DOI https://doi.org/10.30525/978-9934-26-597-6-25

REAL-TIME AUDIO STREAM AGGREGATION IN BLUETOOTH PERSONAL NETWORKS

Khotunov Vladyslav, Kulyk Oleg, Burmistrov Serhii
Cherkasy State Business-College, Ukraine
**Corresponding author's e-mail: vkhotunov@gmail.com, sergij.burmistrov@ukr.net*

Abstract

This article explores an innovative approach to real-time audio stream aggregation in Bluetooth personal area networks (BPAN). Standard Bluetooth implementations are limited to one-to-one communication schemes, preventing the integration of multiple simultaneous audio sources. The proposed architecture introduces a scatter network model with an intermediary aggregator node that processes, synchronizes, and forwards combined audio streams to a receiver. The prototype was implemented using ESP32 microcontrollers, FreeRTOS, and the A2DP Bluetooth profile. Simulations and laboratory tests confirm the feasibility of the solution, showing reduced latency, improved synchronization, and efficient bandwidth usage. The results suggest that this method can enhance Bluetooth applications in multimedia systems, IoT devices, and smart environments.

Keywords: Bluetooth networks, audio streaming, aggregation, ESP32, FreeRTOS, A2DP, scatternet

Introduction

Bluetooth technology is central to wireless multimedia and IoT communication, but its conventional architecture limits dynamic multi-source streaming. While Bluetooth Multipoint allows for switching between devices, true simultaneous audio reception remains unsupported. This article addresses the gap by presenting a system capable of aggregating audio from several sources within a scatter network topology. The system employs an aggregator node that acts as both a slave and master in distinct pico networks. The primary motivation is to expand real-time multimedia capabilities using low-cost, widely available hardware.

The proposed solution involves creating a Bluetooth scatternet where several audio sources are connected to an aggregator. This device collects, synchronizes, and processes PCM audio streams, merging them into a single signal that is transmitted to a Bluetooth audio receiver. The hardware setup uses an ESP32-LyraT board, chosen for its support of A2DP, audio codecs, and dual-core processing. Software implementation leverages FreeRTOS multitasking and ESP-ADF for audio handling.

The core algorithm includes modules for device pairing, audio buffering, PCM mixing, and adaptive bitrate management. Synchronization across input streams is handled using shared semaphores and timestamp alignment. Special attention is given to latency management and buffer overflow prevention.

A simulation model was developed using Wokwi to prototype real-time data processing. The system was tested with virtual audio generators and playback devices, validating the mixing and transmission pipeline. Additionally, a physical testbed using two smartphones and a Bluetooth headset demonstrated successful multi-source audio playback with acceptable quality.

Power consumption measurements indicate that the system is suitable for battery-powered use, consuming less than 100 mW in active mode. This makes it viable for portable and wearable applications. The modular software design allows future support for LE Audio, LC3 codec integration, and remote control via AVRCP.

To ensure the system's performance in real-world use cases, several optimization strategies were implemented. These included efficient memory management routines for audio buffer allocation and dynamic task scheduling within FreeRTOS. By assigning different priorities to streaming, mixing, and transmission tasks, the system minimized latency spikes and ensured real-time processing even under high CPU load. Additionally, ring buffers were employed for audio data to allow seamless reading and writing operations.

Extensive tests were carried out to measure synchronization accuracy. By analyzing time-stamped audio packets, it was determined that stream alignment jitter remained within ±5 ms–an acceptable margin for human perception in stereo playback scenarios. When used with voice input, buffer sizes were adjusted to prioritize low-latency transmission over fidelity. Furthermore, codec configurations were tuned to match various audio types, with SBC used for speech and AAC for music.

A multi-device use case was tested wherein one source transmitted speech while another played background music. The aggregator successfully balanced volume levels, prevented signal clipping, and ensured intelligibility of speech. This paves the way for deployment in smart assistant systems, where voice commands and alerts may need to be layered over ambient audio.

Another innovation is the inclusion of dynamic bitrate adaptation. By monitoring the Bluetooth link quality (RSSI) and throughput statistics, the aggregator dynamically switches between audio quality presets to reduce dropouts. This feature was tested in scenarios with increased RF interference, such as near microwave ovens and Wi-Fi hotspots. The aggregator responded effectively by lowering bitrate to maintain connection stability.

Modularity is a central design philosophy. The system can be expanded with additional sensors such as accelerometers for motion-aware audio switching, or infrared sensors for user presence detection. This opens possibilities for context-aware smart audio environments, where playback behavior changes depending on physical presence or activity.

Beyond technical evaluation, the project includes educational and practical outreach. A documentation website was created using MkDocs, complete with setup guides, schematics, and firmware tutorials. Workshops were held at local universities to demonstrate the system, and open-source release on GitHub encourages student-led enhancements. Initial feedback suggests interest from audiophile communities, developers of hearing aids, and DIY makers seeking multi-source Bluetooth capabilities.

Future extensions may involve support for BLE Isochronous Channels (LE Audio), which enables natively synchronized multi-stream playback. A prototype integration with ESP32-C3 chips is underway, evaluating BLE Mesh for distributed audio sharing. Cross-platform support is also planned, allowing Linux-based hosts (e.g., Raspberry Pi) to serve as aggregators, further expanding versatility in embedded and edge computing contexts.

To verify the system's compatibility with consumer electronics, interoperability tests were conducted with commercial Bluetooth devices including wireless earbuds, speakers, and vehicle infotainment systems. The aggregator successfully paired and transmitted audio to all tested devices. In cases where the receiving device did not support A2DP sink role, the system defaulted to mono stream fallback. Bluetooth version compatibility tests revealed that while performance is optimal with Bluetooth 5.0 and above, backward compatibility with 4.2 devices is maintained through adaptive bitrate negotiation.

Power efficiency tests revealed that the system's average current draw remains below 85 mA in active mode. Standby operation with BLE-only scanning consumes less than 5 mA, enabling practical use in mobile and battery-powered environments. Sleep mode routines were tested and confirmed to successfully suspend non-essential system tasks, resuming operation within 200 ms upon audio input detection.

From a software engineering perspective, the firmware architecture follows a modular service-oriented model. Core services include Audio Input Handler, Stream Sync Engine, Audio Mixer, and Output Transmitter. Each module operates in isolated FreeRTOS tasks with intercommunication via message queues and semaphores, improving fault isolation and code maintainability. In case of task failure, the system can automatically restart the affected module without a full reboot.

The project's development process adhered to Agile principles, with sprint-based iterations, version control in Git, and unit tests for core functionality. Continuous integration was simulated using GitHub Actions to build and validate firmware across multiple configurations. System logs are accessible over UART or Wi-Fi serial bridge, facilitating debugging and telemetry collection during field tests.

Use cases for this technology include multi-user conference headsets, distributed smart speaker networks, and even public address systems that require concurrent microphone and background audio playback. Its affordability and flexibility make it suitable for educational labs exploring wireless communication, real-time audio processing, and embedded system design.

Further research will examine integration with AI voice assistants for real-time voice command recognition and contextual audio switching. Additionally, security measures such as AES encryption and secure pairing protocols will be evaluated to protect transmitted audio data, especially in sensitive environments like healthcare and defense.

Conclusions

The developed audio stream aggregation architecture significantly extends the capabilities of Bluetooth-based multimedia systems. It enables real-time, multi-source audio streaming without requiring proprietary standards or hardware. The proposed solution demonstrates potential for commercialization in smart home audio, conferencing systems, and wearable electronics. Future research may explore adaptive audio compression, multi-zone playback, and cross-protocol streaming via Wi-Fi or BLE Mesh.

References

[1] Zeadally, S., Siddiqui, F., & Baig, Z. (2019). 25 Years of Bluetooth Technology. Future Internet, 11(9), 194. https://doi.org/10.3390/fi11090194

[2] Fayyoumi, E., et al. (2014). L2CAP-Based Prototype Media Streaming via Bluetooth Technology. International Journal of Networking and Virtual Organisations, 14(3), 221. https://doi.org/10.1504/IJNVO.2014.065785

[3] Zeadally, S., & Kumar, A. (2005). Design, Implementation, and Evaluation of the Audio/Video Distribution Transport Protocol (AVDTP) for High Quality Audio Support over Bluetooth. Computer Communications, 28(2), 215–223. https://doi.org/10.1016/j.comcom.2004.09.007

[4] Locatelli, P., et al. (2023). Device Discovery and Tracing in the Bluetooth Low Energy Domain. Computer Communications. https://doi.org/10.1016/j.comcom.2023.02.008

[5] Bezkorovainyi, V., Kolesnyk, L., Gopejenko, V., Kosenko V (2024) The Method of Ranking Effective Project Solutions in Conditions of Incomplete Certainty. Advanced Information Systems, 2024, 8(2), pp. 27–38. DOI: 10.20998/2522-9052.2024.2.04