Communication-efficient Embedded FFT Processing for Acoustic Telemetry in LPWAN-based Beehive Monitoring Systems

Sebastian Górecki

Lodz University of Technology, Łódź, Poland

https://doi.org/10.26636/jtit.2026.3.2678

Abstract — Low-power acoustic telemetry remains a significant challenge in Internet of Things monitoring systems deployed in remote environments. This paper presents an embedded fast Fourier transform framework for communication-efficient acoustic monitoring of honey bee colonies. Instead of transmitting raw audio streams, the proposed approach extracts compact spectral descriptors directly on an embedded sensing node and transmits only a small feature vector using a low-power cellular network. The framework is evaluated using labeled queenright and queenless colony recordings. The proposed solution targets resource-constrained ESP32 class IoT nodes operating over LTE-M and NB-IoT networks. The analysis covered such parameters as dominant frequency, peak amplitude, mean spectral amplitude, spectral centroid, spectral entropy, and band energy extracted from the 200 – 400 Hz band. The results showed that dominant frequency alone did not significantly differentiate colony states at the file level, while mean spectral amplitude remained statistically significant. Queenless recordings also exhibited higher dominant frequency variability. The proposed approach reduces the transmitted payload by more than three orders of magnitude while remaining compatible with resource-constrained ESP32 class IoT devices. The results demonstrate that the extraction of embedded acoustic features is a practical method for scalable smart beehive monitoring under strict memory, power, and bandwidth constraints.

Keywords — acoustic telemetry, edge computing, Fast Fourier Transform, Internet of Things, wireless sensor networks

1. Introduction

Acoustic sensing is increasingly used in Internet of Things (IoT) monitoring systems operating in remote, battery-powered, and bandwidth-constrained environments. In such systems, continuous audio acquisition can provide useful diagnostic information, but raw audio transmission is often impractical due to memory usage, energy consumption, and communication cost. Unlike most previous studies focused on biological interpretation, this work addresses the challenge of acoustic data reduction faced in telecommunications settings, to facilitate low power IoT deployments. Unlike cloud-centric acoustic monitoring systems, the proposed framework explicitly prioritizes communication efficiency and embedded feasibility over high-capacity biological classification.

Honey bee colonies represent a particularly interesting application case for distributed acoustic telemetry. The colonies generate sound and vibration patterns associated with thermoregulation, foraging activity, queen presence, preparation of swarms, and colony stress [1]–[3]. Smart beehive systems commonly integrate temperature, humidity, weight, and acoustic detection to support precision apiculture and non-invasive colony monitoring [4]–[6].

Recent studies have highlighted the importance of energy-efficient sensing and edge intelligence in precision beekeeping systems deployed in remote environments [7]–[9]. Additionally, surveys indicate that acoustic sensing is becoming one of the key modalities (in addition to temperature, humidity and hive weight monitoring) relied upon in precision beekeeping platforms [10].

Works published in the telecommunications literature have highlighted the growing importance of edge intelligence and lightweight signal processing in acoustic wireless sensor networks [11], [12]. From the telecommunications and information technology perspective, the key challenge consists not only in biological interpretation of acoustic signals, but also in efficient transmission of relevant information. The process of sending raw audio from each hive is difficult to scale, especially when devices operate over LTE-M, NB-IoT, or other low-power wireless communication technologies. A more efficient solution is to perform local signal processing on the embedded node and transmit only compact descriptors extracted from the frequency domain. Low-power wide-area communication technologies such as NB-IoT and LTE-M have become increasingly important for remote sensing deployments, where bandwidth and energy consumption must be minimized [7].

The reliability of LPWAN communications can be affected by interference and coverage limitations, which makes communication-efficient telemetry particularly important in remote deployments [13].

This paper presents a communication-efficient embedded fast Fourier transform (FFT) framework for acoustic telemetry

Tab. 1. Representative acoustic frequency ranges reported in the literature dealing with honey bee monitoring systems.

Condition / behavior	Frequency range	Interpretation	References
Ventilation and thermoregulation	50 – 190 Hz	Hive ventilation and microclimate regulation	[3]
Queen-related activity	160 – 280 Hz	Acoustic patterns associated with queen presence, queen acceptance, or queenless state	[16]
Foraging and nectar processing	200 – 300 Hz	Active colony work, nectar handling, and foraging related activity	[3], [17]
Intensive colony activity	240 – 400 Hz	Increased colony activity and internal hive logistics	[5]
Pre-swarming activity	400 – 500 Hz	Increased excitation and swarm preparation	[1], [14]
Swarm-related signals	500 – 550 Hz	Acoustic patterns associated with swarm mobilization	[14]

in smart beehive monitoring systems. The main objective is not to develop a production-ready biological classifier, but to evaluate whether compact FFT-derived descriptors retain diagnostically relevant information while substantially reducing communication payload.

The main contributions of this work are as follows:

1) embedded FFT-based acoustic feature extraction framework for smart beehive IoT nodes,
2) communication reduction analysis comparing raw audio transmission with compact acoustic telemetry,
3) experimental evaluation using queenright and queenless honey bee colony recordings,
4) comparison of dominant-frequency, amplitude-based, and spectral-shape descriptors,
5) file-level statistical analysis mitigating the risk of overinterpreting non-independent audio segments.

The rest of the paper is structured as follows. Section 2 reviews related works. Section 3 presents the proposed embedded acoustic telemetry framework. Section 4 describes the FFT-based feature extraction method. Section 5 introduces the experimental data set. Section 6 reports the results, while Section 7 discusses limitations and deployment-related implications. Section 8 concludes the paper.

2. Related Works

Precision apiculture increasingly relies on non-invasive sensing systems that observe colony dynamics without opening the hive. Existing systems measure internal temperature, humidity, hive weight, and acoustic activity, often transmitting the collected telemetry to a cloud platform for analysis [4]–[6].

Acoustic monitoring has received significant attention, as bee-generated sounds can reflect colony activity, swarm preparation, queen status, and stress responses [1], [3]. The authors of [1] demonstrated that the spectral characteristics of the hive sounds can be used for early swarm detection, while paper [2] stated that the vibrational signatures contain biologically meaningful information on colony dynamics. Recent studies have also explored embedded or machine learning-based approaches to bee sound analysis [8], [14].

A comprehensive review of automated beehive acoustic monitoring was recently presented in [15], highlighting the growing interest in signal processing and machine learning approaches.

However, many acoustic monitoring approaches focus primarily on biological classification or offline analysis. Less attention has been paid to the communication and embedded system aspects of large-scale acoustic telemetry. In practical IoT deployments, transmitting raw audio is often not feasible. Therefore, edge-based signal processing is needed to reduce data volume while preserving diagnostically useful information.

Table 1 summarizes the representative frequency ranges reported in previous beehive acoustic studies. These ranges should be interpreted as indicative rather than universal diagnostic thresholds, because measured values may depend on colony condition, microphone placement, hive type, environmental conditions, and sensor characteristics.

Although acoustic monitoring has been extensively investigated, most studies focused on biological interpretation or offline signal analysis. Only limited attention has been paid to embedded implementation constraints and reduction in communications payloads. A comparison of representative approaches is presented in Tab. 2.

As shown in Tab. 2, most previous studies focused primarily on the biological interpretation of hive acoustics, while the proposed approach explicitly addresses embedded implementation constraints and communication reduction.

Tab. 2. Comparison of acoustic monitoring approaches reported in the literature.

Ref.	FFT	Edge	Embedded	Comm.
[1]	Yes	No	No	No
[2]	Yes	No	No	No
[14]	Partial	Partial	No	No
This work	Yes	Yes	Yes	Yes

Fig. 1. Acoustic sensing hardware used in smart beehive monitoring systems: waterproof microphone module (left) and microphone installed at the hive entrance during field deployment (right).

3. Proposed Embedded Acoustic Telemetry Framework

The proposed framework is intended for smart beehive monitoring systems equipped with a microphone, an embedded microcontroller, and a wireless communication module. The system acquires short acoustic buffers, performs local FFT analysis, extracts compact descriptors, and transmits only the resulting feature vector.

Figure 1 shows the acoustic sensing hardware used in the BeeHUB monitoring system. The microphone may be installed near the entrance to or inside the hive, depending on the monitoring objective and deployment constraints.

The conceptual architecture of the proposed acoustic telemetry chain is shown in Fig. 2. Acoustic signals are processed locally on the embedded node. Only extracted descriptors, such as amplitude, energy, centroid, entropy, and dominant frequency, are transmitted to the cloud platform.

The embedded processing workflow is presented in Fig. 3. The device periodically initializes the microphone, acquires a short buffer, computes the FFT, extracts descriptors from the selected frequency band, transmits the compact packet, and returns to sleep mode.

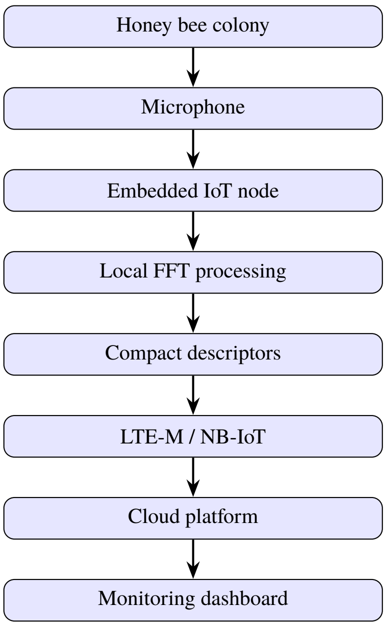

Fig. 2. Architecture of the proposed embedded acoustic telemetry framework.

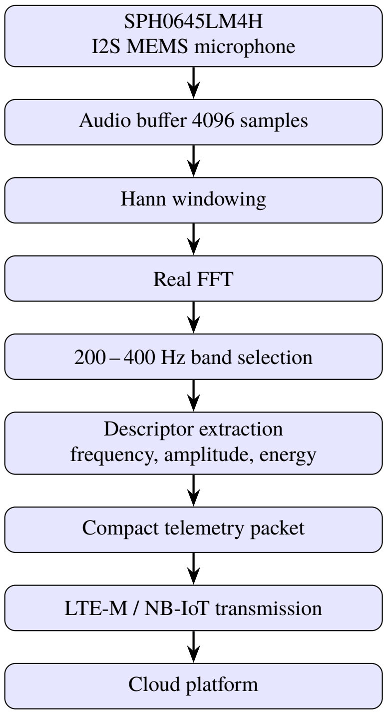

Fig. 3. Signal processing pipeline of the proposed integrated acoustic telemetry framework.

4. FFT-based Feature Extraction Method

4.1. Signal Processing

Fast Fourier Transform (FFT) remains one of the most widely used methods for frequency domain analysis in embedded signal processing systems due to its computational efficiency and suitability for resource-constrained hardware [18]–[20].

For each audio segment, a real-valued FFT was calculated. The discrete Fourier transform of the input signal is defined as:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N} , \tag{1}$$

where $x(n)$ is the time-domain acoustic signal, N is the number of samples, and $X(k)$ is the complex frequency-domain representation.

The frequency resolution is given by:

$$\Delta f = \frac{f_s}{N} , \tag{2}$$

where f_s is the sampling frequency. For the target embedded configuration with $f_s = 8192$ Hz and $N = 4096$, the frequency resolution is 2 Hz.

The FFT window duration is:

$$T_w = \frac{N}{f_s} . \tag{3}$$

For $N = 4096$ and $f_s = 8192$ Hz, one FFT window corresponds to 0.5 s of audio. Multiple short windows may be analyzed within one measurement cycle to observe a longer acoustic interval without storing a full raw audio recording.

Prior to FFT computation, a Hann window was applied to each audio segment to reduce spectral leakage resulting from finite-length signal acquisition.

4.2. Extracted Descriptors

The analysis focused on the 200 – 400 Hz frequency band. This band was selected because it overlaps with the acoustic activity ranges reported for foraging, nectar processing, colony work and selected queen- or stress-related states.

The dominant frequency was calculated as follows:

$$f_{dom} = f\left(\arg \max_{f\in[f_{min},f_{max}]} \left|X(f)\right|\right) . \tag{4}$$

In addition to dominant frequency, the following descriptors were evaluated:

1) peak amplitude,
2) mean spectral amplitude,
3) spectral centroid,
4) spectral entropy,
5) band energy.

The spectral centroid was calculated as:

$$C = \frac{\sum_i f_i A_i}{\sum_i A_i} , \tag{5}$$

where f_i is the frequency bin and A_i is the corresponding spectral magnitude.

The normalized spectral entropy was calculated from the probability distribution of spectral magnitudes:

$$H = -\sum_i p_i \log_2(p_i) , \tag{6}$$

where

$$p_i = \frac{A_i}{\sum_i A_i} .$$

Band energy was calculated as:

$$E = \sum_i A_i^2 . \tag{7}$$

The FFT stage dominates the computational complexity of the proposed pipeline with complexity $O(N \log N)$, whereas the remaining descriptors are computed in linear time. Owing to this computational profile, real-time execution remains feasible on modern microcontrollers such as ESP32 [18], [19].

4.3. Embedded Implementation Parameters and Benchmark

Table 3 summarizes the target embedded implementation parameters used in the proposed FFT-based acoustic telemetry framework. The selected configuration is compatible with ESP32-class IoT nodes and avoids raw audio transmission by extracting compact spectral descriptors locally.

To address the feasibility of the proposed FFT-based processing pipeline on the target hardware platform, an embedded benchmark was designed for the ESP32-class microcontroller used in the BeeHUB acoustic sensing node. The benchmark measures the execution time of one FFT processing cycle, determines the maximum memory usage, the size of the descriptor payload, and the communication reduction achieved by local feature extraction.

Tab. 3. Implementation parameters and resource requirements.

Parameter	Value
Microcontroller class	ESP32-based IoT node
Target sampling frequency	8192 Hz
FFT length	4096 samples
Frequency resolution	2 Hz
Window duration	0.5 s
Analysis duration	12 s
Number of FFT windows	24
Analysis band	200 – 400 Hz
Candidate features	Frequency, amplitude, energy, centroid, entropy
Input audio buffer	8 KB
FFT working buffer	8 KB
Typical feature vector	8 – 24 bytes
Estimated RAM requirement	$\leqslant$ 18 KB
Estimated firmware Flash	$\leqslant$ 100 KB
Wireless technology	LTE-M / NB-IoT

The benchmark procedure consists in acquiring 4096 sample audio windows at 8192 Hz, applying FFT-based spectral analysis, extracting compact descriptors from the 200 – 400 Hz band, and preparing the resulting telemetry payload for wireless transmission.

The same processing configuration as in the offline analysis is used to ensure consistency between the experimental validation and the target embedded implementation. The benchmark reused the original implementation of the production firmware without modifying the FFT processing path, ensuring that the reported execution times accurately reflect the deployed BeeHUB sensing node.

The firmware-level benchmark was performed using the FFT implementation of the ESP32 acoustic pipeline. The benchmark compiled the production firmware FFT implementation in a host side test harness and reproduced the processing structure used by the embedded acoustic acquisition routine.

The measured FFT runtime was 0.017 ms per a 4096-sample window, while the total processing time including descriptor extraction was 0.227 ms. The estimated peak heap usage was 98 344 bytes, and the FFT object Flash footprint was 4578 bytes. The measured processing time of 0.227 ms represents less than 0.05% of the 0.5-second acquisition window, confirming that the proposed processing pipeline easily satisfies real-time constraints on the target embedded platform (see Tab. 4).

Consequently, the benchmark evaluates the identical FFT implementation and processing path used by the production BeeHUB firmware, although timing measurements were obtained using a host side benchmarking harness. The measured peak heap usage remains well below the available RAM of typical ESP32 devices, confirming that the proposed implementation is compatible with low-cost embedded hardware.

Tab. 4. Firmware-level FFT benchmark and payload reduction results.

Metric	Value
Benchmark type	Host-side firmware benchmark
FFT implementation	ESP32 firmware FFT code
FFT size	4096 samples
FFT runtime	0.017 ms
Total processing time	0.227 ms
Peak heap usage	98 344 bytes
Measured FFT module Flash footprint	4 578 bytes
Single descriptor packet payload	4 bytes
Raw I2S buffer payload	32 768 bytes
Reduction factor	8192×
Estimated LTE-M airtime	0.085 ms vs. 699.1 ms
Estimated energy reduction	8192×
Measured board sleep current	62.25 µA
Processing utilization	0.045% of acquisition window

Tab. 5. Experimental audio dataset and recording characteristics.

Parameter	Value
Queenright recordings	6
Queenless recordings	8
Total recordings	14
Queenright segments	36
Queenless segments	48
Total analyzed segments	84
Original recording duration	829 s
Analyzed duration	168 s
Sampling rate	44.1 kHz
Resampled analysis rate	8192 Hz
Segment duration	2 s
Analysis window per file	12 s

The raw I2S buffer used by the embedded implementation contains 4096 frames with two 32-bit channels, corresponding to 32 768 bytes per FFT window. In contrast, the minimal transmitted acoustic descriptor packet consists of a 16-bit frequency value and a 16-bit amplitude value, corresponding to 4 B. In the extended feature-vector configuration used for communication analysis, six 32-bit descriptors correspond to 24 bytes. This corresponds to an 8192-fold reduction in the transmitted payload. Under a simplified LTE-M payload-only airtime model at 375 kbps, this corresponds to a reduction from approximately 699.1 ms for the raw buffer to 0.085 ms for the descriptor payload.

The measured embedded processing time of 0.227 ms occupies only 0.045% of the 0.5 s acquisition interval and remains more than three orders of magnitude lower than the estimated LTE-M transmission time for raw audio (699.1 ms). This confirms that communication latency dominates the overall sensing cycle.

In addition, a board-level measurement of the low power state performed with all sensors connected and without the programmer attached showed a sleep current of approximately 62.25 µA. This value characterizes the quiescent operation of the sensing node between measurement cycles and should not be confused with the average active current reported for the complete sensing cycle power profile in Section 6.

5. Experimental Dataset and Statistical Analysis

The experimental data set consisted of WAV recordings obtained from honey bee colonies labeled as queenright and queenless. The recordings were originally collected as audio files and processed offline in Python using NumPy and SciPy. Files beginning with `q` were treated as queenright recordings, whereas those beginning with `nq` were treated as queenless recordings.

5.1. Audio Preprocessing

Before feature extraction, all recordings were processed using a standardized preprocessing workflow to improve comparability between measurements collected under field conditions. The audio recordings were converted to uncompressed WAV format and analyzed using a common processing pipeline implemented in Python.

The original recordings were acquired at a sampling frequency of 44.1 kHz. To ensure consistency with the intended embedded implementation, all recordings were resampled to the target sampling frequency of 8192 Hz prior to feature extraction. This resampling configuration matches the proposed ESP32-based deployment and provides a frequency resolution of 2 Hz for a 4096-point FFT. The resampled signals were subsequently divided into fixed-length 2-second segments for descriptor extraction (Tab. 5).

The target embedded implementation uses a SPH0645LM4H digital MEMS microphone with an I2S interface [21]. According to the manufacturer's specification, the microphone operates over a frequency range of approximately 50 to 15 kHz and supports a supply voltage of 1.6 to 3.6 V. The digital I2S interface eliminates analog gain-stage variability, although microphone placement and hive-specific acoustic coupling may still influence the measured amplitude.

To reduce the influence of recording-level amplitude differences caused by microphone placement or acquisition conditions, amplitude normalization was applied before FFT analysis. Normalization was performed independently for each analyzed segment while preserving the relative spectral characteristics within the selected frequency band.

Each segment was subsequently windowed using a Hann window to reduce spectral leakage before FFT computation. The frequency domain descriptors were then extracted from the 200 – 400 Hz band, including the dominant frequency,

Tab. 6. Effect of segment duration on extracted acoustic descriptors.

State	Duration	Mean amplitude	Dominant frequency
Queenright	2 s	115508.73	248.88 Hz
Queenright	12 s	128201.93	250.88 Hz
Queenright	30 s	122356.53	251.50 Hz
Queenless	2 s	425703.50	271.08 Hz
Queenless	12 s	1967941.00	250.80 Hz
Queenless	30 s	2066559.00	224.00 Hz

the peak amplitude, the mean spectral amplitude, spectral centroid, the spectral entropy, and the band energy.

Because the recordings originated from practical field deployments rather than controlled laboratory measurements, absolute amplitude values should be interpreted with caution. Consequently, amplitude-based descriptors are considered relative diagnostic features rather than calibrated sound-pressure measurements.

In general, the analyzed dataset comprised 14 recordings, including six queenright and eight queenless colonies. Each file was analyzed over the first 12 s and divided into six 2 second segments. This resulted in 36 queenright and 48 queenless segments, for a total of 84 analyzed segments and 168 s of analyzed audio. The original recordings were sampled at 44.1 kHz.

Because segments extracted from the same recording are not fully independent observations, two levels of statistical analysis were performed. First, segment-level analysis was used to examine the behavior of the descriptor in all 84 short windows. Second, file-level aggregation was performed by averaging descriptors within each audio recording, resulting in 14 independent file-level observations.

The Mann-Whitney U test was used for non-parametric comparison between queenright and queenless groups. The effect size was quantified using Cliff's delta. Bootstrap 95% confidence intervals were calculated for the amplitude descriptors to provide uncertainty estimates. Because multiple acoustic descriptors were evaluated, the statistical significance of the obtained p-values was additionally interpreted using the Bonferroni correction.

6. Results

6.1. Example FFT Spectrum

Figure 4 presents an example of the FFT spectrum extracted from a beehive recording. The highlighted 200 – 400 Hz region was used for the extraction of the descriptor.

6.2. Segment Duration Analysis

Table 6 summarizes the effect of the duration of the segment on the extracted descriptors. Queenright recordings showed relatively stable dominant frequency estimates across segment durations, while queenless recordings exhibited greater variability.

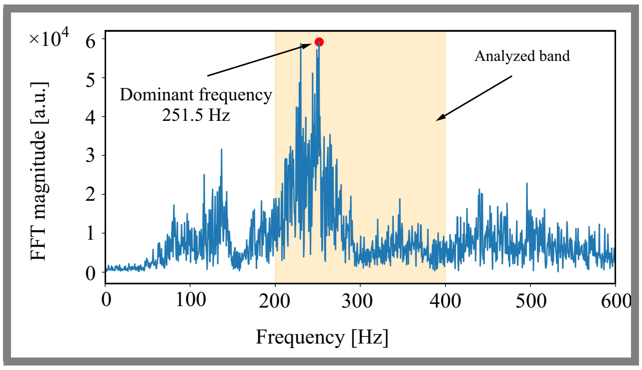

Fig. 4. Example FFT spectrum extracted from a beehive recording. The 200 – 400 Hz band was used for descriptor extraction.

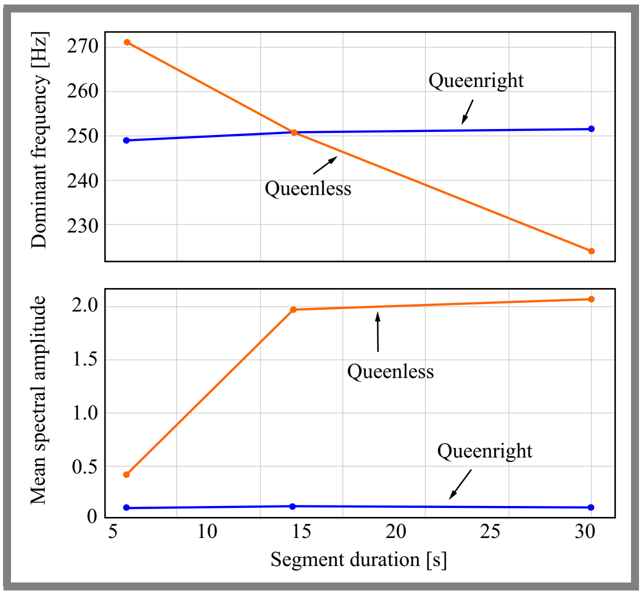

Fig. 5. Influence of segment duration on dominant frequency and mean spectral amplitude descriptors.

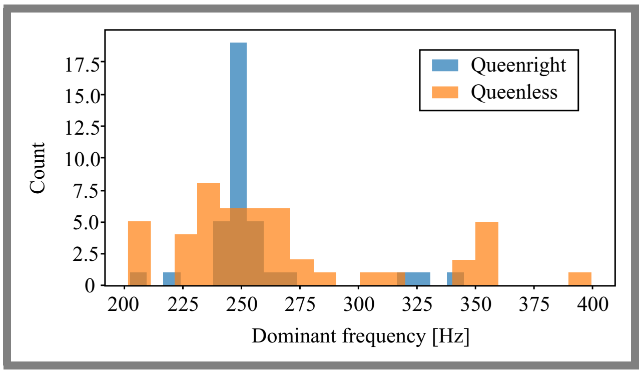

Fig. 6. Distribution of dominant frequencies extracted from queenright and queenless recordings in the 200 – 400 Hz band.

6.3. Dominant Frequency and Mean Amplitude

The extracted dominant frequencies were compared between queenright and queenless recordings. At the segment level, the dominant frequency did not significantly differentiate the groups. However, the amplitude-based descriptors showed clearer separation.

Tab. 7. Segment-level comparison of FFT-derived descriptors between queenright and queenless recordings.

Descriptor	Queenright mean	Queenless mean	Queenright SD	Queenless SD	p	δ
Dominant freq. [Hz]	254.76	265.93	26.00	48.98	0.674	–0.054
Peak amplitude	7.50×10^5	1.44×10^6	2.54×10^5	3.39×10^6	2.92×10^{-8}	0.711
Mean amplitude	1.38×10^5	3.09×10^5	4.43×10^4	6.57×10^5	1.02×10^{-8}	0.734
Spectral centroid [Hz]	273.82	293.15	9.58	16.40	5.40×10^{-8}	–0.697
Spectral entropy	8.17	8.28	0.13	0.24	8.87×10^{-6}	–0.569
Band energy	1.45×10^{13}	2.98×10^{14}	9.00×10^{12}	8.15×10^{14}	7.84×10^{-9}	0.740

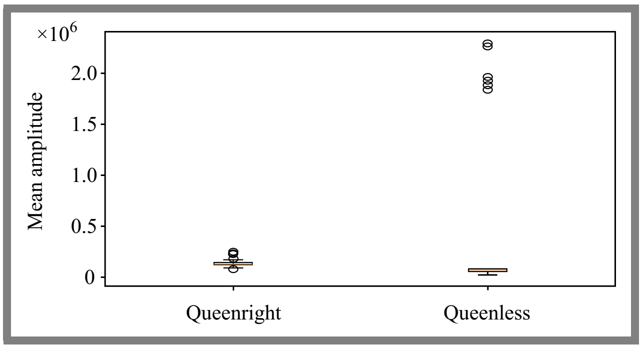

Fig. 7. Comparison of the mean spectral amplitude in the 200 – 400 Hz band for queenright and queenless recordings.

6.4. Extended FFT-derived Descriptors

Table 7 summarizes the segment-level statistical comparison of the FFT-derived descriptors. The dominant frequency was not statistically significant, while peak amplitude, mean amplitude, spectral centroid, spectral entropy, and band energy differed significantly between the two groups.

The strongest segment-level effect sizes were observed for band energy, mean amplitude, and peak amplitude. These results indicate that energy-related and spectral-shaped descriptors contain substantially more diagnostic information than dominant frequency alone.

The file-level effect size for mean spectral amplitude was large (Cliff's $\delta = 0.75$), indicating a substantial practical difference between queenright and queenless recordings despite the limited dataset size.

6.5. File-level Analysis

To reduce the risk of overinterpreting dependent segments, the descriptors were also aggregated at file level. Table 8 summarizes the statistical comparison.

At the file level, the mean dominant frequency was 254.76 Hz for queenright recordings and 265.93 Hz for their queenless counterparts. This difference was not statistically significant ($p = 0.216$). In contrast, mean spectral amplitude differed

Tab. 8. File-level statistical comparison of acoustic descriptors.

Feature	p-value	Cliff's δ	Result
Dominant frequency	0.216	–0.417	n.s.
Mean amplitude	0.0236	0.750	Significant

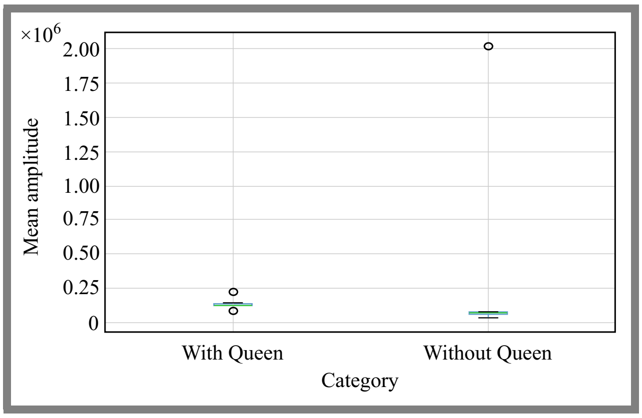

Fig. 8. File-level comparison of mean spectral amplitude. Each point represents an independent audio recording.

significantly between the groups. Queenright recordings had a mean amplitude of 138262, with a bootstrap 95% confidence interval of [124661, 153062], whereas queenless recordings had a mean amplitude of 308576, with a bootstrap 95% confidence interval of [141550, 508733] ($p = 0.0236$, Cliff's $\delta = 0.750$).

6.6. Frequency Stability

In addition to mean dominant frequency, stability was evaluated using the coefficient of variation (CV). Queenright recordings exhibited a CV of 3.75%, whereas queenless recordings showed considerably higher variability (CV of 6.97%). This suggests that frequency stability may provide additional diagnostic information beyond the mean dominant frequency alone.

6.7. Communication and Energy Efficiency

A major advantage of embedded acoustic processing is the reduction of communication load. The transmission of raw audio requires substantially more data than the transmission of compact FFT-derived descriptors. For mono-audio sampled at $f_s = 8192$ Hz with 16-bit resolution over an analysis duration of $T = 12$ s, the raw payload size is:

$$S_{raw} = f_s \cdot T \cdot \frac{b}{8}\ ,$$

where b is the number of bits per sample. This gives:

$$S_{raw} = 8\,192 \cdot 12 \cdot 2 = 196\,608\ \text{bytes}\ .$$

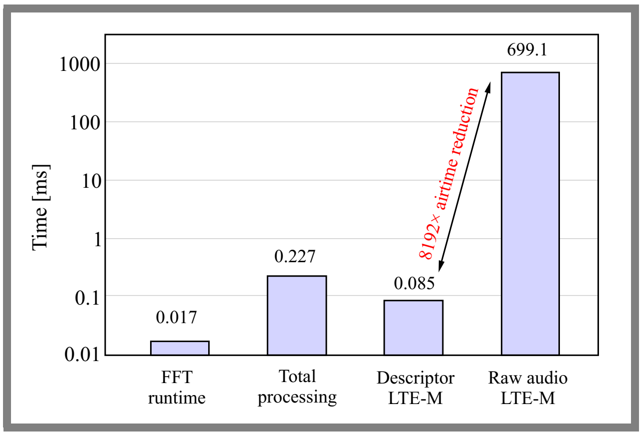

Fig. 9. Comparison of embedded processing latency and estimated LTE-M transmission time, illustrating the communication advantage of local FFT processing.

In contrast, the proposed embedded approach transmits only a small descriptor vector containing frequency and amplitude domain features. If each descriptor is encoded as a 32-bit floating-point value and six descriptors are transmitted, the resulting payload is:

$$S_{desc} = 6 \cdot 4 = 24 \text{ B} .$$

Therefore, the payload reduction ratio is:

$$R = \frac{S_{raw}}{S_{desc}} .$$

For the analyzed configuration, this corresponds to a reduction ratio of approximately 8 192:1.

Figure 9 shows that the execution time for local FFT processing is several orders of magnitude shorter than the time required to transmit raw audio over LTE-M. The total embedded processing time of 0.227 ms remains more than three orders of magnitude lower than the estimated LTE-M transmission time for raw audio of 699.1 ms. Consequently, communication latency dominates the overall sensing cycle, confirming that edge-based spectral feature extraction is substantially more efficient than raw audio transmission.

The substantial reduction in the transmitted payload directly translates into shorter radio activity periods. To complement this theoretical communication analysis, the complete sensing node was experimentally profiled using a Nordic Power Profiler Kit II.

The proposed approach shifts the computational burden from wireless communication to local signal processing, which is advantageous for battery-powered LPWAN deployments. By

Tab. 9. Communication payload reduction achieved by embedded FFT feature extraction.

Representation	Payload size	Relative load
Raw audio, 12 s, 8192 Hz, 16 bit	196 608 bytes	1.0
2 descriptors	8 bytes	4.07×10^{-5}
6 descriptors	24 bytes	1.22×10^{-4}
8 descriptors	32 bytes	1.63×10^{-4}

Tab. 10. Experimental power consumption of the complete sensing node.

Metric	Value
Supply voltage	3.6 V
Average current	93.99 mA
Peak current	1.16 A
Measurement duration	24.54 s
Measured charge	2.31 C
Average power	338.4 mW
Energy per sensing cycle	8.31 J
Measurement scope	Complete node with all sensors connected

Fig. 10. Example of a field observation conducted in an apiary near Konin, Poland, during a period of elevated colony activity preceding a swarming event.

reducing the transmitted payload from 196 608 to 8 – 24 bytes, the proposed framework substantially shortens radio activity duration and decreases communication overhead while preserving diagnostically relevant acoustic information (Tab. 9). Such communication-efficient processing is consistent with the growing trend towards edge intelligence in IoT systems.

Power profiling was performed using a Nordic Power Profiler Kit II at a supply voltage of 3.6 V. The measurement covered one complete sensing cycle with all connected sensors, including sensor initialization, FFT processing, wireless modem activity, and return to idle state. The average current was 93.99 mA for 24.54 s, corresponding to a measured charge of 2.31 C and an energy consumption of approximately 8.31 J per cycle. The peak current reached 1.16 A during short periods of high activity, which is consistent with wireless modem activity and sensor initialization (Tab. 10).

6.8. Field Observation Example

Swarming is one of the most important colony states that motivates automated acoustic monitoring. Figure 10 presents a field observation of increased colony activity during a pre-swarming period.

6.9. Multi-modal Hive Observation

Figures 11 and 12 present an example of multimodal hive monitoring using the BeeHUB telemetry platform.

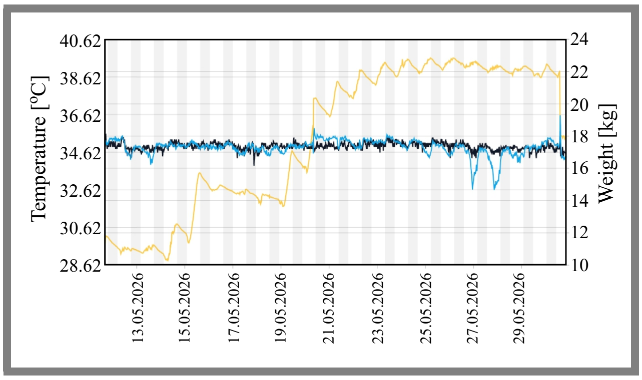

Fig. 11. Hive weight and internal temperature during a field observation period associated with nectar flow and increased colony activity.

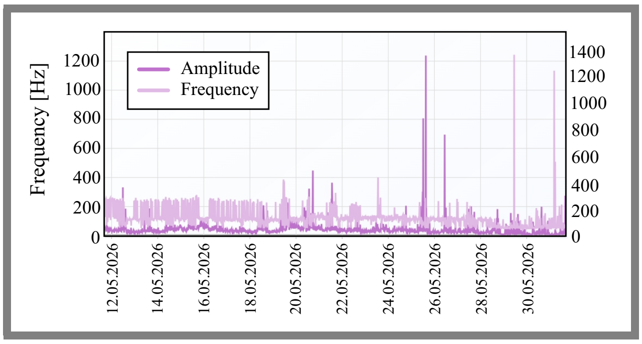

Fig. 12. Acoustic descriptors recorded during the same observation period. Most measurements remained between 200 and 300 Hz, while short-term anomalies exceeded 800 Hz.

During the observation period, the weight of the hive increased from approximately 11 kg to more than 22 kg, indicating intensive accumulation and nectar flow. At the same time, the internal temperature of the hive remained relatively stable between 34 and 35°C, suggesting normal brood thermoregulation.

The acoustic observations revealed dominant frequencies typically ranging between 200 and 300 Hz, which is consistent with colony activity associated with foraging, nectar processing, and ventilation reported in the literature. Several short-duration anomalies exceeded 800 Hz. Such events may indicate periods of elevated colony excitation and could potentially precede swarming-related behavior or other colony disturbances. More studies based on larger datasets are required to verify this hypothesis.

7. Discussion

The obtained results support the main assumption of the proposed framework: useful acoustic descriptors can be extracted locally from short FFT windows and transmitted as compact telemetry variables. This reduces the need for raw audio storage and enables acoustic monitoring in bandwidth-constrained IoT deployments.

The proposed framework follows the broader trend of moving signal processing from cloud platforms to resource-constrained edge devices [8].

Recent studies have investigated TinyML-based approaches for detecting acoustic events directly on embedded devices [11].

Although such methods may provide higher classification capabilities, they usually require trained machine learning models, additional memory resources, and increased computational complexity.

In contrast, the proposed FFT-based approach relies on deterministic signal processing and does not require model training, inference engines, or neural network parameters. As a result, the method remains lightweight, interpretable, and suitable for deployment on low-cost ESP32-class devices operating under strict memory and energy constraints.

An important finding is that dominant frequency alone did not reliably distinguish queenright and queenless recordings. This is relevant because many simple acoustic monitoring approaches focus on tracking a single dominant frequency value. In the present data set, the amplitude- and energy-based descriptors were more informative. At the segment level, band energy, mean amplitude, and peak amplitude showed the largest effect sizes. At the file level, the mean amplitude remained statistically significant despite the small number of independent recordings.

File-level analysis is especially important, since audio segments extracted from the same recording cannot be treated as fully independent observations. The segment-level results should therefore be interpreted as descriptor-level evidence, while the file-level results provide a more conservative assessment.

The study has several limitations. In general, the analyzed material comprised 14 recordings, and the queenright and queenless labels were based on preliminary manual classification. The recordings did not include complete biological metadata such as brood status, nectar flow, Varroa infestation, weather conditions, hive type, or various microphone placements. Therefore, the results should be treated as a validation of the extraction and communication reduction of embedded features rather than as a final biological diagnostic model.

Therefore, this data set should be regarded as a proof-of-concept data set intended for validating the proposed embedded processing framework, rather than an attempt to develop a production-ready biological classifier.

The proposed approach may also be generalized to other resource-constrained IoT acoustic monitoring applications, including environmental sensing, industrial condition monitoring, smart agriculture, wildlife monitoring, and distributed edge intelligence systems operating over LPWAN, LTE-M and NB-IoT networks.

The segment-duration experiment showed that 2-second segments already preserve the dominant acoustic characteristics while minimizing memory requirements and processing latency. Longer segments increased computational cost without providing a proportional improvement in the stability of the descriptor. Therefore, 2-second windows were selected as a compromise between spectral stability and embedded system constraints.

Although the mean dominant frequency did not differ significantly between queenright and queenless colonies, queenless recordings exhibited a dominant-frequency variability that was almost two times higher than in the case of queenright recordings (CV of 6.97% versus 3.75%). One possible explanation is that queenless colonies exhibit less stable collective behavior and increased acoustic heterogeneity.

Although the present study does not attempt to establish biological causality, the observed increase in frequency variability suggests that stability-related acoustic metrics may provide additional diagnostic value beyond mean frequency alone. In practical NB-IoT and LTE-M deployments, such payload reduction directly translates into shorter radio activity periods, lower energy consumption, and improved battery lifetime.

8. Conclusions

This paper presented a communication-efficient embedded FFT framework for acoustic telemetry in smart beehive IoT systems. The proposed approach extracts compact FFT-derived descriptors, including dominant frequency, amplitude-based measures, spectral centroid, spectral entropy, and band energy. The method enables local signal processing on resource-constrained nodes and substantially reduces communication requirements compared to raw audio transmission.

The primary contribution of this work is not the biological classification of colony states, but the demonstration that diagnostically relevant acoustic information can be retained while reducing communication payload by approximately four orders of magnitude (up to 8192-fold).

This enables scalable deployment of acoustic sensing in large-scale IoT monitoring systems operating under bandwidth and energy constraints.

Experimental validation using queenright and queenless recordings showed that dominant frequency alone was not a reliable discriminator. In contrast, amplitude-based and energy-related descriptors were more informative. File-level analysis confirmed that mean spectral amplitude remained statistically significant between groups, while queenless recordings also showed higher dominant frequency variability. The embedded benchmark further demonstrated that the proposed processing pipeline satisfies real-time constraints while requiring only modest computational resources on an ESP32-class microcontroller.

The proposed framework provides a practical and scalable basis for acoustic telemetry in precision apiculture and demonstrates how embedded edge processing can substantially reduce communication overhead in LPWAN-based IoT systems. Future work will include validation on substantially larger multi-season datasets together with direct measurements of embedded execution time, communication energy consumption, and long-term field deployment. The proposed framework may also be adapted to other LPWAN-based embedded acoustic sensing applications where communication efficiency is a primary concern.

Acknowledgments

The author thanks the beekeepers who supported the collection of acoustic recordings and field observations relied upon in this study.

References

[1] S. Ferrari, M. Silva, M. Guarino, and D. Berckmans, "Monitoring of Swarming Sounds in Bee Hives for Early Detection of the Swarming Period", *Computers and Electronics in Agriculture*, vol. 64, pp. 72–77, 2008 (https://doi.org/10.1016/j.compag.2008.05.010).

[2] M. Bencsik *et al.*, "Honeybee Colony Vibrational Measurements to Highlight the Brood Cycle", *PLOS ONE*, vol. 10, art. no. 0141926, 2015 (https://doi.org/10.1371/journal.pone.0141926).

[3] A. Terenzi, S. Cecchi, and S. Spinsante, "On the Importance of the Sound Emitted by Honey Bee Hives", *Veterinary Sciences*, vol. 7, art. no. 168, 2020 (https://doi.org/10.3390/vetsci7040168).

[4] W.G. Meikle and N. Holst, "Application of Continuous Monitoring of Honeybee Colonies", *Apidologie*, vol. 46, pp. 10–22, 2014 (https://doi.org/10.1007/s13592-014-0298-x).

[5] H. Hadjur, D. Ammar, and L. Lefèvre, "Toward an Intelligent and Efficient Beehive: A Survey of Precision Beekeeping Systems and Services", *Computers and Electronics in Agriculture*, vol. 192, art. no. 106604, 2022 (https://doi.org/10.1016/j.compag.2021.106604).

[6] S. Górecki *et al.*, "Improving Bee Living Conditions Through Ecological Thermal Insulation and Remote Early Anomaly Detection-vital Step Towards Preserving Bees Population", *Fibres and Textiles in Eastern Europe*, vol. 32, pp. 1–12, 2024 (https://doi.org/10.2478/ftee-2024-0025).

[7] H. Hadjur, D. Ammar, and L. Lefèvre, "Analysis of Energy Consumption in a Precision Beekeeping System", *Proc. of the 10th International Conference on the Internet of Things*, art. no. 20, 2020 (https://doi.org/10.1145/3410992.3411010).

[8] A. De Simone, L. Barbisan, G. Turvani, and F. Riente, "Advancing Beekeeping: IoT and TinyML for Queen Bee Monitoring Using Audio Signals", *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–9, 2024 (https://doi.org/10.1109/TIM.2024.3449981).

[9] A. Turyagyenda *et al.*, "IoT and Machine Learning Techniques for Precision Beekeeping: A Review", *AI*, vol. 6, art. no. 26, 2025 (https://doi.org/10.3390/ai6020026).

[10] K. Jeong *et al.*, "IoT and AI Systems for Enhancing Bee Colony Strength in Precision Beekeeping: A Survey and Future Research Directions", *IEEE Internet of Things Journal*, vol. 12, pp. 362–389, 2025 (https://doi.org/10.1109/JIOT.2024.3461775).

[11] B.B. Roy, S. Das, and U.K. Mondal, "TinyML-driven Sensor Nodes for Energy-efficient Acoustic Event Detection in Pervasive Acoustic WSNs", *Journal of Telecommunications and Information Technology*, no. 2, pp. 69–77, 2025 (https://doi.org/10.26636/jtit.2025.2.2084).

[12] U. Ghosh and U.K. Mondal, "Pilot Agent-driven Wireless Acoustic Sensor Network for Uninterrupted Data Transmission", *Journal of Telecommunications and Information Technology*, no. 4, pp. 53–60, 2023 (https://doi.org/10.26636/jtit.2023.4.1322).

[13] P. Rajchowski, "Examination of 5G NR, LTE, and NB-IoT Radio Interfaces and Their Vulnerabilities to Interference", *Journal of Telecommunications and Information Technology*, no. 4, pp. 93–100, 2024 (https://doi.org/10.26636/jtit.2024.4.1960).

[14] A. Zgank, "IoT-based Bee Swarm Activity Acoustic Classification Using Deep Neural Networks", *Sensors*, vol. 21, art. no. 676, 2021 (https://doi.org/10.3390/s21030676).

[15] M. Abdollahi, P. Giovenazzo, and T. Falk, "Automated Beehive Acoustics Monitoring: A Comprehensive Review of the Literature and Recommendations for Future Work", *Applied Sciences*, vol. 12, art. no. 3920, 2022 (https://doi.org/10.3390/app12083920).

[16] D. Kanelis *et al.*, "Decoding the Behavior of a Queenless Colony Using Sound Signals", *Biology*, vol. 12, art. no. 1392, 2023 (https://doi.org/10.3390/biology12111392).

[17] A. Farina, "Discovering Ecoacoustic Codes in Beehives: First Evidence and Perspectives", *Biosystems*, vol. 234, art. no. 105041, 2023 (https://doi.org/10.1016/j.biosystems.2023.105041).

[18] J.W. Cooley and J.W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series", *Mathematics of Computation*, vol. 19, pp. 297–301, 1965 (https://doi.org/10.2307/2003354).

[19] A.V. Oppenheim and R.W. Schafer, *Discrete-Time Signal Processing*, 2nd ed., Prentice Hall, 870 p., 1999, (ISBN 9780137549207).

[20] S.W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*, California Technical Publishing, 1997, (ISBN 9780966017632).

[21] Knowles Electronics (Syntiant), MEMS Sisonic Digital Multimode Microphone SPH0645LM4H-B, datasheet specification.

Sebastian Górecki, Ph.D. Candidate
Institute of Electronics
https://orcid.org/0000-0001-5700-4000
E-mail: research@sgorecki.eu
Lodz University of Technology, Łódź, Poland
https://p.lodz.pl