, the presence of such strong signals can easily overload the input amplifier of a radio receiver to a non-linear domain. Therefore, the pre-selected radio filters like the one below should be used.

The actual measurements were made in the 70 cm radio amateur band, with the radio tuned to the MRC-100 [6] satellite, which operates at 436.72 MHz. This satellite was chosen because a reliable, coherent, soft-decision, state-of-the-art GMSK demodulator software called “smogcli2” [7] is compatible with it, which can also output the bit-error ratio of the demodulated packets.

The front of the reception chain consisted of two linear 10-element Yagi-Uda antennas, each with a 30-degree-wide main lobe. These antennas were mounted perpendicularly to each other, so their main lobes pointed in the same direction, but their linear polarization planes were orthogonal to each other.

In addition to differences in polarization, the two signal paths exhibited variation in their hardware stages, specifically through the incorporation of distinct filters. One polarization plane’s path (designated as “Channel A”) included only a Surface Acoustic Wave (SAW) filter, followed by a low noise amplifier. The other path (designated as “Channel B”) consisted of a low-Q, low insertion loss cavity filter, a low insertion loss SAW filter, and the same low-noise amplifier. Both signal paths were terminated in a software-defined radio configured to a coherent sampling rate of 1 MHz.

All the antennas were mounted onto an antenna rotator with a positioning accuracy of approximately ten degrees. It was controlled by software named "Gpredict" [8] using the solutions of Kepler-equations provided by a space tracking service, namely the North American Aerospace Defense Command (NORAD). [9] The rotator was stepped in five degrees.

Other aspects of the reception is presented in [10].

After input filtering, Doppler correction was performed coherently on both channels using the data from NORAD. We should note that, theoretically, the correction could be done without the exact trajectory data. [11] The data were then resampled and saved at a sample rate of 50 kHz, suitable for processing with "smogcli2". The length of specific windows should also be related to this symbol rate. It is essential to note that the satellite sent various packets at different baud rates. In this research, packets sent at a rate of 12500 bauds were only considered in the Results section, as they were dominant during the operation of the MRC-100.

Various methods are investigated, encompassing a wide range of diverse scenarios. The IQ output of each scenario was processed using "smogcli2". Alongside the packets, the corresponding bit-error ratios were recorded.

To provide a reference for comparing the proposed non-linear polarization-diversity methods, several linear and single-channel techniques were also evaluated. The signal processing procedures run in 0.8-second-long overlapped windows, stepped in a raster of 0.2 seconds. Shorter windows resulted in the loss of certain packets even if the experimental procedures studied herein were omitted. In contrast to this, if longer windows had been applied, the incoherence of the reception would have increased (phase incoherency occurring from the change of the angle of incidence).

III. SINGLE-CHANNEL METHODS

In the case of single-channel methods, only one signal path was considered, and the two signal paths were not combined in any way to mutually improve reception quality.

The word "band limiting" in this article references a kind of filtering, which was always performed in the single-channel methods. A Finite Impulse Response (FIR) filter with a 12 kHz cut-off frequency was applied, designed with a 3 kHz transition band using Blackman-window [12]. The bandwidth of such a filter was 24 kHz in the complex spectrum, which included most of the 12500 baud GMSK signal power [1], taking into account also the possible inaccuracy of the Doppler correction.

A. Channel A and Channel B

These cases represent the raw received packets without any preprocessing.

B. Channel data with median filtered phase

This case is a non-linear method, running on a band-limited channel and acting on the phase of the recorded IQ signals. The amplitude of the IQ signal remains unchanged, but its unwrapped phase is median filtered [13] using 3, 5, or 7 sample-wide windows. The window sizes were chosen to fit the trivial median definition; therefore, only odd, wide windows were used. Windows wider than four samples – corresponding to durations longer than a single symbol – begin to introduce intersymbol interference, which gradually leads to a decline in performance. Windows longer than 7 samples were omitted in this study.

In terms of mathematics

$$|Out[k]| = |Input[k]| \tag{1}$$

$$j = \frac{N-1}{2} \tag{2}$$

$$\arg(Out[k]) = Median_{i=k-j}^{k+j} \arg(Input[i]) \tag{3}$$

Where N is the length of the filter, k is the time dimension, and $\arg$ is the unwrapped phase of its argument.

C. Channel data with FIR filtered phase

This case is also a non-linear method, running on a band-limited channel and acting on the phase of the recorded IQ signals like in the "Channel data with median filtered phase" case, but this time, the unwrapped phase is FIR filtered by 3, 5, or 7 sample wide rectangle-shaped windows with unity gain.

In terms of mathematics, using equation 1 and 2:

$$\arg(Out[k]) = \sum_{i=k-j}^{k+j} \frac{1}{N} \arg(Input[i]) \tag{4}$$

IV. MULTI-CHANNEL METHODS

In the case of the multi-channel methods, the two available, coherently sampled, diverse signal paths were combined using different methods.

A. Averaging Channel A and Channel B

In this case, the IQ data of the two channels are averaged sample by sample. Assuming that the same signal is on both channels, this increases the signal-to-noise ratio. This case is essentially equivalent to stacking Yagi-Uda antennas [14], a common practice in satellite tracking ground stations.

In terms of mathematics

$$Out[k] = ChannelA[k] + ChannelB[k] \tag{5}$$

B. Averaging Channel A and Channel B with path correction

In this paper, path correction refers to a phase correction that aligns the phase of channel B with that of channel A, thereby maximizing the average output amplitude of the sum of the two channels.

In this case, the sum of the two channels after the path correction forms the output.

In terms of mathematics

$$Out[k] = ChannelA[k] + ChannelBC[k] \tag{6}$$

Where $ChannelBC$ is rotated in phase to have the highest correlation with $ChannelA$.

C. Averaging Channel A and Channel B in the quadrature-amplitude space

Averaging is not only applicable to the IQ form of the signal but also to its amplitude and phase. Aimed to investigate its effect, the two channels are averaged separately by their amplitudes and phases.

In terms of mathematics

$$|Out[k]| = |ChannelA[k]| + |ChannelB[k]| \quad (7)$$

$$\arg Out[k] = \arg(\exp^{j \arg ChA[k]} + \exp^{j \arg ChB[k]}) \quad (8)$$

D. Median filtered phase of the averaged signals

In this case, the two band-limited and path-corrected channels are averaged sample by sample. The amplitudes remain intact, while the phase of the averaged signals is median-filtered by three and five sample-wide median filters.

Essentially, this entire case represents a concatenation of methods IV-B and III-B.

E. Independent Component Analysis

In earlier solutions, blind source separation techniques have been used on GMSK signals in Automatic Identification Systems [2], [3] and multi-input multi-output systems [4].

In this case, the band-limited and path-corrected IQ signals are converted to an intermediate frequency, producing real signals for each channel. These real signals are processed with the FastICA [15] algorithm, which separates the original signals into a non-gaussian component with the highest variance and an additional residual component, both of which are treated as output.

F. Independent Component Analysis with median filtered phase

In this case, the output of the FastICA algorithm is downconverted to zero intermediate frequency again, and the quadrature components of the results are median filtered with three and five sample wide windows (as in case III-B).

V. RESULTS

The results were examined by both the number of successfully decoded packets and the bit-error rate of them. Venn diagrams were created to visualize the independence and performance of the algorithms based on the found packets. The algorithms were compared by the bit-error rate of the simultaneously found packets with the aid of correlograms. A 23-day observation of the MRC-100 satellite signals was conducted for evaluation, collected between October 21, 2023, and November 12, 2023. [16]

The results can be interpreted both quantitatively and qualitatively. We can measure how many packets with good checksums the demodulator found after the bit-error correction and investigate whether an algorithm found much more CRC-correct data in a raw radio recording than other algorithms. Furthermore, we should also note if an algorithm has found not as many data packets as others, but a significant number of them is unique among the listed methods. On the other hand, we can compare specific methods by examining the quality (in terms of the bit-error rate) of commonly found packets using correlograms.

We used the FastICA implementation in Scikit-Learn 1.6.0 [17] with the default parameters. The random seed was set to 42 to ensure reproducibility.

A. Single-input methods

According to the Venn diagrams [18], no unique packets were found when the 5- and 7-bin moving average FIR filter was applied to the phase. To enhance the readability of Fig. 1, they were omitted.

Fig. 1 shows that most of the packets uniquely found by these methods resulted from the median filtering of the unwrapped phase with three and five sample-wide windows. Using those algorithms, the number of packets increases by 101 on Channel A (4.3% increase relative to demodulating only the raw data) and by 101 on Channel B (3.7% increase relative to demodulating only the raw data).

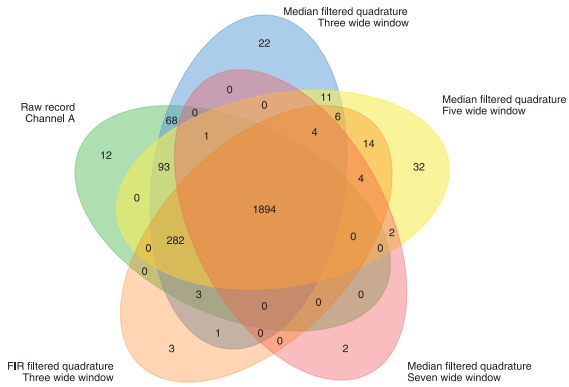


Fig. 1. Comparison of single-input methods on recordings. The number inside each bubble represents the number of packets detected using the corresponding method. 101 of 2453 packets were missed without the proposed methods.

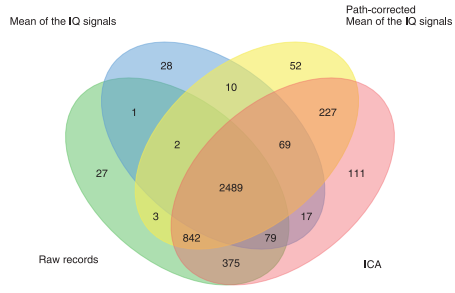


Fig. 2. Comparison of multi-input linear methods. 215 of 4322 packets were missed when only the raw radio signal recordings were considered.

B. Multiple input methods

Fig. 2 shows a significant number of unique packets (514, 13.5% relative increase to demodulating only the raw data

on the two channels) when using the following methods: averaging the IQ signals, averaging the path-corrected IQ signals, and the ICA.

Fig. 3 shows that median filtering the phase on the averaged IQ signals, median filtering the phase of the ICA results, and quadrature averaging methods also significantly increase the number of packets (551, 14.4% increase relative to demodulating only the raw data on the two channels).

The summary of the highlighted methods is shown in Fig. 4. The group "single channel methods" include the raw recordings and the filtered phase raw recordings. The group "mean like methods" covers the addition of the two signals and the path-corrected addition of the two signals. The group "quadrature-oriented methods" includes the quadrature-amplitude averaging method with and without a phase median filter on its output. 10.7% of the packets were missed without the proposed multi-path methods related to the cumulated performance of raw recordings and the single-channel methods.

The Venn diagrams show the number of packets detected by these methods from the raw recordings and measure their contribution to the amount of data downloaded. There are several packets that appear as a result of multiple decoding methods, these can be compared by their quality. One objective measure of the quality is the bit-error ratio.

Fig. 5 and 6 show correlations [19] between the bit-error ratio of the commonly decoded packets. Co-decoded packets, mainly scattered near the equal performance line, suggest that the two methods strongly correlate in terms of bit-error ratios. If most of the packets deviate below the equal performance line, then we can conclude that co-decoded packets exhibit a lower bit-error rate on the vertical axis. Therefore, comparing the co-decoded packets, the method labeled horizontally (in the row) outperforms the one labeled vertically (in the column).

The cumulative gain of these algorithms is 611 packets (16% relative to the packets from the non-preprocessed recordings). The Venn diagram and the correlograms both show that these methods are partly independent.

C. Processing time

The implementation was done in NumPy [20] environment. For one – approximately 10 minutes long – pass of the MRC-100 satellite, evaluation of all the listed methods took an hour on a standard desktop computer (i7-9700KF, 30 GiB RAM), which was less than the revisit period. Preliminary tests show that by utilizing parallel computation methods (such as GPUs), quasi-real-time operation is achievable.

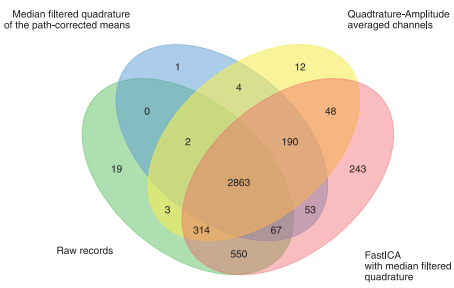


Fig. 3. Comparison of multi-input phase filtering methods. 551 of 4369 packets were missed without the proposed methods.

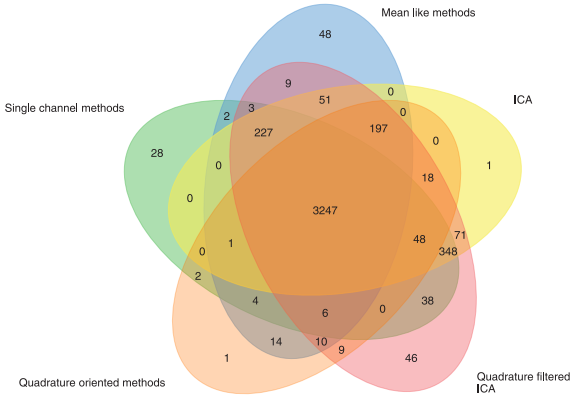


Fig. 4. Comparison of all presented methods. The number inside each bubble represents the number of packets detected using the corresponding method groups. Numerous packets are found only by different methods, furhermore these methods are partly independent.

VI. Discussion

The results allow us to evaluate the effectiveness of various methods and compare their performance. The Venn diagrams highlight that a significant number of packets were detected and successfully decoded exclusively by the proposed methods. Using the bit-error-rate correlograms, we can analyze the performance of multiple methods on the common packets. If there is no correlation between the bit-error rates, we can infer that the algorithms were practical under different scenarios. Conversely, a high correlation would suggest that the two methods perform with similar efficiency. Additionally, correlograms where points fall consistently above or below the equal bit-error-ratio line indicate that one method outperforms the other in terms of bit-error rate.

As Fig. 5 shows, the seven bin wide windows underperformed the other methods according to bit-error-ratios. Also, Fig. 1 shows that only two new data packets were detected using the seven-wide filters. This further suggests that more than five wide windows could be omitted, most probably due to intersymbol interference.

Although single-channel algorithms are easy to implement in software-defined radio systems, the now-rare hardware radios might need further adjustments or upgrades. Applying multichannel methods to polarity-diverse reception stations with coherent sampling radio is also straightforward.

VII. Conclusion

In this paper, we investigated the performance of different atypical filters in a real environment on GMSK signals from space. Furthermore, we introduced and compared trivial and non-trivial polarization diversity methods. The FastICA algorithm and median filtering of signal phases are highly adaptable for use in diverse reception systems and recommended for

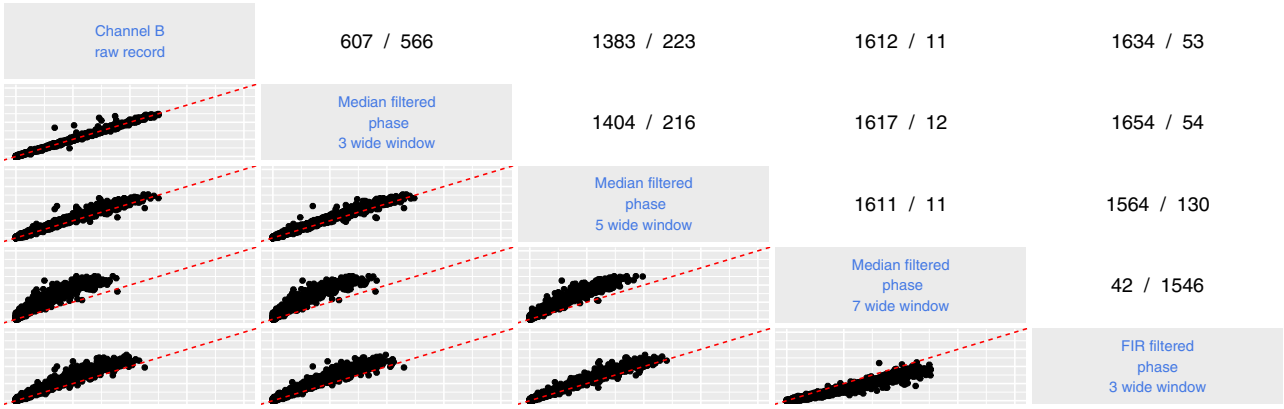


Fig. 5. Correlation between the bit-error rate in different single-input methods. The vertical and horizontal scales range from 0 to 20% on every subgraph, indicating the bit-error ratio. The horizontal axis is labeled along the same column; the vertical axis is labeled along the same row. A dashed equal-performance line is plotted for reference. The bottom-left corner contains the packets without bit error. This shows that the three sample long median filtering does not alter the BER for the co-found packets. The upper triangle shows how many co-found packets have better BER in the method named along the row (first number) and how many are named along the column (second number).

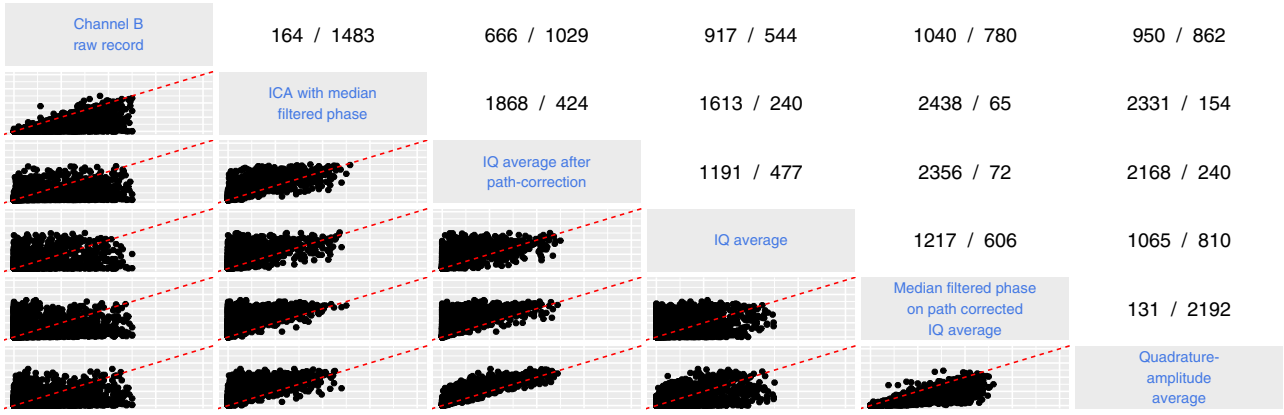


Fig. 6. Correlation between the bit-error-rate in different multi-input methods. Square-like scatter plot indicates independency, while triangle like ones indicate better BER using one method than the other.

adaptation in existing systems. While ICA effectively covers the problem of multi-antenna receptions as a clear use case, the underlying mechanism of median filtering is less clear and requires further research. Other algorithms were also found to be useful, though less effective. With sufficient processing time, a ground station could apply all these methods to the received signals to gather more data. In the future, other blind source separation algorithms, such as JADE or Infomax, should be considered and compared. In summary, implementing these techniques significantly increased the volume of data downloaded in the system by 16% as described above. These algorithms (especially the single-channel ones) can be easily implemented and integrated on GMSK receiver sites to increase downloading performance.

ACKNOWLEDGMENT

The reception system supplying the data for this research was provided by IKIKK, Centre of Excellence for Interdisciplinary Research, Development and Innovation of the University of Szeged, and Csiha Innovation and Technology Plc., Szeged, Hungary.

The research was supported by project TKP2021-NVA-09, which has been implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-NVA funding scheme.

REFERENCES

[1] K. Murota and K. Hirade, "GMSK modulation for digital mobile radio telephony," *IRE Trans. Commun. Syst.*, vol. 29, no. 7, pp. 1044–1050, Jul. 1981. DOI: 10.1109/TCOM.1981.1095089.

[2] "A steady FastICA algorithm based on modified-m estimate function," in *Proceedings of 2016 the 6th International Workshop on Computer Science and Engineering, WCSE*, 2016. DOI: 10.18178/wcse.2016.06.026.

[3] A. Krishna, A. Nimbal, A. Makam A, and V. Sambasiva Rao, "Implementation of fast independent component analysis on field-programmable gate array for resolving the slot collision issue in the space-based automatic identification system," en, *Int. J. Satell. Commun. Netw.,* vol. 38, no. 6, pp. 480–498, Nov. 2020. DOI: 10.1002/sat.1362.

[4] M. Yu, C. Li, B. Xu, and Y. Li, "GMSK modulated DSSS signal separation based on principal component analysis," in *2020 IEEE 20th International Conference on Communication Technology (ICCT)*, Nanning, China: IEEE, Oct. 2020. DOI: 10.1109/ICCT50939.2020.9295670.

[5] E. Laport, E. Tilton, and R. Rowe, "Amateur radio," *IEEE Communications Magazine*, vol. 19, no. 4, pp. 16–24, 1981. DOI: 10.1109/MCOM.1981.1090543.

[6] Y. A. I. Humad and L. Dudás, "Resonant radar reflector on vhf / uhf band based on bpsk modulation at leo orbit by mrc-100 satellite," Infocommunications journal, vol. 16, no. 1, pp. 26–34, 2024, ISSN: 2061-2079. DOI: 10.36244/icj.2024.1.4. [Online]. Available: http://dx.doi.org/10.36244/ICJ.2024.1.4.

[7] M. Maroti and P. Horváth, Smogcli2, 2023. DOI: 10.5281/zenodo.10207615.

[8] A. R. Periyapatna et al., "Automation of a satellite earth station," in *2023 IEEE 8th International Conference for Convergence in Technology (I2CT)*, Lonavla, India: IEEE, Apr. 2023, pp. 1–6. DOI: 10.1109/I2CT57861.2023.10126424.

[9] S. Speretta, P. Sundaramoorthy, and E. Gill, "Long-term performance analysis of norad two-line elements for cubesats and pocketqubes," Apr. 2017.

[10] Á. Kiss, A. Tönköly, and R. Mingesz, "Descriptor: Thermal sensor comparison on the mrc-100 satellite (mrc100szte)," *IEEE Data Descriptions*, pp. 1–9, 2025. DOI: 10.1109/IEEEDATA.2025.3624344.

[11] T. Herman and L. Dudás, "Picosatellite identification and doppler estimation using passive radar techniques," *Infocommunications Journal*, vol. 15, no. 3, pp. 11–17, 2023, ISSN: 2061-2079. DOI: 10.36244/icj.2023.3.2. [Online]. Available: http://dx.doi.org/10.36244/ICJ.2023.3.2.

[12] S. Chakraborty, "Advantages of blackman window over hamming window method for designing fir filter," *International Journal of Computer Science & Engineering Technology*, vol. 4, no. 08, 2013. [Online]. Available: https://api.semanticscholar.org/CorpusID:14186524.

[13] B. I. Justusson, "Median filtering: Statistical properties," in *Topics in Applied Physics*, Berlin/Heidelberg: Springer-Verlag, 2006, pp. 161–196. DOI: 10.1007/ BFb0057597.

[14] L. Dudas, L. Papay, and R. Seller, "Automated and remote controlled ground station of masat-1, the first hungarian satellite," in *2014 24th International Conference Radioelektronika*, Bratislava, Slovakia: IEEE, Apr. 2014. DOI: 10.1109/Radioelek.2014.6828410.

[15] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," en, *Neural Netw.*, vol. 13, no. 4-5, pp. 411–430, May 2000. DOI: 10.1016/s0893-6080(00)00026-5.

[16] Á. Kiss, Doppler-corrected iq signals from the mrc-100 satellite, 2024. DOI: 10.21227/8mr2-jw13.

[17] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[18] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing in science & engineering*, vol. 9, no. 03, pp. 90–95, 2007.

[19] J. W. Emerson et al., "The generalized pairs plot," en, *J. Comput. Graph. Stat.*, vol. 22, no. 1, pp. 79–91, Jan. 2013. DOI: 10.1080/10618600.2012.694762.

[20] C. R. Harris et al., "Array programming with NumPy," *Nature,* vol. 585, no. 7825, pp. 357–362, Sep. 2020. DOI: 10.1038/s41586-020-2649-2. [Online]. Available: DOI: 10.1038/s41586-020-2649-2.

Ádám Kiss is a PhD student at the Department of Physiology, University of Szeged. He earned his degree in electrical engineering, specializing in telecommunications, from the Budapest University of Technology and Economics. His research focuses primarily on physiological signal processing and radio communications.

László Schäffer obtained his B.Sc. degree in engineering information technology in 2013, his M.Sc. degree in computer science and technology in 2015, and his PhD in computer science in 2022 from University of Szeged. Currently he is an assistant professor at the Department of Technical Informatics, University of Szeged. His research interests include FPGA based prototype development, real-time neurophysiological and digital signal processing, and embedded systems.

AMC-Transformer: Automatic Modulation Classification based on Enhanced Attention Model

Yuewen Xu

***Abstract*—High-accuracy automatic modulation classification (AMC) is essential for spectrum monitoring and interference-aware access in future 6G systems [1]. We propose AMC-Transformer, which tokenizes raw I/Q sequences into fixed-length patches, augments them with learnable positional embeddings, and applies multi-layer, multi-head self-attention to capture global temporal–spatial correlations without handcrafted features or convolutions. On RadioML2018.01A, our model achieves 98.8% accuracy in the high-SNR regime (SNR at least 10 dB), showing higher accuracy than a CNN and a ResNet reimplementation by 4.44% and 1.96% in relative terms; averaged across all SNRs, it also improves upon MCformer, CNN, and ResNet baselines. Consistent gains are observed on the RadioML2016.10A dataset, further validating robustness across benchmarks. Ablations on depth, patch size, and head count provide practical guidance under different SNR regimes and compute budgets. These results demonstrate the promise of transformer-based AMC for robust recognition in complex wireless environments.**

***Index Terms*—Modulation Recognition, Deep Learning, Transformer, Attention Mechanism, IQ Signal.**

I. INTRODUCTION

Wireless signal recognition—also known as automatic modulation classification (AMC)—is pivotal across military and civilian scenarios. It enables identifying modulation types from raw RF signals under limited prior knowledge, supporting dynamic spectrum access (DSA), interference detection, spectrum monitoring, and spectrum coexistence. Moving toward 6G, AMC becomes even more critical for improving spectrum utilization, robustness, and low-overhead (pilot-free) communications. [2–3]

Classical AMC approaches fall broadly into *likelihood-based (LB)* and *feature-based (FB)* families. LB methods (e.g., ML/EM-assisted inference, HLRT/QHLRT variants) can achieve high accuracy in favorable conditions but are often sensitive to channel state information and carry significant complexity. FB methods draw on expert features such as cyclostationary statistics and higher-order cumulants, offering lower complexity and near-optimal performance for lower-order schemes, yet they struggle in multipath, overlapping sources, and high-order modulations. [4–7]

Deep learning (DL) has boosted AMC in both supervised CNNs and newer foundation-model-style approaches [8]. CNNs exploit multi-scale structures and constellation geometry effectively (e.g., constellation-image CNNs; robust multi-scale designs under synthetic channel impairments). ResNet-style networks further mitigate vanishing gradients and improve feature reuse; lightweight/binarized ResNets demonstrate competitive accuracy–efficiency trade-offs for edge deployment. [9–13]

Transformers (TRN) provide an alternative by modeling long range dependencies via self-attention with efficient parallelism. Beyond their foundational success, time series surveys highlight their strengths for sequence tasks. In AMC, Transformer variants such as MCformer, CNN Transformer hybrids, and CNN Transformer GNN adaptively weight multi scale patterns and improve robustness and scalability in non-cooperative settings [14-18].

We propose the AMC-Transformer, a transformer-based model designed for time-series IQ samples to improve AMC accuracy. Key contributions include:

1. Learnable Embedding of RF Signal Patches: To represent the time-series nature of raw IQ samples, we design a learnable embedding strategy that combines patch and positional information. RF signals are segmented into fixed-size patches, mapped into the feature space via an MLP, and augmented with positional embeddings to preserve temporal dependencies, enabling effective attention-based modeling.

2. Attention Mechanism on Raw Time-Series IQ Data: We apply self-attention directly to raw IQ data, enabling the model to capture long-range dependencies. This enables the model to capture long-range dependencies and temporal–spatial correlations, extracting global representations without handcrafted features or convolutional operations, thus overcoming the locality limitations of CNN-based methods.

3. Enhanced Diversity and Robustness with Multi-Head Attention: To improve generalization under varying SNR conditions, we employ multi-head attention, allowing feature extraction in multiple subspaces. This enriches representation diversity and enhances robustness against noise and channel impairments, improving classification reliability in realistic wireless environments.

4. Competitive performance on public RadioML datasets: On RadioML2018.01A [19] we obtain 98.8 percent accuracy at SNR at least 10 dB and observe higher average accuracy than MCformer, CNN, and ResNet. Similar trends appear on RadioML2016.10a [19]. We also provide ablations on depth, patch size, and head count.

Department of Engineering, The University of Bristol, Bristol, UK
(E-mail: tu23081@bristol.ac.uk)

DOI: 10.36244/ICJ.2025.4.5

The proposed AMC-Transformer provides a robust attention-based solution for AMC, demonstrating competitive performance across multiple modulation scenarios. At the same time, the increased model complexity introduced by attention mechanisms highlights an inherent accuracy–complexity trade-off, which is particularly relevant for practical 6G deployments. Despite these advantages, transformer models face challenges such as quadratic complexity, higher data requirements, and limited invariance to signal distortions. To enhance practicality and scalability, future work will focus on efficient attention mechanisms, hybrid Conv–Attention architectures, and RF-specific data augmentation strategies.

The paper is structured as follows: Section 2 summarizes related research. Section 3 details the AMC-Transformer architecture. Section 4 presents evaluation results. Section 5 concludes the paper.

II. RELATED WORK

Deep learning–based AMC has advanced markedly in recent years. On the CNN side, one-dimensional residual networks for I/Q sequences and complex-valued convolutions can extract discriminative features while keeping parameter counts manageable; for example, ResNet-style variants tailored to wireless signals and complex depthwise-separable CNNs report strong results on RadioML benchmarks [20,21]. Meanwhile, MCNet—using asymmetric kernels and skip connections—achieves about 93% accuracy at high SNR (20 dB) on RadioML2018.01A, illustrating the upper bound of CNNs in high-SNR regimes [22]. RNN/CRNN and LSTM models have also been used to capture long-range dependencies, but their generalization to unseen channel conditions and modulation parameters remains limited [23–25]. RadioML datasets (e.g., RML2018.01A, RML2016.10a/10b) continue to be the standard benchmarks in this area.

Transformer-based AMC has recently evolved in three directions. First, sequence models operating directly on raw I/Q: Cai et al. apply a Transformer to AMC and report consistent gains over CNN/LSTM baselines—especially at low SNR—with fewer parameters; MCformer embeds each (I, Q) sample via a lightweight 1-D convolution and stacks Transformer encoders, with the notable observation that omitting positional encodings works better; it attains state-of-the-art accuracy on RML2016.10b with only ~10k–72k parameters [16]. Second, hybrid CNN–Transformer designs: CTGNet/CTRNet use convolutions for local invariances and self-attention for long-range dependencies, improving robustness under multiple impairments and non-idealities [26]. Third, ViT on 2-D signal representations: by converting signals to constellation images, MobileViT and related ViT variants improve robustness under noise without an explicit denoising pipeline (e.g., NMformer) [27,28].

In addition, scalability and label efficiency have been advanced via meta-learning and semi/self-supervision: Meta-Transformer provides a general few-shot adaptation framework for previously unseen modulations, and subsequent studies further validate meta-learning for cross-domain generalization [29,30]. Transformer-based contrastive semi-supervised learning and self-supervised RF representation learning (e.g., Self-Contrastive, NextG RF SSL) substantially reduce labeled-data requirements while maintaining accuracy in low-label regimes [31–33].

III. MODEL DESCRIPTION

The AMC-Transformer is tailored to 2-D in-phase/quadrature (IQ) signals and addresses two challenges that limit conventional CNN/ResNet models on RF data: (i) high-frequency noise and (ii) long-range temporal dependencies. As summarized in Fig. 1, the model converts a 2×1024 IQ sample into fixed-length tokens via patching, augments them with positional encodings, and processes the sequence using a Transformer encoder whose output feeds an MLP head for prediction.

A. Input Processing

We adopted a minimal, task-compatible preprocessing pipeline per sample $x \epsilon R^{2\times 1024}$: including: (i) per-channel DC offset removal, (ii) RMS normalization across I/Q channels, and (iii) channel-wise z-score standardization using statistics estimated from the training split only. The same normalization parameters are then applied to validation and test data to avoid information leakage. Explicit filtering or denoising is intentionally avoided to preserve modulation-discriminative spectral and phase characteristics.

$$x'_{c,t} = x_{c,t} - \mathrm{mean}_t\left(x_{c,t}\right)$$
$$\tilde{x}_{c,t} = \frac{x'_{c,t}}{\sqrt{\frac{1}{2T}\sum_c \sum_t \left(x'_{c,t}\right)^2 + \varepsilon}} \quad (1)$$
$$\hat{x}_{c,t} = \frac{\tilde{x}_{c,t} - \mu_c}{\sigma_c + \varepsilon}$$

where (μ_c, σ_c) are the channel-wise mean and standard deviation estimated on the training split after steps (i)–(ii) and then fixed for validation and test sets, and ε is a small constant for numerical stability.

B. Overall Pipeline

The preprocessed signal $\hat{x}$ is treated as a two-channel 2-D array 2×1024. We tokenize it into $N = 1024$ non-overlapping $2 \times P$ patches (covering both I and Q to retain I/Q coherence), add learned positional embeddings, and process tokens with a stack of Transformer encoder blocks (multi-head self-attention and MLP, each preceded by layer normalization and followed by dropout). An MLP head produces the final logits. Fig. 1 is updated to include the preprocessing block.

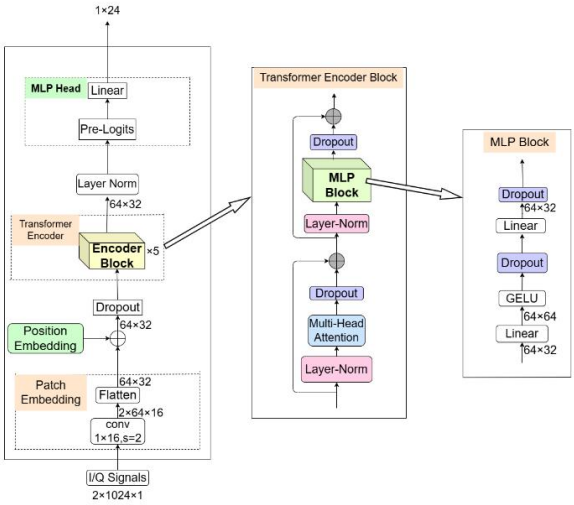


Fig. 1. Model Overview

C. Patch Embedding

To balance context coverage and efficiency, the input is partitioned along the time axis into non-overlapping patches of size 1×16, yielding 64 patches per sample. Each patch contains 2×16 values that are flattened to a 32-D vector and linearly projected to the model width d (we use $d = 32$), producing a token sequence of shape $64 \times d$. This tokenization lets the self-attention mechanism relate local structures (e.g., short-term amplitude/phase transitions) to global patterns influenced by modulation type and SNR. The choice 16 ensures sufficient local resolution under symbol-rate offsets, delays, and noise fluctuations while keeping compute tractable.

D. Positional Encoding

Because Transformer blocks process all tokens in parallel, explicit position information is required to capture temporal dynamics (amplitude, frequency, and phase evolutions). We use learnable positional embeddings of length 64 and dimension d; the positional vector for each patch is added to its token embedding, yielding $E \in R^{64\times d}$. This enables the model to distinguish early/late patches and to learn temporal patterns associated with different modulations and SNRs.

E. Self-Attention

Given the token matrix E, the encoder computes query, key, and value projections

$$Q = EW_Q$$
$$K = EW_K$$
$$V = EW_V$$

where $Q, K \in R^{64\times d_k}$ and $V \in R^{64\times d_v}$.

Scaled dot-product attention (Fig. 2) is

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (2)$$

The output $Z \in R^{64\times d_v}$ preserves sequence length while re-weighting each token by its global correlations. For token i, the vector z_i aggregates values v_j according to the similarity between q_i and k_j, thereby encoding long-range dependencies across the entire 2×1024 signal.

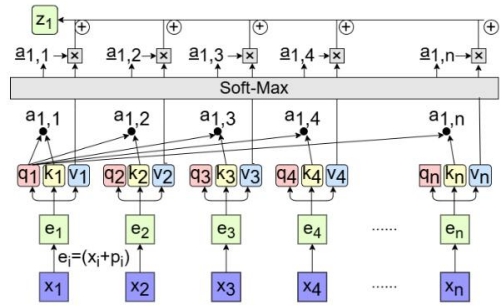


Fig. 2. Scaled Dot-Product Attention

F. Multi-Head Attention and Classifier

To learn complementary temporal and frequency relations. For head i,

$$Q_i = QW_i^Q, K_i = KW_i^K, V_i = VW_i^V, i = 1,2 \qquad (3)$$

with per-head dimension $d_k = d_v = 8$. Each head produces the heads are concatenated and projected:

$$MultiHead(Q,K,V) = Concat(head_1, head_2)W^O \qquad (4)$$

$$head_i = Attention\left(QW_i^Q, KW_i^K, VW_i^V\right) \qquad (5)$$

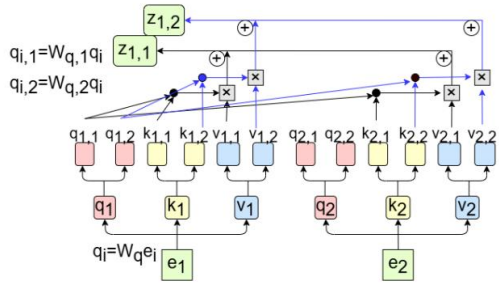


Fig.3. Multi-Head Attention

restoring the model width $d = 32$. Each encoder block applies LayerNorm, multi-head self-attention, dropout, and an MLP with residual connections (see Fig. 3). The encoder output is passed to an MLP head (linear layers with normalization/dropout) to obtain pre-logits and final predictions (e.g., via softmax for classification)

IV. EXPERIMENT AND RESULTS

A. Datasets and Split Protocol

We use the publicly available RadioML 2018.01A dataset, which synthesizes realistic channel impairments—including delay spread, carrier frequency offset, and thermal noise. It covers 24 modulation types across 26 SNR levels from −20 dB to 30 dB in 2 dB steps. Each SNR level contains 4,096 signal examples, yielding a total of 2,555,904 samples, each represented as complex IQ (in-

phase and quadrature) sequences. In addition to training and validation on 2018.01A, we perform cross-dataset evaluation on RadioML 2016 to assess the algorithm's robustness and generalization under distribution shifts and differing channel conditions.

To prevent sample leakage across SNR conditions, we perform a group-aware split of RadioML2018.01A. Let each waveform be indexed by its modulation $m \in M$, SNR $s \in S$, and within-class index $i \in \{0, \ldots, 4095\}$. We define a group as

$$G_{m,i} = \{(m, s, i)|\}$$

i.e., the same base waveform rendered at all SNRs for a given modulation. Splitting is conducted at the group level so that no group $G_{m,i}$ appears in more than one subset, eliminating leakage where the same underlying waveform at different SNRs would otherwise straddle train and evaluation sets.

We adopt a group-aware split to avoid cross-SNR leakage: within each modulation, base examples are grouped across all SNRs and treated as indivisible units. Groups are randomly assigned to 70% / 20% / 10% train/validation/test with a fixed seed (48), stratified by modulation to preserve class priors. Because each group spans the full SNR set, the SNR distribution is preserved across splits by construction. At the per-(modulation, SNR) level this yields approximately 2867 / 819 / 410 samples for train/val/test, respectively (rounded from 4096 per pair).

Class IDs are remapped to [0, 23] following the fixed 24-class list in Sec. A, ensuring a stable label order aligned with the classifier's output layer.

We apply lightweight, task-compatible preprocessing: (i) per-sample DC offset removal on I/Q channels; (ii) per-sample RMS normalization (AGC-style) to unit average power across I/Q; and (iii) channel-wise z-score standardization using training-split statistics (μ_c, σ_c) only. The same normalization parameters are then applied to validation and test data to avoid information leakage. Random seeds and the exact split indices are fixed and recorded to ensure reproducibility.

B. Baselines and training protocol

We re-train all baselines (CNN[34], ResNet[3], MCformer[16]) and our AMC-Transformer under the same preprocessing and training protocol. All models take identical inputs (I/Q, shape 2×1024), use the same loss (multiclass cross-entropy), optimizer and learning-rate schedule, and share the same group-aware data split. Specifically, within each modulation, base examples are grouped across all SNRs, and each group is assigned wholly to train/validation/test (70/20/10), which prevents cross-SNR leakage while preserving class priors. Early stopping and weight decay are applied to mitigate overfitting.

CNN: A 2D ConvNet consisting of four sequential stages including ABlock, BBlock, CBlock1, and CBlock2, followed by global average pooling and a 24-way classifier. The model contains 66,008 parameters.

ResNet: A 1D ResNet with residual connections, featuring an initial Conv1D layer followed by 5 residual blocks with progressive channel expansion from 32 to 64 to 128 channels. The architecture uses kernel size 7, batch normalization, and ReLU activations. The final layers consist of global average pooling followed by dropout and a dense classifier. The model contains 534,104 parameters.

MCformer: A hybrid architecture combining Conv1D with 8 channels and 4 lightweight encoder blocks, followed by temporal aggregation to 4 tokens. The output is processed through flattening, a 128-dimensional fully connected layer, and finally a 24-dimensional classification layer. The parameter counts increases from 10,050 for the original 10-class head to 11,856 for 24 classes, with the increase attributed to the expanded classifier.

AMC-Transformer (ours): The input is reshaped to dimensions 2×1024×1 and divided into 64 patches of 32 dimensions each. The architecture employs an embedding dimension of 96 with positional encoding, followed by 6 encoder layers with 8 attention heads each. The final process consists of flattening followed by a multi-layer perception with layer dimensions 6144, 2048, 1024, and 24. The model contains 15,834,680 parameters.

All models are trained using the AdamW optimizer with a learning rate of 1e-3, cosine decay scheduling with 5-epoch warm-up, weight decay of 1e-4, and gradient clipping at 1.0. The default batch size is 256, with additional results reported using batch size 800. Dropout of 0.1 is applied to MLP and classifier layers. Input data undergoes z-score normalization with no data augmentation applied. Hardware specifications, random seeds, and library versions are documented in the Appendix. All code and training scripts are provided to ensure reproducibility.

TABLE I
MODEL AND ARCHITECTURE OVERVIEW.

Model	Params	Tokens / Patch	Heads	Dim	Blocks
CNN	66,008	–	–	–	A/B/C ×4 → GAP → FC
ResNet-1D	534,104				Conv1D → 5×ResBlock → GAP → FC
MCformer-24 (reimpl.)	11,856	T-agg → 4	–	–	Conv1D → 4×Enc → FC
AMC-Trans (ours)	15,834,680	64 / 16	8	96	6×(MHA+FFN) → MLP

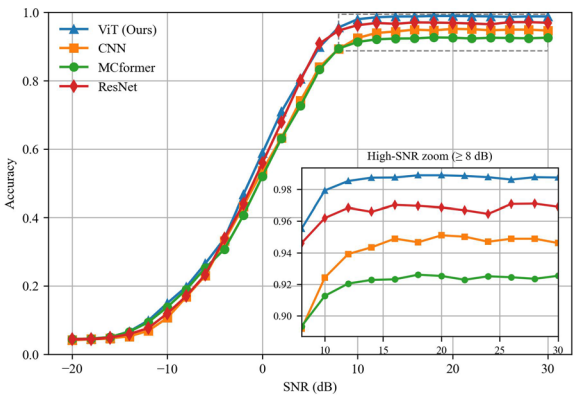


Fig. 4. Accuracy versus SNR on the RML2018.01A dataset

Over the full SNR range, AMC-Transformer achieves an average accuracy of 63.58%, higher than MCformer at 59.02%, CNN at 59.84%, and ResNet at 61.96% (Fig. 4.) In the high-SNR region (SNR of at least 10 dB), AMC-Transformer reaches 98.8% and exhibits a clear saturation plateau, higher than MCformer at 92.29%, CNN at 94.50%, and ResNet at 96.79%. In the low-SNR region (SNR at most -8 dB), AMC-Transformer attains 20.97%, higher than MCformer at 19.19%, CNN at 18.70%, and ResNet at 19.37%. In the mid-SNR range from 2 to 8 dB, AMC-Transformer averages 84.20%, higher than MCformer at 77.07% and CNN at 77.66%, and essentially on par with ResNet at 83.32%.

Overall, AMC-Transformer maintains robustness at low SNR and sustains a consistent performance margin as SNR increases, with a near-saturated accuracy around 98.8% on RML2018.01A in the high-SNR region.

C. Robustness Analyses

1. Accuracy Across SNRs on RML2016.10a

RML2016.10a (Fig. 5). Over the full SNR range, the average accuracies of AMC-Transformer and MCformer are essentially identical (63.48% and 63.48%). In the high-SNR region (SNR of at least 10 dB), AMC-Transformer attains an average accuracy of 93.51%, which is higher than MCformer by 0.82 percentage points and higher than CNN and ResNet by 3.00 and 6.37 percentage points, respectively. In the low-SNR region (SNR at most -10 dB), AMC-Transformer reaches 29.12%, comparable to MCformer at 29.06% and higher than CNN at 25.68% and ResNet at 23.70%

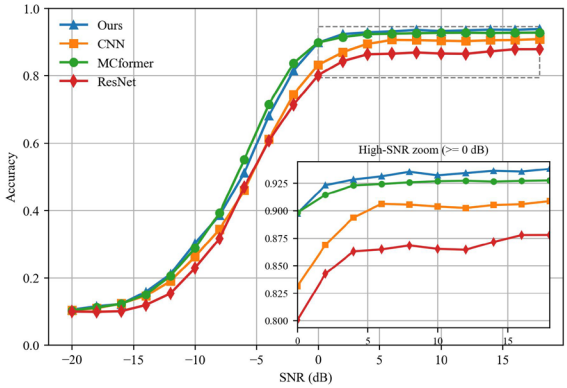


Fig. 5. Evaluation on RML2016.10a Across SNRs

2. Macro-F1 Stability Over Random Seeds

Having established performance trends across SNRs, we next test whether these gains persist under different random initializations. To address the concern that the proposed model only achieves high accuracy under favorable conditions, we further evaluate its robustness across the full SNR range. Fig. 6 shows the Macro-F1 scores from −20 dB to 30 dB, averaged over five independent runs with different random seeds. In addition to the high-SNR regime (20–30 dB), where AMC-Transformer attains near-saturation performance, the model maintains competitive robustness under mid and low SNR conditions. For example, at -10 dB and 0 dB, the Macro-F1 remains above 10.8% and 57.4%, respectively, with narrow confidence intervals, indicating stable generalization across noise levels. This result confirms that the performance of AMC-Transformer is not restricted to high SNRs but extends to more challenging communication environments as well.

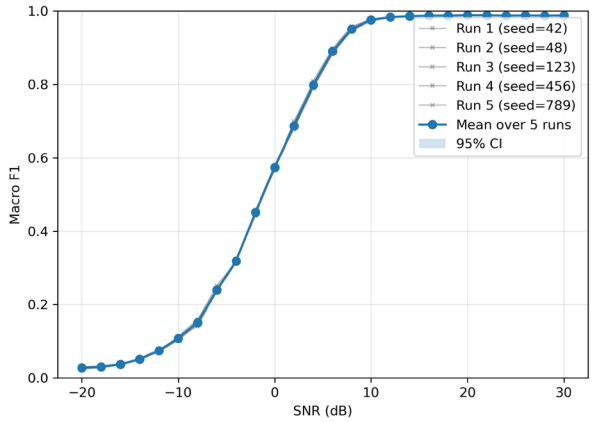


Fig. 6. Robustness to Random Initialization: Macro-F1 over Five Seeds.

3. Per-Class F1 Across SNRs

To further examine the robustness of the AMC-Transformer, we report per-class F1-scores across different SNR levels. The horizontal axis corresponds to SNR values (−20 dB to 30 dB), the vertical axis lists the 24 modulation types, and the color intensity indicates the F1-score.

Overall, fig.7 shows that F1-scores consistently increase with SNR. Low-order modulations such as BPSK and QPSK remain relatively robust even at low SNR (−10 dB), whereas higher-order QAM schemes suffer significant degradation under noise but quickly recover above 0 dB. Importantly, the model maintains competitive per-class F1 performance in the mid-SNR regime (0–10 dB), demonstrating that its effectiveness is not limited to high SNR conditions.

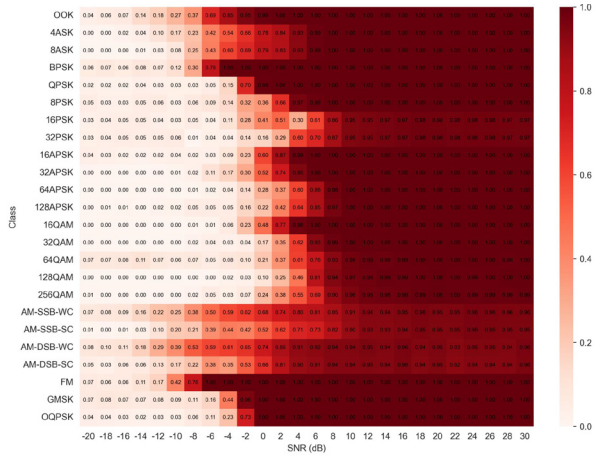


Fig. 7. Per-class F1-scores across SNR levels

D. Analysis of AMC-Transformer Model Parameter Tuning

To systematically evaluate the impact of key hyperparameters on AMC-Transformer performance, we conducted a parameter sensitivity study across four critical dimensions: batch size, transformer layer depth, patch size, and number of attention heads. The baseline configuration used a learning rate of 0.001, a batch size of 256, 100 training epochs, three transformer layers, a patch size of 32, and two attention heads, achieving an overall accuracy of 58.30% and an average accuracy of 89.31% at SNR above 10 dB.

During tuning, only one parameter was varied at a time, while the others were fixed at the baseline values. This single-factor analysis is adopted to provide interpretable sensitivity trends for each design choice under controlled conditions, while we acknowledge that hyperparameters may be coupled. Joint hyperparameter optimization (e.g., Bayesian optimization) could be explored in future work to more efficiently search the coupled space; however, the focus here is to characterize the main effects and practical ranges of key parameters. This single-factor analysis is adopted to provide interpretable sensitivity trends for each design choice under controlled conditions, while we acknowledge that hyperparameters may be coupled. Joint hyperparameter optimization (e.g., Bayesian optimization) could be explored in future work to more efficiently search the coupled space; however, the focus here is to characterize the main effects and practical ranges of key parameters.

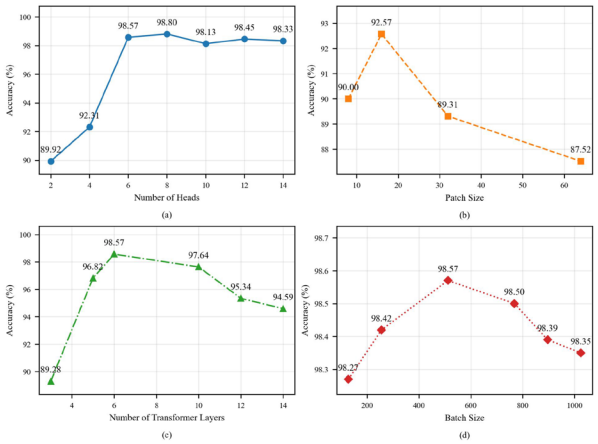


Fig. 8. Hyperparameter sensitivity analysis of AMC-Transformer. (a) Batch size vs. Accuracy; (b) Transformer layer depth vs. Accuracy; (c) Patch size vs. Accuracy; (d) Number of attention heads vs. accuracy

In addition to accuracy, varying the number of layers and attention heads directly changes model complexity (parameter count and compute), whereas batch size mainly affects optimization dynamics and patch size trades temporal resolution against sequence length. Therefore, the following results are discussed from both accuracy and complexity perspectives, which is particularly relevant for low-overhead 6G deployment scenarios.

1) Batch Size

Table 1 shows the relationship between batch size and classification accuracy. A baseline batch size of 256 yields 58.30 percent accuracy. Increasing the batch size gives only marginal gains, with peak performance observed around the 512–732 range; further increases lead to a slight degradation. This suggests the model benefits from more stable gradient updates, but overly large batches reduce helpful stochasticity.

2) Patch Size

Patch size strongly influences feature resolution. As shown in Fig. 8(c), reducing the patch size from 32 to 16 markedly improves accuracy to 92.57 percent, while a patch size of 8 produces 90.00 percent under baseline conditions. Larger patch sizes such as 64 degrade performance to 87.52 percent due to loss of fine-grained temporal features.

3) Transformer Layer Depth

Fig. 8(b) illustrates that accuracy improves steadily as the number of transformer layers increases, up to 10 layers where it reaches about 91.64 percent. Beyond this point, performance plateaus or slightly decreases, reflecting a trade-off between representational capacity and the risk of overfitting. Moreover, deeper stacks increase parameters and attention compute roughly linearly with depth, so the marginal accuracy gains beyond 6–10 layers should be weighed against the added complexity.

4) Multi-Head Attention

The number of attention heads has a pronounced effect. As seen in Fig. 8(d), accuracy rises quickly from 2 heads, where the baseline is 58.30 percent, to 6 heads, which achieves 92.57 percent. It then stabilizes around 8 to 14 heads near 90.8 percent and declines slightly thereafter, indicating that a moderate number of heads captures diverse signal dependencies without introducing redundancy. Since multi-head attention increases projection parameters and compute, the observed saturation beyond 8–14 heads indicate diminishing returns in accuracy relative to complexity.

Combining the best settings from each dimension — batch size 512, transformer layers 6, patch size 16, and attention heads 8 — yields an overall accuracy of 63.87 percent and an average accuracy of 98.80 percent at SNR greater than 10 dB (as shown in Table 2). This represents a notable improvement over the baseline, with an absolute overall gain of 5.57 percentage points and a gain of 9.49 points in the high-SNR regime.

TABLE II
EFFECT OF INDIVIDUAL PARAMETER OPTIMIZATION ON MODEL ACCURACY

Parameter	Baseline	Best Value	Accuracy (Overall)	Accuracy (SNR > 10 dB)
Batch Size	256	512	58.30 → 59.50	89.31 → 90.33
Layers	3	6	58.30 → 61.64	89.31 → 91.64
Patch Size	32	16	58.30 → 62.57	89.31 → 92.57
Heads	2	8	58.30 → 61.23	89.31 → 94.80
Combined	–	(6 layers, 8 heads, patch 16, batch 512)	63.87	98.80

To further illustrate the effect of hyperparameter optimization, Figure 9 presents the classification accuracy across the full SNR range for both the baseline and optimized configurations. While the baseline model saturates around 90% accuracy at high SNR levels, the optimized AMC-Transformer achieves up to 98.8% accuracy at SNR of at least 10 dB and shows consistent improvements in the mid-SNR range from 0 to 10 dB. This confirms that the performance gain is not restricted to very high SNR conditions, addressing concerns about robustness.

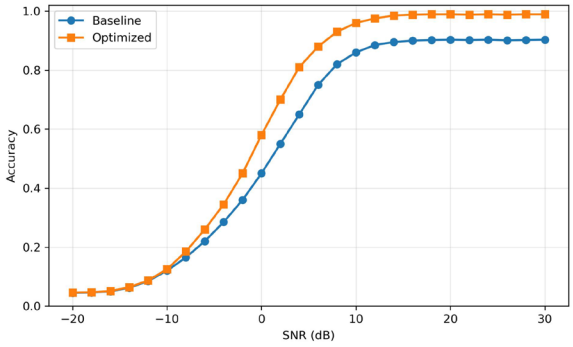


Fig. 9. Accuracy versus SNR for baseline and optimized AMC-Transformer.

To analyze the source of the performance gains, Fig. 10 and Fig. 11 contrasts class-wise confusion matrices before and after hyperparameter optimization across SNR ranges.

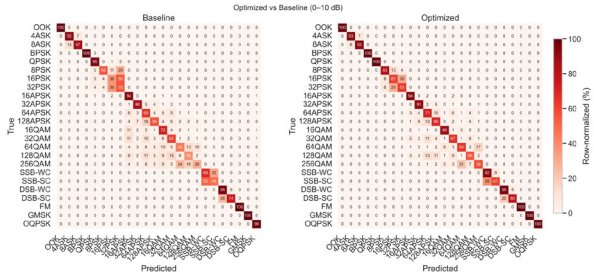


Fig. 10. Confusion matrices comparing baseline and optimized models at mid-to-low SNR range (5-15 dB).

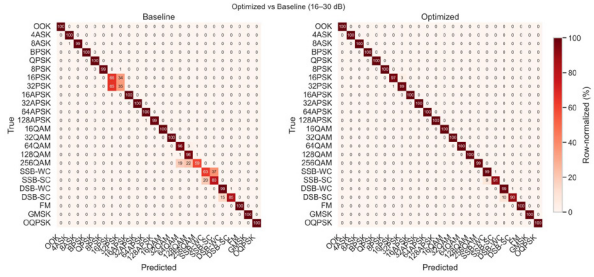


Fig. 11. Confusion matrices comparing baseline and optimized models at high SNR range (10-30 dB).

At higher SNRs (10–30 dB), the optimized model substantially reduces within-family confusion—most notably among high-order QAM constellations (e.g., 256-QAM), between adjacent PSK orders (16- vs. 32-PSK), and between AM subtypes (SSB vs. DSB). In the mid- to low-SNR regime (0–10 dB), the optimization primarily mitigates cross-family confusion, yielding sizable per-class recall gains. Overall, the confusion matrices indicate that the accuracy improvement arises from a systematic attenuation of characteristic misclassification patterns across the entire SNR spectrum, rather than from isolated gains at specific operating points.

Overall, the sensitivity results indicate that most of the achievable gains come from selecting an appropriate patch size and a moderate number of layers/heads, while very deep or heavily multi-headed configurations exhibit diminishing returns. Importantly, the optimized configuration improves accuracy across the full SNR range (Fig. 9) but does so with increased model complexity. This accuracy–complexity trade-off should be considered when targeting resource-constrained receivers and low-overhead 6G deployments.

E. Positional Encoding Strategy

To evaluate the impact of positional encoding methods on AMC-Transformer performance, we compared learnable positional embeddings against fixed sinusoidal encodings while keeping all other hyperparameters constant (6 layers, 8 heads, patch size 16, batch size 512).

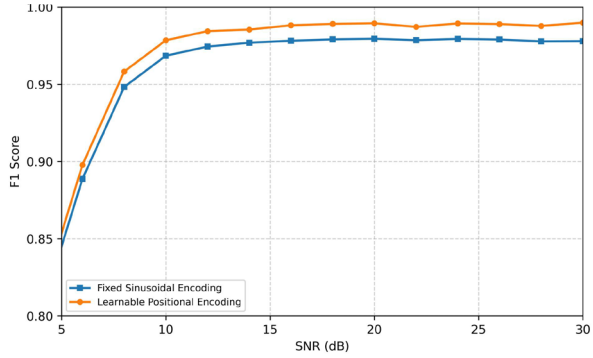


Fig. 12. Comparison of positional encoding strategies across SNR levels

Fig. 12 shows the F1 scores across the full SNR range for both encoding strategies. While both methods achieve comparable performance at low SNRs (below 0 dB), a clear divergence emerges in the mid-to-high SNR regime. Learnable positional encoding consistently outperforms fixed sinusoidal encoding above 10 dB SNR, maintaining an average F1 score of 98.5% compared to 97.3% for the fixed encoding—a relative improvement of 1.2 percentage points. The performance gap is most pronounced between 10-20 dB, suggesting that learnable embeddings better capture the position-dependent temporal patterns specific to modulated signals under favorable channel conditions.

This result indicates that allowing the model to learn task-specific positional representations provides measurable benefits for AMC, particularly when signal quality permits extraction of fine-grained temporal features. The learned embeddings likely adapt to the periodic structures and phase relationships inherent in different modulation schemes, which generic sinusoidal patterns cannot fully capture.

V. CONCLUSIONS

In this paper, we introduced AMC-Transformer, a transformer-based framework for automatic modulation classification that operates directly on raw I/Q time series. The model tokenizes I/Q sequences into fixed-length temporal patches, augments them with learnable positional embeddings, and applies multi-head self-attention to capture both short-range transitions and long-range dependencies in the waveform.

On RadioML2018.01A, our best configuration achieves 98.8% accuracy for SNR of 10 dB or higher and 63.9% on average across all SNRs, showing improved accuracy compared to reimplemented CNN/ResNet and MCformer baselines under the same data splits and training protocol. On RadioML2016.10a, AMC-Transformer maintains competitive accuracy across SNR levels and reaches 93.5% in the high-SNR regime, demonstrating robustness beyond a single dataset. Ablation studies indicate that model capacity and tokenization drive the accuracy–efficiency trade-off: patch size 16, 8 attention heads, and about six encoder layers offer a favorable balance across SNR conditions. These gains, however, come at the cost of increased model complexity, which should be carefully considered for low-overhead and resource-constrained 6G receivers.

Despite these gains, transformer attention scales quadratically and benefits from substantial data. Future work will explore efficient attention (for example, linear or clustered variants), hybrid Conv–Attention designs that inject local inductive bias, and RF-aware augmentation and self-supervision to improve robustness to channel non-idealities while reducing the need for labeled data.

REFERENCES

[1] A. Aboulfotouh, A. Eshaghbeigi and H. Abou-Zeid, "Building 6G Radio Foundation Models with Transformer Architectures," *ICC 2025 – IEEE International Conference on Communications*, Montreal, QC, Canada, 2025, pp. 1888–1893, DOI: 10.1109/ICC52391.2025.11161954.

[2] R. Ding, F. Zhou, H. Zhang, Q. Wu, and Z. Han, "Data- and knowledge dual-driven automatic modulation classification for 6G wireless communications," *IEEE Transactions on Wireless Communications*, 2024 (early access), DOI: 10.1109/TWC.2023.3316197.

[3] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168–179, Feb. 2018, DOI: 10.1109/JSTSP.2018.2797022.

[4] W. Wei and J. M. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *IEEE Transactions on Communications*, vol. 48, no. 2, pp. 189–193, Feb. 2000, DOI: 10.1109/26.823550.

[5] F. Hameed, O. A. Dobre, and D. C. Popescu, "On the likelihood-based approach to modulation classification," *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 5884–5892, Dec. 2009, DOI: 10.1109/TWC.2009.12.080883.

[6] W. Gardner, "Spectral Correlation of Modulated Signals: Part I – Analog Modulation," in *IEEE Transactions on Communications*, vol. 35, no. 6, pp. 584–594, June 1987, DOI: 10.1109/TCOM.1987.1096820.

[7] W. A. Gardner, "Spectral correlation of modulated signals: Part I—Analog modulation," *IEEE Transactions on Communications*, vol. 35, no. 6, pp. 584–594, Jun. 1987, DOI: 10.1109/TCOM.1987.1096820.

[8] C. A. Harper, M. A. Thornton, and E. C. Larson, "Automatic Modulation Classification with Deep Neural Networks," *Electronics*, vol. 12, no. 18, 3962, Sep. 2023, DOI: 10.3390/electronics12183962.

[9] B. Jdid, K. Hassan, I. Dayoub, W. H. Lim, and M. Mokayef, "Machine learning based automatic modulation recognition for wireless communications: A comprehensive survey," *IEEE Access*, vol. 9, pp. 57 851–57 873, 2021, DOI: 10.1109/ACCESS.2021.3071801.

[10] S. Peng, H. Jiang, H. Wang, H. Alwageed, Y. Zhou, M. M. Sebdani, and Y.-D. Yao, "Modulation classification based on signal constellation diagrams and deep learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 3, pp. 718–727, Mar. 2019, DOI: 10.1109/TNNLS.2018.2850703.

[11] T. Huynh-The, V. S. Doan, C.-H. Hua, Q.-V. Pham, and D.-S. Kim, "Chain-Net: Learning deep model for modulation classification under synthetic channel impairment," in *Proc. IEEE GLOBECOM*, 2020, pp. 1–6, DOI: 10.1109/globecom42002.2020.9322394.

[12] N. P. Shankar, D. Sadhukhan, N. Nayak, T. Tholeti and S. Kalyani, "Binarized ResNet: Enabling Robust Automatic Modulation Classification at the Resource-Constrained Edge," in *IEEE Transactions on Cognitive Communications and Networking*, vol. 10, no. 5, pp. 1913–1927, Oct. 2024, DOI: 10.1109/TCCN.2024.3391325.

[13] M. L. Xue, M. Huang, J. J. Yang, and J. D. Wu, "MLResNet: An efficient method for automatic modulation classification based on residual neural network," in *Proc. 2nd Int. Symp. on Computer Engineering and Intelligent Communications (ISCEIC)*, 2021, pp. 122–126, DOI: 10.1109/ISCEIC53685.2021.00032.

[14] A. Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017, pp. 5998–6008, *arXiv:1706.03762*.

[15] Q. Wen, T. Zhou, C. Zhang, et al., "Transformers in time series: A survey," in *Proc. Int. Joint Conf. on Artificial Intelligence (IJCAI)*, 2023, DOI: 10.24963/ijcai.2023/759.

[16] S. Hamidi-Rad and S. Jain, "MCformer: A Transformer-based deep neural network for automatic modulation classification," in *Proc. IEEE GLOBECOM*, 2021, pp. 1–6, DOI: 10.1109/globecom46510.2021.9685815.

[17] W. Ma, Z. Cai, and C. Wang, "A Transformer- and convolution-based learning framework for automatic modulation classification," *IEEE Communications Letters*, vol. 28, no. 6, pp. 1392–1396, Jun. 2024, DOI: 10.1109/lcomm.2024.3380623.

[18] D. Wang, M. Lin, X. Zhang, Y. Huang, and Y. Zhu, "Automatic modulation classification based on CNN–Transformer–GNN (CTGNet)," *Sensors*, vol. 23, no. 16, 7281, Aug. 2023, DOI: 10.3390/s23167281.

[19] Radioml2018.01A. [Online]. Available: https://www.deepsig.ai/datasets

[20] A. Abbas, V. Pano, G. Mainland and K. Dandekar, "Radio Modulation Classification Using Deep Residual Neural Networks," *MILCOM 2022 – 2022 IEEE Military Communications Conference (MILCOM)*, Rockville, MD, USA, 2022, pp. 311–317, DOI: 10.1109/MILCOM55135.2022.10017640.

[21] C. Xiao, S. Yang, and Z. Feng, "Complex-Valued Depth-wise Separable CNN for AMC," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, 2023, DOI: 10.1109/TIM.2023.3298657.

[22] T. Huynh-The, C.-H. Hua, Q.-V. Pham, and D.-S. Kim, "MCNet: An Efficient CNN Architecture for Robust AMC," *IEEE Communications Letters*, vol. 24, no. 4, pp. 811–815, Apr. 2020, DOI: 10.1109/LCOMM.2020.2968030.

[23] K. Liao, Y . Zhao, J. Gu, Y . Zhang, and Y . Zhong, "Sequential Convolutional Recurrent Neural Networks for Fast AMC," *IEEE Access*, vol. 9, pp. 27 182–27 188, 2021, DOI: 10.1109/ACCESS.2021.3053427.

[24] S. Ramjee *et al*., "Fast Deep Learning for Automatic Modulation Classification," *ArXiv, abs/1901.05850*.

[25] T. Wang, Z. Zhang, X. Wang, and Y. Li, "A Survey of Applications of Deep Learning in Radio Signal Modulation Recognition," *Applied Sciences*, vol. 12, no. 23, 12052, Nov. 2022, **DOI**: 10.3390/app122312052.

[26] W. Zhang *et al*., "CTRNet: An Automatic Modulation Recognition Based on CNN with Transformer," *Electronics*, vol. 13, no. 6, 1047, Mar. 2024, **DOI**: 10.3390/electronics13173408.

[27] Q. Zheng *et al*., "A Real-Time Constellation Image Classification Method of AMC Based on MobileViT," *Scientific Reports*, vol. 13, 18656, Oct. 2023, **DOI**: 10.1007/s11571-023-10015-7.

[28] A. Faysal, M. Rostami, R. G. Roshan, H. Wang and N. Muralidhar, "NMformer: A Transformer for Noisy Modulation Classification in Wireless Communication," *2024 33rd Wireless and Optical Communications Conference (WOCC)*, Hsinchu, Taiwan, 2024, pp. 103–108, **DOI**: 10.1109/WOCC61718.2024.10786062.

[29] J. Jang, J. Pyo, Y.-I. Yoon and J. Choi, "Meta-Transformer: A Meta-Learning Framework for Scalable Automatic Modulation Classification," in *IEEE Access*, vol. 12, pp. 9267–9276, 2024, **DOI**: 10.1109/ACCESS.2024.3352634.

[30] X. Hao, Z. Feng, S. Yang, M. Wang and L. Jiao, "Automatic Modulation Classification via Meta-Learning," in *IEEE Internet of Things Journal,* vol. 10, no. 14, pp. 12 276–12 292, 15 July15, 2023, **DOI**: 10.1109/JIOT.2023.3247162.

[31] W. Kong, X. Jiao, Y. Xu, B. Zhang and Q. Yang, "A Transformer-Based Contrastive Semi-Supervised Learning Framework for Automatic Modulation Recognition," in *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 4, pp. 950–962, Aug. 2023, **DOI**: 10.1109/TCCN.2023.3264908.

[32] D. Liu, P. Wang, T. Wang, and T. Abdelzaher, "Self- Contrastive Learning based Semi-Supervised Radio Modulation Classification," in *Proc. IEEE MILCOM*, 2021, pp. 777–782, **DOI**: 10.1109/MILCOM52596.2021.9652914.

[33] K. Davaslioglu, S. Boztaş, M. C. Ertem, Y. E. Sagduyu and E. Ayanoglu, "Self-Supervised RF Signal Representation Learning for NextG Signal Classification With Deep Learning," in *IEEE Wireless Communications Letters*, vol. 12, no. 1, pp. 65–69, Jan. 2023, **DOI**: 10.1109/LWC.2022.3217292.

[34] S.-H. Kim, J.-W. Kim, W.-P. Nwadiugwu and D.-S. Kim, "Deep Learning-Based Robust Automatic Modulation Classification for Cognitive Radio Networks," in *IEEE Access*, vol. 9, pp. 92 386–92 393, 2021, **DOI**: 10.1109/ACCESS.2021.3091421.

Yuewen Xu was born in 2001 in Zhuanglang County, Pingliang City, Gansu Province, China. He enrolled at Xi'an Jiaotong-Liverpool University in 2019 and obtained a bachelor's degree in Electrical Engineering and Automation in 2023. Currently, he is pursuing a master's degree in Communication Network and Signal Processing at the University of Bristol.

From June 2022 to December 2022, he participated in the development of an Artificial Intelligence-based Surgical Instrument Inventory System, where he was responsible for designing the user interface. Between September 2022 and May 2023, he completed a thesis on the design of electric propulsion systems for unmanned aerial vehicles. His research interests include Artificial Intelligence applications, Communication Networks, and Signal Processing. In his free time, he enjoys exploring technological advancements and cultural heritage, striving to apply his expertise to real-world challenges.

Blind Source Separation Spectrum Detection Method Based on Wavelet Transform and Singular Spectrum Analysis

Qian Hu, Zhongqiang Luo, and Wenshi Xiao

***Abstract*—To address the issue of reduced detection performance due to the impaired separation mechanism affected by noise, this paper proposes a blind source separation (BSS) detection method based on Wavelet Transform (WT) and Singular Spectrum Analysis (SSA). Firstly, the input signal is denoised using WT. Then, SSA is employed to denoise and reduce the dimension of the processed signal. Subsequently, the independent component analysis (ICA) based BSS algorithm is employed to separate the mixed signal preprocessed by the previous two ways. Finally, the proposed algorithm and the BSS detection method based on WT are compared in terms of spectrum analysis and separation performance. Simulation results show that the blind source separation detection method based on WT-SSA has a better signal detection performance.**

***Index Terms*—blind source separation; wavelet transform; singular spectrum sensing; independent component analysis**

I. INTRODUCTION

IN WIRELESS communication environments, the wireless channels are always complex, so that transmitting signals will be affected by channel fading and noise, resulting in mixed signals at the receiving end, and most detection algorithms such as traditional energy detection algorithms are unable to distinguish mixed signals to implement spectrum detection. Therefore, blind source separation is an often used promising method for solving this problem thanks to its advanced technical superiority [1][2][3]. Blind source separation refers to the analysis of the unobserved original signal from multiple observed mixed signals. Usually the observed mixed signals come from the outputs of multiple sensors and the output signals of the sensors are independent. Blind source separation has received a great deal of attention as a comprehensive research field that intersects with information theory, signal processing, artificial neural networks, probability theory, and other disciplines. In recent years, blind signal processing has become an important development direction in many research areas such as modern digital signal processing and computational intelligence, and it has great potential for applications in electronic information technology, communications, biomedicine, image enhancement, radar systems, geophysical signal processing and other fields. Blind signal processing relies on the statistical properties of the source signal, and its study has always been a focus of researchers, with blind signal separation being one of the most important research topics. The meaning of "blind" is twofold: on the one hand, the source signal is not known; on the other hand, the source signal mixing method is also not known.

One of the most famous blind source separation algorithms is Independent Component Analysis (ICA). And there have sprung up many algorithms to improve on ICA for theory updates. In [4], the authors designed a fast Power Iterative Independent Component Analysis (PowerICA) noise suppression scheme based on power iteration. First, a single-channel blind separation model is transformed into a multichannel observation model by constructing a pseudo-observation signal through a weighting process. This blind separation algorithm is then used to separate the noise from the source signal. Finally, the effectiveness of the algorithm is verified by experimental simulation. In [5], the authors treat the Lagrange multiplier as a constant in the original derivation of the FastICA algorithm and use a temporary approximation to the Jacobi matrix in the Newton-Raphson update, and then provide an alternative derivation of the FastICA algorithm that does not require an approximation. Based on this, the authors propose a new FastICA power iteration algorithm that is more stable than the fixed-point algorithm when the sample size is not several orders of magnitude larger than the dimensionality. In [6], when the ICA model does not forget to hold in full, the real data used for simulation experiments are not accurate in completing the maximization of the corresponding likelihood. To address this situation, the authors proposed a new algorithm called Picard, which uses only the sparse approximation Hessian as a preprocessor for the L-BFGS algorithm, refining the Hessian approximation from the memory of past iterations. The results demonstrate the superior performance of the proposed technique, especially on real data, by comparing a wide range of values for several algorithms of the same class through simulations.

The algorithms mentioned above in the literature are only a small part of the blind source separation algorithms, but all of them have proved the effectiveness of blind source separation. As the presence of factors such as noise in complex wireless communication environments can affect the effectiveness of signal separation, it is necessary to denoise and reduce the signal dimension prior to implementing blind source separation. In this paper, the Wavelet Transform (WT) is first used for denoising, and then the Singular Spectrum Analysis (SSA) is used for denoising and reducing the dimension. These two methods are applied in many aspects, for example, in [7], the authors investigated the threshold selection problem in the

Qian Hu, Zhongqiang Luo, and Wenshi Xiao are with School of Automation and Information Engineering, Sichuan University of Science and Engineering, Yibin, China; Zhongqiang Luo is also with Intelligent Perception and Control Key Laboratory of Sichuan Province, Sichuan University of Science and Engineering, Yibin, China.

The corresponding author is Zhongqiang Luo (e-mail: luozhongqiang@suse.edu.cn)

DOI: 10.36244/ICJ.2025.4.6

image denoise process based on WT, the wavelet coefficients are obtained by WT of the image signal, suitable threshold values are selected to process the wavelet coefficients, and then the processed wavelet coefficients are inverted by WT to obtain the reconstructed denoised image. Simulation experiments prove that the adaptive threshold denoising technique based on WT has the best denoising result, which can clearly retain the details in the image without sharpening and oversmoothing, and its values of Signal-to-Noise Ratio (SNR) and peak SNR are the largest and the value of mean square error is the smallest, thus improving the overall quality of the image. In [8], the authors used a noise reduction method combining wavelet thresholding and Singular Value Decomposition (SVD) processing, and conducted a study on noise suppression of one-dimensional audio signals received by intelligent navigation for automotive noise pollution, and proposed a joint WT-SVD model algorithm, and finally, through simulated experimental comparison and analysis, the algorithm has good noise suppression for the received signals of intelligent navigation systems under the environment of automotive noise interference adaptive and good suppression performance. In [9], the authors studied the video electromagnetic leakage problem, first of all, the intercepted video electromagnetic leakage signal through the SSA method for noise reduction of the noisy video electromagnetic leakage signal, through simulation experiments found that the singular spectrum analysis method applied to the video electromagnetic leakage signal denoising not only to remove the noise relatively clean, and can well retain the original video electromagnetic signal details and features, The SNR is improved significantly. The above-mentioned papers have demonstrated that WT and SSA have good performance in noise removal, so this paper starts to study this.

From the above-mentioned paper, it can be seen that WT in image processing denoising effect is better, and in the wireless communication environment, the transmission signal will be mixed with the white noise present in the channel, at this time the wavelet denoising effect is not very satisfactory, and the SSA can be used by decomposing the signal into different components, and then according to certain rules, select certain effective components and reconstruct the new signal, to achieve the signal denoising and reducing the signal dimension[10] [11][12]. Based on this, this paper proposes a blind spectrum detection method based on WT and SSA, firstly using WT to denoise the received mixed signal, then using SSA to denoise and reduce the dimension of the processed signal again, and finally using the ICA_p algorithm in blind source separation to separate the mixed signal. The experimental results corroborate the effectiveness of the proposed scheme.

II. SYSTEM MODEL

In wireless communication systems, the situation in the channel is complex and variable, and the observed signals received at the receiver may be affected by noise or mixed with other signals, and many algorithms are unable to distinguish between the mixing sequences, so blind source separation is investigated. Blind source separation is a technique for separating independent source signals from a set of sensor measurements using only the weakly known condition that the source signals are independent of each other, given that the transfer function of the system, the mixing coefficients of the source signals and their probability distribution are unknown.

Assuming that there are m signals that are independent of each other and the length of these signals is T, after channel gain of A and an additive Gaussian white noise channel, a mixed received signal is obtained , i.e. the observed signal

$$X = AS + N. \tag{1}$$

Where A is the $m \times m$ mixing matrix, S is the *m*-dimensional original signal vector, N is the m -dimensional noise vector and X is the *m*-dimensional mixed signal vector, the expressions are as follows

$$S = [s_1, s_2, ..., s_m]. \tag{2}$$

$$N = [n_1, n_2, ..., n_m]. \tag{3}$$

$$X = [x_1, x_2, ..., x_m]. \tag{4}$$

Where $s_i, i = 1, 2, ..., m$ is the signal transmitted by the primary user, $s_i, i = 1, 2, ..., m$ represents the noise generated by the channel and $x_i, i = 1, 2, ..., m$ represents the signal received by the sensor.

The observed mixed signal is separated using the ICA by blind source separation. The aim of blind source separation is to find a separation matrix W, which gives an estimate Y of the source signal S from the observed signal X.

$$Y = WX = WAS. \tag{5}$$

The above equation shows that if the separation matrix W is approximated by A^{-1}, (5) can be written as

$$Y = WAS = A^{-1}AS = IS = S. \tag{6}$$

At this point, Y is approximated by the source signal S.

W can be obtained from the objective function $F(W)$. When the relevant mathematical algorithm is used to make $F(W)$ reach an optimal solution, W is then the separation matrix, and depending on the different definitions of $F(W)$, and the method of finding W, different ICA algorithms can be obtained. The ICA_p algorithm used in this paper is an ICA algorithm using Hessian approximate precoding.

The objective function used in the model is the negative mean log likelihood of the parameterization of the separation matrix $W = A^{-1}$, as follows:

$$F(W) = -\frac{1}{T}\log p(X \mid W^{-1}). \tag{7}$$

Finally, the source signal is recovered by solving for the separation matrix W when the objective function $F(W)$ reaches optimality.

III. BLIND SOURCE SEPARATION DETECTION METHOD BASED ON WAVELET TRANSFORM AND SINGULAR SPECTRUM ANALYSIS

At low SNR, the blind source separation is unsatisfactory, so the observed signal is correlated before the signal is separated.

A. DENOISING OF MIXED SIGNALS USING WT

From a signal science perspective, wavelet denoising is a

signal filtering problem, and although to a large extent wavelet denoising can be seen as low-pass filtering, it is superior to traditional low-pass filters in this respect because it also successfully retains signal features after denoising. As can be seen, wavelet denoising is actually a combination of feature extraction and low-pass filtering functions, the flow block diagram of which is shown in Fig. 1.

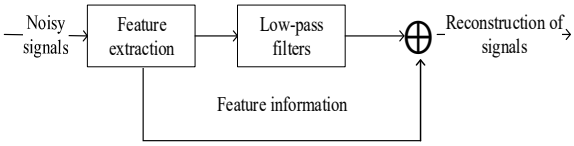


Fig.1. Block diagram of wavelet denoising

In the wireless communication environment, the white noise has the same effect on all wavelet coefficients due to the uneven spatial distribution and small wavelet coefficients, so the WT can be used to denoise the observed signal and thus improve the separation. The WT signal $X^{'}(t)$ is denoted as

$$X^{'}(t)=\left\langle X(t),\psi_{a,b}(t)\right\rangle=\int_{-\infty}^{\infty}X(t)\psi_{a,b}^{*}(t)dt=\int_{-\infty}^{\infty}\left(AS(t)+N(t)\right)\psi_{a,b}^{*}(t)dt \quad (8)$$

Where $\psi_{a,b}(t)$ is the wavelet basis and $\psi_{a,b}{}^{*}(t)$ is the conjugate of the wavelet basis. In this paper, the Daubechies wavelet is chosen as the wavelet basis because it is continuously orthogonal and it has the smallest branching.

B. THE WT SIGNAL $X^{'}(t)$ IS PROCESSED AGAIN USING SSA TO RECONSTRUCT THE SINAL

The main steps in SSA include embedding, decomposition, grouping, and reconstruction. In this paper, SSA is used to process signals because it includes singular value decomposition (SVD) as one of the steps in SSA, as SVD decomposes the signal, a process that involves reducing the dimensionality of the signal by retaining the significant components and removing the insignificant ones.

The first step is to form the trajectory matrix by lag-sorting the WT signal $X^{'}(t)$ through a suitable window length

$$X^{'}=\begin{bmatrix}x_1^{'} & x_2^{'} & \cdots & x_{T-L+1}^{'}\\ x_2^{'} & x_3^{'} & \cdots & x_{T-L+2}^{'}\\ \cdots & \cdots & \cdots & \cdots\\ x_L^{'} & x_{L+1}^{'} & \cdots & x_T^{'}\end{bmatrix}. \quad (9)$$

Where T is the sequence length and L is the window length, usually taken as $L<\frac{T}{2}$. Let $K=T-L+1$, then the trajectory matrix $X^{'}$ can be rewritten as a matrix of $L\times K$

$$X^{'}=\begin{bmatrix}x_1^{'} & x_2^{'} & \cdots & x_K^{'}\\ x_2^{'} & x_3^{'} & \cdots & x_{K+1}^{'}\\ \cdots & \cdots & \cdots & \cdots\\ x_L^{'} & x_{L+1}^{'} & \cdots & x_T^{'}\end{bmatrix}. \quad (10)$$

The trajectory matrix $X^{'}$ is then decomposed by SVD, and the resulting sequence is grouped and reconstructed.

The reconstruction requires the calculation of the projection of the hysteresis sequence $X_i^{'}$ onto U_m

$$a_i^m=X_i^{'}U_m=\sum_{j=1}^{L}x_{i+j}^{'}U_{m,j},0\le i\le T-L. \quad (11)$$

Where $X_i^{'}$ denotes the i th column of the trajectory matrix $X^{'}$, U_m is the eigenvector corresponding to the eigenvalue λ_m, and a_i^m is the weight of the time-evolving pattern reflected by $X_i^{'}$ at time $x_{i+1}^{'},x_{i+2}^{'},...,x_{i+L}^{'}$ of the original series, called the temporal principal component (TPC).

The signal is then reconstructed by means of a temporal empirical orthogonal function and temporal principal components, and the specific reconstruction process is as follows

$$x_i^{'k}=\begin{cases}\frac{1}{i}\sum_{j=1}^{i}a_{i-j}^{k}U_{k,j},1\le i\le L-1\\ \frac{1}{L}\sum_{j=1}^{L}a_{i-j}^{k}U_{k,j},L\le i\le T-L+1\\ \frac{1}{T-i+1}\sum_{j=i-T+L}^{L}a_{i-j}^{k}E_{k,j},T-L+2\le i\le T\end{cases} \quad (12)$$

The sum of all reconstructed sequences should be equal to the original sequence, i.e.

$$x_i^{'}=\sum_{k=1}^{L}x_i^{'k},i=1,2,...,T. \quad (13)$$

C. SUBSTITUTING THE RECONSTRUCTED SIGNAL $x_i^{'k}$ INTO THE SYSTEM MODEL

To bring $x_i^{'k}$ into the system model:

$$x_i^{'k}=AS+n. \quad (14)$$

Since the signals assumed in this paper are independent of each other, the probability of A

$$p(X_i^{'k}\mid A)=\prod_{t=1}^{T}\frac{1}{|\det(A)|}\prod_{i=1}^{m}p_i([A^{-1}x_i^{'k}k]_i(t)). \quad (15)$$

Where $p_i(\cdot)$ is the i th signal probability density function.

Bringing (15) into the objective function (7), we get

$$F(W)=-\log|\det(W)|-E[\sum_{i=1}^{T}\log(p_i(y_i(t)))]. \quad (16)$$

Among them $Y=WX$.

The next solution minimizes the objective function $F(W)$ with respect to W. This corresponds to solving the ICA problem in a maximum likelihood sense. The variation of $F(W)$ with respect to W can be expressed by the Tait expansion of $F((I+\varepsilon)W)$

$$F((I+\varepsilon)W)=F(W)+\langle G\mid\varepsilon\rangle+\frac{1}{2}\langle\varepsilon\mid H\mid\varepsilon\rangle+O(\|\varepsilon\|^3). \quad (17)$$

The first-order terms are controlled by matrices of order $T \times T$ and are called relative gradients. The second-order term depends on the tensor H of $T \times T \times T \times T$ and is called the relative Hessian matrix. Both of these quantities can be obtained from the second order expansions of $\log\det(\cdot)$ and $\log p_i(\cdot)$

$$\log|\det(I+\varepsilon)| = Tr(\varepsilon) - \frac{1}{2}Tr(\varepsilon^2) + O(\|\varepsilon\|^3). \quad (18)$$

$$\log p_i(y+e) = \log p_i(y) - \psi_i(y)e - \frac{1}{2}\psi_i'(y)e^2 + O(\|\varepsilon\|^3). \quad (19)$$

Where ε is a small matrix of $T \times T$ and e is a very small number. where $\psi_i = -\frac{p_i'}{p_i}$, in general, is $\tanh(\frac{\cdot}{2})$. The collection and rearrangement of a to produce the classical expression of first order

$$G_{ij} = E[\psi_i(y_i)y_i] - \delta_{ij}. \quad (20)$$

Or write it as another expression

$$G(Y) = \frac{1}{T}\psi(Y)Y^T - Id. \quad (21)$$

The second-order relative Hessian matrix can be written as

$$H_{ijkl} = \delta_{il}\delta_{jk}E[\psi_i(y_i)y_j] + \delta_{ik}E[\psi_i'(y_i)y_jy_l]. \quad (22)$$

The Hessian matrix approximation is discussed on the basis of the following moments

$$\begin{cases} \hat{h}_{ijl} = E[\psi_i'(y_i)y_jy_l] & 1 \le i,j,l \le N \\ \hat{h}_{ij} = E[\psi_i'(y_i)y_j^2] & 1 \le i,j \le N \\ \hat{h}_i = E[\psi_i'(y_i)] & 1 \le i \le N \\ \hat{\delta}_i^2 = E[y_i^2] & 1 \le i \le N \end{cases}. \quad (23)$$

Therefore, the relative Hessian matrix is

$$H_{ijkl} = \delta_{il}\delta_{jk}E[\psi_i(y_i)y_i] + \delta_{ik}\hat{h}_{ijl} \qquad i \ne j. \quad (24)$$

The first approximation of H lies in the substitution of $\hat{h}_{ijl}$ for $\delta_{jl}\hat{h}_{ij}$ and $\tilde{H}^2$ for this approximation

$$\tilde{H}^2_{ijkl} = \delta_{il}\delta_{jk}E[\psi_i(y_i)y_i] + \delta_{ik}\delta_{jl}\hat{h}_{ij} \qquad i \ne j. \quad (25)$$

The second approximation is represented by $\tilde{H}^1$, going one step further and replacing $\hat{h}_{ij}$ with $\hat{h}_i\delta_j^2$

$$\begin{cases} \tilde{H}^1_{ijkl} = \delta_{il}\delta_{jk}E[\psi_i(y_i)y_i] + \delta_{ik}\delta_{jl}\hat{h}_i\delta_j^2 & i \ne j \\ \tilde{H}^1_{ijkl} = 1 + \hat{h}_{ii} \end{cases}. \quad (26)$$

Finally, the approximate Hessian matrix is derived

$$H_{ijkl} = \delta_{il}\delta_{jk}E[\psi_i(y_i)y_j] + \delta_{ik}\delta_{jl}E[\psi_i'(y_i)], i \ne j. \quad (27)$$

Substituting (26) into (17) gives

$$\langle G|\varepsilon\rangle + \frac{1}{2}\langle\varepsilon|\tilde{H}|\varepsilon\rangle = \sum_{i<j}(G_{ij} - G_{ji})\varepsilon_{ij} + \frac{\hat{k}_i + \hat{k}_j}{2}\varepsilon_{ij}^2. \quad (28)$$

Where $\{\varepsilon_{ij}, 1 \le i \le j \le N\}$, $\hat{k}_i$ are defined as

$$\hat{k}_i = E[\psi_i(y_i)y_i] - E[\psi_i'(y_i)]. \quad (29)$$

When $\hat{k}_i + \hat{k}_j > 0$, so that $\varepsilon_{ij} = -(G_{ij} - G_{ji})/(\hat{k}_i + \hat{k}_j)$ minimizes (28), the resulting quasi-Newton step

$$W_{k+1} = e^D W_k$$
$$D_{ij} = -\frac{2}{\hat{k}_i + \hat{k}_j} \cdot \frac{G_{ij} - G_{ji}}{2}. \quad (30)$$

In summary, this paper uses WT and SSA to process the mixed signal, and then uses ICA_p to separate the signal after processing, the specific process is shown in Fig. 2, the specific steps are as follows.

1. Denoising of the observed signal using the WT.
2. Denoising and dimensionality reduction of the processed signal again using SSA.
3. Correlation of signal removal with whitening.
4. Separation of the signal after preprocessing using the ICA_p algorithm.

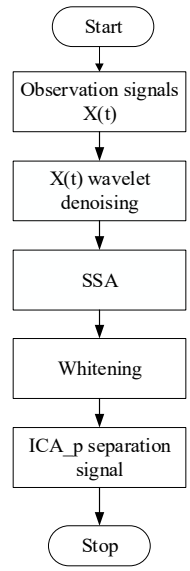


Fig. 2. Blind source separation process based on WT and SSA

IV. SIMULATION EXPERIMENTS AND ANALYSIS

To demonstrate more intuitively the process of blind source separation and spectrum analysis, the ICA_p algorithm is chosen to separate and reconstruct the signal at a SNR of 2 dB. Assuming the presence of 2 channels in the communication system and a sampling frequency f_s of 1KHz, the sampling period $t = 1/f_s$.

Fig. 3 and Fig. 4 show the time domain waveforms of the source signal, where the signal of channel 1 is $s1 = \sin(2\pi \times 20t)$ and the signal of channel 2 is $s2 = (1 + 0.5\sin(2\pi \times 5t)) \times \sin(2\pi \times 50t)$.

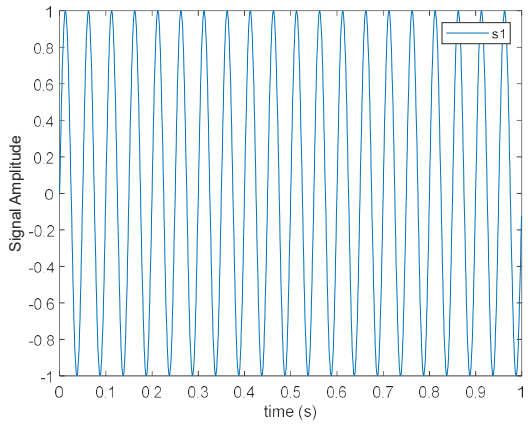


Fig. 3. Time domain waveform of source signal s1

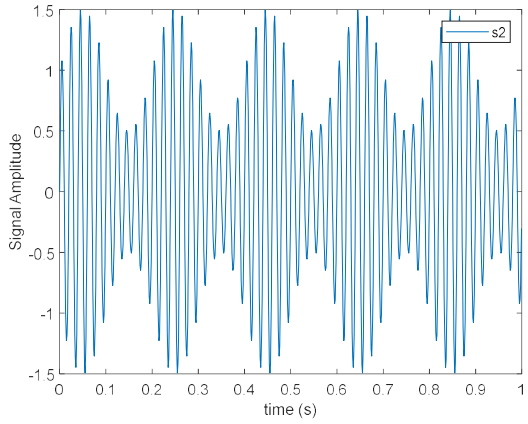


Fig. 4. Time domain waveform of source signal s2

Fig. 5 shows the spectrum corresponding to the source signal.

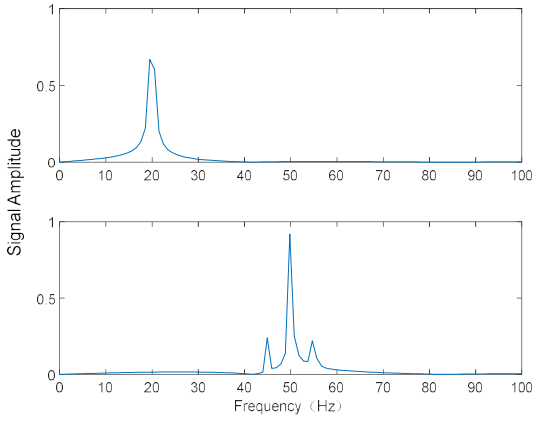


Fig.5. Source signal spectrum

In Fig. 5, it can be seen that the center frequencies of the two source signals are 20 Hz and 50 Hz, respectively.

The mixed signal obtained after the source signal has been passed through the mixing matrix is shown in Fig. 6 and Fig. 7.

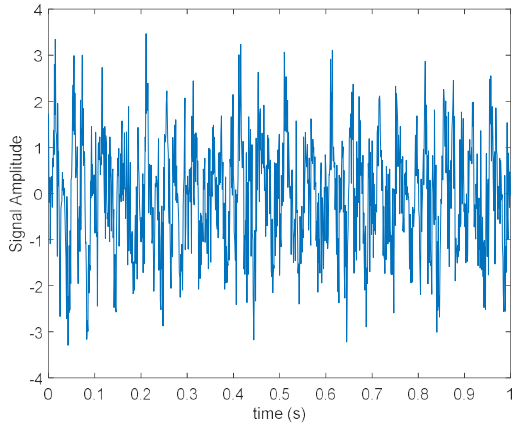


Fig. 6. Time domain waveform of mixed signal

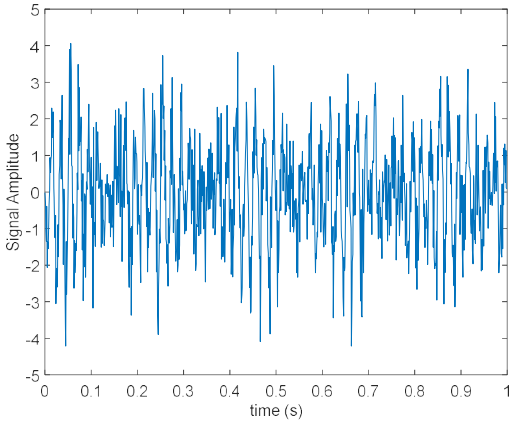


Fig. 7. Time domain waveform of mixed signal

A comparison of Fig. 3 and Fig. 4 with Fig. 6 and Fig. 7 shows that after the source signals have passed through the Gaussian channel, the waveform of the received observation signal is very different from the original signal, the waveform of the observation signal has changed due to the mixing of the signal during transmission.

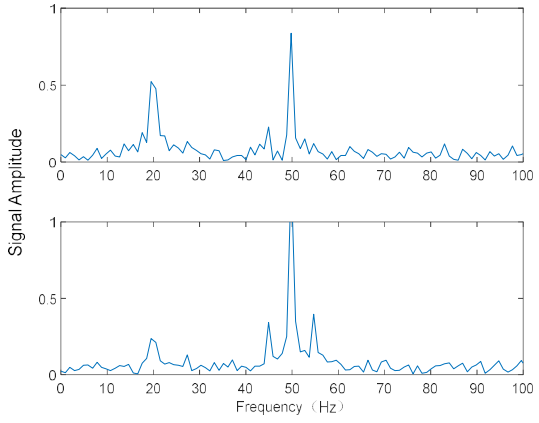


Fig. 8. Spectrogram of mixed signals

Fig. 8 shows the spectrum of the mixed signal. From Fig. 8, it can be seen that the center frequencies of signal 1 are 20 Hz and 50 Hz, and the center frequencies of signal 2 are 20 Hz and 50 Hz.

Compared with Fig. 5, it shows that the center frequency of signal 1 has not only the center frequency of the source signal of 20 Hz, but also an additional center frequency of 50 Hz, and signal 2 has a new center frequency of 20 Hz. From this result in the Fig. 8, it can be found that these source signals, after passing through the mixing matrix, produce a certain spectral overlap in the spectrum and thus it is difficult to distinguish multiple different signals through the spectrogram.

Fig. 9 and Fig. 10 show the waveforms of the signal separated from the mixed signal using the wavelet transform followed by the ICA_p algorithm.

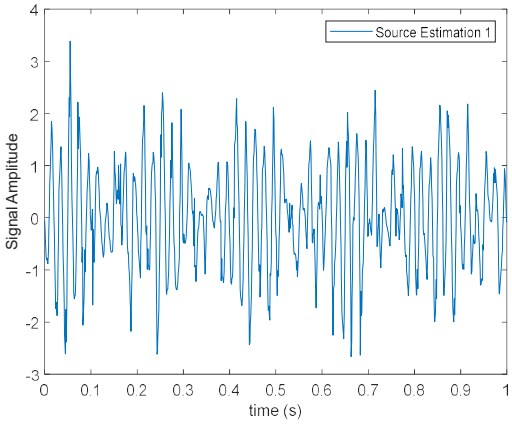


Fig. 9. WT-ICA_p separated signal waveform 1

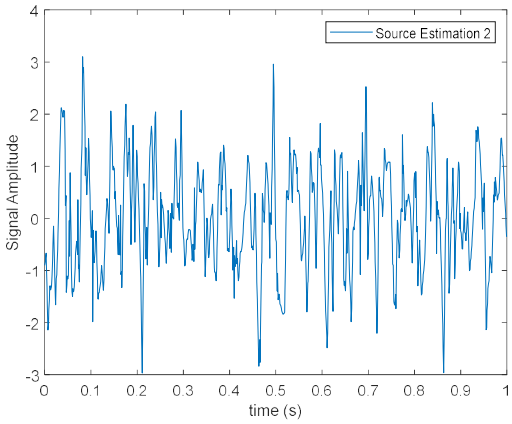


Fig. 10. WT-ICA_p separated signal waveform 2

Compared with Fig. 3 and Fig. 4, it can be found that although the WT-ICA_p algorithm is able to separate the signals, the waveforms obtained by the separation are still somewhat different from the source signals, especially the separated signal 2.

Fig. 11 shows the spectrum of the ICA_p separated signal after WT.

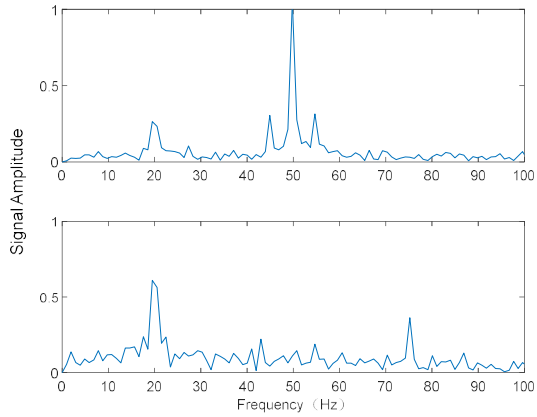


Fig. 11. Spectrum of WT-ICA_p separated signal

From Fig. 11, we can see that the center frequencies of signal 1 are 20 Hz and 50 Hz, and the center frequencies of signal 2 are 20 Hz and 75 Hz. Comparing with Fig. 5, we find that signal 1 has a new center frequency of 20 Hz and signal 2 has a new center frequency of 75 Hz, so the algorithm of separating the signals by ICA_p after wavelet transform processing is not very effective, and there is still some spectral It is also difficult to distinguish several different signals through the spectrogram.

Fig. 12 and Fig. 13 show the time domain plots after the signal has been first WT, then SSA, and finally separated using the ICA_p algorithm.

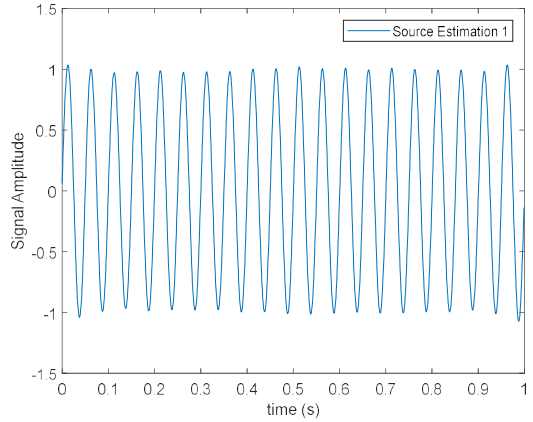


Fig. 12. WT-SSA-ICA_p separated signal waveform 1

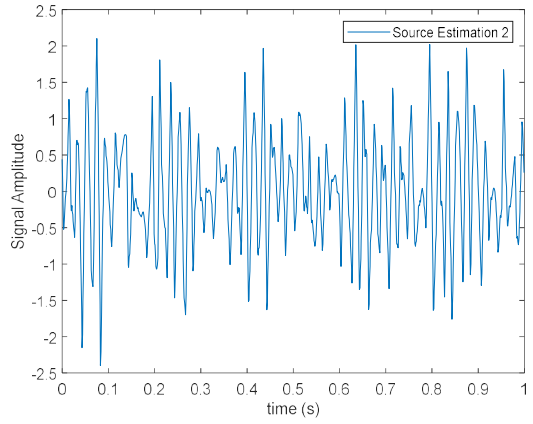


Fig. 13. WT-SSA-ICA_p separated signal waveform 2

Compared with Fig. 3 and Fig. 4, it can be found that the WT-SSA-ICA_p algorithm can separate the signals better, and the waveforms of the separated signals are basically the same as those of the source signals, indicating that the algorithm has better separation. As the signals are independent of each other, the WT is used to preprocess the signals before separation, which can effectively remove the influence of noise, and then the preprocessed signals are again processed with SSA to reduce the peace and remove noise, and finally the separation is carried out with the ICA_p algorithm, and the separation effect obtained is not disturbed by the spectral overlap, thus obtaining a waveform close to the original signal. The proposed algorithm can effectively separate these spectrally overlapped signals from their blind sources.

Fig. 14 shows the spectrum of the WT-SSA-ICA_p separated signal.

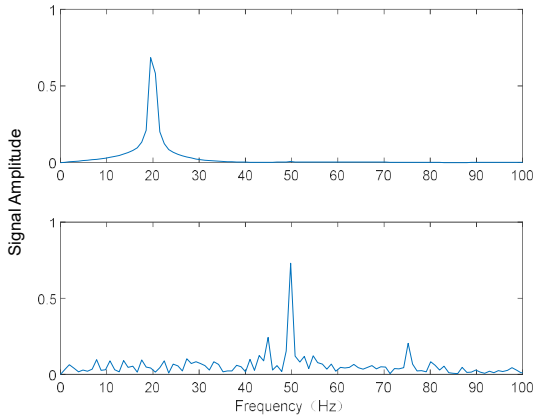


Fig. 14. Spectrum of WT-SSA-ICA_p separated signal

From Fig. 14, we can see that the center frequency of signal 1 is 20Hz, and the center frequencies of signal 2 are 50Hz and 75Hz. Comparing with Figure 4, we find that only signal 2 has a new center frequency of 75Hz, so the algorithm of separating the signals after WT processing and then using SSA and finally ICA_p works better, although only one signal 2 has spectral overlap on the spectrum, but signal 1 is very well separated and the signal can be completely distinguished from the spectrogram.

V. PERFORMANCE ANALYSIS

To evaluate the performance of the blind source separation algorithm, the correlation coefficient C is introduced as a performance indicator by comparing the similarity of the separated signal with the input signal.

$$C(x,y)=\frac{\operatorname{cov}(x,y)}{\sqrt{\operatorname{cov}(x,x)}\sqrt{\operatorname{cov}(y,y)}}. \tag{31}$$

Where $\operatorname{cov}(x,y)$ is the covariance of x and y, with the expression shown in (32).

$$\begin{aligned}\operatorname{cov}(x,y)&=E[x-E(x)]E[y-E(y)]\\&=E(xy)-2E(x)E(y)+E(x)E(y)\\&=E(xy)-E(x)E(y)\end{aligned}. \tag{32}$$

Where $E(\cdot)$ is the expected value and the correlation coefficient takes a range of $0\le C(x,y)\le 1$. When x and y are not correlated $C(x,y)$ takes 0. The closer $C(x,y)$ is to 1, the greater the correlation between x and y, and the more similar x and y are, the better the separation algorithm.

To analyze the performance of the algorithms, 3000 experiments were conducted using (31) to measure the similarity between the input source signals and the separated signals, and the average correlation coefficients of the two blind source separation algorithms were obtained, as shown in Fig. 15.

As can be seen from Fig. 15, although the correlation coefficient of the WT-ICA_p algorithm increases as the SNR increases, the correlation coefficient of the WT-SSA-ICA_p algorithm remains above 0.95 throughout and the value is very stable, which indicates that the WT-SSA-ICA_p algorithm can separate the mixed signal well and the separated signal is very similar to the source signal This shows that the WT-SSA-ICA_p algorithm can separate the mixed signal well and the separated signal is very similar to the source signal.

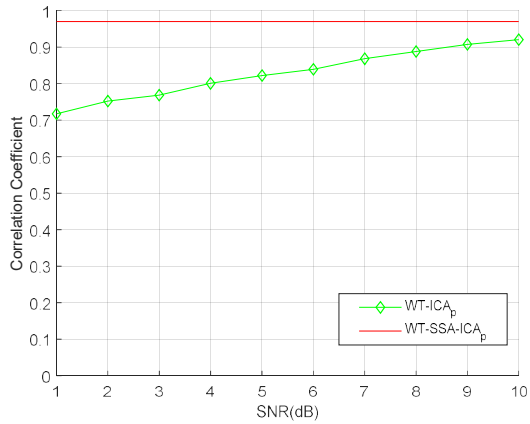


Fig.15. Correlation coefficient with SNR

It is important for communication systems to ensure the stable and reliable transmission of information. In this experimental simulation set when the correlation coefficient C is closer to 1, it is considered as a successful separation, when the value of C is too small or no solution, then the separation fails and the point is considered as a lost point, i.e. a point of nonconvergence. The probability of separation failing and the nonconvergence occurring when the two algorithms were statistically tested for 3000 trials is shown in Fig.16.

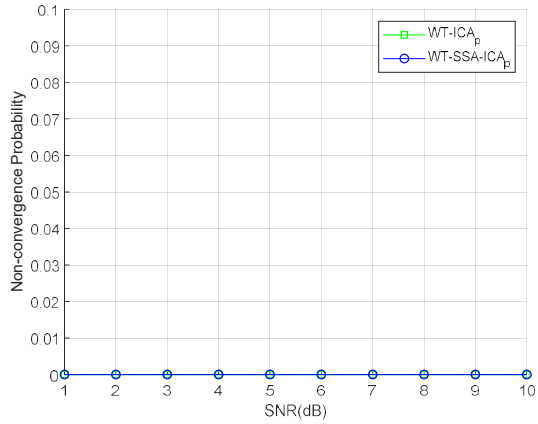


Fig.16. Variation of the probability of nonconvergence with SNR

As can be seen in Fig. 16, the results are stationary. Both

algorithms are uniform convergence regardless of the SNR change.

VI. CONCLUSIONS

In this paper, the WT-ICA_p algorithm is improved on the existing basis, i.e., the WT, by adding SSA after the WT and proposing the WT-SSA-ICA_p algorithm. Before the blind source separation, the mixed signal is first denoised with the WT, then the processed signal is denoised and reduced dimension with SSA, and finally the mixed signal is separated with the ICA_p algorithm. Then, the simulation experiments were carried out to analyze the spectrum of the signals. From the time and frequency domain plots obtained from the simulation, it can be seen that the WT-SSA-ICA_p algorithm has a better separation effect than the WT-ICA_p algorithm. Finally, the performance of the original WT-ICA_p algorithm and the WT-SSA-ICA_p algorithm is compared. The simulation results show that the two algorithms have better stability as the SNR changes, and the correlation coefficient of the WT-SSA-ICA_p algorithm is higher than that of the WT-ICA_p algorithm throughout, indicating that the WT-SSA-ICA_p algorithm can separate the mixed signals very well. The WT-SSA-ICA_p algorithm has better stability than the WT-ICA_p algorithm throughout. In the future work, the dynamic channel model and undertermined mixture model will be further considered for spectrum detection.

REFERENCES

[1] J. Ni, Z. Zhou, "Blind source separation and unmanned aerial vehicle classification using CNN with hybrid cross-channel and spatial attention module," *Scientific Reports*, vol. 15, 21905, 2025. DOI: 10.1038/s41598-025-07946-y

[2] Z. Luo, C. Li, L. Zhu, " A comprehensive survey on blind source separation for wireless adaptive processing: Principles, perspectives, challenges and new research directions," *IEEE Access*, vol. 6, pp. 66 685–66 708, 2018. DOI: 10.1109/ACCESS.2018.2879380

[3] Y. Xue, Y. Xu, A. Li, "Research on Acoustic Signal Denoising Algorithm Based on Blind Source Separation," *Procedia Computer Science*, vol. 243, pp. 380–387, 2024. DOI: 10.1016/j.procs.2024.09.047

[4] Y. Zhang, Z. Luo, "A review of research on spectrum sensing based on deep learning," *Electronics*, vol. 12, no. 21, 4514, 2023. DOI: 10.3390/electronics12214514

[5] W. Zhang, Z. Luo, X. Xiong, "Impulse noise suppression based on power iterative ICA in power line communication," *International Journal of Electronics and Telecommunications*, vol. 65, no. 4, pp. 651–656, 2019. DOI: 10.24425/ijet.2019.129824

[6] S. Basiri, E. Ollila and V. Koivunen, "Alternative Derivation of FastICA With Novel Power Iteration Algorithm," *IEEE Signal Processing Letters*, vol. 24, no. 9, pp. 1378–1382, 2017. DOI: 10.1109/LSP.2017.2732342.

[7] P. Ablin, J. F. Cardoso, A. Gramfort, "Faster independent component analysis by preconditioning with Hessian approximations," *IEEE Transactions on Signal Processing*, vol. 66, no. 16, pp. 4040–4049, 2018. DOI: 10.1109/TSP.2018.2844203

[8] G. Liu, Y. Cao, W. Feng, E. Zhao, C. Xing,"Research on adaptive threshold image denoising based on wavelet transform," *Journal of Anhui Electronic Information Vocational Technology College*, vol. 21, no. 01, pp.1–5, 2022.

[9] L. Zhao. Research on automotive noise suppression based on wavelet transform and SVD. Inner Mongolia University of Science and Technology, 2021. DOI: 10.27724/d.cnki.gnmgk.2021.000332.

[10] N. Kang, "Research on video electromagnetic leakage signal processing method based on singular spectrum analysis," *Taiyuan University of Science and Technology*, 2021. DOI: 10.27721/d.cnki.gyzjc.2021.000235.

[11] F. Xie, M. Li, J. Tian, X. Yue, Q. Yang, " A wavelet transform-based blind separation method for single-channel signals," *Modern Electronics Technology*, vol. 44, no. 07, pp. 56–59, 2021. DOI: 10.16652/j.issn.1004-373x.2021.07.011.

[12] C. Yi, W. Yuan, "Application analysis of time series models based on wavelet analysis and singular spectrum analysis," *Geomining and Mapping*, vol. 35, no. 04, pp. 6–9, 2010. DOI: 10.16864/j.cnki.dkch.2019.0050.

Qian Hu received the B S. degree in communication engineering from the Sichuan University of Science and Engineering (SUSE), Zigong, China, in 2020. She is currently pursuing the master's degree with the Sichuan University of Science and Engineering, Yibin. Her research interests include machine learning and cognitive radio.

Zhongqiang Luo received the B.S. and M.S. degrees in communication engineering and pattern recognition and intelligent systems from Sichuan University of Science and Engineering, Zigong, China, in 2009 and 2012, respectively. He received the Ph.D. degree in communication and information systems from University of Electronic Science and Technology of China (UESTC), in 2016. Since 2017, he has been with the Sichuan University of Science and Engineering, where he is currently a professor. From December 2018–December 2019, he was a visiting scholar with Department of Computer Science and Electrical Engineering of University of Maryland Baltimore County (UMBC). His research interests include spectrum sensing, blind source separation, signal processing for wireless communication system and intelligent signal processing.

Wenshi Xiao received a bachelor' s degree in elec- trical engineering and automation from the school of science and information technology of Hebei University of Engineering in 2021. Since 2021, she has been studying in Sichuan University of light and chemical technology, where she is a graduate student. Her research direction is the design of communication physical layer based on artificial intelligence.

Performance Evaluation of FBMC with optimal subcarrier spacing for 5G & beyond Communications

T. Padmavathi[1]*, Kusma Kumari Cheepurupalli[2], and R. Madhu[3]

***Abstract*—In 5G New Radio (NR) Release 15, the 3rd Generation Partnership Project (3GPP) Physical Layer Modulation for downlink and uplink communications, the Cyclic Prefix OFDM (CP-OFDM) is used. A wide range of potential use cases will describe future wireless networks. In order to achieve this, time-frequency resources must be dynamically assigned, which is difficult for traditional Orthogonal Frequency Division Multiplexing (OFDM). Therefore, OFDM improvements like filtering or windowing are needed. On the other hand, a multicarrier method like Filter Bank Multi-Carrier (FBMC) can be employed. Several prototype filters, including Hermite, PHYDYAS, and Root Raise Cosine (RRC), are used in this work to develop the framework for FBMC. Time-frequency efficiency will be determined for each user in the same band by adjusting the subcarrier spacing. The performance of the Signal to Interference Noise Ratio (SIR) is calculated for FBMC using varying subcarrier spacing and compared with different multicarrier transmission methods such as f-OFDM (filtered OFDM), CP-OFDM, UFMC (Universal Filter Multi Carrier), Weighted Overlap and Add (WOLA). FBMC outperforms CP-OFDM, UFMC, f-OFDM, and WOLA in terms of Signal-to-Interference-plus-Noise Ratio (SIR), especially when subcarrier spacing is short (15 kHz, 30 kHz), where spectral leakage is most noticeable. The PHYDYAS filter performed better than the other ones, reducing inter-carrier interference and increasing spectral efficiency by 20–30% even in asynchronous transmission scenarios. Furthermore, FBMC improved bandwidth economy by maintaining excellent performance without requiring a cyclic prefix. According to these findings, FBMC is a strong contender for upcoming 5G upgrades and 6G networks that require flexible waveform design, low out-of-band emissions, and support for a variety of service classes, such as mMTC, URLLC, and non- orthogonal transmissions.**

***Index Terms*—time-frequency resources, prototype filters, Filter Bank Multi-Carrier (FBMC), subcarrier spacing, Signal to Interference Noise Ratio (SIR).**

I. INTRODUCTION

Future wireless and cellular communication systems have many potential uses, including massive Machine Type Communications (mMTC), ultra-reliable low-latency communication (URLLC), and enhanced Mobile BroadBand (eMBB), particularly in areas like industrial automation and vehicle networks [1–3]. In order to meet these various and demanding needs, spectrum usage must be extremely flexible and efficient, which frequently calls for improvements in multicarrier modulation techniques.

Orthogonal Frequency Division Multiplexing (OFDM), which is widely used in 4G and 5G networks because its simplicity and robustness, provides effective spectrum to usage but has substantial out-of-band emissions and poor adaptability to changing spectral environments [4-6]. Various alternative systems have been proposed to overcome these restrictions, including filtered-OFDM (f-OFDM), windowed-OFDM (W-OFDM), universal filtered multicarrier (UFMC), and filter bank multicarrier (FBMC) [7–9]. Because of its superior spectral containment, removal of cyclic prefix, and enhanced robustness in asynchronous environments, FBMC with Offset Quadrature Amplitude Modulation (OQAM) has shown considerable promise among these [10–12]

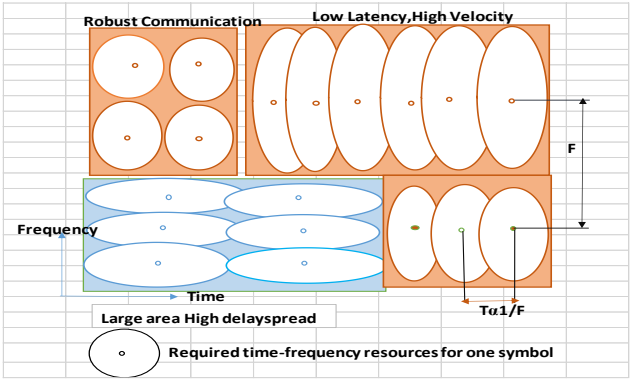


Figure.1. flexible assignment of future wireless system [11]

Several investigations have reviewed OFDM and FBMC in an array of channel conditions, with an emphasis on implementation complexity, inter-symbol interference (ISI), and spectral efficiency [13–17]. Additionally, previous research has investigated how prototype filters, like Root Raised Cosine (RRC), Hermite, and PHYDYAS, affect FBMC's performance [18–20]. The impact of subcarrier spacing optimization on the Signal-to-Interference Ratio (SIR) of various FBMC filter types, especially in scenarios that are pertinent to real-world implementation, like environments with high interference and different guard band sizes, is still largely unknown. The majority of current research ignores filter-specific performance under adaptive subcarrier configurations, which is crucial for dynamic spectrum access in applications like as eMBB, mMTC, and URLLC.

[1]* Department of Electronics and Communication Engineering, CVR College of Engineering, Ibrahimpatnam, Hyderabad, India, *Corresponding author: (e-mail: padmatp41@gmail.com)
[2] Department of Electronics and Communication Engineering at Gayatri Vidya Parishad College of Engineering (Autonomous), Visakhapatnam, Andhra Pradesh, India
[3] Department of Electronics and Communication Engineering, University College of Engineering Kakinada, Jawaharlal Nehru Technological University Kakinada (JNTUK), Kakinada, India

DOI: 10.36244/ICJ.2025.4.7

In order to bridge this gap, the best subcarrier spacing in FBMC systems using several prototype filters is evaluated in this paper. We compare performance to various multicarrier schemes including CP-OFDM, f-OFDM, W-OFDM, and UFMC and examine the effects of spacing on SIR. According to these results, FBMC is an interesting choice for flexible waveform designs in the future since it can achieve nearly optimal performance with simple one-tap equalizers when subcarrier spacing is carefully specified. The main objective of this research is to assess the effects of subcarrier spacing on the Signal-to-Interference Ratio (SIR) for various multicarrier waveforms, with a focus on FBMC under various prototype filters. The key objective is SIR-based interference analysis to guide waveform and filter selection in real-world deployment environments, even though associated issues like spectral efficiency and latency are essential.The rest of this paper is organized as follows: The primary multicarrier methods and prototype filters are introduced in Section 2, the simulation framework and SIR analysis are presented in Section 3, and the subcarrier optimization process and performance evaluation will be addressed in Section 4.

II. MULTICARRIER COMMUNICATIONS

The information is transmitted in the form of pulses for multicarrier system, where the pulses overlap in frequency and time. Due to the small bandwidth of the pulse, the frequency selective channels transform to many flat subchannels with less interference. An equalizer with just one tap is adequate, when the signal is transmitted in the presence of Gaussian noise with maximum likelihood symbol detection.

The time domain multi carrier system transmitted signal s(t) is given as

$$s(t) = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} g_{l,k}(t) x_{l,k} \quad (1)$$

Where $x_{l,k}$ is the transmitted symbol, l is subcarrier position, k represents time position. L denotes the number of subcarriers and K represents the number of multicarrier symbols while transmitting basis pulse. $g_{l,k}(t)$ is basis pulse which is expressed as

$$g_{l,k}(t) = p(t - kT) e^{j2\pi lF(t-kT)} e^{j\theta_{l,k}} \quad (2)$$

The basis pulse of prototype filter $p(t)$ is shifted with time spacing (T) and frequency spacing (F) and phase shift of $\theta_{l,k}$.

The signal is transmitted through Additive White Gaussian Noise (AWGN) Channel, then the signal can be retrieved from the received signal, which is denoted by the symbol r(t).

$$y_{l,k}(t) = < r(t), g_{l,k}(t) > = \int_{-\infty}^{\infty} r(t) g_{l,k}^{*}(t) dt \quad (3)$$

The similar basis pulse is used in the receiver to maximizes the Signal to Noise Ratio (SNR)

The Balian-Low theorem [20] states that the desired properties for the existence of multi carrier systems may not be fulfilled at the same time: i.e

- Orthogonality $< g_{l1,k1}(t), g_{l2,k2}(t) > = \delta_{(l_2-l_1)(k_2-k_1)}$
- Time-localization $\sigma_t < \propto$
- Frequency- localization $\sigma_f < \propto$
- Maximum symbol density $TF = 1$

δ denotes the Kronecker delta function. Time localization σ_t, and frequency localization σ_f are defined as

$$\sigma_t = \sqrt{\int_{-\alpha}^{\alpha} (t - \bar{t})^2 |p(t)|^2 dt}\, \sigma_f \quad (4)$$

$$= \sqrt{\int_{-\alpha}^{\alpha} (f - \bar{f})^2 |P(f)|^2 df} \quad (5)$$

The basis pulse $p(t)$ is normalized to get energy with mean time

$$\bar{t} = \int_{-\alpha}^{\alpha} t |p(t)|^2 dt \quad (6)$$

And mean frequency is given as

$$\bar{f} = \int_{-\alpha}^{\alpha} f |P(f)|^2 df \quad (7)$$

$|p(t)|^2$ and $|P(f)|^2$ represents Probability Density Function (PDF).

The Balian-Low theorem states that one of the desired characteristics can be neglected while implementing the multicarrier systems. Different multi carrier transmissions are compared and shown in Table.1. which indicates Filtered/windowed OFDM and CP-OFDM provide full bi-orthogonality with good time localization, while FBMC variants offer superior frequency localization. FBMC-QAM allows full complex-domain bi-orthogonality at the expense of pilot flexibility, while FBMC-OQAM is restricted to real-domain orthogonality.

CP-OFDM

The common multicarrier transmission is called CP-OFDM technique used in 4G, Wireless LAN, and LTE. In CP-OFDM the computational complexity can be reduced because of transmit and receive pulses.

The prototype filter of transmitter and receiver can be expressed as

$$p_{TX}(t) = \begin{cases} \frac{1}{\sqrt{T_O}} & ; \quad if \left(-\frac{T_O}{2} + T_{CP}\right) \le t \le \frac{T_O}{2} \\ 0 & ; \quad otherwise \end{cases} \quad (8)$$

$$p_{RX}(t) = \begin{cases} \frac{1}{\sqrt{T_O}} & ; \quad if -\frac{T_O}{2} \le t \le \frac{T_O}{2} \\ 0 & ; \quad otherwise \end{cases} \quad (9)$$

Localization: $\sigma_t = \frac{T_O + T_{CP}}{2\sqrt{3}}; \sigma_f = \propto$ (10)

Bi-Orthogonal: $T = T_O + T_{CP}; F = \frac{1}{T_O}$ (11)

T_O signifies the time scaling parameter determined by the required time spacing and subcarrier spacing.

Important time-domain pulse shapes and associated parameters are introduced by equations (8)– (11). Equations (8) and (9), which define the normalized transmit and receive pulse shapes over periods involving T_O (symbol time) and T_{CP} (cyclic prefix duration), respectively, are represented by the symbols $p_{TX}(t)$ and $p_{RX}(t)$. Equation (10) specifies the temporal localization σ_t, while σ_f indicates the frequency localization, which for rectangular pulses is unlimited (∞). The time-frequency spacing in bi-orthogonal systems is given by equation (11) where the subcarrier spacing is $F = 1/T_o$and the symbol duration is $T = T_O + T_{CP}$.

Because of the poor frequency domain localization of rectangular pulses, there is a lot of Out of Band (OOB) emission in CP-OFDM. Additionally, the CP reduces spectral efficiency but makes receiver equalization for frequency-selective

channels simpler. 3GPP considered windowing [21] and filtering [22], [23] to reduce OOB emission in OFDM with Weighted Over Lap and Add (WOLA) i.e windowed OFDM scheme[24-25].

A smoother function (windowing) replaces the borders of the rectangular pulse at the transmitter, and adjacent WOLA symbols overlap in time. Windowing is also used by receivers to prevent inter band interference, however overlapping and adding operations are performed within the WOLA symbol. F. Schaich.et.al [26] proposed two methods to implement f-OFDM (filtered-OFDM). Initially, Universal Filtered Multi-Carrier (UFMC) employs a Dolph-Chebyshev window-based subband-wise filtering technique. The filter design includes 12 subcarriers per subband, having a time-frequency spacing that is orthogonal of $TF = 1.07$ which is same as in LTE. By choosing $TF = 1.14$, orthogonality is ensured for a time-frequency spacing. By decreasing the spacing of time-frequency to $TF = 1.09$, with slight self-interference (≈65 dB) spectral efficiency can be enhanced. The f-OFDM scheme is the 3GPP's second filter-based OFDM technique. A Hann window is multiplied by a sinc pulse (perfect rectangular filter) for generating f-OFDM. The same time-frequency spacing as UFMC but with longer filter lengths allow f-OFDM to experience the same self-interference(65 dB).

Additionally, windowing and filtering could reduce significant OOB emissions from CP-OFDM at the expense of reduced spectral efficiency. Moreover, FBMC continues to provide lower OOB emissions than filtering and windowing.

FBMC-OQAM

The desired properties of Balian-Low theorem, can be satisfied by replacing the strict the complex orthogonality condition $g_{l1,k1}(t), g_{l2,k2}(t) = \delta_{(l_2-l_1)(k_2-k_1)}$ with less real strict condition $\mathbb{R}\{g_{l1,k1}(t), g_{l2,k2}(t)\} = \delta_{(l_2-l_1)(k_2-k_1)}$. FBMC-OQAM can be implemented with the following filters as given in next sub section.

TABLE I
COMPARISON OF MULTI CARRIER TRANSMISSION OVER AWGN CHANNEL

S.No.	Multi carrier transmission	Maximum pilot density	Time localization	Frequency localization	Bi-Orthogonal
1	CP-OFDM	Yes	Yes	No	Yes
2	filtered/Windowed OFDM	No	Yes	Yes	Yes
3	FBMC-OQAM	Yes	Yes	Yes	Real only
4	FBMC-QAM	No	Yes	Yes	Yes

Hermite polynomials

Hermite filter can be designed by

1. Design a prototype filter based on Hermite polynomials Hn(·)[27]

$$p(t) = \frac{1}{\sqrt{T_O}} e^{-2\pi\left(\frac{1}{T_O}\right)^2} \sum_{i=\{0,4,8,12,16,20\}} a_i H_i\left(2\sqrt{\pi}\frac{t}{T_O}\right) \quad (12)$$

For which the co-efficients can be

$$a_o = 1.412692577 \qquad a_{12} = -2.2611X10^{-9} \quad (13)$$

$$a_4 = -3.0145\,X10^{-3} \qquad a_{16} = -4.4570X10^{-15} \quad (14)$$

$$a_4 = -8.8041X10^{-6} \qquad a_{16} = 1.8633X10^{-16} \quad (15)$$

With $p(t) = p(-t)$

2. Real orthogonal can be obtained by taking time-frequency spacing factor of two i.e , $T = T0/2$ and $F = 1/T0$.
3. The phase shift of the induced imaginary interference is shifted with

$$\theta_{l,k} = \frac{\pi}{2}(l+k) \quad (16)$$

Orthogonal : $T = T0;\ F = 2/T0 \rightarrow TF = 2$ (17)

Localization : $\sigma_t = 0.2015\,T_O;\ \sigma_f = 0.403\,T_O^{-1}$ (18)

PHYDYAS prototype filter

The PHYDYAS prototype filter is another important filter [28], designed by a basis function

$$p(t) = \begin{cases} 1 + 2\sum_{i=1}^{O-1} b_i cos\left(\frac{2\pi t}{OT_0}\right); & if\ \frac{-OT_O}{2} < T \le \frac{OT_O}{2} \\ 0 \qquad\qquad\qquad\quad ; & otherwise \end{cases} \quad (18)$$

The calculation of co-efficient b_i [19] depends on the overlapping factor O.

Let $O = 4$ then the co-efficients b_i is expressed as

$$b_1 = 0.97195983 \qquad b_2 = \frac{\sqrt{2}}{2} \qquad b_3 = 0.23514695 \quad (19)$$

Orthogonal: $T = T_O;\ F = 2/T_O \rightarrow TF = 2$ (20)

Localization: $\sigma_T = 0.2745\,T_O;\ \sigma_f = 0.328\,T_O^{-1}$ (21)

The PHYDYAS filter provides stronger frequency-localization but not as good as time-localization than the Hermite prototype filter.

RRC Filter

The frequency domain representation of the signal transmitted by the RRC filter is provided by [29]

$$P(f) = G(f) = \sqrt{|X_{rc}(f)|}e^{-j2\pi f T_O} \quad (22)$$

With $G(f) = P^*(f)$ where T_O is a delay.

And in the time domain $p(t)$ is given by

$$p(t) = \frac{1}{\sqrt{T_O}} \frac{sin\frac{(\pi t(1-\beta))}{T_O} + \frac{4\beta t}{T_O} cos\frac{(\pi t(1+\beta))}{T_O}}{\frac{\pi t}{T_O}\left(1-\left(\frac{4\beta t}{T_O}\right)^2\right)} \quad (23)$$

Where β is the Roll-off factor

Orthogonal: $T = T_O;\ F = 2/T_O \rightarrow TF = 2$ (24)

Localization: $\sigma_T = 0.2745\,T_O;\ \sigma_f = 0.328\,T_O^{-1}$ (25)

III. SIGNAL TO INTERFERENCE RATIO COMPUTATION

The SIR of FBMC-OQAM can be computed from the receiving symbol at l^{th} subcarrier position, k^{th} time position, and set noise to zero, then the received signal is expressed as [30]

$$y_{l,k} = g_{l,k}^H\, H\, G\, x = \left((G\,x)^T \otimes g_{l,k}^H\right) vec\{H\} \quad (26)$$

By combining all basis pulse vectors to form a transmit matrix $G \in \mathbb{C}^{N\times LK}$, and transmit symbol vector $x \in \mathbb{C}^{LK\times 1}$ is represented as

$$G = [g_{0,0} \cdots g_{L-1,0} \;\; g_{0,1} \cdots g_{L-1,K-1}], \quad (27)$$

$$x = \left[x_{0,0} \cdots x_{L-1,0} \;\; x_{0,1} \cdots x_{L-1,K-1}\right]^T \quad (28)$$

The impulse response of channel matrix is $H \in \mathbb{C}^{N\times N}$ and the operator $vec\{.\}$ is used to simplify the operation. To determine SIR in the FBMC-OQAM case

Let $\Gamma \in \mathbb{C}^{LKXLK}$ and it is given by

$$\Gamma = \left(G^T \otimes g_{l,k}^H\right) R_{vec\{H\}} \left(G^T \otimes g_{l,k}^H\right)^H, \quad (29)$$

Where Γ is decomposed with the following equation

$$\Gamma = \Omega\Omega^H \quad (30)$$

Where Ω is auxiliary matrix. $\Omega \in \mathbb{C}^{LKXLK}$ Then compute

$$\left[\check{\Omega}_i\right]_{u,v} = [\Omega]_{u,v} \frac{|[\Omega]_{i,v}|}{[\Omega]_{i,v}},$$

$$\check{\Gamma}_i = \Re\{\Omega_i\}\Re\{\Omega_i\}^H$$

Finally, SIR can be computed as

$$SNR_i^{OQAM} = \frac{[\check{\Gamma}_i]_{i,i}}{tr\{\check{\Gamma}_i\} - [\check{\Gamma}_i]_{i,i}} \quad (31)$$

IV. Effect of Optimal subcarrier spacing while computing SIR

Within the same band, FBMC may efficiently support various subcarrier spacing. Subcarrier spacing needed to be designed so that [31-32]

$$\frac{\sigma_t}{\sigma_f} = \frac{T_{rms}}{V_{rms}} \quad (32)$$

Where time-localization σ_t and frequency localization σ_f

Two use cases are considered in the proposed work i.e., user1 with subcarrier spacing $F_1 = 15kHz$ and user2 with subcarrier spacing $F_2 = 120\ kHz$. These two different subcarrier spacings are included to transmit the signal for various channel conditions. Low latency transmissions are improved by the larger subcarrier spacing. Low subcarrier spacing improves bandwidth efficiency while also making the system more robust to delays.

V. Results

A baseband multicarrier system model was used in MATLAB to carry out the simulations. PHYDYAS, Hermite, and RRC filters were used to analyze each waveform, including CP-OFDM, WOLA, UFMC, f-OFDM, and FBMC, under various subcarrier spacings (15 kHz, 120 kHz, and 480 kHz) and guard band ratios. SIR was calculated by averaging 1000 Monte Carlo runs in an AWGN channel and calculating the ratio of the intended signal power to adjacent subcarrier interference. A 96-tap length and an overlapping factor of 4 were used for implementing FBMC filters. To ensure equity, the same modulation (QAM), symbol durations, and total bandwidth were employed for all waveforms.

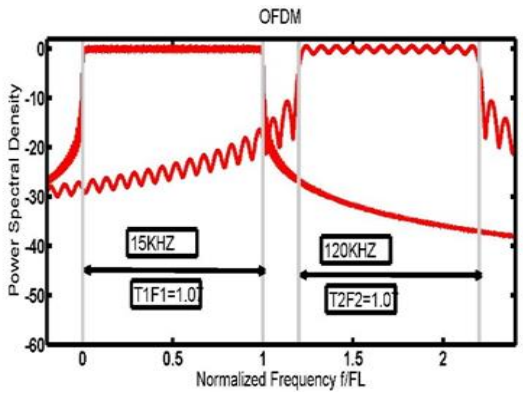
The transmitted signal from user1 is characterized by G1 with Number of subcarriers $L1 = 96$ with $F_1 = 15kHz$ subcarrier spacing resulting transmission bandwidth of $F_1L_1 = 1.44MHz$. Similarly, the transmitted signal from user2 is characterized by G2 with Number of subcarriers $L2 = 12$ with $F_2 = 120kHz$ subcarrier spacing resulting transmission bandwidth of $F_2L_2 = 1.44MHz$. Furthermore, User G2 frequency is shifted by $F_1L_1 + F_G$. A guard band of $F_G = 0.2F_1L_1$ is used between user1 and user2. A time-frequency spacing is used in WOLA, f-OFDM, and UFMC to lower the OOB, $T_1F_1 = 1.09$ for user1, $T_2F_2 = 1.27$ for user2 are assumed. FBMC with Hermite, OFDM, UFMC, WOLA, f-OFDM, PHYDYAS, and RRC with two users are calculated and presented in the Figure.2. (a) (b), (c), (d), (e), and (f)

The overall spectral performance of several multicarrier waveforms at different subcarrier spacings is shown in Figures 2.a through 2.f, highlighting their use for future wireless systems. While UFMC (Figure 2.b.) exhibits some improvement through sub-band filtering, OFDM (Figure 2.a.) suffers from excessive out-of-band emissions (OOBE), which get larger with increasing subcarrier spacing. However, residual side lobes still exist, particularly at higher spacings. WOLA (Figure 2.c.) still shows leakage under large spacing, although time-domain windowing helps with smoother roll-off. Though f-OFDM (Figure 2.d.) is sensitive to spectral overlap, it provides moderate suppression by subband filtering. However, at both 15 kHz and 480 kHz spacings, FBMC (Figures 2.e and 2.f) exhibits excellent spectrum confinement with negligible side lobes, showing its efficiency in minimizing leakage and maintaining orthogonality for high-bandwidth, interference-sensitive systems.

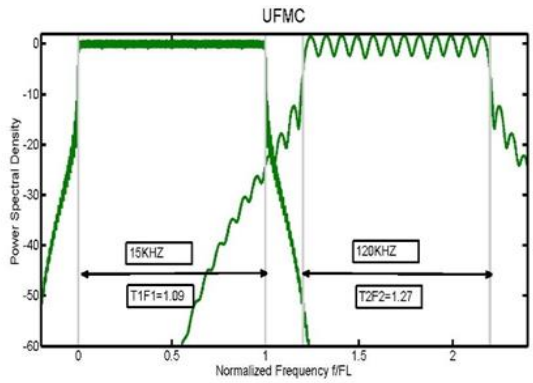
For 2-user case SIR can be given as

$$SIR_{total-2\ user} = \frac{L_1K_1 + L_2K_2}{||\Re\{G_1^HG_2\}||_F^2 + ||\Re\{G_2^HG_1\}||_F^2}$$

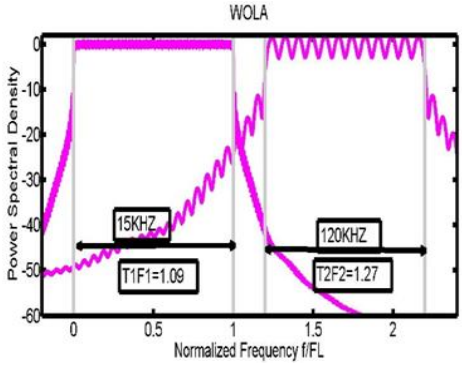
Where $||.||_F$ operation represents Frobenius norm.The $\Re$ in the above equation disappears for CP-OFDM, WOLA, UFMC and f-OFDM because it operates in complex domain. More interference may be experienced by nearby subcarriers than by distant ones. Figure.2(a) to 2(d) shows that if guard band is more then there is a less interference.



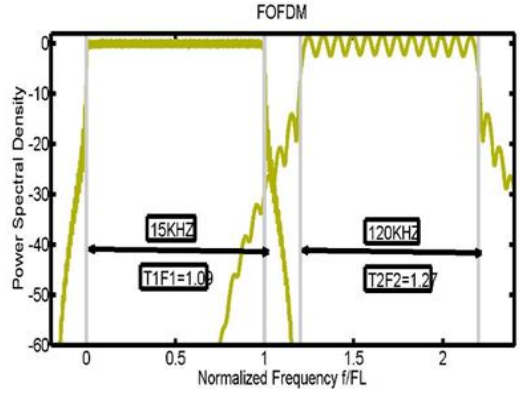
(a) OFDM



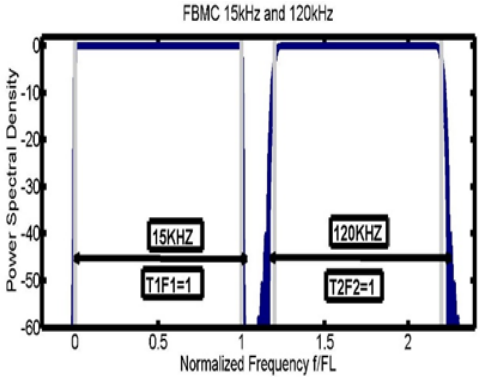
(b)UFMC



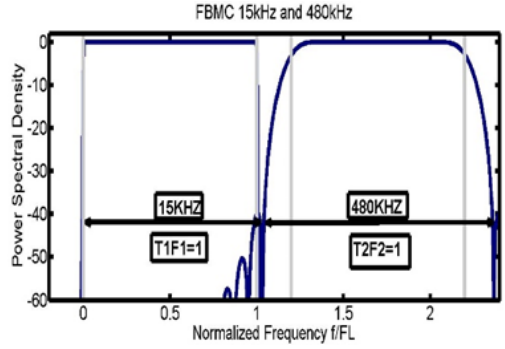
(c)WOLA



(d)f-OFDM



(e) FBMC with 15KHz and 120 KHz



(f) FBMC with 15KHz and 480 KHz

Figure.2. PSD of multicarrier transmission

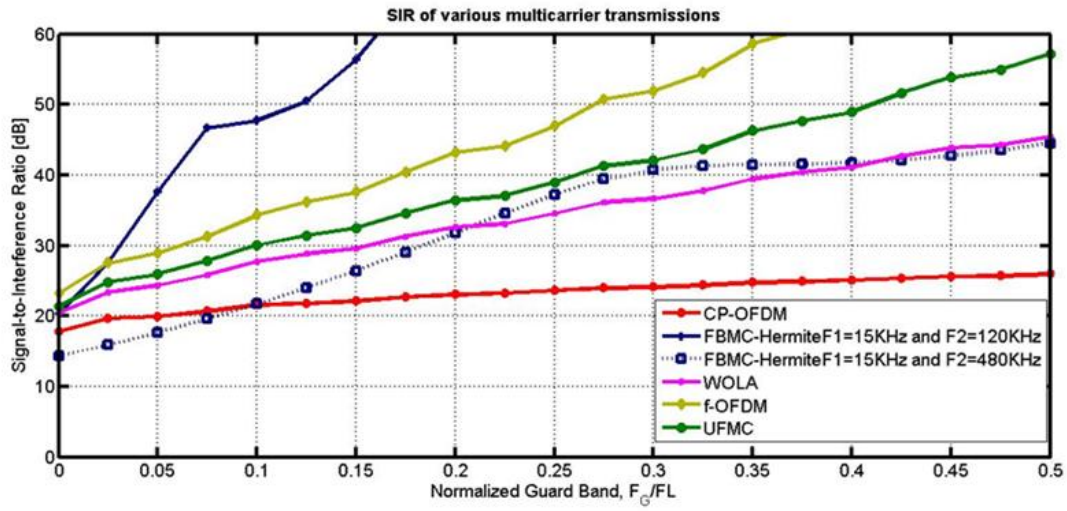


Figure.3. SIR characteristics of Multi carrier transmission techniques with FBMC-Hermite

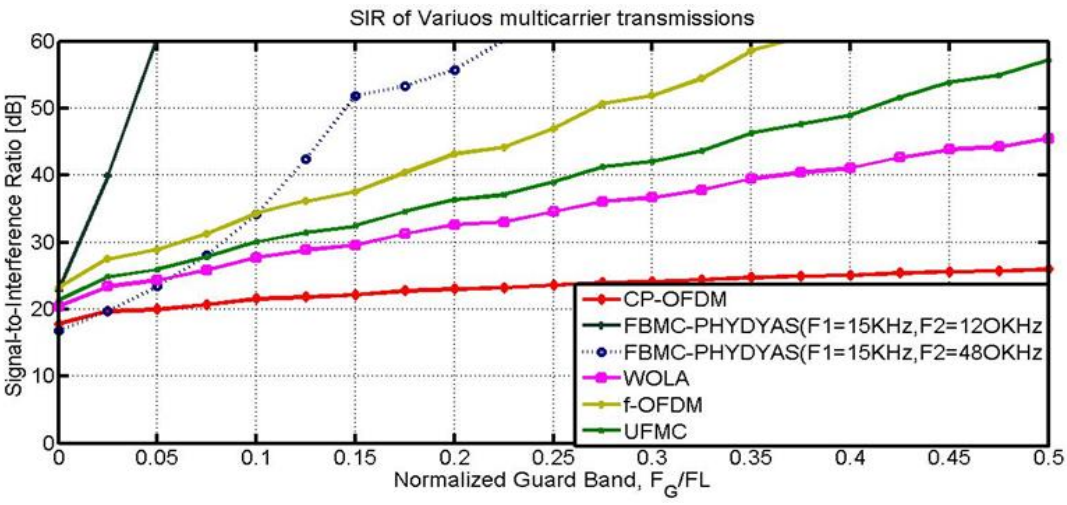


Figure.4. SIR characteristics of Multi carrier transmission techniques with FBMC-PHYDYAS

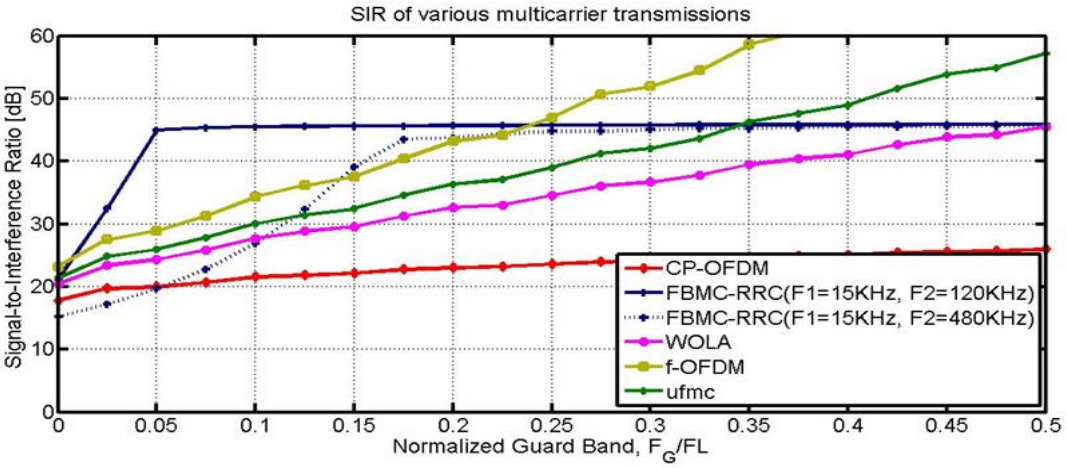


Figure.5. SIR characteristics of Multi carrier transmission techniques with FBMC-RRC

TABLE II
SIR FOR DIFFERENT F_G/F_L RATIOS

S.No	SIR (dB)	CP-OFDM	WOLA	UFMC	f-OFDM	FBMC-Hermite	FBMC-PHYDYAS	FBMC-RRC
1	$F_G/F_L = 0.05$	20	24	27	29	38	60	45
2	$F_G/F_L = 0.1$	22	28	30	34	48	60	45
3	$F_G/F_L = 0.17$	22	30	32	38	60	60	45

After receiving the signal SIR is calculated for various multicarrier transmission techniques with proposed optimal SIR is computed for FBMC with proposed 2 usecases with different prototype filters. FBMC usecase1 with $F1 = 15\ KHZ$ and $F2 = 120\ KHZ$, FBMC usecase2 with $F1 = 15\ KHZ$ and $F2 = 480\ KHZ$. For these usecases Hermite, PHYDYAS, RRC prototype filters are incorporated in FBMC, and computed SIR is shown in Figure 3 to 5. and it reveals that FBMC works significantly better than other systems in every scenario, with PHYDYAS and Hermite filters reaching up to 60 dB SIR at broader guard bands. While WOLA performs marginally better than CP-OFDM, which consistently yields the lowest SIR across all scenarios, UFMC and f-OFDM exhibit moderate gains.

FBMC - PHYDYAS with $F1 = 15KHZ$ and $F2 = 120KHZ$ achieves $60dB$ of SIR with least subcarrier spacing of $\frac{F_G}{F_L} = 0.05$. FBMC - PHYDYAS with $F1 = 15KHZ$ and $F2 = 480\ KHZ$ achieves $60dB$ of SIR with least subcarrier spacing of $\frac{F_G}{F_L} = 0.23$

Consider the required SIR is 45 dB, in case of f-OFDM, WOLA, UFMC the guard band of $F_G = 0.24F_L$ is required. Then for user2 time-frequency efficiency can be evaluated by using following equation

$$\rho = \frac{KL}{(KT+T_G)(F_L+F_G)} \tag{29}$$

The time-frequency efficiency ρ helps to choose modulation format, such that the available time-frequency resources can be maximized.

For user2 time-frequency efficiency can be computed using Equation .28, $\rho = \frac{1}{1.24X1.27} = 0.64$. Whereas in FBMC, ρ is 0.97 which is morethan the f-OFDM. This shows that user2 ($F2 = 120\ KHz$)i.e large subcarrier spacing empowers low latency transmission. If the subcarrier spacing increases by four-fold, the FBMC experiences the same delay as OFDM. If there are multiple subcarriers $L = 12$ in FBMC is decreases to L=3, requires a large guard band ($F_G = 0.13F_L\ for\ 45\ dB\ SNR$). However, time frequency efficiency of ρ=0.8 is still about 40% greater than the f-OFDM. By increasing the subcarrier spacing FBMC is still suitable for minimal latency transmission. This also increases the sensitivity, delay spread and time-offsets.

The Signal-to-Interference Ratio (SIR) performance of several multicarrier waveforms under various guard-to-lobe frequency ratios ($\frac{F_G}{F_L}$) is shown in the Table.2. It demonstrates that FBMC with PHYDYAS filtering continuously attains the highest SIR in every case, highlighting its better interference suppression, whereas WOLA and CP-OFDM perform comparatively more severe particularly as spectral congestion increases.

FBMC variations, especially those that use PHYDYAS and Hermite filters, consistently produce higher SIR values across all subcarrier spacings, as illustrated in Figures 3-5. Remarkably, FBMC-PHYDYAS achieves 60 dB SIR at a guard ratio of 0.17, but WOLA and CP-OFDM peak below 30 dB, indicating that they have limited capacity to reduce interference. While FBMC-PHYDYAS achieves 48 dB, showing improved spectrum isolation, the current implementation shows 30 dB, compared to [18], where UFMC obtained 35 dB SIR at a 0.1 guard ratio. Unlike [20], which use fixed subcarrier spacing, the results presented here show that FBMC SIR increases from 45 dB to 60 dB when spacing is increased from 15 kHz to 480 kHz.

The findings of evaluating SIR across multicarrier systems demonstrate that FBMC performs noticeably better than OFDM-based techniques, particularly at higher subcarrier spacings, demonstrating its applicability in situations where interference is a problem. PHYDYAS constantly offers the highest SIR for filter type evaluation, indicating its exceptional spectral confinement.

In order to provide dense spectrum access, FBMC optimizes guard ratios while maintaining >45 dB SIR even at narrow bands.The performance boost made here in comparison to [18] and [20] confirms the goal of proving filter-spaced optimization in FBMC systems.

While utilizing the proposed optimal spacing, FBMC outperforms f-OFDM by about 15% when high bandwidth is allocated per user (10.08 MHz as opposed to 1.44 MHz). Comparative SIR analysis of several multicarrier systems, including CP-OFDM, UFMC, WOLA, f-OFDM, and FBMC variations, that extend a broad range of subcarrier spacings and guard ratios is our main contribution. The results show that FBMC reaches up to 60 dB SIR, significantly outperforming traditional OFDM-based systems, particularly when used with PHYDYAS and Hermite filters. This confirms the significance of FBMC for applications with interference and limited spectrum, achieving the goal of the research, which was to find reliable waveform candidates for next-generation communication systems.

Conclusions

When the subcarriers are more, OFDM-based multi carrier transmissions like, f-OFDM, UFMC, and WOLA have a great spectral efficiency. However, not all future potential wireless transmission will require such a large number of subcarriers. FBMC is substantially more effective than OFDM for a small number of subcarriers, especially if many use cases share the transmission band. One-tap equalizers can be appropriate when the spacing of subcarrier (pulse shape) matches the channel statistics. The improvement in Signal-to-Interference Ratio (SIR) and signal integrity outweighs any potential minor trade-off in spectral efficiency. Furthermore, by increasing subcarrier spacing, FBMC can facilitate low-latency transmission while maintaining compatibility with one-tap equalizers, provided that the pulse shaping is in line with the channel statistics. A robust contender for upcoming wireless systems like 5G and beyond, the SIR analysis across many schemes validates FBMC's superiority in eliminating interference. These findings highlight how adaptable waveform strategies, such as FBMC, with filter customization, can handle a variety of deployment scenarios, especially in high-interference, dense environments. This investigation can be expanded in future research by assessing FBMC performance in actual wireless applications, including as fading and mobility conditions. Additional flexibility could be achieved by investigating AI-based subcarrier optimization, adaptive filtering, and integration with MIMO systems. These directions are especially relevant to 6G networks, which require strong interference control and spectral flexibility. This will assist in directing future research and implementation toward physical layer designs that are more adaptive and interference aware.

REFERENCES

[1] Dao, Nhu-Ngoc, *et al*. "A review on new technologies in 3GPP standards for 5G access and beyond." *Computer Networks* (2024): 110370. **DOI**: 10.1016/j.comnet.2024.110370

[2] Schwarz, Stefan, Tal Philosof, and Markus Rupp. "Signal processing challenges in cellular-assisted vehicular communications: Efforts and developments within 3GPP LTE and beyond." *IEEE Signal Processing Magazine* 34.2 (2017): 47–59. **DOI**: 10.1109/MSP.2016.2637938

[3] Degambur, Lavanya-Nehan, *et al*. "Resource allocation in 4G and 5G networks: A review." *International Journal of Communication Networks and Information Security* 13.3 (2021): 401–408.

[4] Bölcskei, Helmut. "Orthogonal frequency division multiplexing based on offset QAM." Advances in Gabor analysis. Boston, MA: Birkhäuser Boston, 2003. 321–352. **DOI**: 10.1007/978-1-4612-0133-5_12

[5] Shrivastava, Sandeep, Alok Jain, and Ram Kumar Soni. "Design of optimal FBMC subcarrier index modulation based on optimal prototype filter." *Analog Integrated Circuits and Signal Processing* (2022): 1–11. **DOI**: 10.1007/s10470-021-01933-9

[6] Choi, Kwonhue. "FD-FBMC: A Solution for Multicarrier Full Duplex Cellular Systems." *IEEE Communications Letters* 25.2 (2020): 617–621. **DOI**: 10.1109/LCOMM.2020.3028375

[7] Khudhair, Saif A., and Mandeep Jit Singh. "Review in FBMC to Enhance the Performance of 5G Networks." *J. Commun*. 15.5 (2020): 415–426. **DOI**: 10.12720/jcm.15.5.415-426

[8] F. Schaich and T. Wild, "Waveform contenders for 5G: OFDM vs. FBMC vs. UFMC," in *Proceedings of the 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, pp. 457–460, Athens, Greece, May 2014. **DOI**: 10.1109/ISCCSP.2014.6877912

[9] Arjun, R., *et al*. "Low complexity FBMC for wireless MIMO systems." *Physical Communication* 47 (2021): 101332. **DOI**: 10.1016/j.phycom.2021.101332

[10] Conceição, Filipe, Marco Gomes, Vitor Silva, Rui Dinis, Adão Silva, and Daniel Castanheira. "A survey of candidate waveforms for beyond 5G systems." *Electronics* 10, no. 1 (2020): 21. **DOI**: 10.3390/electronics10010021

[11] R. Gerzaguet, N. Bartzoudis, L. G. Baltar *et al*., "5G candidate waveform race: a comparison of complexity and performance," *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, p. 13, 2017. **DOI**: 10.1186/s13638-016-0792-0

[12] "TSG RAN; study on new radio access technology; physical layer aspects; (release 14)," http://www.3gpp.org/DynaReport/38802 htm, Mar. 2017

[13] J. Vihriälä, A. A. Zaidi, V. Venkatasubramanian *et al*., "Numerology and frame structure for 5G Radio access," in *Proceedings of the IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–5, Valencia, Spain, September 2016. **DOI**: 10.1109/PIMRC.2016.7794610

[14] Raja, S. KanagaSuba, A. Bamila Virgin Louis, and G. Arul Dalton. "Optimal Resource Allocation Scheme in Wireless 5G Networks." *TEST Engineering & Management* 83 (2020): 18 529–18 535.

[15] A. Yazar and H. s. Arslan, "Flexible multi-numerology systemsfor 5G New Radio," *Journal of Mobile Multimedia*, vol. 14, no. 4, pp. 367–394, 2018. **DOI**: 10.13052/jmm1550-4646.1442

[16] M. Bellanger, D. Le Ruyet, D. Roviras, M. Terré, J. Nossek, L. Baltar, Q. Bai, D. Waldhauser, M. Renfors, T. Ihalainen *et al*., "FBMC physical layer: a primer," *PHYDYAS*, January, 2010.

[17] Subalatha, M., *et al*. "Low complexity maximum likelihood FBMC QAM for improved performance in longer delay channels." *Wireless Personal Communications* 117 (2021): 3051–3066. **DOI**: 10.1007/s11277-020-07427-7

[18] Bellanger, M. (2012). Specification and design of a prototype filter for filter bank based multicarrier transmission. In *PHYDYAS Project Report*. **DOI**: 10.1109/ICASSP.2001.940488

[19] Siohan, P., Siclet, C., & Lacaille, N. (2002). Analysis and design of OFDM/OQAM systems based on filterbank theory. *IEEE Transactions on Signal Processing*, 50(5), 1170–1183. **DOI**: 10.1109/78.995073

[20] Bogucka, H., & Kryszkiewicz, P. (2015). Dynamic spectrum access: From cognitive radio to network radio. Wiley.

[21] A. Sahin, I. Guvenc, and H. Arslan, "A survey on multicarrier communications: Prototype filters, lattice structures, and implementation aspects," *IEEE Communications Surveys Tutorials*, vol. 16, no. 3, pp. 1312–1338, December 2012. **DOI**: 10.1109/SURV.2013.121213.00263

[22] M. Fuhrwerk, J. Peissig, and M. Schellmann, "Channel adaptive pulseshaping for OQAM-OFDM systems," in IEEE European Signal Processing Conference (EUSIPCO), 2014, pp. 181–185.

[23] H. G. Feichtinger and T. Strohmer, Gabor analysis and algorithms: Theory and applications. Springer Science & Business Media, 2012. 24. Qualcomm Incorporated, "Waveform candidates," in 3GPP TSG-RAN WG1 84b, Busan, Korea, April 2016. **DOI**: 10.1109/ISETC.2018.8583966

[25] X. Zhang, M. Jia, L. Chen, J. Ma, and J. Qiu, "Filtered-OFDM- enabler for flexible waveform in the 5th generation cellular networks," in *IEEE Global Communications Conference (GLOBECOM)*, 2015, pp. 1–6. **DOI**: 10.1109/GLOCOM.2015.7417854

[26] H. Nam, M. Choi, S. Han, C. Kim, S. Choi, and D. Hong, "A new filter-bank multicarrier system with two prototype filters for QAM symbols transmission and reception," *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 5998–6009, 2016. **DOI**: 10.1109/TWC.2016.2575839

[27] M. Bellanger, D. Le Ruyet, D. Roviras, M. Terré, J. Nossek, L. Baltar, Q. Bai, D. Waldhauser, M. Renfors, T. Ihalainen *et al*., "FBMC physical layer: a primer," *PHYDYAS*, January, 2010.

[28] Dumari, Hise Teferi, *et al*. "BER and PSD improvement of FBMC with higher order QAM using Hermite filter for 5G wireless communication and beyond." *Journal of Electrical and Computer Engineering* 2023.1 (2023): 7232488. **DOI**: 10.1155/2023/7232488

[29] Padmavathi, T., Kusma Kumari Cheepurupalli, and R. Madhu. "Evaluation of FBMC Channel Estimation using multiple Auxiliary symbols for high throughput and low BER 5G and beyond communications." *Infocommunications Journal* 16.2 (2024): 25–32. DOI: 10.36244/ICJ.2024.2.4

[30] Abdel-Atty, Heba M., Walid A. Raslan, and Abeer T. Khalil. "Evaluation and analysis of FBMC/OQAM systems based on pulse shaping filters." *IEEE Access* 8 (2020): 55 750–55 772. DOI: 10.1109/ACCESS.2020.2981744

[31] Khrouf, Wafa, Mohamed Siala, and Fatma Abdelkefi. "How much FBMC/OQAM is better than FBMC/QAM? A tentative response using the POPS paradigm." *Wireless Communications and Mobile Computing* 2018 (2018). DOI: 10.1155/2018/4637181

[32] Han, Hyungsik, Girim Kwon, and Hyuncheol Park. "MMSE-interference canceling receiver for QAM-FBMC systems." *IEEE Communications Letters* 24.11, 2589–2593 (2020): DOI: 10.1109/LCOMM.2020.3007895

T Padmavathi has done her Ph.D in the field of Wireless Communications & Signal processing, from JNTU, Kakinada in 2024. She is currently working as Associate Professor with the Department of Electronics and Communication Engineering, CVR College of Engineering, Ibrahimpatnam, Hyderabad, India. She has published more than 12 research papers in various reputed national and international Journals and conferences. Her research interests include Wireless communication, signal processing and Deep learning in channel estimation.

Kusma Kumari Cheepurupalli working as Associate Professor in the Department of Electronics and Communication Engineering at Gayatri Vidya Parishad College of Engineering (Autonomous), Visakhapatnam, Andhra Pradesh, India. She has graduated with PhD in Electronics and Communication Engineering from Andhra University in 2014. Her Research activity focuses on the development and applications of VLSI System Design, Machine Learning, Communications and Embedded systems. She has co-authored 26 Journal publications and 11 International and National conferences. Received Four best paper awards for the Conferences publications. She published 3 patents granted in UK Agency and IP Australian. She served as Principal Investigator for the Project work "Robust Signal Processing Techniques for RADAR/SONAR Communications", sanctioned by Department of Science and Technology (DST), New Delhi under Women Scientist Scheme (WOS-A), with a sanctioned amount of Rs 20 lakhs.

Ramarakula Madhu received the Bachelor of Engineering degree in electronics and communication engineering from Osmania University, Hyderabad, India, in 2003, the Master of Technology degree in ECE with the specialization of Communication Systems from Jawaharlal Nehru Technological University Hyderabad, India, in 2009, and the Ph.D. degree in Electronics and Communication Engineering from Andhra University, Visakhapatnam, India, in 2014. He is currently an Assistant Professor with the Department of Electronics and Communication Engineering, University College of Engineering Kakinada, Jawaharlal Nehru Technological University Kakinada (JNTUK), Kakinada, India. He has twelve years of teaching experience. He has published more than 45 research papers in various reputed national and international journals and conferences. His research interests include Mobile and Cellular Communications, Antennas, Satellite Communications, and GNSS. He is a Member of IEEE.

Dynamic XTEA Optimization and Secure Key Management for Embedded Microcontroller-Based SDN Systems in Smart Cities

Sunil Kumar Shah[1]*, Raghvendra Sharma[2], and Neeraj Shukla[3]

Abstract—**The rapid expansion of smart city infrastructures necessitates robust and efficient security mechanisms for embedded processor-based Software-Defined Networking (SDN) nodes. Hence, this research introduces the Adaptive Secure XTEA for Embedded Microcontrollers (ASX-EM), a novel encryption method designed to address these environments' unique security and performance needs. Existing encryption implementations often neglect proper padding validation, leading to vulnerabilities such as the Padding Oracle Attack (POA). The proposed Context-Aware Key Expansion and Secure Padding Validation (CAKE-SPV) technique customizes key scheduling based on node-specific parameters and employs a robust padding verification mechanism, significantly enhancing encryption security. Moreover, the computational demands of XTEA, with its multiple rounds of operations, are inefficient on resource-constrained 8-bit microcontrollers, leading to increased latency and reduced system responsiveness. To optimize performance, the Adaptive Round and Parallel Processing (ARPP) method is developed that dynamically adjusts encryption rounds based on system metrics and employs bit-slice processing with precomputed lookup tables for efficient arithmetic operations. The results show that the proposed model has a low encryption time of 2.7s a decryption time of 2.6s, and a high encryption throughput of 22,000 KB/sec and, a decryption throughput of 17600 KB/sec, compared to other existing models.**

Index Terms—**Smart cities, Padding Oracle Attack, Software-Defined Networking, Extended Tiny Encryption Algorithm, precomputed lookup tables.**

I. INTRODUCTION

The infrastructure of smart cities is being revolutionized by Software-Defined Networking (SDN), which provides a programmable and centralized method of controlling intricate and diverse metropolitan networks. SDN offers the required agility and control in the context of a smart city, where the integration of several IoT devices, traffic management systems, energy grids, and public safety networks produces a highly dynamic environment. SDN design allows for centralized control via a software-based controller by severing the network control plane from the data plane. Because it is in charge of the whole network, this controller enables dynamic resource allocation, automatic traffic management, and real-time monitoring. Because SDN is programmable, network policies may be quickly altered in response to shifting circumstances or new threats, guaranteeing ongoing security and optimization. To improve network flow and lower latency for vital applications like emergency response systems, SDN, for example, dynamically redirects data from bottleneck locations in traffic management. By modifying the distribution in response to real-time supply and demand data, SDN in energy management can help make it easier to integrate renewable energy sources. Furthermore, by offering a centralized platform for executing security procedures and identifying irregularities throughout the network, SDN's centralized architecture improves cybersecurity [1-4].

Enormous networks of interlinked sensors, devices, and communication systems underpin smart cities, and secure data transfer is necessary to preserve confidentiality, integrity, and trust. The XTEA (Extended Tiny Encryption Algorithm) is a durable and lightweight symmetric key block cipher that is well-suited for use in smart city applications. Because of its small size and ability to function on 64-bit blocks with a 128-bit key, XTEA is perfect for resource-constrained contexts, such as embedded systems and Internet of Things devices, which are commonly found in smart city infrastructure. Strong cryptographic security is provided by its straightforward structure, which consists of a sequence of bitwise shifts, XOR operations, and modular additions, this offers little computing cost. Because of its iterative method, which typically consists of 64 rounds, XTEA is more resistant to cryptanalysis attempts, which makes it a dependable option for protecting sensitive data such as traffic data, utility use statistics, and personal information. Although XTEA has many benefits, there are security risks and vulnerabilities associated with its usage in smart city infrastructure that need to be properly addressed [5-8].

One major issue is that XTEA is subject to differential cryptanalysis, especially if it is not implemented correctly and with enough rounds (64 rounds is the ideal amount, but implementations with fewer rounds are still vulnerable). Additionally, if weak keys are utilized, XTEA's key scheduling technique is extremely basic and vulnerable to cryptanalytic assaults. Physical assaults, like side-channel attacks, present a significant concern in the context of smart cities since gadgets frequently function in unsupervised and sometimes unsafe settings. To determine the encryption key, these attacks make use of data that is disclosed during the encryption process, such as power usage or electromagnetic emissions. Another risk

[1]* Amity University Madhya Pradesh, Maharajpura (Opposite Airport) Gwalior, Madhya Pradesh, India (e-mail: sunil.ggits@gmail.com)

[2] Amity University Madhya Pradesh, Maharajpura (Opposite Airport) Gwalior, Madhya Pradesh, India (e-mail: rsharma3@gwa.amity.edu)

[3] Gyan Ganga College of Technology, Jabalpur, Madhya Pradesh, India (e-mail: neerajshukla28@gmail.com)

DOI: 10.36244/ICJ.2025.4.8

is replay attacks, in which a hacker intercepts encrypted data and sends it again to trick the target system into carrying out commands or actions that are not authorized. Robust key management procedures, such as regular key rotation and the use of powerful, randomly generated keys, must be put in place to reduce these risks [9–12]. Smart cities may use XTEA's advantages while reducing possible weaknesses by tackling these risks with all-encompassing security solutions.

The XTEA implementation in embedded microcontrollers inside an SDN framework for smart city infrastructure has several technical difficulties that need to be carefully considered. One major problem is that embedded microcontrollers have restricted resources by nature; they include memory, computational power, and energy availability. Despite the lightweight nature of XTEA, cryptographic operations can still put a load on these few resources, which might affect the responsiveness and performance of vital smart city applications. In an SDN environment, where real-time data flow management and fast reconfiguration are critical, this is especially pertinent. The computational expense of XTEA causes latency problems, impacting the speed at which data is sent and decisions are made in the SDN, particularly when the necessary 64 rounds for security are implemented. The integration of XTEA inside the SDN control plane and data plane separation paradigm is another major hurdle. Embedding XTEA encryption techniques that dynamically adapt to the network's changing topology and traffic patterns is necessary to provide smooth and secure communication across various planes [13–15]. It is imperative to optimize the implementation of XTEA for low-power operations to overcome these technological obstacles. This can be achieved, for example, by using software libraries that are specifically designed for microcontroller architectures or by using hardware acceleration. Ultimately, an integrated approach that achieves a balance between security, performance, and resource limitations while guaranteeing the flexibility and scalability of the whole infrastructure is needed for the effective implementation of XTEA in embedded microcontrollers inside an SDN framework for smart cities.

A. Main objective of this paper

The following methodological and experimental contributions have been achieved by this paper:

- To mitigate POA vulnerabilities, CAKE-SPV is implemented within the XTEA encryption, in which Context-Aware Key Expansion Scheduling algorithm is utilized for customizing the key scheduling process, and OAEP with HIPV mechanism is utilized for secure padding with integrated validation, thereby ensuring secure communication and preventing information leakage through POA vulnerabilities.
- To optimize the XTEA for 8-bit Microcontrollers-based smart city infrastructure, an ARPP is presented, which utilizes adaptive round adjustment for an optimal number of encryption rounds and parallel processing for dividing the data into smaller bit-level blocks thereby enhancing the efficiency of XTEA encryption methods to operate effectively on 8-bit microcontrollers, reducing chip area, power consumption, and processing time.

B. Organization of study

The arrangement of the paper is as follows. In Section 2, relevant literature is reviewed; in Section 3, the methodology of the proposed system is explained; in Section 4, experiments, datasets, comparison, and evaluation methodologies are covered; in Section 5, suggestions for future developments and limits of the approach are made.

II. LITERATURE SURVEY

For the protection of sensitive data in a variety of applications, including RFID systems and smart cities, the security and effectiveness of encryption methods are essential. Enhancing the XTEA algorithm and its variations to handle certain security flaws and performance limitations in various scenarios has been the subject of recent research.

Ahmed et al [16] employed an enhanced S-box to boost security and thwart a variety of assaults, resulting in a new and reliable version of the original XXTEA. To achieve the one-time pad idea and provide an extra degree of protection, the M-XXTEA was also combined with a chaotic key-generating system. In contrast to the original XXTEA and AES, the cipher keys were dynamically updated for every block of plaintext throughout the encryption process, offering a more reliable security method. The M-XXTEA works with multiple text block sizes and key sizes in addition to improving data security. To compare the M-XXTEA's performance with that of the original XXTEA and AES, many experiments were carried out. The results showed that M-XXTEA surpassed AES by 60% in terms of encryption and decryption time efficiencies. The addition of new elements, including the chaotic key generation, results in unanticipated weaknesses, even if the M-XXTEA already counters several assaults.

Manikandan et al [17] addressed the XTEA's security issues by using domain-specific customization, random number generation, and hidden key renewal processes. RXMAP-1 and RXMAP-2, two different encoder architectures for the Renovated XTEA Mutual Authentication Protocol (RXMAP), were proposed. Their foundation was the replacement of accurate computational blocks with approximations. The proposed RXMAP protocol's computational and storage overhead was evaluated, and it was tested against a variety of security threats using BAN logic in both formal and informal verification. The proposed encoder designs are simulated for functional verification, and ASIC implementation is carried out on a 132 nm manufacturing node. However, because of the customization, use of random numbers, and key renewal procedures, the suggested protocol resulted in computational and storage overhead.

Zeesha Mishra and Bibhudendra Acharya [18] constructed optimal lightweight ciphers to implement the cipher in hardware by modeling the design characteristics. To accomplish the intended result, the TEA, XTEA, and XXTEA ciphers were developed, put into practice, and optimized utilizing specialist hardware platforms including Application

Specific Integrated Circuit (ASIC) and Field Programmable Gate Array (FPGA). Through the execution of designs for four hardware architectures TEA (T1), XTEA (T2), XXTEA (T3), and hybrid model (T4) many elements, including block sizes, implementation rounds, and crucial scheduling components, have been explored. The percentage gains in frequency for T1, T2, and T3 using a pipelined method are 75.9%, 162%, and 89%, respectively. Nevertheless, when optimizations are carried out, their scope and effects on other aspects like as security or resource use are not thoroughly investigated.

Neha Khute et al [19] proposed a round-based XXTEA-192-bit architecture to reduce the implemented hardware's space. This design had cheap cost and small space required, and it was meant for RFID applications. Simple shifting, addition, and XOR operations are among the fundamental and logical operations used by XXTEA. These simple activities allowed the architecture to be low-area and extremely efficient by design. Performance analysis was carried out on several FPGA device families, including Spartan-3, Virtex-7, Virtex-5, and Virtex-4, assessing variables including throughput and efficacy. Nevertheless, further optimization is needed in terms of speed, power usage, or reduced area.

Dzaky Zakiyal et al [20] developed a distributed MQTT (message queuing telemetry transmission) brokers-optimized architecture. For edge resources, a distributed MQTT broker might reduce latency and network traffic by managing only topics that were consumed on the network. An integer non-linear code was created to optimize container placement and minimize the wastage of edge computing resources. This architecture with the existing distributed MQTT middleware design with random and greedy container placement was simulated through rigorous modeling. When it came to lowering synchronization overhead, power use, network utilization, and deployment failure rates, this design fared better than the others. Nevertheless, the limited memory, processing power, and storage of edge devices affect the solution's viability and efficiency.

Keshari et al [21] suggested using the Grey Wolf Optimization Affinity Propagation (GWOAP) algorithm to arrange many controllers in smart city networks. The network's linked smart devices' traffic was controlled by the controllers. The OS3E network architecture is used to mimic the suggested method. To minimize processing delays and regulate the controller's traffic load, the controller deploys in the OS3E network topology by executing AP and GWO optimization algorithms that split the network into subdomains. IoT-enabled smart switches are better distributed throughout clusters using GWOAP, and node equalization was distributed evenly among all controllers in the deployed architecture. The traffic load of IoT-enabled devices in smart city networks is intelligently balanced across controllers by employing the suggested technique.

Anusha, and Shastrimath [22] developed and put into use a low-cost FPGA RFID-Mutual Authentication (MA) system with XTEA security. By offering Reader's and Tag's challenge and Response utilizing XTEA security, the RFID-MA incorporated Reader and Tag authentication. The RFID-MA procedure was completed faster overall because of XTEA's pipelined design, which combined parallel execution of key scheduling with encryption and decryption processes. RFID incorporated the XTEA with Cypher block chaining (CBC) for protected MA applications. Based on the challenge and response between the Reader and Tag utilizing XTEA-CBC, the authentication procedure was successful. The security of XTEA is constrained by its vulnerability to complex cryptographic techniques such as differential cryptanalysis. For long-term security, more powerful encryption algorithms and frequent upgrades are required.

Chen et al [23] focused on DDoS attack traceback techniques in SDN-based SC. Relevant reports from the past few years were analysed, and it was discovered that the current approaches were less adaptable overall and require more time and resources. As a result, this research provided a simple traceback system based on anomaly trees. By examining network traffic fluctuations, this approach created an anomaly tree. It then calls on several detection algorithms that satisfy the necessary conditions to reduce the tree and ultimately identify the attack path. The main weakness in the method is that it is vulnerable to erroneous data from hacked base station nodes, which might result in imprecise anomaly identification and traceability of attack paths. It is additionally susceptible to noise and inconsistencies since it depends on consistent network traffic patterns.

Abdulkadhim et al [24] presented a more advanced, lightweight Modified XTEA Algorithm that protected against node abuse attacks and side-channel vulnerabilities. Provide a design in this work that used chaotic systems to create encryption keys, making them more unpredictable and random. This research's main goal is to strengthen security protocols against a variety of modern attack methods, ensuring complete defense, unpredictable behavior, and resilience. The purpose of implementing strategic defenses and strategies is to protect important resources from potential harm. Even with these improved security measures, the complexity of chaotic key generation still causes the updated XTEA method to operate poorly on very limited hardware.

Ragab et al [25] demonstrated that the XXTEA lightweight block cypher used fewer memory and computing cycles, so it is a better fit for usage in IoT smart devices for message encryption. Additionally, the elliptic curve cryptography (ECC) asymmetric cipher was chosen over RSA because it provides a higher level of bit security at smaller key sizes. To ensure authenticity, integrity, and non-repudiation, the ECC cipher was employed. For secrecy, the XXTEA block cipher was employed. Additionally, each time data is encrypted, the script hashing algorithm is utilized to confirm data integrity and produce numerous keys. By combining ECC, XXTEA, and script, the suggested hybrid cryptosystem satisfies the four primary requirements of cryptography: secrecy, authenticity, integrity, and non-repudiation. However, the physical setup of the suggested hybrid cryptosystem needs to be addressed.

From this review, it is noted that [16] introduces unexpected weaknesses that can impact reliability, in [17] results in increased computational and storage overhead, and in [18]

observed that their optimized hardware implementations while improving performance, did not thoroughly address the trade-offs between security and resource usage. [19] it still requires further optimization for speed, power consumption, and area reduction, in [20] faces challenges due to the limited memory and processing power of edge devices, in [21] highlighted that their GWOAP-based controller arrangement, while balancing IoT traffic, could be limited by the complexity of the optimization process, in [22] vulnerable to advanced cryptographic attacks and requires more robust encryption solutions. [23] reported that their anomaly tree-based DDoS traceback method, while simplifying detection, is susceptible to inaccuracies from erroneous data and noise, in [24] suffers from poor performance on highly constrained hardware due to the complexity of chaotic key generation.

III. Adaptive Secure XTEA for Embedded Microcontrollers in Smart Cities

As smart cities rely more on networked devices for crucial urban infrastructure, the demand for strong security measures grows. XTEA is a lightweight encryption solution designed for resource-constrained contexts, making it a good fit for embedded processor-based SDN nodes. However, using XTEA in these systems involves specific obstacles that need to be overcome to maintain effective security. Hence, an **Adaptive Secure XTEA for Embedded Microcontrollers (ASX-EM)** is proposed to address the goals of mitigating vulnerabilities caused by POA, optimizing performance for 8-bit microcontrollers, and developing customized XTEA encryption methods for embedded processor-based SDN nodes in smart cities. Many existing embedded processor-based SDN controllers ignore adequate padding validation to improve decryption performance. This carelessness makes them vulnerable to Padding Oracle Attacks, which allow attackers to change ciphertext and exploit incorrect answers from the SDN controller. By iterating through potential modifications, attackers can decrypt sensitive data block by block, compromising the integrity and confidentiality of the communication. To mitigate POA vulnerabilities, the Context-Aware Key Expansion and Secure Padding Validation (CAKE-SPV) are introduced. Here, A Context-Aware Key Expansion Scheduling technique is developed, which customizes the key scheduling process to node-specific factors such as MAC address and node ID. This ensures that even if one key is compromised; the security of other nodes remains intact. Additionally, the OAEP with HIPV mechanism provides secure padding and integrated validation. OAEP adds random padding and a cryptographic hash to the plaintext before encryption. During decryption, the HIPV checks hash consistency, aborting if the message is tampered with, thus ensuring ciphertext integrity and preventing information leakage through padding oracle attack vulnerabilities. This technique adds an extra layer of security to XTEA encryption, safeguarding against potential POA risks.

Furthermore, smart city infrastructure frequently relies on a large number of interconnected devices, many of which are powered by 8-bit microcontrollers since they are inexpensive and consume little power. Such devices manage important tasks including public safety, environmental monitoring, and traffic control. Many existing XTEA designs, which include multiple rounds of complicated arithmetic and bitwise operations, are not appropriate for 8-bit microcontrollers. These devices struggle with computational overhead, causing considerable delays during encryption and decryption. The restricted processing power and memory increase latency difficulties, preventing real-time data transfer. As a result, essential applications in smart city infrastructure may encounter delays, jeopardising the security and efficiency of sensitive data exchange. This inefficiency presents a substantial difficulty for implementing strong encryption in resource-limited contexts. Hence, an Adaptive Round and Parallel Processing (ARPP) method is introduced to optimize XTEA for 8-bit microcontrollers in smart city infrastructure. The Dynamic Round Adjustment technique uses Threshold-based Adaptive Control Logic to monitor system parameters such as CPU load, memory, and network traffic, and then adjusts the number of encryption rounds in real-time. This lowers the need for huge buffers or storage spaces, reducing the necessary chip area. During periods of low system load or restricted resources, fewer rounds accelerate processing without risking security. Furthermore, Bit-Slice Processing with Precomputed Lookup Tables accelerates encryption/decryption by processing bit-level blocks concurrently and obtaining precomputed values. This reduces arithmetic operations while dramatically increasing XTEA efficiency on 8-bit microcontrollers, making it suitable for smart city applications.

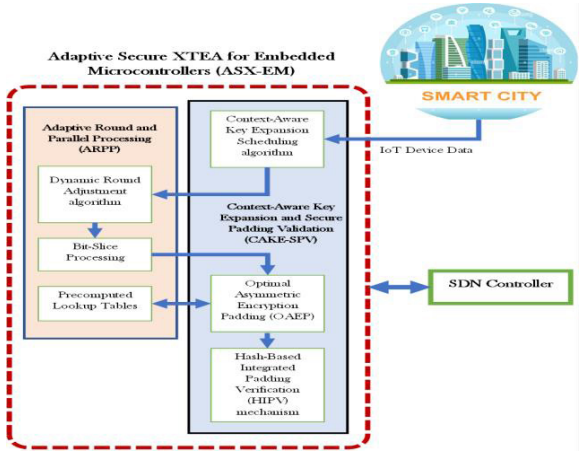


Fig.1: Overall flow diagram of the proposed model

The overall flow diagram of the suggested model is illustrated in the figure 1. IoT devices in the smart city transmit data to the ASX-EM framework. Initially, the Context-Aware Key Expansion method produces unique keys for each node depending on predefined factors. The Dynamic Round Adjustment program then analyses system parameters and modifies the number of encryption rounds accordingly. The XTEA algorithm is used to encrypt data, which is optimized with Bit-Slice Processing and Precomputed Lookup Tables. Furthermore, OAEP provides padding and a hash for safe transmission, while the HIPV technique maintains integrity

during decryption. Finally, the encrypted data is decrypted, and the HIPV detects manipulation. If the data is genuine, it is processed; otherwise, the operation is aborted. The SDN Controller manages the whole process, guaranteeing optimal resource allocation and network management.

A. Context-Aware Key Expansion Scheduling

The Context-Aware Key Expansion Scheduling method is proposed to improve the security of the XTEA encryption system by creating unique encryption keys for specific nodes in a dispersed network. This approach incorporates node-specific factors into the key scheduling process, resulting in unique, pseudo-random encryption keys for each node. This customization enables each node to produce unique encryption keys, which improves security, particularly in dispersed networks. This solution dramatically enhances the security of XTEA encryption by preventing a key breach from influencing the security of other nodes.

The algorithm begins by generating a base key K_b from node-specific parameters. let MAC be the node's MAC address and NodeID be its unique identification. These parameters serve as the basis for generating unique keys for each node. The unique key for each node is generated by combining the base key with the node-specific parameters, which is mathematically represented as in equation (1)

$$K_b = f(MAC, NodeID) \quad (1)$$

Where K_b is the base key. f is a cryptographic function that combines the base key with the node-specific parameters, which is designed using a hash function to ensure that the output is pseudo-random and unique for each node. Once the base key is generated, the key scheduling process begins. The key scheduling function K_{sch} takes the base key K_b and a context parameter C_i. The context parameter is a combination of node-specific parameters or an additional random value for added security.

$$K_{sch}(K_b, C_i) \rightarrow \{K_b^{(1)}, K_b^{(2)}, K_b^{(3)}, \dots, K_b^{(n)}\} \quad (2)$$

The basic key K_b is extended into round keys $\{K_b^{(1)}, K_b^{(2)}, K_b^{(3)}, \dots, K_b^{(n)}\}$ for XTEA encryption rounds. The Context-Aware Key Expansion Scheduling algorithm generates these round keys in a manner that incorporates node-specific context. The key scheduling is adaptive; this is change based on the operational context of the node. To further strengthen the uniqueness of the keys, the algorithm modifies the round keys based on contextual information:

$$C_i = h(MAC, NodeID, V_i) \quad (3)$$

Where V_i is a nonce to ensure the uniqueness of the context parameter, and h adjusts the round key based on the current context, ensuring that even under similar conditions, the keys remain distinct. The key scheduling function K_{sch} uses a pseudo-random number generator (PRNG) seeded with the base key K_b and context parameter C_i to produce the round keys. The PRNG ensures that the round keys are unique and pseudo-random.

$$K_f = PRNG(C_i, V_i) \quad (4)$$

Here, the V_i is a unique number used to prevent replay attacks, ensuring that each key generated is distinct even if the same base key is used. The Context-Aware Key Expansion Scheduling approach generates round keys that are subsequently employed in the XTEA encryption process.

Algorithm 1: Context-Aware Key Expansion Scheduling algorithm

Input: MAC Address, Node ID, Context Parameters C_i, Nonce V_i, and
Step 1: Generate a unique key for node i
Step 2: Generate the context parameter $C_i$using the function h
Step 3: Initialize RoundKeys array
Step 4: Key scheduling using PRNG seeded with base key K_b and context parameter C_i
for i from 0 to n do
$K[i] \leftarrow K_{sch}((K_b, C_i))$
end for
Step 5: Key scheduling function K_{sch} uses a PRNG: $K[i] = PRNG(C_i, V_i)$
Step 6: Return unique context-aware key K_f
Output: A series of unique, context-aware keys for encryption.

This promises that even if one node's key breaches, the security of other nodes is preserved through unique sub-keys formed from node-specific factors. This pseudo-random and unique key expansion considerably improves the security of the XTEA encryption technique. This customization makes it difficult for an attacker to derive keys for other nodes even if one key is compromised.

B. Dynamic Round Adjustment Algorithm

The Dynamic Round Adjustment technique optimizes the XTEA encryption process for 8-bit microcontrollers used in smart city infrastructures. With this approach, XTEA encryption security and performance are optimally balanced since the number of encryption rounds is constantly adjusted based on real-time system parameters, even under varying computational loads and resource availability. Threshold-based adaptive control logic monitors various system metrics and makes decisions about the number of encryption rounds required at any given time.

In the first step, the system predefined the number of encryption rounds minimum (R_{min}) and maximum (R_{max}). The intended security level and the microcontroller's capabilities are used to define these values. Additionally, threshold values are set for network traffic, CPU load, and memory availability. The number of encryption rounds will be increased or decreased based on these thresholds. These thresholds help the algorithm decide when to adjust the number of encryption rounds.

The algorithm continuously monitors real-time data on the identified system metrics (CPU load, available memory, and network traffic). The monitored values are compared against predefined thresholds to evaluate the system's current state.

CPU Load Evaluation

- If the CPU load is below a low threshold, it indicates low processing demand, allowing the system to afford more encryption rounds for enhanced security.
- If the CPU load is above a high threshold, it suggests high processing demand, prompting a reduction in the number of rounds to free up processing power.

Memory Evaluation

- High available memory allows for more encryption rounds without risking memory overflow or significant slowdowns.
- Low available memory necessitates reducing the number of rounds to conserve resources.

Network Traffic Evaluation

- Low network traffic permits more encryption rounds as the system can handle additional processing without impacting transmission speed.
- High network traffic requires fewer encryption rounds to maintain timely and efficient data transmission.

The algorithm calculates the optimal number of encryption rounds (R_{op}) based on a weighted function of the monitored metrics:

$$R_{op} = R_{min}\left(\frac{\text{CPU load+memory+Traffic factor}}{3}\right) \times (R_{max} - R_{min}) \quad (5)$$

The calculated R_{op} is then used to update the number of encryption rounds in real time. The number of encryption rounds is dynamically adjusted in real time based on the ongoing assessment of system metrics. This adjustment helps keep the encryption process both efficient and secure, even as system conditions change. Finally, the algorithm applies the updated number of rounds to the XTEA encryption process. This continuous modification ensures that the system keeps running without any problems, dynamically adjusting to the present situation. The encryption and decryption operations now proceed with the adjusted rounds, ensuring that the system operates efficiently without compromising security. The Dynamic Round Adjustment Algorithm is shown in the following algorithm 2.

Algorithm 2: Dynamic Round Adjustment Algorithm

1. Start

2. Initialize Parameters

Set R_{min}, R_{max}, and Define thresholds for CPU load, memory availability, and network traffic

3. Monitor System Metrics

Continuously collect data on CPU load, available memory, and network traffic

4. CPU Load Evaluation

If CPU load < low threshold, increase encryption rounds

If CPU load > high threshold, decrease encryption rounds

5. Memory Evaluation

If available memory is high, increase encryption rounds

If available memory is low, decrease encryption rounds

6. Network Traffic Evaluation

If network traffic is low, increase encryption rounds

If network traffic is high, decrease encryption rounds

7. Calculate Optimal Number of Rounds

$$R_{op} = R_{min}\left(\frac{\text{CPU load} + \text{memory} + \text{Traffic factor}}{3}\right) \times (R_{max} - R_{min})$$

Adjust the number of encryption rounds to R_{op}

Implement the adjusted number of rounds in the XTEA encryption process

8. Repeat from step 3

9.End

By altering the number of encryptions rounds dynamically, the Dynamic Round Adjustment algorithm method reduces the need for huge buffers or storage locations. This is especially important for 8-bit microcontrollers, which have limited memory and computing capability.

Once the number of encryption rounds is updated, the Bit-Slice Processing is used to divide data into smaller bit-level segments, allowing several encryption processes, which is explained in the following section 3.3.

C. Bit-Slice Processing

Conventional byte-oriented processing processes data in 8-bit (or larger) chunks, resulting in inefficiencies while performing concurrent activities. Hence the bit-slice processing is used in this research, which divides data into discrete bits such that many bits carry out operations concurrently. This method greatly increases efficiency and speed by enabling the simultaneous execution of many encryptions or decryption operations.

The specific architecture of the 8-bit microcontroller determines how bits are efficiently processed in parallel. This includes the availability of parallel execution units and the capability to handle bit-level operations. By dividing the data into slices, multiple bit-wise operations are executed simultaneously, improving throughput and reducing processing time on 8-bit microcontrollers. The Bit-Split function is defined as follows in equation (6)

$$B_s(Q(y)) = B_k \oplus B_{k-1} \oplus, \ldots, \oplus B_i \quad (6)$$

Here, $\oplus$ denotes the bitwise XOR operation, and B_k, B_{k-1},…, B_i represent the different segments obtained by splitting the bit sequence. These bit-level blocks are then processed simultaneously.

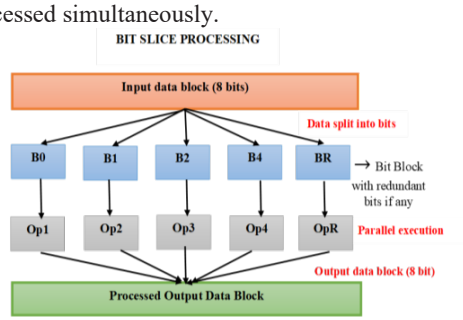


Fig.2: Bit-Slice Processing

Figure 2 depicts the bit slice processing of the proposed model. This implies that multiple encryption/decryption operations are performed at the same time, making use of parallel processing units within the microcontroller. This bit-slice processing offers a flexible and scalable approach to processor architecture, allowing for effective execution of arithmetic and logical operations in parallel. This method is especially useful in resource-constrained situations, such as 8-bit microcontrollers, where optimal performance and resource utilization are crucial. Bit-slice processing improves flexibility, speed, and efficiency in microprocessor designs by breaking down the data route into smaller slices. Once the bit-slicing operation is done, OAEP is utilized for proper encryption, which is explained in the following section 3.4.

D. Optimal Asymmetric Encryption Padding

To provide a secure padding, the OAEP is utilized for encrypting the plaintext message from the IoT device. OAEP is a padding strategy that is frequently used in conjunction with asymmetric encryption algorithms to increase security, particularly against POAs and other cryptographic vulnerabilities. The process of OAEP requires numerous phases, from appending padding to the plaintext to validating the message's integrity after decryption.

Begin with the plaintext message needs to be encrypted. Let M be the original plaintext message. The first step is to determine the required length of the padding. The total length of the padded message needs to match the block size of the encryption algorithm. Apply padding to M to ensure the total length is a multiple of the block size n required by the encryption algorithm. Extra bytes are added to the plaintext to ensure that it fits the encryption algorithm's required block size. This stage also includes adding random bytes to the message to ensure that the same message encrypted several times produces distinct ciphertexts, hence increasing security. Let $P(M)$ denote the operation of adding padding P to M:

$$P(M) = P \oplus M \tag{7}$$

Where $\oplus$ denotes concatenation, and P is the random padding added to make the message length compliant with block size requirements. The padding is random, enhancing security by making it difficult for attackers to predict the padding structure. Then compute a cryptographic hash $H(M)$ of the original plaintext M, which is expressed in the following equation ()

$$h = H(M) \tag{8}$$

This hash h is important for verifying the integrity of the message during the decryption process. The generated hash value is attached to the message along with the random padding. Concatenate the padded message $P(M)$ with the hash h, which is expressed in the following equation ()

$$D_M = P(M) \oplus h \tag{9}$$

This combined message D_M includes both the padded plaintext and the hash, ensuring both data integrity and security. Encrypt the concatenated message D_M using the XTEA encryption algorithm with a key K_f.

$$C = E_{K_f}(D_M) \tag{10}$$

Here, K_{K_f} is the encryption function, and C is the resulting ciphertext. The encryption key K_f is uniquely generated for each node using the Context-Aware Key Expansion Scheduling algorithm, ensuring that the keys are pseudo-random and node-specific. The encryption process transforms the combined message into ciphertext, ensuring its confidentiality during transmission. The encrypted message (ciphertext C) is transmitted to the intended recipient. By including random padding in the message, OAEP promises that even if the same plaintext is encrypted numerous times, the resultant ciphertext is unique each time. This randomization makes it far more difficult for an attacker to anticipate or manipulate the ciphertext. OAEP is designed to work effectively within the constraints of 8-bit microcontrollers. The use of a lightweight hash function and efficient padding mechanisms ensures that the encryption process remains fast and resource-efficient. The algorithm for OAEP is explained in Algorithm 3.

Algorithm 3: Optimal Asymmetric Encryption Padding

Inputs: Plaintext message M, Block size n of the encryption algorithm, encryption key K_f generated using the Context-Aware Key Expansion Scheduling algorithm, and Hash function H

Output: Ciphertext C

1. Compute the length of the padding P needed to make M fit the block size n.
2. Generate a random padding P of appropriate length.
3. Concatenate the random padding P with the plaintext message M:
4. Compute the cryptographic hash h of the original plaintext M
5. Combine the padded message P(M) with the hash h
6. Encrypt the combined message D_M using the XTEA encryption algorithm:

The resulting ciphertext C is returned.

Send the ciphertext C to the intended recipient.

To optimize the XTEA encryption algorithm by using precomputed lookup tables, reducing real-time computational overhead, and improving processing efficiency on 8-bit microcontrollers. Frequently used arithmetic operations in the XTEA algorithm are precomputed and stored in lookup tables. During encryption or decryption, instead of performing the computation in real time, the algorithm retrieves the result from the table. The main purpose is to speed up encryption and decryption by avoiding repetitive computations, especially for computationally expensive operations. By accessing precomputed values, the need for real-time arithmetic operations is minimized, leading to faster encryption and decryption. Minimizes processor cycles and memory usage, making it suitable for 8-bit microcontrollers with limited resources.

The OAEP method is enhanced by HIPV, which checks the integrity of the decrypted message. By recalculating the hash after decryption and comparing it to the initial hash, the system

assures that the message has not been altered, which is explained in the following section 3.5.

E. Hash-Based Integrated Padding Verification

HIPV is designed as an additional layer for improving the security of encrypted data by including padding validation directly into the decryption process. It aims to prevent vulnerabilities corresponding to the POA by ensuring that only valid ciphertexts are decrypted.

Upon receiving the ciphertext, the SDN controller decrypts it using the XTEA decryption algorithm. This restores the padded and hashed plaintext message. The recipient receives the ciphertext that needs to be decrypted. Decrypt the incoming ciphertext to get the padded plaintext and appended hash. The ciphertext C is decrypted using the XTEA decryption function $D_{K_f}(\cdot)$ with key K_f, restoring the combined message M'.

$$M' = D_{K_f}(C) \tag{11}$$

The decrypted message is divided into two parts: the original message with padding and the hash value. This extracted hash was appended during the encryption phase and serves as a reference for integrity verification. The decrypted message M' is split into the padded plaintext P' and the extracted hash H_P'.

Then the random padding bytes are removed from the decrypted message to extract the original plaintext. This step restores the plaintext to its original form before padding and hashing. The padding bytes q are removed from P', yielding the extracted plaintext P. A new hash is calculated from the extracted plaintext using the same cryptographic hash function as in the encryption phase. This recalculated hash (H_P'') is compared with the extracted hash to verify the integrity of the message. Calculate the hash of the extracted plaintext P', which is expressed in the following equation ()

$$H_P'' = H(P') \tag{12}$$

Compare the recalculated hash H_P'' with the extracted hash H_P'. If the recalculated hash matches the retrieved hash, it means that the message was not tampered with. The plaintext is regarded as valid. If the hashes do not match, it means that the message was tampered with. In this situation, the decryption operation is terminated, and the plaintext is not utilized. This hash consistency check detects ciphertext manipulation and padding issues during decryption. If the hashes do not match, the decryption is halted, ensuring the message's integrity

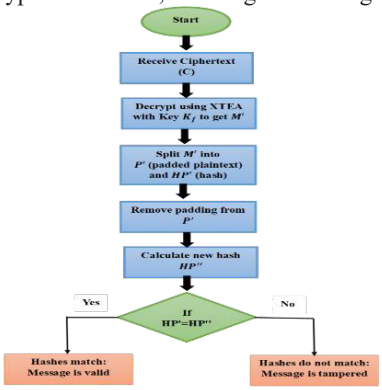


Fig.3: Flow chart for Hash-Based Integrated Padding Verification

Figure 3 illustrates the flow chart for HIPV. HIPV offers an adequate framework for maintaining the integrity and security of encrypted messages. By combining hashing and padding verification, it overcomes significant vulnerabilities including the POA while remaining efficient, making it ideal for embedded processor-based smart city systems. This structured security technique ensures that only genuine and untampered ciphertexts are handled, which considerably improves the encryption scheme's overall resilience. The OAEP with HIPV approach improves the security of encrypted data by ensuring that the plaintext is securely padded and verifiable.

Overall, the proposed approaches solve a variety of difficulties while considerably improving the security, performance, and efficiency of XTEA encryption for integrated processor-based SDN nodes in smart cities.

IV. RESULT AND DISCUSSION

The performance of the proposed system and the implementation findings are explained in depth in this section, which also includes a comparison section to verify that the suggested technique is appropriate for data security in embedded processor-based SDN nodes in smart cities.

A. System Configuration

The proposed data security methodology has been simulated in MATLAB. The evaluation is conducted by varying the data size correspondingly.

Software	: MATLAB
OS	: Windows 10 (64-bit)
Processor	: Intel i5
RAM	: 8GB RAM

The simulation in MATLAB was conducted using an embedded microcontroller model, which includes hardware-supported operations such as bitwise logic, arithmetic computations, and memory access. The microcontroller supports fixed-point arithmetic and instruction-level optimizations, enabling efficient execution of encryption algorithms.

B. Performance of the proposed model

This section discusses the experimental results from the initial setup of the suggested model for ASX-EM to mitigate POA vulnerabilities, and optimize performance for 8-bit microcontrollers in embedded processor-based SDN nodes in smart cities.

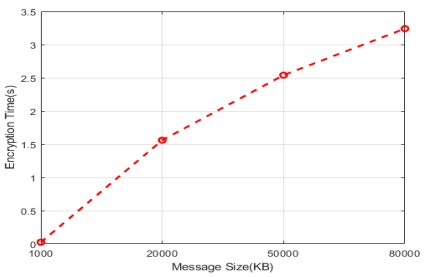


Fig.4: Encryption time of the suggested model

The suggested model's encryption time is displayed in Figure 4. When the message size is 80000KB, the suggested model achieves an encryption time of 3.25s, while when the message size is 1000KB, it achieves an encryption time of 0.07s. If the message size increases, the encryption time of the proposed model also increases. OAEP uses a lightweight hash function and efficient padding procedures to keep the encryption process speedy and resource-efficient, thereby reducing the encryption time.

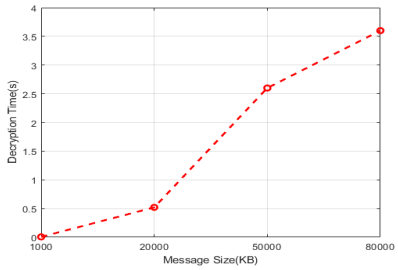


Fig.5: Decryption time of the suggested model

Figure 5 shows the suggested approach's decryption time. The suggested model achieves the fastest decryption time of 3.6 seconds when the message size is 80000KB and the fastest encryption time of 0.07 seconds when the message size is 1000KB. By processing many bits at once and utilizing parallel execution units, the bit slice process technique shortens the time required for encryption and decryption without increasing clock speed.

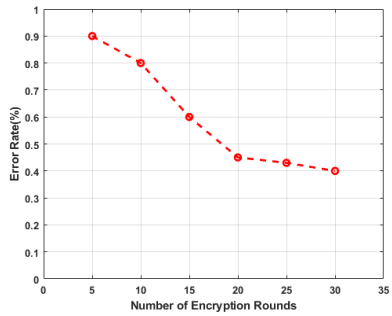


Fig.6: Error rate of the suggested model

Figure 6 presents the error rate of the suggested model across different encryption rounds. At 25 encryption rounds, the model exhibits an error rate of 0.9%, whereas at 5 encryption rounds, the error rate decreases to 0.4%. By verifying that only authentic and unaltered ciphertexts are decrypted, CAKE-SPV in conjunction with HIPV lowers the error rate associated with improper decryption.

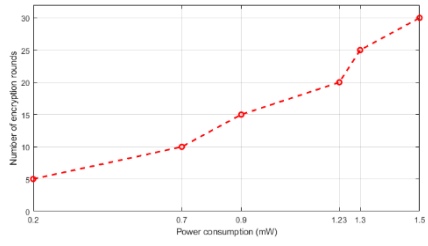


Fig.7: Power consumption of the suggested model

Figure 7 displays the power consumption of the recommended model for varying encryption rounds. The graph indicates that as encryption rounds increase, so does the power consumption. When the encryption rounds are at 5, the power consumption is 0.2mW, whereas when the encryption rounds are increased to 30, the power consumption of the suggested model is 1.5mW. By reducing the computational complexity and number of required operations, these Precomputed Lookup Tables help in maintaining lower power consumption.

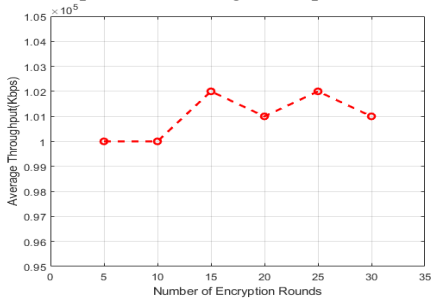


Fig.8: Average throughput of the suggested model

Figure 8 depicts the suggested model's average throughput. The suggested model achieves an average throughput value of 1.00000 Kbps when the number of encryption rounds is 5 and also attains an average throughput value of 1.03000 Kbps when the encryption rounds are 25. By optimizing the number of rounds based on system conditions using dynamic round adjustment, the average throughput is maintained at an optimal level, balancing security needs and system performance.

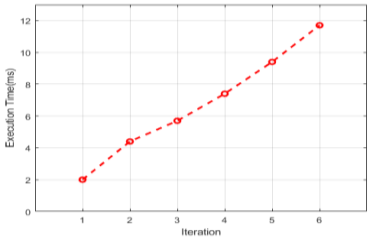


Fig.9: Execution time of the suggested model

The execution time of the suggested model is illustrated in the above figure 9. When the iterations are 6 and 1, the suggested method achieves the execution time of 11.8 ms, and 2 ms respectively. The proposed model's execution time expands with the number of iterations. The dynamic round adjustment algorithm reduces execution time by having logic constantly check system parameters and decide the ideal amount of encryption rounds in real time.

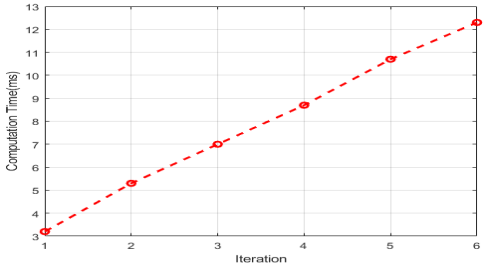


Fig.10: Computation time of the suggested model

The suggested model's computation time is represented in Figure 10. The computation time increases linearly with each iteration, starting from around 3.1 ms in the first iteration to 12.4 ms in the sixth iteration. Precomputed lookup tables and parallel processing are significantly reducing computation times by eliminating needless calculations and accelerating encryption and decryption processes.

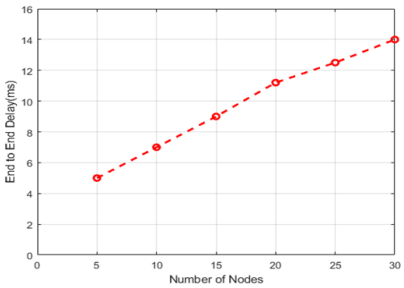


Fig.11: End-to-end delay of the suggested model

The suggested model's end-to-end delay is shown in Figure 11. The mode's end-to-end delay increases as the number of nodes rises. When the number of nodes is 5 the suggested approach attains a delay of 5 ms, also when the number of rounds is 30 the suggested approach attains a delay of 14 ms respectively. The suggested ASX-EM's ability to control end-to-end latency over a range of node densities emphasizes its applicability for scalable smart city applications, ensuring safe and effective communication even as the network increases.

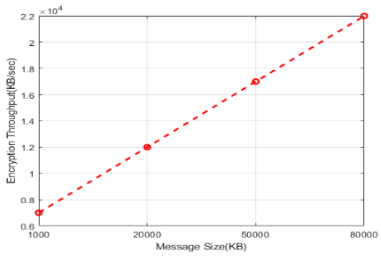


Fig.12: Encryption throughput of the suggested model

Figure 12 shows the encryption throughput across various iterations. The proposed model achieves 22000KB/sec and 6000KB/sec encryption throughput, respectively, at 80000KB and 1000KB message sizes. The encryption throughput of the proposed model rises with message size. The proposed ASX-EM maintains a high level in terms of encryption throughput, which indicates that the optimizations, such as Bit-Slice Processing, effectively sustain throughput performance without degradation over multiple iterations.

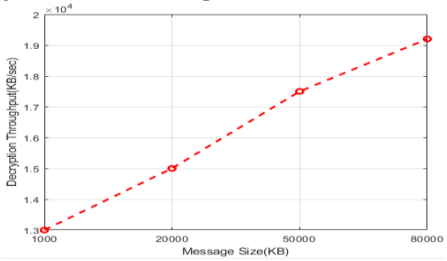


Fig.13: Decryption throughput of the suggested model

Figure 13 shows the decryption throughput across various iterations. When the message size is 80000KB and 1000KB, the suggested approach achieves a decryption throughput of 19200KB/sec and 13000KB/sec, respectively. The proposed ASX-EM maintains a high level in terms of decryption throughput, which indicates that the optimizations, such as Bit-Slice Processing, effectively sustain throughput performance without degradation over multiple iterations.

C. Comparative analysis of the proposed model

In this section, a detailed explanation of the effectiveness of the suggested technique and the achieved outcome were explained. According to the evaluation, the following metrics have been considered: encryption throughput, decryption throughput, End-to-End delay, encryption time, decryption time, throughput/area, latency, and execution time**.**

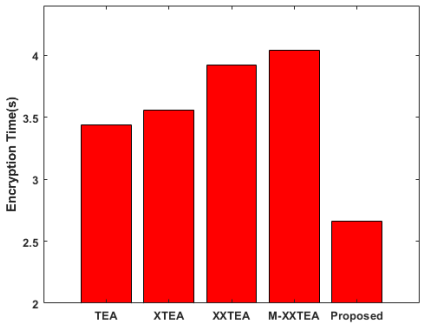


Fig.14: Comparison of encryption time

A comparison of the encryption time of the suggested model with existing models at 50000 KB message size is shown in Figure 14. The encryption time is used to measure how quickly plaintext data can be converted into ciphertext, which is essential for ensuring efficient and real-time secure communication in smart city applications. The existing models [16] such as TEA, XTEA, XXTEA, and M-XXTEA attain an encryption time of 3.43s, 3.55s, 3.91s, and 4.03s, Whereas the proposed model achieves an encryption time of 2.7s. Compared to previous approaches, the proposed model has less encryption time.

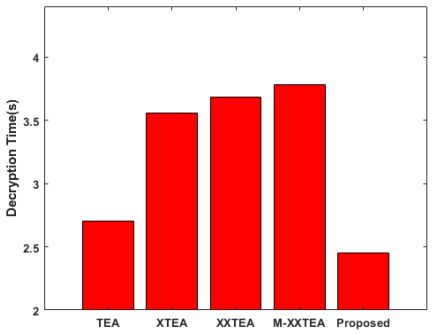


Fig.15: Comparison of decryption time

A comparison of the decryption time of the suggested model with existing models at 50000 KB message size is shown in

Figure 15. The decryption time is essential for assessing how quickly an encrypted message can be converted back to its original form, which is particularly crucial for real-time data access in smart city applications. The existing models [16] such as TEA, XTEA, XXTEA, and M-XXTEA attain a decryption time of 2.7s, 3.57s, 3.66s, and 3.78s, Whereas the proposed model achieves a decryption time of 2.4s. Compared with all the above existing models the proposed model attains a low decryption time.

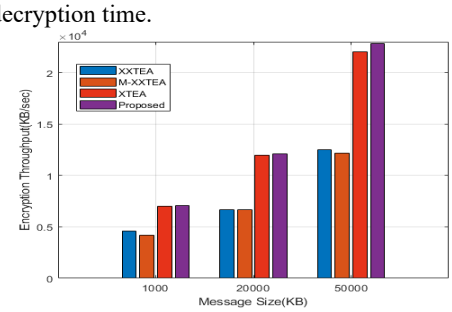


Fig.16: Comparison of encryption throughput

Figure 16 illustrates the encryption throughput of the suggested model with the existing model [16]. Encryption throughput measures the rate at which data is encrypted per unit of time, typically in KB/sec or Mbps. It is crucial for smart city applications where large volumes of data need to be securely processed in real-time. In encryption throughput, the proposed model demonstrates superior performance, especially for larger message sizes (50,000 KB), reaching around 22,000 KB/sec compared to existing models such as 13,000 KB/sec for XXTEA, 12,500 KB/sec for M-XXTEA, and 20,000 KB/sec for XTEA respectively.

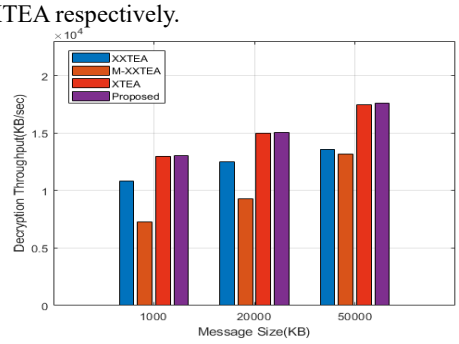


Fig.17: Comparison of decryption throughput

Figure 17 illustrates the decryption throughput of the suggested model with the existing model [16]. Decryption throughput measures how efficiently an encrypted message can be converted back to its original form per unit of time. This is crucial in smart city applications, where real-time data access is essential for traffic control, environmental monitoring, and public safety. The suggested model outperforms existing models in terms of decryption throughput, particularly for larger message sizes (50,000 KB), reaching roughly 17600 KB/sec compared to 13,800 KB/sec for XXTEA, 13,000 KB/sec for M-XXTEA, and 17,500 KB/sec for XTEA.

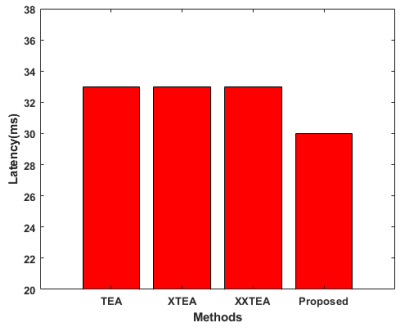


Fig.18: Comparison of latency of the suggested model

A latency comparison between the suggested model and the current models is shown in Figure 18. Latency measures the time delay between the input of a data packet and its corresponding output after encryption or decryption. The latency for TEA, XTEA, and XXTEA [18] in the existing models is 33ms, 33ms, and 33ms, respectively. The lowest latency of the suggested model is 30 ms when compared to the existing models.

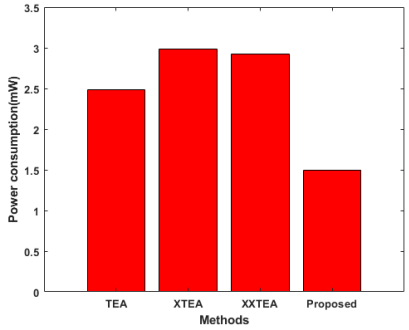


Fig.19: Comparison of power consumption

The power consumption comparison of the suggested model with the existing models is represented in Figure 19. Power consumption is used to evaluate the energy efficiency of the encryption model, which is crucial for embedded microcontroller-based SDN nodes in smart cities. The various existing [18] models including TEA, XTEA, and XXTEA attain a power consumption value of 2.5mW, 3mW, and 2.92mW respectively. Compared with existing models the suggested model achieves a low power consumption of 1.5mW.

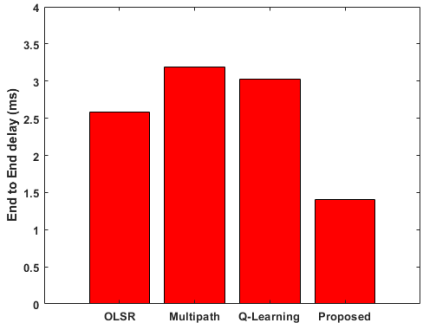


Fig.20: Comparison of end-to-end delay

The comparison of the suggested model's End-to-End Delay with existing models is shown in Figure 20. End-to-end delay measures the total time taken for a data packet to travel from the source to the destination, including encryption, transmission, processing, and decryption delays. The existing models [27] such as OLSR, Multipath, and Q-Learning are attaining an end-to-end delay value of 42ms, 18ms, and 15ms. Compared with existing models the suggested model achieves the lowest end-to-end delay of 14ms.

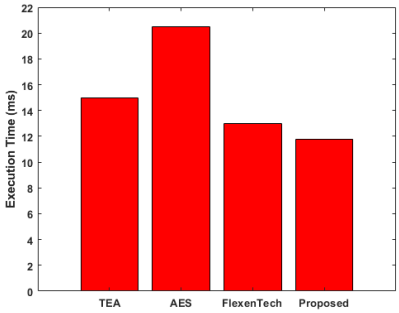


Fig.21: Comparison of execution time

A comparison of the execution time of the suggested model with various existing models is depicted in Figure 21. Execution time measures the total time required for the encryption and decryption processes to complete. This metric is crucial for embedded microcontroller-based SDN nodes in smart cities, as they have limited computational resources and operate in real-time environments. Various existing models [26] such as TEA, AES, and FlexenTech have an execution time of 15ms, 20.5ms, and 13ms, conversely, the suggested mode attains an execution time of 11.8ms. Compared with existing models the suggested model attains the lowest execution time.

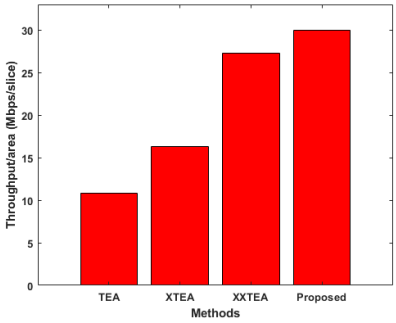


Fig.22: Comparison of throughput/area

Figure 22 illustrates the comparison of the throughput/area of the suggested model with the existing models. Throughput measures the rate at which data is successfully processed (encrypted or decrypted) over time. This metric is crucial for embedded microcontroller-based SDN nodes in smart cities, as they process large volumes of real-time data from connected devices like traffic management systems, surveillance cameras, and environmental sensors. The existing models [18] such as TEA, XTEA, and XXTEA attain a throughput/area value of 10.87 Mbps/slice, 16.27 Mbps/slice, and 27.30 Mbps/slice. Compared with existing models the suggested model attains the highest throughput/area value of 30 Mbps/slice.

TABLE I
COMPARISON OF THE PROPOSED MODEL WITH OTHER EXISTING APPROACHES

Metric	Proposed Model	TEA	XTEA	XXTEA	M-XXTEA
Encryption Time (s)	2.7	3.43	3.55	3.91	4.03
Decryption Time (s)	2.4	2.7	3.57	3.66	3.78
Encryption Throughput (KB/sec)	22,000	-	20,000	13,000	12,500
Decryption Throughput (KB/sec)	17,600	-	17,500	13,800	13,000
Latency (ms)	30	33	33	33	-
Power Consumption (mW)	1.5	2.5	3	2.92	-
Throughput/ Area (Mbps/slice)	30	10.87	16.27	27.30	-

Table 1 proposed model outperforms TEA, XTEA, XXTEA, and M-XXTEA in encryption and decryption time, achieving the fastest execution of 2.7s and 2.6s, respectively. It also demonstrates superior encryption and decryption throughput of 22,000 KB/sec and 17,600 KB/sec while maintaining the lowest power consumption of 1.5 mW. Additionally, the proposed model achieves the highest throughput per area of 30 Mbps/slice, highlighting its efficiency in resource utilization.

TABLE II
COMPARISON OF THE PROPOSED MODEL END-TO-END DELAY AND EXECUTION TIME WITH OTHER EXISTING APPROACHES

Metric	Proposed Model	TEA	OLSR	Multipath	Q-Learning	AES	Flexen Tech
End-to-End Delay (ms)	14	-	42	18	15		
Execution Time (ms)	11.8	15	-	-	-	20.5	13

Table 2 shows the proposed model achieves the lowest end-to-end delay of 14 ms compared to OLSR of 42 ms, Multipath of 18 ms, and Q-Learning of 15 ms, ensuring faster data transmission. It also outperforms TEA and AES in execution time, completing tasks in 11.8 ms, which is faster than TEA of 15 ms and AES of 20.5 ms. These results highlight the proposed model's efficiency in both latency and computational performance.

Overall, in the results section, the proposed model is compared to existing models, and the performance is explained using graphs. This shows that the technique that is used in the novelty Dynamic XTEA Optimization and Secure Key Management for Embedded Microcontroller-based SDN for smart cities has comparatively higher encryption throughput and decryption throughput, and low decryption time, encryption time, and execution time than the previous techniques that are taken for the comparison.

V. CONCLUSION

In conclusion, the proposed ASX-EM tackles major security and performance concerns for SDN nodes in smart cities. By using CAKE-SPV, the suggested approach successfully mitigates POA vulnerabilities while providing secure communication between nodes. Moreover, the ARPP method addresses limitations about chip space, power consumption, and processing time to optimize XTEA for 8-bit microcontrollers. Efficient encryption and decryption operations are made possible by the Dynamic Round Adjustment method and Bit-Slice Processing with Precomputed Lookup Tables, which balance security needs with efficiency. These combined strategies make ASX-EM a highly efficient and secure encryption method suitable for the constrained environments of smart city SDN nodes. Compared with existing models TEA, XTEA, XXTEA, AES and M-XXTEA the proposed model achieves a high encryption throughput of 22,000 KB/sec, decryption throughput of 17600 KB/sec, and low execution time of 11.8ms, power consumption of 1.5mW, encryption time of 2.7s and decryption time of 2.6s. The proposed solution not only mitigates prevalent security risks but also ensures the smooth and efficient operation of smart city infrastructures, paving the way for future advancements in secure and efficient embedded systems.

REFERENCES

[1] A. Al Hayajneh, M. Z. A. Bhuiyan, and I. McAndrew, "Improving internet of things (IoT) security with software-defined networking (SDN)," *Computers*, vol. 9, no. 1, p. 8, 2020.

[2] W. Iqbal, H. Abbas, P. Deng, J. Wan, B. Rauf, Y. Abbas, and I. Rashid, "ALAM: Anonymous lightweight authentication mechanism for SDN enabled smart homes," *Journal of Network and Computer Applications*, p. 103 672, 2023.

[3] S. Atiewi, A. Al-Rahayfeh, M. Almiani, S. Yussof, O. Alfandi, A. Abugabah, and Y. Jararweh, "Scalable and secure big data IoT system based on multifactor authentication and lightweight cryptography," *IEEE Access*, vol. 8, pp. 113 498–113 511, 2020.

[4] N. Anand, and M. A. Saifulla, "EN-LAKP: Lightweight Authentication and Key Agreement Protocol for Emerging Networks," *IEEE Access*, vol. 11, pp. 28 645–28 657, 2023.

[5] A. Ragab, G. Selim, A. Wahdan, and A. Madani, "Robust hybrid lightweight cryptosystem for protecting IoT smart devices. In Security, Privacy, and Anonymity in Computation, Communication, and Storage: SpaCCS 2019 International Workshops, Atlanta, GA, USA, July, 2019, Proceedings," *Springer International Publishing*, vol. 12, no. 14–17, pp. 5–19, 2019.

[6] N. N. Josbert, M. Wei, W. Ping, and A. Rafiq, "A look into smart factory for Industrial IoT driven by SDN technology: A comprehensive survey of taxonomy, architectures, issues and future research orientations," *Journal of King Saud University-Computer and Information Sciences*, p. 102 069, 2024.

[7] L. Vishwakarma, A. Nahar, and D. Das, "Lbsv: Lightweight blockchain security protocol for secure storage and communication in sdn-enabled iov," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 5983–5994, 2022.

[8] S.K. Shah, R. Sharma, and N. Shukla, "Data Security in IoT Networks using Software-Defined Networking: A Review. In 2022 IEEE World Conference on Applied Intelligence and Computing (AIC)," *IEEE*, pp. 909–913, 2022 June.

[9] Q. Zhou, J. Yu, and D. Li, "A dynamic and lightweight framework to secure source addresses in the SDN-based networks," *Computer Networks*, vol. 193, p. 108 075, 2021.

[10] S. Majhi, and P. Mitra, "Lightweight Cryptographic Techniques in 5G Software-Defined Internet of Things Networking. In Lightweight Cryptographic Techniques and Cybersecurity Approaches," IntechOpen, 2022.

[11] M. Rana, Q. Mamun, and R. Islam, "Current lightweight cryptography protocols in smart city IoT networks: a survey," *arXiv preprint arXiv:2010.00852*, (2020).

[12] I. S. Olimov, and X. I. Ibrohimov, "Analysis of Lightweight Cryptographic Algorithms. Golden Brain," vol. 1, no. 18, pp. 189–197, 2023.

[13] A. Diro, H. Reda, N. Chilamkurti, A. Mahmood, N. Zaman, and Y. Nam, "Lightweight authenticated-encryption scheme for internet of things based on publish-subscribe communication," *IEEE Access*, vol. 8, pp. 60 539–60 551, 2020.

[14] I. Laassar, and M. Y. Hadi, "Intrusion detection systems for internet of thing based big data: a review," *International Journal of Reconfigurable and Embedded Systems*, vol. 12, no. 1, p. 87, 2023.

[15] J. O. Olaide, "Internet of Things Security: Encryption Capacity Comparison for IoT Based on Arduino Devices," 2020.

[16] A. A. M. Ragab, A. Madani, A. M. Wahdan, and G. M. Selim, "Design, analysis, and implementation of a new lightweight block cipher for protecting IoT smart devices," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–18, 2023.

[17] M. Nagarajan, M. Rajappa, Y. Teekaraman, R. Kuppusamy, and A. R. Thelkar, "Research Article Renovated XTEA Encoder Architecture-Based Lightweight Mutual Authentication Protocol for RFID and Green Wireless Sensor Network Applications," 2022.

[18] Z. Mishra, and B. Acharya, "High throughput novel architectures of TEA family for high speed IoT and RFID applications," *Network*, vol. 61, p. 102 906, 2021.

[19] N. Khute, Z. Mishra, A.S. Rajput, and O. Parmar, "Optimized Hardware Implementation of XXTEA-192 for Resource Constrained Applications. In 2022 IEEE International Conference on Current Development in Engineering and Technology (CCET)," *IEEE*, pp. 1–6, 2022 December.

[20] D. Z. Fawwaz, S. H. Chung, C. W. Ahn, and W. S. Kim, "Optimal distributed MQTT broker and services placement for SDN-edge based smart city architecture," *Sensors*, vol. 22, no. 9, p. 3431, 2022.

[21] S. K. Keshari, V. Kansal, and S. Kumar, "A cluster based intelligent method to manage load of controllers in SDN-IoT networks for smart cities," *Scalable Computing: Practice and Experience*, vol. 22, no. 2, pp. 247–257, 2021.

[22] R. Anusha, and V. Shastrimath, "FID-MA XTEA: Cost-Effective RFID- Mutual Authentication Design Using XTEA Security on FPGA Platform," *International Journal of Electronics and Telecommunications*, vol. 67, 2021.

[23] W. Chen, S. Xiao, L. Liu, X. Jiang, and Z. Tang, "A DDoS attacks traceback scheme for SDN-based smart city," *Computers & Electrical Engineering*, vol. 81, p. 106 503, 2020.

[24] A. A. A. Abdulkadhim, A. S. Mahmood, and M. R. Ghanim, "Block of Data Encryption Using the Modified XTEA Algorithm," *Ingénierie des Systèmes d'Information*, vol. 29, no. 3, 2024.

[25] A. A. M. Ragab, A. Madani, A. M. Wahdan, and G. M. Selim, "Hybrid cryptosystems for protecting IoT smart devices with comparative analysis and evaluation. In Proceedings of the Future Technologies Conference (FTC) 2019," *Springer International Publishing*, vol. 1, pp. 862–876, 2020.

[26] O. A. Khashan, R. Ahmad, and N. M. Khafajah, "An automated lightweight encryption scheme for secure and energy-efficient communication in wireless sensor networks," *Ad Hoc Networks*, vol. 115, p. 102 448, 2021.

[27] L. El-Garoui, S. Pierre, and S. Chamberland, "A new SDN-based routing protocol for improving delay in smart city environments," *Smart Cities*, vol. 3, no. 3, pp. 1004–1021, 2020.

Sunil Kumar Shah is curently PhD scholar in Electronics Engineering at Amity University Madhya Pradesh, Gwalior, India. He did B. Tech from The Institution of Engineers (India) and M. Tech from Gyan Ganga Institute of Technology and Sciences, Jabalpur, India, in Embedded Systems andVLSI Design. His area of specialization includes IoT systems and VLSI Design for digital circuits.

Raghavendra Sharma received his B. Tech in Electrical Engineering from Dayalbagh University Agra, M. Tech in Electronic Design & Technology from Central Government Institute CEDTI, Aurangabad (MS), and PhD from Dayalbagh University, Agra. He is working as a professor and head at the Department of Electronics & Communication Engineering, Amity University Madhya Pradesh, Gwalior. He is a Fellow of the Institution of Electronics and Telecommunication Engineers (IETE), and the Institution of Engineers (IEI), and a member of many professional bodies such as IEEE, IETI (USA), and IET, etc. He has filed sixteen patents, out of which fifteen are published. His current research interests include Digital Signal Processing, Image Processing, VLSI Design, Antenna Design, and Soft Computing. He can be contacted at e-mail: rsharma3@gwa.amity.edu.

Neeraj Shukla received his PhD in Computer Science and Engineering from MANIT Bhopal in 2014. He has over two decades of academic and research experience, having served as Professor at Gyan Ganga Institute of Technology and Sciences, Jabalpur, and previously at Gyan Ganga College of Technology and Shri Ram Institute of Technology. His research interests include Image Processing, Internet of Things (IoT), Machine Learning, Network Security, Wireless Sensor Networks, and Big Data applications.

The Effect of Large Jumps in a Ring-like Quantum Network

Botond L. Márton, *Student Member, IEEE*, and László Bacsárdi, *Member, IEEE*

***Abstract*—Quantum communications promises major changes in today's communication networks by sending qubits over long distances. These qubits enable large-scale quantum computing or information-theoretically secure distribution of symmetrical keys. One of the main enablers is quantum teleportation, which makes sending quantum information between two nodes possible even when they are far apart. From these nodes, one can build a larger quantum network, but due to the nature of quantum physics, certain tasks that are well understood in classical networks, such as routing, cannot be handled in a similar way. Our work focuses on modifying a previously created model for a ring-like quantum network and assessing the effect of introducing a new node type. Our results show that this node can alter properties of the underlying network. We also look at the possibilities of modeling the capacity of the network as well as the availability of the newly introduced edges, which open interesting questions for future research.**

***Index Terms*—quantum communication networks, entanglement, routing**

I. Introduction

QUANTUM computing and communication use the laws of quantum physics to create a new field of technology, which in theory enables us to perform specific tasks better than what is achievable with our current classical systems. For example, in computing, quantum machines can efficiently break the RSA cryptosystem [1] by solving the factoring problem using Shor's algorithm [2]. Quantum random number generation [3], quantum sensing [4], and simulation [5] are also promising fields, which allow one to create good random numbers, to make more precise measurements, and to better understand certain physical systems [6]. For quantum communication, there are also promising new solutions, like quantum key distribution (QKD) [7], which makes it possible to share a secret key between two parties in such a way that they can even detect if an eavesdropper is present [8].

Another aspect of quantum communications is to connect nodes powered by quantum technologies, therefore creating a network that transmits quantum information. For example, the current implementations of quantum computers are in the NISQ era (Noisy Intermediate-Scale Quantum) and only offer a few hundred quantum bits (qubits), which are prone to suffering from different errors during computation and readout, significantly reducing their effectiveness. A possible way to tackle this problem (apart from creating more efficient and scalable error-correcting codes) is to connect these small-scale computers through a network, therefore creating a distributed but larger-scale machine [9]. A network like this is quite different from its classical counterpart and can enable new services for users while promising new challenges for researchers as well.

If one wants to send information encoded into a qubit (the information-carrying entity in quantum technologies) to another party, the most straightforward way is to send it through the appropriate medium. For example, if we encode the quantum information into the polarization of a single photon, then one can use an optical fiber for transmission. This can only work for a certain distance, because the so-called No-Cloning Theorem prohibits the perfect copying of an unknown qubit; therefore, we cannot use conventional optical amplifiers. To circumvent this problem, the communicating parties can utilize the quantum teleportation protocol [10]. This approach uses an entangled qubit pair and involves sending only two classical bits from one side to the other. This enables the receiving party to successfully recreate the unknown quantum state, but the original one will be destroyed (so the No-Cloning Theorem is not violated).

With this procedure, we can connect nodes, and they will be able to teleport qubits to their neighbors. If two non-neighboring nodes wish to exchange information, they will use the entanglement swapping protocol [11], where an entangled pair between the end nodes will be created with the help of the intermediate nodes, which can later be used for teleportation. From this, we create a quantum network consisting of nodes, which are connected by links capable of sending and receiving qubits (as well as classical information) and also performing certain operations on these qubits to implement teleportation and entanglement swapping.

As we have seen, quantum systems behave differently compared to their classical counterparts, which is also evident in quantum networks. One of the main problems in a communication network is routing, where one wants to find the best path under given circumstances from one node to another. In classical networks, this is a well-known problem with several different solutions. In the quantum case, we have to approach this task differently. Here, our main goal is to have an entangled pair of qubits between the source and the destination node. This means finding the right path along which entanglement swapping can be performed. However, the qubits used in the network must be handled carefully, so their state does not change during operations or transmission. Additionally, the time for which one can store them without losing their state is also limited. These requirements have to

B. L. Márton and L. Bacsárdi are with the Department of Networked Systems and Services, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Budapest, Hungary. (e-mail: bmarton@hit.bme.hu, bacsardi@hit.bme.hu)

The research was supported by the Ministry of Culture and Innovation and the National Research, Development and Innovation Office within the Quantum Information National Laboratory of Hungary (Grant No. 2022-2.1.1- NL-2022-00004).

DOI: 10.36244/ICJ.2025.4.9

be taken into consideration, which means that the currently used classical algorithms need to be updated or new solutions have to be developed.

The physical implementation of a quantum network has been recently demonstrated, but with only a few nodes [12] [13]. Researchers also demonstrated the possibility of a quantum router [14]. The main challenges are to maintain the state of the qubits for extended periods of time, so they can be used during the protocols (this is the field of quantum memories) and to efficiently implement the required operations, which enables one to perform the necessary steps without introducing errors. Although the technology is not yet at a stage where large-scale quantum networks (like a quantum internet) can be created, it is nonetheless important to study them and to create algorithms for managing and controlling this new network type.

Our contribution is an extension of a previous model with a new type of node, which is motivated by real-world network topologies and improves specific characteristics of the underlying quantum network. We also carry out availability and capacity simulations to study the dynamic behavior of the network.

The structure of this paper is the following. Section II introduces the necessary background and the related work regarding routing in quantum networks. In Sec. III, we present the base model that we used in our work and also our extensions to it. Sec. IV details our results regarding the average route length and swap counts in Sec. IV-A, while Sec. IV-B introduces our capacity and availability modeling. At the end, Sec. V concludes our work.

II. Overview of quantum internet protocols

To create a large quantum network, we need protocols, both quantum and classical, as well as quantum memories [15]. On the classical side: routing, signaling, or synchronization protocols [16] are necessary to guide and manage the quantum links. In contrast, the quantum protocols are, of course, essential for transmitting quantum information. Here, we introduce two quantum protocols and the related work regarding the routing problem.

I. Quantum teleportation: During teleportation [10], Alice would like to send a quantum state to Bob, without sending it over directly (note that this state can be unknown to Alice). For this operation, they will use a previously shared entangled pair of qubits, with one qubit being at Alice's side and the other at Bob's. Alice will first entangle the quantum state with their half of the pair and then measure them in a given basis. The measurement results will be sent to Bob on a classical channel (this means two bits), and based on these, Bob will perform certain operations on their qubit. At the end of the procedure, the qubit on Bob's side will be in the same state as the qubit that was meant to be teleported. As this state was measured by Alice, its superposition is lost and cannot be recovered. The act of entangling qubits and then measuring them (in a specific basis) is called the Bell measurement.

It is essential to note that we cannot create a faster-than-light communications protocol using teleportation, as the information necessary for Bob to perform the given operations must be sent over a classical channel. Without it, Bob will only have the right state with probability less than one.

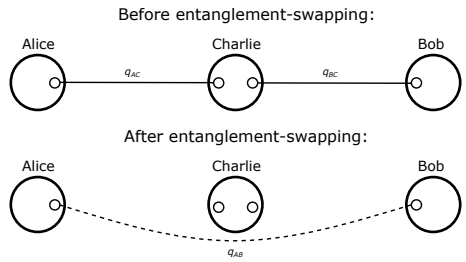


Fig. 1. Entanglement swapping with three participants. Alice and Charlie, as well as Charlie and Bob, have previously created entangled pairs denoted by q_{AC} and q_{BC}. After Charlie performs the Bell measurement, Alice and Bob will have an entangled pair q_{AB} between them, and they will be directly connected to each other.

II. Entanglement swapping: In the context of entanglement swapping [11], we have three parties: Alice, Bob and Charlie. Alice and Bob are not directly connected, but both of them have Charlie as their neighbor. A connection here means that they share a quantum and a classical channel, on which they can perform the previously introduced quantum teleportation protocol. The consequence of this requirement is that they have a shared entangled pair of qubits for every connection. The starting position of the protocol can be seen at the top of Fig. 1.

Charlie will perform the Bell measurement on their qubits, belonging to separate entangled pairs. After the measurements, the results will be sent to either Alice or Bob, allowing them to make the necessary corrections. After these steps are done, Alice and Bob will have a shared entangled pair. This means that they will be connected, even though they were not neighbors beforehand. In this setup, Charlie can be thought of as a quantum relay node. The final state of the three nodes is depicted at the bottom of Fig. 1. Using this technique, if there is a line of nodes in a quantum network, the nodes at the ends of this path can be easily connected by completing the swap protocol on the intermediate nodes.

One of the first works about routing entanglement through a quantum network was presented in [17], where the authors extended Dijkstra's algorithm to maximize the number of entangled pairs between two endpoints. After this, most work focused on the same problem in the context of a specific network structure. In [18] the researchers used ring and sphere-like networks (this work is based on this model) to create a framework for entanglement routing, the authors of [19] studied a diamond-shape structure, while [20] and [21] looked at specific lattice and grid-based topologies, where routing decisions can be easier. Quite recently, there were a number of publications regarding routing in an arbitrary graph while also supporting multi-path routing (in this problem, we have to serve multiple source-destination pairs) [22]: [23] uses fidelity as its base metric, while [24] introduces the "cost-vector analysis".

Another approach is the stochastic modeling of the quantum network, which can incorporate the different imperfections

of the implementation as well as the probabilistic nature of quantum physics. In [25], the authors studied a line of repeaters, while in [26] they focused on a star-like topology.

It is also beneficial to look at specific working environments, where entanglement routing might be used, as it might provide more possibilities for routing. For example, routing in a satellite-based network [27] or in the case of faulty quantum memories [28].

As the physical implementation and the theoretical study of quantum networks progress, it is also important to create a common framework or network stack, which can be used to manage larger networks and upon which applications can be built [29] [30].

III. The Base Model and Our Extensions

A. The base model

The base model that we modified during our work was created by Schoute et al. [18], and the authors' main goal was to create a simple network structure that makes the routing decision easier. They created a ring and sphere-like network, but we focus only on the ring-based approach.

The graph $G_n = (V_n, E_n),\ n \in \mathbb{N}$ is a network with $N = 2^n = |V_n|$ nodes and undirected edges. The nodes in the graph are labeled mod 2^n, that is $V_n = \{0, 1, \ldots, N-1\}$ and an edge $e = (\alpha, \beta)$ with $\alpha, \beta \in V_n$ is part of E_n if:

$$|\alpha - \beta| \equiv \gcd_2(\alpha, \beta) \mod 2^n,$$

where $\gcd_2(\alpha, \beta)$ is the largest power of two that divides both α and β.

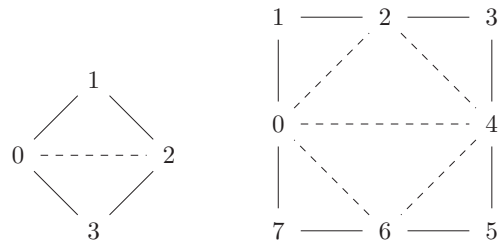
From this it can be seen that the N-long circle C_N is a subgraph of G_n and if α is divisible by 2^k, then there is and edge to $\alpha \pm 2^k \pmod N$, which makes it possible to skip 2^k nodes on C_N. The authors of [18] also grouped the edges into two categories. The edges of C_N correspond to physical links in the network, that is, a connection where quantum teleportation can be performed. Any other edge not present in C_N is a virtual quantum link (or VQL), which can be created from the physical links with the help of entanglement swapping. On Fig. 2 we can see the graph G_2 and G_3. The solid edges represent physical links, while the dashed ones are VQLs created from these links.

It can also be proven (and it is evident from the example of G_2 and G_3) that G_{n-1} is also a subgraph of G_n. Furthermore there is an elegant recursion step which can be used to generate G_n from G_{n-1}: For every physical link $e_{\alpha,\beta}$, $\forall \alpha, \beta \in V_{n-1}$ and $|\alpha - \beta| = 1$ in G_{n-1} create a new node γ and two more edges (α, γ) and (γ, β). After relabeling the nodes from 0, these new edges will be the physical links in G_n, and the "old" physical links will become new VQLs.

In [18] a number of different properties for G_n was proven, but the most important is about the diameter of the graph (the longest path from all possible shortest paths), which is only $d(G_n) = O(\log N) = O(n)$ compared to $d(C_N) = O(N)$.

B. Extensions to the base model

We extended the base model in the following way. Although the diameter of the graph significantly changed compared to



(a) The graph G_2 (b) The graph G_3

Fig. 2. The graphs G_2 and G_3 created by the definitions given in [18]. The solid edges represent physical links, while the dashed edges are VQLs (Virtual Quantum Links) created by entanglement swapping along the physical ones.

C_N, there are still nodes, the ones with an odd label, which are not part of a smaller circle. To overcome this property, we extend the base model with a new node, called the central node C. The central node can connect to any of the nodes in G_n through a new physical link, and with the help of entanglement swapping, it can also create new VQLs. Our primary goal with this node type is to further reduce the diameter (similar to the small-world property) and also the number of swap operations required for a given source-destination pair. If we think of the nodes in G_n as base stations in a telecommunications network, the original physical links and VQLs can be viewed as common interfaces, while the new central node represents a possible connection to the backbone network.

A central node can have at most $k \leq N$ edges, and as the subgraph of physical links belonging to C has a star topology, the new VQLs generated by entanglement swapping create a complete graph with k nodes. An example of G_4 with a central node C is presented in Fig. 3. The original VQLs of G_4 are not shown for better readability.

It is also important to discuss the extensions needed during the routing process. In [18], the main part of the routing algorithm is the `path`(α, β) function, which gives back the shortest path between nodes α and β. This function calls the `path2`(α, β) and the `bestMove`(α, β) subroutines. The `path2` function gives back the shortest path of length at most 2, if there is any, while the `bestMove` gives back the next

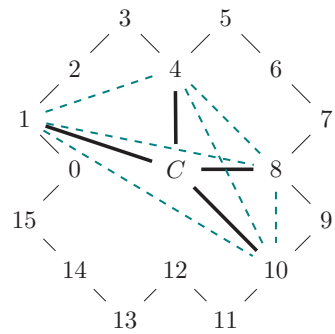
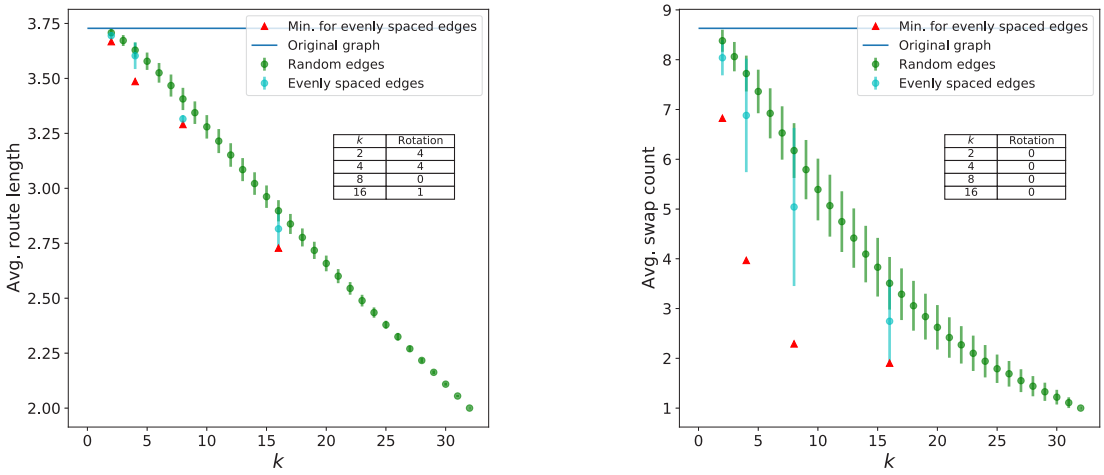


Fig. 3. A possible central node (C) with four new physical edges (solid lines) and all possible VQLs (teal dashed lines) in G_4. Note that the original VQLs of G_4 are omitted.



(a) (b)

Fig. 4. The average route length (a) and swap count (b) with their standard deviation for G_5 with two central node types. The horizontal line shows the values for the original G_5. For the case of a central node with evenly spaced edges, the optimal rotation value is also shown on the figures and summarized in a table. Sample size: 1000

node, which is guaranteed to be on the shortest path between α and β.

With the introduction of the central node, it is also necessary to extend these two subroutines. The `path2` will now check if the nodes are neighbors through C, while the `bestMove` will also examine the possibility of whether we can get closer to β if we go through the central node (this is only considered if α is connected to C).

IV. Results and Discussion

A. The effect on average route length and swap counts

To measure the effect of a central node on G_n, we first looked at two important properties of the newly created quantum networks. The first one is the average route length (or average path length), which is different from the diameter of the graph. This is an important characteristic of any network, as it measures how easily one can navigate the network and reach other nodes. A lower average route length indicates that most nodes are accessible from any point in the network in fewer steps, which helps the flow of information. The second one is the average swap count required to create an entangled pair between a given source-destination pair. It is desirable to keep the average swap count as low as possible, as the entanglement swapping protocol is a complex task and needs additional classical communication to succeed, which adds delay to the overall communication step. Calculating the swap count on the original G_n for a given route can be done in the following way: If the path is $v_1, v_2, v_3, \ldots, v_k$ with $v_i \in V_n \ \forall i = 1, \ldots, k$, then the swap count that is necessary to "get" from v_1 to v_2 is distance of the two nodes on the ring C_N minus 1.

If we extend the graph with a central node, then we have to check for every consecutive node in the path whether they are connected through the central node or not. If they are, the swap count needed is just 1. After this, the swap count for the entire path can be calculated by adding up the swaps for the subsequent steps and adding one, as a final swap is necessary at the end to create the desired connection.

During our simulation, we looked at two types of central nodes. The first one is a node with $k \leq N$ random edges, which connect to different nodes on G_n. The second one is a central node that has $k \leq N/2$ evenly spaced edges connecting to the nodes. This means that the central node is connected to nodes $0, \frac{N}{k}, 2 \cdot \frac{N}{k}, \ldots, (k-1) \cdot \frac{N}{k} (\text{mod } N)$ with k in the form $2^i \quad i = 1, \ldots, n-1$ or one of its possible circular rotations (e.g. starting from a different node).

For the random central node, we sampled possible edges for a given k. For the evenly spaced case, we examined possible values of k and their corresponding circular rotations. After this, we calculated the route length and the swap count for every possible source-destination pair in G_n and averaged the results, which can be seen in Fig. 4. The results show that the introduction of the central node has reduced both the average route length and the average number of swaps. For large k (close to N), this is not surprising, because the addition of a node with this many edges creates an almost complete or a complete graph of VQLs on the nodes in the network. It is also important to notice that the standard deviations of the average route length and swap count are relatively small for central nodes with random edges. This is not the case for evenly spaced edges, as the different rotation values have a strong influence on the achievable average swap counts.

We can observe that in the lower interval of k, the best results for both values occur when the central node has $N/2$ evenly spaced edges, meaning that every second node is connected to C. In this case, the rotations (there are two

possibilities) do not significantly alter the average route length, and we can calculate the diameter of the network, as stated in Proposition 1, along with a sketch of the proof. It is also important that for the evenly spaced edges, the average swap count is minimal if we only connect even nodes. The reason for this comes from the structure of G_n. Even nodes are connected to each other through VQLs and have high degree counts; therefore, they are important during path selection. But the VQLs connecting them can only be realized by a high number of swap operations. For example, realizing the dashed line between node 0 and 4 in G_3 (depicted in Fig. 2) requires three swap operations. With the help of a central node that has evenly spaced edges, we can create new VQLs between the even nodes, which require only one swap, thereby reducing the average swap count needed during routing.

Proposition 1. *The diameter of G_n extended with a central node C that connects all odd or even nodes is 3.*

Proof. In G_n, there are three possible source-destination types: even-to-even, odd-to-odd, and even-to-odd (odd-to-even is the same because of the symmetry of this particular network). For odd-to-odd, if C connects all odd nodes, then the route length is 1. For even-to-odd nodes, we can proceed from the source to one of the neighboring odd nodes and then jump to the destination, resulting in a route length of 2 or 1 if they are originally directly connected. This is also true if C connects all even nodes, as in that case we can always choose an even node neighboring the destination. For the even-to-even routes, there are three possibilities:

1) If the source and destination are connected by a VQL, then the route length is 1.
2) If their distance was 2 on G_n, then the route length remains the same.
3) If their distance was ≥ 3, then we can once again go to one of the odd neighbors, then jump to the odd node nearest to the destination (there are just two possibilities) and do one more step to the destination itself. This is a path of length 3.

The same can be said about odd-to-odd pairs if C connects the even nodes.

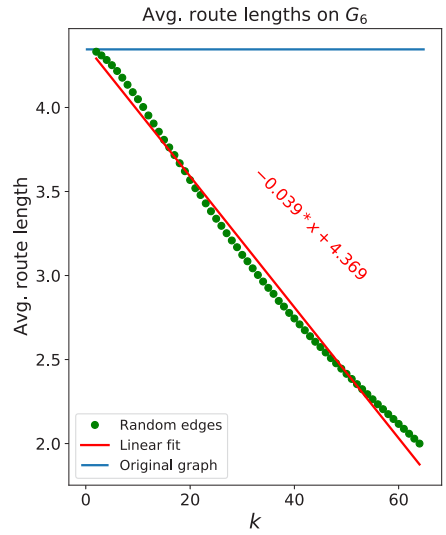
□

Another important finding is regarding the shape of the two graphs corresponding to the random edge case. This is more evident on larger networks. The average route length exhibits a linear trend, while the average swap count follows a decaying exponential pattern. This is shown in Fig. 5, where the values are plotted for G_6. The average route length was fitted with a linear function of the form $ax + b$, and for the average swap count, we used the form $ae^{-bx} + c$. Both of them provide a good fit to the data points (the exact values can be seen in the description of Fig. 5). The exact reason for the shape of these graphs is a promising future question.

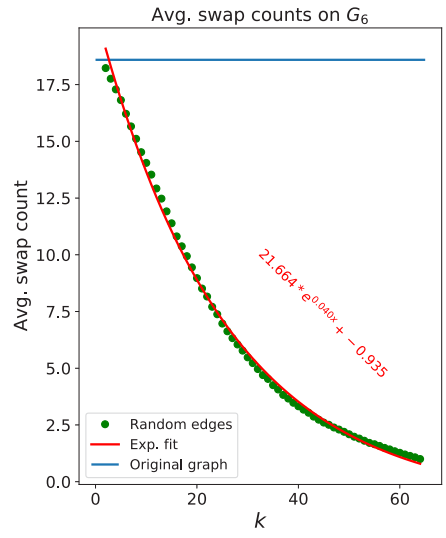
B. Capacity and availability modeling

As we saw in the previous section, the central node with a large degree has significant effects on the properties of the underlying network. This means that the central node has to have a large enough quantum memory to hold at most N qubits and a long enough decoherence time (the time after which the qubits lose their state) for using them as VQLs. However, the current physical implementation of quantum memories can only hold a few qubits at a time [15]. This means that the more edges the central node has, the less usable the VQLs become.

To model this phenomenon, we propose two simple methods. The first one gives a capacity E to the central node. This means that it can serve $\leq E$ nodes reliably. If C has k edges, the current usage is given by E/k. To show the



(a)



(b)

Fig. 5. Curve fit for the average route length (a) and swap count (b) for G_6. For the average route length, we used a linear function, and for the swap count, an inverted exponential function. In the case of the average route, the standard deviation errors for the parameters are 0.0004 and 0.0163, while for the average swap count, they are: 0.1351, 0.0008, and 0.1567. Sample size: 1000

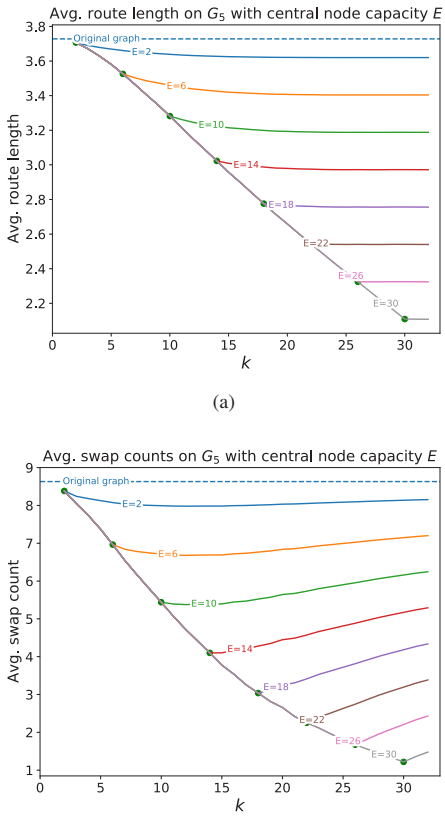


Fig. 6. Capacity modeling of the central node C with different capacity values $E = 2, 6, \ldots, N-2$. In (a), we can see the average route length, while (b) shows the average swap count. Sample size: 5000.

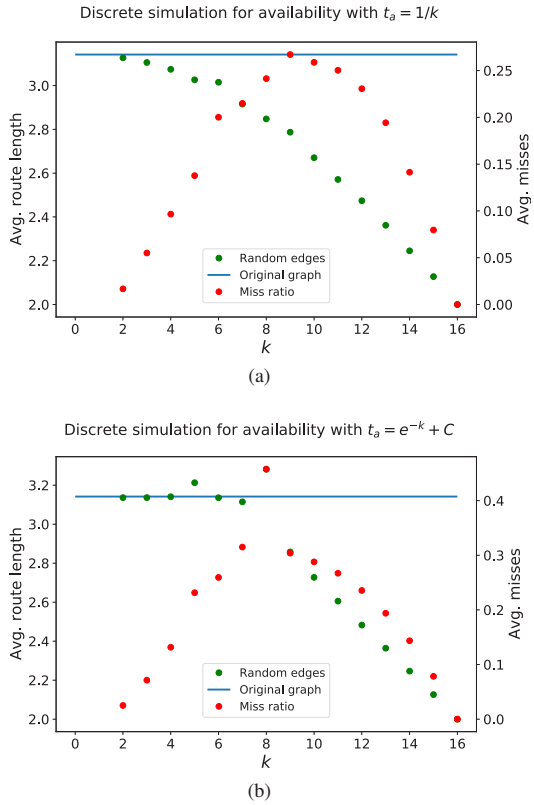


Fig. 7. Discrete time simulation results for the availability of the VQLs on G_4. On (a) the function of the availability time is $t_a = 1/k$, while on (b) it is given by $t_a = e^{-k} + c$, where c is a constant to limit the value of t_a as $k \to N$. In both (a) and (b) $t_r = 1$. Sample size: 5000

unreliability of larger quantum memories, we take the linear combination of the average route length of the original graph G_n and the average route length of G_n extended by a central node with k random edges using weights $1-e_k$ and e_k, where $e_k = \min(1, E/k)$. This combination ensures that a central node with large capacity can serve all its neighbors, but if the capacity is lower (e.g., the quantum memory used is not reliable), the properties of the network get closer to the original graph as k increases. The effect of this method with different E capacities can be seen in Fig. 6.

Regarding the average route length, the graph follows the previous results until $k \leq E$; after this, it does not improve significantly, but we can still observe a slight drop as k gets larger. This is caused by the previously described large gains introduced by a central node with random edges. The average swap count behaves differently. Here, after k gets larger than E, the graphs take a steep incline towards the original value for G_n. This effect is more noticeable when k is closer to N, because in this interval the average swap counts are in a slower decline.

For the availability modeling, we created a discrete simulation. The central node with k edges can serve requests for t_a time, after this, it takes t_r time to make the VQLs available again. The value of t_a and t_r can depend on the edges connected to the central node. We calculate the route step by step, which takes 1 unit of time. If the current node in the path calculated that the next hop is through the central node, it tries to access the VQLs. If it succeeds, the routing takes this VQL; if the request fails, we take the link that is given by the next hop on the original graph. We gather the average route length as well as the ratio of missed resource requests. The results of the simulation can be seen in Fig. 7 with different functions for t_a.

The most significant outcome in the graphs is that as the degree of the central node increases, the number of unsuccessful requests also grows. This has an effect on the average route length, as we have to use the VQLs that are present on the original graphs in a larger portion of the time. At larger k values, this negative effect quickly fades away. The reason for this is that at larger degree values, there is a higher probability that the source-destination pair is connected directly or that they are only a few hops away. Since the simulation starts at zero, the first request is successful.

V. CONCLUSION

In our work, we focused on entanglement generation in a quantum network. We introduced an extension to an existing model that adds a central node and examined the impact it had on various properties of the underlying network. Our

results show that these effects are significant, reducing both the average route length and the swap count required to connect a source-destination pair. As the current physical implementations of quantum memories are not yet at a stage where they can be reliably used in a quantum network, we have created two simple models to incorporate the capacity of a node or the availability of a quantum link. Based on these models, our simulations show that even a central node with a few edges can improve the network.

ACKNOWLEDGMENT

B. L. Márton would like to thank Balázs Solymos and Ágoston Schranz for the helpful discussions and comments on the first version of the manuscript.

REFERENCES

[1] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, no. 2, pp. 120–126, Feb. 1978. [Online]. Available: DOI: 10.1145/359340.359342

[2] P. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," in *Proceedings 35th Annual Symposium on Foundations of Computer Science*, ser. SFCS-94. IEEE Comput. Soc. Press, 1994, pp. 124–134. [Online]. Available: DOI: 10.1109/sfcs.1994.365700

[3] A. Schranz, B. Solymos, and M. Telek, "Stochastic performance analysis of a time-of-arrival quantum random number generator," *IET Quantum Communication*, vol. 5, no. 2, pp. 140–156, Dec. 2023. [Online]. Available: DOI: 10.1049/qtc2.12080

[4] D. Chandra, P. Botsinis, D. Alanis, Z. Babar, S.-X. Ng, and L. Hanzo, "On the road to quantum communications," *Infocommunications Journal*, vol. 14, no. 3, pp. 2–8, 2022. [Online]. Available: DOI: 10.36244/icj.2022.3.1

[5] I. M. Georgescu, S. Ashhab, and F. Nori, "Quantum simulation," *Reviews of Modern Physics*, vol. 86, no. 1, pp. 153–185, Mar. 2014. [Online]. Available: DOI: 10.1103/revmodphys.86.153

[6] A. J. Daley, I. Bloch, C. Kokail, S. Flannigan, N. Pearson, M. Troyer, and P. Zoller, "Practical quantum advantage in quantum simulation," *Nature*, vol. 607, no. 7920, pp. 667–676, Jul. 2022. [Online]. Available: DOI: 10.1038/s41586-022-04940-6

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

[8] E. Udvary, "Integration of qkd channels to classical high-speed optical communication networks," *Infocommunications Journa*l, vol. 15, no. 4, pp. 2–9, 2023. [Online]. Available: DOI: 10.36244/icj.2023.4.1

[9] M. Caleffi, A. S. Cacciapuoti, and G. Bianchi, "Quantum internet: from communication to distributed computing!" in *Proceedings of the 5th ACM International Conference on Nanoscale Computing and Communication*, ser. NANOCOM '18. ACM, Sep. 2018. [Online]. Available: DOI: 10.1145/3233188.3233224

[10] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels," *Physical Review Letters*, vol. 70, no. 13, pp. 1895–1899, Mar. 1993. [Online]. Available: DOI: 10.1103/physrevlett.70.1895

[11] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, ""event-ready-detectors"' bell experiment via entanglement swapping," *Physical Review Letters*, vol. 71, no. 26, pp. 4287–4290, Dec. 1993. [Online]. Available: DOI: 10.1103/physrevlett.71.4287

[12] M. Pompili, C. Delle Donne, I. te Raa, B. van der Vecht, M. Skrzypczyk, G. Ferreira, L. de Kluijver, A. J. Stolk, S. L. N. Hermans, P. Pawełczak, W. Kozlowski, R. Hanson, and S. Wehner, "Experimental demonstration of entanglement delivery using a quantum network stack," *npj Quantum Information*, vol. 8, no. 1, Oct. 2022. [Online]. Available: DOI: 10.1038/s41534-022-00631-2

[13] S. Kucera, C. Haen, E. Arenskötter, T. Bauer, J. Meiers, M. Schäfer, R. Boland, M. Yahyapour, M. Lessing, R. Holzwarth, C. Becher, and J. Eschner, "Demonstration of quantum network protocols over a 14-km urban fiber link," *npj Quantum Information*, vol. 10, no. 1, Sep. 2024. [Online]. Available: DOI: 10.1038/s41534-024-00886-x

[14] X. X. Yuan, J.-J. Ma, P.-Y. Hou, X.-Y. Chang, C. Zu, and L.-M. Duan, "Experimental demonstration of a quantum router," *Scientific Reports*, vol. 5, no. 1, Jul. 2015. [Online]. Available: DOI: 10.1038/srep12452

[15] R. K. Ramakrishnan, A. B. Ravichandran, I. Kaushik, G. Hegde, S. Talabattula, and P. P. Rohde, "The quantum internet: A hardware review," *Journal of the Indian Institute of Science*, vol. 103, no. 2, pp. 547–567, Sep. 2022. [Online]. Available: DOI: 10.1007/s41745-022-00336-7

[16] J. Illiano, M. Caleffi, A. Manzalini, and A. S. Cacciapuoti, "Quantum internet protocol stack: A comprehensive survey," *Computer Networks*, vol. 213, p. 109 092, Aug. 2022. [Online]. Available: DOI: 10.1016/j.comnet.2022.109092

[17] R. Van Meter, T. Satoh, T. D. Ladd, W. J. Munro, and K. Nemoto, "Path selection for quantum repeater networks," *Networking Science*, vol. 3, no. 1–4, pp. 82–95, Dec. 2013. [Online]. Available: DOI: 10.1007/s13119-013-0026-2

[18] E. Schoute, L. Mancinska, T. Islam, I. Kerenidis, and S. Wehner, "Short-cuts to quantum network routing," *arXiv preprint arXiv:1610.05238*, 2016.

[19] S. Pirandola, "End-to-end capacities of a quantum communication network," *Communications Physics*, vol. 2, no. 1, May 2019. [Online]. Available: DOI: 10.1038/s42005-019-0147-3

[20] M. Pant, H. Krovi, D. Towsley, L. Tassiulas, L. Jiang, P. Basu, D. Englund, and S. Guha, "Routing entanglement in the quantum internet," *npj Quantum Information,* vol. 5, no. 1, Mar. 2019. [Online]. Available: DOI: 10.1038/s41534-019-0139-x

[21] S. Das, S. Khatri, and J. P. Dowling, "Robust quantum network architectures and topologies for entanglement distribution," *Physical Review A*, vol. 97, no. 1, Jan. 2018. [Online]. Available: DOI: 10.1103/physreva.97.012335

[22] S. Shi and C. Qian, "Concurrent entanglement routing for quantum networks: Model and designs," in *Proceedings of the Annual conference of the ACM Special Interest Group on Data Communication on the applications, technologies, architectures, and protocols for computer communication*. ACM, Jul. 2020. [Online]. Available: DOI: 10.1145/3387514.3405853

[23] J. Li, M. Wang, K. Xue, R. Li, N. Yu, Q. Sun, and J. Lu, "Fidelity-guaranteed entanglement routing in quantum networks," *IEEE Transactions on Communications*, vol. 70, no. 10, pp. 6748–6763, Oct. 2022. [Online]. Available: DOI: 10.1109/tcomm.2022.3200115

[24] H. Leone, N. R. Miller, D. Singh, N. K. Langford, and P. P. Rohde, "Qunet: Cost vector analysis & multi-path entanglement routing in quantum networks," *arXiv preprint arXiv:2105.00418*, 2021.

[25] M. Caleffi, "Optimal routing for quantum networks," *IEEE Access*, vol. 5, pp. 22 299–22 312, 2017. [Online]. Available: DOI: 10.1109/access.2017.2763325

[26] G. Vardoyan, S. Guha, P. Nain, and D. Towsley, "On the stochastic analysis of a quantum entanglement switch," *ACM SIGMETRICS Performance Evaluation Review,* vol. 47, no. 2, pp. 27–29, Dec. 2019. [Online]. Available: DOI: 10.1145/3374888.3374899

[27] A. Mihály and L. Bacsárdi, "Optical transmittance based store and forward routing in satellite networks," *Infocommunications Journal*, vol. 15, no. 2, pp. 8–13, 2023. [Online]. Available: DOI: 10.36244/icj.2023.2.2

[28] L. Gyongyosi and S. Imre, "Adaptive routing for quantum memory failures in the quantum internet," *Quantum Information Processing*, vol. 18, no. 2, Jan. 2019. [Online]. Available: DOI: 10.1007/s11128-018-2153-x

[29] A. Dahlberg, M. Skrzypczyk, T. Coopmans, L. Wubben, F. Rozpundefineddek, M. Pompili, A. Stolk, P. Pawełczak, R. Knegjens, J. de Oliveira Filho, R. Hanson, and S. Wehner, "A link layer protocol for quantum networks," in *Proceedings of the ACM Special Interest Group on Data Communication,* ser. SIGCOMM '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 159–173. [Online]. Available: DOI: 10.1145/3341302.3342070

[30] W. Kozlowski, A. Dahlberg, and S. Wehner, "Designing a quantum network protocol," in *Proceedings of the 16th International Conference on Emerging Networking EXperiments and Technologies*, ser. CoNEXT '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–16. [Online]. Available: DOI: 10.1145/3386367.3431293

Botond L. Márton (M'23) Received both his B.Sc. and M.Sc. degrees in computer engineering from Budapest University of Technology and Economics (BME). He is currently pursuing his P.hD. at BME. He is involved in a quantum key distribution project at the university. His research interests are quantum computing and quantum communications.

László Bacsárdi (M'07) received his MSc degree in Computer Engineering from the Budapest University of Technology and Economics (BME) in 2006 and his PhD degree in 2012. He is a member of the International Academy of Astronautics (IAA). Between 2009 and 2020, he held various positions at the University of Sopron, Hungary, including Head of the Institute of Informatics and Economics. Since 2020, he has been an Associate Professor at the BME Department of Networked Systems and Services, where he also serves as Head of the Department and is an active member of the BME Mobile Communications and Quantum Technologies Laboratory. His current research interests include quantum computing and quantum communications. He is a past chair of the Telecommunications Chapter of the Hungarian Scientific Association for Infocommunications (HTE), a member of IEEE and HTE, and an alumni member of the United Nations–established Space Generation Advisory Council (SGAC).

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

IECON 2026

Doha, Qatar | 18-21 October

The 52nd Annual Conference of the IEEE Industrial Electronics Society

IEEE Industrial Electronics Society

Call for Papers

IMPORTANT DATES

Paper Submission deadline: April 15th, 2026

Acceptance Notification: May 15th, 2026

Final Submission and registration: July 15th, 2026

IECON 2026 is the 52nd Annual Conference of the IEEE Industrial Electronics Society (IES), focusing on contemporary industry topics ranging from power electronics, power systems, controls, manufacturing, to computational intelligence and communications. IECON is the flagship annual conference of IES. It aims to create a forum for scientists and practicing engineers throughout the world to present the latest research findings and ideas in the areas of Industrial Electronics, and possible contributions toward sustainable development and environment preservation. The objectives of the conference are to provide high quality research and professional interactions for the advancement of science, technology, and fellowship. Papers with new research results are encouraged for submission.

Regular Sessions: The regular sessions of the conference are covered but not limited to the following technical tracks:

- Power Systems and Smart Grid
- Power Electronics Converters
- Electric Machines and Industrial Drives
- Renewable Energy and Energy Storage Systems
- Resilient Control Architectures for Energy Systems
- Smart Building Technologies
- Transportation Electrification and Automotive Technologies
- Control Systems and Applications
- Mechatronics and Robotics
- Computational Intelligence and Signal and Image Processing
- Sensors, Actuators and Micro-Nanotechnology
- Electronic Systems on Chip and Embedded Systems
- Cyber-Physical Systems and Internet of Things in Industry
- Communication for Industrial and Factory Automation
- Industrial Informatics, Cloud Computing, Big Data

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: HÍRKÖZLÉSI 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*