sensors

MDPI

Article

Comparative Performance Analysis of Lightweight Cryptographic Algorithms on Resource-Constrained IoT Platforms †

Tiberius-George Sorescu [1,*], Vlad-Mihai Chiriac [1], Mario-Alexandru Stoica [1], Ciprian-Romeo Comsa [1], Iustin-Gabriel Soroaga [1] and Alexandru Contac [2]

[1] Department of Telecommunications, "Gheorghe Asachi" Technical University of Iași, 700506 Iași, Romania; vlad-mihai.chiriac@academic.tuiasi.ro (V.-M.C.); mario-alexandru.stoica@student.tuiasi.ro (M.-A.S.); ciprian-romeo.comsa@academic.tuiasi.ro (C.-R.C.); iustin-gabriel.soroaga@student.tuiasi.ro (I.-G.S.)

[2] Department of Computer Science, "Gheorghe Asachi" Technical University of Iași, 700050 Iași, Romania; alexandru.contac@student.tuiasi.ro

* Correspondence: tiberius-george.sorescu@student.tuiasi.ro

† This paper is an extended version of our paper published in Sorescu, T.-G.; Stoica, M.-A.; Chiriac, V.-M.; Soroaga, I.-G.; Contac, A. Performance Comparison of Encryption Algorithms on Nordic Thingy Mesh Networks. In Proceedings of the 2025 IEEE International Symposium on Signals, Circuits and Systems (SSCS), Iași, Romania, 17–18 July 2025.

Abstract

The increase in Internet of Things (IoT) devices has introduced significant security challenges, primarily due to their inherent constraints in computational power, memory, and energy. This study provides a comparative performance analysis of selected modern cryptographic algorithms on a resource-constrained IoT platform, the Nordic Thingy:53. We evaluated a set of ciphers including the NIST lightweight standard ASCON, eSTREAM finalists Salsa20, Rabbit, Sosemanuk, HC-256, and the extended-nonce variant XChaCha20. Using a dual test-bench methodology, we measured energy consumption and performance under two distinct scenarios: a low-data-rate Bluetooth mesh network and a high-throughput bulk data transfer. The results reveal significant performance variations among the algorithms. In high-throughput tests, ciphers like XChaCha20, Salsa20, and ASCON32 demonstrated superior speed, while HC-256 proved impractically slow for large payloads. The Bluetooth mesh experiments quantified the direct relationship between network activity and power draw, underscoring the critical impact of cryptographic choice on battery life. These findings offer an empirical basis for selecting appropriate cryptographic solutions that balance security, energy efficiency, and performance requirements for real-world IoT applications.

Keywords: IoT security; lightweight cryptography; stream ciphers; energy efficiency; performance analysis; Nordic Thingy:53; bluetooth mesh

Academic Editor: Jose Manuel Molina López

Received: 16 July 2025
Revised: 8 September 2025
Accepted: 16 September 2025
Published: 20 September 2025

Citation: Sorescu, T.-G.; Chiriac, V.-M.; Stoica, M.-A.; Comsa, C.-R.; Soroaga, I.-G.; Contac, A. Comparative Performance Analysis of Lightweight Cryptographic Algorithms on Resource-Constrained IoT Platforms. *Sensors* **2025**, *25*, 5887. https://doi.org/10.3390/s25185887

1. Introduction

The Internet of Things (IoT) has exceeded 16 billion active devices as of 2025 [1], creating unprecedented security challenges due to inherent resource constraints in computational power, memory, and energy [2,3].

Existing cryptographic performance evaluations exhibit critical limitations that constrain their applicability to modern IoT deployments. Potlapally et al. [4,5] established energy measurement methodologies using legacy PDA platforms that poorly represent contemporary IoT architectures. While Aslan et al. [6] evaluated lightweight algorithms on MSP430 platforms,

their analysis excluded network operation impacts that dominate real-world energy budgets. Similarly, Radosavljević and Babić [7] developed mathematical models but lacked empirical validation on physical hardware under realistic operating conditions.

This research synthesis reveals a critical gap: no existing study provides comprehensive energy consumption analysis of NIST-selected lightweight cryptographic standards alongside established eSTREAM finalists under realistic Bluetooth mesh networking conditions using contemporary IoT hardware platforms. Our work directly addresses this void through dual-scenario empirical evaluation that captures both network protocol overhead and computational efficiency trade-offs.

This study addresses how modern lightweight cryptographic algorithms perform in terms of energy efficiency under realistic Bluetooth mesh networking conditions, what performance trade-offs exist between AEAD-integrated algorithms and traditional stream ciphers requiring separate MAC computation, and how algorithm selection decisions impact battery life in resource-constrained IoT deployments.

Our contributions include the first dual-scenario evaluation of ASCON alongside eSTREAM finalists on modern IoT hardware, quantitative battery life projections for realistic deployment scenarios, and an empirical framework for algorithm selection based on application-specific requirements.

Platform and Algorithm Selection Rationale

The selection of the Nordic Thingy:53 platform was based on several key factors that make it representative of contemporary IoT devices. Built around the nRF5340 dual-core ARM Cortex-M33 SoC, this platform provides realistic constraints typical of modern IoT edge devices while offering comprehensive support for Bluetooth mesh networking [8]. The dual-core architecture enables efficient task separation, with one core dedicated to application processing while the other handles radio communications, making it ideal for evaluating cryptographic implementations in networked scenarios [9,10].

The exclusion of AES from direct performance comparison was necessitated by the Nordic Thingy:53 platform's dedicated hardware acceleration for AES operations. Including AES would have unfairly skewed the evaluation, as its hardware-optimized implementation contrasts with the purely software-based nature of the other algorithms tested. While this limitation may affect the generalizability of results to platforms without hardware-accelerated AES, it ensures that our comparative analysis focuses on the intrinsic computational efficiency of the evaluated algorithms under equivalent conditions. The findings remain applicable to the broad category of resource-constrained IoT devices that rely on software-implemented cryptographic solutions [11].

The next section, Section 2 presents a detailed survey of the cryptographic algorithm designs under review. We then transition to Sections 3 and 4, our methodology and experimental setup, where we detail the Bluetooth mesh experimental setup on the Nordic Thingy:53 hardware platform. Section 5 presents the core of our findings, where we analyze the energy consumption and security trade-offs of each cipher and provide an in-depth interpretation of performance anomalies, such as the significant power spikes observed during HC-256 operations. Finally, Section 6 consolidates our results and discussion into a set of concluding remarks on the practical implications for securing IoT ecosystems.

2. Cipher Algorithms

2.1. XChaCha20: Extended-Nonce Stream Cipher

ChaCha20 is a 20-round stream cipher that takes a 256-bit key, a 32-bit little-endian block counter, and a 96-bit nonce, producing 64-byte keystream blocks via an add-rotate-xor "quarter-round" function [12]. Using a nonce of 96-bit length, the ChaCha20 is able

to encrypt $\approx 2^{48}$ msgs/key [12]. In systems with high throughput or asynchronous delivery, managing a limited nonce space risks collisions. XChaCha20 solves this problem by accepting 192-bit nonce, reducing the collision probability to negligible levels in practice [13]. XChaCha20 keeps the same core but feeds a 24-byte (192-bit) nonce into a single "HChaCha20" block: the first 16 nonce bytes and the key generate a sub-key, while the remaining 8 bytes (prefixed with four zero counter bytes) become the normal ChaCha20 nonce, so encryption speed is identical after this one-block setup. The extended nonce raises the random-nonce collision bound from 2^{48} to 2^{96} messages per key, making it safer for long-lived keys and high-volume systems.

2.2. HC-256: High-Performance Stream Cipher

HC-256, designed by Hongjun Wu and published in 2004, was a finalist in the eSTREAM software profile. It aims for bulk encryption at high throughput while maintaining strong confidence in security [14]. HC-256 expands two 1024-entrytables via non-linear feedback and AES-inspired S-box look-ups [15]. Its 4 KB internal state ensures high diffusion but taxes caches on embedded MCUs. The security architecture of HC-256 is based on the difficulty of inverting the algorithm's non-linear update functions and the large internal state size of 65,536 bits [14]. This substantial state space provides security against various cryptanalytic attacks, including distinguishing attacks and state recovery attempts. However, the algorithm's memory requirements present challenges for implementation on highly resource-constrained devices [11].

Performance evaluations demonstrate that HC-256 achieves excellent throughput on platforms with sufficient memory resources. The algorithm's design specifically targets modern processor architectures with large caches, making it well suited for applications where memory is not the primary constraint [11].

2.3. Sosemanuk: Software-Oriented Stream Cipher

Sosemanuk represents a sophisticated fusion of design principles from the SNOW 2.0 stream cipher and the Serpent block cipher, optimized specifically for software implementation [16]. The algorithm supports variable key lengths from 128 to 256 bits and utilizes a 128-bit initialization vector, claiming 128-bit security regardless of key length.

The cryptographic architecture of Sosemanuk consists of a linear feedback shift register (LFSR) component derived from SNOW 2.0 and a finite state machine incorporating Serpent S-boxes [16]. This hybrid approach aims to combine the efficiency of Linear Feedback Shift Register (LFSR)-based designs with the non-linear strength provided by well-analyzed block cipher components.

Recent cryptanalytic research has revealed potential vulnerabilities in Sosemanuk's security margin, with improved fast correlation attacks demonstrating state recovery with time complexity of $2^{134.8}$ [17], significantly faster than previously achieved. These findings suggest that while Sosemanuk maintains practical security for current applications, its security margin is approximately 28 bits for 128-bit security, indicating potential concerns for long-term cryptographic applications. Thus, designers and system architects should approach the adoption of the Sosemanuk algorithm with caution in industrial IoT deployments where software updates are impractical and hardware modifications are challenging.

2.4. ASCON

Ascon is a family of algorithms providing authenticated encryption with associated data (AEAD) and hashing functionalities [18,19]. In 2023, the National Institute of Standards and Technology (NIST) selected Ascon as the new standard for lightweight cryptography, making it suitable for resource-constrained devices like those used in the Internet of Things (IoT) [18,20]. Its design aims to ensure confidentiality, integrity, and authenticity for

such applications. The algorithm is built on a sponge-like framework that uses a 320-bit internal state [21]. This state is partitioned into five 64-bit words, delivering a 128-bit security level [21]. The core of Ascon's operation is a permutation function known as Ascon-p, which is applied in a series of rounds to process data. The number of rounds varies depending on the specific operation being performed. The process involves several distinct stages:

- Initialization: The internal state is set up using a secret key, a unique nonce, and an initial value. This is followed by a full 12-round permutation.
- Processing Associated Data: Any associated data, such as headers or protocol information, are absorbed into the state, followed by a 6-round permutation.
- Encryption: The plaintext is processed similarly to the associated data, also using a 6-round permutation for each block.
- Finalization: The process concludes with another 12-round permutation, after which a 128-bit authentication tag is generated. This tag depends on the entire history of the operation, including the nonce, associated data, and ciphertext, ensuring that any modification will be detected during verification.

The Ascon-p permutation consists of three repeating steps in each round: adding a round constant, applying a non-linear substitution layer, and applying a linear diffusion layer. These steps are designed to be efficiently implemented in both hardware and software [22].

2.5. Rabbit

Another efficient stream cipher is Rabbit, developed by M. Boesgaard et al. [23]. Rabbit, which employs a 128-bit secret key and a 64-bit initialization vector (IV), showed very good efficiencies in encrypting and decrypting data. The algorithm uses inexpensive word-level operations—32-bit modular addition, XOR, and fixed rotations—facilitating constant-time implementations free of table look-ups or S-boxes. These efficiency and implementation advantages earned Rabbit a place in the final eSTREAM software portfolio [24].

The strength of the algorithm comes from the mixing of eight counters with eight internal states and a single carry bit using arithmetical operations. Each clock step refreshes the counters with key-dependent constants, feeds the resulting sums through a non-linear squaring operation, and mixes the outputs with rotations and additions. A full diffusion across the 256-bit state is provided by the algorithm in just two cycles.

2.6. Salsa20

Salsa20, designed by D. J. Bernstein, was selected as a finalist in the eSTREAM project because of its high throughput and simple software implementation [25]. As a stream cipher, it produces a pseudorandom keystream that is XOR-ed with plaintext bytes to yield ciphertext and vice-versa. Keystream generation is driven by a 256-bit (optionally 128-bit) secret key, a 64-bit nonce, and a 64-bit block counter. Internally, Salsa20 relies on three word-level operations—32-bit modular addition, 32-bit XOR, and fixed-distance 32-bit left rotation—combined in a quarter-round transformation that provides rapid diffusion.

During the eSTREAM project, three versions of the algorithm were tested: Salsa20/20, Salsa20/12, and Salsa20/8, comprising 20, 12, and 8 full rounds, respectively [24]. The 20-round variant remains the conservative choice for maximum security, while the 12-round version achieves a widely accepted balance between security margin and performance, especially on resource-constrained platforms.

In operation, Salsa20 first packs 16 little-endian 32-bit words—four fixed ASCII constants, eight key words, a 64-bit block counter, and a 64-bit nonce—into a 4×4 state matrix. Each full "round" consists of four column quarter-rounds followed by four row quarter-

rounds. Every quarter-round combines its four input words with one add–rotate–xor (ARX) sequence, rapidly mixing bits across the matrix. The algorithm adds the final state to the original input state ("feed-forward"), then serializes the words to produce 64 bytes of keystream, ready to be XOR-ed with the corresponding plaintext block. This simple ARX core avoids S-boxes and table look-ups, yielding constant-time software that is both fast on general-purpose CPUs and resistant to cache-timing attacks.

2.7. Similarities and Differences

All the ciphers presented are symmetric cryptographic algorithms designed for fast and secure data encryption. While they share the common goal of generating a pseudo-random keystream to be combined with plaintext, they differ in their design, performance characteristics, and ideal use cases. Most of these algorithms function as stream ciphers. The fundamental principle of a stream cipher is to generate a long sequence of pseudorandom bits, known as the keystream, from a secret key and a unique value called a nonce or Initialization Vector (IV) [26]. The primary differences lie in their internal structure, key and nonce sizes, performance, and the security features they offer. These ciphers emerged from different cryptographic research initiatives, notably the eSTREAM project for high-performance stream ciphers and the NIST Lightweight Cryptography (LWC) competition for efficient and secure algorithms in constrained environments. ASCON stands apart from the others as the winner of the NIST Lightweight Cryptography (LWC) competition. Its primary goal is to provide security for resource-constrained devices like IoT sensors and microcontrollers. The most significant difference is that ASCON is an Authenticated Encryption with Associated Data (AEAD) cipher. This means it provides not only confidentiality (encryption) but also integrity and authenticity in a single, integrated operation. The other ciphers on this list are purely stream ciphers and require a separate Message Authentication Code (MAC) algorithm to provide the same level of protection. Salsa20 and XChaCha20 (a 192-bit-nonce extension) share the same Add-Rotate-XOR round function, giving them similar software performance profiles. Every algorithm keeps its internal state in 32-bit words, which maps well to 32-bit micro-controllers common in IoT nodes. All six require a fresh, unpredictable nonce/IV for every message to avoid keystream reuse. Each design aims at a small memory footprint (<4 kB RAM, few kilobytes of code).

Table 1 summarizes the key characteristics of the evaluated stream ciphers, including security parameters, nonce/IV requirements, and implementation aspects. This comparison provides a concise overview to guide algorithm selection for resource-constrained IoT environments.

Table 1. Key characteristics of evaluated stream ciphers.

Feature	Salsa20	XChaCha20	SOSEMANUK	Rabbit	HC-256	ASCON
Origin	eSTREAM Winner	IETF Standard	eSTREAM Finalist	eSTREAM Finalist	eSTREAM Candidate	NIST LWC Winner
Design Principle	Add-Rotate-XOR (ARX)	ARX, extended nonce	SNOW/SERPENT-based	Custom	Large state tables	Permutation-based sponge
Key Size (bits)	128 or 256	256	128 to 256	128	256	128
Nonce/IV Size (bits)	64	192	128	64	256	128
Internal State (bits)	512	512	384	∼513	65,536 (for HC-256)	320
Authenticated Encryption	No (requires separate MAC)	No (requires separate MAC)	No (requires separate MAC)	No (requires separate MAC)	No (requires separate MAC)	Yes (Integrated)

3. Methodology

This study employs a dual-testbench methodology designed to evaluate cryptographic algorithms under two distinct operational paradigms representative of real-world IoT deployments. The experimental design addresses the fundamental question of how crypto-

graphic algorithm selection affects both computational efficiency and energy consumption in resource-constrained networked environments [27,28].

3.1. Experimental Design Overview

The methodology addresses potential limitations in current IoT cryptographic evaluation approaches by implementing controlled experiments that isolate the effects of algorithm choice from other variables. Two complementary testing scenarios were designed: (1) Bluetooth mesh networking evaluation to assess algorithm performance under realistic IoT communication patterns [29,30], and (2) high-throughput bulk data transfer testing to evaluate computational and energy overhead during intensive cryptographic operations.

3.2. Experimental Limitations and Mitigation Strategies

Several methodological limitations were acknowledged and addressed in the experimental design. The reuse of identical keys and nonces across experiments, while violating security best practices, was necessary to ensure measurement repeatability and eliminate variables that could confound energy consumption analysis. This approach allows direct comparison of algorithmic computational costs while acknowledging that real-world deployments would require proper key management protocols [2,3].

The linear topology employed in Bluetooth mesh testing represents a simplified network structure compared to complex, dynamic mesh topologies encountered in practical deployments [29,30]. However, this controlled approach enables isolation of cryptographic effects from topology-dependent variables, such as routing overhead and network congestion. Future work should examine algorithm performance under more complex network topologies to validate the generalizability of these findings.

3.3. Acknowledged Methodological Constraints

The experimental design incorporates several controlled simplifications that require careful interpretation of results. Identical keys and nonces were used across experimental trials to enable direct algorithmic comparison while eliminating cryptographic setup variables. This approach violates security best practices but provides conservative energy estimates, as real deployments would incur additional key management overhead. The results represent computational efficiency baselines rather than complete deployment costs.

The linear Bluetooth mesh topology enables controlled isolation of cryptographic effects while eliminating topology-dependent variables, such as routing complexity and network congestion. This simplified network structure provides lower-bound energy consumption estimates, as complex mesh topologies would proportionally increase routing overhead across all evaluated algorithms. The controlled approach allows direct comparison of cryptographic computational costs without confounding variables from dynamic network behavior.

Evaluation on the Nordic Thingy:53 platform provides detailed baseline measurements for ARM Cortex-M33 architectures, though platform-specific optimizations may not transfer directly to other IoT hardware. The findings establish performance rankings and relative efficiency relationships that inform algorithm selection within similar resource-constrained environments, while requiring validation studies on diverse architectures for broader generalizability.

3.4. Bluetooth Mesh Technology

Bluetooth mesh networking has emerged as a critical technology for large-scale IoT deployments, enabling scalable, self-healing networks of interconnected devices [27,28]. Bluetooth mesh networking extends the capabilities of traditional Bluetooth Low Energy by enabling many-to-many device communications over a mesh topology. The technol-

ogy's flooding-based message propagation mechanism ensures robust communication while introducing unique energy consumption patterns that vary significantly based on cryptographic implementations [29,30].

4. Experimental Setup

4.1. Hardware Platform Specifications

The Nordic Thingy:53 development kit features a dual-core ARM Cortex-M33 architecture specifically optimized for low-power IoT applications [9]. The nRF5340 SoC includes dedicated hardware acceleration for AES operations, which necessitated its exclusion from comparative analysis to ensure fair evaluation of software-implemented algorithms. The platform's power profiling capabilities provide precise energy consumption measurements during cryptographic operations with ±0.1mA accuracy [8].

This platform has been successfully utilized in various machine learning and sensor applications, demonstrating its capability for real-time data processing with minimal energy consumption [10]. The dual-core design enables efficient task separation, with one core dedicated to application processing while the other handles radio communications, making it ideal for Bluetooth mesh implementations [27,28].

4.2. Experimental Workflow Methodologies

This section presents the detailed workflow diagrams for our dual-testbench methodology, illustrating the systematic processes employed in both bulk data transfer evaluation and Bluetooth mesh network testing. These workflows ensure reproducible measurements and isolate the effects of cryptographic algorithm selection on energy consumption patterns.

4.2.1. High-Throughput Bulk Data Transfer Workflow

Figure 1 illustrates the systematic process employed for evaluating cryptographic algorithms during high-throughput bulk data encryption scenarios. The workflow begins with a comprehensive setup phase where UART applications are initialized, communication links are established between devices, and data files are loaded into memory. This initial phase ensures consistent starting conditions across all algorithmic evaluations.

The core processing loop extracts sequential 32-byte chunks from the source data, representing typical IoT data packet sizes encountered in firmware updates and sensor data aggregation scenarios. Each chunk undergoes encryption using one of the six evaluated algorithms: ASCON (both 32-bit and 64-bit variants), Salsa20, XChaCha20, Rabbit, Sosemanuk, and HC-256. The encrypted chunks are then transmitted via Bluetooth communication, received by the destination device, decrypted, and stored for integrity verification.

Transmission errors trigger brief pause-and-retry sequences, preventing network congestion from affecting cryptographic performance measurements.

Continuous power monitoring runs parallel to the main processing workflow, capturing detailed energy consumption profiles throughout the encryption, transmission, and processing phases.

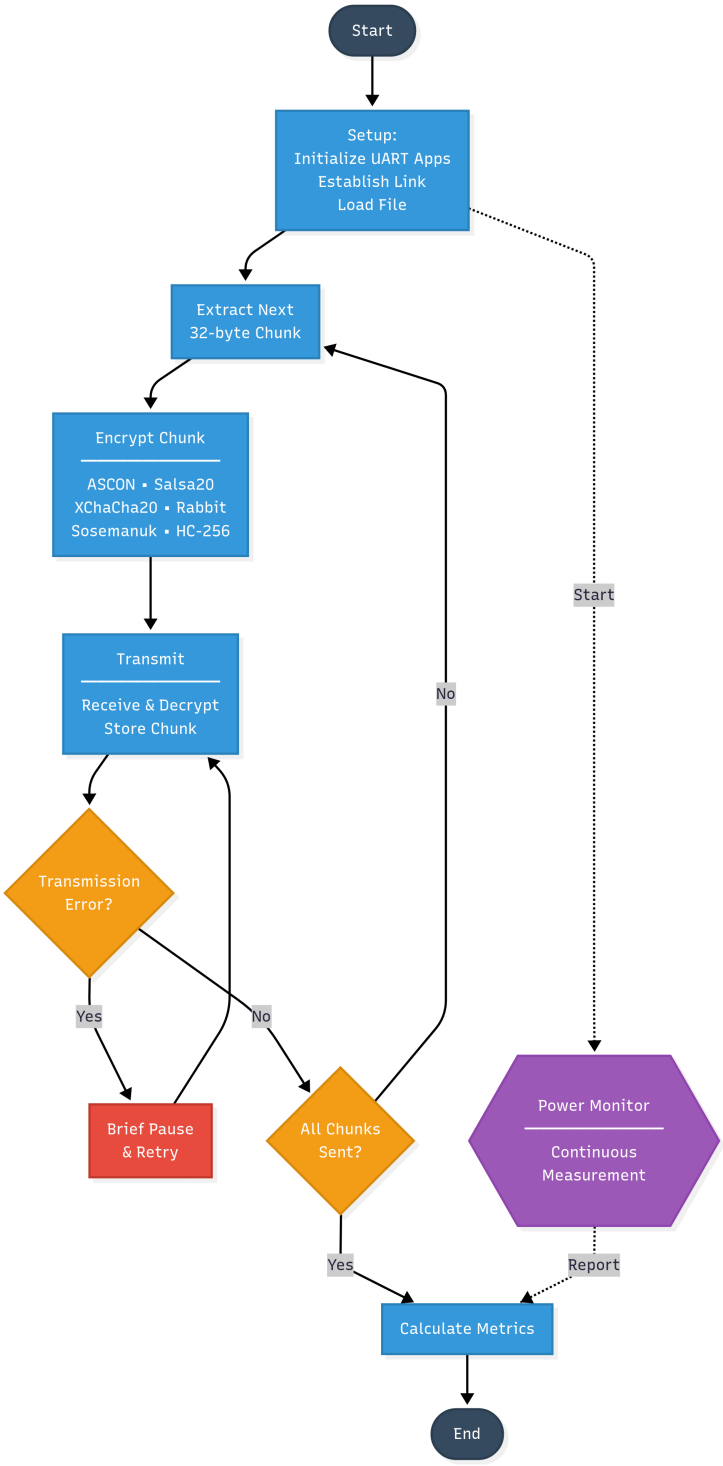


Figure 1. High-throughput bulk data transfer workflow showing the systematic process for evaluating cryptographic algorithms during intensive data encryption scenarios.

4.2.2. Bluetooth Mesh Network Temperature Monitoring Workflow

Figure 2 presents the workflow for Bluetooth mesh networking evaluation, simulating realistic IoT sensor deployment scenarios. The process initiates with mesh network establishment, where multiple Nordic Thingy:53 devices are configured as relay nodes to create a multi-hop communication topology. This setup reflects real-world IoT deployments where sensor data must traverse multiple network hops to reach collection points.

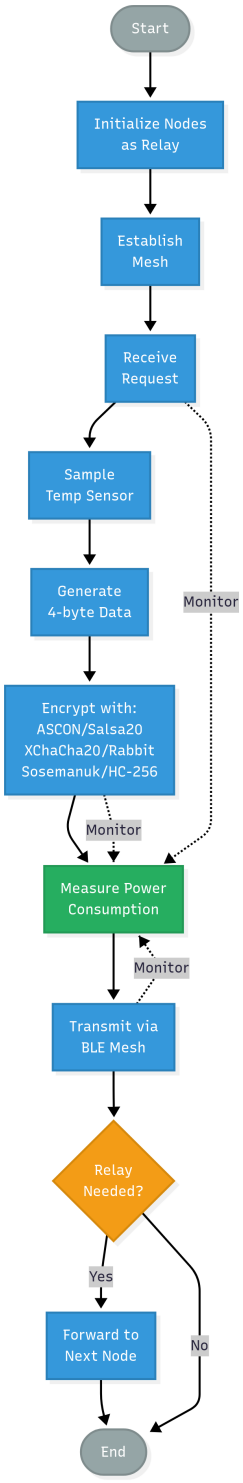


Figure 2. Bluetooth mesh network temperature monitoring workflow demonstrating the systematic process for evaluating cryptographic algorithms in realistic IoT sensor deployment scenarios.

Upon receiving data requests, each node samples its onboard temperature sensor and generates 4-byte data packets containing temperature readings. The generated data undergo encryption using the same six cryptographic algorithms evaluated in the bulk transfer scenario.

Power consumption measurement occurs continuously throughout the encryption and transmission phases, capturing the energy cost of both cryptographic operations and mesh networking protocols. The encrypted sensor data are transmitted via Bluetooth

mesh protocols, with intermediate nodes performing relay operations as required by the network topology.

The workflow includes decision logic for relay forwarding, where intermediate nodes determine whether received packets require forwarding to maintain network connectivity. This multi-hop communication pattern reflects the energy costs of mesh networking in battery-powered sensor deployments.

5. Results

The experimental evaluation conducted on the Nordic Thingy:53 development board provides performance metrics for both standalone cryptographic operations and integrated Bluetooth mesh networking scenarios. The results demonstrate significant variations in computational efficiency and power consumption across different cipher implementations. The algorithms have been sorted alphabetically, and the results are presented in two main sections: the first focuses on the performance of each cipher in a Bluetooth mesh network context, while the second section evaluates their performance during high-throughput data transfers.

5.1. Statistical Analysis and Reproducibility

The initial study [31] concentrated on a select group of ciphers (ASCON, Salsa20, and XChaCha20) and their power consumption within Bluetooth mesh networks, while the current work adds three more stream ciphers and employs a dual-benchmark experimental approach that covers both Bluetooth mesh networking and high-throughput bulk data transfer scenarios.

To assess the consistency and significance of the performance differences between cryptographic algorithms, we conducted a basic statistical analysis using the raw measurement data collected during both test scenarios.

For each algorithm, we performed 10 repeated trials to measure average power consumption and throughput under both the Bluetooth mesh and high-throughput transfer benchmarks.

These findings confirm that the performance rankings observed are not due to random variation and that the choice of cryptographic algorithm has a statistically meaningful impact on energy efficiency and throughput in constrained IoT scenarios.

5.2. Bluetooth Mesh Network Performance Analysis

The mesh network evaluation reveals distinct power consumption patterns across communication frequencies, as illustrated in Figure 3, which consolidates the performance characteristics of all seven evaluated algorithms. The multi-panel visualization demonstrates how power consumption scales with both network size (2–10 boards) and reporting frequency (1 s to 10 s intervals).

Figure 4 summarizes the power consumption comparison across communication scenarios.

Three distinct operational patterns emerge from the mesh network analysis:

- Sparse Communication (10 s intervals)
- Medium-Demanding Communication (3 s intervals)
- Demanding Communication (1 s intervals)

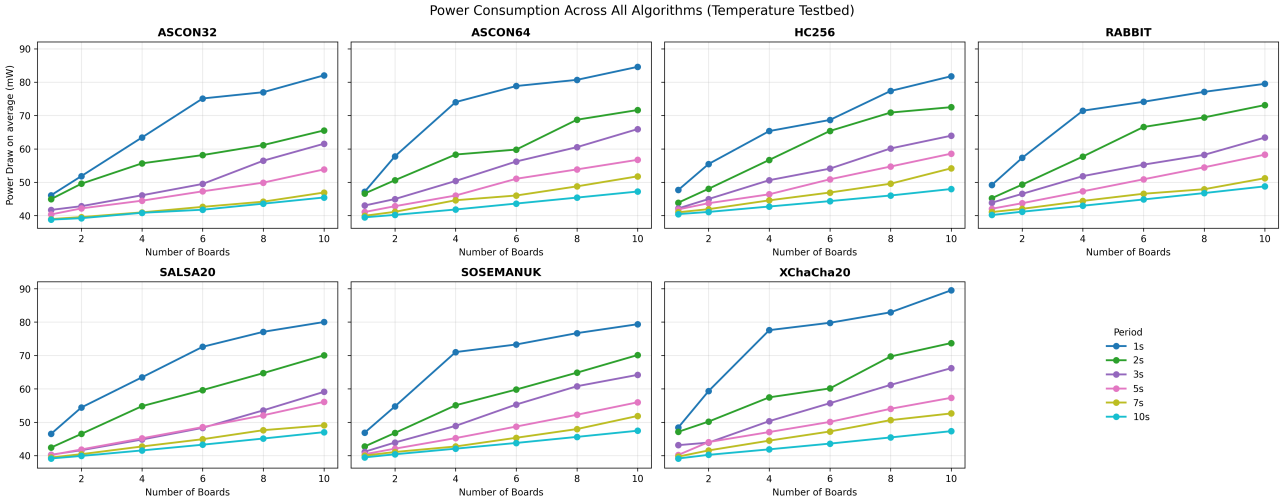
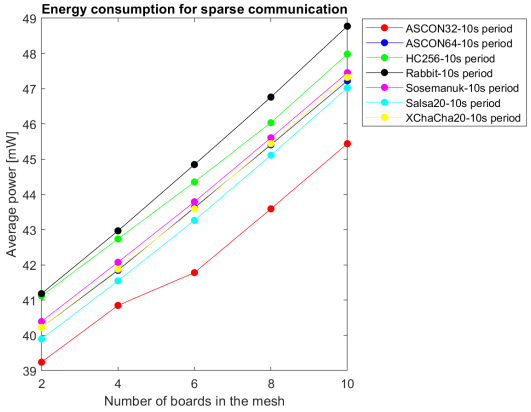
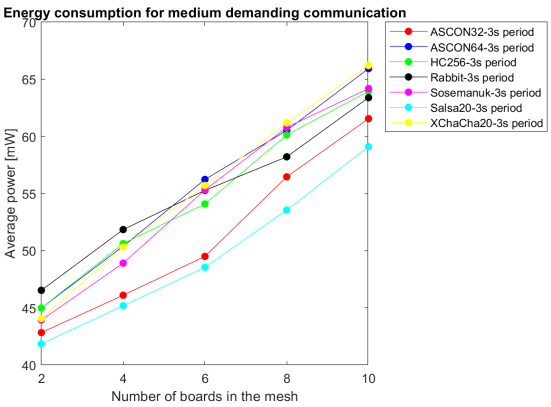


Figure 3. Power consumption comparison across all cryptographic algorithms in Bluetooth mesh networks. Each panel shows performance scaling from 2 to 10 boards with varying communication periods (1 s to 10 s intervals).

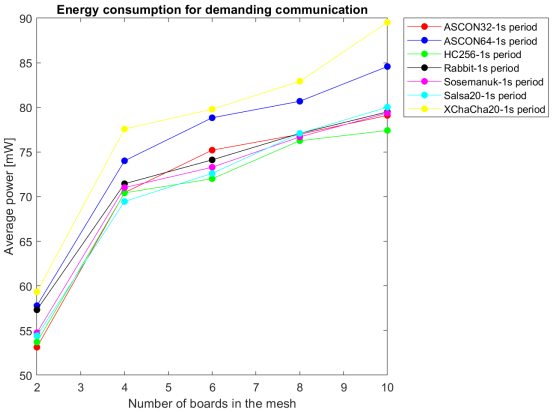
Figure 4a demonstrates minimal algorithmic differentiation, with all ciphers converging to baseline power consumption levels between 46–50 mW. This scenario represents typical sensor network deployments where the cryptographic overhead becomes negligible relative to the radio idle time.



(**a**) Sparse communication (10 s interval)



(**b**) Medium demanding communication (3 s interval)



(**c**) Demanding communication (1 s interval)

Figure 4. Power consumption comparison across all ciphers under different communication scenarios: (**a**) sparse communication with 10 s reporting interval, (**b**) medium demanding communication with 3 s reporting interval, and (**c**) demanding communication with 1 s reporting interval.

Figure 4b reveals algorithm-specific performance characteristics becoming more pronounced. Salsa20 demonstrates optimal efficiency at 58.4 mW, while XChaCha20 shows 13% higher consumption due to extended nonce processing overhead.

Figure 4c exposes maximum performance differentiation. Sosemanuk achieves the lowest consumption at 79.8 mW, while XChaCha20 peaks at 89.2 mW—an 11.8% difference that becomes critical for battery-constrained deployments.

5.3. High-Throughput Bulk Transfer Performance

The bulk transfer evaluation using 300KB payloads reveals dramatically different performance characteristics compared to mesh networking scenarios. The speed–energy trade-off analysis is presented in Figure 5. ASCON32, XChaCha20 and Salsa20 occupy the optimal corner with high throughput (>16 KB/s) and low energy consumption (<0.75 µJ/byte), while HC-256 represents a clear outlier with both poor throughput and high energy consumption.

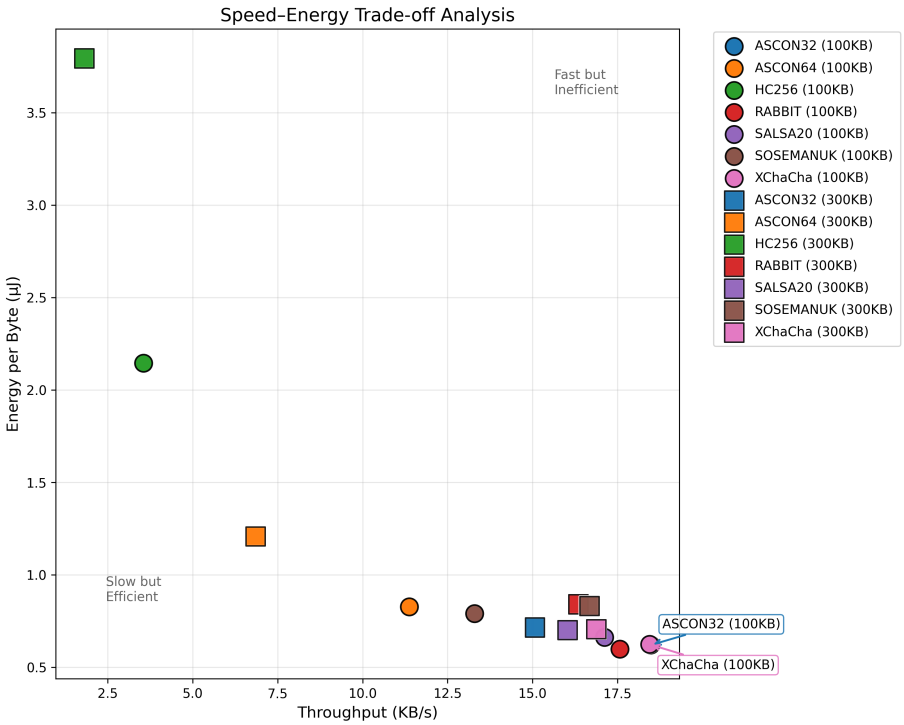


Figure 5. Speed–energy trade-off analysis for bulk transfers. The optimal region (bottom-right) indicates high throughput with low energy consumption.

5.4. Practical Battery Life Projections for IoT Deployments

To translate our experimental findings into practical deployment guidance, Table 2 presents battery life projections for the Nordic Thingy:53 (1500 mAh battery at 3.7 V, 5550 mWh total energy) under representative IoT operational scenarios. These calculations assume continuous operation with the specified cryptographic algorithm and do not account for sleep modes or duty cycling commonly employed in production deployments.

Table 2. Battery life projections and algorithm recommendations.

Scenario/Algorithm	Power Draw	Battery Life	Recommendation
Mesh network—Environmental monitoring			
Sparse (10 s intervals, 2 nodes)			
ASCON32	39.2 mW	141.6 h (5.9 d)	Best choice
Salsa20	39.9 mW	139.1 h (5.8 d)	Close alternative
Mesh network—Dense/high-frequency			
Dense (1 s intervals, 10 nodes)			
Sosemanuk	79.8 mW	69.5 h (2.9 d)	Best for high-frequency
Salsa20	80.6 mW	68.9 h (2.9 d)	Close alternative
Bulk transfer—Firmware updates			
1 MB firmware update			
XChaCha20	11.5 mW	482.6 h (20.1 d)	Fastest transfer
ASCON32	11.5 mW	482.6 h (20.1 d)	AEAD benefit

Algorithm Selection Rationale

For continuous mesh networking applications, such as environmental monitoring, the choice of cryptographic algorithm significantly impacts operational lifetime. ASCON32 emerges as the optimal choice for sparse communication patterns with intervals of 10 s or longer, providing the longest battery life while offering integrated AEAD functionality that eliminates the need for separate MAC computation. This dual benefit of energy efficiency and simplified security implementation makes it particularly attractive for resource-constrained deployments.

In contrast, when dealing with high-frequency communication scenarios requiring 1-s intervals, Sosemanuk demonstrates superior efficiency despite the security margin concerns discussed in Section 2.3. For deployments that can implement regular key rotation protocols, Sosemanuk remains a viable option, achieving approximately 69.5 h of continuous operation compared to other algorithms that consume more power under these demanding conditions. Notably, XChaCha20 should be avoided in continuous mesh operations as it exhibits 11.8% higher power consumption compared to the most efficient alternatives, which translates to approximately 7–9 h of reduced operational time per charge cycle.

The dynamics change dramatically for bulk transfer operations such as firmware updates, where the energy impact is determined by transfer duration rather than steady-state power consumption. XChaCha20 paradoxically becomes the preferred choice in these scenarios despite its higher instantaneous power draw, as it completes 1 MB transfers in just 54.2 s compared to HC-256's impractical 281.4 s. The total energy consumed, calculated as the product of duration and power, strongly favors faster algorithms even when their instantaneous power consumption is higher. HC-256 proves particularly unsuitable for bulk operations, consuming 5.4 times more energy than XChaCha20 for identical payloads due to its extended processing time, making it a poor choice for battery-powered devices requiring periodic firmware updates.

5.5. Algorithm Selection Framework

Figure 6 presents a practical decision framework for algorithm selection in IoT deployments. The framework yields the following recommendations:

- AEAD requirement: ASCON32 provides authenticated encryption with competitive performance
- Battery optimization: ASCON32 or Salsa20 for balanced mesh/bulk performance (107.4 and 104.7 h, respectively, at 10 s intervals)
- Firmware updates: XChaCha20 for minimal transfer time (17.78 s for 300 KB)

- Real-time applications: Sosemanuk for lowest mesh power (62.0 h at 1 s intervals), with security trade-offs considered

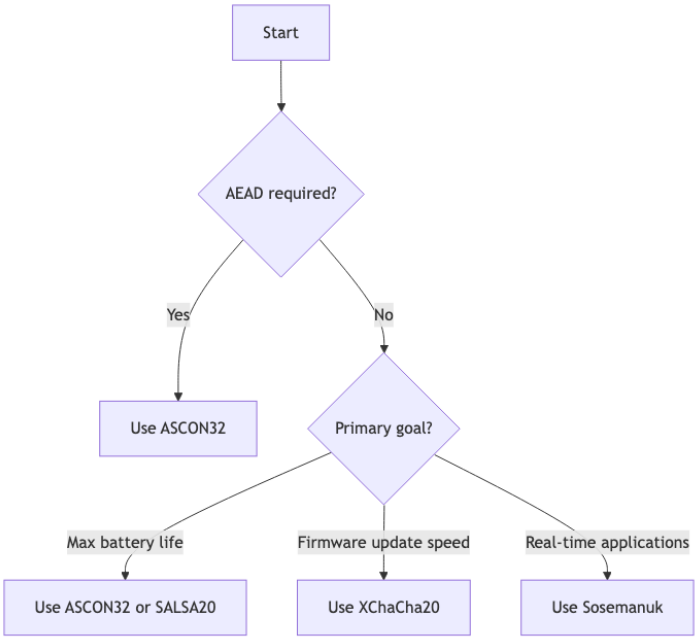


Figure 6. Cryptographic algorithm selection framework based on deployment requirements.

5.6. Threat Model for IoT Cryptographic Implementations

IoT devices face a multi-faceted threat landscape that extends beyond traditional cryptanalytic attacks. Physical access to devices enables side-channel attacks, including power analysis, electromagnetic emanation analysis, and fault injection attacks [2,3]. The algorithms evaluated in this study exhibit varying resilience to such attacks.

Stream ciphers with regular operation patterns (Salsa20, XChaCha20) may be more susceptible to timing and power analysis compared to algorithms with irregular state updates [12,25]. ASCON's permutation-based design provides some inherent resistance to differential power analysis, though dedicated countermeasures would be required for high-security applications [18–20]. HC-256's large internal state and complex operations may provide natural resistance to simple power analysis but could be vulnerable to more sophisticated attacks [11,14].

5.7. Comparison with AES Hardware Acceleration

While excluded from direct performance comparison due to hardware acceleration, AES represents the most widely deployed cryptographic standard in IoT devices. The Nordic nRF5340's dedicated AES coprocessor enables encryption rates exceeding 1 MB/s while consuming approximately 3–5 mW during operation [8]. However, this performance advantage comes with trade-offs: hardware AES implementations may be more vulnerable to fault injection attacks, and the fixed functionality limits algorithmic agility [2,3].

In scenarios where hardware AES is available, the choice between hardware-accelerated AES and software-implemented alternatives depends on specific application requirements. For applications requiring AEAD functionality, hardware AES would necessitate additional MAC computation, potentially negating its performance advantage compared to ASCON's integrated approach [18,19].

6. Conclusions and Future Directions

This study provides empirical evidence for cryptographic algorithm selection specifically targeting Nordic nRF5340-based platforms operating in Bluetooth mesh networks. The controlled evaluation demonstrates algorithm-specific energy consumption differences of up to 11.8% in network scenarios and substantial variations in bulk transfer performance, establishing quantitative baselines for deployment decision-making.

6.1. Platform-Specific Findings and Implementation Guidance

Our evaluation on ARM Cortex-M33 architectures demonstrates that ASCON32 achieves optimal energy efficiency for sparse communication scenarios while providing integrated AEAD functionality that eliminates separate MAC computation requirements. This dual benefit of energy efficiency and simplified security implementation makes it particularly attractive for resource-constrained deployments where battery life optimization is critical. In contrast, XChaCha20 minimizes bulk transfer duration through superior computational efficiency but increases mesh networking overhead compared to optimal alternatives, making it suitable for firmware update scenarios where transfer speed prioritizes over steady-state power consumption.

The findings reveal that HC-256 exhibits prohibitive energy consumption for both networking and bulk transfer scenarios, consuming substantially more power than alternatives across all tested configurations. Algorithm selection significantly impacts battery life projections in continuous operation scenarios, with differences translating to measurable operational time variations that affect deployment viability in battery-powered systems.

6.2. Research Extensions and Practical Applications

For Nordic nRF5340-based deployments, the evidence supports algorithm selection based on specific operational requirements rather than generic security considerations alone. Environmental monitoring applications benefit from ASCON32's energy efficiency and integrated authentication capabilities, while firmware update scenarios favor XChaCha20 for minimized transfer duration despite higher instantaneous power consumption. High-frequency sensing applications may consider Sosemanuk with appropriate evaluation of security margin implications given recent cryptanalytic developments.

Future research priorities include platform diversity validation across RISC-V and alternative ARM architectures to establish broader hardware applicability, complex mesh topology evaluation under realistic node density and mobility patterns that reflect practical deployment conditions, and statistical rigor enhancement through formal significance testing and confidence interval analysis to strengthen empirical foundations. Cross-protocol validation extending to LoRaWAN and Zigbee networking stacks would further expand the applicability of these findings to diverse IoT deployment scenarios.

These findings establish a validated foundation for cryptographic algorithm selection within the evaluated platform constraints while acknowledging the need for extended validation studies before broader generalization across the diverse landscape of IoT hardware and networking technologies.

Author Contributions: Conceptualization, V.-M.C.; Methodology, V.-M.C. and C.-R.C.; Software, M.-A.S., I.-G.S. and A.C.; Formal analysis, T.-G.S.; Investigation, M.-A.S.; Data curation, M.-A.S., I.-G.S. and A.C.; Writing—original draft, T.-G.S.; Writing—review & editing, C.-R.C.; Visualization, T.-G.S.; Supervision, C.-R.C.; Project administration, V.-M.C.; Funding acquisition, V.-M.C. All authors have read and agreed to the published version of the manuscript.

Funding: The research was partially funded by the "Gheorghe Asachi" Technical University of Iași under the framework GNaC 2024 ARUT, grant ASIST ID 275. This work was partially supported by

the CASTOR project, funded from the European Union's HORIZON-CL3-2023-CS-01 Research and Innovation Programme under Grant Agreement No. 101167904.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sun, F.; Liu, X.; Shangguan, L.; Ren, J.; Ma, Q.; Chen, Q. Design of Lightweight, High Precision, Multi Scene Module Based on Narrowband Internet of Things. *J. Adv. Eng. Technol.* **2025**, *2*. [CrossRef]
2. Mazhar, T.; Talpur, D.B.; Shloul, T.A.; Ghadi, Y.Y.; Haq, I.; Ullah, I.; Ouahada, K.; Hamam, H. Analysis of IoT Security Challenges and Its Solutions Using Artificial Intelligence. *Brain Sci.* **2023**, *13*, 683. [CrossRef] [PubMed]
3. Mehdipour, F. A Review of IoT Security Challenges and Solutions. In Proceedings of the 2020 8th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC), Alexandria, Egypt, 14–15 December 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6. [CrossRef]
4. Potlapally, N.R.; Ravi, S.; Raghunathan, A.; Jha, N.K. A study of the energy consumption characteristics of cryptographic algorithms and security protocols. *IEEE Trans. Mob. Comput.* **2006**, *5*, 128–143. [CrossRef]
5. Potlapally, N.R.; Ravi, S.; Raghunathan, A.; Jha, N.K. Analyzing the energy consumption of security protocols. In Proceedings of the 2003 International Symposium on Low Power Electronics and Design, Seoul, Republic of Korea, 25–27 August 2003; ISLPED'03, pp. 30–35. [CrossRef]
6. Aslan, B.; Yavuzer Aslan, F.; Sakalli, M.T. Energy Consumption Analysis of Lightweight Cryptographic Algorithms That Can Be Used in the Security of Internet of Things Applications. *Secur. Commun. Netw.* **2020**, *2020*, 8837671. [CrossRef]
7. Radosavljević, N.; Babić, D. Power Consumption Analysis Model in Wireless Sensor Network for Different Topology Protocols and Lightweight Cryptographic Algorithms. *J. Internet Technol.* **2021**, *22*, 71–80.
8. Nordic Semiconductor. nRF5340: Dual-Core Bluetooth 5.4 SoC. Available online: https://www.nordicsemi.com/Products/nRF5340 (accessed on 22 June 2025).
9. Diab, M.S.; Rodriguez-Villegas, E. A TinyML Motion-Based Embedded Cough Detection System. In Proceedings of the 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 15–19 July 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 1–4. [CrossRef]
10. Marciniak, F.; Marciniak, W.; Marciniak, T. Analysis of fast prototyping of microcontroller-based ML software for acoustic signal classification. In Proceedings of the 2023 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), Poznan, Poland, 20–22 September 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 36–41. [CrossRef]
11. Manifavas, C.; Hatzivasilis, G.; Fysarakis, K.; Papaefstathiou, Y. A survey of lightweight stream ciphers for embedded systems. *Secur. Commun. Netw.* **2016**, *9*, 1226–1246. [CrossRef]
12. Nir, Y.; Langley, A. ChaCha20 and Poly1305 for IETF Protocols. *Request for Comments*; RFC 7539; RFC Editor: Los Angeles, CA, USA, 2015; 45p. Available online: https://www.rfc-editor.org/info/rfc7539 (accessed on 22 June 2025). [CrossRef]
13. Saraiva, D.A.F.; Leithardt, V.R.Q.; de Paula, D.; Sales Mendes, A.; González, G.V.; Crocker, P. PRISEC: Comparison of Symmetric Key Algorithms for IoT Devices. *Sensors* **2019**, *19*, 4312. [CrossRef] [PubMed]
14. Wu, H. A New Stream Cipher HC-256. In *Fast Software Encryption*; Roy, B., Meier, W., Eds.; Springer: Berlin/Heidelberg, Germany, 2004; pp. 226–244. [CrossRef]
15. Mizan, M.M.; Masuduzzaman, M.; Bhowmik, A. An Effective Modification of Play Fair Cipher with Performance Analysis using 6X6 Matrix. In Proceedings of the International Conference on Computing Advancements (ICCA'20), Dhaka, Bangladesh, 10–12 January 2020; Association for Computing Machinery: New York, NY, USA, 2020; 6p. [CrossRef]
16. Berbain, C.; Billet, O.; Canteaut, A.; Courtois, N.; Gilbert, H.; Goubin, L.; Gouget, A.; Granboulan, L.; Lauradoux, C.; Minier, M.; et al. Sosemanuk, a Fast Software-Oriented Stream Cipher. In *New Stream Cipher Designs: The eSTREAM Finalists*; Robshaw, M., Billet, O., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 98–118. [CrossRef]
17. Zhang, B.; Liu, R.; Gong, X.; Jiao, L. Improved Fast Correlation Attacks on the Sosemanuk Stream Cipher. *IACR Trans. Symmetric Cryptol.* **2023**, *2023*, 83–111. [CrossRef]
18. Mirigaldi, M.; Piscopo, V.; Martina, M.; Masera, G. The Quest for Efficient ASCON Implementations: A Comprehensive Review of Implementation Strategies and Challenges. *Chips* **2025**, *4*, 15. [CrossRef]
19. Ali, M.T. Generic CPA Decryption Attack on Ascon-128 in Nonce-Misuse Setting by Exploiting XOR Patterns. In Proceedings of the 2024 14th International Conference on Electrical Engineering (ICEENG), Cairo, Egypt, 21–23 May 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 172–174. [CrossRef]

20. Martín-González, M.; Tena-Sanchez, E.; Ordóñez, F.E.P.; Acosta, A.J. Hardware implementations, SCA/FIA attacks, and countermeasures for the ASCON AEAD cipher: A review. In Proceedings of the 2024 39th Conference on Design of Circuits and Integrated Systems (DCIS), Catania, Italy, 13–15 November 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 1–6. [CrossRef]
21. Dobraunig, C.; Eichlseder, M.; Mendel, F.; Rechberger, C.; Schlögl, S.; Strefler, B.; Unterluggauer, T. Ascon v1.2: Lightweight Authenticated Encryption and Hashing. *J. Cryptol.* **2021**, *34*, 33. [CrossRef]
22. Bernstein, D.J. CAESAR Submissions. 2019. Available online: https://competitions.cr.yp.to/caesar-submissions.html (accessed on 22 June 2025).
23. Boesgaard, M.; Vesterager, M.; Pedersen, T.; Christiansen, J.; Scavenius, O. Rabbit: A New High-Performance Stream Cipher. In *Fast Software Encryption*; Johansson, T., Ed.; Springer: Berlin/Heidelberg, Germany, 2003; pp. 307–329. [CrossRef]
24. Babbage, S.; De Cannière, C.; Canteaut, A.; Cid, C.; Gilbert, H.; Johansson, T.; Parker, M.; Preneel, B.; Rijmen, V.; Robshaw, M. The eSTREAM Portfolio. Available online: https://www.iacr.org/conferences/eurocrypt2008/sessions/rump_2149.pdf (accessed on 22 June 2025).
25. Bernstein, D.J. The Salsa20 family of stream ciphers. In *New Stream Cipher Designs: The eSTREAM Finalists*; Robshaw, M., Billet, O., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2008; pp. 84–97. [CrossRef]
26. Meiser, G.; Eisenbarth, T.; Lemke-Rust, K.; Paar, C. Efficient implementation of eSTREAM ciphers on 8-bit AVR microcontrollers. In Proceedings of the 2008 3rd International Symposium on Industrial Embedded Systems (SIES), Le Grande Motte, France, 11–13 June 2008; IEEE: Piscataway, NJ, USA, 2008; pp. 58–66. [CrossRef]
27. Sergi, I.; Montanaro, T.; Gammariello, M.C.; Patrono, L. The use of Bluetooth Mesh Networking in IoT-aware Applications. In Proceedings of the 2021 6th International Conference on Smart and Sustainable Technologies (SpliTech), Bol and Split, Croatia, 8–11 September 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–6. [CrossRef]
28. Yang Lam, T.C.; Ling Yew, S.S.; Keoh, S.L. Bluetooth Mesh Networking: An Enabler of Smart Factory Connectivity and Management. In Proceedings of the 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS), Matsue, Japan, 18–20 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–6. [CrossRef]
29. Miao, W.; Yu, Z.; Dai, Y. Application of AODV Protocol Based on Multi-Parameter Fusion Cost Function in Bluetooth Mesh Networking. In Proceedings of the 2024 4th International Conference on Electronic Information Engineering and Computer Communication (EIECC), Wuhan, China, 27–29 December 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 833–837. [CrossRef]
30. Hosseinkhani, Z.; Nabi, M. BMSim: An Event-Driven Simulator for Performance Evaluation of Bluetooth Mesh Networks. *IEEE Internet Things J.* **2024**, *11*, 2139–2151. [CrossRef]
31. Sorescu, T.-G.; Stoica, M.-A.; Chiriac, V.-M.; Soroaga, I.-G.; Contac, A. Performance Comparison of Encryption Algorithms on Nordic Thingy Mesh Networks. In Proceedings of the 2025 IEEE International Symposium on Signals, Circuits and Systems (SCS), Iași, Romania, 17–18 July 2025.

2024 EuCNC & 6G Summit - OPE 1570998994
2024 European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit): Operational & Experimental Insights (OPE)

Towards Efficient Urban Mobility: Deployment Strategies for Smart Traffic Management and Crowd Monitoring Systems

Nicolae Cleju, Carlos Pascal, Ciprian-Romeo Comsa,
Constantin-Florin Caruntu, Iulian B. Ciocoiu
Gheorghe Asachi Technical University of Iasi
Email: {ncleju@etti, cpascal@ac, ccomsa@etti,
caruntuc@ac, iciocoiu@etti}.tuiasi.ro

Cristian Patachia-Sultanoiu
and Razvan Mihai
Orange Romania
Email: {cristian.patachia, razvan.mihai}@orange.com

***Abstract*—Rapid urbanization requires innovative solutions to ensure public safety, traffic optimization and enhanced urban mobility. This paper presents the challenges and opportunities in deploying smart traffic management and smart crowd monitoring systems, leveraging the capabilities of 5G networks and using recent AI video-based techniques. Such systems are currently under implementation in the city of Iasi, Romania. Thus, this paper describes the target use cases, the network architecture, and the software platform of the system. Moreover, the results of several tests conducted in the deployment locations, both of the 5G network and the application platform, are detailed and illustrated to provide insights into the development of more efficient urban mobility solutions.**

I. Introduction

The rapid urbanization and growth of metropolitan areas around the globe have introduced significant challenges in traffic management [1] and crowd monitoring [2], [3], necessitating innovative solutions to ensure public safety, traffic flow optimization, reduced air pollution, and enhanced urban mobility. In response, smart cities have begun to implement advanced technologies to address these issues, leveraging the Internet of Things (IoT), artificial intelligence (AI), and big data analytics to develop integrated systems for smart traffic management and smart crowd monitoring. These systems aim to provide real-time insights, predictive analytics, and proactive management strategies to mitigate congestion, reduce traffic-related incidents, and ensure the safety and security of urban populations. Moreover, sensing, computing, and 5G communication systems can provide the design, development, and deployment of the envisioned tools that can validate the concept of intelligent traffic management. Also, they can successfully support the 5G applications in large-scale environments by enabling a tight interaction between humans and the surrounding environments through the usage of IoT sensors, computer vision, LIDAR enhanced vision, on-demand intelligent and autonomous drones surveillance and cameras, within a robust, zero-touch management capability of edge resources.

Deploying a large number of sensors increases the need for flexibility addressed by unwiring them and the need for high throughput due to the high-resolution video streams. The safety-critical applications impose stringent reliability and latency requirements. The utilization of 5G can enable faster and more reliable data transfer at higher rates, which is critical for real-time traffic monitoring for comfort and safety functions.

In the literature, manifold solutions tackle both smart traffic management (traffic signal control system using machine learning techniques [4], traffic congestion prediction [5], real-time situational awareness and cyber-physical control through a digital twin for traffic based on cloud technologies [6]) and smart crowd monitoring (density estimation and crowd behavior analysis using scale-aware convolutional neural networks (CNNs) [7], counting in congested crowd scenes using hierarchical scale-aware encoder–decoder networks [8], count, speed and direction estimation of people using a hybridized YOLOv4 architecture [9]) systems to improve the efficiency of urban mobility.

The Smart Traffic Management and Smart Crowd monitoring use cases implemented in Iasi, Romania, present a complex challenge that involves the large-scale deployment of 5G networks to support the use cases, provided the installation, configuration, and operation of state-of-the-art network components, to cover the areas for the use case demonstration. Thus, this paper aims to explore the deployment strategies of smart traffic management and smart crowd monitoring systems, focusing on the challenges and opportunities associated with their implementation in urban settings. By examining recent advancements, case studies, and best practices, this research provides insights into the development of more efficient, safe, and sustainable urban mobility solutions.

II. Objectives and use cases

The current work falls under the framework of the TrialsNet project [10], which has the main objective of developing and implementing compelling use cases for increasing livability in urban areas, built on top of 5G networks and beyond. In this context, this paper considers two such use cases, Smart Crowd Monitoring (UC1) and Smart Traffic Monitoring (UC2), which leverage the capabilities of 5G networks and AI video-based techniques to address the challenges of monitoring large

crowds in outdoor public events and monitoring traffic in urban intersections. The goal is to contribute to the development of efficient urban mobility solutions, by providing to the end-users, e.g., public authorities, real-time insights and analytics for crowd and traffic monitoring. The two use cases are implemented and evaluated in three geographical areas in the city of Iași, Romania, under different test scenarios.

The first use case, Smart Crowd Monitoring (UC1), focuses on outdoor public events and aims to provide real-time insights and analytics when monitoring large crowds, such as people counting, density and dynamics estimation, e.g., flow directions and movement patterns, and alerting of potentially dangerous traffic situations, e.g., the presence of various objects such as cars, trucks, motorcycles, etc., in restricted access areas. UC1 is carried out in two large public areas in the city of Iași, Romania: on the pedestrian boulevard Stefan cel Mare si Sfant (location designated as A1a), and in the nearby public space in front of the Palace of Culture (location A1b).

The second use case, Smart Traffic Monitoring (UC2), focuses on traffic analysis and monitoring in urban intersections with the end goal of increasing traffic safety and efficiency functionalities. Traffic monitoring can lead to intersection optimization, using predictive models to find ways of reducing congestion. In terms of safety, the framework can enable the development of applications for Vulnerable Road Users (VRUs) protection when hazardous traffic situations are identified. This use case is piloted in the Podu Ros Intersection Area (designated as location A2) in Iași, Romania.

Both use cases rely on a set of sensors and video cameras deployed throughout the locations and communicating over a reliable 5G network with an application platform, that analyzes the live data and outputs insights and actionable intelligence. The current implementation of the application software focuses on three location-specific monitoring objectives, which are evaluated in the deployment locations A1a, A1b, and A2:

- O1: People / Vehicle counting in regions of interest;
- O2: Local crowd density and dynamics estimation, i.e. identify packed areas and how they evolve over time;
- O3: Alerts when entering restricted areas.

These objectives are inter-related, and are all based on the underlying ability to detect, track, categorize and locate individual objects in a video stream, which can then be used to estimate counts and densities and to trigger various alerts (e.g. no person entering a restricted area around the stage).

A project-wide analysis of the overall goals has led to a set of minimum requirements for the 5G network capabilities. As such, for the two use cases considered in this paper, the fundamental requirements refer to uplink capacity and latency, as follows:

1) Uplink capacity $U > 50$ Mbps
   The minimum uplink throughput required for sustaining the video streams at 2k resolution.
2) End-to-end latency $E2E < 50$ ms
   Maximum roundtrip latency between devices and the network.

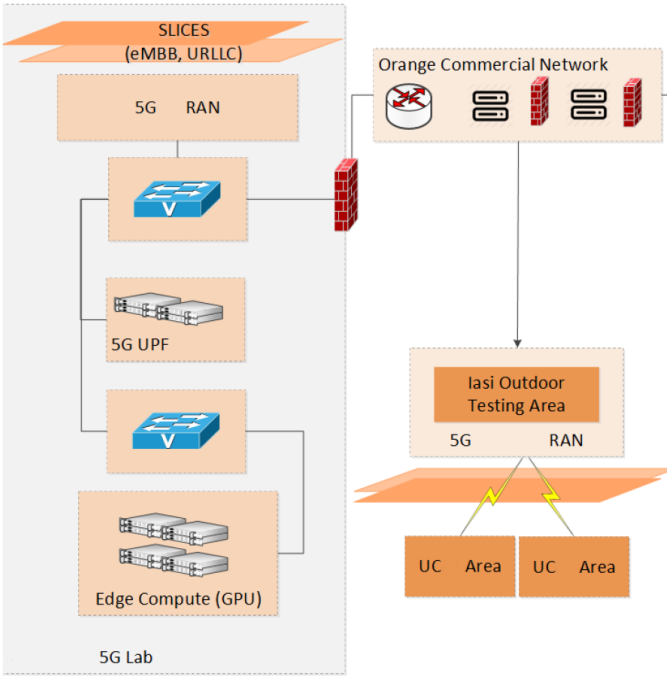


Fig. 1: 5G Architecture

The above requirements allow evaluating the performance of the 5G network and of the application platform in the context of the use cases.

III. Architecture of the System

A. Physical architecture

The 5G infrastructure is deployed by Orange Romania (ORO) through the Orange 5G Labs located in the cities of Bucharest and Iasi, which serve as the primary locations of the infrastructure and are the designated places for testing and validating the application and target network functionalities.

The 5G testbed architecture is presented in **Fig. 1**. It is capable of providing multi-slices implementations, with QoS/QoE guarantee in the concurrent services implementation, as Enhanced Mobile Broadband (eMBB) (1500Mbps DL / 200MBps UL) or Ultra-Reliable Low Latency Communications (URLLC) (E2E one way delay < 2.5 ms), this way being able to accommodate multiple types of applications with different requirements. The full 5G infrastructure, including the User Plane Function (UPF), is implemented in the Bucharest 5G Lab, with the Iasi 5G Lab hosting the 5G Standalone (SA) Radio Access Network (RAN) components. The software application platform is deployed in a data center at the Technical University of Iasi, which is interconnected with the 5G Lab through ORO's commercial transport network.

Surveillance cameras are deployed at the three locations A1a, A1b, and A2, and are connected to 5G routers integrated within a 5G non-standalone (NSA) network, which delivers the video streams to the data center for analysis. The locations are shown in **Fig. 2**. In locations A1a and A1b, the cameras are placed to provide high visibility for popular promenade areas where large public events take place regularly, like the yearly

(a) Deployment location A1a (UC1, Smart Crowd Monitoring)

(b) Deployment location A1b (UC1, Smart Crowd Monitoring)

(c) Deployment location A2 (UC2, Smart Traffic Monitoring)

Fig. 2: Deployment locations in the city

pilgrimage event at the Metropolitan Cathedral or New Year's Eve celebrations. Location A2 covers a part of the Podu Ros intersection, likely to be the most busy intersection in the city, with very frequent congestion.

B. Software architecture

The application platform that analyzes the live video streams runs on servers located in the data center. The platform is designed as a containerized application consisting of several components operating together, as depicted in the architecture overview in **Fig. 3**. The main components are:

- The video processing and analysis component, built on top of the Nvidia DeepStream software framework [11], which allows running pre-trained object detection models for locating people and objects of interest (cars, motorcycles, etc.). The service outputs the analytics extracted from video streams (people count in a selected area, heatmaps estimating people density, and bounding box coordinates of objects of interest), as well as output video feeds with the rendered detections.
- Confluent Kafka [12] message broker, for serving output analytics to other applications in real-time, with minimal delay.
- Grafana [13] dashboard visualization tool, for displaying real-time analytics data to the user. Note that, the Grafana dashboard pulls the data from the Kafka server.
- Several utilities, e.g., a video streaming server, responsible for delivering the output video feed to the dashboard, and tools for configuring and monitoring the Kafka server.

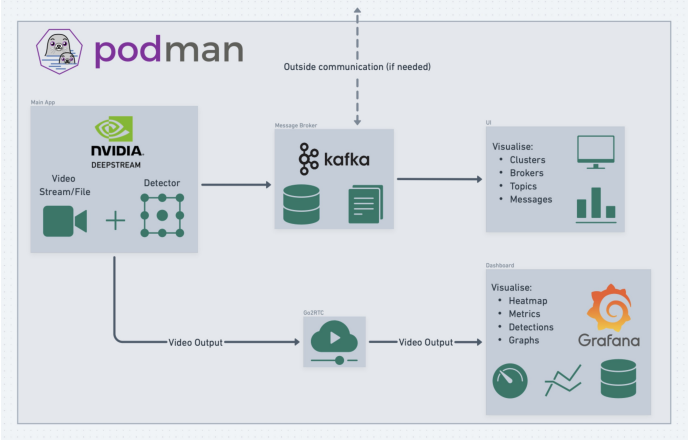


Fig. 3: Architecture of the application platform, showing the main components: DeepStream-based video processing, Kafka message broker and Grafana dashboard

IV. Application Tests

The three specific objectives listed in Section II (O1, O2, O3) are tested in the deployment locations A1a, A1b and A2, using the system described in Section III.

A description of the test for each objective is provided below, along with details regarding applicability to the use cases, followed by screenshots of the results.

A. Objective O1: People / vehicle counting

The target of the first test is to count the number of people and/or vehicles in a predefined region of interest, in every frame.

Depending on the use case, we use different object detector models in the video analytics framework. For UC1 (crowd monitoring), where the only objects of interest are people, we use the P2PNet crowd counting model [14], pre-trained on the ShanghaiTech Part A dataset [15], able to reliably detect the heads of people. For this test videos were collected from the large public ceremony of Saint Parascheva, on October 14th, 2023, with thousands of people gathered on the Stefan cel Mare pedestrian street in Iași. For UC2 (smart traffic monitoring), where both pedestrians and vehicles are present, we use the YOLOv7 [16] object detection model, pre-trained on the COCO dataset [17], since we are interested in different object classes.

The resulting number of people and/or vehicles in every frame are smoothed using an exponential moving average filter, with configurable factor, in order to filter out jitter caused by fluctuations in detection results. The resulting values are published on the Kafka messaging system, for consumption by other clients. In our test, the data is read by the dashboard software which displays the time series to the user.

A screenshot of a stream from the public ceremony at location A1a is shown in Fig. 4 (left), with 782 people counted in the image. Every detected person's head is marked with a light blue dot.

Fig. 5 displays a screenshot from location A2 (Podu Ros intersection) taken during the particularly snowy day of 2024-01-08, showing the video stream and the corresponding pedestrian and vehicle counters. Note that, because the raw counter values undergo smoothing filtering, the resulting values in the time-series are floating-point values. This is helpful for trend estimation but may confuse a human operator expecting to see only integers; in upcoming releases, we shall provide rounded integer values as well.

B. Objective O2: Density and dynamics estimation

The second goal is to estimate the density of people and/or vehicles and their dynamics. For this purpose, the spatial centroids of the detected objects' bounding boxes are computed and then transformed from the image space (pixel coordinates) into GPS coordinates. For this purpose, the camera must first undergo a calibration process, in which the extrinsic parameters are found (camera location and orientation angles). This calibration procedure uses the CameraTransform package [18], which estimates the camera parameters by matching the pixel coordinates of a set of visible landmarks points on the image with their GPS coordinates taken from Google Maps.

Once the GPS coordinates of the detection centroids are available, they are published on the Kafka messaging system, for every frame. In our test, the data is ingested by the dashboard software, which plots the points on a map and displays a heatmap, illustrating the dynamics of the detections in time. This is illustrated in Fig. 4 (right), which displays the crowd heatmap corresponding to the video image on the left side.

Further analysis of the dynamics needed to extract insights in movement patterns are not yet implemented, but these are planned for later stages of the project.

C. Objective O3: Alert when entering restricted areas

The third objective is to detect and alert a human operator of potentially dangerous situations, defined as follows:

- Vehicles are present inside a pedestrian-only area;
- Pedestrians crossing the street outside the crosswalks.

For this test, we use the YOLOv7 neural network model for object detection. The exclusion zones and the object classes that are not allowed in each zone are configured beforehand. When a detected object overlaps with an exclusion zone, the system raises a visible alarm on the dashboard video, thus alerting a human operator.

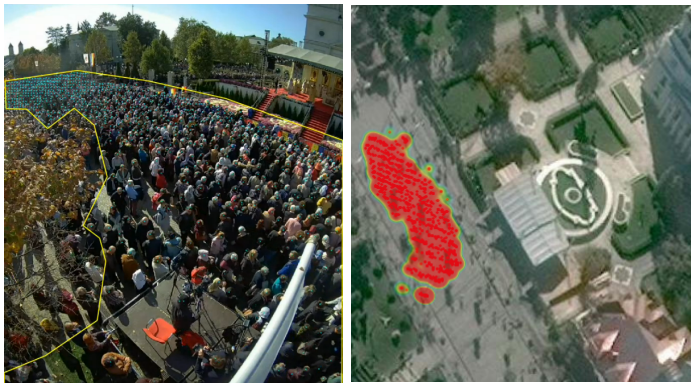


Fig. 4: Visualization for tests O1 and O2 in location A1a for UC1 (crowd monitoring), showing the video stream and people detections (left) and corresponding crowd density heatmap (right). The yellow polygon is the ROI.

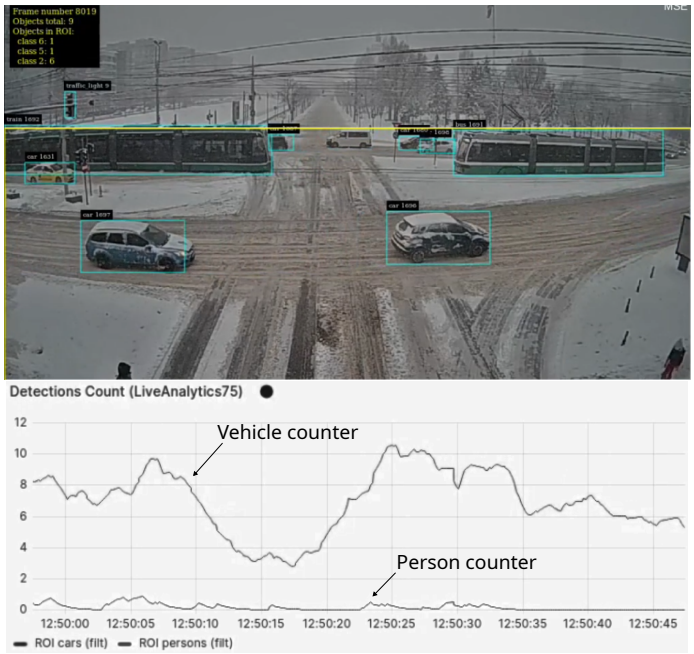


Fig. 5: Visualization for tests O1 and O2 in location A2 for UC2 (traffic monitoring), showing the detections (top) and the live counters (bottom). The values are floating-point due to smoothing filtering. The yellow rectangle marks the ROI.

An example is displayed in Fig. 6, which shows an alert being raised that a pedestrian is crossing the street outside the crosswalk.

V. Network Measurements

A series of network measurements were conducted to evaluate the performance of the cellular network during continuous, stable usage of the application.

Two types of measurements were performed. First, the coverage of the 5G cells in the areas A1a and A1b was assessed, to estimate signal quality and potential handover areas over the entire region. Second, the throughput and latency values were estimated at the locations where the 5G

Fig. 6: Visual alert is raised for an unlawful street crossing (restricted area marked in yellow)

routers are mounted, to verify the requirements necessary to sustain the use cases, listed in Section II.

A. Coverage measurements

The signal quality and coverage of the 5G cells in the overall area around locations A1a and A1b were determined using an independent user equipment (UE) for industrial and commercial applications. This UE supports 3GPP Release 15 SA/NSA operation through a Quectel RM500Q-GL module, capable of a maximum throughput of 2.1/2.5 Gbps download and 900/650 Mbps upload. The cellular connection for a device is supported through its evaluation board (5G-M2-EVB) over USB or WLAN/Ethernet. Network information was collected from the Quectel UE, such as physical cell ID, reference signals received power (RSRP), signal-to-noise ratio (SNR), and other relevant data, All measurements were collected at a frequency of 1 Hz.

Fig. 7 displays the coverage results over the entire area around A1a and A1b locations. The region is covered by 3 cellular antennas, displayed here in different colors, belonging to the same 5G base station (marked BS on the figure). The color intensity represents the 5G RSRP strength, with darker colors signifying better values. We can observe that the A1a and A1b router locations are well covered by the 5G cells, and far from the handover areas, with a good signal quality.

B. Throughput and latency measurements

Several throughput and latency measurements were performed at the locations A1a and A1b, where the 5G routers are installed. Every test was done on a different day, with different conditions. The results are summarized in **Table I**.

The measurements were conducted in two different ways:

- With the same 5G Quectel UE as in the coverage measurements, using the StarTrinity continuous internet speed test tool. The values are averaged over 10 minutes (600 samples), with the standard deviation shown in parentheses. The throughput and latency measurements were made separately, because the latency increases when the UE network is under stress.

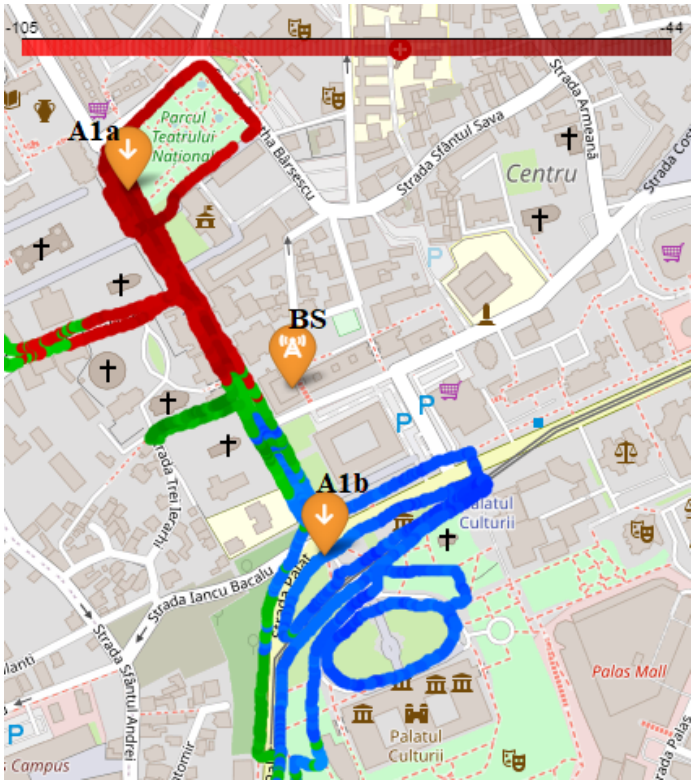


Fig. 7: 5G coverage measurements in A1a and A1b areas. The three cells are displayed in red, green, and blue. A darker color means better RSRP. BS is the location of the Base Station.

- Using the built-in testing capabilities of the 5G routers installed at the locations, which in turn rely on the Speed Test by Ookla tool. The latency values were also manually validated using ping probes between the 5G routers and a virtual machine from the edge-compute infrastructure.

The measured values of the 5G network parameters in the tested locations are capable of sustaining the use cases. For uplink capacity, all the measured throughput values are higher than the minimum required 50 Mbps, with the lowest value being 62.40 Mbps recorded at location A1a. The measured latency values are lower than the maximum required 50 ms, with the worst value being 39.218 ms at location A1a, during the same test.

Note that, there are some differences between the measurements, depending on measurement days and the test equipment. Thus, Test #2 was performed during the public religious service of Saint Parascheva, when the network was congested due to thousands of people being gathered in the area, which explains the lower throughput and higher latency. Concerning the equipment, the 5G routers are placed on street poles at a height of approximately 5m, with better line-of-sight than the 5G Quectel UE, which is handheld at a height of around 1m, and this may explain the differences between the values of Test #1 and Test #3.

The large variation in the throughput and latency values suggests that the network performance is highly dependent

TABLE I: 5G network measurements

Test #	Location	Conditions	Downlink [Mbps]	Uplink [Mbps]	E2E latency [ms]	RX Jitter [ms]	TX Jitter [ms]
#1	A1a	5G UE, height $\approx$ 1m regular day	205.81 (8.90)	76.09 (9.80)	20.44 (3.70)	10.79 (2.00)	12.27 (2.17)
#2	A1a	5G router, height $\approx$ 5m during crowded public event	130.45	62.40	39.218	10.815	-
#3	A1a	5G router, height $\approx$ 5m regular day	720.51	112.46	28.487	1.903	-
#4	A1b	5G UE, height $\approx$ 1m regular day	177.6 (15.00)	76.78 (9.90)	22.11 (4.80)	10.77 (3.35)	12.11 (2.15)

on external factors like the number of users connected to the network. It is worth mentioning that, similar conclusions may be drawn from the results observed for the A2 location (Podu Ros intersection).

VI. Conclusions

In this paper, the implementation of a 5G-based video analytics system was presented for crowd and traffic monitoring, deployed in the city of Iași, Romania. The two use cases presented in the paper were tested in three locations with good results, and the system was able to count people and vehicles in designated regions of interest, estimate their density, and raise alerts when restricted areas were entered. Also, some measurements of the 5G network were conducted to assess its capability of sustaining the use cases, with the obtained throughput and latency values being better than the minimum requirements. The results of the tests are promising, and the system is being further developed with the aim of deploying it in other cities and public spaces.

An alternative approach worth considering for crowd-monitoring is to use WiFi connection reports from the cellular network. Although this is outside the scope of the current paper, we find it worth mentioning as a convenient complementary solution to the video-based data employed in the current system.

Acknowledgment

This work was partially supported by the TrialsNet project, funded from the European Union's Horizon-JU-SNS-2022 Research and Innovation Programme under Grant Agreement No. 101095871 and partially by the Romanian Executive Unit for Research, Development and Innovation Higher Education Funding (UEFISCDI) under Programme 5.8, grant PN-IV-P8-8.1-PRE-HE-ORG-2023-0037.

References

[1] Y. Modi, R. Teli, A. Mehta, K. Shah, and M. Shah, "A Comprehensive Review on Intelligent Traffic Management using Machine Learning Algorithms," *Innovative Infrastructure Solutions*, vol. 7, no. 1, Dec. 2021. [Online]. Available: http://dx.doi.org/10.1007/s41062-021-00718-3

[2] Y. Jiang, Y. Miao, B. Alzahrani, A. Barnawi, R. Alotaibi, and L. Hu, "Ultra Large-Scale Crowd Monitoring System Architecture and Design Issues," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10 356–10 366, 2021.

[3] X. Li, Q. Yu, B. Alzahrani, A. Barnawi, A. Alhindi, D. Alghazzawi, and Y. Miao, "Data Fusion for Intelligent Crowd Monitoring and Management Systems: A Survey," *IEEE Access*, vol. 9, pp. 47 069–47 083, 2021.

[4] M. Ali, G. Lavanya Devi, and R. Neelapu, "Intelligent Traffic Signal Control System Using Machine Learning Techniques," *Lecture Notes in Electrical Engineering*, p. 611–619, Jun. 2020. [Online]. Available: http://dx.doi.org/10.1007/978-981-15-3828-5_63

[5] M. Akhtar and S. Moridpour, "A Review of Traffic Congestion Prediction Using Artificial Intelligence," *Journal of Advanced Transportation*, vol. 2021, p. 1–18, Jan. 2021. [Online]. Available: http://dx.doi.org/10.1155/2021/8878011

[6] H. Xu, A. Berres, S. B. Yoginath, H. Sorensen, P. J. Nugent, J. Severino, S. A. Tennille, A. Moore, W. Jones, and J. Sanyal, "Smart Mobility in the Cloud: Enabling Real-Time Situational Awareness and Cyber-Physical Control Through a Digital Twin for Traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 3, p. 3145–3156, Mar. 2023. [Online]. Available: http://dx.doi.org/10.1109/TITS.2022.3226746

[7] V. K. Sharma, R. N. Mir, and C. Singh, "Scale-Aware CNN for Crowd Density Estimation and Crowd Behavior Analysis," *Computers and Electrical Engineering*, vol. 106, p. 108569, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790622007844

[8] R. Han, R. Qi, X. Lu, L. Huang, and L. Lyu, "Counting in Congested Crowd Scenes with Hierarchical Scale-Aware Encoder–Decoder Network," *Expert Systems with Applications*, vol. 238, p. 122087, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417423025897

[9] M. H. K. Khel, K. A. Kadir, S. Khan, M. Noor, H. Nasir, N. Waqas, and A. Khan, "Realtime Crowd Monitoring—Estimating Count, Speed and Direction of People Using Hybridized YOLOv4," *IEEE Access*, vol. 11, pp. 56 368–56 379, 2023.

[10] "TRials Supported By Smart Networks Beyond 5G (TrialsNet)." [Online]. Available: https://trialsnet.eu/

[11] N. Corporation, "DeepStream SDK." [Online]. Available: https://developer.nvidia.com/deepstream-sdk

[12] "Confluent | Apache Kafka Reinvented for the Cloud." [Online]. Available: https://www.confluent.io/

[13] "Grafana: Dashboard anything. Observe everything." [Online]. Available: https://grafana.com/grafana/

[14] Q. Song, C. Wang, Z. Jiang, Y. Wang, Y. Tai, C. Wang, J. Li, F. Huang, and Y. Wu, "Rethinking Counting and Localization in Crowds: A Purely Point-Based Framework," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct. 2021, pp. 3345–3354, iSSN: 2380-7504.

[15] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma, "Single-Image Crowd Counting via Multi-Column Convolutional Neural Network," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 589–597.

[16] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2023, pp. 7464–7475.

[17] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common Objects in Context," in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham: Springer International Publishing, 2014, pp. 740–755.

[18] R. C. Gerum, S. Richter, A. Winterl, C. Mark, B. Fabry, C. Le Bohec, and D. P. Zitterbart, "CameraTransform: A Python Package for Perspective Corrections and Image Mapping," *SoftwareX*, vol. 10, p. 100333, Jul. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352711019302018

2024 16th International Conference on Electronics, Computers and Artificial Intelligence (ECAI) | 979-8-3503-7115-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/ECAI61503.2024.10607552

UWB Indoor Localization: Accuracy Evaluation in a Controlled Environment

Tiberius-George Sorescu,
Andreea-Valentina Militaru,
Gheorghe Asachi Technical University of Iasi
Email: tiberius.sorescu@etti.tuiasi.ro,
andreea.militaru@etti.tuiasi.ro

Vlad-Mihai Chiriac, Ciprian-Romeo Comsa,
and Iolanda-Elena Alecsandrescu,
Gheorghe Asachi Technical University of Iasi
Email: vchiriac@etti.tuiasi.ro,
ccomsa@etti.tuiasi.ro, ialex@etti.tuiasi.ro

***Abstract*—This paper presents a comprehensive study on the precision of Ultra-Wideband (UWB) technology for indoor positioning in controlled environments. In this study, we evaluate the positioning accuracy of Decawave DWM1001 sensors based on the IEEE 802.15.4 standard within two different spatial configurations: a 3x3 meter area and a 3x12 meter area.**

Our method involved deploying a network of four anchors and one tag in these two setups and conducting a series of empirical measurements to assess the technology's performance. Our findings demonstrate the high accuracy of UWB localization, particularly in smaller spaces. However, our experiments uncovered challenges in larger areas, where environmental factors play a significant role in influencing the accuracy and reliability of positioning measurements, compared to the high accuracy promised by the UWB technology positioning and shown by our simulations in ideal conditions. We experimented a high level of accuracy in the smaller 3x3 meter area, with an average error of 6.950 cm at a 1-meter height. The larger 3x12 meter area exhibited increased error rates, with an average error of 9.339 cm at the same height.

***Keywords*—UWB (Ultra-Wideband), indoor localization or positioning, DWM1001, localization error estimation**

I. Introduction

In recent decades, the rapid advancement of technology has fundamentally reshaped the way we interact with our surroundings. Accurate localization, both indoors and outdoors, is essential in many industries, including navigation, asset monitoring, robotics, and augmented reality applications. The capacity to detect the exact position of items or people in real-time improves efficiency, increases safety, and allows for seamless interaction in smart configurations.

Indoor localization has emerged as a captivating and extensively explored topic in the current era of growing technology and the rising potential of broad Internet of Things (IoT) services. Several technologies have been proposed to provide indoor positioning services, such as Wi-Fi, Bluetooth, Radio Frequency Identification (RFID), and Ultra-Wideband (UWB) [1], [2]. Among them, UWB is considered one of the most promising technologies due to its high accuracy, robustness to multipath, and low power consumption [3], [4].

The applications of indoor positioning extend across a diverse array of domains, each benefiting from the precision and real-time tracking capabilities it offers. To choose only one example, in healthcare settings, indoor positioning systems enable the accurate monitoring of medical equipment, personnel, and patients, facilitating streamlined workflows and enhancing patient care [5].

UWB technology operates by transmitting short-duration pulses across a wide spectrum of frequencies, allowing for accurate distance measurements between nodes. The UWB-based positioning has emerged as a prominent solution for indoor localization due to their high precision and reliability. However, various factors influence UWB localization accuracy and reliability, including clock synchronization issues, Non-line-of-sight (NLOS) circumstances, channel noise, and interference. Given the limitations of Global Positioning System (GPS) due to signal strength, it proves unsuitable for indoor settings like buildings and tunnels. UWB sensors provide effective solutions in contexts where target and anchor nodes are used to identify position. UWB can be used with wireless techniques like Time of Flight (ToF), Time of Arrival (ToA), and Time Difference of Arrival (TDoA) to determine the location of radio sources [6], [7].

Our study focuses on assessing the accuracy of indoor positioning utilizing Decawave DWM1001 [8] sensors based on IEEE 802.15.4 [9] in various scenarios. A configuration was employed comprising four anchors and one tag to evaluate positioning accuracy within two distinct setups: a 3x3 meter area and a 3x12 meter area. By systematically deploying these sensors in real-world settings, we aimed to comprehensively evaluate the performance of the localization system across different spatial dimensions. This approach enables us to provide insights into the effectiveness of the technology in both constrained and extended indoor spaces, offering valuable data for optimizing deployment strategies and informing decision-making processes in practical applications.

The remainder of the paper is organized as follows. In section II, a literature work analysis is presented. The next section describes real-world measurements in two configurations. Section III presents the same two scenarios evaluated at the simulation level to assess the accuracy promises of the technology in ideal scenarios and to understand the implications of the differences in a real-world environment. Finally, Section IV presents conclusions.

979-8-3503-7115-4/24/$31.00 ©2024 IEEE

II. Related work

Localization systems attempt to detect the position of the objects using stationary nodes and mobile computer devices. The applications of indoor positioning extend across a diverse array of domains, each benefiting from the precision and real-time tracking capabilities it offers. In healthcare settings, indoor positioning systems enable the accurate monitoring of medical equipment, personnel, and patients, facilitating streamlined workflows and enhancing patient care [5]. Similarly, in retail environments, indoor positioning technologies support inventory management, personalized marketing, and efficient store navigation, ultimately improving the shopping experience for customers [10].

Furthermore, indoor positioning finds utility in industrial settings, aiding in asset tracking, personnel safety, and workflow optimization. The integration of indoor positioning with emerging technologies like augmented reality (AR) and robotics opens up further possibilities, enabling immersive experiences and enhancing automation in various sectors. Overall, the versatility of indoor positioning systems makes them indispensable tools across industries, driving efficiency, safety, and innovation in diverse applications.

Multiple techniques that are used for indoor localization are presented and compared in [6] such as: multilateration, trilateration, fingerprinting, proximity, vision analysis or different techniques based on the measurement of the propagation time (e.g., ToF, ToA, TDOA, Round-Trip Time of Flight (RTOF)), angle-based techniques (Angle of Arrival (AOA)) and signal strength based (Received Signal Strength (RSS)). Moreover, the combination of TOF and TDOA approaches offer an improved accuracy in localization estimation using a device with low power consumption energy. Thus, in [11] the authors experimentally demonstrated the enhanced localization accuracy achievable through the fusion of ToF and TDoA approaches. This fusion not only refines the precision of position estimation but also proves particularly advantageous for devices with low power consumption requirements. By capitalizing on the complementary strengths of these techniques, the resulting localization system achieves robust performance while mitigating energy consumption concerns, making it well-suited for prolonged deployment in resource-constrained environments.

Our experiment builds upon existing research in the field of indoor positioning systems, focusing on the evaluation of ultrawide-band (UWB) network devices in indoor environments. Previous studies have explored various localization techniques and technologies, including Wi-Fi, Bluetooth, and Radio Frequency Identification (RFID), highlighting the growing importance of accurate indoor positioning across a range of applications. While UWB technology has shown promise for indoor localization, limited empirical evidence exists regarding its performance in real-world scenarios. The results of a precision localization experiment, which yielded a 10 cm accuracy and a 30cm positioning error in most of the cases, were reported in studies [12], [13]. In [7] multiple anchors configurations (7 setups) are presented with a detailed analysis of the UWB localization accuracy for ad-hoc deployments. One experiment revealed that the average distance estimation error between two DWM1001 nodes is 8.6 cm.

Our experiment addresses the need of an analysis assessing the accuracy and reliability of UWB network devices in two distinct layouts: a three-by-three meter area and a three-by-twelve meter area. By carefully controlling the environment and collecting data systematically, our experiment aims to provide empirical evidence on the effectiveness of UWB technology for indoor positioning.

Our experiment extends beyond conventional assessments of localization accuracy by incorporating advanced analysis techniques to comprehensively evaluate positioning errors. In addition to traditional metrics such as deviation, our evaluation includes the generation of heatmaps, providing visual representations of areas with varying levels of precision. By mapping the distribution of errors across the test environments, these heatmaps offer valuable insights into the spatial variability of localization performance. Areas with higher error rates are identified, allowing for targeted interventions to improve system robustness and reliability. Conversely, regions characterized by lower error rates can be leveraged to enhance the overall accuracy of indoor positioning systems.

III. UWB location estimation in a simulated environment

In order to assess the performance of the real-world experiments conducted in the laboratory, simulations were carried out in Matlab [14] environment to evaluate the effectiveness of UWB localization estimation techniques, specifically TOF and TDOA. One simulation was designed for 3x3 meters scenario and another one for larger measurements for 3x12 meters. In both scenarios, a total number of 30 simulation iterations were run, with the tag positions maintained on a fixed grid of 50x50 centimeters between them. Simulation estimates localization via one-way ranging (OWR) [15] frame transmission in the uplink direction, which is accomplished by transmitting broadcast blink signals.

Additionally, the TDOA method leverages the signal transmitted by the tag to four synchronized anchors for pinpointing the device's location. For every test point at which the tag is positioned, the method calculates the time difference as the signal reaches pairs of anchors. From these time differences, hyperbolas representing possible locations are generated. The point where these hyperbolas intersect gives a precise estimate of the tag's position.

Fig. 1 provides a simulation representation of the UWB localization experiment for 3x3 meters setup, while Fig. 2 describes the 3x12 meters setup. The four anchor nodes define the space in which the location of multiple tag points is analyzed. The considered distance between test points is 50 centimeters. Moreover, the blue circles represent the actual tag location while the red stars indicate the estimated tag locations by the algorithm. The localization algorithm specified in IEEE 802.15.4z enables the identification of the intersection points between all the hyperbolic curves determined by pairs of anchors. The task at hand involves estimating the tag location based on these intersections and computing the localization error for each as Root Mean Square Error (RMSE).

The graph (see Fig. 3) depicts an in-depth illustration of UWB localization, focusing on a single tag test point to

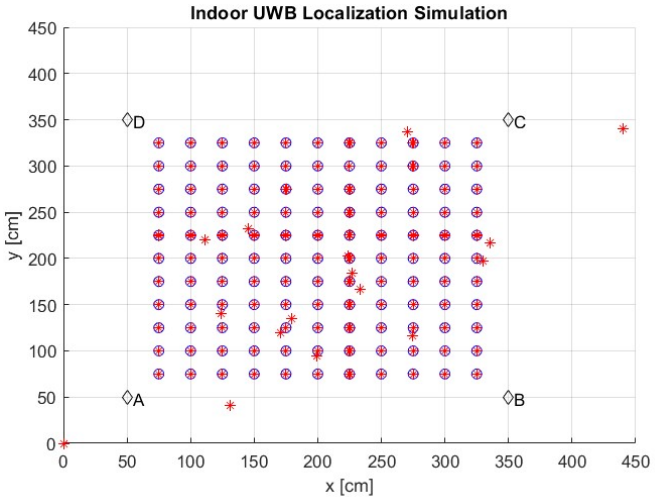


Fig. 1. UWB localization with multiple tag points and RMSE for a 3x3 [m] anchors determined setup

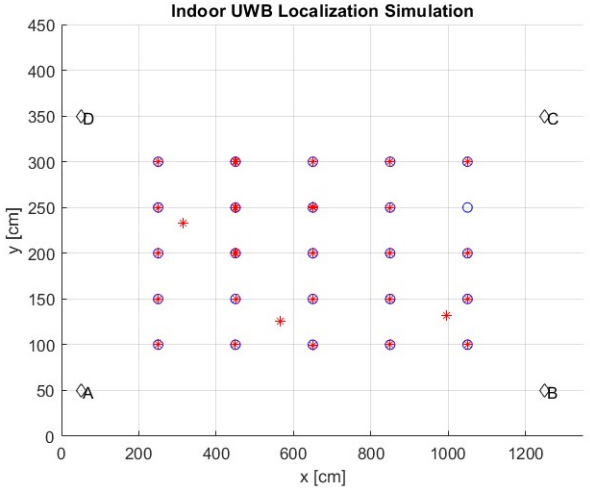


Fig. 2. UWB localization with multiple tag points and RMSE for a 3x12 [m] anchors determined setup

highlight the localization errors, with an emphasis on the centimeters error margin.

The heatmap graphs shown in Fig. 4 and Fig. 5 illustrate the spatial distribution of localization errors in UWB positioning system. The X and Y axes represent the physical space, while the color gradient reflects the magnitude of localization errors, with warmer colors indicating higher error values.

The analysis of heatmap for the 3x12 meters scenario reveals a notable increase in error localization, compared to the heatmap for the 3x3 meters scenario. These findings indicate that the expanded area introduces more challenges for accurate position estimation, resulting in increased localization errors. On the other hand, the heatmap for the 3x3 meters scenario displays fewer and more subdued clusters of errors, underscoring the challenges posed by larger spaces in achieving accurate position estimations with UWB technology. The accuracy of localization is further compromised by the increasing spatial dimensions, which probably increases problems associated

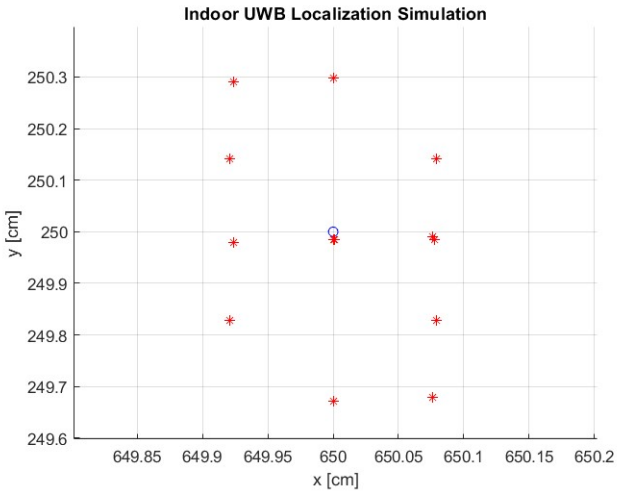


Fig. 3. Fine-grained UWB localization with centimeter-level precision

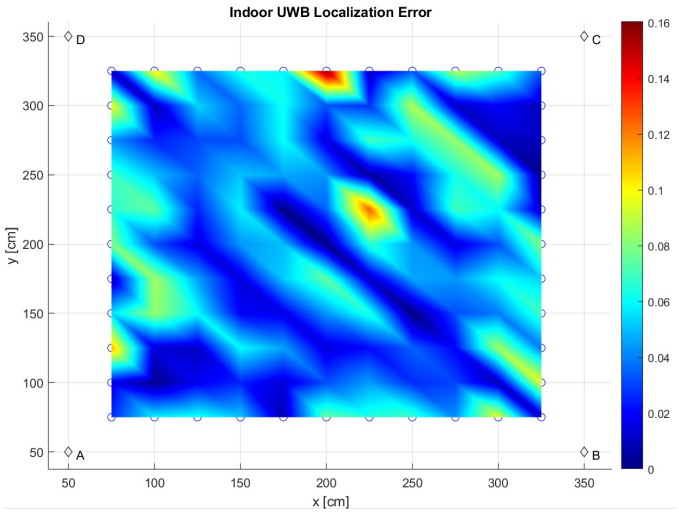


Fig. 4. Localization error distribution heatmap for 3x3 [m] setup

with multipath propagations.

IV. Real-world empirical test-bench

The experiment focused on creating a controlled indoor environment with well-defined dimensions. This allowed for the assessment of ultrawide-band network devices in various scenarios commonly encountered in indoor positioning applications. By carefully controlling the environment and collecting data systematically, the experiment aimed to provide empirical evidence on the accuracy and reliability of ultrawide-band network devices for indoor positioning.

A. Real-word measurements setup

In order to calculate positioning using the Time Difference of Arrival (TDoA) localization technique, four anchor nodes were strategically placed in the test area, as seen in Fig. 6. The role allocation included an initiator module that synchronized timing across all other UWB modules. This was necessary for accurate TDoA measurements among the four active anchor

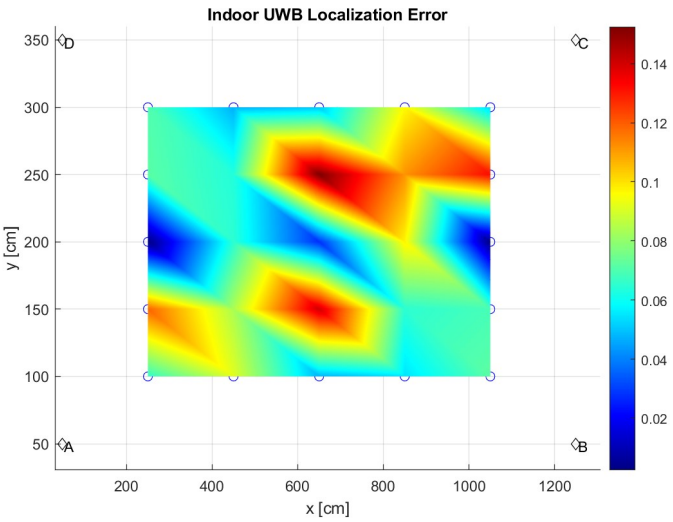


Fig. 5. Localization error distribution heatmap for 3x12 [m] setup

Fig. 6. Laboratory setup illustrating the arrangement of equipment for data collection.

nodes, which were deployed at fixed coordinates in both experimental layouts, placed at ground level in one scenario and at a constant height of 1 meter in a second scenario in two different configurations: a 3x3 meters layout and a 3x12 meters layout. The distance between the UWB modules in both layouts was measured and maintained to ensure accurate positioning. One tag under observation was tracked through multiple iterative runs per setup. A passive listener was incorporated, whose primary function was to collect signal data without actively participating in transmission processes. This allowed for post-processing analysis.

The DWM1001 modules used in this experiment were pre-calibrated with the manufacturer's default calibration. This calibration guarantees precise time measurements and consistent performance across all the modules. The physical layer configuration, MAC layer configuration, localization algorithms, and synchronization settings were all kept at their default values, as provided by the DWM1001 modules.

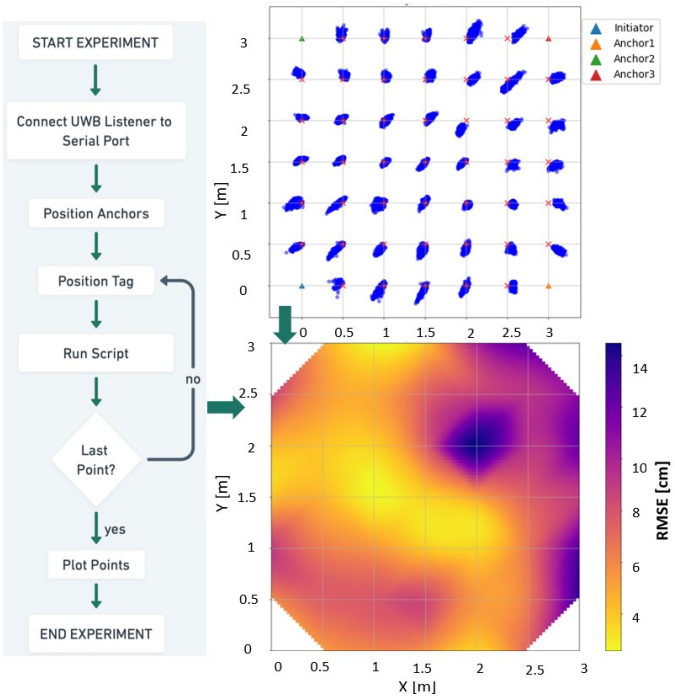


Fig. 7. Spatial distribution of measurement points in a 3x3 meter area

The system used One-Way Ranging (OWR) to estimate the distance between the tag and anchors. OWR was chosen because it is the default ranging method in the DWM1001 modules and provides a simple and efficient way to determine the distances needed for localization. OWR eliminates the need for two-way communication, which reduces the complexity and power consumption of the system compared to Two-Way Ranging (TWR) methods. To ensure accuracy and reliability, 1000 measurements were taken for each measured point in both layouts. The data collection, processing, and plotting were performed using a Python script, and the entire experimental process can be visually examined in the accompanying flowcharts from Fig. 7 and Fig. 8.

B. Results

In the initial layout measuring 3x3 meters, a notable concentration of measurement points around the actual positions was observed. The average error was 6.950 cm for the test with all modules placed at a height of 1 meter. This low average error shows a high level of accuracy in a small, confined space, suggesting that the UWB system, using the TDoA localization technique, performs well in such environments. The dense clustering of measurements, as shown in 7, implies that the system is less prone to errors in smaller areas, possibly because of fewer obstacles or wall reflections causing multipath effects or signal attenuation.

Conversely, the second layout, measuring 3x12 meters, presented a different scenario. Although the overall trend of clustering around the true positions was still clear, the spread of measurements was larger, as depicted in 8. As shown in Table I the average error was 17.871 cm for the experiment with all nodes placed at ground level and 9.339 cm at a height of 1 meter. The increase in error, particularly for ground-level

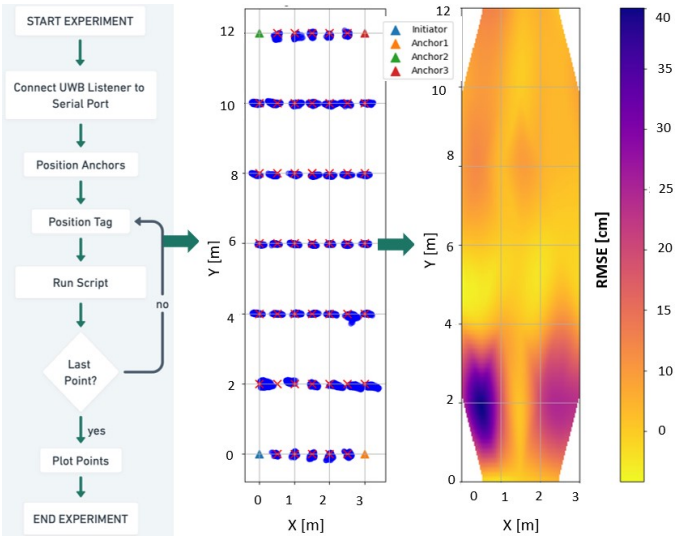


Fig. 8. Spatial distribution of measurement points in a 3x12 meter area

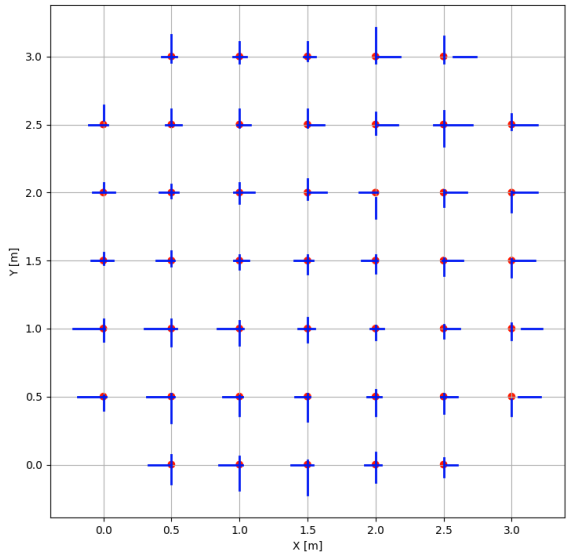


Fig. 9. Spatial distribution of measurement points in a 3x3 meter area

Height	Layout	Avg. Error
0 m	3x3	7.505 cm
0 m	3x12	17.871 cm
1 m	3x3	6.950 cm
1 m	3x12	9.339 cm

TABLE I. AVERAGE ERROR FOR UWB MEASUREMENTS

measurements, could be attributed to the larger area causing more signal dispersion or a greater influence of environmental factors on signal propagation.

A difference in error magnitude between the two heights in the larger layout is observed. One would expect the error to remain consistent. The data shows a decrease in average error when the measurements are taken at a height of 1 meter. This suggests that certain interferences or multipath effects from the floor are less relevant when the system operates at an elevated position.

Further analysis of the UWB indoor measurement results reveals several additional observations. The plots illustrate the distribution of measured positions concerning the actual positions of the anchors. These plots provide valuable insights into the behavior of UWB signals in indoor environments.

In the 3x3 meter layout, the consistent results at both heights indicate that the UWB system is relatively robust in small, controlled spaces. The dense clustering of measurements around the true anchor positions indicates that the system can maintain high precision. This suggests that it could be effectively used for applications that require fine-grained localization, such as asset tracking or robotic navigation in confined areas.

In the 3x12 meter layout, the measurements exhibit a random distribution (see Fig.9, Fig.10). However, they form elongated clusters along the Y-axis. This pattern might indicate a systematic bias or a directional influence, possibly due to the geometry of the space or the orientation of the UWB antennas. Further investigation is necessary to identify the underlying cause and address it to improve the system's performance in larger indoor spaces.

Interestingly, a difference in error magnitude was observed between the two heights in the larger layout. The data shows a decrease in average error when the measurements are taken at a height of 1 meter compared to ground level. This suggests that certain interferences or multipath effects from the floor are less significant when the system operates at an elevated position, potentially because of reduced ground reflections and improved line-of-sight conditions.

Further analysis of the UWB indoor measurement results reveals additional insights into the system's performance. The plots in Fig.9, Fig.10 illustrate the distribution of measured positions relative to the actual positions of the anchors, providing valuable information about the behavior of UWB signals in indoor environments.

In the 3x3 meter layout, the consistent results at both heights show that the UWB system, using the TDoA technique, is relatively robust in small, controlled spaces. The dense clustering of measurements around the true anchor positions suggests that the system maintains high precision in confined areas, making it suitable for applications requiring fine-grained localization, such as asset tracking or robotic navigation.

However, in the 3x12 meter layout, the measurements exhibit a more dispersed distribution. The elongated clusters along the Y-axis might show a systematic bias or directional influence, possibly because of the geometry of the space or the orientation of the UWB antennas. This observation highlights the need for further investigation to identify the underlying cause and develop strategies to mitigate its impact on the system's performance in larger indoor spaces.

The increased error observed at ground level in the larger layout could be explained by factors such as ground bounce, where the UWB signal reflects off the floor, introducing multipath interference. The reduction in error at a height of 1 meter suggests that elevating the UWB system above the

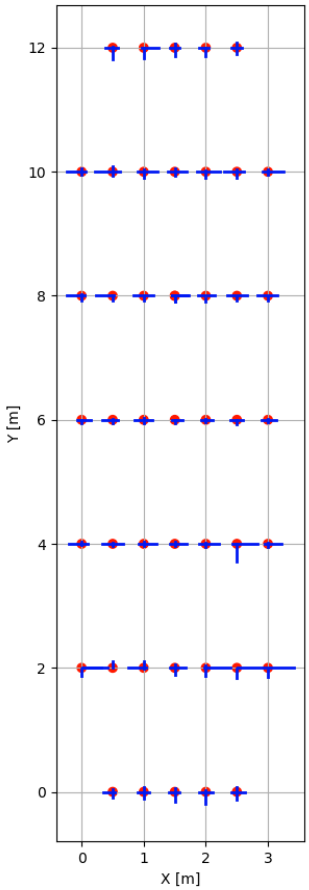


Fig. 10. Spatial distribution of measurement points in a 3x12 meter area

ground may mitigate certain wireless propagation phenomena, leading to improved positioning accuracy.

While the average error is a useful indicator of system reliability, it is equally important to consider the distribution of errors. The way data points are spread around the average value of the error can reveal the system's precision. A small spread indicates that the system consistently produces results close to the average error, which is highly desirable for applications requiring predictability. The shape and orientation of the error clusters can provide insights into multi-path propagation, where signals take multiple paths to reach the receiver. This phenomenon is common in indoor environments because reflections off walls, floors, and objects lead to destructive interference and signal fading.

V. Conclusion

In summary, the findings highlight the high accuracy potential of UWB technology in indoor positioning, especially in smaller spaces. However, the performance in larger areas suggests that environmental factors, signal multipath, and the system's configuration all contribute significantly to the accuracy and reliability of the measurements. Acquiring a thorough understanding of these factors will be crucial in optimizing UWB systems for broader indoor positioning applications.

Further research may focus on refining system configurations, mitigating the effects of signal multipaths, and developing advanced algorithms to enhance localization accuracy in larger indoor environments. Additionally, investigating strategies to address environmental factors such as interference and signal attenuation could further improve the robustness of UWB-based indoor positioning systems.

Acknowledgment

Partial funding for the work was provided by ORANGE ROMANIA S.A. through grant number 29547 / 25.08.2023. Also, the research was partially supported by the Gheorghe Asachi Technical University of Iasi under the framework GNaC 2023 ARUT, grant ASIST.

References

[1] F. Zafari, A. Gkelias, and K. K. Leung, "A Survey of Indoor Localization Systems and Technologies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.

[2] A. Poulose and D. S. Han, "UWB Indoor Localization Using Deep Learning LSTM Networks," *Applied Sciences*, vol. 10, p. 6290, Sep 2020.

[3] C. Wang and Z. Lin, "Precise Indoor Positioning Based on UWB and Deep Learning," *arXiv*, 2022.

[4] F. Che, Q. Z. Ahmed, P. I. Lazaridis, P. Sureephong, and T. Alade, "Indoor Positioning System (IPS) Using Ultra-Wide Bandwidth (UWB)—For Industrial Internet of Things (IIoT)," *Sensors*, vol. 23, p. 5710, Jun 2023.

[5] R. Bharadwaj, S. Swaisaenyakorn, C. G. Parini, J. C. Batchelor, and A. Alomainy, "Impulse radio ultra-wideband communications for localization and tracking of human body and limbs movement for healthcare applications," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 7298–7309, 2017.

[6] A. A. Khudhair, S. Q. Jabbar, M. Q. Sulttan, and D. Wang, "Wireless Indoor Localization Systems and Techniques: Survey and Comparative Study," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 3, p. 392, Aug 2016.

[7] J. P. Queralta, C. M. Almansa, F. Schiano, D. Floreano, and T. Westerlund, "UWB-based System for UAV Localization in GNSS-Denied Environments: Characterization and Dataset," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, Oct 2020.

[8] Decawave Limited, "DWM1001 Data Sheet." https://www.decawave.com/sites/default/files/resources/dwm1001-datasheet-v1.6.pdf, 2024. Accessed: February, 2024.

[9] T. Ye, M. Walsh, P. Haigh, B. O'Flynn, and J. Barton, "Experimental impulse radio IEEE 802.15. 4a UWB based wireless sensor localization technology: Characterization, reliability and ranging," 06 2011.

[10] M. Paolanti, M. Sturari, A. Mancini, P. Zingaretti, and E. Frontoni, "Mobile robot for retail surveying and inventory using visual and textual analysis of monocular pictures based on deep learning," in *2017 European Conference on Mobile Robots (ECMR)*, pp. 1–6, 2017.

[11] R. Mazraani, M. Saez, L. Govoni, and D. Knobloch, "Experimental results of a combined TDOA/TOF technique for UWB based localization systems," in *2017 IEEE International Conference on Communications Workshops (ICC Workshops)*, IEEE, May 2017.

[12] R. Simedroni, E. Puschita, T. Palade, P. Dolea, C. Codau, R. Buta, and A. Pastrav, "Indoor Positioning using Decawave MDEK1001," in *2020 International Workshop on Antenna Technology (iWAT)*, pp. 1–4, 2020.

[13] E. Puschita, R. Simedroni, T. Palade, C. Codau, S. Vos, V. Ratiu, and O. Ratiu, "Performance Evaluation of the UWB-based CDS Indoor Positioning Solution," in *2020 International Workshop on Antenna Technology (iWAT)*, pp. 1–4, 2020.

[14] MathWorks, "UWB localization using IEEE 802.15.4z." https://ch.mathworks.com/help/comm/ug/uwb-localization-using-ieee-802.15.4z.html. Accessed: February, 2024.

[15] MathWorks, "UWB Ranging Using IEEE 802.15.4z." https://ch.mathworks.com/help/comm/ug/uwb-ranging-using-ieee-802.15.4z.html. Accessed: February, 2024.

Validation of Real-world Measurements through Ray Tracing Simulation for 5G Urban Scenarios

Andreea V. Militaru[1], Razvan G. Lazar [2], Carlos Pascal [2],
Ciprian R. Comsa [1], Ion Bogdan [1], Constantin F. Caruntu [2]
[1] *Department of Electronics, Telecommunications and Information Technology*,
[2] *Department of Automatic Control and Applied Informatics*,
Gheorghe Asachi Technical University of Iasi, Iasi, Romania
E-mail: andreea-valentina.militaru@student.tuiasi.ro

2024 International Symposium on Electronics and Telecommunications (ISETC) | 979-8-3503-9086-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISETC63109.2024.10797276

***Abstract*—This paper presents an approach to validate ray tracing simulations using real-world measurements in 5G urban scenarios. The measurements were obtained by traversing specific routes and recording parameters such as position, Signal-to-Interference-plus-Noise Ratio (SINR), Reference Signal Received Power (RSRP), and Reference Signal Received Quality (RSRQ) using the Quectel RM500Q 5G module as user equipment (UE). Subsequently, these real-world measurements are used to validate a MATLAB-based ray tracing simulation. Simulating such scenarios in a controlled environment allows for easy manipulation and study of various parameters, enabling comprehensive verification of the simulator's accuracy. The advantage of this approach lies in its ability to ensure the reliability of the simulation tool, enabling robust testing of diverse scenarios and parameters for deployments in real-world use cases.**

***Index Terms*—5G technology, signal propagation, urban scenarios, real-world measurements, ray tracing simulation, SINR, RSRP, RSRQ, received signal power.**

I. Introduction

A major advancement in mobile communications, 5G technology promises faster data rates, lower latency, and improved connectivity. To achieve these characteristics, accurate signal propagation modeling is essential, especially in complex urban contexts where elements like buildings, human and vehicle movement, and other structures can significantly affect signal behavior [1]. Precise propagation models are necessary for efficient Next Generation Node Base Stations (gNB) placement and network design, guaranteeing reliable and high-quality service for end-users.

Accurate propagation models are fundamental in the design and deployment of 5G networks. They enable precise prediction of signal coverage, interference, and capacity, which are critical for optimizing network performance and ensuring reliable connectivity. In urban environments, where signal behavior is influenced by a multitude of factors such as building materials, heights, and density, accurate modeling becomes even more vital. The effect of small objects producing significant diffraction and reflection of waves leads to a lower received power. A comprehensive survey of ray tracing (RT) propagation modeling for small-cell (mmWave) systems, with a focus on unresolved issues and potential applications, is presented in [2]. The paper also assesses RT's potential for multidimensional prediction, including Multi-Input Multi-Output (MIMO) performance, and emphasizes the need for accurate material parameters and proper diffuse scattering modeling for reliable results. The paper also presents the increasing use of RT models in communication systems for multipath channel profiling and characterization. These models are important for multidimensional channel characterization at (mmWave) and terahertz frequencies, and for planning future wireless networks. The application of RT models has expanded for 5G networks, which operate across multiple frequency bands, ranging from sub GHz to mmWave frequencies between 24 GHz and 100 GHz.

The importance of having a proper communication channel and every user connected to an intelligent infrastructure was studied in [3]. Moreover, planning for network design and gNBs deployment is directly impacted by the results of precise propagation models. Thus, the purpose of this study is to compare the results of real-world measurements with those from MATLAB simulations, where a variety of parameters, including the amount of reflections and diffraction, may be precisely adjusted. We evaluated the quality of our theoretical models and used real-world measurements to detect any outliers that require improvement by contrasting these two sets of data. Thus, the contributions of this study include: the validation of theoretical models, enhanced ray-tracing simulation, and insights for network planning.

The remainder of this paper is organized as follows: Section II provides a summary of the RT method in the context of signal propagation. In Section III, the real-world experiment, including multiple measurements along a specific path, is presented. Section IV details the simulation setup and the obtained results. The comparison and analysis of the results are presented in Section V. Finally, Section VI concludes the paper and suggests directions for future research.

II. Ray-tracing signal propagation methodology

Propagation models are essential for optimizing communication protocols, improving network reliability, and enabling the creation of efficient applications for Intelligent Transportation Systems (ITSs) by simulating the behavior of radio waves in dynamic vehicular environments.

Regarding wireless communication, the ray-tracing propagation model is a deterministic method used to simulate and model the propagation of radio waves in various complex

indoor and outdoor environments. In urban scenarios, these waves interact with multiple objects, e.g., buildings and other obstacles, affecting signal strength, coverage, and quality. Thus, RT helps to predict signal characteristics like path loss, multi-path fading, and shadowing effects by considering the reflections, diffraction, scattering, and absorption of waves [4].

The channel impulse response (CIR) using the ray-tracing method is calculated as follows:

$$h_{R_x}(t) = \sum_{i=1}^{M} a_i \delta(t - \tau_i) e^{-j\phi_i}, \tag{1}$$

where $h_{R_x}(t)$ represents the CIR composed of M time-delayed impulses (rays). Each of these rays is characterized by parameters such as amplitude a_i, arrival time τ_i, and phase ϕ_i, determined by applying Snell's laws [5].

To determine the received signal, one must convolve the transmitted signal $s(t)$ with the channel impulse response $h_{R_x}(t)$. The received signal $r(t)$ is given by:

$$r(t) = (s * h_{R_x})(t) = \int_{-\infty}^{\infty} s(\tau) h_{R_x}(t - \tau) d\tau. \tag{2}$$

This convolution accounts for how the transmitted signal is modified by the propagation environment, considering the multipath effects described by the CIR.

Electromagnetic wave propagation involves various types of rays that interact with the environment, contributing to the overall signal characteristics in wireless communication systems. These rays include direct rays, reflected rays, and diffracted rays. Regarding the ray-tracing algorithms, Image, Shooting and Bouncing Ray (SBR), and hybrid methods are the most well-known as discussed in detail in [6]. Some of the most important aspects and applications of the ray-tracing method in wireless communication include gNB design and placement, wireless network planning, 5G and mmWave systems, virtual environments, and simulation [7], [8].

III. Real-world measurements

This section outlines the data collection methodology, detailing the hardware setup based on the Quectel RM500Q module used as user equipment (UE) and the script developed for communication with the gNBs.

The 5G cellular network's relevant parameters are measured with the setup containing the 5G module connected by a 5G development board to a tablet/laptop linked through an USB-C connection for the power supply and data/command transmission.

The experiment was conducted within a cellular system part of the 5G network used in this study involving two gNBs located at points A and B on the map shown in Fig. 1. The blue marker indicates the position of gNB A, while the orange marker represents gNB B. The measurements were carried out using an UE which was moved along a predetermined route through a densely populated area with high user activity and featuring buildings of various heights. The measurements were executed under the 5G Non-Standalone (NSA) architecture compliant with 3GPP Release 15, similar to the experiments

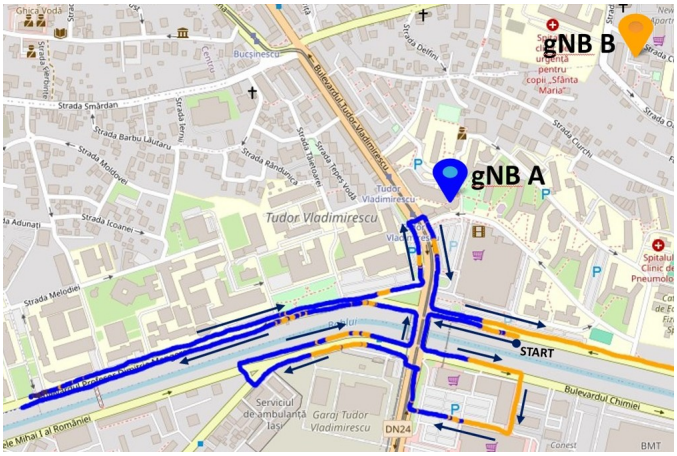


Fig. 1: Signal Strength Distribution from gNB A and gNB B in Urban Environment

performed in [9]. Likewise, the experimental studies conducted in [10], [11] specifically evaluate the performance of 5G communication between a UE and a 5G base station (gNB) at various distances in a dense urban area.

During the experiment, the UE was programmed to measure the received signal strength at regular intervals of 1 second while a pedestrian was walking at 5 km/h. These measurement points are marked on the map in Fig. 1 with orange and blue dots. The blue dots indicate the points where the UE received a strong signal from gNB A, while the orange dots represent the points where the UE was connected to gNB B based on the signal strength.

A key aspect of this experiment is the handover process, which is a standardized and essential mechanism in 5G networks. Note that, the module automatically triggers a handover between the two gNBs as it moves through the coverage area, based on an internal algorithm. The mobile receiver dynamically connects to either gNB A or gNB B depending on the signal parameters such as signal strength and signal power at its current location. This adaptive mechanism ensures optimal connectivity and performance by selecting the gNB providing the best signal at any given moment.

Throughout the measurements, the UE operated within the n78 frequency band for 5G communication, which falls within Frequency Range 1 for 5G New Radio (NR) bands. Employing the Time Division Duplex (TDD) mode, the 5G NR band n78 spans a frequency range from 3300 MHz to 3800 MHz, with a bandwidth of 500 MHz for both UL and DL transmissions according to Orange Romania spectrum - mobile network frequency plan [12]. The actual bandwidth utilized in the measurement area was 120 MHz, which is consistent with the resources the operator had available in this location.

To quantify the signal level, several parameters received by the UE were analyzed without differentiating between the gNBs from which the signal was received. Reference Signal Received Power (RSRP) is a key parameter in assessing the quality of a wireless communication link. It is defined as the average received power of a single resource element carrying

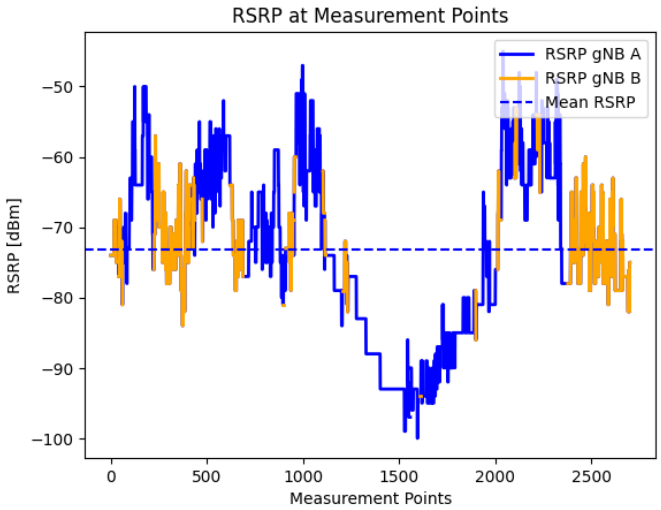


Fig. 2: UE RSRP measured values of the signal received from either gNB A or B (selected according to the system handover algorithm) providing a comprehensive view of the overall signal strength experienced along the path.

the reference signal [13].

Fig. 2 illustrates the RSRP measurements taken at various location positions registered at the UE level. The x-axis represents the location position number, while the y-axis displays the RSRP values in decibels-milliwatts (dBm), from -99 dBm to -47 dBm. The data was collected during real-world measurements in an urban environment, where the UE was moved along a predetermined route as previously described. It is noticeable that there are periods marked with yellow where the UE's RSRP level drops substantially when it is connected to gNB A. These were reported as the UE handed over to the gNB B moments as it can be observed in Fig. 1.

The figure illustrates the variability in RSRP values due to environmental factors such as building obstructions, reflections, and diffraction. Peaks in the graph indicate positions with strong signal strength, likely due to a clear line-of-sight with the transmitter or favorable propagation conditions. Conversely, troughs represent locations where the signal strength was weaker, possibly due to obstructions or multipath effects causing signal degradation.

IV. Simulation Results

In urban environments, RT for outdoor scenarios involves complex interactions of rays with various obstacles such as buildings, trees, and other structures. Diffracted rays play an essential role in accurately modeling these interactions, especially in dense urban environments where line-of-sight paths are frequently obstructed. When a ray encounters an edge or corner of a building, it undergoes diffraction, bending around the obstacle and creating secondary paths that can reach otherwise shadowed areas. This phenomenon is essential for simulating realistic lighting and radio wave propagation, affecting both visual rendering and communication signals. By incorporating diffracted rays into ray-tracing algorithms, simulations can better predict the distribution of light and electromagnetic waves, resulting in more accurate representations of urban environments for applications such as virtual city planning, autonomous vehicle navigation, and urban wireless network design.

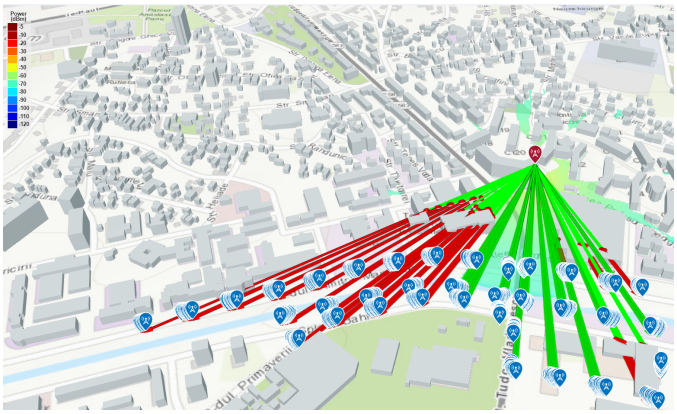


Fig. 3: Simulated propagation signal for gNB A

In this study, the real-world experiment was replicated (see Fig. 3) in a simulated environment using MATLAB and ray-tracing concepts. The primary goal was to identify areas with lower signal strength or regions where the signal is completely absent. Including 3 diffraction rays in the simulations allowed for more accurate modeling of signal behavior in complex urban environments where obstacles can significantly affect signal propagation. As such, the simulation offers a detailed view of signal distribution, highlighting weak or missing signal areas, which aids in more effective planning and optimization of communication networks.

The experiment uses a 3D OpenStreetMap map including buildings offering a detailed topographic view of the urban landscape. Subsequently, transmitter sites are defined based on the gNB A and gNB B coordinates, height, and technical features (transmitter power, transmitter frequency). Further, a ray-tracing propagation model designed for 5G urban scenarios is implemented to enable the analysis of signal behavior in complex urban environments, particularly under non-line-of-sight conditions. The SBR ray-tracing method is employed to visualize coverage areas, with variations in the number of reflections, diffraction, and launched rays studied to understand their impact on received power levels.

Fig. 3 illustrates 3D visualizations of radio wave propagation in an urban environment, obtained through RT analysis. The transmitter gNB A in Fig. 3 is denoted by the red pin, from which signal rays propagate through the route. The rays are color-coded, with green indicating strong signal paths associated with higher received power, and red representing paths where signal level is lower, due to significant attenuation from buildings. Blue markers depict specific locations where signal power was measured, capturing data at varying distances and orientations relative to the transmitter.

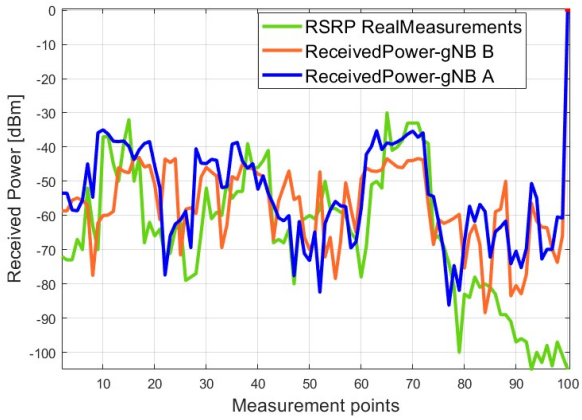


Fig. 4: Received power with 3 reflections

V. Comparison and Discussion

The corresponding illustration obtained using MATLAB (see Fig. 4) shows the received power in dBm for both empirical data and modeling with concrete materials considering buildings and terrain in the analysis. The signal power distribution between the transmitter and receiver illustrates that the power is generally decreasing along longer distances from the transmitter. The real measurements (green line) show notable variations in received power at various site positions, suggesting that possible obstructions and environmental conditions can cause signal strength variations.

The modeled received power using gNB B concrete materials (orange line) and gNB A concrete materials (blue line) also display variability but with different characteristics compared to the RSRP real measurements (see Fig. 5). However, several discrepancies between the real measurements and the model predictions can be observed at certain locations, likely due to the fact that the simulation did not account for frequency-selective fading. For example, the model using gNB A materials has several peaks and drops to 0 dBm, indicating complete signal loss at some locations, which may suggest areas of potential improvement for the model.

These findings have noticeable implications for understanding wireless signal channel behavior and indicate success in providing a model with reliable results. However, there is still a need for additional measurements to compare with the model outcomes to further improve it and validate its usability.

VI. Conclusion

This paper showcased the significance of validating ray tracing simulations using real-world measurements in 5G urban scenarios and the variations that signal propagation suffers from transmitter to receiver. Leveraging real-world measurements for simulator validation provides a robust framework for testing various scenarios and parameters, thereby enhancing confidence in the simulation results and facilitating well-defined decision-making in 5G network planning and optimization efforts. Going forward, this validated simulation approach can be fundamental for further advancements in 5G network design and deployment strategies.

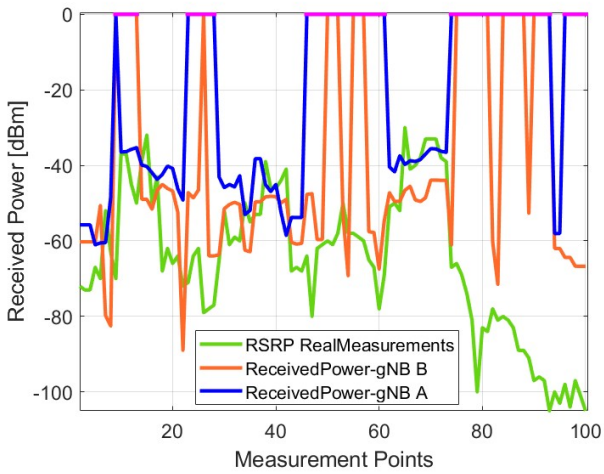


Fig. 5: Received power when concrete mode is enabled for considering building and terrain materials

References

[1] F. Lu, K. Yunoki, and N. Suzuki, "Influence of human body movement on mm wave communication in bullet train cars," in *IEEE International Conference on Ubiquitous Wireless Broadband*, 2016, pp. 1–4.

[2] F. Fuschini, E. M. Vitucci, M. Barbiroli, G. Falciasecca, and V. Degli-Esposti, "Ray tracing propagation modeling for future small-cell and indoor applications: A review of current techniques," *Radio Science*, vol. 50, no. 6, pp. 469–485, 2015.

[3] C. F. Caruntu, A. V. Militaru, and C. R. Comsa, "Cyber-Physical Systems-based Architecture for Signalized Traffic Corridors: Monitoring and Synchronized Coordination," in *23rd International Conference on Control Systems and Computer Science*, 2021, pp. 314–321.

[4] S. Jiang, W. Wang, Y. Miao, W. Fan, and A. F. Molisch, "A survey of dense multipath and its impact on wireless systems," *IEEE Open Journal of Antennas and Propagation*, vol. 3, pp. 435–460, 2022.

[5] I. Bogdan, *Antennas and Propagation: Techniques for Determining User Position in Mobile Communication Networks (in Romanian)*. Venus Publishing House, 2007. [Online]. Available: https://books.google.ro/books?id=RfzHrQEACAAJ

[6] Z. Yun and M. F. Iskander, "Ray Tracing for Radio Propagation Modeling: Principles and Applications," *IEEE Access*, vol. 3, pp. 1089–1100, 2015.

[7] M. Lecci, P. Testolina, M. Giordani, M. Polese, T. Ropitault, C. Gentile, N. Varshney, A. Bodi, and M. Zorzi, "Simplified Ray Tracing for the Millimeter Wave Channel: A Performance Evaluation," in *Information Theory and Applications Workshop*, 2020, pp. 1–6.

[8] S. Hussain and C. Brennan, "An efficient ray tracing method for propagation prediction along a mobile route in urban environments," *Radio Science*, vol. 52, no. 7, pp. 862–873, 2017.

[9] R. G. Lazar, A. V. Militaru, C. F. Caruntu, C. Pascal, and C. Patachia-Sultanoiu, "Real-Time Data Measurement Methodology to Evaluate the 5G Network Performance Indicators," *IEEE Access*, vol. 11, pp. 43 909–43 924, 2023.

[10] R. G. Lazar, A. V. Militaru, C. F. Caruntu, and C. Patachia–Sultanoiu, "Performance analysis of 5G communication based on distance evaluation using the SIM8200EA-M2 module," in *26th International Conference on System Theory, Control and Computing*, 2022, pp. 37–42.

[11] C. Li, S. Wang, J. Yu, W. Chen, and K. Yang, "Distance-dependent V2I wireless channel characteristics and performance in 5G small cell based on measurements," p. 3661–3668, Dec. 2020.

[12] Spectrum-Tracker.com, "Romania orange mobile spectrum," 2024, accessed: 2024-06-17. [Online]. Available: https://www.spectrum-tracker.com/Romania/Orange

[13] 3GPP TS 36.214, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer; Measurements," *3rd Generation Partnership Project (3GPP)*, vol. version 16.4.0, 2020.

Application Design Principles for Road Users' Safety

Andreea V. Militaru*,†, Constantin F. Caruntu*,†, Ciprian R. Comsa*,†
*Gheorghe Asachi Technical University of Iasi, Iasi, Romania
†Continental Automotive Romania, Iasi, Romania
Email: andreea-valentina.militaru@student.tuiasi.ro

2021 International Symposium on Signals, Circuits and Systems (ISSCS) | 978-1-6654-4942-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISSCS52333.2021.9497370

***Abstract*—The increase of safety for the vulnerable road users (VRUs) can be managed by an effective traffic system that implies users' interactions through a smart application. This paper proposes an interactive traffic application based on the strong connectivity of all traffic participants and with the smart infrastructure. Thus, the proposed architectural solution is based on two main components: one part running on the user equipment (UE) with a mobile application that has an intuitive interface, and the other one running in the backbone network, where the pieces of information are collected from different sensors, fused and processed in a proper amount of time.**

***Index Terms*—Vulnerable road users, connected traffic participants, increased safety, multi-access edge computing**

I. Introduction

Due to the large number of vehicles that are sharing the urban streets with the other traffic participants including trams, buses, trucks, motorcycles, bicycles, mopeds, and in the absence of an efficient management system that can fit the current mobility context, inevitable problems do occur. The pedestrians, cyclists and motorized two-wheel operators belong to the group of vulnerable road users (VRUs) and they are the most exposed to traffic accidents. According to the World Health Organization (WHO), each year, 1.35 million people die being victims of traffic accidents, while between 20 and 50 million persons suffer non-fatal injuries. VRUs represents more than half of all traffic deaths [1].

By 2020, a questionnaire was shared among Iasi citizens with the sole purpose of revealing the real difficulties faced by the traffic participants. The survey was completed by 665 persons within one month. The following statistics were generated based on the received answers, Figs. 1 - 2 illustrating most people's choices revealing the intersections prone to congestion and accidents, as well as the most relevant current difficulties in day-to-day Iasi road users' (RUs) life, mentioning that in the case of bicyclists the biggest problem is the insufficiency of proper bicycle tracks.

Also, in Iasi, up to 202 accidents were caused in 2019 by pedestrians, who either did not cross the street on the crossing marks or walked along the road. The cause for 147 accidents was the failure to adapt the speed to the road conditions. Other accident causes are the deviation of the bicyclists from their tracks, non-compliance of the legal distance between vehicles, not giving priority to cars/pedestrians, or driving under the influence of alcohol

*This work was partially supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS/CCCDI - UEFISCDI, project number PN-III-P2-2.1-PED-2019-5296, within PNCDI III.

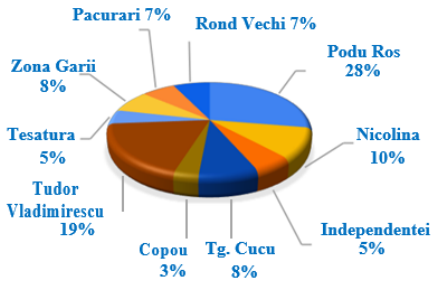


Fig. 1. Frequently congested and prone to accidents intersections from Iasi

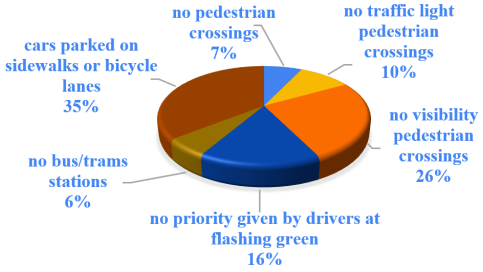


Fig. 2. Pedestrian problems in traffic

[2]. As it can be seen from Fig. 1, the big intersections of Iasi (Podu Ros and Tudor Vladimirescu) bring special difficulties to the RUs. Moreover, as seen in Fig. 2, many of the perceived traffic issues are related to pedestrians and bicyclists, that fall under the category of VRUs.

Thus, there is an urgent need to improve the safety of VRUs, especially in intersections, because they lack protection during collisions when compared to vehicle occupants. This is also highlighted by the huge number of VRUs among traffic fatalities [3]. As such, a mobile application would be useful to allow the VRUs to exchange relevant information (position, velocity, acceleration, heading) with vehicles and infrastructure to prevent and avoid critical traffic situations [4]. Therefore, the goal of this paper is to introduce the design of an application intended to increase the safety of VRUs, by augmenting the awareness of the drivers through the exchange of real-time information between all traffic participants and the smart infrastructure.

II. VRUs Protection

In the last couple of years, the technology advancement in the automotive industry experienced a tremendous leap towards connected and autonomous vehicles (CAVs). But, although there is still much progress ahead to meet the full safety and security requirements for autonomous vehicles, there appears the opportunity of developing safety applica-

tions for VRUs based on their connectivity with vehicles [5]. Regarding the possible use cases, in [6] they are categorized in three broad scenarios: *(i)* VRU high risk zones, *(ii)* interactive communications between VRUs and vehicles, *(iii)* VRU safety messages and algorithms. The first two focus on delivering warnings in certain less-critical situations, while the latter concerns the collision alerts.

In such critical situations, the most stringent problems are related to the latency of the transmitted messages and the localization precision requirements of VRUs for collision warning and prevention. [7] mentions a minimum end-to-end latency time less than 300 ms, while the messages should be generated with a data sampling rate of 10 Hz. In [8], an investigation was performed regarding the impact of communication delay to detect impending collisions between VRUs and vehicles. The study revealed that the average latency for cases where the density is less than 50 people is less than 100 ms, and the delay increases when the sending frequency of the messages is higher, but, generally, Pedestrian-to-Vehicle delay is less than 500 ms in most cases. Moreover, the same analysis emphasized that the localization precision for the VRUs influences greatly the maximum overall delay for detection and communication.

The latest developments in the area of CAVs also lead to the possibility of sharing information regarding the surroundings between traffic participants and even with the smart infrastructure, called *collective perception* [9]. While the autonomous vehicles create an environment model (EM) of their own with the support of data from sensors installed on the vehicle, such a model can be definitely improved to create a comprehensive EM with information that could be hardly perceived or even totally unperceivable from their perspective [10], thanks to the advancement in vehicle-to-everything (V2X) communication technologies.

The roadside infrastructure can take a decisive part in improving the safety of VRUs, but it needs to be embedded with various sensors, e.g., cameras, radars, LiDAR (Light Detection and Ranging) sensors, ultra-wideband (UWB) sensors, GNSS sensors, for passive awareness [11]. For creating VRU awareness through the passive approach, the smart infrastructure uses leading tools such as computer vision and artificial intelligence (AI), which play an important role in object detection and classification, pose estimation, object tracking, and behaviour prediction for all traffic participants.

The following characteristics of the 5G V2X technology will definitely help in improving VRUs protection, satisfying the aforementioned requirements: *(i)* ultra-reliable low-latency communication (URLLC) to support low-latency transmissions of small payloads with very high reliability from a limited set of terminals, e.g., all the traffic participants in an intersection, *(ii)* ultra-high throughput and ultra-low latency connectivity to enable the exchange of raw or processed data gathered through local sensors, *(iii)* very precise positioning and ranging offered by the 5G Positioning and Location services to accurately localizing of user equipments (UEs), *(iv)* ultra-high throughput to build local, dynamic maps based on camera and other sensor data, *(v)* very large number of simultaneous connections in a small geographic area - enables each road user to gather more information about its immediate surroundings [12].

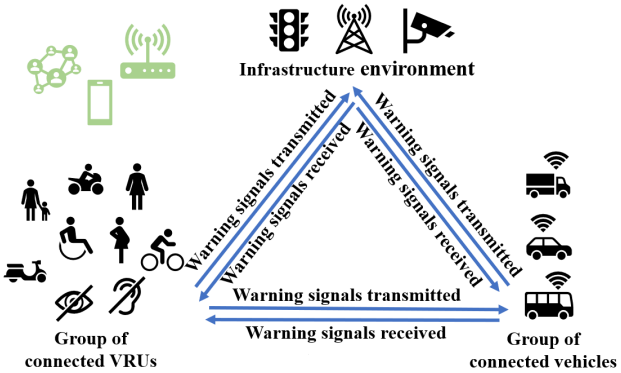


Fig. 3. Diagram illustrating the communication between traffic users

To support all of these characteristics, the cellular vehicle-to-everything (C-V2X) solution includes several types of communications (see Fig. 3): vehicle-to-infrastructure (V2I) - communication between vehicles and road infrastructure, e.g., road sign, traffic light, and vehicle-to-pedestrian (V2P) - communication between vehicles and pedestrians, e.g., sensing nearby persons or cyclists [13], [14].

The inputs from the smartphones of the VRUs and sensors from vehicles and infrastructure represent a huge amount of data that needs a large bandwidth to transmit and great computing capabilities to process. The virtualized cloud resources have impressive processing and computing capabilities, but the multiple-access edge computing (MEC) capabilities enable the placement of storage and computation resources at the network edge, in the proximity of the radio access network (RAN). By processing data locally and accelerating data streams through various techniques, MEC reduces both the end-to-end latency, as well as the traffic overhead towards the core network [6], [15]. Moreover, at crowded intersections, the presence of clusters/groups of VRUs that are using smartphones (or not) can be signalized by the static roadside units (RSUs) that can send alerts further into the network (see Section IV).

Considerable efforts are currently being made to define protocols and to standardize interoperability in the exploitation of information from different technologies. The minimum cooperative awareness message (CAM) frequency for VRU applications is set to 1 Hz by the European Telecommunication Standards Institute (ETSI), while for the V2X transmission the ETSI in Europe and the Basic Safety Message (BSM) in the USA, the frequency of data transmission was standardized between 1 to 10 Hz. The vehicles transmit to the VRUs and infrastructure information about their heading, position and speed, through CAM, while the information from sensors used to increase the precision and accuracy is transmitted through collective perception messages (CPMs). The transmission rate of the sensor information container through CPMs is fixed to a frequency of 1Hz [9], [16].

III. Application Design - Front-end

Nowadays, almost each VRU owns a smartphone that has an internet connection, is outfitted with advanced technolo-

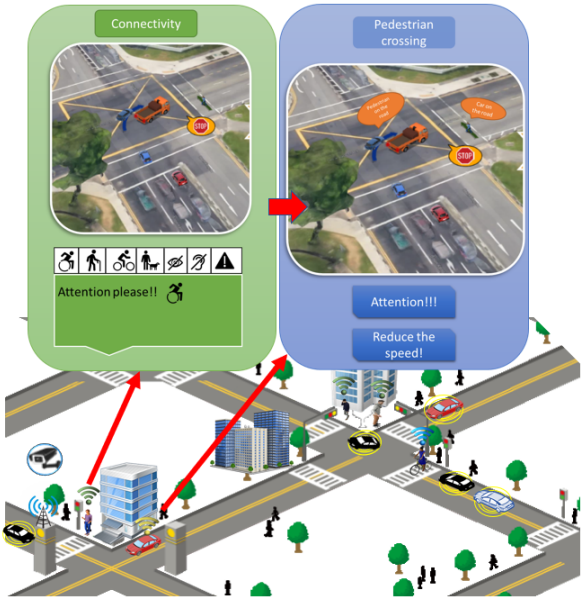


Fig. 4. Interface Application usage in a traffic scenario

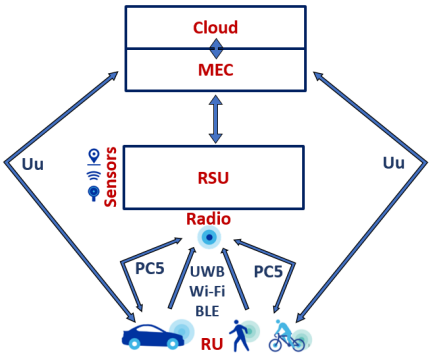


Fig. 5. Physical Architecture

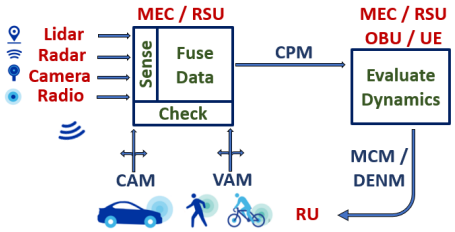


Fig. 6. Functional Architecture

gies, and GPS sensor. Using their smartphones, the VRUs can be always connected with the drivers and the smart infrastructure. Thus, the proposed solution is based on two main components: one part running on the UE with a mobile application that has an intuitive interface, and the other one running in the backbone network, where the pieces of information are collected from different sensors, fused and processed in a proper amount of time. This section focuses on the first part, describing the interface design of the applications for the VRUs and drivers, while the later is addressed in Section IV.

Depending on the type of the RU, the application has a custom design (see Fig. 4) and distinct features, as detailed in what follows. For the VRUs, the attention is focused on the following aspects (green rectangle in Fig. 4): planning a route - the application marks the crowded intersections and unsignalized crosswalks to avoid them; signalizing their intention to cross the street and sending warnings towards approaching drivers when the crosswalk is not signalized or does not exist; choosing a predefined profile if they have any disability, or if they are adults with children needing more time to cross the street; for active persons that prefer to walk or ride bicycles, the application offers the possibility of planning a route that is more on the active part, like combining bicycle riding with walking.

For the drivers, the emphasis is on the following characteristics (blue rectangle in Fig. 4): planning a route; allowing to be warned by the application about crowded intersections and unsignalized crosswalks; receiving proper warning alerts depending on reduced visibility conditions determined by the weather or if their visibility is obstructed by other obstacles; receiving warning signals from the VRUs to be informed in time of what action they will take.

The infrastructure equipped with various sensors based on AI to detect, predict and analyse the traffic situations is not efficient when the VRUs are not in the line of sight or in case of bad weather conditions. Thus, a fusion between data merged by the infrastructure network (described in Section IV) and data collected from the traffic users and analysed by the application is essential for the most accurate results.

IV. System Architecture - Back-end

The proposed application exploits a smart infrastructure that has to provide the proper support for both front-end and back-end sides, respectively on the road user equipment (RUE) and in the Cloud / MEC or RSU. As such, Fig. 5 depicts the physical architecture of the system, while Fig. 6 illustrates the functional architecture.

Due to RUs' detection and classification, precise localization, low latency processing requirements, [3], [5], [7], [17], the application needs to make use of as much as possible data collected by different technologies, e.g., lidar; radar; camera; short range communications, such as Bluetooth Low Energy (BLE), Wi-Fi, UWB, V2X, e.g., PC5 interface for direct device-to-device; or long range communications, such as Uu interface for 5G communication between the UE and the RAN, [6], [11], [13], [14]. With the new connectivity development, the on-board units (OBUs) of the vehicles will be able to communicate with the RSUs or directly with MEC or Cloud server within the 5G network by using specialized messages like the CAM. In a similar way, the connected VRUs will be able to send VRU Awareness Messages (VAMs). These messages contain fields populated at different frequencies with information related to user type, position, positioning precision, heading, speed, lane position, acceleration, environment and other [18], [19]. This information is first validated and checked for confidence, consistency and timing, Fig. 5, before being fed to the Data Fusion unit. While the fusion can take place in either RSU or MEC, its scope is to combine the information related to the nearby RUs coming actively from the RUs themselves or sensed by individual sensors of the smart infrastructure, [20]–[22]. This contributes to the creation of an Augmented

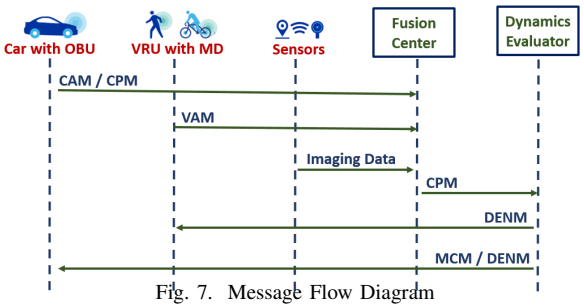


Fig. 7. Message Flow Diagram

Environment Model (AEM), by detecting, classifying and precise localizing the objects representing the RUs. This AEM can be forwarded to the Cloud for further more computationally intensive or global data dependent mobility services, or it can be used locally for developing further mobility functions, or it can be transmitted (e.g., by Uu or PC5) to RUE (like OBUs or UEs) for objects dynamics evaluation and further functions development. The latter two apply to the case of the proposed application that has the scope to implement traffic comfort, efficiency or safety functions [4], [15], [23]. These include diverse functionalities: object tracking, intent and trajectory prediction, abnormal traffic event detection, collision risk evaluation. The proper structuring of the information of AEM is to follow the CPM format that offers the necessary fields for the information on the sensors and perceived objects, on top of the CAM information [9], [16], [21], [24]. Once the safety risk was evaluated as high, in order for the RUE host application to react, a Decentralized Environmental Notification Message (DENM) is vehiculated, containing information on the type of the detected event, as well as on its location. In the future, other type of messages might be used, like advisory Maneuver Coordination Message (MCM), [25], or Infrastructure to Vehicle Information Message (IVIM) [20].

Fig. 7 shows a message flow diagram to clarify the information flux from fusion the AEM based on communicated (CAM, CPM or VAM messages via Uu or short range communication techniques) and sensed imaging information (lidar, radar, camera), to the situation evaluation and warning generated (e.g., via DENM) in case of elevated risk of undesired event detection.

V. Conclusions

This paper tackles the traffic mobility problems in urban areas by proposing an application-based solution for the connectivity between traffic users via smart infrastructure with the aim of increasing the safety of VRUs. The focus is on the architectures needed to implement the proper scene for the smart application that allows a real-time information stream between all the traffic participants.

References

[1] Worls Health Organisation, "Save lives: a road safety technical package," 2017. [Online]. Available: https://apps.who.int/iris/handle/10665/255199/

[2] M. Manole, O. Duma, G. Zanoschi, L. Iliescu, M. Mihai, and A. Manole, "Patterns of road traffic accident in North-East Romania," *Revista medico-chirurgicala a Societatii de Medici si Naturalisti din Iasi*, vol. 120, pp. 1673–679, 2016.

[3] P. Mannion, "Vulnerable road user detection: state-of-the-art and open challenges," 2019.

[4] I. Vourgidis, L. Maglaras, A. S. Alfakeeh, A. H. Al-Bayatti, and M. A. Ferrag, "Use of smartphones for ensuring vulnerable road user safety through path prediction and early warning: An in-depth review of capabilities, limitations and their applications in cooperative intelligent transport systems," *Sensors*, vol. 20, no. 4, p. 997, 2020.

[5] A. Tahmasbi-Sarvestani, H. Nourkhiz Mahjoub, Y. P. Fallah, E. Moradi-Pari, and O. Abuchaar, "Implementation and evaluation of a cooperative vehicle-to-pedestrian safety application," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 4, pp. 62–75, 2017.

[6] "Vulnerable Road User Protection," 5GAA Automotive Association, 2020.

[7] "Intelligent Transport System (ITS); Vulnerable Road Users (VRU) awareness; Part 2: Functional Architecture and Requirements definition; Release 2, ETSI Standard TS 103 300-2 V2.1.1 (2020-05)," ETSI, 2020.

[8] M. Morold, Q.-H. Nguyen, M. Bachmann, K. David, and F. Dressler, "Requirements on delay of VRU context detection for cooperative collision avoidance," in *IEEE 92nd Vehicular Technology Conference*, Victoria, BC, Canada, 2020, pp. 1–5.

[9] F. A. Schiegg, I. Llatser, D. Bischoff, and G. Volk, "Collective perception: A safety perspective," *Sensors*, vol. 21, p. 159, 2021.

[10] M. Herbert, A. Varadi, and L. Bokor, "Modelling and examination of collective perception service for V2X supported autonomous driving," in *11th International Conference on Applied Informatics*, Eger, Hungary, 2020, pp. 138–149.

[11] "Study of Vulnerable Road User awareness," Car 2 Car – Communication Consortium, 2021.

[12] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5G Wireless Network Slicing for eMBB, URLLC, and mMTC: A Communication-Theoretic View," *IEEE Access*, vol. 6, pp. 55 765–55 779, 2018.

[13] L. Miao, J. J. Virtusio, and K.-L. Hua, "PC5-based Cellular-V2X evolution and deployment," *Sensors*, vol. 21, no. 3, p. 843, 2021.

[14] A. Kanavos, D. Fragkos, and A. Kaloxylos, "V2X communication over cellular networks: Capabilities and challenges," *Telecom*, vol. 2, no. 1, pp. 1–26, 2021.

[15] S. Barmpounakis, G. Tsiatsios, M. Papadakis, E. Mitsianis, N. Koursioumpas, and N. Alonistioti, "Collision avoidance in 5G using MEC and NFV: The vulnerable road user safety use case," *Computer Networks*, vol. 172, p. 107150, 2020.

[16] "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Analysis of the Collective Perception Service (CPS); Release 2, ETSI Standard TR 103 562 V2.1.1 (2019-12)," ETSI, 2019.

[17] Q.-H. Nguyen, M. Morold, K. David, and F. Dressler, "Car-to-pedestrian communication with MEC-support for adaptive safety of vulnerable road users," *Computer Communications*, vol. 150, pp. 83–93, 2020.

[18] "Intelligent Transport System (ITS); Vulnerable Road Users (VRU) awareness; Part 1: Use Cases definition; Release 2, ETSI Standard TR 103 300-1 V2.1.1 (2019-09)," ETSI, 2019.

[19] "Intelligent Transport Systems (ITS); Vulnerable Road Users (VRU) awareness; Part 3: Specification of VRU awareness basic service; ETSI Standard TS 103 300-3 V2.1.1 (2020-11)," ETSI, 2020.

[20] A. Jesus, R. Martinez, R. Vilalta, M. Condoluci, Y. Li, A. Kousaridas, P. Spapis, M. Mahlouji, and T. Mahmoodi *et al.*, "Final design and evaluation of the 5G V2X system level. architecture and security framework - deliverable D4.1 of – Fifth Generation Communication Automotive Research and innovation, v1.1," 5GCAR, 2019.

[21] F. de Ponte Muller, E. M. Diaz, J. Perul, and V. Renaudin, "Urban vulnerable road user localization using GNSS, inertial sensors and ultra-wideband ranging," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, NV, USA, 2020, pp. 1846–1852.

[22] A. Ferdowsi, U. Challita, and W. Saad, "Deep learning for reliable mobile edge analytics in intelligent transportation systems: An overview," *IEEE Vehicular Technology Magazine*, vol. 14, no. 1, pp. 62–70, 2019.

[23] M. Bieshaar, G. Reitberger, S. Zernetsch, B. Sick, E. Fuchs, and K. Doll, "Detecting intentions of vulnerable road users based on collective intelligence," 2018.

[24] M. Shan, K. Narula, Y. F. Wong, S. Worrall, M. Khan, P. Alexander, and E. Nebot, "Demonstrations of cooperative perception: Safety and robustness in connected and automated vehicle operations," *Sensors*, vol. 21, no. 1, p. 200, 2020.

[25] A. Correa, R. Alms, J. Gozalvez, M. Sepulcre, M. Rondinone, R. Blokpoel, L. Lücken, and G. Thandavarayan, "Infrastructure support for cooperative maneuvers in connected and automated driving," in *IEEE Intelligent Vehicles Symposium*, Paris, France, 2019.

Redundancy Based V2V Communication Platform for Vehicle Platooning

Adrian Abunei[1], Ciprian R. Comşa[1,2], Constantin F. Caruntu[1,2], Ion Bogdan[1]
aabunei @ etti.tuiasi.ro,
[1] Technical University "Gheorghe Asachi" of Iasi, Romania
[2] Continental Automotive Romania SRL, Iasi, Romania

***Abstract* — Vehicle platooning is one solution to cope with issues such as increased risk of accidents, fuel consumption, and pollution, vehicle components wear, increased number of traffic jams, driving stress and discomfort for the passengers, increased journey duration, large vehicle-to-vehicle gaps. Advanced platooning approaches involve the need for reliable, short latency vehicle-to-vehicle communication (V2V). This work presents results obtained by V2V communication experiment with a low cost, customizable platform that supports different VANET standards for 5.9 GHz and 700 MHz frequency bands. [1]**

***Keywords* — Vehicle platooning, V2V platform, DSRC, OpenC2X.**

I. INTRODUCTION

According to World Health Organization [1], more than 1.35 million people die yearly as a result of traffic incidents. Vehicle platooning could be one solution to improve safety as well as to the other stringent problems caused by the increased number of vehicles on the roads. As such, it has been already proved that platooning reduces fuel consumption with approximately 10%, lowers CO2 emissions with approximately 10%, combats congestion by increasing the road capacity and maintaining optimum vehicle-to-vehicle gaps, reduces the operating costs for the transport sector, reduces the driving stress and passenger discomfort and shortens the journey duration. As vehicles add more sensors, vehicle-to-vehicle communication (V2V) modules are expected to become standard car equipment [2], enabling also advanced vehicle platooning [3].

At the moment, V2V solutions integrate technologies for 4G and 5G network access, Dedicated Short Range Communication (DSRC) and Cellular-V2X (C-V2X) communication [2] . While C-V2X shows promising results [4, 5], DSRC is widely adopted for applications that have high reliability and low latency requirements. For applications such as platooning, an end-to-end application layer latency of about 40 ms or less may be needed, [6], achieved by DSRC, [6]. However, the DSRC reliability suffers due to phenomena typical to wireless communications. In particular, in urban and sub-urban environments, radio propagation effects can significantly reduce the range of radio transmissions. To cope with this, in [7] we demonstrated by simulation that using a supplementary lower frequency, e.g., around 700 MHz, the Received Signal Strength (RSS) is improved, enabling the communication in areas where obstacles obstruct the communication at 5.9 GHz. High costs of commercial devices make practical experiments for DSRC quite expensive. In [8] a cost-effective V2V platform is proposed. The hardware platform is built by cheap components widely available on the market, is customizable, and allows operation on the 5.9 GHz band and supplementary on the 700 MHz band to support the implementation proposed in [7].

Current work demonstrates by real world experiments the feasibility of the solution simulated in [7] by employing the platform presented in [8] for the use case of vehicle platooning by using the speed information. As such, Section II presents the use case of velocity-based vehicle platooning. Section III describes the platform and set-up, while the results of the experiment and conclusions close the paper.

II. VELOCITY-BASED VEHICLE PLATOONING

Vehicle platooning is one of the most important innovations of the automotive industry nowadays, which assumes the forming of clusters of vehicles instead of considering them individually on highways. This smart technology implies that the vehicles travel very closely to one another maintaining a small safety distance between them even at high speeds, compared to how human drivers would behave, i.e., maintaining a larger safety distance that allows them to sense the immediate danger and react in case the vehicle in front suddenly brakes.

The vehicle platoon has a leader vehicle that controls its velocity and direction and the other vehicles follow the leader vehicle at the desired safety distance. The follower vehicles respond to the movement of the leader vehicle through their longitudinal controller by accelerating and braking precisely. For this almost synchronous movement of all the vehicles in the platoon, the control strategies of one follower usually make use of variables related to the leader or to its relative leader, e.g., position, velocity, acceleration, steering angle, inertial heading, forecasts of them all, and sometimes also information shared whether to speed up or slow down. Some of these variables can be determined locally by using the sensors of the ego vehicle and others are transmitted through wireless communication networks (as illustrated in Fig. 1). As such, an onboard computer is connected to a V2V communication device that receives and transmits data using a DSRC system. This exchange of data between vehicles provides enhanced information so that the V2V equipped vehicles can follow their preceding vehicles (relative leaders) with higher accuracy, faster response, and shorter distance gaps.

This was funded in part by Continental Automotive Romania, Iasi.

978-1-7281-3896-1

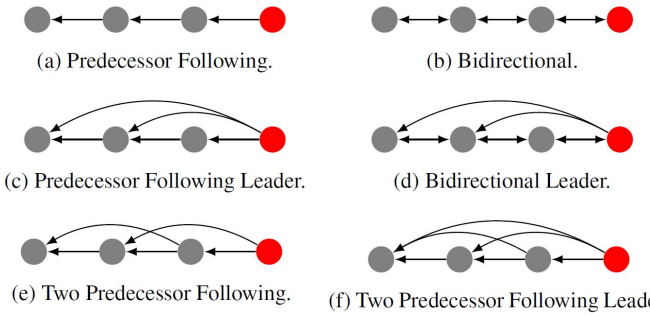


Figure 1.Information flow topologies in a four-vehicle platoon. The red node indicates the platoon leader [9, 10]

The coordinated movement of the digitally connected vehicles has several advantages, e.g., increased highway capacity, improved safety by reducing the risk of accidents, increased efficiency by reducing the fuel consumption up to 10%, which implies also the decrease of exhaust gas emissions, more steady-state traffic flow, reduced travel time, but it also comes with problems related to ensuring the stability of the whole platoon, which is usually referred to as string stability.

The vehicle platooning technology is based on two strategies to determine the distance that should be kept between two vehicles termed as follows:

- *distance-headway*, for which the follower is supposed to maintain a constant distance between itself and the relative leader without considering their velocities;
- *time-headway*, for which the follower must maintain a constant time gap between itself and the relative leader, which means that the distance to be maintained is directly related to the velocity of the vehicles.

In the literature, there are several control methodologies, denoted as cooperative adaptive cruise control (CACC), for the follower vehicles based on information related to the velocity of their relative leader: linear controller [11], optimal controller [12], H_∞ controller [13, 14], sliding-mode controller [15], model predictive controller [9, 16].

III. HARDWARE AND SOFTWARE SETUP

To successfully complete communications tests on both radio bands, the project was divided into several phases. First, we chose a hardware platform with the following requirements: low cost, embedded device, allowing easy and open source customization. Also, this platform should allow radio communication on the 5.9 GHz radio band according to the 802.11p ITS-G5 standard, and on the supplementary lower frequency, around 760 MHz. In order to meet these requirements, the Mikrotik RB433UAH hardware platform was chosen. This allowed the installation of the Ubiquiti XR7 miniPCI radio that operates at 760 MHz band with 10 MHz channel width corresponding to the channels of WiFi 802.11g 2.4 GHz, [8]. This platform can work both as a Road Side Unit and OnBoard Unit. For installation as OBU on vehicles, a GPS USB unit – Ublox 7 was selected. Vehicle ECU (Electronic Control Unit) is formed via USB – OBD II ELM327 module.

Figure 2. Testbed hardware platform components.

The Mikrotik board allows voltage power supply in the 8-30v range with a consumption of approx. 10 W for complete equipment (Fig. 2), thus allowing power from the cigar lighter socket. Also, in order to compare the results after conducting radio tests, antennas with the same gain of 3dB were chosen for both radio modules. The second phase of the project was the development and installation of the open source software [17] on the two OBU vehicles units. Several open source packages that implement the ETSI ITS-G5 protocol have been studied for installation: OpenC2X [18], Vanetza [19], GeoNet Stack [20]. In other works all of these are compiled on LINUX systems with enough hardware and software resources. Embedded systems have limited resources, so the software package needs to be optimized. The Mikrotik RB433UAH runs a proprietary router operating system. It has been replaced with OpenWRT, which is a Linux distribution for embedded devices and allows the installation of V2V communication packages. Although in previous work we used Vanetza software package, currently the most complete and developed open source project for ETSI ITS-G5 communications is OpenC2x (Fig. 3) developed by researchers from the University of Paderborn, Germany. The software package also implements OBDII communication, and the data is stored in the Local Data Map SQL database. Further, the data can be processed, and are very useful when performing real field tests. However, installing the operating system on embedded systems requires cross-compilation on a Linux Ubuntu operating system. The cross-compilation was done in Ubuntu 14, and the resulting binary file that was installed is LED 17.01. Besides dedicated IVC applications, a dedicated software package was developed in order to perform communication parameter measurements for the two radio modules in the 5.9 GHz and 760 MHz Physically, the modules are based on different radio drivers: Atheros ath5k for Ubiquiti XR7 and Atheros ath9k for the 5 GHz radio module. Also, we modified the access layer to allow Atheros cards communication on both ath5k and ath9k drivers for latency results, the TCP/IP protocols v4 stack was installed. The OpenC2X project implements in detail the main service layer features such as DCC (Decentralized Congestion Control), allowing a deeply analysis of the relationship between these services. Also, the OpenC2X protocol layer access layer can

978-1-7281-3896-1

perform data transmission through an OCB-WLAN interface configured through the 802.11p standard.

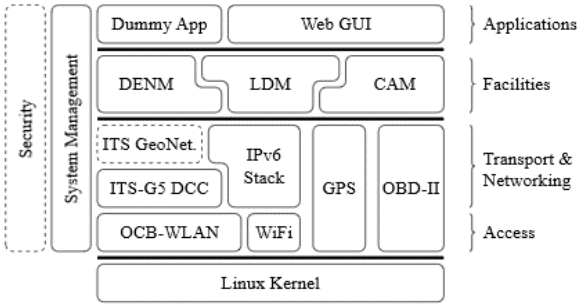


Figure 3. Protocol stack of the ETSI ITS G5 standard for OpenC2X.

We set the first card to work on 5900 MHz central frequency and the second on 768 MHz central frequency. We set identical 21 dB transmitted power on both cards. The OBDII standard is used to access the ECU of the vehicle, to obtain speed information. The GPS service is used to collect vehicle location data. The CAM protocol defined in ETSI 302 637-2 has the function of allowing the exchange of messages between stations (either OBU or RSU) to know at any time the behaviour of the vehicles on the road. CAM (Cooperative Awareness Messages) provide basic awareness service, by means of periodic sending of status data (location, speed, heading) to neighbour nodes. In our case CAM sending frequency is one per second. The data for all neighbour OBU or RSU is collected in Local Data Map database. Besides CAM and LDM (Local Dynamic Map) data, we collect, through our new package, information regarding radio signals RSS and Noise. Also, we stored measured data about communication latency.

IV. REAL FIELD TESTING

The third phase of the project was allocated to real field test. To do this, we equipped two passenger cars, Renault Laguna (leader car) and Opel Zafira (follower car), with the above-mentioned devices. We install the modules on the vehicles roof, for radio visibility, and we connect: the power to the car lighter socket and the OBD adapter through USB cable to the car OBD socket. The tests were conducted on a public road located near Bucharest. We follow the road, one vehicle behind other, with a speed less than 40 km/h, and a distance between vehicles less than 200 m. We choose this value to avoid losing connection on any of the radio bands, knowing that communication on 5.9 GHz band does not allow large distances nor NLOS. Also, the maximum distance was selected to serve a platooning use case. The route was repeated 3 times within a 30-minute period. We recorded data transmitted via the ETSI ITS-G5 protocol on both frequency bands, and we set the access layer on both wireless cards to work only on OCB mode (Out of Context BSS) to allow fast connectivity and low latency.

Fig. 4 represents a capture screen from OpenC2X web GUI application. The leader car is represented with red pin with Cam message, and the follower car has a blue pin with ReceivedCam message.

Figure 4. Screen web caption of OpenC2X application.

V. RESULTS

After completing the tests, we collected and processed the data both from OpenC2X package and from the radio signal and latency measurement application.

In Fig. 5 is represented the first-time interval collected from module on the leader car. The first graphic shows that, the received signal SNR (computed by reporting the RSS to the noise level) for 700MHz (blue line) is greater than 5.9 GHz (red line). The RSS for both frequencies depends the distance between cars, lower distance results in better signal. For distance greater than 130 m the 5.9 GHz SNR is quite low, while the 700 MHz SNR is about 15 dB higher. That means that communication at 700 MHz can be used in harsh environment, in critical situations, even when 5.9 GHz communication may fail.

Fig. 6 represents the latency for both signals in the same time interval, as measured at the leader car, together with graph representing the speed and distance between vehicles.

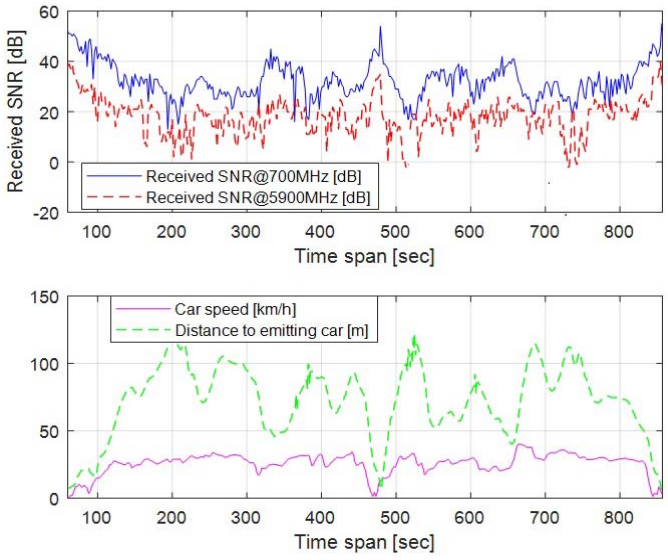


Figure 5. For the leader car: received SNR, speed and distance between cars

978-1-7281-3896-1

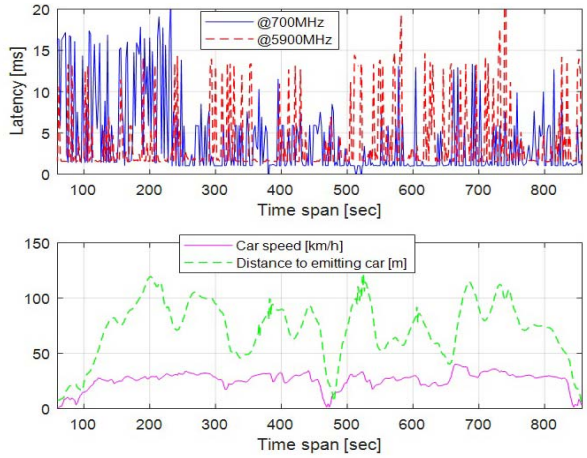


Figure 6. Latency for the two signals together with distance and speed.

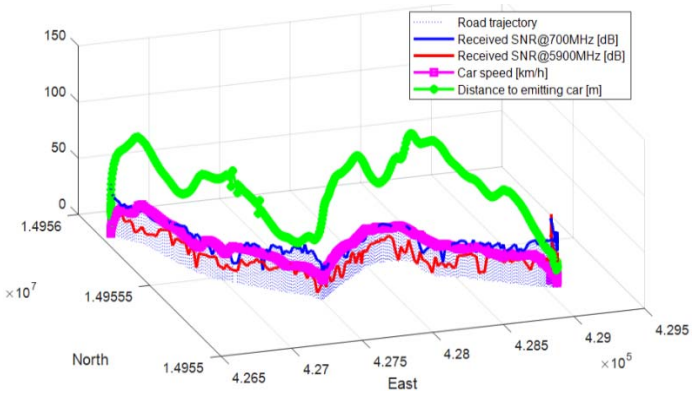


Figure 7. 3D view - Road trajectory map together with received SNR signals, car speed and distance (from the leader car).

In applications like vehicle platooning, latency is a critical parameter, [6]. The latency was calculated from ICMP (Internet Control Message Protocol) messages using Linux Ping command on TCP/IP v4 stack. It shows that latency is lower than 30 ms for both communication signals, similar to the results obtained in [7]. As conclusion, 700MHz communication can be used as redundancy for critical vehicular messages.

Fig. 7 shows in 3D together with road trip map, all the parameters used to compare the results for two types of ITS-G5 communication, showing the dependency of the received SNR (or RSS) to both the vehicle velocity and distance between cars.

For the follower car, measurement data was recorded and results similar as for the leader car can be shown.

VI. CONCULSIONS

This paper describes a real test with two OBUs operating simultaneously on the two DSRC frequency bands, 760 MHz and 5.9 GHz. According to the results, it is shown that the 760 MHz channels can be used to supplement the ITS 5.9 GHz communication.

One important parameter used in vehicle platooning besides position is the speed of vehicles. Through V2V communication data exchange between vehicles forming a platoon can easily used for specific algorithms. It is shown that the latency is the same for the two frequency bands, and OBU unit can switch easily on 760 MHz when 5.9 GHz CAM messaging fails.

REFERENCES

[1] World Health Organization, "Global status report on road safety 2018,"

[2] Continental Automotive. *V2X Communication for Real Time Information Exchange Between Vehicles, the Infrastructure and the Cloud*. Available: https://www.continental-automotive.com/en-gl/Passenger-Cars/Technology-Trends/V2X-Communication

[3] C. Bergenhem, E. Hedin, and D. Skarin, "Vehicle-to-vehicle communication for a platooning system," *Procedia-Social and Behavioral Sciences,* vol. 48, pp. 1222-1233, 2012.

[4] C. Campolo, A. Molinaro, G. Araniti, and A. O. Berthet, "Better platooning control toward autonomous driving: An LTE device-to-device communications strategy that meets ultralow latency requirements," *IEEE Vehicular Technology Magazine,* vol. 12, pp. 30-38, 2017.

[5] 5GAA Automotive Association. (2019, March 2019). *V2X Functional and Performance Test Report; Test Procedures and Results, 30.10.2018,* Available: http://5gaa.org/news/5gaa-report-shows-superior-performance-of-cellular-v2x-vs-dsrc/

[6] S. Gao, A. Lim, D. J. D. C. Bevly, and Networks, "An empirical study of DSRC V2V performance in truck platooning scenarios," vol. 2, pp. 233-244, 2016.

[7] A. Abunei, C.-R. Comsa, and I. Bogdan, "RSS improvement in VANETs by auxiliary transmission at 700 MHz," in *2015 International Symposium on Signals, Circuits and Systems (ISSCS)*, 2015, pp. 1-4.

[8] A. Abunei, C.-R. Comşa, and I. Bogdan, "Implementation of a Cost-effective V2X hardware and software platform," in *Communications (COMM), 2016 International Conference on*, 2016, pp. 367-370.

[9] Y. Zheng, S. E. Li, K. Li, F. Borrelli, and J. K. Hedrick, "Distributed model predictive control for heterogeneous vehicle platoons under unidirectional topologies," *IEEE Transactions on Control Systems Technology,* vol. 25, pp. 899-910, 2017.

[10] J. Guanetti, Y. Kim, and F. Borrelli, "Control of connected and automated vehicles: State of the art and future challenges," *Annual Reviews in Control,* 2018.

[11] A. A. Peters, R. H. Middleton, and O. Mason, "Leader tracking in homogeneous vehicle platoons with broadcast delays," *Automatica,* vol. 50, pp. 64-74, 2014.

[12] A. Alam, A. Gattami, K. H. Johansson, and C. J. Tomlin, "Guaranteeing safety for heavy duty vehicle platooning: Safe set computations and experimental evaluations," *Control Engineering Practice,* vol. 24, pp. 33-41, 2014.

[13] G. Guo and W. Yue, "Autonomous platoon control allowing range-limited sensors," *IEEE Transactions on vehicular technology,* vol. 61, pp. 2901-2912, 2012.

[14] J. Ploeg, D. P. Shukla, N. van de Wouw, and H. Nijmeijer, "Controller synthesis for string stability of vehicle platoons," *IEEE Transactions on Intelligent Transportation Systems,* vol. 15, pp. 854-865, 2014.

[15] J.-W. Kwon and D. Chwa, "Adaptive bidirectional platoon control using a coupled sliding mode control method," *IEEE Transactions on Intelligent Transportation Systems,* vol. 15, pp. 2040-2048, 2014.

[16] W. B. Dunbar and D. S. Caveney, "Distributed receding horizon control of vehicle platoons: Stability and string stability," *IEEE Transactions on Automatic Control,* vol. 57, pp. 620-633, 2012.

[17] J. Fernández Pastrana, "802.11 p standard and V2X applications on commercial Wi-Fi cards," 2017.

[18] S. Laux, G. S. Pannu, S. Schneider, J. Tiemann, F. Klingler, C. Sommer, *et al.*, "Demo: OpenC2X—An open source experimental and prototyping platform supporting ETSI ITS-G5," in *Vehicular Networking Conference (VNC), 2016 IEEE*, 2016, pp. 1-2.

[19] R. Riebl. (2017). *Vanetza, Open-source implementation of the ETSI ITS-G5 protocol stack*. Available: https://github.com/riebl/vanetza

[20] A. Voronov, J. De Jongh, D. Heuven, and A. Severinson, "Implementation of ETSI ITS G5 GeoNetworking stack," *Java: CAM-DENM/ASN. 1 PER/BTP/GeoNetworking*.

978-1-7281-3896-1

ECAI 2018 - International Conference – 10th Edition
Electronics, Computers and Artificial Intelligence
28 June -30 June, 2018, Iasi, ROMÂNIA

Safety Concepts for Body Control Automotive Functionalities

Vlad-Mihai Chiriac
Interior, Body and Security
Continental Automotive Romania
Iași, Romania
Vlad.Chiriac @ gmail.com

Ciprian-Romeo Comșa, Dănuț Burdia
Faculty of Electronics, Telecommunications and Information Technology
Iași, Romania
ccomsa, burdia @ etti.tuiasi.ro

***Abstract* – Automotive industry faces big challenges in providing safe functionalities which in case of malfunction do not put user or others at risk. These safety functionalities need to be developed according to a safety concept which incorporates safety strategies and safety mechanisms. Safety strategies prevent failures leading to dangerous situations, and safety mechanisms bring the system to a safe state in case of failures. Without any direct guidance, the process of creating safety concept can be a challenge. Starting from requirements recommended by automotive safety standard ISO26262, in this paper are proposed safety concepts for power window and low beam functionalities.**

Keywords – Functional Safety Management; ISO 26262; Body Control Module

I. INTRODUCTION

Nowadays, safety is one of the most important issues of the automobile development. With the advancement in the automotive industry, more and more safety related electronically controlled units (ECU) are integrated into vehicle that now consists of larger system architecture with complex interaction and interfaces. The development and integration of these ECUs will strengthen the need for safe system development processes and the need to provide evidence that all reasonable product safety objectives are satisfied. On the other hand, with the trend of increasing technological complexity, software and hardware implementation, there are increasing risks from both systematic failures and random hardware failures. All these aspects pushed the automotive industry to develop functional safety standards as guidelines to keep risk emanated from the system functions at an acceptance level in any condition. ISO 26262 [1], introduced in 2011, provides guidance to reduce these risks by failures avoidance and control by appropriate requirements and processes.

By having certification of ISO 26262, automotive manufactures promote high confidence for customers to purchase automobiles in which prevention of accidents and the reduction of risks is at an acceptable level [2].

ISO 26262 consists of ten parts, as shown in Figure 1. It starts by describing the vocabulary and management of functional safety. Then, it covers from concept phase to the different level of product development which includes system, hardware and software. Automotive Safety Integrity Level (ASIL) decomposition, analysis of dependent failures and safety analyses are explained in part 9.

In the concept phase, hazard analysis and risk assessment (HARA) is performed to determine the safety goals and their ASIL for the system function by a systematic evaluation of hazardous events. Based on severity, probability of exposure and controllability, ASIL is classified into five different levels (QM, A, B, C and D) where level D constitutes the highest level of safety integrity and level A the lowest. QM (quality managed) level corresponds to not safety relevant events. Table 1 contains examples of ASIL classification for some systems [3].

For the development phase, the standard provides requirements to be applied to avoid unreasonable residual risks and, also, requirements for validation and confirmation measures to ensure a sufficient and acceptable level of safety.

In case of safety relevant systems, the additional actions which are needed to reduce the risk to an accepted level are performed to avoid the systematic faults and, also, to control of random hardware faults and systematic faults. The actions for avoidance of systematic faults are implemented by safety management, processes and supporting processes. The control of random and systematic faults is realized by technical requirements within a safety concept which contains safety measures, including the safety mechanisms, to comply with the safety goals.

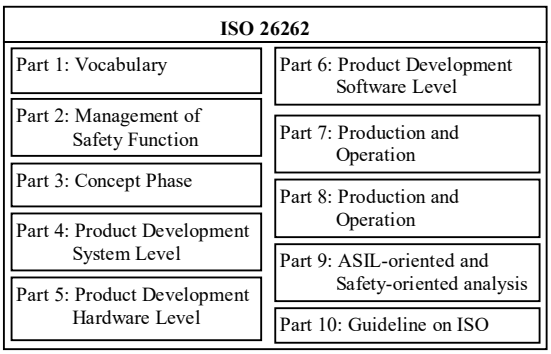


Figure 1. Parts involved in ISO 26262

978-1-5386-4901-5/18/$31.00 ©2018 IEEE

TABLE I. EXAMPLES OF ASIL CLASSIFICATION

System	Hazards	Safety Goal	ASIL
Window Lifter	Pinching limit	Avoid unintended closing	A
Low Beam	Loss of road illumination	Maintain sufficient road illumination	B
Electronic Stability Program (ESP)	Activation of faulty break	Avoid unintended breaking	C
Electronic Steering Column Lock	Activation of faulty locks while driving	Avoid unintended locking	D

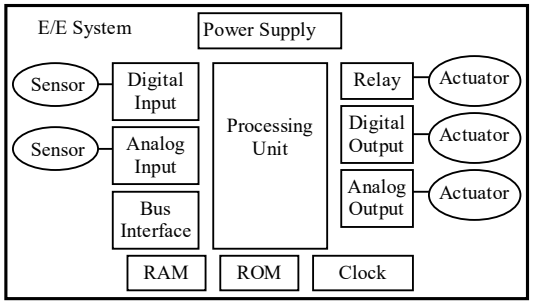


Figure 2. Generic hardware of an electronic control system

At first, functional safety requirements are derived and are allocated to elements based on preliminary architectural assumption of the items. Afterwards, technical safety requirements are refined into software and hardware level requirements.

Starting from guidelines and requirements from automotive safety standard ISO26262, in this paper safety concepts for power window and low beam functionalities are proposed. After a briefly presentation of the Body Control Module (BCM) in Section II, the proposed safety concepts for Power window and Low beam functions are presented in Section III and IV, respectively. Remarks and conclusions are provided in Section V.

II. BODY CONTROL MODULE

Modern cars have more than 30-50 Electronic Computer Units (ECU) to provide numerous functionalities, grouped in four areas [4]: body functionalities (e.g., access, lightning), chassis functionalities (e.g., braking, steering), powertrain functionalities (e.g., ignition, traction control), and infotainment and connectivity (e.g., navigation).

For body functionalities, Body Control Module (BCM) is the central ECU, communicating with many sensors and actuators through tens of input/output interface lines, and with other ECUs through Controller Area Network (CAN), Local Interconnect Network (LIN), FlexRay, and/or Ethernet [5]. BCM can provide many functions like: door lock, power window, exterior lighting, interior lighting, wiper, heating ventilation and air conditioning (HVAC) (see Figure 1). Functions like power window, wiper, and headlamp control for exterior lighting can pose hazards to people if these functions malfunction. In the next two sections, safety mechanisms defining the safety concepts for power window and low beam control functions are proposed.

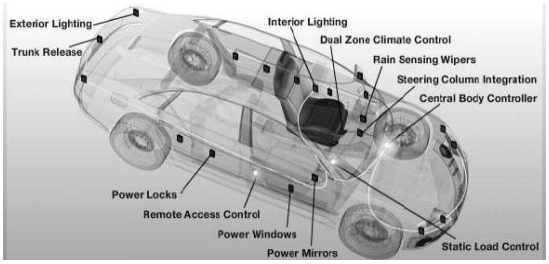


Figure 3. Functionalities provided by BCM [5]

III. POWER WINDOW SAFETY CONCEPT

This section considers power window functionality as a case study in analyzing and developing safety mechanisms and safety strategies.

Power window represents a standard functionality on every car, functionality which permits a user to move window in the car up or down automatically or manually. The request to move window can come from an analog switch or through a Controller Area Network (CAN) message. Based on the analog switch position or on CAN message, a control module actuates an H-bridge controlled motor to move window up or down. Due to a potential malfunction in the actuation system, a hazard can happen by closing the window unintendedly while parts of a human are outside through the window. In [6], power window was analyzed by performing hazard analysis and risk assessment, and it resulted that power window is a safety functionality with ASIL A level, and with safety goal "avoid unintended closing". To prevent closing situations leading to injuries, safety mechanisms and safety strategies need to be implemented in window functionality.

In Figure 4, a block diagram of a possible power window implementation is proposed. The technical implementation is split into four areas: input, processing system, output, and monitoring system. Starting from Figure 2 Generic hardware of an electronic control system, and from guidelines presented in Annex D from ISO 26262 FE Part 5, safety mechanisms and safety strategies for power window are proposed.

Input area contains input signals which can request power window movements, i.e., analog switch signal, and window CAN messages. One of the common faults of switches is represented by stuck-at faults (e.g. short to Ground, short to Power supply), faults which can lead to an unintended closing of window.

A safety mechanism to detect stuck faults is to use coded resistive switches which provide different voltage ranges when the switch is pressed, pulled, or released. Using this type of safety mechanism, a diagnostic coverage of 90% can be obtained.

An unintended actuation of power window can result from erroneous CAN messages containing actuation requests. CAN messages can be affected by different faults (e.g. repetition, deletion, insertion, corruption, etc) [7], faults which can lead to closing window unintendedly. To protect from such faults, the window CAN requests must be protected by a Cyclic Redundant Check (CRC) and a counter message.

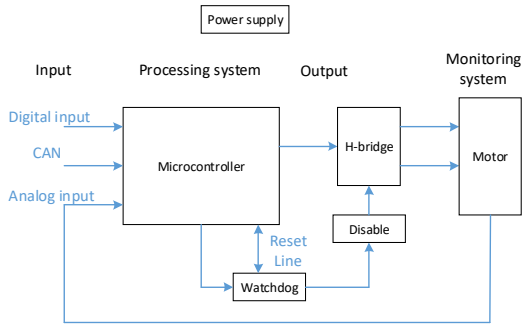


Figure 4. Power window block diagram

CRC represents a code added to data which is used to detect corruption of data during transmission. The counter message protects against repetition, deletion, and insertion faults. These safety mechanisms lead to a diagnostic coverage of 99%.

Processing or controlling system is formed by microcontroller which need to verify that the input signals are not damaged, to actuate H-bridge, and to detect and to react to pinch situations. A pinch situation occurs when during an upwards movement, the window exerts to an obstacle a force higher than a permissible threshold. ISO 26262 FE [1] presents the following failures which could affect the SW running in microcontroller: wrong coding, wrong or no execution, execution out of order, execution too fast or too slow, stack overflow/underflow. In case of these failures, it is possible that microcontroller could not detect pinch situations or could not react to pinch situations. Thus, an external watchdog monitors the correctness execution of program sequence, and in case of faults it resets microcontroller and concurrently, disables the H-bridge, stopping any window movement. Moreover, the external watchdog has a separate time reference than the microcontroller, and thus watchdog could detect faults in microcontroller clock like incorrect frequency, or period jitter.

The output area is represented by the microcontroller signals which control H-bridge, by H-bridge and by motor. To check that the output area is working correctly, the monitoring system provides information to microcontroller related to the motor rotation, the rotation direction, and if the torque force is increasing, as it happens in a pinch situation. The monitoring system can be implemented using hall sensors, for example.

The monitored information is provided to the microcontroller through analog digital converters, which can be tested during window rest phases with a reference signal to detect static failures (stuck-at failures) and cross-talk.

Using this safety mechanism (i.e., Test pattern Table D.7 — Analogue and digital I/O [1]), a diagnostic coverage of 90% for detecting failures in ADC convertors can be obtained. Moreover, because the digital conversion depends on the reference voltage signal, microcontroller needs to monitor the reference signal for over voltage and under voltage situations.

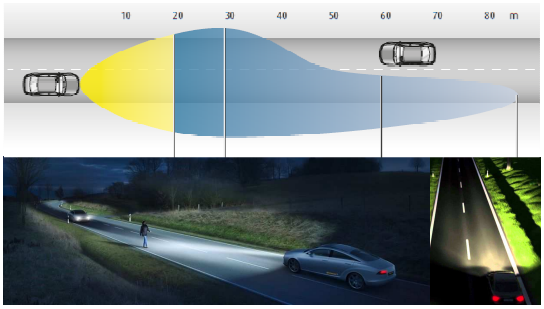


Figure 5. Road illumination by Low Beam

IV. Low Beam Lighting Safety Concept

One of the main classes of BCM functions is the control of exterior front lighting. Within this class, from safety point of view, we propose an analysis of the low beam lighting. The low beam lighting is ensured by the two, left and right, headlights. The low beams provide a light distribution to give adequate forward and lateral illumination without dazzling the oncoming traffic, as depicted in Figure 5. The low beams remain switched on when activated in order to provide illumination of the road for at least 50 meters such that to be able to drive safely even under critical (high speed in the dark) environmental conditions [8]. The activation can be performed manually by the driver through a switch or automatically, in case the car provides this functionality.

Out of the Hazard Analysis and Risk Assessment [9], a typical Safety Goal says that "while driving, loss of sufficient illumination level of the road shall be prevented“ with ASIL B. A typical Safe State is “At least one low beam is ON”. As emergency operation, it is required to inform the driver whenever possible via the Instrument Cluster. From this point on, the design of a Safety Concept is an open topic that we address in the followings.

Starting from [10] and [11] our proposed block diagram for the implementation of the low beam functionality is given in Figure 6. The Safety Concept includes a safety strategy to ensure the intended functionality and prevent violation of the safety goal, as well as a set of safety mechanisms to detect and control potential failures that could lead to the violation of the safety goal. Using guidelines presented in Annex D from ISO 26262 FE Part 5 [1], we propose for each block appropriate safety mechanisms.

For low beam control, the power supply is redundant in the sense that there are more than one physical input supply lines from the car battery. At the same time, to detect faults like *drift and oscillation*, *under and over voltage*, *power spikes*, safety mechanisms like *D.2.8. Voltage or current control* [1] shall be used for the supply chain to ensure proper functionality of the component blocks, especially the microcontroller and the output driver.

As the triggers for turning the two low beams (left and right) are the light switch and the ignition switch, the inputs coming from these switches have to be checked for *consistency*, *integrity* or *stuck condition*, by measures like *filtering*, *debouncing* and *pattern testing* according to measure set *D.2.6* [1].

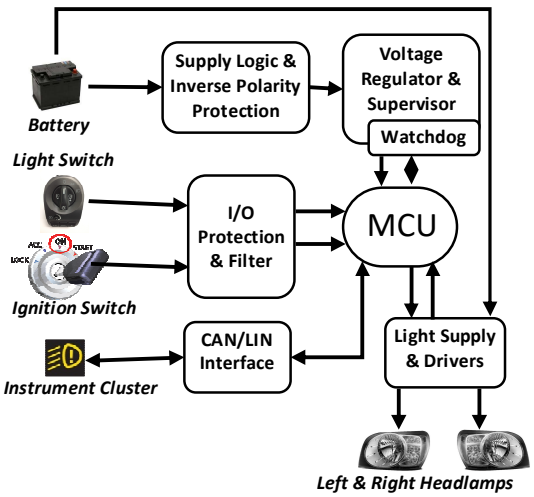


Figure 6. Low Beam control system architecture

The communication protocols, be it CAN, LIN or SPI, have to include mechanisms to cope with situations like *failures of communication peer*, *message delay*, *corruption* or *loss*, *unintended message repetition* or *insertion* [7]. The mechanisms shall include the so called end-2-end (E2E) protection by combining *D.2.7.6. Information redundancy*, *D.2.7.7. Frame counter*, and *D.2.7.8. Timeout monitoring* [1].

For the execution of the designed algorithms, the units that have to be monitored for possible failures are *D.4. Processing unit*, *D.10 Clock*, *D.5. ROM*, *D.6. RAM*. For the first two, the safety mechanisms proposed from section *D.2.9* of ISO 26262 Part 5 [1] are the *watchdog with separate time base* and the *logical monitoring of the program sequence*. For the memory related failures, several mechanisms have to be selected from sections *D.2.4* and *D.2.5*, e.g., *error-detection-correction codes*, *checksum*, *parity bit*, *block replication*, *double-inverted storage*, *RAM pattern/ march test*.

The blocks that are actually responsible for actuating the low beams are the ones related to voltage supply and output drivers, one set for each side, left and right. Moreover, with modern technologies, like LED or laser lighting, and increased complexity of the lighting functions, many manufacturers design separate ECUs especially for driving each of the headlamps. Although there are two such units, we state that ASIL decomposition is not feasible in general due to the fact that the units are similar, i.e., same design, same development. This situation falls into the category of the so-called homogeneous redundancy.

No matter they are separate units or blocks within same ECU, the light drivers have to satisfy same safety requirements. As such, the communication with the light drivers has to be reliable, the processing unit of the drivers has to also be reliable and they have to perform *diagnosis of the outputs* to insure proper command response. Thus, mechanisms for communication, supply and processing similar to those previously described have to be implemented.

In addition, monitoring and feedback mechanisms have to be necessarily involved to detect possible failures like *open load*, *short circuits*, *voltage level*, *insufficient current intensity*.

V. Conclusions

In this paper we determined what measures should be implemented within a Body Control Module to ensure the functional safety of the window lifter and low beam automotive functionalities. While a repository of general requirements is given within the ISO26262 standard, which measure is appropriate for a specific function is a difficult engineering topic that we tackled within this paper.

Acknowledgment

Part of this work was supported by Continental Automotive Romania.

References

[1] International Standards. (2011). *ISO 26262 Functional safety for road vehicles*. Geneva, Switzerland.

[2] A. Ismail, L. Qiang, “ISO 26262 automotive functional safety: issues and challenges”, International Journal of Reliability and Applications, vol. 15, no. 2, pp. 151-164, 2014

[3] J. Schwarz, J. Buechl, “Preparing the Future for Functional Safety of Automotive E/E-Systems”, 21st (ESV) International Technical Conference on the Enhanced Safety of Vehicles, pp. 1–3., 2009.

[4] B. Groza, H.E. Gurban, and M. Pal-Stefan, "Designing security for in-vehicle networks: a Body Control Module (BCM) centered viewpoint," on the 46th Annual IEEE/IFIP International Conference on *Dependable Systems and Networks Workshop*, 2016.

[5] T. Martinez, “Body Control Module,” Freescale, 2009, accessed 2018.04.01, https://www.nxp.com/files-static/training_pdf/WBNR_LA_AUTO_BCM_SPANISH.pdf

[6] Schwarz, J., “Functional Safety and Automotive Software – Introduction ISO 26262 Daimler”, 10th Workshop of Critical Software System, Tokyo, pp. 27-28, September 2012.

[7] J. Hedberg, A. Söderberg, T. Malm, M. Kivipuro, and H. Sivencrona, “Methods for Verification & Validation of time-triggered embedded systems,” Technical report NT TR 600, December 2005.

[8] Valeo, *Lighting Systems. From light to advanced vision technologies*, accessed 2018.04.01, https://static.valeoservice.systems/sites/default/files/catalog/lighting_systems_from_light_to_advanced_vision_technologies_technical_handbook_valeoscope_en_998542_web.pdf

[9] Autosar Standard, *Safety Use Case Example AUTOSAR CP Release 4.3.1*, accesed 2018.04.01, https://www.autosar.org/fileadmin/user_upload/standards/classic/4-3/AUTOSAR_EXP_SafetyUseCase.pdf

[10] Texas Instruments, *Body Control Module. Reference Design*, accessed 2018.04.01, http://www.ti.com/solution/automotive_central_body_controller

[11] Renesas, *Body Control Module (BCM) / LED Head Light Unit*, accessed 2018.04.01, https://www.renesas.com/en-eu/solutions/automotive/body/body-control.html

PERFORMANCE BOUND FOR TIME DELAY AND AMPLITUDE ESTIMATION FROM LOW RATE SAMPLES OF PULSE TRAINS

Ciprian R. Comsa [*,†], *Alexander M. Haimovich* [†]

[*] Telecommunications Dept, "Gheorghe Asachi" Technical University of Iasi, Romania
[†] ECE Dept, New Jersey Institute of Technology, USA

ABSTRACT

The problem considered is estimating the time delays and amplitudes of a train of pulses of known shape. It was recently shown that in the absence of noise, if the pulses are short relative to the interval separating them, the stream of pulses can be sampled at rates lower than the Nyquist rate without any loss of information. This paper investigates the performance of time delay and amplitude estimation from samples taken at low (sub-Nyquist) rates when the stream of pulses is affected by additive white Gaussian noise. To this end, the Cramér-Rao bound (CRB) is developed. For a particular setup, the CRB for time delay estimation is inverse proportional to the cube of the sampling rate, while the CRB for amplitude estimation is inverse proportional to the sampling rate. Numerical simulations confirm this result.

Index Terms — Cramér-Rao bound, time delay and amplitude estimation, low rate sampling, pulse train

1. INTRODUCTION

In many applications, signals can be uniquely described by a small number of parameters. For example, in radar and wireless communications signal can be constituted of a stream of short pulses of known shape, stream that is defined by the time delays and amplitudes of the pulses. Estimation of these parameters is the subject of this work. Time delay estimation and amplitude estimation are conventionally performed from samples taken at rates higher than the Nyquist rate, which is twice the signal's bandwidth. This approach is required when the only knowledge on the signal is that it is bandlimited. Other priors on signal structure can lead to more efficient sampling. An interesting class of structured signals was considered in [1], where the signals have a finite number of degrees of freedom per unit time. These signals were termed to have finite rate of innovation (FRI). A special case of FRI signals consists of a stream of K pulses of known shape per time interval T. For such a stream, any of its segments of length T is uniquely determined by no more than $2K$ parameters, i.e., K time delays and K amplitudes. It is then said that the signal's local rate of innovation is finite and equals $2K/T$.

For streams of pulses of short duration relative to the reference interval T, standard sampling methods require very high sampling rates. However, the rate of innovation per interval T is much smaller than the number of samples taken at the Nyquist rate. This observation was exploited to set the grounds for sampling at rates lower than the Nyquist rate. To this end, a mechanism to sample at low rates streams of Diracs can be found in [1], and the references therein. A scheme for recovering the original stream from the samples was also proposed. More recent work, [2, 3], generalizes the approach to sampling at low rates streams of pulses of known arbitrary shape, rather than just Diracs, by considering a multichannel sampling setup. Although the proposed multichannel sampling scheme allows perfect recovery of the original signal from its noiseless samples, it was found in [4] that in the presence of noise the performance of the signal recovery from low rate samples deteriorates significantly relative to the signal recovery from samples taken at the Nyquist rate. However, with low rate sampling, increasing the sampling rate above the rate of innovation brings substantial performance improvement.

This work considers low rate sampling of streams of pulses as in [2, 3], but rather than recover the signal itself, the goal is to estimate specific signal parameters, specifically, delay and amplitude. The signal recovery in [2, 3] is equivalent to estimating all the unknown signal parameters, e.g., time delays and corresponding amplitudes. By contrast, estimation of individual parameters may be performed independently, e.g., delay estimation does not require the estimation of amplitudes, and thus its performance may vary with system parameters differently than that of signal recovery.

The paper is organized as follows: Section 2 provides the system model for a multichannel sampling scheme. In Section 3, CRB expressions are developed for TDE and amplitude estimation. Discussion of our results vis-à-vis other work is included. Section 4 contains numerical results, and conclusions wrap up the paper.

2. SYSTEM MODEL

The problem considered is estimation of a vector $\boldsymbol{\theta}$ that parameterizes a FRI segment $x(t)$ of length T_0 from samples taken at low rates from a noisy observation $y(t)$. To give practical relevance to the problem considered, $x(t)$ is viewed as the product of transmitting a signal $x_T(t)$ through a multipath channel $h(\tau,t)=\sum_{k=1}^{K}\alpha_k\delta(t-\tau_k)$, where α_k are the complex valued amplitudes and τ_k the time delays associated with each propagation path. The transmitted signal $x_T(t)$ consists of a train of pulses. That is, a pulse of known shape $g(t)$ is emitted periodically at a constant rate $1/T$, after being modulated by some arbitrary, known sequence $x_M[n]$. Vector $\boldsymbol{\theta}$ contains then the time delays and amplitudes as the unknowns that parameterize the signal $x(t)$. It is assumed that $T_0 \gg T > \tau_k$, for any $k \in \{1,\dots,K\}$. Also, the multipath channel is assumed time invariant. Thus, the amplitudes α_k do not change from one interval $\mathrm{I}_n=[(n-1)T, nT]$ to another. Additionally, the received signal is shift invariant, i.e., the time delays also do not change from one interval I_n to another, with respect to the beginning of each interval. The received signal is then a complex valued, noisy semi-periodic train of pulses:

$$y(t)=\sum_{n=1}^{N}\sum_{k=1}^{K}a_k[n]\,g(t-\tau_k-nT)+\eta(t), \qquad (1)$$

where $a_k[n]=\alpha_k x_M[n]$ and $N=T_0/T$. The signal is defined by K time delays τ_k and by NK amplitudes $a_k[n]$. However, any segment of $x(t)$ of length T is defined by only K time delays and K amplitudes. Moreover, with knowledge of the sequence $\{x_M[n]\}$, the whole observed signal $x(t)$, of length T_0, is defined by K time delays τ_k and K amplitudes α_k. Thus signal $x(t)$ is FRI.

The received signal is affected by complex valued additive white Gaussian noise (AWGN), $\eta(t)$, of power spectral density (PSD) $\bar{\Phi}_\eta$. For later use, let $\sigma_\eta^2=\bar{\Phi}_\eta/T$ denote the power of the noise after passing through an ideal low pass filter of cut-off frequency $1/2T$.

The system model for the problem considered is depicted in Fig. 1a. The continuous-time signal $y(t)$ is passed through a *Sampling* block to obtain the set of samples $\boldsymbol{c}$. Based on samples $\boldsymbol{c}$, an estimate of $\boldsymbol{\theta}$ is obtained with the *Estimation* block. For the *Sampling* block a multichannel structure is considered. With a multichannel sampling setup the signal $y(t)$ is convolved with P different functions $s_1^*(-t)$, …, $s_P^*(-t)$, and the output of each channel is sampled at a rate $1/T$, [2, 3]. The set of samples is given by

$$c_p[n]=\int_{-\infty}^{\infty}y(t)s_p^*(t-nT)dt\ ,\ \ p=1,\dots,P. \qquad (2)$$

The system is said to have a total sampling rate P/T.

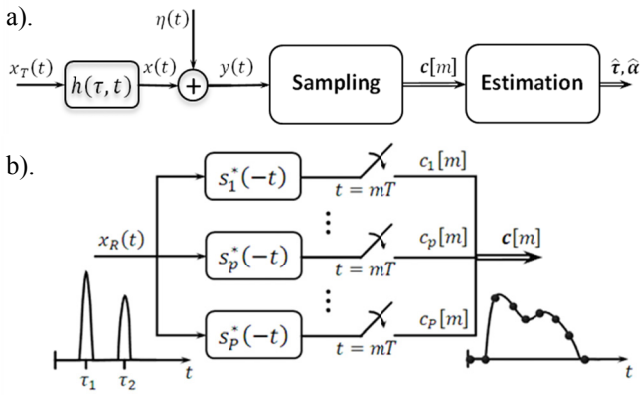


Fig. 1. a). System model; b). Filter-bank sampling block.

Two main multichannel sampling methods were proposed recently in the literature, i.e., the filter-bank method in [2] and the modulator-bank method in [3]. The modulator-bank sampling scheme can be used to treat different FRI signals under the assumption that the pulse $g(t)$ is compactly supported and pulses do not overlap boundaries of interval I_n. In contrast, the filter-bank sampling scheme can accommodate arbitrary pulse shapes $g(t)$, including infinite-length functions. Also, the filter-bank scheme is well suited for sampling long ($N\to\infty$) semi-periodic signals of form (1). For the rest of the paper, the filter-bank sampling is assumed, with the note that similar analysis can be carried out for the modulator-bank scheme.

The filter-bank scheme is illustrated in Fig. 1b. On each branch, the signal is filtered by $s_p^*(-t)$, followed by actual sampling at a rate $1/T$. The signal model is handled easier in the frequency domain. For long observation intervals, $T_0\to\infty$, the discrete-time Fourier transform (DTFT), $C_p(e^{j\omega T})\triangleq\sum_{n\in\mathbb{Z}}c_p[n]e^{-j\omega nT}$, can be applied. With this, (2) becomes

$$C_p(e^{j\omega T})=\frac{1}{T}\sum_{m=-\infty}^{\infty}Y\left(\omega-\frac{2\pi}{T}m\right)S_p^*\left(\omega-\frac{2\pi}{T}m\right), \qquad (3)$$

where $Y(\omega)=X(\omega)+\mathcal{N}(\omega)$ and $X(\omega)$, $Y(\omega)$, $\mathcal{N}(\omega)$, and $S_p^*(\omega)$ denote the Fourier transforms of $x(t)$, $y(t)$, noise $\eta(t)$, and sampling filters $s_p^*(-t)$, respectively. Note that, based on (1), $X(\omega)=G(\omega)\sum_{k=1}^{K}A_k(e^{j\omega T})\,e^{-j\omega\tau_k}$, where $G(\omega)$ is the Fourier transform of $g(t)$.

Focusing, as proposed in [2], on sampling filters with finite support in the frequency domain, e.g., contained in the range $\mathcal{F}=[-P\pi/T\ ,\ P\pi/T]$, (3) becomes

$$C_p(e^{j\omega T})=\frac{1}{T}\sum_{q=1}^{P}S_p^*(\omega+\omega_q)Y(\omega+\omega_q), \qquad (4)$$

where $\omega_q=2\pi(q-1-P/2)/T$. Note that all the

expressions in the frequency domain are $2\pi/T$ periodic and ω is restricted to the interval $[0, 2\pi/T)$.

If $\boldsymbol{c}$ is a column vector collecting $C_p(e^{j\omega T})$ from all the P sampling channels,

$$\boldsymbol{c} = \boldsymbol{S}\boldsymbol{y} = \boldsymbol{S}\boldsymbol{G}\boldsymbol{N}(\boldsymbol{\tau})\boldsymbol{D}\boldsymbol{a} + \boldsymbol{S}\boldsymbol{\eta}, \tag{5}$$

where $\boldsymbol{S}$ is a $P \times P$ matrix of elements $\frac{1}{T}S_p^*(\omega + \omega_q)$, $\boldsymbol{G} = \text{diag}\{G(\omega + \omega_1), \dots G(\omega + \omega_P)\}$, $\boldsymbol{N}(\boldsymbol{\tau})$ is a $P \times K$ matrix of elements $e^{-j\omega_q\tau_k}$, $\boldsymbol{D} = \text{diag}\{e^{-j\omega\tau_1}, \dots e^{-j\omega\tau_K}\}$, $\boldsymbol{a}^T = [A_1(e^{j\omega T}), \dots A_K(e^{j\omega T})]$, $A_k(e^{j\omega T})$ is the DTFT of $a_k[n]$, $\boldsymbol{y}$ is a column vector collecting elements $Y(\omega + \omega_q)$, and $\boldsymbol{\eta}$ is a column vector collecting elements $\mathcal{N}(\omega + \omega_q)$. Although it not explicitly evident in the notation, note the dependence on $e^{j\omega T}$ of the all the arrays in (5) except $\boldsymbol{N}(\boldsymbol{\tau})$. Also note that the unknown time delays τ_K are embedded in $\boldsymbol{N}(\boldsymbol{\tau})$ and $\boldsymbol{D}$, while the unknown amplitudes α_K are in the amplitudes $A_k(e^{j\omega T})$. Matrices $\boldsymbol{S}$ and $\boldsymbol{G}$ are known.

We briefly review the operations performed within the *Estimation* block, as proposed in. First, (5) is normalized by $\boldsymbol{W}^{-1}$, with $\boldsymbol{W} = \boldsymbol{S}\boldsymbol{G}$. Thus, let $\boldsymbol{\mu} = \boldsymbol{W}^{-1}\boldsymbol{c} = \boldsymbol{N}(\boldsymbol{\tau})\boldsymbol{D}\boldsymbol{a} + \boldsymbol{W}^{-1}\boldsymbol{\eta}$. By denoting $\boldsymbol{b} = \boldsymbol{D}\boldsymbol{a}$ and $\overline{\boldsymbol{\eta}} = \boldsymbol{W}^{-1}\boldsymbol{\eta}$ (dependence on $e^{j\omega T}$ continues to be implicit) and taking the inverse DTFT,

$$\boldsymbol{\mu}[n] = \boldsymbol{N}(\boldsymbol{\tau})\boldsymbol{b}[n] + \overline{\boldsymbol{\eta}}[n]. \tag{6}$$

Matrix $\boldsymbol{N}(\boldsymbol{\tau})$ has a Vandermonde structure and thus super-resolution techniques, [5, 6], traditionally used in spectral estimation, can be employed with (6) to estimate the time delays $\boldsymbol{\tau}$, where the number of multipaths K is a priori known, [2]. Once $\boldsymbol{\tau}$ is known, the vector $\boldsymbol{a}$ can be found using the linear relation $\boldsymbol{a} = \boldsymbol{D}^{-1}\boldsymbol{N}^{\dagger}(\boldsymbol{\tau})\boldsymbol{\mu}$, where is the Moore-Penrose pseudo-inverse of $\boldsymbol{N}(\boldsymbol{\tau})$. Further, with knowledge of the sequence $\{x_M[n]\}$, the amplitudes α_k can be determined.

In the absence of noise $\boldsymbol{\eta}$, the filter-bank sampling scheme supports recovery of the signal $x(t)$ only if the number of samples P per interval T is at least $2K$. The number of samples cannot exceed the number at Nyquist rate, $2T\mathcal{B}_g$, where $\mathcal{B}_g$ is the single side bandwidth of the pulse shape (if band-limited), [2, 7]. Combined, these lead to condition $2K \le P \le 2T\mathcal{B}_g$ for the filter-bank scheme to work.

Also note that the values of $\boldsymbol{\mu}[n]$ are obtained in (6) by taking the inverse DTFT of $\boldsymbol{W}^{-1}\boldsymbol{c}$. Thus, the matrix $\boldsymbol{W}$ needs to be stable invertible. One example of sampling filters satisfying this condition is

$$S_p(\omega) = \begin{cases} T, \text{for } \omega \in [\omega_p,\ \omega_p + 2\pi/T] \\ 0, \text{otherwise.} \end{cases} \tag{7}$$

3. CRAMER-RAO BOUND

The performance of an estimate $\widehat{\boldsymbol{\theta}}$ of a vector parameter $\boldsymbol{\theta}$ from a set of measurements is typically measured by its mean squared error (MSE). A lower bound on the MSE is often used instead to gain insight into the factors affecting the MSE. The Cramér-Rao bound (CRB) is a lower bound on the MSE of any unbiased estimate of $\boldsymbol{\theta}$. For the problem at hand, the available measurements are the samples $\boldsymbol{c}$, obtained as (5) by a filter-bank low rate sampling scheme applied to the FRI signal $y(t)$ of (1). Since $\{x_M[n]\}$ in (1) is assumed known, the unknown parameter vector is $\boldsymbol{\theta} = [\tau_1, \dots, \tau_K, \alpha_1^{\text{R}}, \dots, \alpha_K^{\text{R}}, \alpha_1^{\text{I}}, \dots, \alpha_K^{\text{I}}]^{\text{T}}$, where $\alpha_i^{\text{R}} = \text{Re}\{\alpha_i\}$ is the real part of α_i and α_i^{I} is the imaginary part of α_i. The CRB on the MSE of an element θ_i of $\boldsymbol{\theta}$ is given by the i-i element of a matrix $\boldsymbol{C}(\boldsymbol{\theta})$, i.e., $\text{CRB}(\theta_i) = [\boldsymbol{C}(\boldsymbol{\theta})]_{i,i}$. By definition, the matrix $\boldsymbol{C}(\boldsymbol{\theta})$ is the inverse of the Fisher Information Matrix (FIM), whose elements are given by [8],

$$[\boldsymbol{J}(\boldsymbol{\theta})]_{i,j} = \mathbb{E}\left\{\frac{\partial^2 \ln f(\boldsymbol{c};\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j}\right\}, \tag{8}$$

where $f(\boldsymbol{c};\boldsymbol{\theta})$ is the probability density function (pdf) of the sample vector $\boldsymbol{c}$, parameterized by $\boldsymbol{\theta}$.

The CRB for the signal reconstruction from low rate samples was discussed in [4]. In contrast, we are concerned here with parameter estimation, rather than estimation of the signal itself. Insight into the performance is developed from closed-form results for a particular setting, as stipulated in Theorem 1 below.

We make the following two assumptions:

(A1) The pulse shape shape $g(t)$ is ideal, in the sense that $G(\omega) = 1$ for $\omega \in [-2\pi\mathcal{B}_g,\ 2\pi\mathcal{B}_g]$ and $G(\omega) = 0$ everywhere else.

(A2) $\Phi_x(\omega) = \left|x(e^{j\omega T})\right|^2$, where $x(e^{j\omega T})$ is the DTFT of the sequence $\{x[n]\}$, and $\Phi_x(\omega)$ is assumed constant and known within the frequency range of interest. With this, we define the signal to noise ratio (SNR) as $\Phi_x(\omega)/\sigma_\eta^2$. Note that if $\{x[n]\}_{n=1}^N$ is a long bipolar sequence, where the $+1$ and -1 symbols are equiprobable, it can be further shown that $\Phi_x = N$, [9].

Theorem 1. Let $\boldsymbol{c}$ in (5) consist of the low rate samples of a semi-periodic stream of pulses in AWGN noise, of form (1). Under assumptions (A1) and (A2), with the choice (7) of the sampling filters, and for a multipath free environment, i.e., $K = 1$, the CRBs for delay and amplitude estimation from low rate samples are:

$$\text{CRB}(\tau_1) = \frac{3}{2\pi^2}\frac{1}{SNR|\alpha_1|^2}\frac{T^2}{P^3}, \tag{9}$$

$$\text{CRB}(\alpha_1) = \frac{1}{SNR}\frac{1}{P}. \tag{10}$$

Proof: See Appendix for a sketch of the proof.

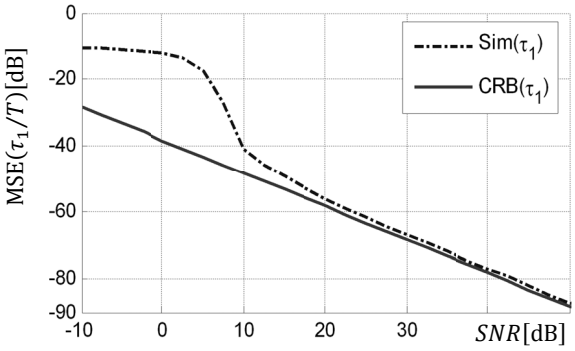


Fig. 2. Accuracy of delay estimation versus SNR.

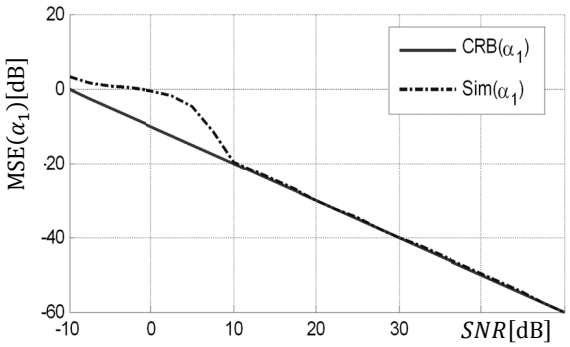


Fig. 3. Accuracy of amplitude estimation versus SNR.

Discussion. The CRB expression (9) shows that the performance of delay estimation improves with the SNR, i.e., the CRB decreases as the SNR increases. The same observation holds for the CRB of amplitude estimation (10). Interestingly, the CRB of delay estimation decreases with the cube of the number of sampling filters, i.e., it is proportional to P^{-3}. This means that increasing the number of sampling filters quickly improves the performance of delay estimation. In contrast, the CRB of amplitude estimation is proportional only to P^{-1}. If we denote $B_P = P/T$, and employ the assumption that $|\alpha_1|^2 = 1$, the CRB can be written as $(3/2\pi^2 B_P^2)(1/B_P\, T\, SNR)$ for delay estimation and $1/B_P T\, SNR$ for amplitude estimation. With these definitions, the expression for the CRB has the same form as the CRB for delay estimation in continuous signals developed in [10], where the term $B_P T\, SNR$ is referred to as postintegration SNR. This is justified by the fact that the product $B_P T$ is the processing gain due to the time bandwidth product of the signal. Thus we observe that the variation with P^3 of the CRB of delay estimation can be split into the effect of bandwidth $P^2$and the effect of processing gain P. In the case of amplitude estimation, the CRB is influenced only by the SNR gain.

4. NUMERICAL EXAMPLES

The MSE of delay estimation (CRB and maximum likelihood simulations) as a function of SNR is shown in Fig. 2 for $N = 100$ symbols, $P = 10$ sampling channels,

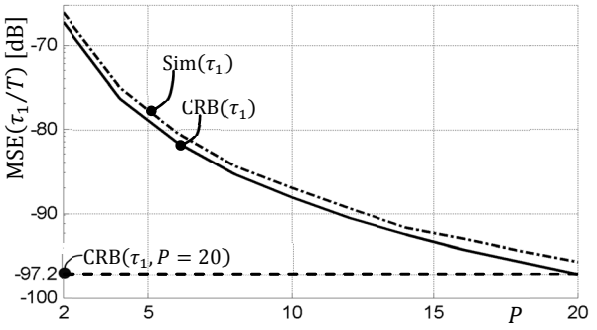


Fig. 4. Delay estimation errors as function of the number of sampling filters.

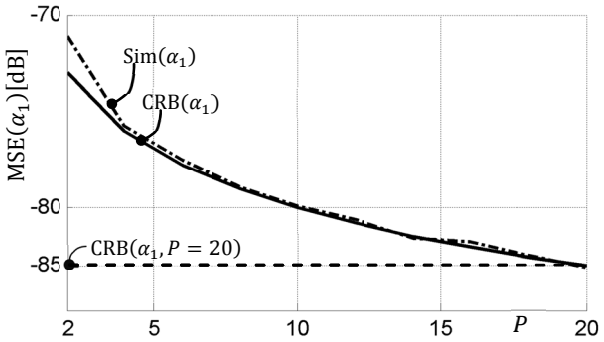


Fig. 5. Amplitude estimation errors as function of the number of sampling filters.

$\alpha_1 = 1$, and $T = 10\mu s$. Results shown are averaged over 1000 runs in which the noise have been chosen randomly (source location was kept fixed). It may be noticed that at high SNR, the MSE from simulations approaches asymptotically the CRB. This confirms that the CRB given by (9) is a tight lower bound. With the logarithmic scaling used in Fig. 2, the CRB decreases linearly with slope -3. In contrast, at low SNR, maximum likelihood estimation experiences a threshold effect, i.e., the MSE increases steeply. This is because, as the SNR decreases, the estimation enters a so called "large-errors" region, where the noise values become dominant over the signal. For the given FRI problem addressed in this work, time delay values are bounded by T andso are the errors, resulting in the plateau region evident at the low end of the SNR scale. In Fig. 3, it is observed that for the same setup as delay estimation, the MSE of amplitude estimation varies linearly with slope -1.

Based on (9) and (10), the performance of delay estimation improves with P^{-3} when P varies from $2K$ to $2TB_g$, while the amplitude estimation performance improves only with P^{-1}. When $P = 2TB_g$, the estimation performance of the filter-bank scheme equals the one of a single filter scheme working at Nyquist rate [1]. This can be observed in Fig. 4, and Fig. 5, where for low rate sampling the simulated MSE and CRB are plotted against the number of sampling filters P. The CRB for Nyquist rate for $TB_g = 10$is the same as the CRB for low rate sampling with $P = 20$.

5. CONCLUSIONS

Performance of time delay and amplitude estimation from low rate samples of pulse trains has been addressed. In particular, the CRB expression for the case of using a filter-bank sampling scheme was developed. With some simplifying assumptions, closed form expressions were determined for multipath free signal. The estimation performance in noise was shown to deteriorate significantly if the number of samples taken is the one indicated by the rate of innovation rather than by the Nyquist rate. Specifically, with the simplifying assumptions considered, the CRB is inverse proportional to the cube of the number of sampling filters for delay estimation and to the number of sampling filters for amplitude estimation.

6. APPENDIX

Due to space considerations, we can provide in this appendix only a sketch of the proof of Theorem 1. We start by expressing the pdf $f(\boldsymbol{c};\boldsymbol{\theta})$ used in (8) in terms of the signal model (1)

$$f(\boldsymbol{c};\boldsymbol{\theta}) = \frac{1}{\det\left(\pi \boldsymbol{K}_{\varepsilon_N}\right)} e^{-(\boldsymbol{c}_N-\tilde{\boldsymbol{\mu}}_N)^{\mathrm{H}}\boldsymbol{K}_{\varepsilon_N}^{-1}(\boldsymbol{c}_N-\tilde{\boldsymbol{\mu}}_N)}, \qquad (11)$$

where $\boldsymbol{c}_N^{\mathrm{T}} = [\boldsymbol{c}^{\mathrm{T}}[1], \dots, \boldsymbol{c}^{\mathrm{T}}[N]\,]$, $\tilde{\boldsymbol{\mu}}_N^{\mathrm{T}} = [\tilde{\boldsymbol{\mu}}^{\mathrm{T}}[1], \dots, \tilde{\boldsymbol{\mu}}^{\mathrm{T}}[N]\,]$, and the observation interval is limited to $[0, NT)$. The matrix $\boldsymbol{K}_{\varepsilon_N}$ denotes the $NP \times NP$ covariance matrix of the noise vector $\boldsymbol{\varepsilon}_N^{\mathrm{T}} = [\boldsymbol{\varepsilon}^{\mathrm{T}}[1], \dots, \boldsymbol{\varepsilon}^{\mathrm{T}}[N]\,]$. Equation (15.52) from [8] is applied to (11), together with the observation that the covariance matrix $\boldsymbol{K}_{\varepsilon_N}$ does not depend on parameter $\boldsymbol{\theta}$ [11], to reduce (8) to

$$[\boldsymbol{J}(\boldsymbol{\theta})]_{i,j} = 2\mathrm{Re}\left\{\frac{\partial \tilde{\boldsymbol{\mu}}_N^{\mathrm{H}}}{\partial \theta_i}\boldsymbol{K}_{\varepsilon_N}^{-1}\frac{\partial \tilde{\boldsymbol{\mu}}_N}{\partial \theta_j}\right\} = \qquad (12)$$

$$= \frac{T}{\pi}\int_{-\pi/T}^{\pi/T} \mathrm{Re}\left\{\frac{\partial(\boldsymbol{\mu}^{\mathrm{H}}\boldsymbol{W}^{\mathrm{H}})}{\partial \theta_i}\left(e^{j\omega T}\right)\boldsymbol{K}_{\varepsilon}^{-1}\frac{\partial(\boldsymbol{W}\boldsymbol{\mu})}{\partial \theta_j}\left(e^{j\omega T}\right)\right\}\mathrm{d}\omega = \qquad (13)$$

$$= \frac{T}{\pi}\int_{-\pi/T}^{\pi/T}\sum_{p=1}^{P}\frac{1}{\sigma_\eta^2}\mathrm{Re}\left\{\frac{\partial \mu_p^*}{\partial \theta_i}(e^{j\omega T})\frac{\partial \mu_p}{\partial \theta_i}(e^{j\omega T})\right\}\mathrm{d}\omega. \qquad (14)$$

Term (13) is obtained from (12) by replacing the data vector $\boldsymbol{c}_N$ by the vector of its Fourier coefficients (obtained by applying the DTFT to $\boldsymbol{c}_N$, when $N \to \infty$) and following a reasoning similar to that in [12]. Matrix $\boldsymbol{K}_\varepsilon$ is a $P \times P$ matrix of elements $[K_\varepsilon(e^{j\omega T})]_{p,q}$ given by the DTFT of the cross-correlation of sequences $\left[\varepsilon_p[1], \dots, \varepsilon_p[N]\right]$ and $\left[\varepsilon_q[1], \dots, \varepsilon_q[N]\right]$. With the choice (7) of sampling filters, $\boldsymbol{K}_\varepsilon$ is $\left(\sigma_\eta^2/T^2\right)\mathrm{diag}\{|S_1(\omega+\omega_1)|^2, \dots |S_P(\omega+\omega_P)|^2\}$. Furthermore, the choice of an ideal pulse shape $g(t)$, in the sense that $G(\omega) = 1$ for $\omega \in \left[-\mathcal{B}_g, \mathcal{B}_g\right]$ and $G(\omega) = 0$ everywhere else, determines that matrix $\boldsymbol{W}(e^{j\omega T})$ is an $P \times P$ identity matrix. This leads to further simplification of (13) to (14).

In (14) we use $\boldsymbol{\theta} = [\tau_1, \dots, \tau_K, \alpha_1^{\mathrm{R}}, \dots, \alpha_K^{\mathrm{R}}, \alpha_1^{\mathrm{I}}, \dots, \alpha_K^{\mathrm{I}}]^{\mathrm{T}}$ and $\mu_p(e^{j\omega T}) = x(e^{j\omega T})\sum_{k=1}^{K}\alpha_k e^{-j(\omega+\omega_p)\tau_k}$ to determine the elements of the FIM. Particularizing it for $K = 1$,

$$\boldsymbol{J}(\boldsymbol{\theta}) = \frac{2P\Phi_x}{\sigma_\eta^2}\left[\begin{array}{c|cc} P^2\pi^2|\alpha_1|^2/3T^2 & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]. \qquad (15)$$

With (15), the equations of Theorem 1 can be easily derived. Moreover, it can be shown that for $K \geq 2$, for well separated multipaths, the CRB for the time delay and amplitude of any k^{th} component is also given by Theorem 1.

7. REFFERENCES

[1] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse Sampling of Signal Innovations," *IEEE Signal Processing Magazine*, vol. 25, pp. 31-40, 2008.

[2] K. Gedalyahu and Y. C. Eldar, "Time-Delay Estimation From Low-Rate Samples: A Union of Subspaces Approach," *IEEE Transactions on Signal Processing*, vol. 58, pp. 3017-3031, 2010.

[3] K. Gedalyahu, R. Tur, and Y. C. Eldar, "Multichannel Sampling of Pulse Streams at the Rate of Innovation," *IEEE Transactions on Signal Processing*, vol. 59, pp. 1491-1504, 2011.

[4] Z. Ben-Haim, T. Michaeli, and Y. C. Eldar, "Performance Bounds and Design Criteria for Estimating Finite Rate of Innovation Signals," *Arxiv preprint arXiv:1009.2221*, 2010.

[5] X. Li and K. Pahlavan, "Super-Resolution TOA Estimation with Diversity for Indoor Geolocation," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 224-234, 2004.

[6] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, pp. 720-741, 1989.

[7] W. U. Bajwa, K. Gedalyahu, and Y. C. Eldar, "Identification of Parametric Underspread Linear Systems and Super-Resolution Radar," *IEEE Transactions on Signal Processing*, vol. 59, pp. 2548-2561, 2011.

[8] S. M. Kay, *Fundamentals of Statistical Signal Processing*: *Estimation theory*: Prentice-Hall PTR, 1993.

[9] L. W. Couch II, *Digital and analog communication systems*: Prentice Hall PTR, 2000.

[10] A. Weiss and E. Weinstein, "Fundamental limitations in passive time delay estimation--Part I: Narrow-band systems," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, pp. 472-486, 1983.

[11] M. L. Fowler and X. Hu, "Signal models for TDOA/FDOA estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, pp. 1543-1550, 2008.

[12] A. Zeira and A. Nehorai, "Frequency domain Cramer-Rao bound for Gaussian processes," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, pp. 1063-1066, 1990.

SOURCE LOCALIZATION USING TIME DIFFERENCE OF ARRIVAL WITHIN A SPARSE REPRESENTATION FRAMEWORK

Ciprian R. Comsa[*], *Alexander M. Haimovich*[*], *Stuart Schwartz*[†], *York Dobyns*[†], *and Jason A. Dabin*[‡]

[*]CWCSPR Lab, ECE Dept, New Jersey Institute of Technology, Newark, NJ, USA
[†]EE Dept, Princeton University, Princeton, NJ, USA
[‡]U.S. Army Communications-Electronics Research, Development and Engineering Center Fort Monmouth, NJ, USA
email: ccomsa@ieee.org, web: www.ccspr.njit.edu

ABSTRACT

The problem addressed is source localization via time-difference-of-arrival estimation in a multipath channel. Solving this localization problem typically implies cross-correlating the noisy signals received at pairs of sensors deployed within reception range of the source. Correlation-based localization is severely degraded by the presence of multipath. The proposed method exploits the sparsity of the multipath channel for estimation of the line-of-sight component. The time-delay estimation problem is formulated as an $\ell 1$-regularization problem, where the $\ell 1$-norm is used as a channel sparsity constraint. The proposed method requires knowledge of the pulse shape of the transmitted signal, but it is blind in the sense that information on the specific transmitted symbols is not required at the sensors. Simulation results show that the proposed method delivers higher accuracy and robustness to noise compared to conventional or even super-resolution MUSIC time-difference-of-arrival source localization methods.

Index Terms — Source localization, time-difference-of-arrival, sparse multipath channel, $\ell 1$-regularization

1. INTRODUCTION

Accurate localization of a signal source is a problem of interest in various applications, [1]. The current work addresses the problem of localizing a signal source by sensors distributed over an area. Typically, the source location is estimated in two stages. During the first stage, a measure of the received signal, usually the propagation time-delay, is estimated at each sensor. In the second stage, the actual location is computed from the time delay estimates. Time-delay-estimation (TDE) becomes challenging in multipath propagation environments, where the line-of-sight (LOS) signal component becomes obscured by multipath reflections. Hence, accurate localization requires techniques capable of resolving the LOS signal component. When the transmitted signal and its transmission time are known at a sensor, the time of arrival (TOA) can be estimated by a variety of techniques. A classical method is to estimate the TOA from the timing of the peak of the cross-correlation (CC) between the transmitted and received signals, [2]. The resolution of the TOA estimated in this case is limited by the width of the main lobe of the time autocorrelation function of the transmitted signal. This limitation makes the method unable to distinguish between the LOS signal and a reflected component when they are spaced closer than the resolution limit. Over the years, various techniques have been proposed to overcome this limitation. An example is the root-MUSIC method, belonging to the larger class of subspace methods, also referred to as super-resolution methods due to their high resolution capabilities, [3].

Recently, some potentially even higher resolution estimation techniques have been proposed, based on the observation that many propagation channels associated with multipath environments tend to exhibit a sparse structure in the time domain, i.e., the number of multipaths is much smaller than the number of samples of the received signal. This sparsity has been exploited in TOA estimation, [4], and other TOA-related applications, such as compressed channel sensing, [5, 6], or underwater acoustic channel deconvolution, [7]. TOA estimation requires the transmitted signal to be known to the sensors. In many applications, the source may be non-cooperative or otherwise the signal and timing information may not be available at the receiving sensors. The common approach for such a case is to take one of the sensors as reference and measure the time-difference-of-arrival (TDOA) at each of the other sensors with respect to the chosen reference sensor. A method for TDOA estimation for sparse non-negative acoustic channels is presented in [8].

In this paper, a method for high resolution TDOA estimation for complex-valued sparse multipath channels is developed and applied to source localization. The proposed method casts the TDOA estimation as a convex optimization problem that can be efficiently solved by conventional algorithms, [9]. In particular, the problem is formulated as an $\ell 1$-regularization problem, i.e., the $\ell 1$-norm is used to impose a sparsity-constraint on the channel. While the proposed approach does not require the transmitted signal to be known at the sensors side, as is the case in [4-7], the pulse shape is assumed known. Also, for simplicity, the reference sensor is considered single-path, i.e., the reference sensor receives only the LOS signal. In [10], TOA estimation is carried out also based on knowledge of only the pulse shape. Unlike our approach, where sparsity is exploited to improve localization performance, in [10] the sparsity of the channel is applied to reduce the sampling rate of the received analog signal. The actual time delay estimation is performed by a subspace method.

The remainder of this paper is organized as follows. Sec. 2 introduces the signal model. In Sec. 3, the $\ell 1$-regularization method for TDOA estimation is proposed and discussed. In Sec. 4, numerical simulations are conducted to compare the performance of our method with other techniques for source localization. Conclusions are listed in Sec. 5.

2. SIGNAL MODEL

With the localization problem, the unknown x-y location, ξ_0, of a signal source has to be estimated based on the signals collected by a number M of sensors. The source is assumed to transmit an unknown lowpass signal, $s(t)$, of bandwidth B. The receiving sensors are widely dispersed within a surveillance area, at arbitrary but precisely known locations, ξ_k. The signal received at any sensor is expressed as the convolution between the transmitted signal, $s(t)$, and the channel impulse response (CIR), $h_k(t)$:

$$r_k(t) = (s * h_k)(t) + w_k(t), \tag{1}$$

where $w_k(t)$ is additive white Gaussian noise (AWGN), with variance σ_k^2. The multipath channel is modeled

$$h_k(t) = \sum_{p=1}^{P_k} \alpha_{kp}\delta\big(t - \tau_{kp}\big), \tag{2}$$

where $\delta(\cdot)$ denotes the delta function, P_k is the number of paths of the channel observed at sensor k, and α_{kp} is the complex valued channel gain. The channel parameters P_k and α_{kp} are unknown to the sensors.

The localization method proposed here is based on the estimation of the TDOA at pairs of sensors, $\Delta\tau_{kl}(\xi_0) = \tau_{k0}(\xi_0) - \tau_{l0}(\xi_0)$. The LOS propagation delay between the source located at ξ_0 and any sensor at ξ_k, $\tau_{k0}(\xi_0)$, is proportional to the source-to-sensor distance: $\tau_{k0}(\xi_0) = (1/c)\sqrt{(x_k - x_0)^2 + (y_k - y_0)^2}$, where c is the speed of light. A TDOA measurement localizes the source on a hyperboloid with a constant range difference between the two sensors, k and l. Since the source can occupy only one point on the hyperbolic curve, TDOA measurements from the other sensors are used to resolve the location ambiguity. Processing is carried out at a fusion center, assumed linked via ideal communication links to the sensors. Ideal time synchronization is assumed between the sensors. One of the sensors, say $l = 1$, is chosen as reference such that the sensor pairs used for TDOA estimation are $\{k, 1\}$, for $k = 2, \dots, M$.

3. TDOA ESTIMATION FOR SPARSE CHANNELS

The focus of this paper is on improving the TDOA resolution and accuracy and, implicitly, the source localization resolution and accuracy. Recent work has shown that channel estimation can be improved through sparsity regularization, [4-7]. In this section, we propose an $\ell 1$-regularization method for TDOA estimation, exploiting the sparsity of multipath channels.

3.1. $\ell 1$-regularization method

Assuming for simplicity of presentation that the time-delays of the CIR are integer multiples of the sampling rate, define the received signal vector $\boldsymbol{r}_k = [r_k(1), \dots, r_k(Q + L - 1)]^T$, the CIR vector $\boldsymbol{h}_k = [h_k(1), \dots, h_k(L)]^T$, and the noise vector $\boldsymbol{w}_k = [w_k(1), \dots, w_k(Q + L - 1)]^T$, where $Q + L - 1$, L, and Q are the lengths of the received signal vector, channel and transmitted signal vector, respectively. With these definitions, the signal model (1) can be written

$$\boldsymbol{r}_k = \boldsymbol{S}\boldsymbol{h}_k + \boldsymbol{w}_k, \tag{3}$$

where $\boldsymbol{S}$ is the $(Q + L - 1) \times L$ matrix relating the received signal vectors $\boldsymbol{r}_k$ to the channel vectors $\boldsymbol{h}_k$. Since typically $L \gg P_k$, the CIR, $\boldsymbol{h}_k$, is a sparse vector. Sparsity of the CIR vector can be enforced by minimizing its $\ell 0$-norm, i.e., the number of non-zero elements. Minimization of the $\ell 0$-norm of $\boldsymbol{h}_k$ is a non-convex optimization problem and it is NP-hard, which means that no known algorithm for solving this problem is significantly more efficient than an exhaustive search over all subsets of entries of $\boldsymbol{h}_k$. In lieu of the $\ell 0$-norm, an approximation, e.g., the ℓq-norm, can be used with $0 < q \leq 1$. While smaller q implies better approximation of the $\ell 0$-norm, $q = 1$ is often used because minimization of the $\ell 1$-norm is a convex problem, and it can be efficiently solved by standard algorithms. Thus, assuming that the transmitted signal, and hence the matrix $\boldsymbol{S}$, are known, the CIR estimation can be formulated as an $\ell 1$-regularization problem [7],

$$\min_{\boldsymbol{h}_k} \ \|\boldsymbol{r}_k - \boldsymbol{S}\boldsymbol{h}_k\|_2^2 + \lambda_k\|\boldsymbol{h}_k\|_1, \tag{4}$$

where $\|\boldsymbol{v}\|_q = \sqrt[q]{\sum_i |v_i|^q}$ denotes the ℓq-norm of vector $\boldsymbol{v}$.

The estimate of the CIR can be used to find the TOA as the timing of the earliest peak of the CIR. We now seek to formulate the problem of TDOA estimation. The TDOA has to be determined from a sufficient statistic involving signals received at two sensors. A common such statistic is cross-correlation of the received signals, implying that the TDOA has to be estimated from the cross-correlation of the CIR of two channels, e.g., from $h_{kl}(t)$. For a single channel, the TOA is determined as the time of the first path of the estimated channel. However, when cross-correlating two CIR's, the time of the first path in the cross-correlation does not necessarily correspond to the TDOA. Assuming that for each of the channels, the line-of-sight path is the strongest, the TDOA can be found from the time of the strongest component of the cross-correlation. Here, to simplify the situation, we assume that one of the sensors does not experience multipath, and use this sensor as reference for TDOA estimation. In this case, the TDOA is given by the delay of the first time component of $h_{kl}(t)$. Cross correlating the signal received at sensor k with the reference sensor l, and dropping the noise term for simplicity, we have:

$$\widetilde{\boldsymbol{y}}_{kl} = \boldsymbol{\Gamma}_S\boldsymbol{h}_{kl}, \tag{5}$$

where $\widetilde{\boldsymbol{y}}_{kl} = \boldsymbol{F}\boldsymbol{y}_{kl}$, $\boldsymbol{F}$ is the unitary discrete Fourier transform (DFT) matrix, and $\boldsymbol{y}_{kl}$ is the cross-correlation sequence of the received signal vectors $\boldsymbol{r}_k$ and $\boldsymbol{r}_l$. Let $N = 2Q + 2L - 3$ be the number of elements of the vector $\boldsymbol{y}_{kl}$. The matrix $\boldsymbol{\Gamma}_S$ is a $N \times N$ transformation matrix relating the frequency domain cross-correlations of the received signals, $\widetilde{\boldsymbol{y}}_{kl}$, to the time-domain cross-correlations of the channels, $\boldsymbol{h}_{kl}$. It can be verified that $\boldsymbol{\Gamma}_S = \mathrm{diag}\{\widetilde{\boldsymbol{u}}_S\}\,\boldsymbol{F}$, where $\widetilde{\boldsymbol{u}}_S$ is the power spectral density of the transmitted signal padded with $L - 1$ zeros.

The problem of TDOA estimation can be formulated as an $\ell 1$-regularization problem:

$$\min_{\boldsymbol{h}_{kl}} \ \|\widetilde{\boldsymbol{y}}_{kl} - \boldsymbol{\Gamma}_S\boldsymbol{h}_{kl}\|_2^2 + \lambda_{kl}\|\boldsymbol{h}_{kl}\|_1, \tag{6}$$

which may be efficiently solved with conventional convex optimization algorithms, [9]. Note the presence of the auto-correlation of the transmitted signal within the cost function. The proposed TDOA method utilizes auto-correlation information (for uncorrelated symbols, pulse shape information is sufficient), but the method is blind in the sense that it does not require knowledge of the transmitted symbols.

Formulating (3) with a denser sampled channel has the potential of a higher resolution TDOA estimate, but increases the complexity of the optimization algorithms. An iterative grid refinement approach is adopted to keep the complexity of the optimization algorithms in check. Initially, (6) is solved for the samples corresponding to a desired range of delays. A refined grid is obtained by taking a second set of samples focusing on a range of delays that are indicated by the first iteration to contain multipath. This corresponds to a higher sampling rate of the smaller area of interest. Samples of the second set are obtained by interpolating the original samples. The transformation matrix $\boldsymbol{\Gamma}_S$ is also recalculated to match the refined sample support of the correlation sequences. With the refined grid, (6) is solved again and a new TDOA estimate, of higher resolution, is obtained. The grid refinement procedure can be repeated until a desired resolution is attained.

3.2. Discussion

The cost function to be minimized in $\ell 1$-regularization problems, e.g., (4) and (6), has two terms: the first term is a *measurement fidelity* (or reconstruction error); the second term is a *regularization* (or penalization) *term*, that imposes sparsity on the estimate by using its $\ell 1$-norm. The factor λ is a *regularization parameter*. The sparsity of the solution is governed by the choice of λ, which balances the fit of the solution to the measurements versus sparsity, [11]. A small regularization parameter corresponds to a good fit to the measurements, while too much regularization (over-penalization through a large λ) produces sparser results, but may fail to explain the measurements well. A number of methods have been studied in the literature for automated *choice of* λ (see [12] and the references therein). However, in practice an optimal value of λ is difficult to select by any of these methods, and usually the choice of λ resorts to semi-empirical means, [11].

It is known that super-resolution methods, such as root-MUSIC, have the capability of asymptotically achieving optimal performance. However, in practice, with limited number of samples, or with highly correlated signal components, the accuracy performance often degrades away from the theoretical lower bounds, due to *resolution* limitations, [13]. In recent works, [7, 14], it was found that the $\ell 1$-regularization method may offer higher resolution than the super-resolution methods. Moreover, the sparse regularization has the advantage of producing good accuracy even at low signal-to-noise ratios (SNR), i.e., it exhibits good *robustness to noise*, as it has been noted in [14]. In fact it has been proved (see [15] and the references therein) that there is a fundamental connection between *robustness and sparsity*. Specifically, if some disturbance is allowed into the transformation matrix $\boldsymbol{\Gamma}_S$ or the measurements vector $\boldsymbol{y}_{kl}$, finding the optimal solution in the worst case sense is equivalent to solving the problem in the $\ell 1$-regularization formulation, which imposes sparse solutions.

The $\ell 1$-regularization continuously shrinks the estimate elements toward 0 as λ increases, leading to sparse solutions. However, the $\ell 1$-regularization shrinkage results in a small bias in the non-zero elements of the estimate, since the estimation of these elements is based on the measurement fidelity term, [16]. Thus solving the problem of estimating $\boldsymbol{h}_{kl}$ from the measurements $\boldsymbol{y}_{kl}$, (5), by employing an $\ell 1$-regularization formulation, (6), may lead to a sub-optimal solution for the non-zero elements of the estimate. However, despite this downside, with a reasonable choice of λ, the $\ell 1$-regularization method still produces better TDOA estimation (and hence source localization) accuracy than conventional techniques, especially at low SNR, as demonstrated by the results presented in the next section. Moreover, the proposed method doesn't necessary require knowledge of the number of the multipath components, as root-MUSIC does.

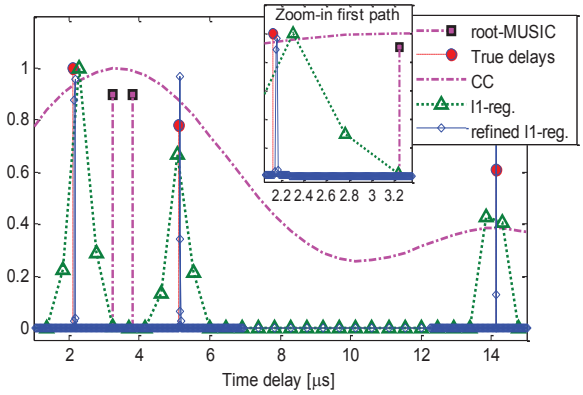


Fig. 1. True and estimated multipath components.

When the power spectral density of the transmitted signal is flat across the frequencies of interest, $\boldsymbol{\Gamma}_S$ in (5) has the form of a DFT matrix. In this case, the sparse estimate can be found with fewer equations than in (6), reducing the required computational effort. A procedure for selecting a subset of equations among those in (6) and the sufficient number of equations in the subset to ensure that the solution is not altered, can be found in [17].

4. SIMULATION RESULTS

In this section, we present some simulation results to demonstrate the performance of the proposed $\ell 1$-regularization method. For a typical setup, we show significant improvement in localization accuracy, compared to conventional methods, such as cross-correlation and root-MUSIC. We considered a system in which a number $M = 8$ of sensors are approximately uniformly distributed around the source, on an approximately circular shape of radius 1 km. The source, which location is to be estimated, transmits a Gaussian Minimum Shift Keying (GMSK) modulated signal of bandwidth $B = 200$ kHz that is received by the M sensors through different multipath channels. For each sensor $k = 2, \ldots M$, the TDOAs are measured relative to the chosen reference sensor, $l = 1$. The pulse shape is known at the sensors side and used to generate the auto-correlation of the transmitted signal. The wireless channels between the source and each of the sensors are modeled as (2). Specifically, a three-paths model is used, as in [4]. The first two paths are spaced well below the bandwidth resolution, while the separation between the second and the third path is higher, as it can be seen in Fig. 1. The only exception is the reference sensor assumed to be an AWGN channel, i.e., no multipath. The simulation scenario also employs the same noise level across sensors. For $\ell 1$-regularization, the regularization parameter λ is chosen by semi-empirical means, as motivated in Sec. 3.2. Originally, when solving problem (6), a number of 500 received symbols are sampled 8 times the Nyquist rate. An oversampling factor of $\times 50$ is used for grid refinement as explained in Sec. 3.1.

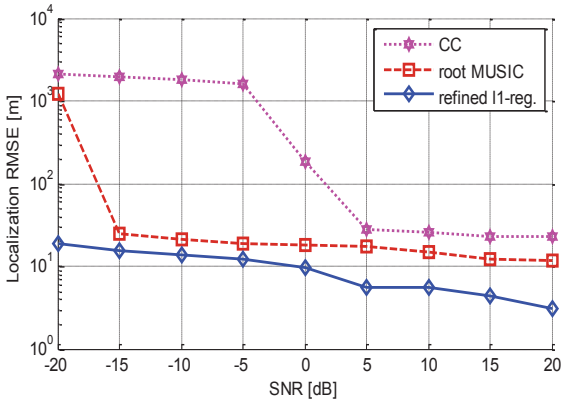


Fig. 2. Source localization accuracy in noise.

Fig. 1 shows the time-delays of the multipath components and their estimates at one of the sensors, for a received SNR of 15 dB per sample. The TDOA is given by the earliest of these components. The estimate by the $\ell 1$-regularization approach with grid refinement is the closest to the true delays, when compared to the other methods considered. The result of the $\ell 1$-regularization without grid refinement is visibly biased due to the limited sampling rate. The CC and the root-MUSIC estimates show significantly larger errors particularly for the two close paths. The TDOA estimates are transformed, by multiplication with the known signal propagation speed, into range difference information for constructing a set of hyperbolic equations. The solution provided by these equations is the estimated location of the source, relying on the knowledge of the sensors locations. In the literature, several methods can be found for solving the hyperbolic equations. We adopt here a recent approach that involves convex relaxation techniques supported by standard, efficient semidefinite programming (SDP) [18].

Fig. 2 illustrates the localization performance, in terms of root mean square error (RMSE) against SNR, of the proposed methods in the aforementioned scenario. The RMSE is obtained from Monte Carlo simulations with 100 runs per SNR value. The plot shows better accuracy of the proposed method over CC and root-MUSIC, at both high and low SNR. At high SNR, given the low separation between the first two multipaths, both CC and root-MUSIC provide biased estimates due to their limited resolution capabilities, though root-MUSIC is better. The reason for the better accuracy at low SNR is that $\ell 1$-regularization is robust to noise, as discussed in Sec. 3.2.

5. CONCLUSIONS

A method for source localization via TDOA estimation in multipath environments was developed. The sparsity of the channels is exploited and a grid refinement procedure was formulated to improve the resolution of the TDOA estimation. The proposed technique compares favorably to the conventional cross-correlation and root-MUSIC techniques, in terms of TDOA and source location accuracy estimation. For dense multipath environments the proposed method may succeed where conventional methods fail to resolve closely separated components. Therefore it is suitable for applications like source localization in multipath. Moreover, simulation results confirmed the noise robustness of the method.

6. REFERENCES

[1] J. F. Bull, "Wireless geolocation," *IEEE Vehicular Technology Magazine,* vol. 4, pp. 45-53, 2009.

[2] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech and Signal Processing,* vol. 24, pp. 320-327, 1976.

[3] X. Li and K. Pahlavan, "Super-resolution TOA estimation with diversity for indoor geolocation," *IEEE Transactions on Wireless Communications,* vol. 3, pp. 224-234, 2004.

[4] J. J. Fuchs, "Multipath time-delay detection and estimation," *IEEE Transactions on Signal Processing,* vol. 47, pp. 237-243, 1999.

[5] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proceedings of the IEEE,* vol. 98, pp. 1058-1076, 2010.

[6] C. R. Berger, Z.-H. Wang, J.-Z. Huang, and S. Zhou, "Application of compressive sensing to sparse channel estimation," *IEEE Communications Magazine,* vol. 48, no. 11, pp. 164–174, Nov. 2010.

[7] W.-J. Zeng, X. Jiang, X.-L. Li, and X.-D. Zhang, "Deconvolution of sparse underwater acoustic multipath channel with a large time-delay spread," *The Journal of the Acoustical Society of America,* vol. 127, pp. 909-919, 2010.

[8] Y. Lin, J. Chen, Y. Kim, and D. D. Lee, "Blind sparse-nonnegative (BSN) channel identification for acoustic time-difference-of-arrival estimation," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2007, pp. 106-109.

[9] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, ver. 1.21. http://cvxr.com/cvx, Feb. 2011.

[10] K. Gedalyahu and Y. C. Eldar, "Time-delay estimation from low-rate samples: A union of subspaces approach," *IEEE Transactions on Signal Processing,* vol. 58, pp. 3017-3031, 2010.

[11] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE, Special Issue on Applications of Sparse Representation and Compressive Sensing,* vol. 98, pp. 948-958, 2010.

[12] R. Giryes, M. Elad, and Y. C. Eldar, "The projected gsure for automatic parameter tuning in iterative shrinkage methods," *submitted to Applied and Computational Harmonic Analysis, 2010.*

[13] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transactions on Acoustics, Speech and Signal Processing,* vol. 37, pp. 720-741, 1989.

[14] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing,* vol. 53, pp. 3010-3022, 2005.

[15] X. Huan, C. Caramanis, and S. Mannor, "Robust regression and lasso," *IEEE Transactions on Information Theory,* vol. 56, pp. 3561-3574, 2010.

[16] H. Zou, "The adaptive Lasso and its oracle properties," *Journal of the American Statistical Association,* vol. 101, p. 12, 2006.

[17] J. Haupt, L. Applebaum, and R. Nowak, "On the restricted isometry of deterministically subsampled fourier matrices," in *The 44th Annual Conference on Information Sciences and Systems (CISS),* 2010, pp. 1-6.

[18] E. Xu, Z. Ding, and S. Dasgupta, "Robust and low complexity source localization in wireless sensor networks using time difference of arrival measurement," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2010, pp. 1-5.

Simulation Model for WLAN and GSM interoperating

Ciprian-Romeo COMŞA[1], Ion BOGDAN[2]
Telecommunications Department, Technical University "Gheorghe Asachi" Iaşi, Romania
[1] ccomsa@etc.tuiasi.ro, [2] bogdani@etc.tuiasi.ro

Abstract — **This paper presents a finite state machine model of a mobile station supporting simultaneously both GSM voice communications and data transmissions in an IEEE 802.11b WLAN network.**

I. INTRODUCTION

Increasing competitiveness in the wireless industry combined with the rapid pace at which technology is evolving challenges equipment vendors to find new and better ways to reduce product development costs. Also, the demand for a multitude of wireless applications leads to the need of having different technologies on the same mobile terminal.

Sharing internet access, transmitting voice over WLANs, managing manufacturing and inventory are only a few of the most common applications motivating the rapid growth of the WLAN technology. There is no doubt about the wide usage of the GSM voice services and the utility of mobile station supporting simultaneously both GSM voice communications and data transmissions in an IEEE 802.11 WLAN network. While terminals implementing WLAN and GSM on separated chips are already used, the power consumption, dimensions and costs are the main arguments for implementing both GSM and WLAN on the same chip.

Based on this idea, we developed a simulation model as a finite state machine for a mobile terminal supporting simultaneously both GSM voice communications and data transmissions in an IEEE 802.11b WLAN network and we integrated this model in a simple IEEE 802.11b WLAN network to monitor the dual WLAN and GSM station behavior.

In elaborating the model of the dual station we had to take into account the following statements: (1) The terminal cannot communicate (send or receive) simultaneously in both of the networks because it has only one transmitter and one receiver both connected to one single antenna. Thus, the simultaneous communication in both of the networks is only apparent to the user: in fact, it is interleaved in disjoint time slots. (2) Because GSM is a TDMA standard and the mobile terminal has to fulfill specific tasks (receive, transmit or measure) in time slots prescribed by network, unconditional priority is given to the GSM communication. There is however one exception, which is the Beacon receiving in the 802.11b network, which blocks, if necessary, any GSM operation.

The WLAN and GSM interoperation schematic is shown in Fig. 1 where the WLAN packets are placed between GSM slots.

II. SYSTEM MODEL

The 802.11 LAN is based on a cellular architecture, where the system is subdivided into cells. Each cell is controlled by a station called Access Point (AP). The functionality of a WLAN is assured by the Medium Access Control (MAC) layer, which can be constructed as a set of diverse services to accomplish information exchange, power control, synchronization and power control.

The *basic access mechanism*, called the Distributed Coordination Function (DCF) is basically a Carrier Sense Multiple Access with Collision Avoidance mechanism and it is used together with a Positive Acknowledge (ACK) scheme [1]. DCF works by a station willing to transmit data, senses the medium first. If the medium is busy, then the station defers its transmission to a later time, but if the medium is free for a specified time (called Distributed Inter Frame Space (DIFS)), the station transmits. The receiving station then checks the CRC of the received packet and sends an acknowledgement (ACK) packet. This receipt indicates to the transmitting station that there were no collisions detected. If the sender does not receive ACK, then it retransmits the last fragment [2], [3].

In order to reduce the probability of two mobile stations (STA) colliding because they cannot hear each other, the standard defines a *Virtual Carrier Sense mechanism*: a STA wanting to transmit a packet first transmits a short control packet called Request To Send (RTS), which includes the duration of the following transaction (that is the packet and the respective ACK), the destination STA responds (if the medium is free) with a response control packet called Clear to Send (CTS), which includes the same duration information [1]. All STAs receiving either the RTS and/or the CTS, set their Virtual Carrier Sense indicator (called Network Allocation Vector – NAV), for the given duration, and use this information together with the Physical Carrier Sense when sensing the medium [2].

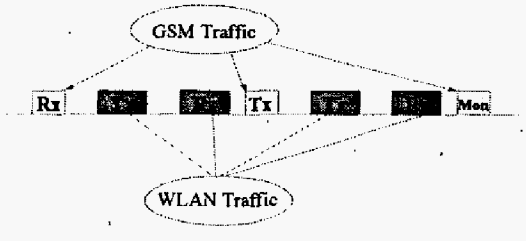


Figure 1. WLAN and GSM time interopperation

0-7803-9029-6/05/$20.00 ©2005 IEEE

The diagram from Fig. 2 shows how the RTS/CTS mechanism works for station *A* as a transmitter and station *B* as a receiver and the NAV settings for their neighbors [2].

Bit error rates on wireless systems (10^{-5}-10^{-6}) are substantially higher than wire-line systems (10^{-12}). Thus, large blocks may approach the number of bits where the probability of a block error occurring may be extremely high. To reduce the possibility of this happening, large blocks may be *fragmented* by the transmitter and *reassembled* by the receiver node. While there is some overhead in doing this – both the probability of an error occurring is reduced and, in the event of an error, the re-transmission time is also reduced [1].

802.11 Specifications require for a Data volume to be transmitted to be split in units denoted as MAC Service Data Unit (MSDU) and allow for every MSDU to be split in smaller units denoted as MAC Protocol Data Unit (MPDU) [5]. The transmission of an MSDU consists in successive transmissions of its MPDUs. Every MPDU transmission should be acknowledged by the receiving part. The Short Inter Frame Space (SIFS) time intervals separate the receiving of an MPDU and the transmission of its ACK as well as the receiving of an ACK and the transmission of the next MPDU. The unacknowledged MPDUs are retransmitted after an AckTimeOut waiting time.

In good propagation conditions the transmission of an MSDU is a continuous succession of the sequence MPDU–SIFS–ACK–SIFS. The transmission of a second MSDU is allowed only after applying the basic access procedure. This way a STA having a huge volume of data to send is not allowed to block the channel for a too long period of time.

There are defined some types of *inter frame spaces*, illustrated in Fig. 3, in order to provide different priorities to stations [3]: (1) Short Inter Frame Space (SIFS) is used to separate transmissions belonging to the single dialog (Fragment-ACK) and it is the minimum inter frame space. There is, at most, one single station to transmit at any given time, therefore giving it priority over all other stations. (2) Point Coordination IFS (PIFS) is used by the Access Point to gain an access over the medium before any other station. The value is SIFS plus one slot time. (3) Distributed IFS (DIFS) is the inter frame space used for a station willing to start a new transmission, which is calculated as PIFS plus one slot time.

Backoff is a well-know method used to resolve contention between different stations waiting to access the media [2]. This method requires each station to choose a random number between 0 and a given limit (CW – Contention Window), and wait this number of slot times before accessing the medium. The slot time is defined as a way a station will always be capable of determining if another station has accessed the medium at the beginning of the previous slot. It reduces the collision probability by half. Each station listens to the network, and the first station to finish its allocated number of slot times begins the transmission. If any other station hears the first station talk, it stops counting down its backoff timer. When the network is idle again, it resumes the countdown.

The 802.11 standard [5] defines an *exponential backoff algorithm* which must be executed in the following cases: when the station senses the medium before the first transmission of the packet, and the medium is busy; after each retransmission; and, after a successful transmission. The only case when this mechanism is not used, is when the station decides to transmit a new packet and the medium has been free for more than DIFS.

On a WLAN infrastructure, the stations are kept *synchronized* by all the STAs updating their clocks according to the AP's clock, using the following mechanism: the AP periodically transmits frames called Beacon frames. These frames contain the value of the AP's clock at the moment of transmission. The receiving STAs check the value of their clocks at the moment the signal is received, and correct it to keep in synchronization with the AP's clock. Beacon frames are intended to be transmitted at regular time intervals (BP – Beacon Period), but the transmission of a particular Beacon could be delayed if the medium is busy at the intended moment (TBTT – Target Beacon Transmission Time) [5].

In mobile applications, *battery power* is an important resource. This is the reason why the 802.11 standard directly addresses the issue of *power saving* (PS) and defines a mechanism which enables STAs to go into sleep mode for long periods of time without losing information, as exemplified in Fig. 4. The main idea behind the PS mechanism is that the AP maintains a continually updated record of every STA working in PS mode, and buffers the packets addressed to it until the STA either specifically request the packets by sending a polling request (PsPoll message) or it changes its operation mode.

As part of its Beacon frames, the AP also transmits information (TIM – Traffic Indication Map) about which PS STAs have MSDUs buffered at the AP. The STAs wake up in order to receive the Beacon frame. If there is an indication that there is an MSDU buffered at the AP waiting for delivery, then the STA stays awake and sends a PS Poll message to the AP to get this MSDU. Some Beacon frames, regularly distributed in time, are followed by multicast and broadcast messages. These Beacon frames are denoted as DTIMs (Delivery TIM).

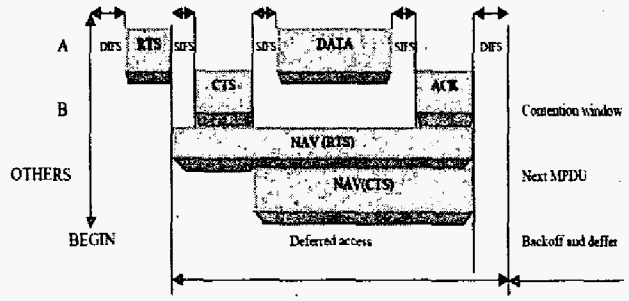


Figure 2. NAV state for the RTS/CTS mechanism

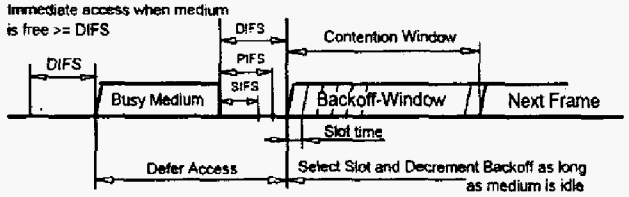


Figure 3. A schematic of the access medium

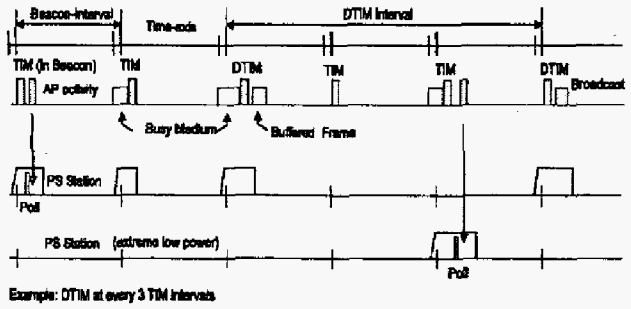


Figure 4. Power management of one STA

III. Simulation Model

The main state diagrams of the modeled Power-Saving Station (Ps_STA) are presented in the followings. The initial state of a normal Ps_STA is DOZE. There are two conditions for exiting (wake-up) this state:

A. First is if the STA has to wake up to listen to the network activity and there are two situations: (1) When approaching a TBTT scheduled for wake-up (according to STA Listen Interval): a waiting state (WaitBcBMData) is activated and the channel state is monitored. The STA transits to RECEIVE state any time the channel becomes Busy. When exiting this state the packet is checked for errors. If there are errors in receiving it the STA discards the packet and returns to the WaitBcBMData state in order to receive the next packet. When the received packet is valid (FCS=1) its type is established through successive conditional transitions. If the packet is Beacon, its TIM part is analyzed to check for possible data stored in the AP having the STA as the final destination. (2) When approaching a DTIM beacon: at DTIM moments, the STA continues to wait after the Beacon for the scheduled broadcast data.

In both cases, if the received packet is of type Data STA checks for its final destination discarding it if the STA is not the intended destination or interpreting it, otherwise. If the Data are intended for the STA it stores the information, waits for a SIFS time and sends an ACK. If the received packet is an ACK the STA waits for a SIFS time and sends the next packet (MPDU) if there are still Data to transmit (Dts=1) or returns to DOZE state (Dts=0). Dts is an internal parameter used to indicate the application has still data to be transmitted by the STA. The transition to DOZE state is postponed if there are still Data to receive from the AP (DtR=1).

When the TIM element in the Beacon signals the existence of an MSDU having the STA as the destination (DtR=1) the STA sends a PsPoll message. This message is sent after a DIFS time if it is the only STA to receive a message, or by activating the back-off procedure if there are messages waiting for more than one STA. If, after a scheduled waking-up, neither a Beacon, nor an expected broadcast data packet is received, the STA returns to the WaitBcBMData state until these events happen.

B. The second condition for the station to wake up is the need to transmit data from a higher-level application, that is transmission of an MSDU (Dts=1). These moments are randomly generated following an exponential distribution. The basic access procedure is activated for the transmission of the first MPDU and successive SEND-SIFS-ACK-SIFS transitions are used until the entire MSDU is sent towards AP. The STA waits for an ACKTimeOut time for an ACK. It repeats the transmission of the last MPDU if the waiting time is longer than this.

Based on the GSM standard features [4] and aiming at using a worst case analysis the following rules were adopted in building the GSM part of a mobile transceiver (Dual Station).

There are 4 possible operational states for an MS (Mobile Station): SEND, RECEIVE, MEASURE and SILENT. While in SEND state the MS transmits towards its own BS on the allocated frequency and time Slot. During the RECEIVE state the MS receives the information transmitted by the BS (Base Station) on the allocated frequency and time Slot. In the MEASURE state the MS tunes itself to the frequencies broadcast by the network and takes measurements of the local electromagnetic field strength. Finally, the SILENT state is the time period when the MS has no task to fulfill.

While engaged in a call (CA=1) and the discontinuous transmission feature is not activated (DTX=0) the time evolution of the MS states is the one presented in Fig. 5 and it is almost the same as the standard specifies. The only difference is that the transmitting Slot does not occur at exactly 3 Slots after the beginning of a receiving Slot, but earlier with a Ta (timing advance) time. This timing advance Ta is needed for the compensation of the propagation delay and it is used for a precise superposition of the MS's transmissions at the BS level. In order for all transmissions associated with a time Slot to arrive at the BS precisely at the beginning of the given Slot it is mandatory for the MSs at a greater distance from the BS to begin their transmissions earlier than MSs positioned at a smaller distance from the same BS. According to the GSM Specifications the timing advance is a multiple of the bit period (which is approx. 3.69μs). The maximum value of the multiple is 63 and so, the maximum value of the timing advance is approx. 232ms. Obviously the minimum value of the timing advance is 0.

The time and duration of the mandatory measurements during a call in progress are implementation dependent of both the network and the mobile terminal. The network requests the mobile terminal to take periodic measurements on the frequency pilot of their own cell and, also, on frequency pilots of all the neighboring cells. We considered for the simulation that the measurement window lasts for a time Slot duration (approx. 577μs) and it is randomly positioned around the center of the time interval separating a transmitting Slot and the next receiving one. The time offset is no greater than half of Slot.

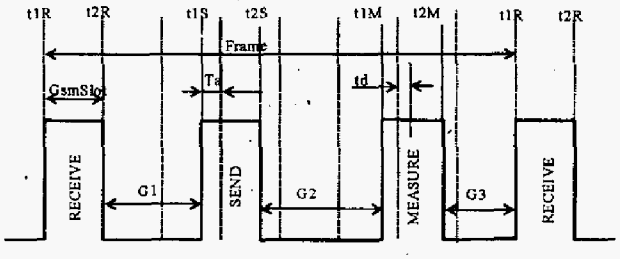


Figure 5. GSM state sequence when CA = 1 and DTX = 0

When the DTX feature is activated the mobile terminal remains in its SILENT state during the allocated transmitting Slots, except for the 8 specified Slots in the Frames no. 52-59 in a group of successive 4 Traffic Multiframe.

While in idle mode the MS stays in its SILENT state most of the time, except for the 4 no. 0 Slots in the Frames included in its Paging group.

The Ps_STA part of the Dual STA has to have some supplementary features in order to interact with the GSM part. As the GSM standard is a pure implementation of the TDMA concept mobile terminals has to obey all the control commands it receives from the network like to transmit and to receive on the frequency channel and the time slot allocated by the network. As a consequence a dual terminal has to treat its tasks in the GSM network with greater priority than the ones in the WLAN. The sole exception is the receiving of the Beacon, which has absolute priority on all the events in GSM part.

The reason of implementing this exception is that after loosing a Beacon due to a simultaneous GSM task the Dual station continues to wait the Beacon all the time interval until the next Beacon, thus blocking all its traffic activity in the WLAN network. This has a great negative influence on its throughput and on its power saving efficiency. Loosing a slot or even two successive slots in GSM network has no other consequence than stealing a very small part of the voice traffic with little effect on its intelligibility.

The absolute priority given to GSM events means that any WLAN activity could be interrupted in the following situations: MPDU transmission, ACK transmission, PS Poll transmission, RTS transmission; Backoff, Receive, Waiting states. This behavior is obtained in the model by setting a variable (ExitGsm) anytime a WLAN state is abnormally terminated. The GSM part of the Dual STA contains, also, an extra state (GsmBusy) as compared to a normal STA in order to take into account this new feature.

In Fig. 6 is presented an example of 16ms of activity in the simulated network, while the Beacon Period is set to 10ms. The diagrams represent the codified states of the AP, Ps_STA and DualSTA (including WLAN and GSM), where in the WLAN diagrams values of 20 to 30 signifies Tx states (30 – Beacon B, 27 – Broadcast Message BM, 28 – RTS, 29 – CTS, 25 – Ps-Poll, 26 – MPDU Data, 24 – ACK), values of 10 to 20 means Rx states (13 – Beacon B, 12 – Broadcast Message BM, 18 – RTS, 19 – CTS, 15 – Ps-Poll, 16 – MPDU Data, 14 – ACK) and values from -5 to 10 are codes for Waiting states (-5 – Idle/Doze, -3 – GSMBusy, 0 – BackOff, 1 – SIFS, 3 – DIFS, 4 – WaitData), while in the GSM diagram, 0 means SILENT, 1 – RECEIVE (Rx), 2 – MEASURE (Mon) and 3 – SEND (Tx).

It may be observed in the example considered that there is 1 MSDU formed by 6 MPDUs transmitted from the Dual_STA to the WLAN Ps_STA trough AP and the Ps_STA transmits to the Dual_STA via AP 1 MSDU formed by IV MPDUs. The WLAN part of Dual_STA is interrupted and delayed by the GSM activity, which has greater priority. Also, during the GSMBusy state, the Ps_STA takes control over the WLAN channel and when the DualSTA resumes from GSMBusy, it has to monitor and wait for the channel to become idle.

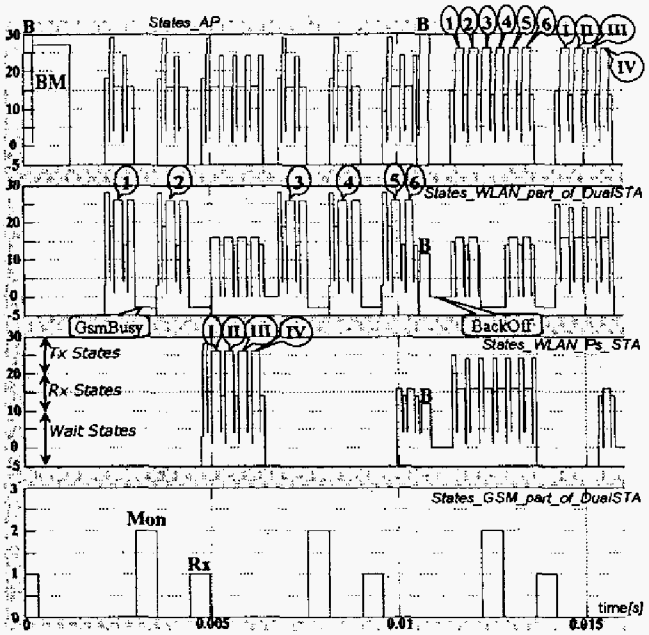


Figure 6. An example of activity in the simulated network

One can observe that the MPDU transmission interrupted by GSM is not acknowledged and that particular MPDU is retransmitted. Also the first Beacon is followed by a Broadcast Message and the second is delayed until the AP finishes the receiving of the MPDUs transmitted by one STA at the TBTT moment. After receiving the Beacon (at 11ms), because both stations have data to receive, they enter in competition for the channel, using the BackOff mechanism. First station that sends a PsPoll message gains control over the channel, receive its corresponding data from the AP, while the other station waits its turn to use the channel and receive its data from AP.

IV. Conclusions and Future Work

The simulation model developed using Simulink Stateflow is of great help in studying the opportunity of using a Dual Station including both WLAN and GSM standards on a single chip. The throughput and the delay in transmissions are relevant for such a DualSTA implementation and both may be determined, and then analyzed, using our simulation model. But this analyze suppose a lot of simulations with different parameters and this is just our next future work.

References

[1] J. Heiskala, J. Terry, "OFDM wireless LANs: A theoretical and practical guide", SAMS Publishing, 2002.

[2] P. Nedeltchev, "Wireless Local Area Networks and the 802.11 Std.", http://www.cisco.com/warp/public/784/packet/jul01/pdfs/whitepaper.pdf, 2001.

[3] Ph. Belanger, W. Diepstraten, "802.11 tutorial. 802.11 MAC entity: MAC basic access mechanism, privacy and access control", http://grouper.ieee.org/groups/802/11/Tutorial/MAC.pdf, 1996.

[4] 3GPP TS 45.002 V6.6.0 (2004-04) - "Technical Specification Group GSM/EDGE: Multiplexing and multiple access on the radio path".

[5] IEEE 802.11, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, ANSI/IEEE Std. 802.11, 1999.