NTRICS
2026
3rd Edition

Proceedings of the 3rd International Workshop

on New Trends in Image, Communication and Signal Processing

NTRICS'2026

January 15–16, 2026

SUP'COM, Tunis, Tunisia

Organized by

COSIM Research Lab
Higher School of Communication of Tunis (SUP'COM)

Preface

We are delighted to present the proceedings of the **3rd International Workshop on New Trends in Image, Communication and Signal Processing (NTRICS'2026)**, held on January 15–16, 2026, at the Higher School of Communication of Tunis (SUP'COM), Tunisia. Organized by the **COSIM Research Lab**, this workshop continues its tradition of gathering researchers, practitioners, and graduate students to discuss cutting-edge advancements in Artificial Intelligence, Wireless Communication, IoT, and Signal Processing.

This edition is particularly rich, featuring keynote seminars and technical sessions by distinguished international experts and local leaders in the field. We are honored to host **Pr. Moncef Gabbouj** (Tampere University, Finland) for the keynote session on AI in medical diagnostics. We also extend a special thanks to **Pr. Noureddine Lasla** (ENSIA, Algeria) for his valuable participation and expertise in secure systems.

We would like to express our sincere gratitude to all our distinguished speakers who enriched the technical program with their insights into AI, NTN, Smart Systems, and Communication Networks: **Pr. Sofia Ben Jebara, Pr. Fethi Choubani, Dr. Imen Nasr, Dr. Abderrahman Trichili, Pr. Fatma Rouissi, Pr. Soumaya Hamouda, Pr. Wael Jaafar, Dr. Rahma Gharslaoui, Dr. Yosra Abbes, Hend Baklouti, Akrem Hadji, Dr. Anis Youssef, Dr. Ibtissem Malouche,** and **Dr. Hakim Ghazzai**.

We also thank all the authors for their valuable contributions and the technical program committee for their diligent review process. A special thanks goes to our student participants in the Poster Competition, whose innovative work represents the promising future of our field.

We deeply appreciate the unwavering support of our co-organizers and partners, without whom this event would not have been possible: **Agence Nationale des Fréquences (ANF), Instance Nationale des Télécommunications (INT), Tunisie Telecom, Sup'Com Alumni, Federation of Arab Scientific Research Councils, Centre d'Etudes et des Recherches des Télécommunications (CERT), Global Training and Consulting Services (GTC),** and **ACTIA ES**.

We hope that these proceedings will serve as a valuable reference for the scientific community and inspire future research collaborations.

Tunis, January 2026

The Organisation Committee
Hichem Besbes, Sameh Najeh, Hakim Ghazzai, and Leila Najjar

جـامعـة قـرطـاج
Université de Carthage

SUP'COM
المدرسـة العليـا للمواصـلات بتونـس
Ecole Supérieure des Communications de Tunis
Higher School of Communications of Tunis

COSIM
Communications, Signal and Image

NEW TRENDS IN IMAGE COMMUNICATION AND SIGNAL PROCESSING WORKSHOP
(NTRICS 2026, 3rd Edition)
15-16 January 2026, SUP'COM Tunisia

NTRICS 2026
3rd Edition

NTRICS Program

DAY 1: Thursday, January 15th, 2026

Opening

09:00–09:15 **Pr. Ridha Bouallegue (SUP'COM, Tunisia)** – SUPCOM Director Opening Speech

09:15–09:30 **Pr. Hichem Besbes (SUP'COM, Tunisia)** – COSIM Director Opening Remarks

Keynote Session

09:30–10:30 **Pr. Moncef Gabbouj (Tampere University, Finland)** – Echocardiography-based Computer-Aided Diagnosis of Myocardial Infarction

10:30–11:30 ***Coffee Break & Poster Competition****: 1st Session*

Session 1: AI Enables & Smart Systems

11:30–12:00 **Pr. Sofia Ben Jebara (SUP'COM, Tunisia)** – Multimodal Cough Detection: Use of Signal Processing and Artificial Intelligence

12:00–12:30 **Pr. Fethi Choubani (SUP'COM, Tunisia)** – Beyond Smart Agriculture: A General and Comprehensive Approach

12:30–13:00 **Dr. Imen Nasr (SUP'COM, Tunisia)** – Embedded AI for Energy Intelligence: An AI-Driven Smart Home for Energy Disaggregation and P2P Energy Trading

13:00–14:00 Lunch

Session 2: Communication & Networks

14:00–14:30 **Pr. Noureddine Lasla (ENSIA, Algeria)** – Decentralization for Secure Systems

14:30–15:00 **Dr. Abderrahman Trichili (Northumbria University, UK)** – Measuring VLC Wiretap Channels

15:00–15:30 **Pr. Fatma Rouissi (SUP'COM, Tunisia)** – Cooperative Communication in Impulse Noise Impaired PLC Channel: AI-based Solutions

15:30–16:00 **Pr. Soumaya Hamouda (SUP'COM, Tunisia)** – 5G Lessons Learned: Implications for 6G

DAY 2: Friday, January 16th, 2026

Session 3: NTN Session

09:00–09:30 **Pr. Wael Jaafar (ETS, Canada)** – Non-Terrestrial Networks for Ubiquitous and Reliable Communications, Computing, and AI

09:30–10:00 **Dr. Rahma Gharslaoui (ANF, Tunisia)** – NTN: Frequencies and Associated Equipment

10:00–10:30 **Dr. Yosra Abbes, Hend Baklouti, Akrem Hadji (INT, Tunisia)** Exploring the Sky's Limitless Opportunities: Regulation and the Evolution of Non-Terrestrial Networks

10:30–11:30 ***Coffee Break & Poster Competition****: 2nd Session*

Session 4: Industrial Session

11:30–12:00 **Dr. Anis Youssef (Telnet, Tunisia)** – Sensing Space: AI-Driven IoT Networks for Nanosatellite Earth Observation

12:00–12:30 **Dr. Ibtissem Malouche (ACTIA, Tunisia)** – AI-Driven Transformation at ACTIA ES: Scaling Innovation Across All Business Units

12:30–14:00 Lunch

Session 5: Presentations and Closing Ceremony

14:00–15:00 **Student Pitches** – Presentations of the Selected Posters

15:00–15:30 **Dr. Hakim Ghazzai (KAUST, Saudi Arabia)** – Bridging AI Research and Real-World Smart City Applications

15:30–16:00 **Awards and Closing Remarks**

Program

Co-Organizers & Partners

الوكالة الوطنية للترددات
Agence Nationale des Fréquences

Tunisie Telecom

اتحاد مجالس البحث العلمي العربية
Federation of Arab Scientific Research Councils (FASRC)

GTC
Global Training and Consulting Services

CERT
مركز الدراسات و البحوث للاتصالات
Centre d'Etudes et des Recherches des Télécommunications
Research and Studies Telecommunications Centre

الهيئة الوطنية للإتصالات
Instance Nationale des Telecommunications

ACTIA®

List of Accepted Papers

NTRICS'2026 Proceedings

ACCEPTED PAPERS

Predicting Human Attention States from EEG Signals using K-means Clustering to Drive Conventional Features-Based Classifiers

Moez Zouari[1], Ibtihel Zayani[1], Sofia Ben Jebara[1,2] Malek Abidi[3,4] Giovanni de Marco [3,4]
[1] University of Carthage, Higher School of Communication of Tunis (SUPCOM), Tunis Tunisia
[2] COSIM Laboratory, SUPCOM
[3] Laboratory LINP2, UPL, Paris Nanterre University
[4] COMUE Paris Lumières University
moez.zouari@supcom.tn, ibtihel.zayenne@supcom.tn, sofia.benjebara@supcom.tn,
m.bida@parisnanterre.fr, gdemarco@parisnanterre.fr

Abstract

This article presents an approach for classifying levels of human attention using Electroencephalography (EEG) and machine learning techniques, specifically combining unsupervised and supervised learning. We used the K-means algorithm on the EEG Alpha-to-Beta power ratio to automatically define three distinct levels of attention (High, Medium, Low) and we classified these defined states using conventional supervised machine learning models applied to key EEG features. The results demonstrate that the Random Forest algorithm achieve excellent classification performance, confirming the value of exploiting both frequency and spatial information within EEG signals for robust attention level assessment.

Keywords: Attention prediction, EEG, Alpha-to-Beta power ratio.

1 Introduction

Attention is a fundamental cognitive function which permits to select and process relevant information in the environment while filtering out distractions. When driving, for example, sustained attention is essential to stay focused on the road while ignoring the scenery and other distractions. In learning, it is essential for the initial encoding of information and its transfer to memory.

This study focuses on action control, which enables the planning and targeted execution of movements and behaviors based on visual-spatial instructions and alerts. Subjects must visualize rectangles of different colors and listen to a signal of constant or variable intensity to alert them to the imminent task to be performed. As a result, they must press a key on a keyboard. The choice of keyboard key depends on the type of alert he has been given. The aim is to identify their level of attention based on the accuracy of the task performed: whether or not they pressed the correct key.

ElectroEncephaloGraphy (EEG) is a non-invasive technique that records fluctuations in the brain's electrical activity, providing a direct measure of cortical processing. The state-of-the-art literature showcases a variety of approaches and features derived from EEG signals for attention recognition [1]. For instance, methodologies based on energy descriptors within specific frequency bands have emerged as key indicators of the cognitive activation state [2,3]. Both supervised classification techniques (e.g., [2]) and unsupervised classification techniques (e.g., [5]) have been successfully employed in this domain.

In this study, we explored the use of a well established EEG feature (the energy ratio between alpha and beta frequency bands), which is widely recognized as a robust marker reflecting the status of cognitive activation [4]. We used unsupervised classification to identify attention levels (Low, Medium, High) and introduced classification methods which use handcradt features and conventional machine learners.

2 Methodology

2.1 Experience and Data

To conduct the study, a device called MEDELOPT®V1 developed by SEENEL Imaging was used [8]. It includes 8 EEG electrodes which where placed at occipital, parietal and frontal lobes.The data was collected during an experiment conducted on 24 individuals. Each participant sat in front of a computer and had to press one of two keys on the keyboard (/ at the bottom right or z at the top left) depending on what was displayed on the screen. The experiment included 256 trials with varying warning stimuli and conditions. Two types of conditions (named congruence) were studied: when the border of the rectangle is green, the response must be made on the same side as the rectangle (congruent task) whereas when the border is red, the response must be made on the opposite side (incongruent), involving additional cognitive effort. In addition, a white noise-type sound signal was used as an alert to prepare participants for the task. It could have a constant intensity throughout the experiment (endogenous mode named ENDO) or an intensity that gradually increased over time (hydrid mode of endogenous and exogenous named HYB). Furthermore, the delay between the alert and the target stimulus, called SOA (Stimulus Onset Asynchrony), varies randomly between four values: 0 ms (no alert), 100 ms, 250 ms, and 850 ms [7].

All of these different experimental conditions (16 in total) ensure that a significant cognitive load is required and that a high level of attention is demanded to each participant. This is what we are seeking to determine through this study.

2.2 EEG Preprocessing

In order to ensure the quality of the EEG signals and enable reliable analysis, preprocessing phase was applied to all recordings :

- Bandpass filtering between 0.5 Hz and 50 Hz.
- Independent Component Analysis (ICA) to reduce artifacts.
- Extraction of frequency bands related to attention : Alpha ($8 - 13\,\text{Hz}$) indicating relaxation, alert calmness and reduction in sensory activity and Beta band ($13 - 30\,\text{Hz}$) related to attention, concentration and active cognitive activation.
- For each EEG segment, the spectral energies in the Alpha and Beta bands (namely E_α and E_β) across all 8 EEG channels are calculated and merged in a ratio form :

$$R = \frac{E_\alpha}{E_\beta} = \frac{\sum_{i=1}^{8} E_{\alpha,i}}{\sum_{i=1}^{8} E_{\beta,i}}. \tag{1}$$

2.3 Unsupervised Attention Level Labeling

To automatically and objectively categorize the participant's attention state from the EEG signals, it's necessary to associate each segmented trial with a discrete label reflecting the cognitive load. We employed an unsupervised clustering approach based on the Alpha-to-Beta Energy Ratio. This specific ratio is recognized as a robust index of cortical activation and engagement, with higher values typically indicating a decline in alertnes or low-attention state, and lower values indicating higher engagement and attention [4, 9].

The K-means clustering algorithm was applied directly to the $\frac{E_\alpha}{E_\beta}$ feature values across all trials. The number of clusters was determined via the elbow method/silhouette analysis [10], which indicated that the optimal reduction in within-cluster variance (or optimal separation) was achieved at three partitions.The resulting cluster assignments provided the ground-truth labels for the subsequent supervised classification stage.

2.4 EEG Features for Supervised Classification

Following the creation of these attention labels, the next objective was to develop a predictive model capable of automatically classifying a participant's attention level based on a broader range of EEG features.

The data preparation involved five key steps: *i* labeling based on k-means clustering results, *ii* data cleaning to remove outliers, *iii* standardization (zero mean, unit variance) to equalize scales, *iv*imbalance handling using SMOTE principle to generate synthetic examples for minority classes, and *v* data splitting using 5-fold stratified cross-validation to maintain class balance in all partitions.

To ensure complementarity and non-redundancy between the unsupervised labeling method and the classification model, the features selected for the supervised learning phase were deliberately chosen to exclude the $\frac{E_\alpha}{E_\beta}$ ratio and to exhibit a low correlation with this cluster-defining variable. This strategy ensures that the classification approach is not merely replicating the clustering results but is instead leveraging novel information derived from the EEG signal to predict the categorized states.

The univariate ANOVA test was utilized as the feature selection mechanism to identify the most discriminative features (from an initial pool of 184 calculated descriptors) that showed significant variance across the three established attention groups (Low, Medium, High).

The final set of features retained for the classifier includes key parameters across the five standard frequency bands (Delta, Theta, Alpha, Beta, Gamma) are:

- Entropy per frequency band and per channel, serving as a measure of the uncertainty or complexity of the neural oscillations.
- Maximum power of each frequency band, reflecting the peak intensity of brain activity in the respective frequency ranges.
- Peak-to-peak amplitude within each frequency band, providing an indication of the magnitude of the underlying neural oscillations.
- Total energy of each frequency band, representing the overall amount of electrical activity concentrated in these frequencies.

3 Results and Discussion

3.1 Attention Level Labelling

An unsupervised k-means clustering method was applied to all extracted $R = \frac{E_\alpha}{E_\beta}$ values (6,144 data points) from all EEG segments (all subjects and conditions combined). The number of clusters was set at $k = 3$, corresponding to three levels of attention defined as follows: sustained attention(low R), fluctuating attention (intermediate R) and impaired attention (high R.).

Tab. 1 gives cluster distributions related to attention levels across experimental conditions. Its analysis revealed that participants exhibited consistently high proportions of high attention level across all modes, congruence levels, and SOA values, generally above 85%. Lower proportions were observed for low and medium attention levels, typically below 10% and 7%, respectively. Notably, incongruent trials tended to slightly reduce the proportion of high attention states and increase low/medium states, while longer SOAs generally favored high attention, suggesting that both task difficulty and preparatory cue timing modulate attentional engagement.

A three-way ANOVA with repeated measures was performed on the proportion of high attention, treating the mode (ENDO/HYB), congruence (C/I), and SOA (0, 100, 250, 850 ms) as factors. The results indicated a significant main effect of congruence (F= 448,112, p< 0.001) and SOA (F = 325.59, p < 0.001), whereas the main effect of mode was not significant (F= 1.53, p = 0.346). These findings suggest that the cognitive load induced by incongruent trials and the timing of alerting cues significantly modulated the likelihood of observing high attention states, while the attentional mode (ENDO vs HYB) did not strongly influence these proportions.

Mode	Congruence	SOA	High	Low	Medium
ENDO	C	0	86.4	7.8	5.7
HYB	C	0	86.6	7.9	5.5
ENDO	I	0	85.6	7.8	6.5
HYB	I	0	86.5	8.3	5.2
ENDO	C	100	84.3	8.6	7.0
HYB	C	100	86.6	8.4	5.0
ENDO	I	100	84.6	8.4	7.1
HYB	I	100	86.2	8.4	5.5
ENDO	C	250	87.0	8.1	4.9
HYB	C	250	86.9	8.4	4.7
ENDO	I	250	86.2	8.4	5.5
HYB	I	250	85.4	8.4	6.3
ENDO	C	850	85.4	8.4	6.3
HYB	C	850	86.9	8.4	4.7
ENDO	I	850	85.9	8.6	5.5
HYB	I	850	87.2	8.1	4.7

Table 1: Cluster distributions across experimental conditions.

Additionally, a global chi-square test on the full contingency table confirmed that the distribution of cluster assignments was significantly dependent on the experimental conditions (Chi^2 = 10,777.51, p < 0.001), corroborating the ANOVA findings and indicating that both SOA and congruence systematically affected attentional states across trials.

Note that the ANOVA applied here is appropriate for repeated-measures designs, allowing the assessment of within-subject effects, while the chi-square test provides a complementary non-parametric confirmation of overall dependence between clusters and conditions. Overall, these results quantitatively support the interpretation that task congruence and preparatory timing influence attentional dynamics, consistent with the patterns observed in the descriptive percentages.

3.2 Attention Level Prediction based on EEG Features

The performance of various classification algorithms was compared using several key metrics: Accuracy, Precision, Recall, and F1-score, as presented in Table 2 :

- The Random Forest (RF) method provided the best overall performance, achieving a high accuracy of 98.14% and a high F1-score of 0.9468. Its low stan-

Table 2: Comparison of classification algorithm performances.

Algo	Accuracy	F1-score	Precision	Recall
RF	0.9814	0.9468	0.9409	0.9530
SVM	0.7063	0.6135	0.6306	0.7563
LR	0.7961	0.6746	0.6486	0.8176
KNN	0.7837	0.6324	0.5999	0.7521
DT	0.9680	0.9111	0.8977	0.9257

dard deviation (±0.0097) confirms its robustness and reliability.

- The Decision Tree (DT) method yielded results very close to those of the Random Forest, with an accuracy of 96.80% and good stability (±0.0077). While it is slightly less performant than the Random Forest, its superior interpretability offers a strong trade-off between performance and simplicity for analyzing feature importance.
- The Support Vector Machine (SVM), Logistic Regression (LR) and K-Nearest Neighbors (KNN) methods showed significantly lower performance and stability (accuracy around 70% to 80%). While they provide high Recall, their poor Precision and low F1-scores suggest they struggled with the task and likely produced a high number of false positives for the minority classes.

4 Conclusion

This study demonstrated an hybrid approach for classifying three levels of human attention (Low, Medium, High). The methodology combined unsupervised learning to objectively generate attention labels, with supervised classification strategies. The results validated the hidg performances of Decision Tree and Random Forest models based on EEG features. Moving forward, this framework suggests that future work should explore the fusion of this approach with complementary modalities like functional Near-Infrared Spectroscopy (fNIRS) to integrate real-time hemodynamic data and further enhance the robustness of attention monitoring systems.

References

[1] Marcantoni, I.; Assogna, R., Del Borrello, G., Di Stefano, M., *et al.* "Ratio Indexes Based on Spectral Electroencephalographic Brainwaves for Assessment of Mental Involvement: A Systematic Review," *Sensors*, 2023, 23(13):5968, 2023. doi: 10.3390/s23135968.

[2] Nuamah, J. K., Seong, Y. and S. Yi, "Electroencephalography (EEG) Classification of Cognitive Tasks based on Task Engagement Index," *IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, Savannah, GA, USA, 2017, pp. 1-6. doi: 10.1109/COGSIMA.2017.7929581.

[3] You, S.D. , "Classification of Relaxation and Concentration Mental States with EEG," *Information*, 2021, 12(5):187. doi: 10.3390/info12050187.

[4] Alirezaei, M. and H. Sardouie, S., "Detection of Human Attention Using EEG Signals," *24th National and 2nd International Iranian Conference on Biomedical Engineering (ICBME)*, 2017, Tehran, Iran, pp. 1-5, doi: 10.1109/ICBME.2017.8430244.

[5] Orhan, U., Hekim, M. and Ozer, M., "EEG Signals Classification using the K-means Clustering and a Multilayer Perceptron Neural Network Model," *Expert Systems with Applications*, 38(22), 2011. doi: 10.1016/j.eswa.2011.04.149.

[6] Posner, M. I., Petersen, S. E., " The Attention System of the Human Brain," *Annual Review of Neuroscience*, 1990, 13(1), pp. 25-42. doi: 10.1146/annurev.ne.13.030190.000325. PMID: 2183676.

[7] McCormick, C. R., Redden, R. S., Hurst, A. J. and Klein, R. M., "On the Selection of Endogenous and Exogenous Signals," *R. Soc. Open Sci.6190134*, 2019. doi: 10.1098/rsos.190134.

[8] Seenel Imaging. (n.d.). *Medelopt.* https://seenel-imaging.com/medelopt-product (Accessed October 3, 2025).

[9] Lubar, J.F., "Discourse on the Development of EEG Diagnostics and Biofeedback for Attention-Deficit/Hyperactivity Disorders," *Biofeedback Self Regul*, 1991, 16(3):201-25. doi: 10.1007/BF01000016. PMID: 1932259.

[10] Thorndike, RL., " Who Belongs to the Group?," *Psychometrika*, 1953, 18, 267-276. https://doi.org/10.1007/BF02289263.

Quantifying the impact of co-registration and satellite jitter filtering on the accuracy of multi-temporal elevation differences

Imen Brini[1,2], Denis Feurer[1], Riadh Tebourbi[2], Fabrice Vinatier[1], Riadh Abdelfattah[2]

[1]LISAH, Univ Montpellier, AgroParisTech, INRAE, Institut Agro, IRD, Montpellier, France
[2]Higher School of Communications of Tunis, COSIM Lab, University of Carthage, Tunis, Tunisia
`imen.brini@supcom.tn`

Abstract

Multi-temporal satellite-based change detection analysis requires addressing biases related to co-registration and satellite jitter. This study compares two workflows combining two co-registration strategies with jitter filtering using Pléiades imagery from 2015 and 2023. The Time-SIFT approach shows improved horizontal accuracy and comparable vertical accuracy compared to the classical approach without requiring GCPs or pre-georeferenced data, reducing elevation difference dispersion from 1.71 to 1.32 m in terms of NMAD. After jitter correction, dispersions reach 1.13 m and 0.96 m NMAD for the classical and the Time-SIFT approaches, respectively. These results demonstrate that the Time-SIFT co-registration quality yields improved elevation difference after filtering.

Keywords: Co-registration, satellite jitter, Pléiades, 3D change detection

1 Introduction

Multi-temporal surface elevation data from recurrent acquisitions enable change detection through differencing. Optical satellite systems can deliver very high-resolution elevation products with minimal errors. However, two key issues affect the reliability of detecting surface changes: co-registration biases [1,2] and biases related to the satellite acquisition geometry, in particular satellite attitude variations known as jitter [3].

Co-registration biases manifest as horizontal misalignment, where elevation differences exhibit terrain-like patterns due to slope effects [3], and/or vertical offsets characterized by predominantly positive or negative values [4]. Two main strategies exist for DEM co-registration. The first involves co-registering independently generated DEMs using methods such as iterative closest point variants [5,6] or terrain-based approaches [3,7], requiring georeferenced or coarsely co-registered DEMs. The second integrates co-registration during DEM generation by jointly processing multi-temporal images. All images from different epochs are aligned together in a single bundle adjustment routine, producing inherently co-registered DEMs that do not require further alignment, herein referred as the Time-SIFT approach [8,9].

Biases associated with the satellite acquisition geometry can be substantial enough to require correction. Such corrections are applied to data products, as DEMs or elevation difference, where jitter effects are clearly observable. The jitter component is modeled as a sinusoidal function [10], expressed as:

$$\phi(t) = A\sin(2\pi f t + \phi) \tag{1}$$

where $\phi(t)$ is the jitter component of satellite attitude at imaging time t, and A, f, ϕ are the amplitude, frequency and phase, respectively.

Filtering this noise can be applied either in the spatial domain [4,11,12] or based on frequency domain using transformations such as the Fourier transform [13]. However, [11] reported that spatial filtering alone does not fully eliminate along-track undulations, while [13] demonstrated improved results using frequency-domain–based filtering. The filtering is performed using a finite-impulse-response (FIR) high-pass filter, which effectively suppresses the along-track oscillations.

While co-registration approaches could ensure geometric consistency across epochs, they do not address systematic errors introduced by satellite jitter. Complete multi-temporal processing workflows require combining co-registration strategies with jitter filtering. Two workflow can be employed: classical co-registration (e.g., [3]) followed by jitter filtering [13], or Time-SIFT co-registration [9] followed by the same jitter filtering. The relative performance of these workflows remains unquantified, yet understanding their trade-offs is essential for selecting optimal processing strategies for multi-temporal studies.

This study compares the classical and the Time-SIFT co-registration approaches, both followed by the jitter filtering. The accuracy is evaluated at multiple processing stages to quantify each method's contribution to the final elevation difference quality.

2 Methodology

2.1 Dataset

The tri-stereo Pléiades images were ordered in 2015 and 2023 covering the Lebna watershed, located in the coastal plain in northeastern Tunisia, through the DINAMIS platform. The images ordered in 2015 were captured with Pléiades-1A, while the images ordered in 2023 were captured with Pléiades-1B, a platform subject to jitter since 2021 [14].

2.2 Data analysis

In order to evaluate the performance of both workflows, some metrics are used to assess their impact on the co-registration and the jitter filtering. For the co-registration, the relative accuracy was determined by quantifying the offset between both models in xy and z directions, using 9 ground control points (GCP). For the jitter filtering, stable terrain is an indicator of the quality of the corrections. The Normalized Median Absolute Deviation (NMAD) is employed as a robust metric to evaluate accuracy after co-registration and jitter filtering, allowing the performance of each processing step to be quantitatively assessed. It is defined as:

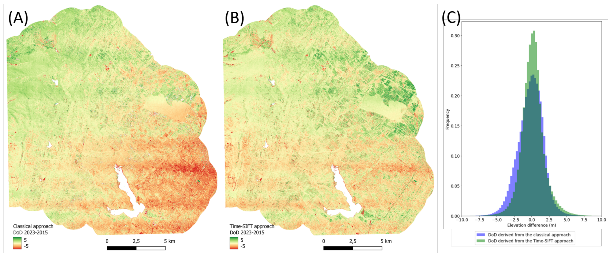
$$\mathrm{NMAD}(\Delta H) = 1.4826 \cdot \mathrm{median}(|\Delta H - \mathrm{median}\Delta H|) \tag{2}$$

where ΔH is the elevation difference.

3 Results and Discussion

3.1 DEM co-registration

As illustrated in figure 1, the elevation differences after co-registration with the classical approach exhibit a positive trend in the NW direction and negative trend in the SE direction. However, the Time-SIFT co-registration attenuates both this directional trend and the jitter magnitude, reducing the elevation differences dispersion from 1.71 to 1.32 m in terms NMAD. In figure 1(C), the histograms of the elevation differences derived from the time-SIFT approach tended mostly to 0, whereas the classical approach revealed higher deviations.



Contains information © CNES (2015, 2023), Distribution Airbus DS, all rights reserved. Commercial use is prohibited

Figure 1: DoD derived from (A) the classical approach, (B) the Time-SIFT approach, and (C) histograms of the elevation differences.

The classical approach achieves relative accuracies of 0.76 m in xy direction and 0.46 m in z direction, while Time-SIFT achieves 0.52 m in xy direction and 0.49 m in z direction. Both fall within the typical 0.2–1 m range for relative DEM accuracy [3], suitable for multi-temporal change detection.

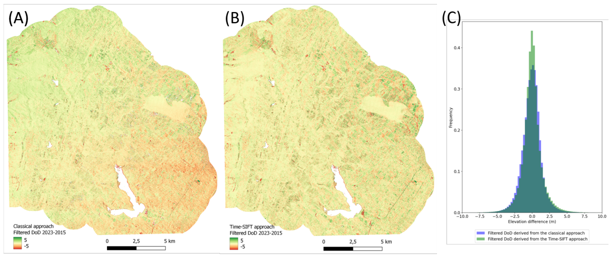
The Time-SIFT approach shows improved horizontal accuracy and comparable vertical accuracy compared to the classical approach without requiring GCPs or pre-georeferenced data, promoting its application in areas with limited reference information.

3.2 Satellite jitter correction

The elevation difference derived from the classical and the Time-SIFT approaches show an along-track jitter effect with an amplitude A reaching on average 2 m, and an undulation wavelength λ of about 2 km.

After filtering, the dispersion of elevation differences derived from the classical and the Time-SIFT approach in terms of NMAD is 1.13 and 0.96 m, respectively. As shown in figure 2, in the filtered DoD obtained from the Time-SIFT approach, the artifacts linked to the satellite jitter are further attenuated compared to the classical approach. The histogram of the filtered DoD obtained from the Time-SIFT approach indicates tighter elevation difference dispersion compared to the histogram of the filtered DoD obtained from the classical approach.

As both workflows apply the same jitter filtering, the qualitative and quantitative improvement reflects the co-registration quality of the Time-SIFT approach, demonstrating that it produces more reliable



Contains information © CNES (2015, 2023), Distribution Airbus DS, all rights reserved. Commercial use is prohibited

Figure 2: Filtered DoD derived from (A) the classical approach, (B) the Time-SIFT approach, and (C) histograms of the elevation differences.

elevation difference for detecting surface changes.

4 Conclusion

Although both workflows achieve suitable accuracy for multi-temporal change detection, the Time-SIFT approach offers significant advantages. With no need for GCPs, it substantially reduces processing time. Since the Time-SIFT approach produces inherently co-registered DEMs, applying jitter filtering on it yields imporved elevation differences dispersion compared to filtering classical co-registration outputs. This shows that improved initial co-registration quality lead to improved final products. However, the enhancement could vary across different landscapes and acquisition conditions.

Acknowledgment

This publication was made possible through support provided by the IRD.
This work has been supported by the Programme National de Télédétection Spatiale (PNTS, Grant No. PNTS-2023-05).
Data from the DINAMIS device were funded by CNES, CNRS, IGN, IRD, INRAE, and CIRAD.

References

[1] Aicardi, I., Nex, F., Gerke, M., & Lingua, A. M. (2016). An image-based approach for the co-registration of multi-temporal UAV image datasets. *Remote sensing*, 8(9), 779.

[2] Ball, J. E., Anderson, D. T., & Chan, C. S. (2017). Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community. *Journal of applied remote sensing*, 11(4), 042609-042609.

[3] Nuth, C., & Kääb, A. (2011). Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change. *The Cryosphere*, 5(1), 271-290.

[4] Li, T., Hu, Y., Liu, B., Jiang, L., Wang, H., & Shen, X. (2022). Co-registration and residual correction of digital elevation models: a comparative study. *The Cryosphere Discussions*, 2022, 1-22.

[5] Di, K., Hu, W., Liu, Y., & Peng, M. (2012). Co-registration of Chang'E-1 stereo images and laser altimeter data with crossover adjustment and image sensor model refinement. *Advances in Space Research*, 50(12), 1615-1628.

[6] Micheletti, N., Chandler, J. H., & Lane, S. N. (2015). Investigating the geomorphological potential of freely available and accessible structure-from-motion photogrammetry using a smartphone. *Earth Surface Processes and Landforms*, 40(4), 473-486.

[7] Peduzzi, P., Herold, C., & Silverio, W. (2010). Assessing high altitude glacier thickness, volume and area changes using field, GIS and remote sensing techniques: the case of Nevado Coropuna (Peru). *The Cryosphere*, 4(3), 313-323.

[8] Feurer, D., & Vinatier, F. (2018). Joining multi-epoch archival aerial images in a single SfM block allows 3-D change detection with almost exclusively image information. *ISPRS journal of photogrammetry and remote sensing*, 146, 495-506.

[9] Brini, I., Feurer, D., Tebourbi, R., Vinatier, F., & Abdelfattah, R. (2025). Evaluating the potential of the Time-SIFT approach using Pléiades satellite imagery for 3D change detection. *Journal of applied remote sensing*, 19(2), 024501-024501.

[10] Pan, J., Che, C., Zhu, Y., & Wang, M. (2017). Satellite jitter estimation and validation using parallax images. *Sensors*, 17(1), 83.

[11] Berthier, E., Lebreton, J., Fontannaz, D., Hosford, S., Belart, J. M. C., Brun, F., ... & Blondel, C. (2024). The Pléiades Glacier Observatory: high-resolution digital elevation models and ortho-imagery to monitor glacier change. *The Cryosphere*, 18(12), 5551-5571.

[12] Deschamps-Berger, C., Gascoin, S., Berthier, E., Deems, J., Gutmann, E., Dehecq, A., ... & Dumont, M. (2020). Snow depth mapping

from stereo satellite imagery in mountainous terrain: evaluation using airborne lidar data. *The Cryosphere Discussions*, 2020, 1-28.

[13] Brini, I., Feurer, D., Tebourbi, R., Vinatier, F., & Abdelfattah, R. (2025b). Correction of the jitter effect in Pléiades satellite elevation data for enhanced 3D change monitoring. In: Blanc-Talon, J., Delmas, P., Takahashi, H., Yasuhiro, M. (eds) Advanced Concepts for Intelligent Vision Systems. ACIVS 2025. To be published in *Lecture Notes in Computer Science* in October 2025.

[14] Berthier, E., Lebreton, J., Fontannaz, D., Hosford, S., Belart, J. M. C., Brun, F., ... & Blondel, C. (2024). The Pléiades Glacier Observatory: high-resolution digital elevation models and ortho-imagery to monitor glacier change. *The Cryosphere*, 18(12), 5551-5571.

Real-Time Non-Intrusive Load Monitoring on Raspberry Pi: An Edge-Cloud BERT4NILM Implementation

Imen NASR[1,2], Anis NSIR [1], Hichem, Besbes[1]

[1]University of Carthage, SUPCOM, COSIM Research Lab, Ariana, Tunisia
[2]University of Sousse, ESSTHS, Hammam Sousse, Tunisia
imen.nasr@supcom.tn, anis.nsir@supcom.tn, hichem.besbes@supcom.tn

Abstract

This project addresses the challenge of effective energy consumption management due to the lack of detailed, per-appliance visibility from traditional smart meters. Our solution leverages Non-Intrusive Load Monitoring (NILM) to disaggregate the total electrical consumption into individual appliance contributions. The core of the system is the BERT4NILM architecture, an innovative adaptation of the Transformer model that utilizes bidirectional attention to capture complex temporal dependencies in electrical time series. This architecture is deployed on a Raspberry Pi 4 Model B acting as an edge computing device to perform real-time inference. Crucially, the performance of BERT4NILM is benchmarked against established NILM baselines, including Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. The disaggregation results are communicated via a robust, low-latency pipeline using the MQTT protocol, stored in the InfluxDB time-series database, and presented to the user through a real-time dashboard on Grafana. Experimental validation, including cross-dataset fine-tuning (REFIT to REDD), confirmed the model's robustness and high performance, achieving an F1-Score of 55.54% on the REFIT dataset and demonstrating a significant improvement over the baselines. This edge-cloud architecture provides a practical, scalable, and energy-efficient solution for intelligent energy management.

Keywords: BERT4NILM, NILM, Raspberry Pi, Edge Computing, Grafana, IoT.

1 Introduction

The growing imperative for energy efficiency is often constrained by the ubiquitous smart meters, which typically provide only aggregated consumption data. This limitation prevents end-users from identifying the most energy-intensive appliances and optimizing their usage. Non-Intrusive Load Monitoring (NILM) offers a solution by computationally disaggregating the total load into individual appliance contributions [1].

1.1 Problem Definition

Traditional NILM methods, such as those based on Hidden Markov Models (HMM) or basic neural networks (CNN, RNN), often struggle to capture the complex, long-term temporal dependencies inherent in electrical load profiles, limiting their generalization capabilities [1]. There is a need for a model that is both highly accurate and deployable on resource-constrained, low-power embedded devices for a practical, real-time solution.

1.2 Contribution

This work's main contributions are:

- Implementation and optimization of the BERT4NILM architecture, a powerful Transformer-based model for NILM.
- A comprehensive comparative analysis of BERT4NILM against established deep learning NILM baselines (CNN and LSTM).
- Deployment of the real-time inference model on a low-power Raspberry Pi 4 Model B for local edge processing.
- Development of a complete, distributed Edge-Cloud pipeline using MQTT, InfluxDB, and Grafana to ensure real-time data flow and visualization.

2 Methodology

This section explains how the research was conducted.

2.1 BERT4NILM Architecture

The BERT4NILM (Bidirectional Encoder Representations from Transformers for Non-Intrusive Load Monitoring) model [2] is a revolutionary approach that adapts the bidirectional attention mechanisms of the Transformer architecture [4] to energy time series. In this context, energy measurement points are treated as ”temporal tokens,” and a time window of consumption is treated as an ”energy phrase”. This enables the model to leverage multi-headed attention to simultaneously analyze context from both the past and future of a given time step, significantly improving the accuracy of appliance state and power prediction over sequential or simpler neural network models. The model was pre-trained on a large dataset like REFIT and validated with a fine-tuning strategy on diverse domains like REDD.

2.2 Baseline Models for Comparison

To contextualize the performance of BERT4NILM, we implemented two widely adopted deep learning architectures for NILM as baselines:

- **Convolutional Neural Networks (CNN):** A one-dimensional CNN model is employed to extract local temporal features from the input power sequence. CNNs have shown strong performance in NILM by identifying distinct load shapes [5].
- **Long Short-Term Memory (LSTM):** A recurrent neural network (RNN) variant designed to capture long-term dependencies in time-series data. The LSTM baseline is used to assess how the recurrent sequential processing compares to the attention-based, bidirectional context modeling of BERT4NILM [5].

All models were trained and evaluated under the same conditions (datasets, window size, and evaluation metrics) to ensure a fair comparison.

2.3 Edge Computing Platform

The inference engine is deployed on a Raspberry Pi 4 Model B (4GB), chosen for its excellent performance/consumption ratio. The Broadcom BCM2711, Quad-core Cortex-A72 @ 1.5GHz processor provides sufficient power for PyTorch inference with low latency (average 247ms for a batch size of 1), meeting the real-time requirements of energy monitoring. The software stack on the edge device includes Raspberry Pi OS (64-bit), Python, PyTorch (CPU version), and the Mosquitto MQTT broker.

3 Real-Time Pipeline and Visualization

The system utilizes a distributed Edge-Cloud architecture to ensure optimal performance and scalability

3.1 Data Flow via MQTT

The Raspberry Pi, acting as the edge layer, performs the real-time BERT4NILM inference. The disaggregation results—including the timestamp, aggregate_power, predicted_power, appliance, and confidence score—are encapsulated in a JSON message. This data is published locally via the Mosquitto MQTT broker. The lightweight and efficient MQTT protocol [3] is central to the communication, enabling low-latency data transfer to the cloud server.

3.2 Storage and Presentation with Grafana

On the cloud layer, the Telegraf agent subscribes to the MQTT topics, validates and transforms the raw JSON messages, and writes the time-series data into InfluxDB. InfluxDB is used as the time-series database due to its high-performance architecture for ingesting and querying high-frequency temporal data.

Finally, Grafana serves as the primary presentation layer. It is configured to query the InfluxDB database and display the load disaggregation results in a real-time dashboard. The visualization allows users to monitor aggregated consumption, view per-equipment predictions, and observe the model's confidence scores.

3.3 Qualitative Analysis

Experimental evaluation confirmed the superior performance and robustness of the BERT4NILM model.

3.3.1 Comparative Performance Analysis

To strengthen our evaluation, we benchmarked the BERT4NILM model against the CNN and LSTM baselines on the REFIT dataset. The results, summarized in Table 1, highlight the superior performance of the attention-based approach.

Table 1: Comparison of Model Performance on the REFIT Dataset (F1-Score %)

Appliance	**CNN [5]**	**LSTM [5]**	**BERT4NILM**
Microwave	48.11	51.05	**56.78**
Kettle	72.50	75.31	**78.91**
Dishwasher	35.22	37.89	**41.45**
Fridge	58.90	60.12	**65.01**
Average	53.68	56.09	**60.54**

The results clearly indicate that BERT4NILM significantly outperforms both the CNN and LSTM baselines across all tested appliances. The average F1-Score of 60.54% for BERT4NILM is a notable improvement over the LSTM's 56.09% and the CNN's 53.68%, which validates the strength of the bidirectional attention mechanism in capturing complex, long-range dependencies crucial for accurate NILM.

3.3.2 Robustness and Transferability

Our optimized BERT4NILM architecture achieved a reference F1-Score of 55.54% (on all appliances). Moreover, the cross-dataset fine-tuning strategy proved effective, showing an improvement of +14.13% in accuracy when adapting the model from the REFIT training domain to the REDD target domain, validating the model's transferability to heterogeneous environments.

4 Conclusion

In conclusion, the project successfully implemented and validated the BERT4NILM model on a resource-constrained Raspberry Pi, establishing a robust and efficient edge-cloud pipeline using MQTT, InfluxDB, and Grafana. The comprehensive experimental evaluation demonstrated that the BERT4NILM architecture achieved superior disaggregation performance compared to conventional deep learning baselines (CNN and LSTM), validating the efficacy of the Transformer's bidirectional attention for NILM. This architecture provides a robust, low-power solution that delivers detailed, real-time, per-appliance energy data, opening new avenues for intelligent energy management systems.

References

[1] M. Zeifman and K. Roth. "Nonintrusive appliance load monitoring : Review and outlook." IEEE transactions on Consumer Electronics, vol. 57, no. 1, pp. 76–84, 2011.

[2] Z. Yue, C. R. Witzig, D. Jorde, and H.-A. Jacobsen. "Bert4nilm : A bidirectional transformer model for non-intrusive load monitoring." In Proceedings of the 5th International Workshop on Non-Intrusive Load Monitoring, pp. 89–93, 2020.

[3] R. Silva and P. Costa. "Mqtt for iot applications : Performance analysis and optimization." Computer Networks, vol. 186, pp. 107–120, 2020.

[4] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. "Attention is all you need." Advances in neural information processing systems, vol. 30, 2017.

[5] J. Kelly and W. Knottenbelt. "Neural nilm: Deep neural networks applied to energy disaggregation." In Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Buildings, pp. 55-64, 2015.

Intelligent Energy Monitoring: Optimization and Implementation of a NILM Algorithm on STM32

Imen Nasr[1,2], Nour Krichen [1], Hichem, Besbes[1]

[1]University of Carthage, SUPCOM, COSIM Research Lab, Ariana, Tunisia
[2]University of Sousse, ESSTHS, Hammam Sousse, Tunisia
`imen.nasr@supcom.tn, nour.krichen@supcom.tn, hichem.besbes@supcom.tn`

Abstract

The challenge of modern energy management lies in the lack of detailed appliance-level consumption data from conventional smart meters. This work proposes an optimized solution for Non-Intrusive Load Monitoring (NILM), which computationally disaggregates the total electrical load into individual appliance profiles. The core of the methodology is the Super-State Hidden Markov Model (SSHMM) combined with the Sparse Viterbi algorithm, specifically chosen to mitigate the exponential increase in computational complexity associated with the SSHMM in multi-appliance scenarios. The primary contribution is the successful migration and optimization of this sophisticated NILM prototype, initially designed for a PC, to a real-time embedded platform, the STM32MP157F-DK2 development board. Validation on the REFIT dataset demonstrates the efficiency of the overall pipeline, with the SSHMM achieving an F1-score exceeding 93% and an accuracy above 95% on specific appliances. The system successfully demonstrates real-time visualization of the disaggregation results on the ThingSpeak cloud platform. This project proves the feasibility of deploying advanced NILM techniques on resource-constrained microcontrollers for effective, on-device intelligent energy monitoring.

NILM, SSHMM, Viterbi, STM32, Embedded System, Load Disaggregation, Edge Computing

1 Introduction

The widespread adoption of smart metering provides consumers with global visibility into their energy use. However, without knowing the specific consumption of each electrical device, it remains challenging to implement effective strategies for energy reduction and optimization [1]. Non-Intrusive Load Monitoring (NILM) is a field dedicated to estimating the consumption of individual appliances from a single aggregate measurement point [2].

NILM applications are often hindered by several factors, including signal noise, signature similarity between devices, and high computational complexity [3]. These constraints are particularly severe when attempting to integrate existing NILM methods onto embedded platforms, such as microcontrollers, which are limited in both memory and processing power [4]. Addressing this requires an optimized algorithm capable of delivering high accuracy under stringent hardware limitations.

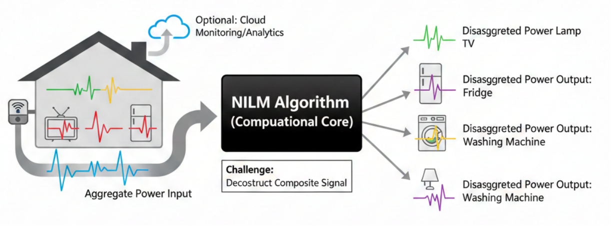


Figure 1: Conceptual Diagram of the NILM Challenge.

This work's main contributions are:

- Optimization and adaptation of the Super-State Hidden Markov Model (SSHMM) using the Sparse Viterbi algorithm to ensure computational efficiency.
- Demonstration of a real-time embedded solution by migrating the optimized SSHMM/Viterbi prototype from a PC environment to the STM32MP157F-DK2 card.
- Validation of the system's performance on a real-world dataset (REFIT) and integration of real-time visualization via the ThingSpeak platform.

2 Methodology: SSHMM and Sparse Viterbi

2.1 Super-State Hidden Markov Model (SSHMM)

The chosen approach for disaggregation is the SSHMM [5]. In this model, it is assumed that each appliance possesses a discrete number of operating states. The "super-state" corresponds to the cartesian product of all appliance states at a given instant, representing the household's global energy profile. The challenge with this model is that the number of super-states grows exponentially with the number of monitored appliances.

To formalize this complexity, let N be the total number of monitored appliances, and S_i be the number of discrete states for appliance i. The total number of super-states, S_{total}, is given by the formula:

$$S_{total} = \prod_{i=1}^{N} S_i$$

For instance, with $N = 5$ appliances, each having $S_i = 2$ (ON/OFF) states, the number of super-states is $2^5 = 32$. However, if N grows to 10 appliances, S_{total} explodes to $2^{10} = 1024$. This exponential growth clarifies the severity of the computational problem the Sparse Viterbi algorithm addresses.

2.2 Sparse Viterbi Optimization

To overcome the exponential complexity of the SSHMM, the system utilizes the Sparse Viterbi algorithm for decoding the optimal state sequence [6]. The Sparse Viterbi algorithm achieves its significant reduction in complexity by leveraging the inherent sparsity of the transition matrix in a real-world SSHMM. In typical usage scenarios, many super-state transitions are physically impossible (e.g., three appliances cannot all switch ON simultaneously between two consecutive time steps). This means the transition matrix contains many zero-value entries. The Sparse Viterbi algorithm avoids computing the probability of these impossible (zero-probability) transitions. By considering only the $\sim 2N$ valid transitions (i.e., those involving a single appliance change) instead of $S_{total} \times S_{total}$ potential transitions, the computational complexity of the Viterbi decoding is reduced from $O(S_{total}^2)$ to a complexity that is closer to linear in the number of appliances, $O(N \cdot S_{total})$, making the disaggregation process much more efficient in terms of calculation time and memory usage, which is critical for successful embedded deployment.

2.3 Event Detection Pipeline

The overall pipeline involves several preprocessing steps, followed by an event detection phase. This phase, which identifies transitions between ON/OFF states, is performed using the Cumulative Sum (CUSUM) standard algorithm [7]. Evaluation confirms the robustness of CUSUM compared to other methods, establishing it as the preferred choice for segmenting the power sequences and extracting relevant events prior to SSHMM modeling.

3 Implementation and Results

3.1 Embedded Platform and Deployment

The solution is deployed on the STM32MP157F-DK2 board, an advanced development kit, to achieve a real-time, on-device NILM solution. This is a crucial step towards Edge Computing for energy disaggregation. The deployment involves adapting the reference Python code for the embedded Linux environment on the STM32, ensuring performance meets the low-latency requirements of real-time monitoring.

3.2 Visualization

For user interaction and monitoring, the real-time disaggregation results (e.g., aggregate power, predicted power, and appliance states) are published to the ThingSpeak cloud platform. A ThingSpeak dashboard is configured to query the data and provides a clear, real-time view of the energy breakdown.

3.3 Experimental Results

The model is evaluated primarily on the REFIT dataset.

- **Event Detection:** The CUSUM standard algorithm demonstrates high reliability for event detection, achieving a precision of 0.87, a recall of 0.93, and an F1-score of 0.90.

- **Load Disaggregation:** The SSHMM combined with Sparse Viterbi yields highly satisfactory and consistent results. For specific household appliances (like the TV), the model achieves an F1-score of over 93% and an overall accuracy exceeding 95%. This performance validates the optimization and the feasibility of using this advanced NILM model on an embedded device.

4 Conclusion

This work successfully demonstrates the optimization and real-time implementation of a high-performance Non-Intrusive Load Monitoring (NILM) solution on a resource-constrained embedded platform, the STM32MP157F-DK2. The core innovation lies in adapting the Super-State Hidden Markov Model (SSHMM) and mitigating its exponential complexity by employing the Sparse Viterbi algorithm, whose complexity reduction is essential for on-device operation. The system is validated on the REFIT dataset, where it achieves high metrics, such as an F1-score exceeding 93% for specific loads, and integrates real-time results visualization via the ThingSpeak cloud platform. By effectively tackling the memory and processing constraints of microcontrollers, this project validates the possibility of shifting complex NILM tasks to the network edge, providing a practical, scalable, and cost-effective solution for intelligent energy management and detailed appliance-level monitoring. Future research will explore integrating advanced deep learning models within this optimized embedded framework for enhanced accuracy and handling of complex appliance signatures.

References

[1] M. Zeifman and K. Roth. ”Nonintrusive appliance load monitoring: Review and outlook.” *IEEE transactions on Consumer Electronics*, vol. 57, no. 1, pp. 76–84, 2011.

[2] G. W. Hart. ”Residential energy monitoring and service diagnostics.” *LCS TR-358*, 1985.

[3] P. A. Schirmer and I. Mporas, ”Non-Intrusive Load Monitoring: A Review,” in IEEE Transactions on Smart Grid, vol. 14, no. 1, pp. 769-784, Jan. 2023, doi: 10.1109/TSG.2022.3189598.

[4] S. Mari, G. Bucci, F. Ciancetta, E. Fiorucci and A. Fioravanti, ”An Embedded Deep Learning NILM System: A Year-Long Field Study in Real Houses,” in IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1-15, 2023.

[5] Z. Wu, C. Wang, W. Peng, W. Liu, H. Zhang, ”Non-intrusive load monitoring using factorial hidden markov model based on adaptive density peak clustering”, Energy and Buildings, Vol. 244, 2021,ISSN 0378-7788.

[6] S. Makonin, F. Popowich, I. Baji´c, B. Gill, L.Bartram and I. Bajic, ”Exploiting HMM Sparsity to Perform Online Real-Time Nonintrusive Load Monitoring (NILM)”,*IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 1765-1772, July 2016.

[7] V. Montes De Oca, D.R. Jeske, Q. Zhang, C. Rendon, M. Marvasti, ”A cusum change-point detection algorithm for non-stationary sequences with application to data network surveillance”, Journal of Systems and Software, Vol. 83, Issue 7, 2010, pp 1288-1297, ISSN 0164-1212,.

Comparative Analysis of Q-Learning and PSO for AoI Minimization in Hybrid NOMA/OMA Systems

Akram HAMDI[1], Ammar BOUALLEGUE[1], Sameh NAJEH[2]
[1]SysCom Research Laboratory, ENIT, University of Tunis El Manar
[2]COSIM Research Laboratory, SUP'COM, University of Carthage
`akram.hamdi@etudiant-enit.utm.tn`

Abstract

This paper presents a comparative study of Q-learning and Particle Swarm Optimization (PSO) for resource allocation in hybrid NOMA/OMA systems, focusing on Age of Information (AoI) minimization. We develop a unified framework where a base station dynamically switches between NOMA and OMA modes while optimizing power allocation. Through extensive simulations, we demonstrate that Q-learning achieves superior performance with an average weighted AoI of 2.60 slots and throughput of 13.31 bps/Hz, outperforming PSO which achieves 3.45 slots and 12.87 bps/Hz respectively. Q-learning's model-free approach and long-term optimization capabilities make it more suitable for dynamic environments, while PSO offers faster convergence in static scenarios.

Keywords: Age of Information, NOMA, PSO, Q-Learning, Resource Allocation

1 Introduction

1.1 Background

The proliferation of real-time applications in 5G-Advanced/6G networks has emphasized the importance of information freshness, quantified by the Age of Information (AoI) metric [1]. Hybrid NOMA/OMA systems combine the spectral efficiency of Non-Orthogonal Multiple Access with the reliability of Orthogonal Multiple Access, making them suitable for AoI-sensitive applications [2].

1.2 Problem Definition

Existing approaches often optimize AoI, throughput, and fairness independently. A comprehensive comparison of learning-based methods for joint optimization in hybrid NOMA/OMA systems remains unexplored. The challenge involves dynamic mode selection and power allocation under stochastic channel conditions.

1.3 Contributions

- Comparative analysis of Q-learning and PSO under identical conditions
- Multi-objective optimization framework for AoI and throughput
- Practical guidelines for algorithm selection

2 System Model

2.1 Age of information metric:

To evaluate the freshness of the information received by each client, we use the Age of Information (AoI) metric. AoI quantifies the time elapsed since the generation of the most recently received update at the destination. Formally, the AoI of client i at time t, denoted $\triangle_i(t)$, is defined as $t - u_i(t)$, where $u_i(t)$ is the generation time of the latest successfully received status update at time t. Under the *generate-at-will* model adopted at the BS, if client i successfully receives its update during a time slot, its AoI is reset to 1; otherwise, the AoI increments by 1 [3], [4]. This dynamic can be expressed mathematically as:

$$\triangle_i(t+1) = \begin{cases} \triangle_i(t) + 1, & \text{if } u_i(t) = 0 \\ 1, & \text{if } u_i(t) = 1 \end{cases} \tag{1}$$

where $u_i(t)$ is an indicator equal to 1 when client i receives its information from the BS in time slot t and $u_i(t)$= 0 otherwise. The weighted sum of the expected AoI of the two clients is adopted to measure the network-wide information timeliness, which is given by

$$\bar{\triangle} = \lim_{T\to\infty} \frac{1}{T} E[\sum_{i=1}^{2} \sum_{t=1}^{T} w_i \triangle_i(t)] \qquad (2)$$

where w_i is the weight coefficient for $user_i$ with the constraint $\sum w_i = 1$, and the expectation is calculated across all potential system dynamics.

2.2 OMA mode

In the OMA mode, we assume the base station (BS) allocates dedicated time slots to each user for individual transmissions. When a time slot is assigned to a user i, the observation y_i can be mathematically expressed as follows

$$y_i = h_i \sqrt{P} m_i + n_i, \qquad (3)$$

where:

- P is the BS's constant transmission power (in Watts)
- $m_i \in \mathbb{C}$ represents the status update message for user$_i$
- $h_i = \sqrt{d_i^{-\nu}} g_i$,ν denotes the path loss exponent. and
$$g_i \sim \mathcal{CN}(0,,1) \qquad (4)$$
- n_i is the complex additive Gaussian noise with variance σ_i^2.
We assume that $\sigma_i^2 = \sigma^2$ $\forall i$.

After receiving the signal, the information can be decoded in an interference free manner with a SNR [4]:

$$\gamma_i = |h_i|^2 \rho \qquad (5)$$

where

$$\rho = P/\sigma^2 \qquad (6)$$

The rate in OMA case is defined as

$$R_i^{OMA} = log_2(1 + \gamma_i) \qquad (7)$$

Proposition 1:
The outage probability relative to a user i, defined as

$$P_i^{OMA} = P(R_i^{OMA} < R_i) \qquad (8)$$

using OMA is given as follows

$$P_i^{OMA} = 1 - e^{-\frac{(2^{R_i}-1)d_i^{\nu}}{\rho}} \qquad (9)$$

where R_i represents the target rate relative to a user i.

2.3 NOMA mode

In the NOMA mode the observation at user i can be expressed as

$$y_i = h_i(\sqrt{\alpha_1 P} m_1 + \sqrt{\alpha_2 P} m_2) + n_i \qquad (10)$$

where,

- α_i represents the power allocation coefficient,
- α_1 + α_2 =1.

We assume that the BS possesses only stochastic Channel State Information (CSI) about its channels to both users, while the users, as receivers, have perfect knowledge of CSI.
Assumption: Consequently, we maintain $\alpha_1 < \alpha_2$ in accordance with the NOMA principle.
Then, for $user_2$, who is the far user, decodes its message from the BS directly by treating m_1 as interference. The received SINR γ_{22} relative to the $user_2$ is can be written as

$$\gamma_{22} = \alpha_2 |h_2|^2 / (\alpha_1 |h_2|^2 + 1/\rho) \qquad (11)$$

Proposition 2:
The probability of outage P_2^{NOMA} relative to the user 2 under NOMA technique, can be expressed as

$$\begin{aligned} P_2^{NOMA} &= 1 - P(log_2(1+\gamma_{22}) \geq R) \\ &= 1 - e(-\tfrac{(2^R-1)d_2^{\nu}}{\rho(\alpha_2-\alpha_1(2^R-1))}) \end{aligned} \qquad (12)$$

Proposition 3:
The outage probability P_1^{NOMA} relative to a $user_1$ in the NOMA system is given as below

$$\begin{aligned} P_1^{NOMA} &= 1 - P(log_2(1+\gamma_{22}) \geq R\ ,\ log_2(1+\gamma_{11}) \geq R) \\ &= 1 - e(-max\frac{(2^R-1)d_1^{\nu}}{\rho(\alpha_2-\alpha_1(2^R-1))}, \frac{(2^R-1)d_1^{\nu}}{\rho\alpha_1}) \end{aligned} \qquad (13)$$

2.4 Particle Swarm Optimization

Particles represent power allocation solutions, optimized for immediate reward maximization [5]. The PSO algorithm based on :

$$v_i^{k+1} = wv_i^k + c_1 r_1(pbest_i - x_i^k) + c_2 r_2(gbest - x_i^k) \qquad (14)$$

$$x_i^{k+1} = x_i^k + v_i^{k+1} \qquad (15)$$

Where :

- $V_i(t+1)$: The velocity of particle i at the next time step.
- $X_i(t+1)$position of particle i at the next time step.

- w: The inertia weight.
- $V_i(t)$: The velocity of particle i at the current time step.
- $X_i(t)$: The current position of particle i at the current time step.
- c_1, c_2: The cognitive learning factors.
- r_1, r_2: A random numbers between 0 and 1.
- $pbest_i(t)$: The personal best position of particle i at the current time step.
- $gbest(t)$: The global best position.

Algorithm 1 in [5] seeks to maximize the overall throughput by following an optimal policy that maintains information freshness.

2.5 Q-Learning

2.5.1 Fundamental Principles of Reinforcement Learning

Machine learning comprises a range of techniques that allow systems to learn from data, with supervised learning, unsupervised learning, and reinforcement learning (RL) being among the most widely used paradigms.

Supervised learning relies on labeled data, where each input is associated with a target output, allowing the model to learn a mapping function. Unsupervised learning, in contrast, operates on unlabeled data, aiming to discover hidden structures or intrinsic groupings within the data. Reinforcement learning (RL) distinguishes itself from these two approaches by focusing on learning through interaction with a dynamic environment.

2.5.2 Agent-Environment Interaction and Policy

The agent observes the state (s_t) of the environment and selects an action (a_t) according to a policy (π). This action leads to a transition to a new state (s_{t+1}) and generates a reward (r_t). The agent's objective is to learn an optimal policy π^* that maximizes the total expected reward over the long term.

2.5.3 The Bellman Equation

The Bellman equation is central to reinforcement learning. It recursively defines the optimal value of a state ($V^*(s_t)$) or a state-action pair ($Q^*(s_t, a)$). For optimal state values, it is expressed as:

$$V^*(s_t) = \max_{a_t} \left(r(s_t, a_t) + \gamma \sum_{s_{t+1}} P(s_t, a_t, s_{t+1}) V^*(s_{t+1}) \right) \tag{16}$$

where $r(s_t, a)$ is the expected immediate reward for taking action a_t in state s_t, $P(s_t, a_t, s_{t+1})$ is the probability of transitioning to the next state s_{t+1} after taking action a_t in the current state s_t, and γ is the discount factor that determines the importance of future rewards.

2.5.4 The Q-table and Update Rule

In Q-learning, a Q-table is used to estimate the Q-value ($Q(s_t, a_t)$), which represents the expected future reward for taking action a_t in state s_t. The Q-table is updated iteratively using the Q-learning update rule [6]:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left(r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \right) \tag{17}$$

where α is the learning rate that controls the magnitude of the update.

2.5.5 Exploration-Exploitation Trade-off and Convergence

The agent must manage an exploration-exploitation trade-off. Exploration allows for the discovery of new, potentially more rewarding actions, while exploitation involves using current knowledge to maximize immediate rewards. A common strategy to manage this trade-off is the ϵ-greedy approach, where the action a_t at time t is determined as follows:

$$a_t = \begin{cases} a \in \Pi \text{ with probability } Pr = \epsilon \\ \arg\max_a Q(s, a) \text{ with Probability } Pr = 1 - \epsilon \end{cases} \tag{18}$$

Over time and through numerous interactions with the environment, the Q-values converge towards the optimal Q-values ($Q^*(s, a)$), and the agent's policy tends towards an optimal policy (π^*) that maximizes the cumulative reward.

2.5.6 Multi-Objective Reinforcement Learning

While the preceding sections have focused on the single-objective reinforcement learning paradigm, many real-world problems naturally involve the simultaneous optimization of multiple, potentially conflicting objectives. Multi-Objective Reinforcement Learning (MORL) extends the standard RL framework to address such scenarios. In MORL, the agent receives a vector of rewards $\mathbf{r}_t = [r_{t,1}, r_{t,2}, ..., r_{t,m}]^T$ at each time step, where

m represents the number of distinct objectives. In our study, we specifically focus on optimizing two key metrics: maximizing throughput (to ensure efficient data transmission) while simultaneously minimizing the Age of Information (AoI) (to guarantee the freshness of status updates). These objectives often conflict, as aggressive throughput maximization may lead to increased AoI due to network congestion or resource contention, making MORL particularly suited for this trade-off analysis.

2.5.7 Optimization Problem

The multi-objective optimization problem is:

$$\begin{aligned} \underset{\alpha_1(t),\alpha_2(t)}{\text{Maximize}} \quad & \left[R^{\text{Total}}, -\mathbb{E}[r(\hat{s},a)]\right] \\ \text{subject to} \quad & \text{C1: } \sum_{i=1}^{2} p_{i,t} \leq P \\ & \text{C2: } R_{i,t} \geq R_{\min}, \quad \forall i \in \{1,2\} \\ & \text{C3: } \alpha_2(t) \geq \frac{2^{R_{\min}}-1}{2^{R_{\min}}} \end{aligned} \tag{19}$$

where the total rates are:

- **NOMA:**
$$R^{\text{Total}} = \log_2\left(1+\frac{\alpha_1 P g_1}{\alpha_2 P g_1 + N_0}\right) + \log_2\left(1+\frac{\alpha_2 P g_2}{N_0}\right) \tag{20}$$
- **OMA:**
$$R^{\text{Total}} = \log_2\left(1+|h_i|^2\rho\right) \tag{21}$$
- **C1:** represents that the total transmit power of base stations (BS) at each time slot t should not exceed P.
- **C2:** The second constraint ensures that the rate $R_{i,t}$ for each user i at time t meets or exceeds the minimum required rate $R_{\min}$.
- **C3:** Ensures the feasibility of successive interference cancellation (SIC) in NOMA mode. This condition guarantees that $user_1$ can successfully decode $user_2$'s message before applying SIC, as established in [7,8].

3 Results and Discussion

3.1 Performance Comparison

Q-learning outperformed PSO in AoI minimization (2.60 vs 3.45 slots) while maintaining higher throughput (13.31 vs 12.87 bps/Hz).

3.2 Mode Selection Analysis

Q-learning demonstrated intelligent mode selection:

- OMA for user 1: 51%
- OMA for user 2: 18%
- NOMA: 31%

PSO showed more uniform distribution but lower overall performance.

3.3 Computational Efficiency

PSO converged faster (45 iterations vs 120 episodes) but Q-learning achieved better final performance, highlighting the trade-off between convergence speed and solution quality.

4 Conclusion

Q-learning is recommended for dynamic environments requiring long-term AoI optimization, while PSO suits scenarios needing quick deployment in stable conditions. The hybrid NOMA/OMA framework effectively balances spectral efficiency and information freshness.

References

[1] Syed Amaar Ahmad, Abolfazl Hajisami, Hariharan Krishnan, Farid Ahmed-Zaid, and Ehsan Moradi-Pari. V2v system congestion control validation and performance. *IEEE Transactions on Vehicular Technology*, 68(3):2102–2110, 2019.

[2] Akram Hamdi, Ammar Bouallegue, and Sameh Najeh. Fairness for non-orthogonal multiple access based-superposition coding. In *2023 22nd Mediterranean Microwave Symposium (MMS)*, pages 1–6, 2023.

[3] Qian Wang, He Chen, Changhong Zhao, Yonghui Li, Petar Popovski, and Branka Vucetic. Optimizing information freshness via multiuser scheduling with adaptive noma/oma. *IEEE Transactions on Wireless Communications*, 21(3):1766–1778, 2022.

[4] Qian Wang, He Chen, Yonghui Li, and Branka Vucetic. Minimizing age of information via hybrid noma/oma. In *2020 IEEE International Symposium on Information Theory (ISIT)*, pages 1753–1758, 2020.

[5] Akram Hamdi, Ammar Bouallegue, and Sameh Najeh. Age of information in hybrid noma/oma systems based on particle swarm optimization. In Leonard Barolli, editor, *Advanced Information Networking and Applications*, pages 72–85, Cham, 2025. Springer Nature Switzerland.

[6] Alex M Andrew. Reinforcement learning:: An introduction. *Kybernetes*, 27(9):1093–1096, 1998.

[7] Zhiguo Ding, Zheng Yang, Pingzhi Fan, and H Vincent Poor. On the performance of non-orthogonal multiple access in 5g systems with randomly deployed users. *IEEE Signal Processing Letters*, 21(12):1501–1505, 2014.

[8] Stelios Timotheou and Ioannis Krikidis. Fairness for non-orthogonal multiple access in 5g systems. *IEEE Signal Processing Letters*, 22(10):1647–1651, 2015.

Energy Efficiency Optimization in Heterogeneous Networks Using GA- and PSO-Based Switching Mechanisms

Khaled Hmidi[1], Sameh Najeh[1], Ammar Bouallegue[2]

[1]University of Carthage, Sup'Com, LR11TIC01, Research Lab. COSIM, Technopark El Gazala , 2088, Ariana, Tunisia
[2]University of Tunis El Manar, ENIT, LR-99-ES21, Research Lab. SysCom, Belvédère , 1002, Tunis, Tunisia .
`khaled.hmidi@isimg.tn`

December 12, 2025

Abstract

In 5G/B5G Heterogeneous Networks (HetNets), optimizing power allocation and base station (BS) ON-OFF switching is crucial for improving Energy Efficiency (EE) [1]. This paper presents a novel dynamic switching technique for Small BSs, aiming to minimize total power consumption (PC) while serving user traffic. Given that BSs account for 70–80% of network energy use [2], the proposed method employs Artificial Intelligence (AI)–specifically Genetic Algorithms (GA) and Particle Swarm Optimization (PSO)–to develop an optimal cell switch-off scheme. Simulations show the scheme enhances scalability and efficiency, achieving a competitive EE of 7.43 bit/J.

Keywords: HETNETs, Power Allocation (PA), Energy Efficiency (EE), ON-OFF Switching, Genetic Algorithm (GA), Particle Swarm Optimization (PSO).

1 Introduction

1.1 Background

To address the energy demands of increasing 5G traffic and promote environmental sustainability, this research employs dynamic ON-OFF switching of BS. The strategy deactivates idle cells, especially Small BSs (SBSs), during low-traffic periods, forcing them into sleep mode to achieve significant energy savings proportional to the number of cells switched off.

1.2 Problem Definition

The objective is to minimize total power consumption while ensuring uninterrupted service. As detailed in Table 1, key challenges include accurately estimating traffic load to avoid service degradation, managing disruptive user handovers, and preventing coverage holes. This makes the EE optimization a complex, nonlinear, and multidimensional problem [3]. Quality of service is maintained by reserving capacity above a threshold for prioritized traffic, ensuring low latency and high reliability. Users associate with BSs based on the strongest received signal strength and available capacity, addressing the coverage and capacity trade-off highlighted in the table. If a BS nears its capacity threshold, new users may be handed off to neighboring BSs to preserve QoS for existing connections, directly managing the handover challenge identified.

1.3 Contributions

Small cells are essential to 5G HetNets for enhancing network capacity and coverage, particularly in dense urban areas, by offloading traffic from macrocells and enabling efficient spectrum reuse [4]. This dense deployment supports stringent 5G requirements for high data rates and ultra-low latency, forming a foundational layer for ubiquitous connectivity [5].

This study enhances EE in HetNets by applying AI tools (GA and PSO) to dynamically manage small cell switching. A key contribution is a novel, low-complexity distributed algorithm that simplifies the BS selection problem. Simulations show that the proposed approach significantly reduces energy consumption and confirms its scalability for large-scale network deployments. The algorithms employ a threshold to decide when BS can be switched on or off, ensuring the required QoS.

Table 1: Key Issues in Base Station ON-OFF Switching Optimization

Issue	Description
Traffic Load Estimation	Accurate prediction of user demand to determine when BS deactivation is justified, avoiding both under-provisioning (service degradation) and over-provisioning (wasted energy).
Handover Management	Minimization of disruptive and unnecessary user handovers during BS switch-off events to maintain seamless connectivity and service continuity.
Coverage Hole Prevention	Ensuring no areas of insufficient signal strength are created when BSs are deactivated, maintaining adequate coverage for all active users.
QoS Guarantee	Maintaining required QOS metrics (data rate, latency, reliability) for all users following network reconfiguration.
Algorithm Complexity	Addressing the nonlinear, combinatorial optimization problem with computationally efficient methods suitable for real-time or near-real-time operation in large-scale networks.
Scalability	Ensuring the switching solution remains effective as network size increases, handling potentially hundreds of BSs and thousands of users.

2 Methodology

2.1 Theoretical Framework

The system is modeled as a hierarchical HetNet comprising a Macro BS (MBS, $j = 0$), assumed to be always active, and J SBSs ($j \in \{1, ..., J\}$), which can be dynamically activated or deactivated. User connectivity to BS j is defined by the binary variable $x_{k,j}$, and the SBS ON-OFF state is defined by y_j.

The optimization problem seeks to maximize EE (μ), defined as the ratio of the total system throughput (C) to the total PC. The total system throughput C is derived from the fundamental Shannon capacity theorem, which provides the upper bound for the achievable data rate on a communication link under additive white Gaussian noise (AWGN). For a given user k connected to BS j, the throughput $C_{k,j}$ depends on the achieved Signal-to-Interference-plus-Noise Ratio (SINR). The SINR, denoted $\gamma_{k,j}$, for a user k served by BS j is defined as the ratio of the received power from its serving BS to the sum of the received interference power from all other active BSs and the thermal noise. This is expressed as:

$$\gamma_{k,j} = \frac{P_j \, h_{k,j}}{\sum_{\substack{i=0 \\ i \neq j}}^{J} y_i \, P_i \, h_{k,i} + \sigma^2}, \tag{1}$$

where P_j is the transmit power of BS j, $h_{k,j}$ is the channel power gain between user k and BS j, σ^2 is the variance of the AWGN, and the term $\sum_{i \neq j} y_i P_i h_{k,i}$ represents the aggregate inter-cell interference. The bandwidth available for transmission is denoted by B. The achievable throughput $C_{k,j}$ is therefore given by:

$$C_{k,j} = B \, \log_2 \left(1 + \gamma_{k,j}(X, Y, P)\right). \tag{2}$$

Consequently, the total network throughput C, which is the sum of the individual throughputs of all users, is defined as:

$$C = \sum_{j=0}^{J} \sum_{k=1}^{K} x_{k,j} \, C_{k,j}, \tag{3}$$

where K is the total number of users in the system.

The optimization problem seeks to maximize the EE (μ), defined as the ratio of the total system throughput (C) to the total PC:

$$\mu = \max_{X,Y,P} \frac{C}{P_0 + \sum_{j=1}^{J} y_j P_j} \quad , \tag{4}$$

The constraints ensure that each user connects to only one BS, i.e., $\sum_{j=0}^{J} x_{k,j} = 1, \forall j$, and that users only connect to an SBS j if it is active, formally expressed as $x_{k,j} \leq y_j, \forall k, \forall j$. P_0 is the transmit power of the MBS, bounded by its maximum value $P_{\max}$. Similarly, P_j is the transmit power of SBS j, limited by its maximum power $P'_{\max}$.

2.2 Datasets and Tools

The system was evaluated using a Python simulation. The setup featured one MBS and three SBSs serving 29 total users. Key parameters included a 1 MHz bandwidth, with maximum transmit powers of 12 W for the MBS and 1 W for each SBS.

2.3 Algorithm Description

Two AI algorithms were implemented for dynamic SBS switching. The **Genetic Algorithm (GA)** evolves a population of binary-coded solutions, where each bit represents an SBS's ON/OFF state. Solutions are evaluated by a fitness function measuring EE and improved via crossover and mutation.

The **Particle Swarm Optimization (PSO)** uses a swarm of particles, each representing a potential SBS configuration. The algorithm iteratively refines solutions by updating each particle's position based on its own best state and the global best state found by the swarm, guided by an EE objective function.

3 Results and Discussion

The proposed GA and PSO algorithms were evaluated across four parameter scenarios and benchmarked against the established JOS scheme [6], with user association initially based on proximity and balanced via an MBS offloading mechanism. The quantitative results revealed significant performance differences. The GA achieved a superior EE of 7.43 bit/J, substantially outperforming both PSO at 6.58 bit/J and the JOS scheme at 4.2 bit/J (Fig. 1). Furthermore, GA proved computationally more efficient, converging to its optimal solution in just 23 generations, which was twice as fast as PSO's 44 generations. However, scalability tests in larger networks using 128-bit encoding showed that PSO's exploration capability gave it a slight EE advantage over GA, indicating its potential for more complex, dense network deployments. Qualitatively, the steady fitness improvement in GA confirms its effective evolution toward a solution that balances SBS deactivation with QoS requirements. Performance is influenced by factors like user placement and association strategies, and while PSO excels in dense networks, this comes with increased computational demands (Fig. 2).

GA demonstrate superior performance in small, dense networks due to their robust exploration capabilities and disruptive search operators. The crossover mechanism promotes a broad exploration of the complex, constrained solution space, effectively escaping local optima prevalent in high-interference environments. Conversely, PSO tends toward premature convergence as particles rapidly cluster around suboptimal attractors, limiting its search efficacy. In contrast, within ultra-dense network deployments, the search landscape and scalability challenges alter this dynamic. The enhanced particle diversity and gradient-like information flow in PSO can mitigate premature convergence, allowing it to navigate the expanded solution space more efficiently. Consequently, PSO can achieve superior convergence speed and identify global optima with greater reliability than the more computationally exploratory GA in such high-density scenarios.

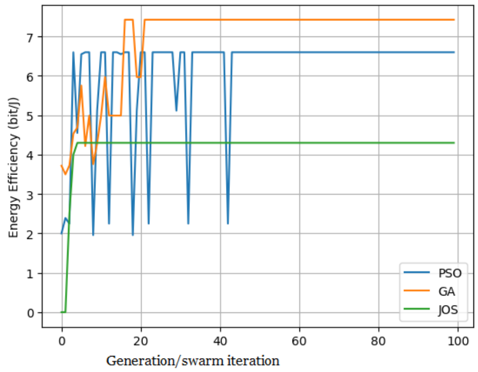


Figure 1: Average Cell Energy Efficiency (32-bit Encoding)

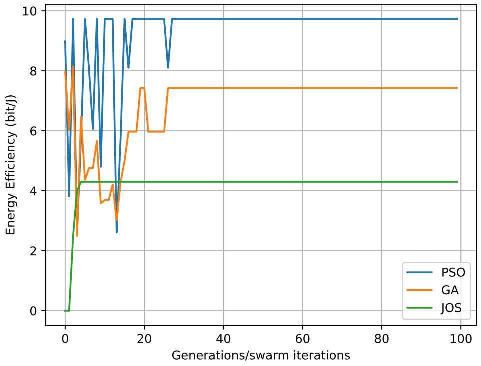


Figure 2: Average Cell Energy Efficiency (128-bit Encoding)

4 Conclusion

This study confirms that AI-driven dynamic SBS switching significantly enhances EE in HetNets. Both GA and PSO effectively reduce PC while maintaining QoS, with GA performing best in the tested scenarios; notably, GA converges faster (23 generations vs. PSO's 44), which helps explain its superior performance in scenarios with fewer SBSs. Future work will focus on integrating advanced AI, such as Deep Reinforcement Learning, with real-time network data to enhance adaptability. furthermore, investigation into cell-free architectures can address scalability for 6G systems. Finally, cross-layer optimization incorporating network slicing cab be also explored for green networking.

References

[1] Natarajan, J., Rebekka, B. (2023). Machine learning based small cell ON/OFF for energy efficiency in 5G heterogeneous networks. Wireless Personal Communications, 130(4), 2367-2383

[2] Chochliouros, I. P., Kourtis, M. A., Spiliopoulou, A. S., Lazaridis, P., Zaharis, Z., Zarakovitis, C., Kourtis, A. (2021). Energy efficiency concerns and trends in future 5G network infrastructures. Energies, 14(17), 5392.

[3] Chen, Y., Xun, L., Zhang, S. (2023). The energy efficiency of heterogeneous cellular networks based on the poisson hole process. Future Internet, 15(2), 56.

[4] Kamel, M., Hamouda, W., Youssef, A. (2016). Ultra-dense networks: A survey. IEEE Communications surveys tutorials, 18(4), 2522-2545.

[5] Andrews, J. G., Buzzi, S., Choi, W., Hanly, S. V., Lozano, A., Soong, A. C., Zhang, J. C. (2014). What will 5G be?. IEEE Journal on selected areas in communications, 32(6), 1065-1082.

[6] Mensah, K. K., Chai, R., Bilibashi, D., Gao, F. (2016). Energy efficiency based joint cell selection and power allocation scheme for HetNets. Digital Communications and Networks, 2(4), 184-190.

Sector-based BEV Fusion in Limited Communication Conditions for Optimized Collaborative Perception

Eya Besbes[1], Hakim Ghazzai[2], Gianluca Setti[2]

[1] Polytechnic School of Tunisia, University of Carthage, Tunis, Tunisia
[2]King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
`eya.besbes@etudiant-fst.utm.tn`

Abstract

Collaborative perception is a fundamental aspect of autonomous driving enabling connected autonomous vehicles (CAVs) to share sensory data and enhance their environmental awareness. This collaborative approach is especially valuable in complex scenarios, such as the detection of occluded objects at intersections. In such cases, limited visibility can compromise safety and hinder real-time decision-making. Various methodologies have been proposed for integrating multi-vehicle data, among which, early fusion of raw LiDAR data has proven particularly effective in achieving high-quality perception. However, despite its robustness, early fusion imposes substantial communication and computational demands. To overcome this limitation, we propose an optimized framework where each CAV generates its own Bird's Eye View (BEV) map and divides it into sectors. A roadside unit (RSU) evaluates then the quality and size of these sectors and selectively instructs vehicles to share only the most valuable ones. This selective sharing strategy minimizes redundant transmissions while preserving accurate perception.

Keywords: Autonomous vehicle, collaborative perception, LiDAR, BEV, V2V communication.

1 Introduction

While autonomous driving relies on robust environmental perception, sensors such as LiDAR remain limited by occlusions, hardware constraints, and dynamic environments [1,2]. Collaborative perception mitigates these issues through real-time sensor data sharing, expanding the collective field of view and improving risk anticipation, which is further enhanced by Bird's Eye View (BEV) maps that fuse and compress multi-sensor data into occlusion-resilient representations for reliable situational awareness [4,5].

1.1 Background

Sensor fusion and AI are critical for autonomous systems, and intelligent transportation [3]. Sensors such as LiDAR, radar, and cameras generate voluminous data that require real-time processing. Individual perception is constrained by occlusions and traffic complexity. Collaborative perception enables CAVs to share data, improving awareness and safety [3]. Common fusion strategies include early fusion of raw data [6], intermediate feature-level fusion [7], and late fusion of independent outputs [8].

1.2 Problem Definition

Although multi-vehicle early fusion ensures high-quality perception, it imposes substantial communication and computational overhead, especially in urban intersections [6], while intermediate and late fusion reduce data load at the cost of accuracy [7,8], highlighting the need for a framework that maintains accurate detection with minimal communication.

1.3 Contribution

This paper:

- Proposes an optimized collaborative perception framework for CAVs at urban intersections.
- Uses sectorized LiDAR-derived BEV maps with RSU-based selective sharing to reduce redundant transmissions.
- Applies BEV fusion to mitigate occlusions and enhance environmental awareness.
- Evaluates performance on a dataset generated with CARLA and OpenCDA, showing near-full collaborative accuracy with lower data exchange.

2 Methodology

2.1 Theoretical Framework

The proposed algorithm employs an Integer Linear Programming (ILP) formulation to optimize collaborative perception among connected vehicles. Each vehicle divides its BEV map into sectors, each characterized by data size and quality score. These metrics are sent to a Roadside Unit (RSU), which solves the ILP to select the optimal combination of sectors to share between vehicles. The objective is to maximize the overall perception quality under a bandwidth constraint, while a distance-based weighting promotes data sharing among distant vehicles, reducing redundant information exchange.

2.2 Datasets and Tools

The dataset is generated using the CARLA simulator integrated with the OpenCDA framework to model cooperative

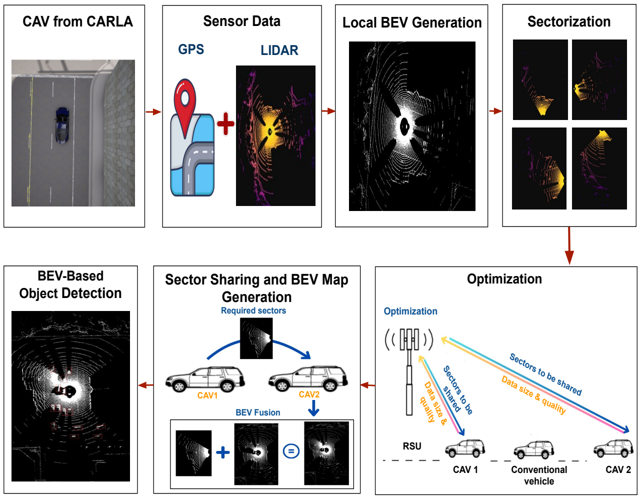


Figure 1: Overview of the proposed architecture for sectorized collaborative perception using early fusion technique.

driving in a four-leg intersection with four connected autonomous vehicles (CAVs) and random 20–40 non-connected vehicles. Each CAV employs a 64-channel LiDAR sensor capturing 3D point clouds within a 120 m range at 1.3 million points per second. The simulation setup requires a 64-bit Ubuntu system (16.04/18.04/20.04 recommended), a GPU with at least 3 GB of memory (8 GB recommended for Unreal Engine 4), 100 GB of disk space, and Python 3.7 or higher.

2.3 Algorithm Description

The framework is structured into five key steps (Fig. 1):

1. **Local BEV Generation:** Encode LiDAR data to a 2D BEV map.
2. **Sectorization:** Divide BEV into S sectors; compute $\sigma_{i,s}$ and $\tilde{S}_{i,s}$.
3. **Optimization:** Select high-quality sectors under budget $\bar{B}$.
4. **Sector Sharing & Fusion:** Share sectors via V2V; fuse locally using max-value strategy.
5. **Object Detection:** Detect obstacles using COOPER [6] with PIXOR on fused BEV maps.

3 Results and Discussion

We evaluate our optimized collaborative perception framework in 57 four-leg intersection scenarios simulated with CARLA and OpenCDA. Each scenario includes 4 CAVs and 20–40 non-CAVs, equipped with 64-channel LiDAR sensors and V2V/RSU connectivity. We compare five strategies: no collaboration, full collaboration, and optimized sharing at 25%, 50%, and 75% of $\bar{B}$ with $S = 4$ and $S = 8$ sectors. Performance is measured using AP@0.5, the number of detected objects, and the in-memory transmitted data volume.

3.1 Quantitative Results

The results show that full collaboration achieves an AP@0.5 of 0.87 and the highest number of detected objects, while no collaboration scores 0.69 with the fewest detections. Optimized strategies approach full performance; for example, using 4 sectors at 75% $\bar{B}$ reaches 0.83. Data transmission is greatly reduced, with optimized sharing requiring only 210.51 KB versus 363.42 KB for full collaboration.

3.2 Qualitative Analysis

The results show that sector-based optimization preserves perception accuracy while greatly reducing communication overhead. Lower thresholds or more granular sectorization slightly reduce AP, highlighting the trade-off between data reduction and sector granularity. Careful tuning of $\bar{B}$ and the sectorization enables near-full collaborative performance efficiently, demonstrating the framework's scalability and effectiveness compared to existing approaches [5, 6].

4 Conclusion

This paper presents an optimized collaborative perception framework for connected autonomous vehicles at urban intersections, where sectorized LiDAR-derived BEV maps reduce data transmission while approaching near-full collaboration, demonstrating the potential for efficient and scalable CAV deployment. Limitations include reliance on accurate sector quality estimation, and future work will explore advanced optimization and dynamic sectorization for diverse traffic scenarios to further enhance smart vehicle collaboration.

References

[1] Hu, S. et al., "Collaborative Perception for Connected and Autonomous Driving," *arXiv:2401.01544*, 2024.

[2] Zheng, Z. et al., "CooperFuse: A Real-Time Cooperative Perception Fusion Framework," *2024 IEEE Intelligent Vehicles Symposium (IV)*, pp. 533–538, 2024.

[3] Xiang, C. et al., "Multi-Sensor Fusion and Cooperative Perception for Autonomous Driving: A Review," *IEEE Intell. Transp. Syst. Mag.*, pp. 36–58, 2023.

[4] Zhao, J., Shi, J., and Zhuo, L., "BEV Perception for Autonomous Driving: State of the Art and Future Perspectives," *Expert Syst. Appl.*, 2024.

[5] Xu, R. et al., "OPV2V: An Open Benchmark Dataset and Fusion Pipeline for Perception with Vehicle-to-Vehicle Communication," *2022 IEEE Int. Conf. Robotics Autom. (ICRA)*, 2022.

[6] Chen, Q. et al., "Cooper: Cooperative Perception for Connected Autonomous Vehicles Based on 3D Point Clouds," *2019 IEEE 39th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, pp. 514–524, 2019.

[7] Wang, T. et al., "V2VNet: Vehicle-to-Vehicle Communication for Joint Perception and Prediction," *Comp. Vis. – ECCV 2020*, pp. 605–621, 2020.

[8] Xu, R. et al., "Model-Agnostic Multi-Agent Perception Framework," *2023 IEEE Int. Conf. Robotics Autom. (ICRA)*, pp. 1471–1478, 2023.

Design of an IoT Architecture for Data Acquisition and Non-Intrusive Load Monitoring

Fathi Chelly[1], Lotfi Charaabi[1], Sejir Khojet El Khil[1], Sofiane Sayahi[2]

[1]National Engineering School of Tunis, University of Tunis El Manar, Tunis, Tunisia
[2]ACTIA Engineering Services, Tunis, Tunisia
`fathi.chellyy@email.com; lotfi.charaabi@gmail.com; sejirkek@gmail.com; sofiane.sayahi@actia-engineering.tn`

Abstract

Non-intrusive load monitoring (NILM) plays a key role in enhancing smart grid technologies, since it addresses the estimation of an appliance's power usage based on whole-house power measurements. To ensure high accuracy of NILM algorithms, real-world smart meter data collection is mandatory. In that sense, this work presents an IoT platform that builds a Tunisian dataset from smart meter data collected and stored locally and in the cloud. A method combining Factorial Hidden Markov Models (FHMM) and Adaptive Density Peak Clustering (ADPC) enables automatic appliance-level disaggregation without manual configuration. Experiments achieved ∼96% accuracy, demonstrating the system's potential for automatic real-time power disaggregation.

Keywords: Energy Monitoring, IoT, Non-Intrusive Load Monitoring, Factorial Hidden Markov Models, Adaptive Density Peak Clustering

1 Introduction

Efficient energy management is increasingly important as electricity demand and environmental concerns grow. NILM addresses this need by estimating individual appliance consumption from aggregate power data, improving energy optimization and user awareness [1]. Recent work has combined FHMM with ADPC to enhance disaggregation performance [2]. However, the absence of Tunisian-specific datasets reflecting local grid conditions and usage patterns limits the applicability of such NILM models. Current systems also face challenges with real-time processing and often require manual configuration. To address these limitations, this work proposes an IoT platform for real-time energy monitoring and the generation of a Tunisian dataset, along with an adapted FHMM–ADPC method that enables automatic appliance-level disaggregation.

2 Methodology

2.1 System Architecture

The system consists of a smart meter that records a household's electrical parameters at a 1Hz resolution. This data is then trasmitted via UART to an ESP32 microcontroller. The latter publishes it using MQTT to a Node red server responsible for the dashboard interface and dual data storage (SQLite and Firebase) as shown in Fig. 1. The ESP32 is also equipped with an SD card for local storage in case of connectivity loss, and once the connection is restored, the stored data is sent in a burst.

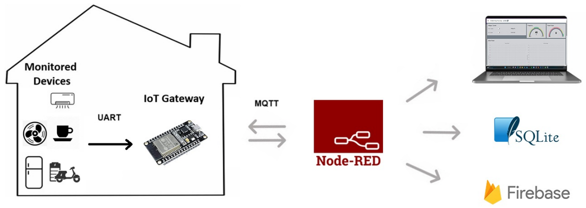


Figure 1: System architecture.

2.2 Theoretical Framework

The aggregate power is the sum of individual appliance consumptions. Each appliance is modeled as a Hidden Markov Model (HMM) with multiple states, each following a Gaussian emission. Let $P_i(t)$ denote the active power of appliance i, $s_i(t)$ its current state, and $(\mu_i^{s_i(t)}, \sigma_i^{s_i(t)})$ the Gaussian parameters:

$$P_{\text{total}}(t) = \sum_{i=1}^{N} P_i(t), \quad P_i(t) \sim \mathcal{N}\left(\mu_i^{s_i(t)}, \sigma_i^{s_i(t)}\right) \quad (1)$$

ADPC marks a power level p as a state if $d_p > d_{TD}$ and $q_p > q_{TD}$, where q_p is its local density and d_p its minimum distance to higher-density levels. In [2], $d_{TD} = 0.1 \max(P_{\text{appliance}})$, but this single threshold often merges low-power states, reducing accuracy. To avoid this, we use two thresholds:

$$d_{TD}^{\text{high}} = 0.1 \max(P_{\text{appliance}}), \qquad d_{TD}^{\text{low}} = 35,$$

applying d_{TD}^{low} for p values below 60 W and d_{TD}^{high} for ones above. The 60 W split is chosen heuristically.

ADPC-FHMM operates in 3 phases. First, ADPC is applied to identify appliances' states. These are then used as parameters for training the FHMM model with the Baum–Welch algorithm. Finally, the trained FHMM is evaluated with the Viterbi algorithm, which computes the most likely sequence of hidden appliance states given the observed power.

2.3 Experimental Setup

The monitoring system recorded data from four appliances (fridge, 100 W lamp, fan, and scooter battery) in the Smart House at LSE-ENIT to build the dataset. However, since the latter has been under construction and not publicly available, we applied ADPC-FHMM to 30,000 measurements from September 1, 2023, covering five appliances from Plegma [3], which contains one year of data from 13 Greek households sampled at 10-second intervals and was chosen for its similarity to the Tunisian context. Evaluation was performed on a separate 30,000-point sample starting at timestamp 39,000.

2.4 Quantitative Results

37,568 readings were collected. Figure 2, for example, shows the data corresponding to a 4-hour charging cycle of an electric scooter battery. As for the model, it reached ~96% accuracy acroos the 5 appliances as shown in figure 3:

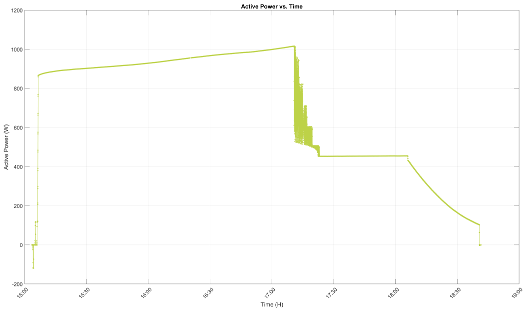


Figure 2: Electric scooter battery charging cycle

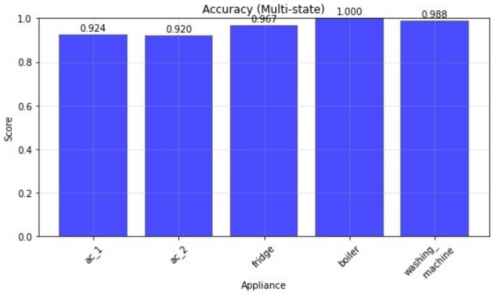


Figure 3: Model accuracy for each of the 5 appliances

2.5 Qualitative Analysis

The ADPC–FHMM approach achieved high accuracy with fully automated disaggregation. Our threshold formulation outperformed [2], whose original thresholds reached only 89% under the same conditions. Limitations lies with overlapping appliance signatures and generalization to diverse datasets.

3 Conclusion

This work presented a data acquisition system capable of building a Tunisian energy dataset, and an ADPC–FHMM model that achieves approximately 96% accuracy in fully automated appliance-level disaggregation. Remaining challenges include appliance overlaps and limited generalization. Future work will focus on parameter optimization, broader dataset evaluation, and real-time deployment.

Acknowledgment

We thank ACTIA Engineering Services and LSE/ENIT for supporting the work.

References

[1] Y. Himeur *et al.*, "Building power consumption datasets: Survey, taxonomy and future directions," *Energy and Buildings*, vol. 227, p. 110404, 2020.

[2] Z. Wu *et al.*, "Non-intrusive load monitoring using factorial hidden Markov model based on adaptive density peak clustering," *Energy and Buildings*, vol. 244, 2021.

[3] S. Athanasoulias *et al.*, "The Plegma dataset: Domestic appliance-level and aggregate electricity demand from Greece," *Scientific Data*, vol. 11, p. 376, 2024.

Aerial and RIS-assisted Edge Computing for Intelligent Transportation Services

Insaf Rzig[1], Wael Jaafar[2], Maha Jebalia[1], Sami Tabbane[1]

[1]MEDIATRON Laboratory, École Supérieure des Communications (Sup'Com), University of Carthage, Tunisia
[2]Department of Software and IT Engineering, École de technologie supérieure (ÉTS), University of Quebec, Canada
{insaf.rzig, maha.jebalia, sami.tabbane}@supcom.tn, wael.jaafar@etsmtl.ca

Abstract

Intelligent Transportation Systems (ITS) require ultra-low-latency and energy-efficient computing to support autonomous vehicles' telematic tasks, such as real-time navigation, traffic monitoring, and vehicular data processing. We propose an aerial edge computing framework using a High-Altitude Platform Station (HAPS) and Reconfigurable Intelligent Surface (RIS)-equipped Unmanned Aerial Vehicles (UAVs) to enhance wireless connectivity and computational resources for vehicular task offloading. RIS-integrated UAVs intelligently shape signal propagation to improve uplink reliability, while the HAPS enables wide-area coordination. Our method jointly optimizes task offloading and RIS configuration to minimize delay and energy consumption for next-generation ITS.

Keywords: UAV, HAPS, RIS, MEC, ITS.

1 Introduction

The advent of truly autonomous and connected vehicles is fundamentally constrained by the limitations of terrestrial networks, which struggle to provide the ultra-low-latency, high-reliability, and energy-efficient computing required for safety-critical applications. To overcome these barriers, our research introduces a novel aerial edge computing architecture built upon a Non-Terrestrial Network (NTN). This framework synergistically combines a High-Altitude Platform Station (HAPS) for wide-area orchestration with agile Unmanned Aerial Vehicles (UAVs) equipped with Reconfigurable Intelligent Surfaces (RIS). The RIS-equipped UAVs act as intelligent relays, actively shaping the wireless environment to guarantee robust communication links for vehicular task offloading, thereby eliminating signal dead zones.

1.1 Background

Recent research demonstrates benefits of integrating UAVs, MEC, and RIS. Earlier work used terrestrial and building-mounted RIS to enhance ground-to-UAV-MEC links [1], showing substantial improvements in static or low-mobility environments. To address mobility challenges, recent studies have investigated aerial RIS architectures with RIS-equipped UAVs (UAV-RIS), enabling flexible deployment, agile beamforming, and enhanced spatial adaptability in MEC systems [2].

1.2 Problem Definition

Despite existing works that use RIS and MEC, they fail to address the high-mobility and scalability demands of vehicular networks, often relying on static RIS or single-UAV systems. A fully integrated, multi-layered aerial architecture is needed.

1.3 Contribution

- A novel multi-layered aerial architecture combining a HAPS for wide-area orchestration with agile, dual-function UAVs.
- A joint optimization framework for task assignment, RIS configuration, and power allocation.
- A comparison with baseline methods.

2 Methodology

Our methodology leverages a cooperative aerial network with a HAPS and dual-function UAVs (equipped with both RIS and MEC) to support autonomous vehicle telematic tasks in remote highway segments with no ground infrastructure. Each UAV performs both RIS-assisted signal shaping and local MEC computation for nearby vehicles, while the HAPS coordinates UAVs and can process aggregated tasks if needed. We formulate a joint optimization problem to minimize a weighted sum of task delay and energy consumption. Given its NP-hard nature, we propose an efficient iterative algorithm that decouples and sequentially optimizes task assignment, RIS configuration, and power control, yielding a high-performing, convergent solution.

2.1 Theoretical Framework

The core of our model is the ability to intelligently manipulate the signal-to-noise ratio (SNR) using UAV-mounted RIS. The received SNR at an aerial node u_n from ground user m, via a reflecting UAV-RIS $u_{n'}$, is given by

$$\mathrm{SNR}^{\mathrm{RIS}}_{m\to n'\to n} = \frac{P_m \left\| (\mathbf{h}^{\mathrm{RIS}}_{n'\to n})^H \, \boldsymbol{\Theta}_{n'} \, \mathbf{h}^{\mathrm{RIS}}_{m\to n'} \right\|^2}{\sigma^2}, \tag{1}$$

where the aerial node u_n can be either a MEC-enabled UAV or the HAPS, depending on the offloading path. Here, P_m is the transmit power of user m, $\mathbf{h}^{\mathrm{RIS}}_{m\to n'}$ is the channel vector from the user to the UAV-RIS, $\mathbf{h}^{\mathrm{RIS}}_{n'\to n}$ is the channel vector from the UAV-RIS to the receiving aerial node, $\boldsymbol{\Theta}_{n'}$ is the diagonal phase-shift matrix of the UAV-RIS, and σ^2 is the receiver noise power at the aerial node.

2.2 Algorithm Description

Fig. 1 illustrates the proposed iterative algorithm, which decouples the joint optimization problem into three tractable subproblems—task offloading, RIS phase configuration, and power control—solved sequentially until convergence.

3 Results and Discussion

3.1 Experimental Setup

The results are averaged over 1,000 independent simulation runs for statistical validity, with all simulations conducted in MATLAB R2024a. Autonomous vehicles enter the highway segment from multiple directions. Each vehicle travels at its preferred speed while dynamically adjusting its velocity based on surrounding traffic to maintain safe distances, capturing realistic ITS mobility patterns.

Our experiments use 1 HAPS node and 3 UAV-RISs, all equipped with MEC capabilities. The UAVs are strategically deployed to cover distinct, non-overlapping areas along a 3 km two-lane highway. The scenario considers 51–64 autonomous vehicles performing telematic tasks. We compare our proposed JOPPO method against several benchmarks: SAO2 with random RIS, Full HAPS Offloading, Equilibrium Offloading, Local UAV, and our previous work (NoRIS-GA).

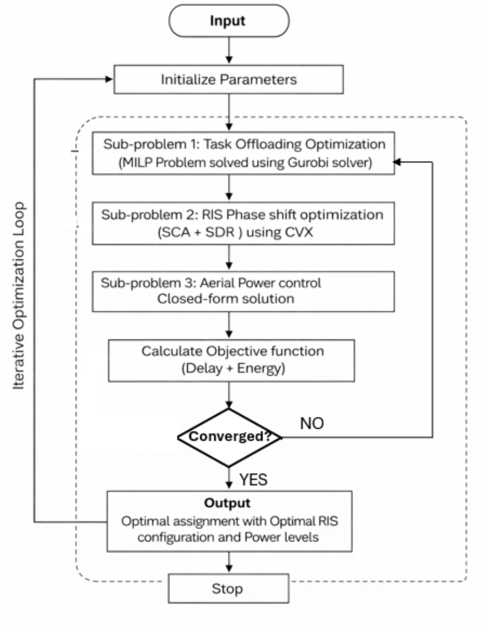


Figure 1: Proposed Algorithm Flowchart.

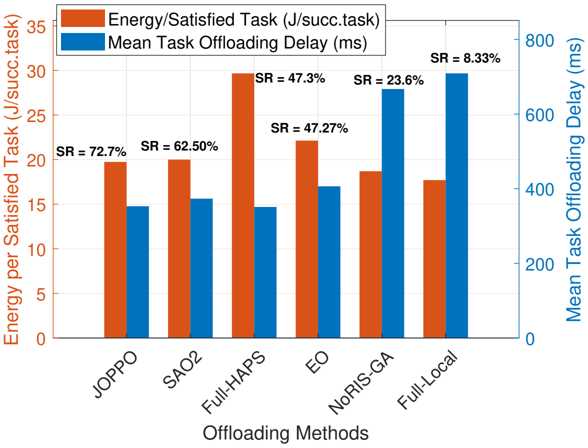


Figure 2: Energy efficiency and mean task offloading delay.

3.2 Quantitative Results

As shown in Fig. 2, our proposed collaborative HAPS-UAV-RIS framework achieves the lowest mean

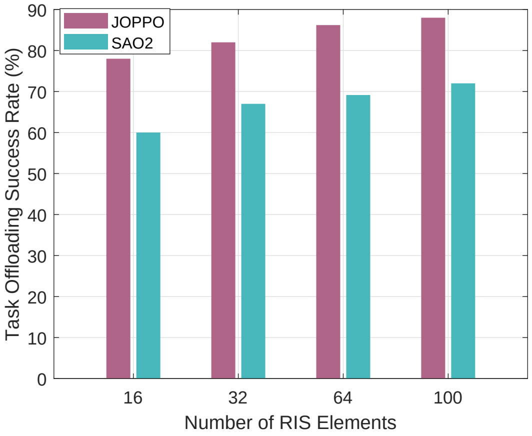


Figure 3: Task offloading success rate versus the number of RIS elements.

task delay and highest energy efficiency, successfully completing 72.7% of autonomous vehicle telematic tasks. It provides an optimal balance, outperforming Full-HAPS (fast but energy-intensive) and NoRIS-GA (slow and less reliable).

Fig. 3 shows the impact of varying the number of RIS elements on the task offloading success rate for our proposed scheme compared to the SAO2 (Random phase-shift). Both methods improve as more RIS elements are deployed, highlighting the benefit of larger RIS configurations for better communication quality. However, JOPPO consistently outperforms SAO2, achieving significantly higher success rates across all configurations. This demonstrates the advantage of jointly optimizing task assignment and RIS phase adjustments for optimal performance.

3.3 Qualitative Analysis

Our results demonstrate a fundamental shift from simple aerial offloading (as in [3]) to actively managing the wireless environment with RIS , which intelligently shapes the channel to boost SNR at the MEC server. This SNR gain directly enables higher data rates, reducing task offloading delay and increasing success rate. The comparison with NoRIS-GA in Fig. 2 quantifies this benefit: RIS is not merely an add-on, but a critical enabler of reliable, low-latency vehicular computing.

4 Conclusion

This work shows that HAPS–UAV networks with RIS-enhanced links improve task success, latency, and energy efficiency through joint optimization. The results emphasize the critical role of intelligent channel management and aerial-assisted computing in enabling reliable and scalable vehicular networks under dynamic traffic conditions. Future research will focus on real-world experimental validation, mobility-aware channel modeling, and the development of decentralized and learning-based optimization algorithms to further improve adaptability and scalability in large-scale ITS deployments.

References

[1] H. Hu, Sheng, and et.al, “Computation capacity maximization for UAV and RIS cooperative MEC system with NOMA,” *IEEE Commun. Lett.*, vol. 28, no. 3, pp. 592–596, 2024.

[2] Z. Zhai and et.al, “Energy-efficient UAV-mounted RIS assisted mobile edge computing,” *IEEE Wireless Commun. Lett.*, vol. 11, no. 12, pp. 2507–2511, 2022.

[3] I. Rzig and et al., “Dependency-aware task offloading in cooperative UAV-HAPS-assisted vehicular networks,” in *Proc. Int. Wireless Commun. and Mob. Comput. (IWCMC)*, 2024, pp. 1541–1546.

Energy Consumption Optimization in a Microgrid Using Artificial Intelligence Techniques

Chaima BEN FREDJ[1], Samir SAKARANI[1], Lotfi CHARAABI[1]

[1]Laboratory of Renewable Energy Systems (LSE), National Engineers School of Tunis (ENIT), Tunis, Tunisia
`chaima.benfredj01@gmail.com`

Abstract

With the increasing integration of renewable energy sources and the need for efficient energy management, microgrids have become essential for sustainable power systems. However, optimizing energy dispatch in real time while respecting physical constraints remains challenging due to the computational complexity of traditional optimization methods. This paper presents an artificial intelligence–based approach for microgrid energy consumption optimization using a Physics-Informed Convolutional Neural Network (Pi-CNN). The proposed model embeds domain-specific physical constraints directly into its loss function, ensuring power balance and generation limits while minimizing operational costs [1]. Using adapted real-world data from the "Ouessant Island microgrid in France", the Pi-CNN was trained to predict optimal dispatch decisions for photovoltaic, wind, and diesel generation units, achieving R^2 values above 0.97 for all active sources. The trained model was quantized and converted to TensorFlow Lite format for prospective deployment on the BeagleBone AI-64 edge platform. Although full DSP execution could not be completed due to software limitations, results confirm the feasibility of physics-informed deep learning for efficient, edge-based microgrid energy optimization.

Keywords: Microgrid, Energy optimization, Edge AI, Physics-informed CNN (Pi-CNN), Real-time dispatch, Embedded systems

1 Introduction

Microgrid energy dispatch requires fast and reliable coordination of multiple generation units under operational constraints. This work investigates a physics-informed learning approach to improve the reliability of AI-based microgrid dispatch.

1.1 Background

A microgrid is a localized energy system composed of renewable sources such as **Photovoltaic (PV)** and **Wind (W)**, conventional generators including **Diesel (DS), Natural Gas (NG),** and **Combined Heat and Power (CHP)**, along with energy storage units. These components must be coordinated to ensure stable operation and cost-effective energy supply [2].

Recent advances in deep learning have enabled learning-based approaches for power system analysis and control [3]. In parallel, Edge AI allows trained models to be executed directly on embedded platforms close to the data source, enabling low-latency and autonomous operation [4]. Such capabilities are particularly relevant for microgrid applications where real-time response and local control are essential.

1.2 Problem Definition

The Microgrid Economic Dispatch (MED) problem aims to minimize generation cost while meeting demand and respecting each unit's operational limits [1]. Classical optimization methods (e.g. linear or quadratic programming), though effective, are computationally intensive and unsuitable for real-time inference on embedded systems. Recent studies have explored deep neural networks to approximate optimal dispatch more efficiently; however, these models are rarely constrained by physical laws and have not been deployed on resource-limited devices. This gap motivates the development of a Physics-Informed Convolutional Neural Network (Pi-CNN) that embeds physical constraints to achieve accurate and efficient energy dispatch.

1.3 Contribution

- Development of a Physics-Informed Convolutional Neural Network (Pi-CNN) integrating physical laws in its loss function.

- Adaptation of real-world Ouessant Island microgrid data to the Pi-CNN input/output format.
- Quantization and conversion of the model to TensorFlow Lite for embedded deployment.
- Experimental validation showing $R^2 > 0.97$ for all generation units.
- Attempted deployment on BeagleBone AI-64, identifying DSP-related software limitations and proposing future solutions.

2 Methodology

2.1 Theoretical Framework

The Microgrid Economic Dispatch (MED) problem aims to determine the optimal power generation schedule that satisfies the energy demand at minimum cost while respecting operational and physical constraints. Mathematically, the MED problem for diverse generation sources including **CHP, NG, DS, PV, and W**, i.e., $G \in \{\text{CHP, NG, DS, PV, W}\}$, can be formulated as follows:

• Objective function (minimize generation cost):

$$\min \sum_{t=1}^{T} \sum_{i=1}^{N_G} F_i^G(t) = \min \sum_{t=1}^{T} \sum_{i=1}^{N_G} c_i P_i^G(t) \tag{1}$$

Where c_i denote the cost coefficient of generation unit i, $P_i^G(t)$ its generated power at time t, and N_G the total number of available units.

• Power balance constraint:

$$\sum_{i=1}^{N_G} P_i^G(t) = P_{load}(t) \tag{2}$$

Equation (2) ensures that total generated power matches the load demand at every time step.

• Operational constraints:

Each generator must operate within its technical limits. Renewable units such as photovoltaic (PV) systems and wind (W) turbines are constrained by minimum and maximum generation limits:

$$P_i^{min} \leq P_i^G(t) \leq P_i^{max}, \quad G \in \{PV, W\} \tag{3}$$

Conventional thermal generators (CHP, NG, Diesel) must additionally satisfy ramp-rate constraints that limit rapid changes in output between consecutive time steps:

$$|P_i^G(t) - P_i^G(t-1)| \leq R_i, \quad G \in \{CHP, NG, DS\} \tag{4}$$

Traditional numerical optimization methods such as linear programming and mixed-integer programming can solve this dispatch problem accurately; however, their computational cost increases significantly in real-time and embedded contexts, especially when high-resolution temporal forecasts and uncertainty are included. To address this limitation, a Physics-Informed Convolutional Neural Network (Pi-CNN) is employed. This model integrates domain physics directly into the learning framework, allowing it to learn dispatch rules that remain consistent with fundamental system behavior.

The Pi-CNN loss function incorporates multiple terms to enforce both learning accuracy and physical feasibility:

$$\mathcal{L}_{PI} = \mathcal{L}_{MSE} + \lambda_1 \left\| \sum_{i=1}^{N_G} P_i^G(t) - P_{load}(t) \right\|_2^2 + \mathcal{L}_{Con} \tag{5}$$

where:

- $\mathcal{L}_{MSE}$ ensures that predictions remain close to optimal dispatch values.
- the second term penalizes any violation of the power-balance constraint in (2).
- $\mathcal{L}_{Con}$ enforces constraints such as generator limits and ramp-rate bounds.

By integrating physics into the loss function, the Pi-CNN enforces energy-balance constraints while learning economic behavior, improving robustness under unseen conditions and supporting real-time smart-grid deployment.

Compared to classical optimization solvers that rely on iterative constrained optimization at each time step, the proposed PI-CNN performs inference through a single forward pass with fixed computational complexity, making it more suitable for real-time and embedded microgrid deployment.

2.2 Datasets and Tools

This work uses real-world microgrid data from the publicly available Ouessant Island 2016 dataset [5], which provides hourly measurements of electrical load and meteorological variables (solar irradiance, temperature, and wind speed) for an isolated microgrid in France. To ensure compatibility with the reference Pi-CNN architecture using 11 input features [1], the dataset was cleaned, validated, and reformatted, with missing or inconsistent entries removed.

2.2.1 Physics-Based Feature Engineering

To reconstruct renewable generation profiles consistent with the reference model assumptions, physical power generation equations were applied. First, photovoltaic (PV) output was estimated through the NOCT (Nominal Operating Cell Temperature)-based temperature model and power production formula, accounting for irradiance and temperature-dependent efficiency effects. Similarly, wind generation was computed using a piecewise turbine power curve that considers cut-in, rated, and cut-out wind speeds. These physics-informed transformations ensured realistic renewable production profiles aligned with true microgrid operation.

2.2.2 Output Label Generation via Economic Dispatch Optimization

Since the Ouessant dataset does not directly include dispatch decisions, ground-truth labels were generated by solving a constrained linear optimization problem at each time step. The objective was to minimize total generation cost while meeting power balance and respecting generator capacity limits. Cost coefficients for PV, wind, and diesel units were adapted from the reference literature, incorporating capital investment, operation and maintenance, and fuel consumption terms. Unlike the original case study, combined heat-and-power (CHP) and natural gas (NG) units are not present in the Ouessant microgrid; therefore, their contributions were set to zero to preserve the 11-input / 5-output architecture.

2.2.3 Temporal Extraction and Final Dataset

To mirror the experimental setup of the reference study which focuses on one month of high-resolution data this work extracted the June 2016 window from the yearly series. Although the sampling frequency remains hourly, this subset captures seasonal fluctuations in renewable availability and load profiles. The resulting dataset, saved as `ED_labeled_hourly_2016-06.csv`, includes:

- Timestamp.
- **11 input features:**

$$[P_{\text{load}}, P_{PV}, P_{Wind}, P_{PV}^{\max}, P_{Wind}^{\max},\\ P_{CHP}^{\min}, P_{CHP}^{\max}, P_{NG}^{\min}, P_{NG}^{\max},\\ P_{DS}^{\min}, P_{DS}^{\max}]$$

- **5 dispatch target outputs:**

$$[P_{CHP}^{\text{disp}}, P_{NG}^{\text{disp}}, P_{DS}^{\text{disp}}, P_{Wind}^{\text{disp}}, P_{PV}^{\text{disp}}]$$

- For compatibility with the original architecture [1], **CHP and NG dispatch values are set to zero** (these units are absent in the Ouessant microgrid).
- A load-shedding indicator for infeasibility cases.

An 80/20 split was adopted for training and validation, ensuring that the Physics-Informed Convolutional Neural Network (Pi-CNN) learns dispatch behavior under realistic operating conditions while maintaining generalization capability.

2.3 Algorithm Description

The proposed Pi-CNN architecture consists of a 1D convolutional layer with 16 filters followed by a max-pooling layer of size 2, then a second 1D convolutional layer with 32 filters and another max-pooling layer of size 2. The extracted features are passed through fully connected layers, ending with a five-neuron output layer representing the dispatch power of CHP, NG, DS, PV, and Wind units. The model was trained using the Adam optimizer for 60 epochs and subsequently quantized to 8-bit integers before conversion to TensorFlow Lite format for prospective BeagleBone deployment.

3 Results and Discussion

3.1 Experimental Setup

Model training and testing were performed on a local x86-based computer (Python 3.10, TensorFlow 2.x).

For deployment on the BeagleBone AI-64 (TDA4VM SoC), model training, quantization, TensorFlow Lite conversion, TIDL compilation, and artifact generation were performed on an x86 host, and the resulting binaries were transferred to the board for inference. However, full DSP execution was not possible due to an error in the firmware (vx_app_rtos_linux_c7x_1.out), which prevented initialization of the C7x runtime. The firmware failure prevents proper allocation of memory carveouts and mapping of shared buffers between the A72 and C7x cores, thereby blocking DSP initialization.

3.2 Results

The proposed PI-CNN effectively captures the nonlinear relationship between renewable availability, load demand, and economic dispatch decisions. Figure 1 shows that the predicted dispatch closely follows the optimal trajectories for photovoltaic, wind, and diesel units, while CHP and NG outputs remain at zero, consistent with the Ouessant microgrid configuration.

In addition to the coefficient of determination, the proposed PI-CNN achieved an average Mean Absolute Error (MAE) of **7.31 kW** across all generation units, with overall average R^2 score of **0.99**. Renewable units exhibit low prediction error (MAE = 1.10 kW for PV and 18.30 kW for wind), while the diesel generator maintains accurate dispatch tracking ($R^2 = 0.9867$).

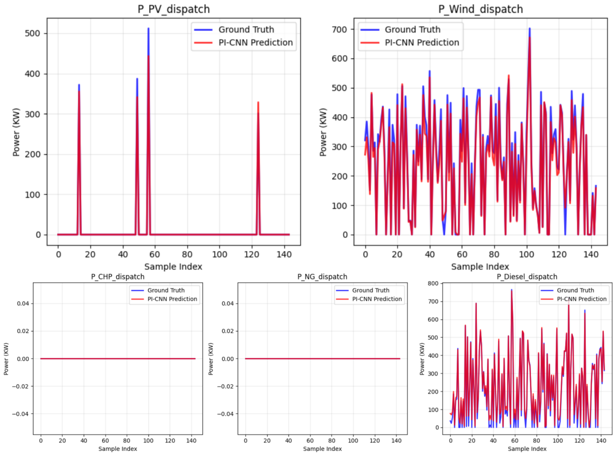


Figure 1: Dispatch predictions vs. ground truth for all generation units (CHP, NG, Diesel, PV, and Wind).

Physics-constraint validation shows zero violations for all generation limits and an average power-balance error of **1.65%**, demonstrating the effectiveness of the physics-informed loss formulation.

Compared to a standard CNN trained without physical constraints, the proposed PI-CNN achieves higher predictive accuracy while strictly satisfying power balance and generation limits. This confirms that embedding physics into the learning process improves both reliability and feasibility, which are critical for real-time microgrid operation.

4 Conclusion

- This work introduced a Physics-Informed Convolutional Neural Network (Pi-CNN) for optimizing microgrid energy dispatch. The model achieved high predictive accuracy ($R^2 > 0.97$) while maintaining power balance and respecting generation constraints, confirming the effectiveness of physics-informed learning for realistic energy management.
- Compared to purely data-driven baselines, the PI-CNN demonstrates superior accuracy and guaranteed physical feasibility while maintaining low computational complexity, making it well suited for real-time edge deployment in microgrid environments.
- The results highlight the potential of combining deep learning with edge computing to enable fair, autonomous, and efficient power allocation in microgrids helping prevent localized blackouts by ensuring equitable energy distribution across all consumers.
- Deployment on the BeagleBone AI-64 was limited by DSP firmware issues, which restricted full utilization of the heterogeneous hardware.
- Future work includes using our own custom datasets, deploying the Pi-CNN on alternative platforms such as NVIDIA Jetson Nano, extending it to larger and more complex microgrids, and applying model compression to enhance inference speed and scalability on edge devices.

References

[1] X. Ge and J. Khazaei, “Physics-informed Convolutional Neural Network for Microgrid Economic Dispatch,” *Sustainable Energy, Grids and Networks*, vol. 40, Art. no. 101525, 2024.

[2] D. E. Olivares, A. Mehrizi-Sani, A. H. Etemadi, C. A. Cañizares, R. Iravani, M. Kazerani, A. H. Hajimiragha, O. Gomis-Bellmunt, M. Saeedifard, R. Palma-Behnke, G. A. Jiménez-Estévez, and N. D. Hatziargyriou, “Trends in Microgrid Control,” *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1905–1919, 2014.

[3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[4] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, “Edge Computing: Vision and Challenges,” *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, 2016.

[5] Microgrids-X, “Microgrids.py: Tools and Examples for Microgrid Simulation and Optimization,” GitHub repository, 2021. [Online]. Available: https://github.com/Microgrids-X/Microgrids.py

NILM Based on Clustering for Load Disaggregation

Wissal BEN MAATOUG[1], Sejir KHOJET EL KHIL[1], Lotfi CHARAABI[1], Sofiane SAYAHI[2]
[1]Université de Tunis El Manar, Ecole Nationale d'Ingénieurs de Tunis, LR11ES15, Tunis, Tunisia
[2]Research and Innovation Department, ACTIA Engineering Services, Tunis, Tunisia
wissal.benmaatoug@etudiant-enit.utm.tn

Abstract

Non-Intrusive Load Monitoring (NILM) enables appliance-level energy disaggregation from aggregate consumption data. A key limitation of traditional methods is their dependence on prior knowledge of appliance signatures. To address this, this paper proposes a robust, two-stage unsupervised hybrid framework that automatically discovers appliance states without such prior knowledge. The methodology consists of an appliance state identification phase, where a comparative study of three density-based clustering algorithms (ADPC, DBSCAN, and HDBSCAN) is presented, followed by a disaggregation phase using a Factorial Hidden Markov Model (FHMM). The study evaluates the framework on the AMPds dataset, measuring clustering quality, parameter sensitivity, and disaggregation accuracy. Results demonstrate that HDBSCAN outperforms both ADPC and DBSCAN, achieving 85.4% disaggregation accuracy with minimal parameter tuning.

Keywords: NILM, load disaggregation, clustering algorithms, HDBSCAN, ADPC, HMM

1 Introduction

Non-Intrusive Load Monitoring (NILM), introduced by Hart in 1992 [1], identifies individual appliances from total power usage without separate meters, providing a cost-effective way to manage energy. Modern NILM uses machine learning, especially clustering, to detect appliance states (OFF, ON, standby), thereby eliminating the need for manual labeling. However, traditional clustering methods such as K-means require the number of clusters K_n to be specified in advance. This is a significant limitation, as K_n varies per appliance and an incorrect specification can severely degrade performance, hindering real-world deployment where appliance characteristics are often unknown.

The main contributions of this work are as follows:

1. The authors propose a hybrid model, combination of clustering algorithms with FHMM. The clustering algorithms are used to identify automatically the operational states of the system's devices, while HMM are trained for each device to capture consumption dynamics. These HMMs are then combined to form an FHMM, a factorial model, to disaggregate power using those state.

2. The authors conduct a systematic comparison of three density-based clustering algorithms—ADPC, DBSCAN, and HDBSCAN—for identifying automatically appliance operating states and associated powers.

3. The authors provide a comprehensive evaluation on the AMPds dataset, measuring clustering quality, parameter sensitivity, and disaggregation accuracy.

2 Methodology

2.1 Theoretical Framework

Let y_t denote aggregate power consumption at time t. NILM estimates individual contributions $x_t^{(n)}$:

$$y_t = \sum_{n=1}^{N} x_t^{(n)} + \epsilon_t \tag{1}$$

where N is the appliance count and ϵ_t represents noise. Each appliance operates in discrete states $s_t^{(n)} \in \{1, 2, ..., K_n\}$, where K_n is the number of operational states for appliance n. The key challenge is determining K_n automatically without prior knowledge.

2.2 Proposed Two-Stage Framework

Stage 1 - Automatic State Discovery via Clustering: Apply density-based clustering to power consumption features to automatically identify operational states K_n for each appliance without supervision. This approach eliminates the manual specification of state numbers required by traditional methods.

Stage 2 - Temporal Modeling via HMM/FHMM: Clustering provides static state identification but ignores temporal dependencies. For each appliance n, we train a Hidden Markov Model (HMM) [2] using discovered cluster centers as initial emission means, with transition matrix $A^{(n)}$ and emission probability $B^{(n)}$:

$$P(s_t^{(n)}|s_{t-1}^{(n)}) = A^{(n)}, \quad P(x_t^{(n)}|s_t^{(n)}) = B^{(n)} \tag{2}$$

where $A^{(n)} = [a_{ij}^{(n)}]$ is the state transition probability matrix with $a_{ij}^{(n)} = P(s_t^{(n)} = j|s_{t-1}^{(n)} = i)$, and $B^{(n)}$ represents the emission probability distribution, typically modeled as Gaussian $\mathcal{N}(\mu_k, \sigma_k^2)$ for state k. Individual HMMs are combined into a Factorial HMM (FHMM) for joint disaggregation .

2.3 Dataset

AMpds dataset [3]: Vancouver residential data (April 2012-March 2013), 1-minute sampling, 11 appliances, 10,080 training samples (one week).

2.4 Clustering Algorithms

The Density Peaks Clustering (DPC) algorithm identifies clusters based on density but requires manual selection of cluster centers from its decision graph [5] . To address this limitation, an ADPC algorithm that automatically determines the cluster centers.

Adaptive Density Peak Clustering (ADPC) computes local density and minimum distance to identify cluster centers:

$$\rho_i = \sum_j \chi(d_{ij} - d_c), \quad \delta_i = \min_{j:\rho_j > \rho_i} (d_{ij}) \tag{3}$$

where d_{ij} represents the Euclidean distance between data points i and j, d_c is the cutoff distance threshold, $\chi(x) = 1$ if $x < 0$ and $\chi(x) = 0$ otherwise, ρ_i denotes the local density of point i, and δ_i represents the minimum distance from point i to any other point with higher density. Points with high density ρ_i and large separation δ_i are selected as cluster centers.

Density-Based Spatial Clustering (DBSCAN) finds density-connected clusters. For a point p, its ϵ-neighborhood is:

$$N_\epsilon(p) = \{q \in D : \text{dist}(p, q) \leq \epsilon\} \tag{4}$$

p is a **core point** if $|N_\epsilon(p)| \geq$ minPts.

Hierarchical DBSCAN (HDBSCAN) operates by first transforming pairwise distances using local density: for each point p_i, its core distance $\delta_k(p_i)$ is the distance to its k-th nearest neighbor and $d(p_i, p_j)$ is the original distance between points p_i and p_j. The mutual reachability distance between two points is then defined as

$$d_{\text{mreach}}(p_i, p_j) = \max\{\delta_k(p_i),\ \delta_k(p_j),\ d(p_i, p_j)\}, \tag{5}$$

which emphasizes denser regions. A minimum spanning tree (MST) is constructed from these distances, and edges are removed in decreasing order to form a cluster hierarchy. Each cluster c in the hierarchy is assigned a persistence value $\lambda(c)$, often as

$$\lambda(c) = \frac{1}{\epsilon_{\text{birth}} - \epsilon_{\text{death}}}, \tag{6}$$

where ϵ_{birth} and ϵ_{death} are density thresholds at which the cluster appears and splits. The condensed tree is obtained by maximizing total persistence

$$\text{TP} = \sum_{c \in \mathcal{C}} \lambda(c), \tag{7}$$

subject to constraints. Finally, for each unassigned point x, the probability of belonging to cluster c_i is computed using a distance-based measure such as

$$P(x \in c_i) = \frac{\exp(-d(x, c_i))}{\sum_j \exp(-d(x, c_j))}. \tag{8}$$

3 Results and Discussion

3.1 Experimental Setup

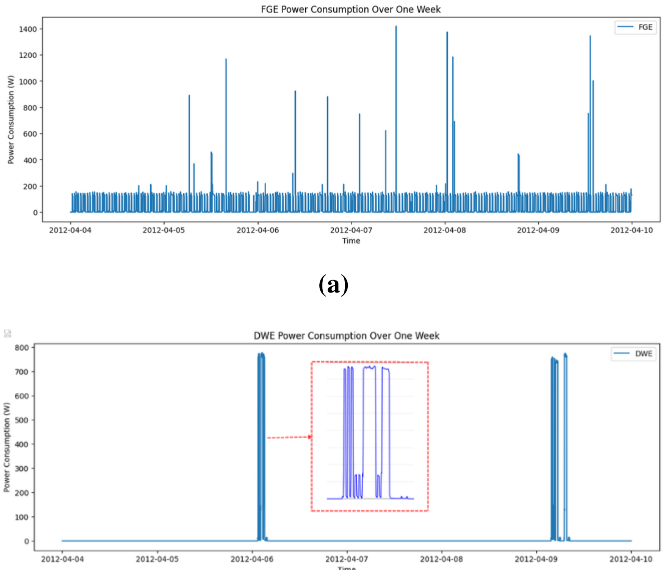
The authors use dynamic parameter tuning for all clustering methods. DBSCAN with adaptive epsilon parameter selection per appliance using k-NN (k=5) and KneeLocator. For HDBSCAN, the authors cross-validate `min_cluster_size` over the range $[5, 30]$ and set `min_samples` using k-NN ($k = 5$). ADPC automatically optimizes its density and distance thresholds. ADPC automatically optimizes its density and distance thresholds.

3.2 Quantitative Results

3.2.1 Identification of device operating states

The identified operating states directly influence the initialization of the HMM parameters and thus affect disaggregation accuracy.

Figure 1 illustrates the raw power consumption patterns for two representative appliances from the AMPds dataset: (a) Fridge and (b) Dishwasher. The authors apply ADPC, DBSCAN, and HDBSCAN algorithms to discover each appliance's operating states.



(a)

(b)

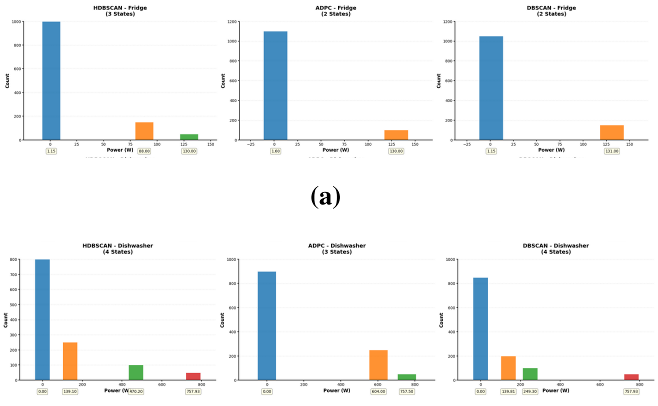
Figure 1: Power consumption of two appliances in AMPds: (a) Fridge, (b) Dishwasher.

The operating states of the appliances are identified using three distinct clustering algorithms. Figure 2 provides a comparative visualization of the states discovered by each algorithm, with a separate set of bars representing each algorithm's output; each bar corresponds to a discovered state and its associated mean power consumption level.

For the refrigerator (FGE) in Figure 2(a), HDBSCAN identifies three states (1.15, 88, 130W), while DBSCAN and ADPC identify only two (1.15, 130W). For the dishwasher (DWE) illustrated in Figure 2(b), both HDBSCAN and DBSCAN identify four operating states, whereas ADPC only detects three. The intermediate power levels differ between HDBSCAN (470.2 W) and DBSCAN (249.3 W), which could correspond to different low-consumption or normal-operation phases. Notably, the fourth state identified by HDBSCAN, at 757.9 W, is consistent with the high-power demand expected for a drying cycle.

3.2.2 Clustering Performance

Table 1 summarizes the clustering silhouette scores, which measure clustering quality by assessing both cluster cohesion and separation. HDBSCAN achieves the highest average silhouette score of 0.9854, indicating well-separated and cohesive clusters, compared to DBSCAN (0.9712) and



(a)

(b)

Figure 2: Comparison of operating states obtained with HDBSCAN, ADPC, DBSCAN algorithms: (a) Fridge, (b) Dishwasher.

ADPC (0.947). Notably, HDBSCAN achieves perfect clustering (score of 1.0) for the clothes dryer and demonstrates substantial improvements for the fridge (0.969 vs. 0.835 for ADPC).

Table 1: Clustering Performance (Silhouette Score)

Appliance	ADPC	DBSCAN	HDBSCAN
Washing Machine	0.979	0.987	0.987
Clothes dryer	0.995	0.994	**1.0**
Heat Pump	0.934	0.969	**0.973**
Fridge	0.835	0.914	**0.969**
Dishwasher	0.992	0.992	**0.998**
Average	0.947	0.9712	**0.9854**

3.2.3 Disaggregation Accuracy

Table 2 reports the appliance-level disaggregation accuracy. HDBSCAN demonstrates superior performance for appliances with distinct consumption patterns, achieving near-perfect accuracy for the washing machine (99.9%) and dishwasher (98.9%), and perfect accuracy (100%) for the clothes dryer. However, the fridge remains low at 41.7%, despite high clustering quality.

Table 2: Appliance-Level Disaggregation Accuracy (%)

Appliance	ADPC +HMM	DBSCAN +HMM	HDBSCAN +HMM
Washing Machine	64.5	63.8	**99.9**
Clothes dryer	20.3	21.2	**100**
Heat Pump	86.5	86.7	86.7
Fridge	47.3	41.7	41.7
Dishwasher	2.6	19.8	**98.9**
Average	44.24	46.64	**85.44**

3.3 Qualitative Analysis

HDBSCAN achieves superior performance because it adapts to different data densities, identifies clusters reliably, and exhibits minimal sensitivity to parameter settings. This leads to more accurate HMM initialization and better FHMM disaggregation performance. DBSCAN struggles with varying densities across appliances, and ADPC performs poorly with high-dimensional NILM data. The dishwasher accuracy improvement (2.6% to 98.9%) demonstrates how robust clustering and low parameter sensitivity enhance both state detection and temporal sequence modeling.

The paradox of high clustering quality (0.969) but low disaggregation accuracy (41.7%) for the fridge stems from two interconnected challenges that the FHMM cannot overcome. First, the fridge's irregular, high-frequency state transitions (cycling rapidly between OFF, compressor ON, and defrost modes) multiple times per hour, violate the FHMM's assumption of stable state durations, making its temporal sequence hard to model. Second, state ambiguity in the aggregate signal is critical: the fridge's intermediate state (88W) fall within a crowded low-power range where other appliances (e.g., heat pump) also operate. While clustering cleanly separates these states in feature space, the FHMM fails to correctly attribute them during disaggregation because its independence assumption cannot resolve overlapping power draws from simultaneously operating devices. Thus, high clustering quality confirms that an appliance's signature is recognizable, but low disaggregation accuracy indicates that its temporal behavior is exceptionally hard to model with standard sequential approaches.

4 Conclusion and Future Work

The authors have introduced a robust, two-stage unsupervised framework for NILM that overcomes a major limitation of traditional methods by automatically discovering appliance states without prior knowledge. By effectively combining density-based clustering with a Factorial HMM (FHMM), the framework enables fully unsupervised load disaggregation. Evaluation on the AMPds dataset demonstrates that HDBSCAN clearly outperforms ADPC and DBSCAN, achieving an average disaggregation accuracy of 85.4%. The framework shows particular strength for appliances with distinct operational patterns, while also providing critical insights into the challenge of modeling appliances with highly complex or irregular temporal dynamics.

Future work will focus on three main directions: (1) constructing a comprehensive NILM dataset within the ENIT smart home laboratory to validate and test the framework's generalization capabilities; (2) applying the proposed framework to this new dataset as a critical benchmark; and (3) exploring the integration of deep learning models, such as recurrent or attention-based architectures, to better capture complex temporal patterns and improve accuracy for appliances with highly variable or transient consumption states.

Acknowledgment

This work was supported by the LR11ES15 Laboratory at the National Engineering School of Tunis and by ACTIA Engineering Services.

References

[1] Hart, G. W. "Nonintrusive Appliance Load Monitoring," *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.

[2] Makonin, S., et al. "Exploiting HMM Sparsity for Online NILM," *IEEE Trans. Smart Grid*, vol. 7, no. 6, pp. 2575–2585, 2016.

[3] Makonin, S., et al., "AMPds," *Proc. IEEE EPEC*, pp. 1–6, 2013.

[4] Klemenjak, C., et al., "Electric Energy Data Set," *Scientific Data*, vol. 7, no. 1, 2020.

[5] A. Rodriguez, A. Laio, Machine learning. clustering by fast search and find of density peaks, Science 10.1126/science.1242072. 344 (2014) 1492–1496, https://doi.org/10.1126/science.1242072.

Evaluating a tracking-by-detection approach across challenging datasets for cow monitoring

Roua Mkadmi[1], Rabaa Youssef-Douss[1,2], Amel Benazza-Benyahia[1]

[1] University of Carthage, SUP'COM, LR11TIC01, COSIM Lab., 2083, Ariana, Tunisia
[2] University of Carthage, INSAT, 1080, Tunis, Tunisia
roua.mkadmi@supcom.tn

Abstract

In precision livestock farming, computer vision solutions have been developed to track cows in farms for monitoring their behaviour and enabling timely interventions. In this work, a tracking by detection approach is considered: it involves YOLOv8n as the detector and BotSort as the tracker. Its performance is evaluated and discussed on three public datasets with different properties. The proposed solution achieves promising multiple object tracking performance even for challenging scenes.

Keywords: cattle, multi-object detection, multi-object tracking, deep neural network.

1 Introduction

In smart cattle rearing, computer vision techniques are gaining considerable attention due to their non-intrusive nature and low cost. In this respect, in [1], a method called CowXNet was developed to track cows in order to decide on their estrus state. Recently [2], we have proposed a novel tracking-by-detection method using YOLOv8n [3] as the detector and BotSort [4] as the tracker of cows in barns. However, the performance was only assessed on a single dataset [1]. The present study aims at studying the generalizability of our tracking-by-detection solution by evaluating its performance on other video datasets with varying contents captured under different conditions. Also, we pay attention to benchmark our tracker-by-detection with the recent method CowXNet [1].

2 Methodology

2.1 Considered datasets

This study considers three open datasets, namely CowXNet [1], MmCows [5] and CoBRA [6]. Their properties are different as summarized in Table 1 where all the frames in each dataset are RGB. Indeed, CowXNet features a relatively small pen and a specific camera configuration, leading to appearance variation challenges, an example frame of this datatset is illustrated in Figure 1a. MmCows has the highest sampling rate, and the pen is illuminated continuously by natural and artificial light during during the 20 hours of recording, an example frame from this dataset is shown in Figure 1b. Lastly, CoBRA dataset departs from the two other ones as it includes images captured by night-vision cameras, Figures 1c and 1d illustrate, respectively, day-captured and night-captured frames. All the datasets suffer from occlusions due to the spatial regrouping of the cows. This aspect is very pronounced in the case of CoBRA. Regarding annotations, the first two datasets provide bounding boxes (BB) and an identifier (ID) for each cow in every frame. The third dataset CoBRA does not provide annotations for the video set. Indeed, it is accompanied by a separate set of 621 still images along with the coordinates of six keypoints in specific parts (nose, head, wither, tail-head, left hip, and right hip) of each cow identified by its ID.

2.2 Data preparation

The adopted tracking by detection method requires training the detector, which entails training a separate YOLOv8n model for each dataset. Such trainings involve still images. Specifically, for both CowXNet and MmCows, we extract frames from the available annotated videos, whereas for CoBRA dataset, the 621 annotated still images are used. It is worth noting that for the latter dataset, since BBs are not provided, we delineate them from the available keypoints. A $70\% - 20\% - 10\%$ partition is followed for the training, validation, and test sets. Table 2-(a) describes the content of each set for the underlying datasets.

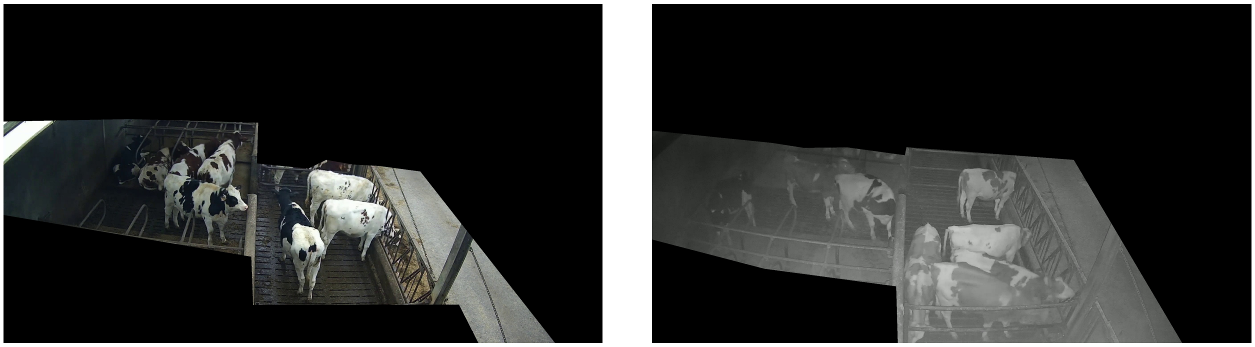
Regarding the tracker, as aforementioned, the BotSort algorithm [4] is retained in inference mode with-

Table 1: Considered dataset specifications. N stands for Natural and A stands for Artificial.

Dataset	Resolution	# Videos	Duration	Sampling rate	# Cows	Pen size	Camera	Illumination
CowXNet [1]	1280 × 960	1	∼17 min	0.5 s	3	3 × 6 m	top-down	N (day)
MmCows [5]	4480 × 2800	1	∼20 hours	15 s	16	20 × 12 m	corner angle	N (day), A
CoBRA [6]	2688 × 1520	100	1 min	0.5 s	8	NA	roof-mounted	N (day, night)

(a) CowXNet frame example.

(b) MmCows frame example.



(c) CoBRA (day) frame example.

(d) CoBRA (night) frame example.

Figure 1: Example frames from the considered datasets.

out needing any learning step. However, for measuring its performance, sequences of consecutive frames (clips) have to be extracted. For both CowXNet and MmCows datasets, the test sets that are unseen during YOLOv8 training are employed to construct these evaluation clips. For CoBRA, we manually annotate four videos: two day videos and two night videos. A subsampling 1:2 is chosen to merely facilitate the annotation. Table 2-(b) gives information on the evaluation clips.

Table 2: Dataset preparation.

(a) Detector training

Dataset	**# train**	**# val**	**# test**
CowXNet	1420	400	180
MmCows	3528	1008	504
CoBRA	416	105	100

(b) Tracker evaluation

Dataset	**# clips**	**# frames/clip**
CowXNet	3	60
MmCows	146	4
CoBRA (day)	2	60
CoBRA (night)	2	60

3 Experimental results

All experiments are conducted on an Ubuntu system equipped with NVIDIA GeForce GTX 1080 Ti with 11 GB dedicated memory. To ensure a fair benchmarking, the CowXNet method [1] is also learned on the aforementioned datasets by keeping the same training and evaluation data (see Table 2). More precisely, its detector (YOLOv4) is trained and, a simple Intersection Over Union (IoU) between BBs in two consecutive frames is used as the tracker.

Table 3 summarizes the performance of our detector (YOLOv8n) and the benchmarked one (YOLOv4). Both detectors achieve strong and comparable results across all datasets in terms the localization metric, the mean Average Precision at a 50% IoU threshold (mAP@50), and classification metric, the F1-score. In most cases, YOLOv8n outperforms YOLOv4, while offering the additional advantage of a considerably smaller model size.

The retained tracker metrics are Multiple Object Tracking Accuracy (MOTA), as defined in Equation 1, which aggregates all tracking errors into a single measure. Multiple Object Tracking Precision (MOTP), as given in Equation 2, quantifies the localization precision of detected objects, and the ID F1-score (IDF1), as expressed in Equation 3, measures how consistently a tracker maintains the same identity for each object throughout a video sequence.

$$\text{MOTA} = 1 - \frac{\sum_t (FN_t + FP_t + IDS_t)}{\sum_t GT_t} \tag{1}$$

where FN_t is the number of false negatives (missed targets), FP_t is the number of false positives (false detections), IDS_t is the number of identity switches (assigned different IDs in two consecutive frames), and GT_t is the number of ground truth objects at frame t.

$$\text{MOTP} = \frac{\sum_{i,t} d_{i,t}}{\sum_t c_t} \tag{2}$$

where $d_{i,t}$ is the IoU between the detected object and the ground truth object i at frame t, and c_t is the number of matches found at frame t.

$$\text{IDF1} = \frac{2 \times IDTP}{2 \times IDTP + IDFP + IDFN} \tag{3}$$

where $IDTP$ is the number of true positive identifications, $IDFP$ is the number of false positive identifications, and $IDFN$ is the number of false negative identifications.

Results are reported in Table 4. Our proposed method outperforms CowXNet method for all the metrics on the three datasets, especially for the challenging CoBRA night data.

Table 3: Performance of the detectors.

Dataset	**YOLOv8**		**YOLOv4**	
	mAP@50	F1-score	mAP@50	F1-score
CowXNet	0.9690	0.9562	0.9988	0.9900
MmCows	0.9930	0.9835	0.9928	0.9700
CoBRA (day)	0.9480	0.9100	0.9088	0.9100
CoBRA (night)	0.8880	0.8569	0.8385	0.8400
CoBRA (overall)	0.9180	0.8834	0.8736	0.8750

Table 4: Tracking on diff datasets.

Method	**Dataset**	**MOTA**	**MOTP**	**IDF1**
Proposed	CowXNet	89.63%	90.62%	82.88%
	MmCows	89.67%	90.74%	91.38%
	CoBRA (day)	65.41%	75.08%	77.89%
	CoBRA (night)	40.62%	68.69%	65.90%
	CoBRA (overall)	53.01%	71.88%	71.89%
CowXNet [1]	CowXNet	69.44%	81.51%	80.62%
	MmCows	81.06%	85.79%	87.73%
	CoBRA (day)	21,88%	67,25%	45,73%
	CoBRA (night)	23,44%	66,97%	44,37%
	CoBRA (Overall)	22.66%	67.11%	45.05%

4 Conclusion

The proposed tracking-by-detection method, consisting of YOLOv8n followed by BotSort, shows robust results across datasets with different visual cues. These include occlusions in high-density cow scenarios, appearance variations due to different camera configurations, and detection difficulties in night-vision setups. Furthermore, the proposed method is evaluated against the CowXNet method. Further work will focus on integrating our tracker by detection into a downstream application to analyze cow behaviour.

References

[1] T. Lodkaew, K. Pasupa, and C. K. Loo. Cowxnet: An automated cow estrus detection system. *Expert Sys. with Appli.*, 211:118550, 2023

[2] R. Mkadmi, R. Youssef-Douss, and A. Benazza-Benyahia. Cow monitoring system based on deep learning models for multiple objects detection and tracking. In *AICCSA*, pages 1–8, 2024.

[3] G. Jocher, A. Chaurasia, and J. Qiu, Ultralytics YOLOv8, `https://github.com/ultralytics/ultralytics`, 2023.

[4] N. Aharon, R. Orfaig, and B. Z. Bobrovsky. Bot-sort: Robust associations multi-pedestrian tracking. *arXiv preprint* arXiv:2206.14651, 2022.

[5] H. Vu, O. C. Prabhune, U. Raskar, D. Panditharatne, H. Chung, C. Choi, and Y. Kim. Mmcows: A multimodal dataset for dairy cattle monitoring. *Advances in Neural Information Processing Systems*, 37:59451–59467, 2024

[6] M. Perneel, I. Adriaens, B. Aernouts, and J. Verwaeren. Consistent multi-animal pose estimation in cattle using dynamic Kalman filter based tracking. *Smart Agri. Tech.*, page 101014, 2025.

Reconfigurable Phoenix-Inspired Unit Cells for X-Band Reflectarray Weather Radar

Ikram Dendani[1,3], Erwan Fourn[1], Saber Dakhli[2], Fethi Choubani[3]

[1]Univ Rennes, CNRS, INSA, IETR - UMR 6164, F-35000 Rennes, France
[2]GeePs Laboratory, CNRS UMR-8507 CentraleSupélec,Université Paris-Saclay, F-91192 Gif-sur-Yvette, France,
[3]Innov'Com Laboratory at SUPCOM, Univ of Carthage, Carthage, Tunisia,
`ikram.dendani@insa-rennes.fr`

Abstract

This paper presents a novel dual-varactor reconfigurable unit cell inspired by the Phoenix passive structure, designed for X-band weather radar reflectarray. The dual-varactor unit cell design achieves approximately 281 degrees, near-linear phase control, and reflection losses around 4 dB arise primarily from the varactor's internal resistance and inductance.

Keywords: Reflectarray, Phoenix cell, varactor diode, X-band, weather radar, reconfigurable.

1 Introduction

High-gain, electronically steerable apertures are central to weather surveillance radars that require precise beam control with compact, low-loss hardware suitable for cost-sensitive deployments. By employing spatial feeding and per-element phase control without RF combiners, reflectarray offer a convincing alternative to large, expensive phased arrays and unwieldy parabolic reflectors [1]. A critical design challenge lies in realizing unit cells capable of wide-range, smooth, and low-loss phase shifting across the operational X-band. This challenge is compounded by resonant dispersion effects that tend to limit bandwidth in medium-sized reflectarrays.

To address the above challenges, this work leverages the Phoenix passive motif [2] to derive reconfigurable unit cells with electronically controlled phase. The proposed unit cells are directly inspired by the concentric square-ring geometry, which provides multiple current flow paths as well as strong capacitive coupling, thus enabling continuous and extended phase control under linear polarization in the X-band [3]. First, by shifting the unit cell reference plane between two neighboring cells by half a period and recombining the facing halves, an equivalent single varactor version is obtained. This first configuration was previously presented in [4]. Then, by using the original shape of the Phoenix unit cell, a reconfigurable topology is obtained by integrating two varactor diodes. The latter are placed symmetrically across the inter-gap to preserve mirror symmetry, allowing for balanced biasing, and minimizing the cross-polarization by suppressing odd-mode currents. This dual-varactor configuration constitutes the second reconfigurable version. Fig. 1 presents the transformation scenarios from the passive Phoenix tiling leading to these two versions. This paper presents the second version in details, starting with the unit cell design and HFSS-based setup. The validation of the structure is then presented in section 3.

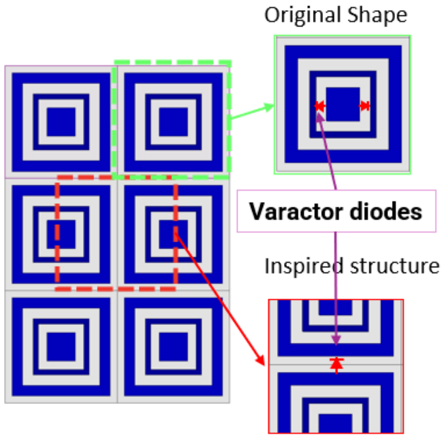


Figure 1: Structural design of single- and dual-varactor Phoenix-inspired unit cells within the complete reflectarray antenna.

2 Simulation Setup

This study focuses on extending this concept to a reconfigurable dual-varactor cell 16 x 16 mm in size, whose performance was evaluated using electromagnetic simulations under HFSS. Simulations were conducted in a central frequency of 10 GHz to accurately model an infinite periodic by exciting the unit cell with a Floquet port while imposing periodic (master-slave) boundaries conditions on the lateral faces under a normal incidence, as presented in Fig. 2. Perfect conductor assumption (PEC) was adopted. The metal patch was printed on a 0.508 mm thick metalized Rogers RO4003 substrate with a relative permittivity $\varepsilon_r = 3.55$ and a loss tangent $\tan\delta = 0.0027$.

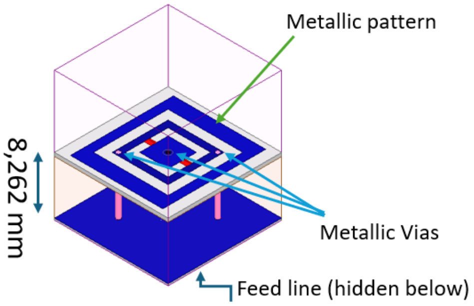


Figure 2: Illustration of the unit cell boundaries with periodic boundary conditions applied.

3 Results and Discussion

The dual-varactor structure unit cell exhibits a wide range and near-linear tuning law at 10 GHz, delivering approximately 281 degrees of continuous phase control, with a capacitance ranging from 0.02 to 0.2 pF as shown in Fig. 3. To span the required capacitance range, the MAVR-011020-1411 was adopted as a varactor device [5] These low-cost varactors introduced losses of 4 dB maximum on the reflected wave because of their internal resistance (R= 0.3 Ohm) and inductance (L= 0.4 nH) and of their DC biasing circuits.

4 Conclusion

This study has introduced the inspiration of two reconfigurable unit cells designed for X-band weather radar applications, with low-losses and fabrication simplicity suitable for low-cost deployable weather surveillance radar reflectarrays. This paper focuses on the dual-varactor diodes cell version that reach an acceptable phase-shifting range of 281 degrees, with reflection losses kept below 4 dB mainly due to the varactor diodes and DC biasing network. It achieve a nearly linear and parallel phase response across 10 GHz. This design is the last step toward developing a complete dual-polarization reflectarray antenna.

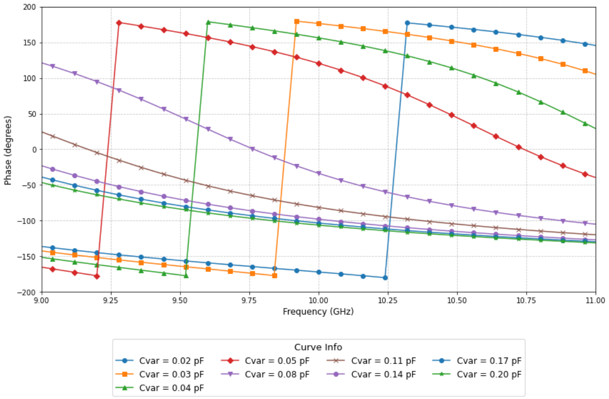


Figure 3: Phase variation of the proposed unit cell under normal incidence, across the practical range of the selected varactor capacitance values.

References

[1] John Huang, Jose Antonio Encinar “Reflectarray Antennas”, Book Type: Wiley-IEEE Press, 2008.

[2] Moustafa, et.al. “The Phoenix Cell: A New Reflectarray Cell With Large Bandwidth and Rebirth Capabilities”, *IEEE Antennas and Wireless Propagation Letters*, pp. 1643-16440, 2018.

[3] Makdissy, Tony and Gillard, Ranhaël. “Single Layer Dual-Band ‘Phoenix’ Reflectarray Unit Cell with Dual-Linear Polarization”, *IEEE International Symposium on Antennas and Propagation USNC/URSI National Radio Science Meeting*, pp. 1643-16440, 2018.

[4] Dendani, Ikram and Fourn, Erwan and Dakhli, Saber and Choubani, Fethi. “A Phoenix Unit Cell with Linear Polarization for Weather Radar Applications,” *2025 19th European Conference on Antennas and Propagation (EuCAP)*, pp. 1-4, 2025.

[5] https://www.macom.com/products/product-detail/MAVR-011020-1411

Cooperative Repeated Game with Q-Learning for Fair Multi-User Underwater Acoustic Networks

Khouloud GHARSALLI[‡][†], Sameh NAJEH[‡], Thierry LE PORS[†],
Pierre-Jean BOUVET[†], Hichem BESBES[‡]
[‡]COSIM, SUP'COM, Carthage University, Arianna, Tunisia
[†]LabISEN, ISEN Yncréa Ouest, Brest, France
{khouloud.gharsalli, sameh.najeh, hichem.besbes}@supcom.tn
{khouloud.gharsalli, thierry.le-pors, pierre-jean.bouvet}@isen-ouest.yncrea.fr

Abstract

This paper proposes a cooperative resource allocation framework for multi-user underwater acoustic (UWA) networks by jointly leveraging repeated game theory and Q-learning. Repeated interactions allow the detection of selfish behaviors and the enforcement of long-term cooperation through punishment strategies, while a Q-learning approach with centralized training and decentralized execution (CTDE) enables adaptive power and subcarrier allocation under dynamic underwater channel conditions. A composite reward function balancing throughput, fairness, and energy consumption is introduced. Simulation results under a realistic UWA channel model demonstrate significant improvements in cooperation level, global utility, fairness, and convergence speed compared to non-cooperative and classical benchmark schemes.

Keywords: Underwater acoustic networks, repeated games, Q-learning, cooperative resource allocation, fairness.

1 Introduction

Underwater acoustic (UWA) networks play a critical role in maritime surveillance, environmental monitoring, autonomous underwater vehicle (AUV) coordination, and offshore exploration. Unlike terrestrial wireless networks, UWA communications suffer from extremely limited bandwidth, long propagation delays due to the slow speed of sound in water, severe frequency-dependent attenuation, and strong multipath effects [1]. These unique characteristics significantly degrade communication reliability and make efficient resource allocation particularly challenging. As a result, power control, spectrum allocation, fairness, and energy efficiency must be jointly considered when designing UWA communication systems.

A considerable amount of prior work has investigated resource allocation in UWA networks using centralized optimization or heuristic-based approaches. While these methods can achieve good performance under static conditions, they often require full channel state information and incur high signaling overhead, which is impractical in dynamic underwater environments. To reduce complexity, non-cooperative game-theoretic formulations have also been proposed, where each user selfishly maximizes its own utility. However, such approaches typically converge to inefficient Nash equilibria, leading to poor global performance in interference-limited and bandwidth-constrained UWA networks.

To overcome these limitations, cooperative and bargaining-based game models have been introduced to improve system efficiency and fairness [2]. Although cooperation enhances global performance, it usually requires explicit coordination and information exchange among users, which can be costly in underwater networks characterized by long delays and limited feedback capabilities.

Repeated game theory has emerged as an effective framework to enforce cooperation among selfish agents without requiring full centralization. By exploiting repeated interactions, agents can detect non-cooperative behavior and apply punishment strategies, thereby sustaining long-term cooperation. Such mechanisms have been widely applied to power control, spectrum sharing, and access coordination in terrestrial wireless networks [3]. However, despite their demonstrated effectiveness in terrestrial systems, there are relatively few studies exploring repeated game models specifically for multi-user underwater acoustic (UWA) networks, where interactions are naturally long-term and delay-tolerant but subject to severe bandwidth limitations and high propagation delays.

In parallel, reinforcement learning (RL), and in particular Q-learning, has gained significant attention for adaptive resource allocation in dynamic and partially unknown environments. RL-based approaches allow agents to learn optimal policies directly from interaction with the environment, without relying on explicit channel models. Recent studies have demonstrated the potential of RL for underwater communications, showing improved adaptability to slow-varying and highly uncertain UWA channels. Moreover, it has been shown that combining game-theoretic frameworks with reinforcement learning can naturally promote cooperative behavior, as agents learn to balance short-term gains with long-term rewards [4].

Despite these advances, the joint integration of repeated game mechanisms and learning-based resource allocation for multi-user UWA networks remains largely unexplored. In particular, how to enforce cooperation, ensure fairness, and adapt resource allocation under realistic underwater channel conditions is still an open research challenge.

Contributions: Motivated by these observations, this paper proposes a novel cooperative resource allocation framework that integrates repeated game theory with Q-learning for multi-user UWA networks. The main contributions of this work are summarized as follows:

- We formulate the multi-user UWA resource allocation problem as a repeated game, enabling the detection of selfish power allocation behaviors and the enforcement of punishment strategies to sustain long-term cooperation under severe bandwidth and propagation delay constraints.
- We develop a Q-learning-based solution with centralized training and decentralized execution (CTDE), which mitigates multi-agent non-stationarity and enables adaptive power and subcarrier allocation over slow-varying, frequency-selective UWA channels.
- We design a composite reward function that jointly accounts for spectral efficiency, Jain's fairness index, and energy consumption, addressing key performance requirements of long-term underwater deployments.
- Through extensive simulations under a realistic UWA channel model, we demonstrate that the proposed framework significantly outperforms non-cooperative learning approaches and classical proportional fairness benchmarks in terms of system throughput, fairness, and convergence speed.

The remainder of this paper is organized as follows. Section 2 describes the system model and problem formulation. Section 3 presents the proposed repeated game and Q-learning framework. Simulation results are discussed in Section 4, and Section 5 concludes the paper.

2 Problem Formulation and Underwater Acoustic Channel Model

The underwater acoustic channel is characterized by frequency-dependent attenuation, long propagation delays, and severe multipath propagation. The channel gain $h_{i,n}(t)$ incorporates spreading loss, absorption loss, and slow temporal variations, following the statistical model proposed in [5]. This model captures the specific physical characteristics of underwater propagation and distinguishes the considered scenario from terrestrial wireless channels.

We consider an underwater acoustic (UWA) network composed of K users, indexed by $i \in \{1, \ldots, K\}$, sharing N_{sc} orthogonal subcarriers indexed by $n \in \{1, \ldots, N_{\text{sc}}\}$. Each user dynamically allocates its transmission power $p_{i,n}(t)$ over the available subcarriers at iteration t in order to maximize its achievable utility.

The utility of user i is defined as

$$u_i(\mathbf{P}, t) = \frac{1}{N_{\text{sc}}} \sum_{n=1}^{N_{\text{sc}}} \log_2\left(1 + \text{SINR}_{i,n}(\mathbf{P}, t)\right), \tag{1}$$

where $\mathbf{P} = \{p_{i,n}\}_{i,n}$ denotes the global power allocation vector.

Each user is subject to a total transmit power constraint:

$$0 \leq \sum_{n=1}^{N_{\text{sc}}} p_{i,n}(t) \leq P_{\max}, \quad \forall i, t, \tag{2}$$

where $P_{\max}$ is the maximum allowable transmit power.

The signal-to-interference-plus-noise ratio (SINR) experienced by user i on subcarrier n is given by

$$\text{SINR}_{i,n}(\mathbf{P}, t) = \frac{p_{i,n}(t)\, h_{i,n}(t)}{\sum_{j \neq i} p_{j,n}(t)\, h_{j,i,n}(t) + \sigma^2}, \tag{3}$$

where $h_{i,n}(t)$ denotes the direct UWA channel gain of user i, $h_{j,i,n}(t)$ represents the interference channel gain from user j to user i, and σ^2 is the noise power.

2.1 Objective and Reward Function

Each agent aims to maximize its long-term cumulative reward over a repeated-game horizon of T_{rep} iterations:

$$\max_{\mathbf{p}_i} \sum_{t=1}^{T_{\text{rep}}} R_i(t). \tag{4}$$

The instantaneous reward $R_i(t)$ balances spectral efficiency, fairness, and energy consumption, and is defined as

$$\begin{aligned} R_i(t) = &\underbrace{W_{\text{RATE}} \frac{u_i(\mathbf{P}, t)}{u_{\max}}}_{\text{Rate utility}} + \underbrace{W_{\text{FAIRNESS}}\, F(t)}_{\text{Fairness}} \\ &- \underbrace{W_{\text{ENERGY}} \left(\frac{P_i(t)}{P_{\max}}\right)^2}_{\text{Energy penalty}}, \end{aligned} \tag{5}$$

where $P_i(t) = \sum_{n=1}^{N_{\text{sc}}} p_{i,n}(t)$ is the total transmit power of user i, and $u_{\max}$ is a normalization constant.

$W_{\text{RATE}}, W_{\text{FAIRNESS}}, W_{\text{ENERGY}} \in [0, 1]$ control the tradeoff between throughput, fairness, and energy efficiency, and satisfy $W_{\text{RATE}} + W_{\text{FAIRNESS}} + W_{\text{ENERGY}} = 1$.

Fairness among users is measured using Jain's fairness index:

$$F(t) = \frac{\left(\sum_{i=1}^{K} u_i(\mathbf{P}, t)\right)^2}{K \sum_{i=1}^{K} u_i^2(\mathbf{P}, t)}, \quad 0 < F(t) \leq 1. \tag{6}$$

3 Proposed Repeated Q-Learning Framework

3.1 State, Action, and Q-Table Structure

Each agent follows a reinforcement learning policy based on a finite state and action space.

Action space $\mathcal{A}$: At each iteration, an agent selects one subcarrier and one discrete transmit power level from a predefined set.

State space $\mathcal{S}$: The state observed by agent i at iteration t is defined as

$$s_i(t) = \big(h_i(t), I_i(t)\big), \tag{7}$$

where $h_i(t)$ and $I_i(t)$ denote the quantized direct channel gain and quantized interference level, respectively.

3.2 Repeated Game Principle

The system is modeled as a repeated game in which agents interact over T_{rep} iterations. At each iteration t, an agent observes the current state s_t, selects an action a_t, receives a reward $R_i(t)$, and generates interference that affects the next state s_{t+1}. This repeated interaction enables agents to detect selfish behaviors and learn cooperative strategies over time.

3.3 Q-Learning Update with CTDE

To mitigate non-stationarity in the multi-agent environment, a *Centralized Training, Decentralized Execution* (CTDE) approach is adopted. Each agent maintains a local Q-table Q_i, while a global critic aggregates information during training:

$$Q_{\text{global}}(s, a) = \frac{1}{K} \sum_{i=1}^{K} Q_i(s, a). \tag{8}$$

The Q-learning update rule for agent i is given by

$$Q_i(s_t, a_t) \leftarrow (1 - \alpha_t)\, Q_i(s_t, a_t) + \alpha_t \Big[R_i(t)$$
$$+ \gamma \max\nolimits_a Q_{\text{global}}(s_{t+1}, a) \Big], \ (9)$$

where α_t is the learning rate and γ is the discount factor.

4 Simulation Setup and Results

4.1 Simulation Setup

Simulations consider $K = 20$ users and $N_{\text{sc}} = 128$ subcarriers over $T_{\text{rep}} = 2000$ iterations. The discount factor is set to $\gamma = 0.9$, and the initial exploration rate is $\epsilon_0 = 1.0$. The UWA channel follows the statistical model described in [5].

4.2 Results: Comparative Performance Analysis

Figure 1 illustrates the evolution of the system's aggregated rate (Total Rate) over the iterations for the different approaches.

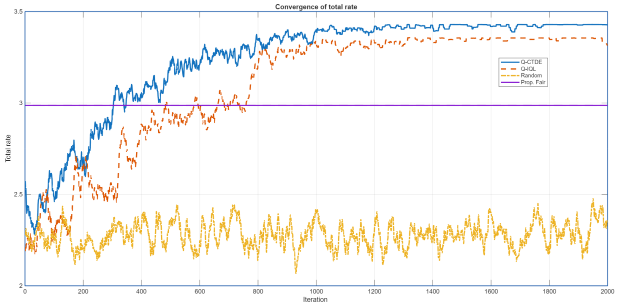


Figure 1: Convergence of total rate (Total Rate)

- **Q-CTDE (Proposed):** This approach achieves the highest overall rate, stabilizing at approximately 3.4. The convergence is relatively fast, occurring around 1000 iterations. The superior performance can be attributed to two key factors: (i) the repeated game structure, which enforces cooperation among agents and discourages selfish behavior through long-term reward shaping, and (ii) the centralized aggregation of Q-values (CTDE), which mitigates the non-stationarity inherent in multi-agent environments and allows agents to better coordinate subcarrier and power allocation. As a result, Q-CTDE exploits the spectrum more efficiently and reaches a cooperative equilibrium that maximizes global utility.

- **Q-IQL (Independent Q-Learning):** Achieves high performance but is lower than Q-CTDE, stabilizing. The fact that each agent learns independently leads to a Nash equilibrium that is globally less efficient than the cooperative optimum found by Q-CTDE. Its convergence is also slower (stabilizing beyond 1400 iterations).

- **Prop. Fair (Proportional Fair):** The reference level for the Proportional Fair solution is 2.8. Both learning schemes (Q-CTDE and Q-IQL) significantly outperform this classical benchmark.

- **Random:** As expected, the random approach maintains a very low and fluctuating rate, situated around 2.2.

Figure 2 shows the evolution of Jain's Fairness Index (which measures fairness among users, with 1 being perfect fairness) over the iterations.

- **Q-CTDE (Proposed):** Achieves the best fairness, converging toward a value close to 1.0 (approximately 0.99), indicating near-perfect fairness among users. This validates the effectiveness of the penalty term in the reward function and the cooperative effect of the repeated game mechanism.

- **Q-IQL (Independent Q-Learning):** Also achieves a very good level of fairness, stabilizing around 0.97. Although its rate was slightly lower than Q-CTDE, its

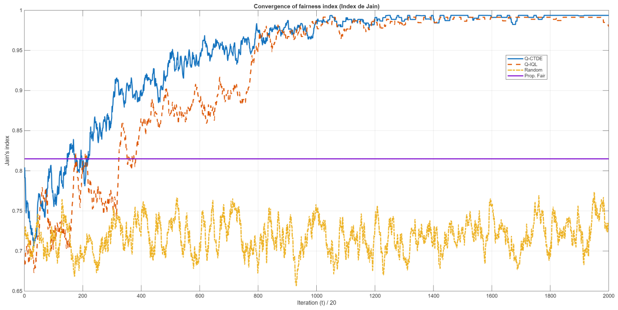


Figure 2: Convergence of Jain's Index

fairness level is very close, showing that even independent learning manages to achieve a good distribution of resources.

- **Prop. Fair (Proportional Fair):** The Proportional Fair fairness reference is set at 0.815. Again, the learning schemes (Q-CTDE and Q-IQL) achieve fairness levels significantly superior to this benchmark.
- **Random:** The random approach exhibits low and fluctuating fairness, settling around 0.72, confirming that without intelligent optimization, fairness is not guaranteed.

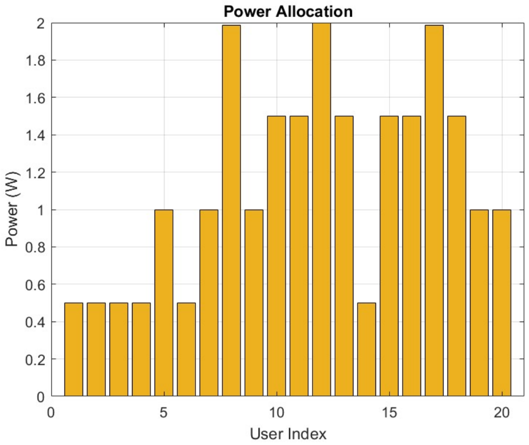


Figure 3: Distribution of allocated power

The power allocation presented by Figure 3 provides further insight into the strategy learned by the system. Power is allocated in a differentiated manner between users: some users receive higher power levels—up to 2 W—while others consume as little as 0.5 W. This dynamic allocation reflects the adaptive behavior of the Q-learning algorithm, where users experiencing poor channel conditions are granted more transmission power to maintain acceptable data rates. Conversely, those with better channel quality can sustain satisfactory performance with reduced power. This demonstrates the energy-aware and throughput-optimized nature of the policy, contributing to the overall system efficiency.

The simulation results clearly demonstrate that the proposed Q-CTDE framework provides the best performance across both key metrics: Total Rate and Jain's Index. It maximizes spectral efficiency while guaranteeing the highest level of fairness and achieving faster convergence compared to the non-cooperative Q-IQL approach.

5 Conclusion

The proposed framework integrates repeated game theory with Q-learning to enhance cooperation and optimize resource allocation in multi-user underwater acoustic networks. By modeling user interactions as a repeated game, the approach enables the detection and penalization of selfish behavior, thereby encouraging sustained cooperation. At the same time, Q-learning allows agents to adapt power and subcarrier allocation strategies to dynamic underwater channel conditions without requiring explicit channel models. Simulation results under realistic UWA environments show that the proposed solution achieves higher system utility, improved fairness, and stable convergence, making it a practical and scalable approach for underwater acoustic communications.

References

[1] Akyildiz, I.F.; Pompili, D.; Melodia, T. Challenges for efficient communication in underwater acoustic sensor networks. *ACM Sigbed Rev.* **2004**, *1*, 2, 3–8. ACM: New York, NY, USA.

[2] D. Muhammed, M. H. Anisi, M. Zareei, C. Vargas-Rosales, and A. Khan, "Game theory-based cooperation for underwater acoustic sensor networks: Taxonomy, review, research challenges and directions," *Sensors*, vol. 18, no. 2, p. 425, 2018.

[3] D. T. Hoang, X. Lu, D. Niyato, P. Wang, D. I. Kim, and Z. Han, "Applications of repeated games in wireless networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2102–2135, Oct. 2015.

[4] Ding, Z.-W.; Zheng, G.-Z.; Cai, C.-R.; Cai, W.-R.; Chen, L.; Zhang, J.-Q.; Wang, X.-M. Emergence of cooperation in two-agent repeated games with reinforcement learning. *Chaos, Solitons & Fractals*, **2023**, *175*, 114032.

[5] Qarabaqi, P.; Stojanovic, M. Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels. *IEEE Journal of Oceanic Engineering*, **2013**, *38*(4), 701–717.

The Design of a miniature circularly polarized antenna for CubeSat

Najet Bougdima[1], Mondher Dhaouadi[2], Fethi Choubani[3]

[1]PhD Student, University of Carthage, Sup'COM, Innov'COM Laboratory, Tunisia
[2]Assistant Professor, University of Carthage,Sup'COM, Innov'COM Laboratory, Tunisia
[3]Professor, University of Carthage, Sup'COM, Innov'COM Laboratory, Tunisia
`najat.bougdima@supcom.tn`, `mondher.dhaouadi@isimg.tn`, `fethi.choubani@supcom.tn`

Abstract

This paper presents a compact printed circularly polarized antenna for 1U CubeSat applications. Deployed in orbit by a miniaturized actuator to a separation of $\lambda/10$ (34.5 mm), it reduces coupling with the satellite structure. Optimized for nadir-pointing radiation, the $100 \times 100 \times 1.6$ mm^3 antenna achieves a maximum gain of 6.12 dB and an axial ratio below 3 dB over 853–879 MHz, demonstrating effective circular polarization and suitability for CubeSat missions.

Keywords: CubeSat, printed antenna, circular polarization, compact design, S-band, axial ratio, gain, space communication.

1 Introduction

Satellite-based IoT requires compact, robust, high-performance antennas [1]. Miniature circularly polarized antennas maintain reliable connections regardless of device orientation and provide high gain for bidirectional communication with satellites. This work presents a high-gain, circularly polarized printed antenna designed for CubeSat applications.

2 Results

Our antenna prototype was simulated on an FR4 substrate with a permittivity of $\varepsilon_r = 4.4$ and a thickness of $h = 1.6$ mm.

In Fig. 2, the antenna is shown separated by 34.5 mm from the CubeSat after actuator deployment. The design focuses radiation toward Earth, maximizing gain for ground stations while maintaining circular polarization. Future work will integrate a miniaturized actuator and study the deployment mechanism for reliable in-orbit operation.

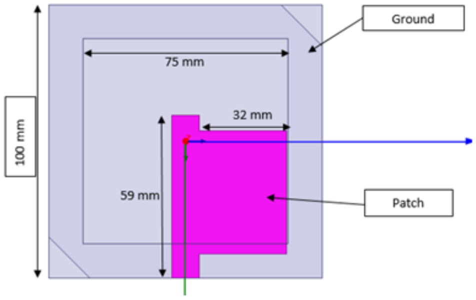


Figure 1: Top view of the designed patch antenna.

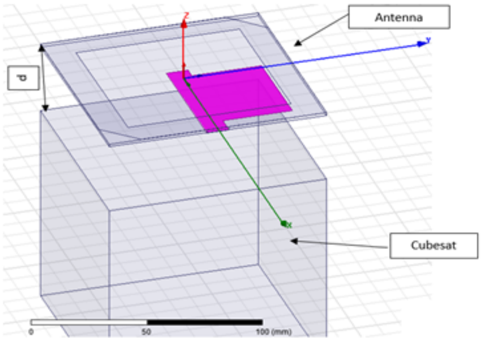


Figure 2: Placement of the proposed patch antenna on the CubeSat in orbit (Top view).

Fig. 3 shows the simulated reflection coefficient. At 871 MHz, a 17 dB return loss is achieved with a 34.5 mm separation, reducing electromagnetic coupling with the CubeSat structure.

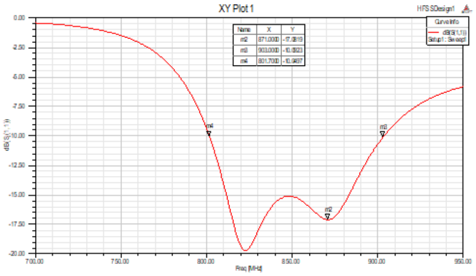


Figure 3: Simulated reflection coefficient of the proposed antenna.

Circular polarization is required in satellite applications to avoid up to 3 dB of mismatch loss and to maintain signal quality when the CubeSat rotates or tumbles in orbit. Optimizing the antenna's axial ratio along the desired direction improves power transfer and link reliability.

Fig. 4 shows the simulated axial ratio (AR) versus frequency along $\theta = \phi = 0°$. The 3 dB AR bandwidth is 25.5 MHz (853.5–879 MHz), covering the target frequency range and accommodating potential Doppler shifts.

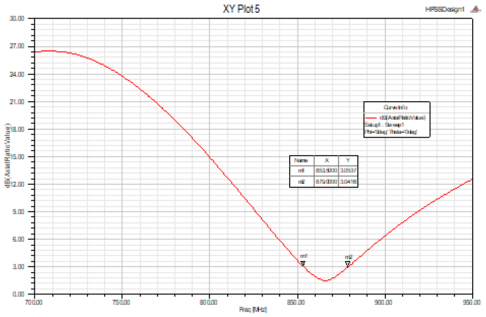


Figure 4: Simulated axial ratio (AR) of the proposed antenna.

Antenna gain is related to directivity, with higher gain achieved when energy is concentrated in the main lobe. Fig. 6 shows the 3D radiation pattern, with a peak gain of 6.12 dB, demonstrating efficient radiation. Such directivity is essential for CubeSat applications to maximize transmitted power.

3 Future Perspectives

Future work may focus on metamaterial-based antennas that can be mounted directly on the CubeSat

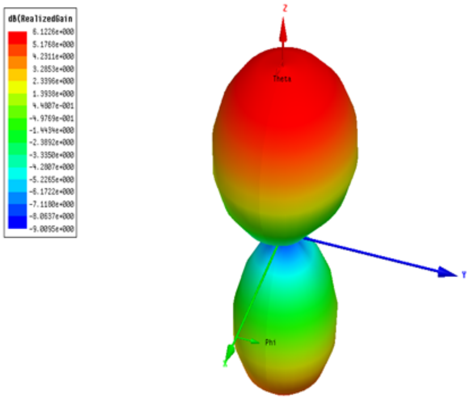


Figure 5: 3D gain pattern of the proposed antenna.

surface, reducing volume and simplifying integration while maintaining performance for a 100 × 100 mm patch [2].

Deployable mechanisms, such as miniaturized actuators or arms, can extend the antenna post-launch. While separation of $\lambda/10$ reduces coupling with the metallic structure, CubeSat space is limited. AMC/EBG metamaterials enable direct surface mounting with high radiation efficiency, and deployable arms can further enhance gain or directivity while minimizing stowed volume.

4 Conclusion

The proposed antenna provides circular polarization with an axial ratio below 3 dB over 853–879 MHz, ensuring stable communication regardless of satellite orientation. Its simulated peak gain of 6.12 dB delivers sufficient power, and its compact $100 \times 100 \times 1.6$ mm^3 size allows easy integration into a 1U CubeSat.

References

[1] S. Abulgasem, F. Tubbal, R. Raad, P. I. Theoharis, S. Lu and S. Iranmanesh, "Antenna Designs for CubeSats: A Review," *IEEE Access*, vol. 9, pp. 45289–45324, 2021, doi: 10.1109/ACCESS.2021.3066632.

[2] T. Saeidi, S. Saleh, and S. Karamzadeh, "Miniaturized and Circularly Polarized Dual-Port Metasurface-Based Leaky-Wave MIMO Antenna for CubeSat Communications," *Electronics*, vol. 14, no. 14, p. 2764, 2025. doi: 10.3390/electronics14142764.

Modeling and Low-Complexity Battery-Aware Energy Management for a Residential Microgrid

Henda Ben Mansour[1],Lotfi Charaabi[2],Khaled Jelassi[3]

[1,2,3]Université de Tunis EL Manar,ENIT, Laboratoire des Systèmes Electriques,LR11ES15
1002 Tunis, Tunisie,
[1]henda.benmansour@enit.utm.tn; [2]lotfi.charaabi@enit.utm.tn; [3]khaled.jelassi@enit.utm.tn

Abstract

This paper presents an efficient and low-complexity energy management strategy (EMS) for a residential microgrid composed of a photovoltaic (PV) array, a battery (BT), and the utility grid. The proposed EMS optimizes the use of PV generation to reduce electricity costs and limit rapid variations in battery power, thereby enhancing overall system reliability. The control algorithm manages energy flow between components, ensuring stable operation across different power flow modes. The microgrid was modeled and simulated in MATLAB/Simulink, and the results demonstrate stable operation and show the effectiveness of the proposed EMS in reducing grid dependency and limiting battery power fluctuations.

Keywords: Residential Microgrid, Energy Management Strategy, Photovoltaic, Li-ion Battery.

1 Introduction

In recent years, microgrids have gained significant attention as an effective solution for integrating renewable energy sources into residential power systems. Their ability to operate in both grid-connected and islanded modes enhances reliability and resilience during disturbances. Among renewable energy sources, photovoltaic (PV) systems are particularly attractive for residential applications due to their sustainability and decreasing installation costs. However, the intermittent and weather-dependent nature of solar energy limits the ability of PV generation to continuously meet load demand. To address this issue, energy storage systems (ESS), especially battery-based solutions, are commonly integrated into residential microgrids. Batteries provide energy balancing over extended periods, improve system reliability, and enable higher utilization of locally generated renewable energy[4-8]. As a result, efficient energy management strategies (EMS) are required to coordinate power flows between PV generation, storage systems, loads, and the utility grid, with the objectives of maximizing renewable energy utilization, ensuring stable operation, and reducing electricity costs. Several studies have investigated EMS approaches for residential and hybrid microgrids, including coordinated control schemes, stochastic optimization, and supervisory control architectures, demonstrating the critical role of energy management in improving microgrid performance and economic efficiency[9-10]. Recently, advanced EMS approaches based on Model Predictive Control (MPC) and artificial intelligence techniques, such as reinforcement learning, have been widely explored[1–3]. Although these methods can achieve high performance, they generally require accurate forecasting, high computational resources, and complex tuning, which limit their practical deployment in low-cost residential systems. In contrast, rule-based EMS remain attractive due to their simplicity, robustness, and ease of real-time implementation.

In this paper, we propose a residential microgrid configuration combining PV generation, battery storage, and the utility grid. The proposed EMS focuses on reducing grid dependency, improving photovoltaic energy utilization, and explicitly limiting battery power fluctuations while maintaining cost-effective energy management.

2 System Description and Modeling

Figure 1 presents the configuration of the proposed residential microgrid system, which integrates multiple distributed energy sources and load units connected to both DC and AC buses. The interconnection between the AC and DC networks is achieved through a bidirectional AC/DC converter, enabling flexible power exchange between the two sides. The overall microgrid architecture is modeled and simulated in MATLAB/Simulink to analyze the dynamic behavior, power flow, and control performance of the system. In this configuration, photovoltaic (PV) arrays are interfaced with the DC bus via a DC/DC boost converter, supplying power to the AC loads through the bidirectional AC/DC converter. A battery energy storage system is also linked to the DC bus through a bidirectional DC/DC converter, providing both charging and discharging functionalities depending on system conditions. The utility grid and AC loads are connected to the AC bus, facilitating energy exchange with the utility grid when necessary.

In this microgrid framework, the battery serves as a key component for maintaining power balance among distributed generation units and load demands, as well as for enhancing voltage stability across the system. The control objectives of various power electronic converters are coordinated by an energy management system, which optimizes the operation of each subsystem. The main converter is responsible for maintaining a stable and high-quality AC bus voltage, while the PV subsystem operates under Maximum Power Point Tracking (MPPT) control to ensure efficient and continuous energy extraction under varying environmental conditions.

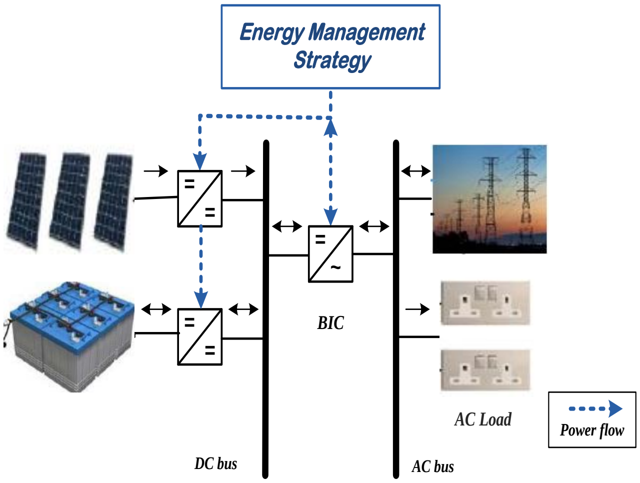


Figure 1: Structure of Residential Microgrid.

2.1 PV Model

To maximize efficiency under varying sunlight and temperature conditions, the PV strings are interfaced with a DC/DC converter controlled by an MPPT algorithm. The total power P_{PV} generated by the PV surface S_{pv} is monitored, as shown in Figure 2, and the overall PV efficiency η_{PV} is calculated using the following equations [10-11]:

$$\begin{cases} P_{pv}(T_a, G_{tot}) = \eta(T_c, G_{tot}) S_{pv} G_{tot} \\ \eta_{pv}(T_c, G_{tot}) = \eta_{manuf}(1 - \beta_{pv}(T_c - T_r)) \\ T_c = T_a + (T_{NOCT} - T_{a,NOCT}) \dfrac{G\ tot}{G\ NOCT} \end{cases} \quad (1)$$

Where T_a is the ambient temperature, G_{tot} is the total solar radiation received by the PV panel, T_c the cell temperature, G_{NOCT} is the nominal solar radiation, T_r is the reference temperature, β_{PV} is the temperature coefficient, and η_{manuf} is the nominal efficiency.

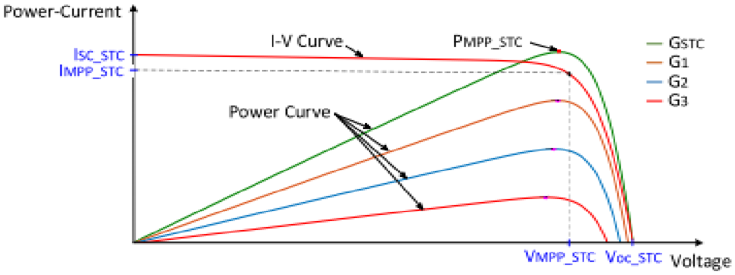


Figure 2: PV characteristic at different solar irradiance levels (25°C)[12].

2.2 Battery Model

To ensure the simulation stability, a filtered battery current, instead of the actual battery equivalent, is used to account for the polarization resistance[13]. The model parameters are derived from datasheets or simple dynamic tests. From Figure3, the battery voltage can be expressed as follows [14]:

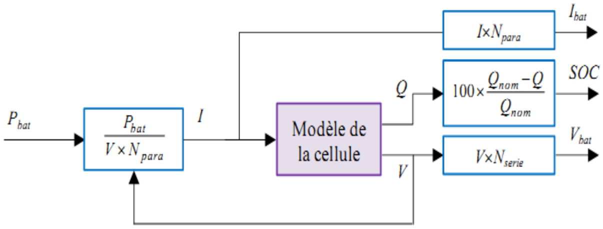


Figure 3: Li-ion Battery Model.

$$\begin{cases} I > 0 \rightarrow V = E_0 - R.I - K \dfrac{Q_{nom}}{Q_{nom} - it}.it - KI^{*}.\dfrac{Q_{nom}}{Q_{nom} - it} + Ae^{-Bit} \\ I < 0 \rightarrow V = E_0 - R.I - K \dfrac{Q_{nom}}{Q_{nom} - it}.it - KI^{*}.\dfrac{Q_{nom}}{it - 0.1Q_{nom}} + Ae^{-Bit} \end{cases} \quad (2)$$

where E_0 is the battery constant voltage (V), K is the polarization constant (V/Ah), Q_{nom} is the nominal battery capacity (Ah), I^* is the filtered battery current (A), A is the exponential zone amplitude (V), B is the exponential zone time constant inverse (Ah^{-1}) and R is the battery internal resistance (Ω). Equation (2) recalls the equations of the Shepherd model for a Li-ion accumulator distinguishing between the charge phase (I <0) and the discharge phase (I> 0).

3 Energy Management Strategy

Present results and The energy flow control in the residential microgrid is divided into two main operating modes: the grid-connected mode and the islanded mode. In the grid-connected mode, two distinct energy flow paths can be identified, while in the islanded mode, two additional energy flow paths are also considered, as illustrated in Figure 4. These operating scenarios will be discussed in detail in the following sections. The control of the different energy flows within the residential microgrid is determined by four main input variables P_{pv}: Power supplied by the photovoltaic (PV) array.

P_{grid}: Power exchanged with the main distribution grid. P_{bat}: Power available from the battery storage system. P_{ld}: Power demanded by the AC load, which consists of two loads, Load 1 and Load 2, with respective powers P_{ld1} and P_{ld2}, such that: $P_{ld} = P_{ld1} + P_{ld2}$.

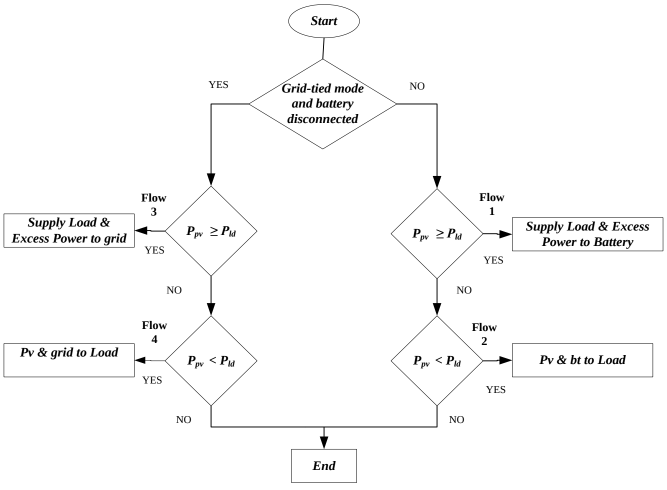


Figure 4: The proposed Energy Management Strategy.

The first mode is the islanded mode where the microgrid and relies on the energy storage system (battery) to ensure energy balance between PV energy production and load demand. In this mode, there are two energy flows as follows:

Flow 1:PV power is greater than or equal to load demand ($P_{pv} \geq P_{ld}$). If the PV power is equal to the load demand ($P_{pv} = P_{ld}$), the PV power only supplies the loads. If P_{pv} is strictly greater than Pld, the battery absorbs the excess power ($P_{bat} = P_{pv} - P_{ld}$).

Flow 2: PV power is less than the load demand ($P_{pv} < P_{ld}$). If P_{pv} is less than P_{ld}, then the battery will supply the loads ($P_{ld} = P_{pv} + P_{bat}$).

In the second mode, where the microgrid is connected to the utility grid, it is based on two power flows as follows:

Flow 3: PV power is greater than or equal to load demand ($P_{pv} \geq P_{ld}$). Flow 3 is identical to flow 1 when the load demand is satisfied by P_{pv}, if the power ($P_{pv} = P_{ld}$), P_{pv} only supplies the loads. If the power P_{pv} is greater than the power P_{ld}, then the excess power is injected into the grid ($P_{grid} = P_{pv} - P_{ld}$).

Flow 4: PV power is less than the load demand ($P_{pv} < P_{ld}$). when the power P_{pv} is less than the power P_{ld}, the loads will be supplied by the utility grid ($P_{ld} = P_{pv} + P_{grid}$).

The main contribution of the proposed EMS lies in its low computational complexity and its explicit consideration of battery power smoothing, making it suitable for real-time residential microgrid applications.

4 Results and Discussion

To validate the proposed design and evaluate the effectiveness of the implemented energy management strategies, the residential microgrid system was modeled using the MATLAB/Simulink environment. During the simulation tests, the power was defined as P_{pv} = 5 kW for the photovoltaic generation unit and P_{ld} = 6 kW for the total load, divided into P_{ld1} = 3 kW and P_{ld2} = 3 kW).

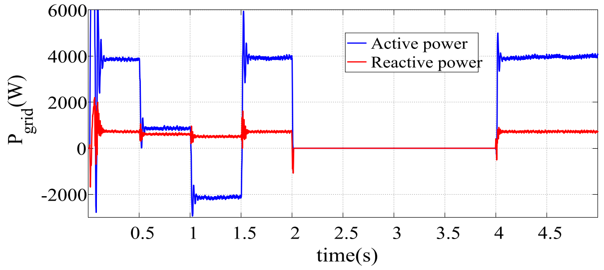


Figure 5:Power exchanged with the utility grid.

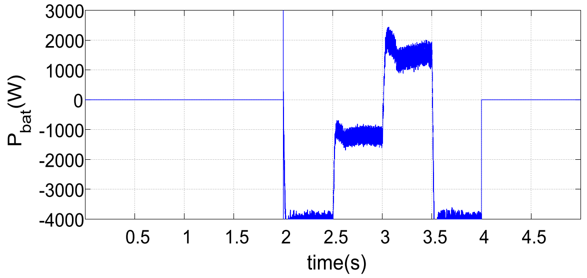


Figure 6:Power exchanged with battery.

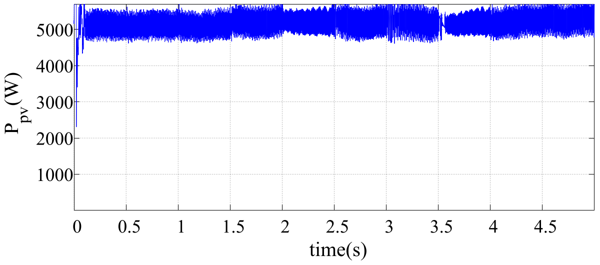


Figure 7:Power of solar photovoltaic.

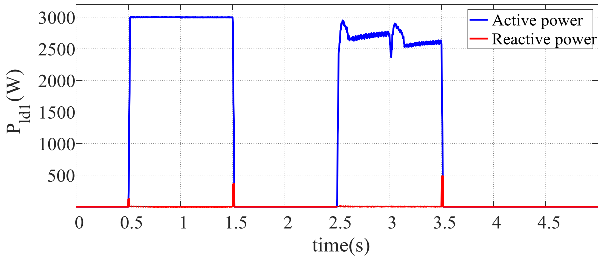


Figure 8:Active power of Load 1 (P_{ld1}).

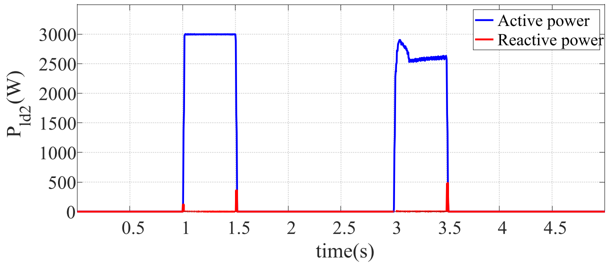


Figure 9: Active power of Load 2 (Pld2).

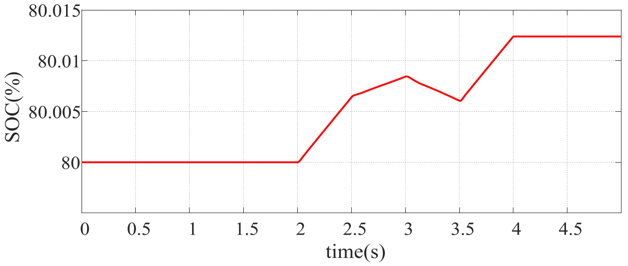


Figure 10:State of charge (SOC) of the Battery.

Mode 1

Energy flow from PV to battery:

At t = {2–2.5 s, 3.5–4 s}, with P_{pv} = 5 kW and P_{ld} = 0 kW, as illustrated in Figures 6 and 10, the power generated by the PV system is entirely absorbed by the battery.

Energy flow from PV to battery and load:

At t = {2.5–3 s}, when P_{pv} = 5 kW and P_{ld} = 3 kW, the PV array supplies the load, and the excess PV power is used to charge the battery, as shown in Figure 10.

Energy flow from PV and battery to load:

At t = {3–3.5 s}, when P_{pv} = 5 kW and Pload = 6 kW, both the PV system and the battery coordinate to supply the load. The battery discharges to meet the additional power demand, as illustrated in Figure 10.

Mode 2

Power flow from PV to the grid:

At t = {0–0.5 s, 1.5–2 s, 4–5 s}, with P_{pv} = 5 kW and P_{ld} = 0 kW, the power generated by the PV system is injected into the utility grid, as shown in Figure 5.

Power flow from PV to grid and load:

At t = {0.5–1 s}, when P_{pv} = 5 kW and P_{ld} = 3 kW, the PV system supplies the load, and the surplus power is injected into the grid, as illustrated in Figure 5.

Power flow from grid and PV to load:

At t = {1–1.5 s}, when P_{pv} = 5 kW and P_{ld} = 6 kW, both the PV system and the grid operate in coordination to supplies the load.

5 Conclusion

This work presented an energy management strategy for a residential hybrid microgrid integrating photovoltaic generation, battery storage, and the utility grid. The proposed strategy effectively coordinates energy flows between these sources to ensure continuous and reliable power supply while improving overall system efficiency. Simulation results confirmed that the integration of battery storage enhances energy stability and reduces grid dependency. Future work will focus on experimental validation and comparison with advanced EMS approaches, such as Model Predictive Control (MPC) based strategies, to further evaluate the trade-off between performance and computational complexity.

Acknowledgment

This paper has been revised with Gemini support.

References

[1] A. Parisio and al, "Model predictive control for residential microgrids: Recent advances and challenges," Renewable and Sustainable Energy Reviews, vol. 181, pp. 113–128, 2023.

[2] Y. Wang and al, "Reinforcement learning–based energy management of residential microgrids: A review," Applied Energy, vol. 343, pp. 121–139, 2024.

[3] Qinglin Meng and al, "An Online Reinforcement Learning Based Energy Management Strategy for Microgrids With Centralized Control ," IEEE Transactions on Industry Applications,vol. 61, 2025.

[4] B. El Barkouki and al, "Energy efficiency and optimal operation of a residential microgrid based on demand side management strategy," 9th International Youth Conference on Energy (IYCE), IEEE, 2024.

[5] N. A. Chowdhury and al "Energy management of AC residential microgrids using advanced fuzzy inference method, IEEE International Conference on Energy Technologies for Future Grids (ETFG), 2023.

[6] L. Meng, E. R. Sanseverino, A. Luna, T. Dragicevic, J. C. Vasquez, J. M. Guerrero, "Microgrid supervisory controllers and energy management systems: A literature review", Renewable and Sustainable Energy Reviews 60, pp.1263 – 1273, 2016.

[7] H .Ben Mansour, and al, "Supervisory control for Energy Management of Islanded Hybrid AC/DC Microgrid," IJCSNS International Journal of Computer Science and Network Security,vol.22 No.3,March 2022.

[8] H. Auer and R. Haas, "On integrating large shares of variable renewables into the electricity system," Energy, 115, pp.1592-1601,2016.

[9] S. F. Thabet and al ,"Intelligent home energy management system based on genetic algorithm for hybrid PV-battery operation under on/off-grid conditions," IEEE Access, 2023.

[10] I. Cristian Hoarcă, " Mathematical modeling and simulation of PV systems Part I: Mathematical modeling and Simulink implementation",International Conference on Applied and Theoretical Electricity (ICATE),2021.

[11] H. Ben Mansour, and al,"Energy Management Supervisory Controller Embedded-Board for Islanded Hybrid AC/DC Microgrids," 6th IEEE International Energy Conference (ENERGYCon),pp.924– 929 ,2020.

[12] P. M. Mutuku and al, "A Systematic Open Source Modeling Method for Energy and Power Systems: Solar PV Application," IEEE Access,vol.13, 2023.

[13] F. Rukavina and al ,"Identification of equivalent circuit model parameters for a Li-ion battery cell," ,11th IEEE International Conference on Systems and Control (ICSC), 2023.

[14] J. Singh and al, "A Review on State of Charge (SoC) Estimation Methods of Li-Ion Batteries Used in Electrified Transportation," International Conference on Cognitive Computing in Engineering, Communications, Sciences and Biomedical Health Informatics (IC3ECSBHI), IEEE,2025.

Optimization of ON/OFF Control for Computing Servers in Non-Terrestrial Networks

Insaf Rzig[1] and Wael Jaafar[2]
[1]MEDIATRON Laboratory, École Supérieure des Communications (Sup'Com), University of Carthage, Tunisia
[2]Department of Software and IT Engineering, École de technologie supérieure (ÉTS), University of Quebec, Canada
`insaf.rzig@supcom.tn, wael.jaafar@etsmtl.ca`

Abstract

The rapid evolution of 5G and the forthcoming 6G technologies has enabled massive connectivity and high data rates for emerging applications such as the Internet of Things (IoT), autonomous vehicles, and smart cities. However, this advancement has also intensified energy consumption and computational demand in modern communication systems. This work presents a multi-layer simulation framework that integrates terrestrial base stations (BS), High-Altitude Platform Stations (HAPS), and Low-Earth Orbit (LEO) satellites to study energy-efficient ON/OFF control of computing servers in Non-Terrestrial Networks (NTNs). The goal is to optimize server activation and task assignment while maintaining high Quality of Service (QoS). Four optimization algorithms, namely Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Differential Evolution (DE), and Q-Learning, are compared with heuristic approaches such as Closest Node and HAPS-only assignment.

Keywords: LEO, HAPS, MEC, ON/OFF Control

1 Introduction

The evolution of 5G and the upcoming 6G networks aims to deliver global connectivity, ultra-low latency, and massive data throughput. However, these advances come at a significant energy cost: mobile networks are projected to account for a growing share of global electricity consumption, raising urgent concerns about sustainability and carbon footprint.

To address these challenges while extending coverage, Non-Terrestrial Networks (NTNs)—including High-Altitude Platform Stations (HAPS) and Low-Earth Orbit (LEO) satellites—are emerging as essential complements to traditional Terrestrial Networks (TN). When equipped with Mobile Edge Computing (MEC) capabilities, NTN nodes can process data closer to users, reducing backhaul traffic and latency. Yet, keeping MEC servers perpetually active across all layers—terrestrial, HAPS, and LEO—leads to substantial idle energy waste, especially in sparsely populated or low-demand regions. This is particularly critical for energy-constrained NTN platforms, such as solar-powered HAPS or battery-limited LEO satellites, where every joule counts. Inspired by cell-switching strategies in terrestrial networks—where inactive base stations are turned off to save energy—this work extends the concept to multi-layer NTN-MEC systems through intelligent ON/OFF control. We propose a multi-tier simulation framework that dynamically activates or deactivates computing servers based on real-time user demand, network load, and QoS requirements. By doing so, we aim to maximize energy efficiency without compromising service reliability, paving the way toward sustainable, green NTN architectures for 6G and beyond.

1.1 Background

Integrating NTNs—such as HAPS and LEO satellites—with terrestrial infrastructure has been widely explored to extend coverage and improve energy efficiency. Prior work highlights NTN–TN hybrid systems' ability to enable traffic offloading, adaptive resource allocation, and reduced reliance on grid power through renewable energy sources like solar. However, challenges persist in managing latency disparities and coordinating heterogeneous nodes, calling for intelligent control mechanisms to jointly optimize performance and energy efficiency [1, 2].

1.2 Problem Definition

We consider a multi-layer NTNs composed of terrestrial base stations, one HAPS, and one LEO satellite, each equipped with a switchable MEC server. Let $\mathcal{U} = \{1, \ldots, U\}$ be the set of user requests, and $\mathcal{N} = \mathcal{N}_{\mathrm{BS}} \cup \{\mathrm{HAPS}\} \cup \{\mathrm{LEO}\}$ the set of computing nodes.

We formulate our problem as follows:

$$\max_{\mathbf{x},\mathbf{y},\mathbf{s}} \alpha \cdot \underbrace{\frac{1}{U}\sum_{u=1}^{U} s_u}_{\text{Success Rate (SR)}} +(1-\alpha)\cdot \underbrace{\frac{1}{E_{\max}}\sum_{n=0}^{N}\left(E_n^{\max}-E_n^{\text{tot}}\right)}_{\text{Remaining Energy (RE)}} \tag{$\mathcal{P}1$}$$

$$\text{s.t.} \quad T_{u,n}^{\text{srv}} \cdot x_{u,n} \leq D_u, \quad \forall u \in \mathcal{U},\ n \in \mathcal{N}, \tag{$\mathcal{P}1$.a}$$

$$E_n^{\text{tot}} = E_n^{\text{comp}} + E_n^{\text{trans}} + E_n^{\text{act}} y_n \leq E_n^{\max}, \quad \forall n \in \mathcal{N}, \tag{$\mathcal{P}1$.b}$$

$$\sum_{n\in\mathcal{N}} x_{u,n} = 1, \quad \forall u \in \mathcal{U}, \tag{$\mathcal{P}1$.c}$$

$$x_{u,n} \in \{0,1\}, \quad \forall u, n, \tag{$\mathcal{P}1$.d}$$

$$y_n \in \{0,1\}, \quad \forall n \in \mathcal{N}, \tag{$\mathcal{P}1$.e}$$

$$s_u \in \{0,1\}, \quad \forall u \in \mathcal{U}, \tag{$\mathcal{P}1$.f}$$

In problem ($\mathcal{P}1$), the decision variables are defined as follows:

- $x_{u,n} \in \{0,1\}$: binary assignment variable, where $x_{u,n} = 1$ if user request u is offloaded to node $n \in \mathcal{N}$ (terrestrial BS, HAPS, or LEO), and 0 otherwise;
- $y_n \in \{0,1\}$: binary activation variable, where $y_n = 1$ if node n is switched **ON** to serve requests, and 0 if kept **OFF** to conserve energy;
- $s_u \in \{0,1\}$: satisfaction indicator, where $s_u = 1$ if request u is successfully completed within its deadline D_u, and 0 otherwise.

The constraints ensure physical feasibility and system requirements: ($\mathcal{P}1$.**a**) enforces that the total service delay $T_{u,n}^{\text{srv}}$ (transmission + computation + queuing) for any assigned request does not exceed its deadline D_u. ($P1$.**b**) limits the total energy consumed by each node E_n^{tot}—comprising computation, transmission, and activation costs—to its maximum available energy budget $E_n^{\max}$. ($\mathcal{P}1$.**c**) guarantees that each user request is assigned to exactly one computing node. ($\mathcal{P}1$.**d**–**e**) enforce binary decisions for assignment and activation. ($\mathcal{P}1$.**f**) defines user satisfaction: a request is marked as satisfied *only* if it is both assigned and completed before its deadline.

1.3 Contribution

- Proposes an integrated architecture using terrestrial, HAPS, and LEO nodes for low-latency, energy-efficient processing of IoT and mobile applications.
- Formulates the request assignment problem with energy, activation cost, and load balancing constraints.
- Compares the proposed approach with baseline methods.

2 Methodology

Our methodology leverages a heterogeneous network of terrestrial, HAPS, and LEO nodes. We model request assignment with constraints on energy, activation cost, and task deadlines. To evaluate performance, we implement a Python-based simulator incorporating transmission and computation energy models. Both online (QL, Closest Node) and offline (GA, PSO, DE) strategies are tested, providing insights on energy consumption, request success rate, and overall energy/QoS efficiency.

2.1 Theoretical Framework

This work is grounded in established models for heterogeneous NTN. The transmission rate between user u and node n is computed using Shannon's capacity:

$$D_{u,n} = B_{u,n} \log_2\left(1 + \frac{P_{u,n}G_{u,n}}{N_0 B_{u,n}}\right), \tag{2}$$

where $B_{u,n}$ is bandwidth, $P_{u,n}$ transmit power, $G_{u,n}$ effective channel gain (including free-space path loss, atmospheric attenuation for HAPS, and Doppler effects for LEO), and N_0 noise spectral density.

2.2 Datasets and Tools

To model the heterogeneous network, we designed and implemented a modular Python-based simulator using libraries such as NumPy, PyQt, pymoo, and pandas.

2.3 Algorithm Description

We evaluate both online and offline strategies for task assignment under OnDemand server activation:

- **Q-Learning (online)**:
  - *State space*: ($N_{\text{active}}^{\text{BS}}, N_{\text{active}}^{\text{HAPS}}, N_{\text{active}}^{\text{LEO}}$) — number of active users per layer.
  - *Action space*: Assign the current task to one of {BS, HAPS, LEO} (3 discrete actions).
  - *Reward function*: Positive reward for successful task completion, penalty for failure; inversely proportional to energy consumed (Joule/bit).
- **GA, PSO, DE (offline)**:
  - *Search vector*: A flattened binary assignment matrix $\mathbf{x} \in \{0,1\}^{U\times N}$, where $x_{u,n} = 1$ if task u is assigned to node n.

  - *Fitness function*: Directly derived from the scalarized objective of ($\mathcal{P}1$): maximize success rate and remaining energy.

- **Heuristics** (Closest Node, HAPS-only) use On-Demand activation but make greedy assignment decisions without learning or global optimization.

3 Results and Discussion

3.1 Experimental Setup

Our experiments consider a multi-layer NTN with four terrestrial BS (1 km coverage each), one HAPS at 20 km, and one LEO satellite at 480 km, all equipped with switchable MEC servers. User requests follow three types: Type 1 (1–3 MB, 0.2 s deadline), Type 2 (3–6 MB, 0.5 s), and Type 3 (6–10 MB, 1.0 s). Node capacities are 3 GHz (BS), 5 GHz (HAPS), and 10 GHz (LEO). Transmit powers are 30 dBm (BS), 33 dBm (HAPS), and 40 dBm (LEO), with antenna gains of 10 dBi, 15 dBi, and 8 dBi, respectively. Activation costs are 5 J, 20 J, and 50 J, and residual energy budgets are 10 KJ (HAPS) and 15 KJ (LEO). The channel uses $N_0 = -174$ dBm/Hz noise.

3.2 Quantitative Results

We define the energy/satisfaction ratio (also referred to as energy/QoS) as the number of successfully completed tasks per joule of total energy consumed across all nodes, providing a unified metric that balances service reliability and energy sustainability.

As shown in Fig. 1, the Q-Learning model converges after approximately 1000 training episodes, achieving a significantly better energy/satisfaction ratio than the ClosestNode heuristic, with Q-Learning requiring approximately 15% less energy per successful task at convergence. This demonstrates Q-Learning's ability to learn energy-aware assignment policies that outperform greedy strategies.

Figure 2 presents the success rate as a function of the number of concurrent users, ranging from 10 to 100. At peak load (100 users), PSO achieves the highest success rate at 94%, followed closely by Q-Learning at 92%, while GA and DE reach 88% and 85%, respectively. In contrast, heuristic methods suffer severe degradation: ClosestNode drops to 48% and HAPSOnly to only 27%. These results confirm that intelligent optimization is essential for maintaining high QoS under heavy demand, whereas static or proximity-based assignment fails to scale in dynamic multi-layer NTN environments.

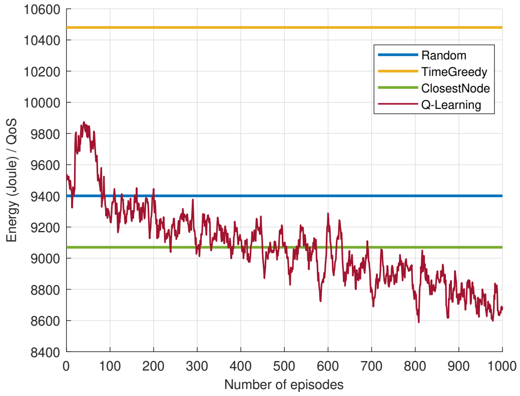


Figure 1: Q-Learning Model Training

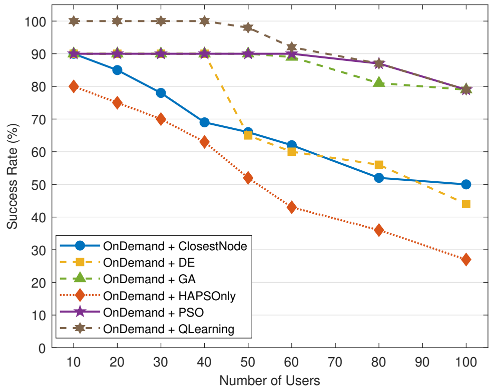


Figure 2: Success Rate vs Users

4 Conclusion

This work demonstrated that intelligent ON/OFF control of computing servers in Non-Terrestrial Networks can significantly improve energy efficiency and service reliability. Among the evaluated methods, PSO and Q-Learning achieved the best balance between success rate and scalability, outperforming heuristic approaches.

References

[1] Sadovaya and et al., "Enhancing service continuity in non-terrestrial networks via multi-connectivity offloading," *IEEE Commu. Lett.*, 2024.

[2] Ozturk and et al., "Integrating terrestrial and non-terrestrial networks for sustainable 6g operations: A latency-aware multi-tier cell-switching approach," *IEEE Network*, 2025.

A Comprehensive Evaluation of Clustering Algorithms for LiDAR Object Detection

Yasmine Maalej[1,2], Ferdaous Chaabane[1], Sofiane Sayahi[2], Hichem Besbes[1]
[1]COSIM, SUP'COM, Carthage University, Ariana, Tunisia
[2]Innovation Department, ACTIA ES, Ariana, Tunisia
`yasmine.maalej@supcom.tn, ferdaous.chaabene@supcom.tn,`
`sofiane.sayahi@actia-engineering.tn, hichem.besbes@supcom.tn`

Abstract

The uneven distribution of LiDAR point cloud (PCL) data frequently leads to poor clustering quality, that significantly affects the accuracy of 3D road object detection. This paper presents a LiDAR-based detection pipeline designed to alleviate this problem. First, the approach limits clustering inconsistencies by defining regions of interest (ROIs) and evaluating several clustering methods. Then, it applies object-oriented bounding box (OBB) fitting strategies adapted to different object classes, using Principal Component Analysis (PCA) to guarantee stable and reliable OBB fitting results. Finally, a Random Forest (RF) model is employed for object classification. Experimental results on the KITTI dataset demonstrate that the proposed method achieves robust and consistent 3D object recognition performance

Keywords: Autonomous driving, Environment perception, LiDAR, Clustering, Vscore

1 Introduction

Accurate 3D object detection from LiDAR point clouds (PCL) is critical for the safety and navigation of autonomous vehicles. However, the PCL distribution variation represents a significant challenge for the 3D OBB detection. Specifically, while the PCL density is highest for points closest to the sensor, it decreases as the square of the distance from the sensor[1]. As a result, as the distance from the sensor increases, the points become sparser and are thus more difficult to cluster using conventional methods based on fixed density thresholds. This leads to miss-classification between neighboring clusters and the consideration of distant clusters as noise due to reduced clustering density [2]. To address this issue, Y. Tian et al. introduced the elevation-reference connected component labeling (ER-CCL) algorithm with a flexible search range [3], though at the cost of high computational complexity. Subsequent work proposed adaptive approaches such as A-DBSCAN and improved L-Shape algorithms for oriented bounding box (OBB) fitting [4]. PointNet [5] introduced a new way to use neural networks for classification and segmentation of point cloud data. The recently updated PointPillars architecture, as noted in [6], uses a combination of multiple pooling and multi-scale fusion in order to improve classification accuracy on the KITTI dataset.

However, one major limitation still persists : the existing method for separation of foreground are poor or incomplete. Without a reliable separation between foreground and background, selection of clustering parameters becomes difficult. In this paper, it is shown that by properly segmenting foreground and background, conventional clustering approaches can continue to be very successful in detecting objects on the roadway and achieve equal or better performance compared to when using complex adaptive methods.

The key contributions of this paper include the following:

- Clustering evaluation on kitti validation set.
- A comprehensive evaluation of the object classification performance.

2 Methodology

2.1 Theoretical Framework

A multilevel DBSCAN approach is used in this pipeline to resolve the problem of points' density variation in LiDAR data. Typically, DBSCAN only uses one threshold value (density) based on the two parameters ε (maximum possible distance between two points that can be considered neighbors) and *minPoints* (the minimum number of points needed to determine whether a cluster exists). Therefore, it will not work properly on LiDAR data that have varying densities of points close together and far away from each other. In order to overcome this issue, a three-level approach is taken to mitigate the effects of this particularity in clustering (it uses 3 layers of DBSCAN). First, relatively small ε values and a high number of *minPoints* are used to find compact clusters of closely related objects. Secondly , we considered larger ε values and fewer *minPoints*, simultaneously, to accommodate for the sparse nature of some of these clusters that may have been missed during the first phase of clustering. In this way, by applying DBSCAN in three levels, the number of instances in which adjacent objects are incorrectly merged, and the potential for incorrectly disrupting large clusters, should be greatly diminished. Parameters of both clustering algorithms are mentioned in Table 1.

OBBs are then computed using Principal Component Analysis (PCA) [7] on the 2D horizontal projection of the 3D point cloud. Let $\mathbf{p}_i^{[xy]} = [x_i, y_i]^T$ denote the 2D projection of the i-th 3D point onto the xy-plane. To perform PCA, these points must first be centered by subtracting the mean of the point cloud. The centered 2D points $\mathbf{p}_i'^{[xy]}$ are calculated as:

$$\mathbf{p}_i'^{[xy]} = \mathbf{p}_i^{[xy]} - \bar{\mathbf{p}}^{[xy]}, \quad \text{where} \quad \bar{\mathbf{p}}^{[xy]} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{p}_i^{[xy]} \tag{1}$$

The covariance matrix $\mathbf{C}$ is then computed from the **centered** points $\mathbf{p}_i'^{[xy]}$ as:

$$\mathbf{C} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{p}_i'^{[xy]}(\mathbf{p}_i'^{[xy]})^T \tag{2}$$

Then, eigenvalue decomposition is applied to $\mathbf{C}$ to extract the principal directions. The orientation angle θ of the OBB is derived from the dominant eigenvector.

2.2 Algorithm Description

The 3D object detection pipeline described in Figure 1 consists of several stages. It starts with generating the PCL annotation given 3D OBB annotations based on the camera's coordinate frame. The OBBs are then transformed from the camera's coordinate system into the LiDAR coordinate system. The LiDAR points in the car, cyclist, and pedestrian boxes were assigned the same label as their corresponding box. Any points that were found outside of the Regions of Interest (ROIs) were considered background points and removed from consideration. The remaining foreground points were then clustered using five separate algorithms (DBSCAN [8], Multilevel DBSCAN, HDBSCAN [9], OPTICS [8], and Spectral Clustering [10]) in order to create a benchmark for comparing the clusters. These five clustering methods support model interpretability and have relatively low computational requirements; entity clustering is relevant to the development of robotically autonomous agents. The clusters that result from the clustering methods were each fitted to a 3D OBB, and subsequently, a set of 10 descriptors were generated for each cluster. It includes cluster centroid coordinates, length, width, and height of each cluster, density, cluster reflectance, occlusion status, and orientation. To provide balanced classification results, training data was equilibrated across the three different classes (car, cyclist, pedestrian) and a machine learning model was trained to predict the classes of the 3D OBB. Performance of the model was evaluated on the KITTI validation dataset.

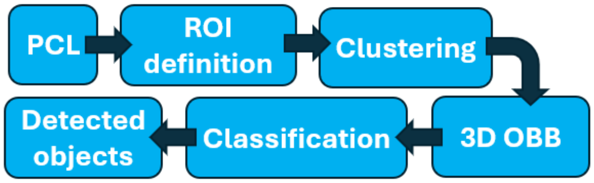


Figure 1: Lidar Pipeline

The KITTI dataset [11] is commonly used in research related to Autonomous Driving. It contains 7,481 unique images which have been labeled. According to [12],the set is split into 3,712 images used for training and 3,769 images for testing. In total, the KITTI dataset has 17,558 labeled objects.

Table 1: Comparison of clustering methods on KITTI LiDAR data (validation set); V: Vscore, H: Homogeneity, C: completeness

Method	V	H	C
DBSCAN (n=8, eps=1)	0.96	0.96	0.98
Multilevel (n=10, 9, 8; eps=0.8, 0.9, 1)	0.95	0.98	0.95
HDBSCAN (n=10)	0.86	0.96	0.88
OPTICS (n=10, eps=0.1, xi=0.05)	0.85	0.96	0.86
Spectral Clustering (n=8, kernel=rbf)	0.58	0.87	0.57

Table 2: Model Performance Metrics with Percentage Breakdown

Metric	**Precision**	**Recall**	**F1-Score**	**TP %**	**FP %**	**FN %**
RF-Multilevel	0.94	0.88	0.91	87.6	5.7	12.4
RF-DBSCAN	0.94	0.87	0.90	86.9	5.3	13.1

Table 3: Performance by Difficulty Level and Class

Metric	**Car**			**Cyclist**			**Pedestrian**		
	Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard
Precision	0.9866	0.9832	0.9812	0.8824	0.6043	0.2703	0.8765	0.8343	0.7494
Recall	0.9156	0.9019	0.9005	0.8824	0.5904	0.2041	0.8656	0.8218	0.7222
F1-Score	0.9498	0.9408	0.9391	0.8824	0.5973	0.2326	0.8710	0.8280	0.7355

3 Results and Discussion

3.1 Experimental Setup

Using a systematic grid search optimization method, optimal values were found for the parameters associated with each of the five clustering algorithms. The results of the grid search optimization are summarized in Table 1. Each algorithms performance was assessed using standard measures: V-measure, homogeneity and completeness. We also compared the performance of three different machine learning classifiers: Support Vector Machine (SVM), eXtreme Gradient Boosting (XGBoost), and Random Forest (RF). Among these, the RF classifier performed the best and was used for all future tests and analyses. The RF model constructed 200 trees of variable depths. Similar to traditional tree-based classification models, the RF model made use of sqrt feature sampling and bootstrap aggregation to create the random patterns observed in the trees. The splits of a given node in the tree required a minimum number of two samples to create a viable candidate for the next branch. This allows single sample leaves to support fine-grained classification patterns. We evaluated classification performance using standard metrics: Precision, Recall, F1-score, as well as detection statistics: True Positive (TP), False Positive (FP), and False Negative (FN). We used the Hungarian algorithm to match ground-truth objects with those predicted by the RF classifier with a maximum distance of two meters.

3.2 Quantitative Results

Based on the clustering and classification results showed in Tables 1 and 2, the proposed multilevel DBSCAN algorithm realized the best performance compared to the other methods in terms of highest TP rates and V-measure scores. It demonstrates that it can provide improved performance by effectively dealing with points distributions variations, by preventing the merging of adjacent objects and the over-segmentation of larger objects. The overall recall of scenes is improved for the proposed method.

The classification results listed in Table 3 further support that this proposed clustering method yields TP rates high enough to be reliable and accurate for classifying objects.

3.3 Qualitative Results

The quality of the detection compared to the ground truth can be seen in Figure 2: the further an object is from a vehicle, the more challenging its 3D OBB accurat prediction (especially their size and orientation). The smaller clusters produced in Tracks 4 and 5 resulted from having fewer PCL available to properly capture the volume of the objects being detected.

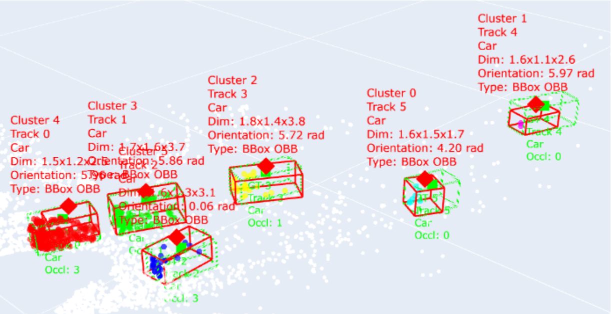


Figure 2: visualisation of OBB detected in frame 8; ground truth: green OBB and detection: red OBB

4 Conclusion

This study developed an effective LiDAR-based 3D detection pipeline, demonstrating that traditional clustering methods, particularly DBSCAN and Multilevel DBSCAN, remain highly competitive when combined with reliable foreground–background separation, achieving V-scores up to 0.9593. The proposed approach to estimate the object orientation by PCA produced geometrically accurate object orientations and yielded to very sufficient results from the entire pipeline across object classes. The vehicle class having an F1score of 0.9498 highlights the performance of the proposed method . However, there are still challenges associated with detecting distant objects and cyclists in difficult conditions as a result of point sparseness and occlusions. Future work will investigate more robust background filtering methods and adaptive clustering strategies to improve the detection reliability from these types of objects.

References

[1] S. Kodors and I. Kangro, "Simple method of lidar point density definition for automatic building recognition," *ENGINEERING FOR RURAL DEVELOPMENT*, pp. 415–424, 05 2016.

[2] Z. Duan, J. Shao, M. Zhang, J. Zhang, and Z. Zhai, "A small-object-detection algorithm based on lidar point-cloud clustering for autonomous vehicles," *Sensors*, vol. 24, no. 16, 2024.

[3] Y. Tian, W. Song, L. Chen, Y. Sung, J. Kwak, and S. Sun, "A fast spatial clustering method for sparse lidar point clouds using gpu programming," *Sensors*, vol. 20, no. 8, 2020.

[4] X. Jin, H. Yang, and Z. Li, "Vehicle detection framework based on lidar for autonoumous driving," in *2021 5th CAA International Conference on Vehicular Control and Intelligence (CVCI)*, Oct 2021, pp. 1–5.

[5] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017, pp. 77–85.

[6] J. Xu, Z. Chen, S. Xiong, K. Wu, and C. Qian, "Improved pointpillars for 3d object detection based on multidimensional feature fusion," in *2025 8th International Conference on Transportation Information and Safety (ICTIS)*, July 2025, pp. 645–650.

[7] S. Gottschalk, M. Lin, and D. Manocha, "Obbtree: A hierarchical structure for rapid interference detection," *Computer Graphics*, vol. 30, 10 1997.

[8] X. Jin, H. Yang, X. He, G. Liu, Z. Yan, and Q. Wang, "Robust lidar-based vehicle detection for on-road autonomous driving," *Remote Sensing*, vol. 15, no. 12, 2023.

[9] B. R. Kiran, L. Roldão, B. Irastorza, R. Verastegui, S. Suss, S. Yogamani, V. Talpaert, A. Lepoutre, and G. Trehard, "Real-time dynamic object detection for autonomous driving using prior 3d-maps," 2019.

[10] Y. Tian, W. Song, L. Chen, Y. Sung, J. Kwak, and S. Sun, "A fast spatial clustering method for sparse lidar point clouds using gpu programming," *Sensors*, vol. 20, no. 8, 2020.

[11] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: the kitti dataset," *The International Journal of Robotics Research*, vol. 32, pp. 1231–1237, 09 2013.

[12] H. Wu, C. Wen, W. Li, X. Li, R. Yang, and C. Wang, "Transformation-equivariant 3d object detection for autonomous driving," 2022.

definition remark

Toward a New Definition of Ergodicity: An application to image texture analysis

Sonia Rezk[1], Zaineb Loksaier[2],
[1]University of Sousse, Noccs Lab, ENISO, Sousse,Tunisia
[2]University of Sfax, Lamha Lab, FSS, Sfax, Tunisia
sonia.rezk@isitc.rnu.tn, loksaie0@gmail.com

Abstract

This work investigates the relationship between continuous transformation groups and their discrete equivalent in image texture analysis using the notion of rational closure $I(H, \Gamma)$. We focus on the behavior of rational and irrational orbits generated by subgroup actions on a discrete lattice $\Gamma = \mathbb{Z}^2$. Rational orbits produce periodic and structured trajectories, while irrational orbits exhibit quasi-ergodic distributions over the discrete grid. These results are illustrated by numerical simulations.

Keywords: Rational colusre, irrational, Ergodic, Heisenberg Group, Nilpotency, Image, Texture.

1 Introduction

Image texture analysis is a central problem in visual information processing, as it provides critical cues for object recognition, surface characterization, and pattern understanding. Textures often combine local regularity with global variability [7] , making them challenging to model within a unified mathematical framework. Although traditional approaches rely on statistical descriptors or frequency-based methods [8], recent advances highlight the need for geometric and algebraic models that can help texture analysis.By introducing the notion of rational closure, defined as the smallest connected and normal subgroup compatible with the discrete structure of Γ, we can determine which continuous transformations can be discretized consistently without disrupting the structural coherence of the representation. This contribution is interpreted in the context of Texture image analysis.

2 Background

In this section, we give a brief review of certain results from rational structures and uniform subgroups of connected simply connected nilpotent Lie groups which will be used in the main part of the paper. The general references used in this section are [3].

Let G be a connected, simply connected nilpotent Lie group with Lie algebra $\mathfrak{g}$. Then the exponential map

$$\exp : \mathfrak{g} \longrightarrow G$$

is a diffeomorphism. Let

$$\log : G \longrightarrow \mathfrak{g}$$

denote the inverse of exp.

Definition 2.1 (Totally irrational subalgebra) *Let G be a connected simply connected nilpotent Lie group with Lie algebra $\mathfrak{g}$, and let Γ be a discrete uniform subgroup of G. A subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is said to be* totally irrational *if $\mathfrak{h}$ is not contained in any proper rational Lie subalgebra of $\mathfrak{g}$.*

Theorem 2.1 *Let G be a connected simply connected nilpotent Lie group with Lie algebra $\mathfrak{g}$ and Γ a discrete uniform subgroup of G. Let H be a connected closed subgroup of G. Then the action of H on G/Γ is ergodic if and only if its Lie algebra $\mathfrak{h}$ is totally irrational in $\mathfrak{g}$.*

Remark 2.1 *The above result is well known in the case where H is a one-parameter subgroup.*

In [1], we give a new criterion of ergodicity of nilflows in terms of $I(H, \Gamma)$.

Theorem 2.2 *Let G be a connected simply connected nilpotent Lie group and Γ a discrete uniform subgroup of G. Let H be a connected closed subgroup of G. Then the nilflow $(G/\Gamma, H)$ is ergodic if and only if $I(H, \Gamma) = G$.*

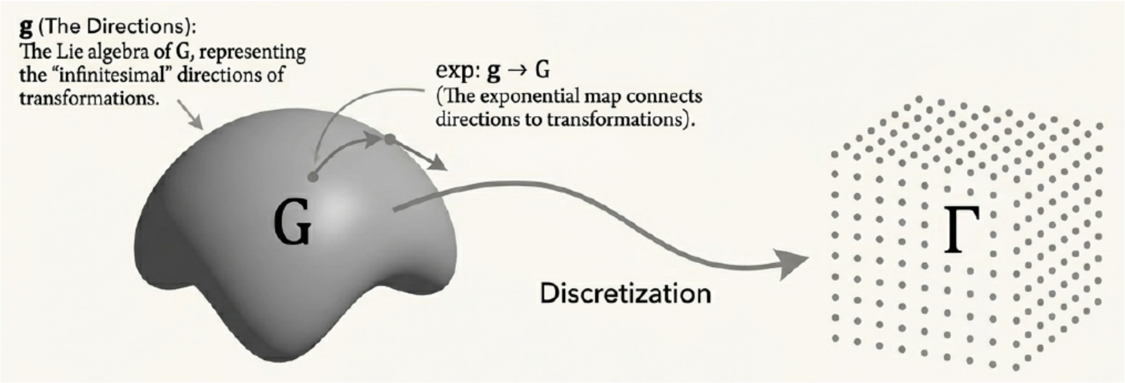


Figure 1: **The continuous space G**: A connected, simply connected nilpotent Lie group. Think of this as the smooth space of all possible transformations(e.g, translations, rotations,..), The discrete grid Γ): a discrete uniform subgroup of G. This represents a structured lattice , analogous to the pixel grid in an image.

Example 2.1 (The 3-dimensional Heisenberg group) *Consider the 3-dimensional Heisenberg algebra* $\mathfrak{g} = \mathbb{R}\text{-}span\{X, Y, Z\}$ *with Lie brackets given by*

$$[X, Y] = Z$$

and the non-defined brackets being equal to zero or obtained by antisymmetry.

Let G *be the connected, simply connected nilpotent Lie group with Lie algebra* $\mathfrak{g}$*. For* $k \in \mathbb{N} \setminus \{0\}$*, let*

$$\Gamma_k = \exp\left(\frac{1}{k}\mathbb{Z}Z\right)\exp(\mathbb{Z}Y)\exp(\mathbb{Z}X).$$

Each uniform discrete subgroup of G *is isomorphic to exactly one* Γ_k*. Let* $\Gamma = \Gamma_k$ *for some* $k \in \mathbb{N} \setminus \{0\}$*.*

1. *Let* $H = \exp(\mathbb{R}X)$ *be a one-parameter subgroup of* G*. Then*

   $$\langle H, \Gamma\rangle = \exp(\mathbb{R}X)\exp(\mathbb{Z}Z)\exp(\mathbb{Z}Y),$$

   and

   $$I(H, \Gamma) = \exp(\mathbb{R}X)\exp(\mathbb{Z}Z) = \exp(\mathbb{Z}Z)H.$$

2. *Let* $H = \exp(\mathbb{R}(X + \alpha Y))$*, where* $\alpha \in \mathbb{R} \setminus \mathbb{Q}$*. It is clear that* $\langle H, \Gamma\rangle = G$ *and so* $I(H, \Gamma) = G$*.*

   Thus

   $$N_G^\Gamma(H) = G,$$

   and H *acts ergodically on* G/Γ*.*

3 Application to image texture analysis

In this section, We demonstrate how the theoretical concepts can be related to an application such as image processing, using the Heisenberg group as the underlying mathematical model.
In the context of digital image processing, the operations applied to an image are performed on a pixel grid, which corresponds to a discrete lattice (or network) $\Gamma \subset G$.

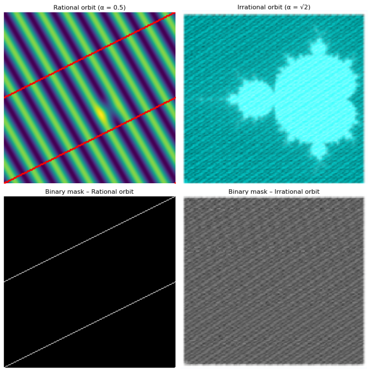


Figure 2: first line: rational orbit with $\alpha = 1/2$ (and irrational orbit with $\alpha = \sqrt(2)$) orbit. Second line: Their corresponding binary mask..

We recall that G is the 3-dimensional Heisenberg group:

$$G = (x, y, t) \in \mathbb{R}^3, [(x, y, t).(x', y', t')] = (0, 0, xy' - yx')$$

describes local transformations in the space of signals or images. Then, we consider its discrete network (or lattice):

$$\Gamma = \{(m, n, p) | m, n, p \in \mathbb{Z}\}$$

modeling pixels grid and phases. Therefore, the rational closure $I(H, \Gamma)$ represents the set of consistent translations with the discrete grid, i.e those that preserve the structure of the pixel and phase network.if we introduce an oblique sub-group:

$$H_\alpha = \{(x, \alpha x, 0) | x \in \mathbb{R}\}$$

then we can easily demonstrate that the ergodicity criterion distinguishes between rationnel and irrationnel transformation:

- if $\alpha \in \mathbb{Q}$ (rationnel transformation), $I(H, \Gamma) \neq G$, why? Let α be a rationnel number (e.g.1/2). The transformation aligns with the discrete underlying lattice Γ. After applying the criterion, calculation shows the rationnel closure $I(H, \Gamma)$ is a proper subgroup of $G(\ I(H, \Gamma) \neq G)$. In this case, the rationnel transformation produces a periodic,structured trajectory. The orbit is sparce and aligns with the underlying grid, creationg a predictable patter,
- if $\alpha \notin \mathbb{Q}$ (irrationnel transformation), $I(H, \Gamma) = G$, why? Let α be a rationnel number (e.g.$\sqrt{2}$). The transformation fundamentally misaligned with the discrete underlying lattice Γ. After applying the criterion, calculation shows the rationnel closure $I(H, \Gamma)$ spans the entire group G $(I(H, \Gamma) = G)$. The action is ergodic. It is space-filling and aperiodic. In this case, the irrationnel transformation produces a a quasi ergodic distribution. The orbit densely fills the space, creating a complex, aperiodic texture without discernable repetition: it is the signature and unfolding of ergodicity,

  This reflects a fundamental distinction between periodic and aperiodic textures in image analysis.

  Simulations results in figure 2 demonstrate the effectiveness of our approach for α =-1/2 (rational) and $\alpha = -\sqrt{2}$ (irrational). The two images definition is 256*256. The orbit is generated for different scenarios by varying the number of iterations n. The results shown in figure 2 are the trajectory reached by the orbit in the last iteration i.e. $n = 256 * 256$.

4 Conclusion

The main contributions of this paper are summarized as follows: we introduced a new elegant criterion in nilpotent Lie group: the action of H is ergodic if only if $I(H, \Gamma) = G$ and we demonstrated that this abstract mathematical condition provides a direct and powerful explanation for the visual difference between periodic and ergodic image texture.

As clear directions for future work, we plan to develop quantitative metrics of ergodicity to move behind the binary ergodic/non-ergodic distinction and implement large-scale simulations to validate the proposed model on real-world image datasets. Finally, we will explore the utility of this framework in other areas of signal processing and pattern recognition.

References

[1] A. Ghorbel and Z. Loksaier, *On the rational closure of connected closed subgroups of connected simply connected nilpotent Lie groups*, Proc. Indian Acad. Sci. (Math. Sci.) **129** (2019), 82.

[2] A. Ghorbel and Z. Loksaier, *On the irreducibility of some restrictions in nilpotent Lie groups*, Bull. Sci. Math., vol. 161, 102855, 2020.

[3] L. Corwin and F. P. Greenleaf, *Representations of Nilpotent Lie Groups and Their Applications, Part 1: Basic Theory and Examples*, Cambridge Studies in Advanced Mathematics, vol. 18, Cambridge University Press, New York, 1989.

[4] L. Auslander, *On radicals of discrete subgroups of Lie groups*, Amer. J. Math. **85** (1963), 145–150.

[5] M. S. Raghunathan, *Discrete Subgroups of Lie Groups*, Springer, 1972.

[6] B. Bekka and P. Driutti, *Ergodic actions on nilmanifolds and restriction of unitary representations to lattices*, Math. Proc. Camb. Phil. Soc. **131** (2001), 91–96.

[7] Leila Kabbai, Mehrez Abdellaoui, Ali Douik: Image classification by combining local and global features. Vis. Comput. 35(5): 679-693 (2019)

[8] M. Mirmehdi, X. Xie and J. Suri, Handbook of texture analysis, London: Imperial college press, 2008.

fNIRS Signal Processing to Study of Attentional Networks in Rapid Event Paradigms

Ali Abidi[1], Sofia Ben Jebara[1,2] Malek Abidi[3,4] Giovanni de Marco [3,4]
[1] University of Carthage, Higher School of Communication of Tunis (SUPCOM), Tunis Tunisia
[2] COSIM Laboratory, SUPCOM
[3] Laboratory LINP2, UPL, Paris Nanterre University
[4] COMUE Paris Lumiéres University
`ali.abidi@supcom.tn, sofia.benjebara@supcom.tn,`
`m.bida@parisnanterre.fr, gdemarco@parisnanterre.fr`

Abstract

The accurate analysis of brain activity during rapid cognitive tasks is crucial for neuroergonomics and brain-machine interfaces. This study investigates the combined effects of type of attentional alerting, cognitive conflict, and temporal delays on cerebral oxygenation using functional Near-Infrared Spectroscopy. We employed a pipeline including signal preprocessing, featuring and modelling. Statistical analysis revealed significant main effects for both congruence and SOA on signal intensity (amplitude and overall oxygen delivery over task time), with no effect observed for mode. For instance, congruence significantly enhanced oxygenation responses at shortest and longest delays.

Keywords: fNIRS, Attentional orienting, GLM, attentional modulation.

1 Introduction

The study of the brain mechanisms involved in decision-making, motor coordination and cognitive load is of major importance in fields such as neuroergonomics, motor rehabilitation and brain-machine interfaces. In this context, it is essential to be able to accurately record and analyse brain activity during simple but rigorously controlled cognitive and motor tasks. Besides established techniques like functional Magnetic Resonance Imaging (fMRI), MagnetoEncephaloGraphy (MEG), Positron Emission Tomography (PET) and electroencephalography (EEG), a widely used modality is functional Near-Infrared Spectroscopy (fNIRS). It measures variations in cerebral oxygenation with high spatial resolution. The portability, low cost, and tolerance to motion of fNIRS make it an invaluable tool for conducting ecological studies, often in combination with EEG, thereby overcoming the constraints inherent to other high-resolution methods like fMRI and MEG [1].

Specifically in the field of attentional research, fNIRS has been instrumental in mapping the neural correlates of working memory, cognitive load, and executive control. In [2], fNIRS was employed to study tasks requiring the inhibition of distracting information and the temporal orienting of attention. In [3], a multi-factor complex paradigm, allowed the study of the hemodynamic cost associated with various cognitive demands, etc.

This study employs a paradigm to investigate the dynamic interplay between anticipation (SOA), attentional alerting (Mode), and cognitive conflict (Congruence) on Oxygenated Hemoglobin (HbO) dynamics [4]. The primary objective is to determine how these combined factors affect brain oxygenation in a rapid-event paradigm by quantifying key features of the hemodynamic response: the Amplitude, reflecting the peak intensity of the brain's response or activation level, and the Area Under the Curve (AUC), representing the cumulative intensity and duration of the response, or the total overall oxygen delivery over the task time.

Paper is organized as follows. Section 2 details the experimental approach. It is followed by fNIRS preprocessing, GLM modelization with SS regressors, features extraction and cleaning. Section 3 validates the GLM model and gives the main statistical findings, focusing on the significant Congruence×SOA interaction. Finally, Section 4 provides the conclusion and perspectives.

2 Methodology

2.1 Experimental paradigm

Fourteen brain optodes were placed in order to target the brain specific areas of interest : frontal lobe (particularly the prefrontal cortex), the parietal lobe (focusing more on its posterior part), and the upper part of the temporal lobe. The data was collected during an experiment conducted on 24 individuals. Each participant sat in front of a computer and had to press one of two keys on the keyboard ('/' at the bottom right or 'z' at the top left) depending on what was displayed on the screen. There were two categories of alerts: endogenous alert (a mechanism where participants prepared to act on their own in presence of an auditive stimuls with fixed intensity) and hybrid alert of endogenous and exogenous (use of an external auditive stimulus with a temporary increase in volume). Each trial consisted of displaying a coloured rectangle with two squares. If the rectangle was green (congruent), the participant had to press / if the square on the right lit up, or z if it was the square on the left lit down. If the rectangle was red (incongruent), the responses were reversed: z for the square on the right and / for the square on the left. In addition, the time interval between the warning signal (sound) and the imperative stimulus, called Stimulus Onset Asynchrony (SOA), varied in order to assess its impact on participants' performance. Four SOA conditions were tested: no delay (0 ms), 100 ms, 250 ms, and 850 ms [4].

2.2 fNIRS preprocessing

Raw fNIRS signals are often affected by noise, particularly due to movement or intrinsic physiological processes. In this work, fNIRS data processing was performed using the Homer3 interface [5]. The preprocessing steps are :

- Data conversion: the raw data was first converted into intensity, then into absorbance changes. Next, optical signals were transformed into concentrations of Oxyhemoglobin (HbO).
- Motion artefact removal: to correct for disturbances caused by participant movement, an automatic artefact detection method, followed by robust local regression were applied. This combination effectively eliminates transient anomalies while preserving the essential characteristics of the hemodynamic signal.
- Bandpass filtering: a filter between 0.03 and 0.15 Hz was applied to eliminate physiological and instrumental noise, while preserving the slow components related to brain activity.

2.3 GLM modelisation

The slowness of the hemodynamic response combined with the proximity of the trials causes activations to overlap, making it difficult to separate them temporally. In fact, the Hemodynamic Response Function (HRF) which reflects the change in blood oxygenation, typically peaks between 5 and 8 seconds after the stimulus onset and can take up to 20-30 seconds to fully return to baseline. Conversely, in this expermiment, the button-press trials are often separated by only a few seconds which causes the hemodynamic responses from successive trials to significantly overlap.

To successfully separate these successive trials and accurately estimate the brain's response to each experimental condition, a process called deconvolution is essential. It is achieved using the General Linear Model (GLM) which is based on a canonical hemodynamic model allowing to estimate the amplitude of the cerebral response for each experimental condition [6]. In presence of external noise, Short-Separation (SS) channels are added as nuisance regressors in the design.

The general formulation of GLM-SS is the following [7] :

$$\mathbf{Y} = \mathbf{X}_{\text{task}}\boldsymbol{\beta}_{\text{task}} + \mathbf{X}_{\text{SS}}\boldsymbol{\beta}_{\text{SS}} + \boldsymbol{\varepsilon} \qquad (1)$$

where $\mathbf{Y}$ is the measured fNIRS signal (here HbO) as a time series of length N, $\mathbf{X}_{\text{task}}$is task-related regressor (stimulus convolved with HRF), $\mathbf{X}_{\text{SS}}$ is short-separation regressor modeling superficial physiological noise, $\boldsymbol{\beta}_{\text{task}}$ are coefficients of interest representing task-related activation and $\boldsymbol{\beta}_{\text{SS}}$ are nuisance coefficients accounting for scalp/systemic components.

2.4 fNIRS based features

In order to characterise the hemodynamic response, nine robust features from the GLM model are extracted [8].They are illustrated in Fig. 1 and resumed as follows: Amplitude (maximum intensity or activation level), Latency or Time To Peak (time duration from the start of the response to peak), Time to Trough (duration from the start to maximum negative intensity), Rise Slope (rate of increase from the onset towards the peak), Fall Slope (rate of decrease from the peak towards the baseline), Peak to Trough (duration from the peak to the trough), Area Under the Curve or AUC (total energy of the activation) and Full Width at Half Maximum or FWHM (duration of activation above 50% of the maximum).

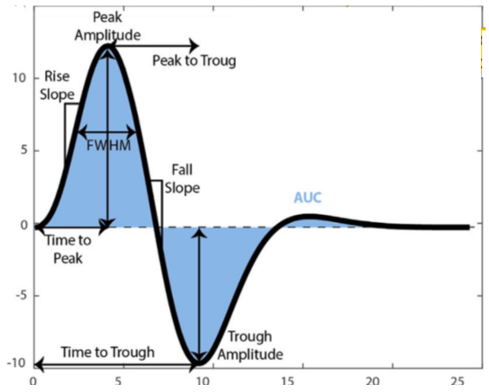


Figure 1: HRF signal and its parameters of interest

.

2.5 Statistical tests

To assess the effects of the experimental manipulations on the HbO hemodynamic response, Linear Mixed-Effects (LME) models were implemented for each extracted feature to know if the selected feature makes a statistical difference between the experimental conditions.This method was chosen in order to analyze data where measurements were repeatedly taken from the same participants (one to eight trials for each participant for each condition).

The LME considered the three experimental factors, namely the mode, the congruence and the SOA, along with their one way and two-way interactions. The one-way effects allow to examine how each variable individually modulated the HbO signal. The two-way interactions provides deeper insight into how these factors jointly influences the neural activation patterns.

To further explore the interaction effects, contrast analyses were performed to test specific hypotheses regarding the congruence factor within each SOA level. The post-hoc comparisons allow to determine whether the difference between congruent and incongruent trials remained constant or varied across temporal intervals. Hence, Bonferroni correction can be applied across all SOA levels.

3 Results

3.1 GLM modelisation

Fig. 2 compares fNIRS signal quality before (No GLM) and after applying a GLM-SS. The dashed lines represent the raw HbO signal (No GLM), while the solid lines show the reconstructed responses estimated by the GLM-SS, which better match the canonical hemodynamic response function (black line). Boxplots show that the GLM+SS correction significantly increases the correlation of HbO with the expected response ($pvalue < 0.001$). The RMSE $= \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \hat{Y}_i\right)^2}$ decreased from 0.19 µMol to 0.08 µMol. These results demonstrate that GLM-SS removes superficial physiological noise, yielding cleaner and more physiologically consistent hemodynamic responses.

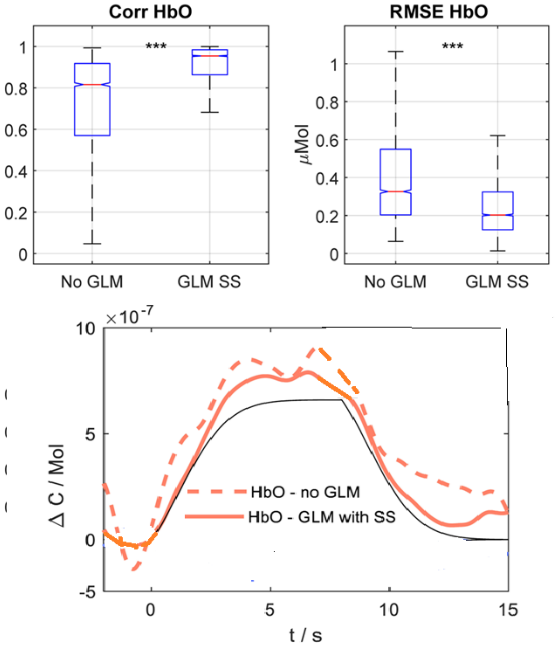


Figure 2: GLM modelisation results.

3.2 Statistical results

Results from these univariate models revealed that Congruence significantly influenced Amplitude ($p = 0.047$), indicating stronger hemodynamic responses in congruent compared to incongruent trials. Additionally, SOA showed a significant main effect on both Amplitude ($p < 0.001$) and RecoverySlope ($p = 0.048$), suggesting a temporal modulation of the hemodynamic response across stimulus onset asynchronies. No significant main effects of Mode were observed for any descriptor. Crucially, features related to temporal aspects of the response (PeakTime, Latency, OnsetTime, RecoveryTime, FWHM) did not show significant variations, possibly due to the temporal proximity of events and the averaging effect of the HRF model.

To investigate the temporal dependence of the attentional effect, a mixed-effects ANOVA was conducted to examine the interaction between Congruence and SOA on the two primary descriptors of intensity. The model revealed a significant main effect of Congruence ($p = 0.044$ for AUC and $p = 0.030$ for Amplitude), as well as a robust significant interaction effect between Congruence and SOA for both descriptors ($p < 0.001$ in both cases). This crucial finding indicates that the effect of congruence on the hemodynamic response depends on the temporal separation between cue and target.

Post-hoc contrasts (coefTest), corrected using a Bonferroni adjustment ($p < 0.0125$), further specified this modulation: the Congruence effect was significant for the shortest (0 ms) and longest (850 ms) SOA levels, for both AUC and Amplitude. These results suggest a bimodal modulation of attentional processing, where congruence enhances neural efficiency at the immediate and delayed stages of attentional deployment.

Fig. 3 illustrates these effects by displaying the mean HbO responses across SOA levels for both congruent and incongruent conditions. As shown, the AUC values are higher in the congruent condition at the shortest SOA, reflecting an early facilitation effect in attentional orienting. For longer SOAs, the difference between conditions diminishes, suggesting that the temporal separation between cue and target reduces the congruence-driven enhancement of the hemodynamic response. In contrast, Amplitude exhibits a similar but less pronounced modulation, indicating that while the total activation (AUC) is sensitive to temporal dynamics, the peak intensity of the response remains more stable across conditions.

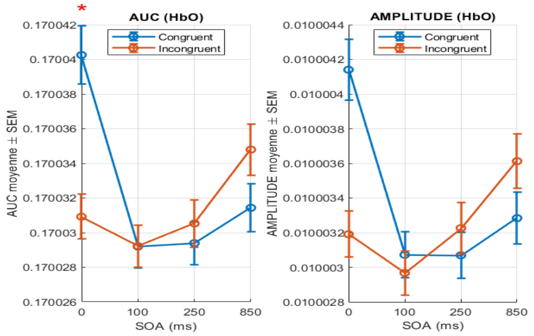


Figure 3: Interaction effects of Congruence and SOA on HbO's Amplitude and AUC.

4 Conclusion

This study investigated the dynamic temporal allocation of attention using fNIRS. It addressed the challenges of a rapid-event paradigm through a pipeline including signal preprocessing tasks and GLM+SS to ensure reliable HbO features. Statistical analysis, performed using LME, demonstrated a significant Congruence × SOA interaction which is maximized at immediate (SOA=0ms) and delayed (SOA=850ms) intervals. For future work, a multimodal integration by co-analyzing these fNIRS findings with EEG data will allow a critical alignment of the slow hemodynamic changes with fast electrophysiological activity, providing a complete spatio-temporal understanding of the observed bimodal attentional control pattern.

References

[1] Ayaz, H., Baker, W.B., Blaney, G., *et al*, " Optical Imaging and Spectroscopy for the Study of the Human Brain: Status Report," *Neurophotonics*, 2022, 9(S2):S24001.

[2] Ehlis, A. C., Schneider, S., Herrmann, M. J., and Fallgatter, A. J., "fNIRS in Psychiatry: a New Tool for the Objective Assessment of Cerebral Correlates of Psychiatric Disorders," *J Psychiatr Res*,2008; 42(7), 611-621.

[3] Quaresima, V., and Ferrari, M., "Functional near-infrared spectroscopy (fNIRS) for Assessing Cerebral Cortex Function during Exercise, Rehabilitation and Sports," *J Biophotonics*, 2012; 5(4), 304-315.

[4] McCormick C. R., Redden R. S., Hurst A. J. and Klein R. M., "On the Selection of Endogenous and hybrid Signals," *R. Soc. Open Sci.6190134*, 2019. http://doi.org/10.1098/rsos.190134.

[5] Boston University Neurophotonics Center. (s.d.), *Homer3: MATLAB application for fNIRS data processing and visualization*, Available at *https://github.com/BUNPC/Homer3.*

[6] Tak, S, Ye JC. "Statistical Analysis of fNIRS Data: a Comprehensive Review of Methodologies and Recent Advances," *Neurophotonics*, 2014;1(1):011002.

[7] Gagnon, L., Pinti P., Tachtsidis, I, *et al*, "Short Separation Channel Location Impacts the Performance of Short Channel Regression in NIRS," *NeuroImage*, 2012;59(3):2518-2528.

[8] West, K.L., Zuppichini, M.D., Turner, M.P,, Sivakolundu, D.K., Zhao, Y., Abdelkarim, D., Spence, J.S., Rypma, B., "BOLD Hemodynamic Response Function Changes Significantly with Healthy Aging," *Neuroimage*, 2019 Mar;188:198-207.