Alango Short Range Sound Modem (SRSM)

Dmitry Dicky
CTO, Alango Technologies

The emergence of the smart home, wearables, and hearing devices in our lives has created a world where data and connectivity are becoming increasingly complex. Global sales of such devices are growing steadily. These include smart speakers, earbuds, watches, headsets, car audio, and more. With this steady proliferation comes an increased need for a non-destructive method of controlling these devices. The requirement is to reconfigure out-of-the-box sealed devices, transfer certain data to them in an RF-restricted environment, or even communicate with equipment that is not RF-enabled.

Even though many devices on the market feature connectivity, not all of them utilize extra software and hardware that support auxiliary data transfer and fine-tuning. More often, connected devices work with a set of standard protocols (Bluetooth®, Wi-Fi, etc.), and few can be reconfigured on-the-fly without expensive dedicated software.

Alango SRSM is designed to overcome these limitations. In mid-2019 we began developing this sound modem specifically for data-over-sound communication, where no other means of connectivity is possible. We focused on transferring data between a computer and a device with a built-in microphone integrating Alango's DSP technologies, to upload configuration data. In other words, to communicate directly with the DSP that runs Alango's software.

Data-Over-Sound

SRSM utilizes similar technology to that of a traditional telephone modem, wherein the transmitter encodes the data and sends it via a communication line. On the other side, the receiver decodes the data with possible feedback to the sender. However, aside from this analogy, not all devices can reply to the sender, and some are capable of one-way communication only.

With SRSM on the transmit side, data is encoded into an acoustic signal and played through the medium of air, but it could as well be a transmit-line like the telephony modem example mentioned previously. On the receiving side, data is collected by the microphone, decoded, and then passed to the data consumer.

The picture below illustrates the typical transmitter-receiver configuration.

Why Data-Over-Sound?

There are already several technologies and solutions, including data-over-sound, that provide communication between devices—QR, NFC, Wi-Fi, Li-Fi, ZigBee and others. Because the technology is integrated into the end-user device, the connectivity protocol is well documented and developed. However, for instance, while a typical headset uses a Bluetooth ("BT") RF link, there is no BT protocol designated to configure and fine-tune the DSP data processing blocks, such as echo cancelling, noise reduction, and other types of processing. The manufacturer must perform the tuning. The only way to change these "internal" parameters is by using a cumbersome disassembly procedure that will provide access to the physical connections.

Another example is transferring data between devices which are not RF-enabled but have some audio features, like toys. They might be able to speak and perhaps even listen, but not to one another.

In all these cases, data-over-sound adds extra functionality to the devices in which no other means of non-destructive communication is possible.

Advantages of Alango's SRSM

Alango's SRSM is a front-end technology. It passes unmodified data to the next level, when no data is acquired and indicates all stages when communication begins: preamble, main data receiving, and end-of-data with possible error code.
The advantages of SRSM are as follows:

- Software-only solution—zero additional BOM for sound-equipped devices
- Low power consumption—no high sampling rate required; SRSM can work with 8kHz sampling rate

- No pairing required—reliable data acquisition and an extremely low false alarm rate
- Integrates with any sound-equipped device
- Supports both one-way and two-way communication—no feedback from the receiver is required for data transfer
- High data rate
- Platform- and core-agnostic
- Requires as little as 1 kB of RAM for ARM Cortex-M and < 8kB of ROM for code plus memory of user's data storage
- Receiver requires only 1.3MIPS (2.2MIPS for harsh environments)
- Transmitter needs even less memory and much fewer MIPS
- Communication rate is up to 1500 bits per second
- Allowed clock skew up to 10^{-4}
- Supports one-to-many transmission

Internals of SRSM

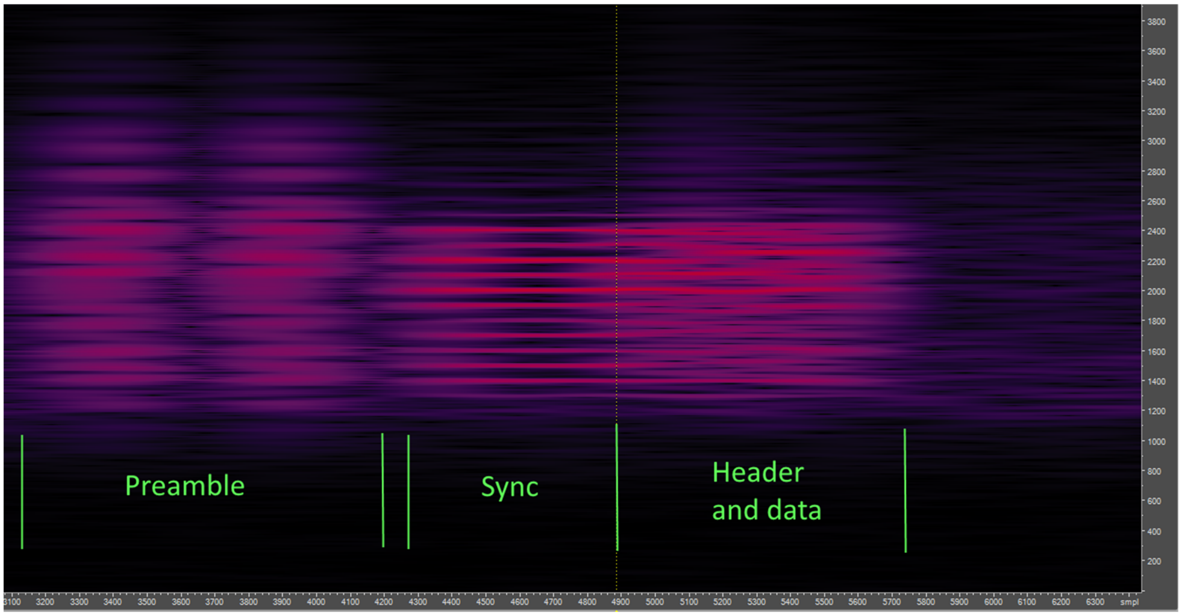
SRSM uses PSK modulation to achieve a high data rate while operating at a low sampling rate and within a short range. In general, PSK is not robust to reverberations, multi-path propagation, and clock skew. SRSM filters odd signal distortions and follows the clock skew, thus making communication reliable.

SRSM maintains a low false alarm rate while preserving high acquisition reliability. No feedback required – data integrity is checked against two, spatially separated, CRCs.

SRSM utilizes low bandwidth (1.2kHz to 2.6kHz) where almost all speakers and microphones provide good frequency response.

Typical sequence's sonogram is shown in the picture below:

:



The transmitter encodes the data provided with 8-channel OFDM and PSK (QAM-4) for every channel, adds preamble (two following extended Barker's code sequences), creates synchronization sequence, adds a header (consisting of data type, length, data CRC, and header CRC), and then adds encoded data. In the end, wave data is sent to the DAC and then played by the speaker.

The transmitter assumed to be working with 8kHz or 16kHz sampling rates.

The receiver is a front-end technology - it must be situated between ADC and any other data processing blocks. It can work with 8kHz and 16kHz sampling rates. In general, any sampling rate is possible:

- When the sampling rate is a multiple of 8, no modifications to transmitter/receiver are required.

- When the sampling rate is **not** a multiple of 8, no changes to transmitter/receiver are needed, but the transmitter's sampling rate must match the receiver's sampling rate.

The receiver provides transparent data transfer with zero-delay to further data blocks. It listens constantly to the incoming data, searching for the preamble. When the preamble is detected, it calibrates its internals to the assumed clocks/timings, receives the header, checks CRCs, and continues to collect the forthcoming data. During reception, it follows the injected pilot signal, thus maintaining the required phase match and ensuring reliable data acquisition. When data reception is complete, it computes the data's CRC and checks it against the one received previously. The receiver indicates possible error codes to the upper level.

The API for both transmitter and receiver is simple; it just provides data storage and calls the main processing block in a cycle. As is typical, the central processing block is called within a cycle that assumes the data is transferred to the user by frames of a particular size, but SRSM can work with frames of ANY size.

Data Rate, Range, Reliability

SRSM utilizes 8-channel OFDM with PSK (QAM-4) for each channel modulation scheme. It encodes 16 bits of data per symbol of 10ms length. This provides a net data rate of about 1.5kb per second. The header and preamble take 250ms.

SRSM is designed for short-range communication. With standard speaker and standard MEMS microphone, it has been tested with a range up to 50cm. Any distance less than this will also work, providing there is no microphone overload.

The SRSM preamble guarantees that it cannot be spoofed with similar sounds unless the sender knows exactly what to send. Neither human or natural sounds can generate that sequence. Computed probability of the false alarm, when the wrong preamble is detected is about 10^{-6}.

Data reliability depends on the Signal-to-Noise Ratio (SNR). To simplify processing and achieve low power consumption, SRSM does not use error correction. With a SNR of about 8dB, this type of modulation provides 10^{-4} BER probability. Reverberations and ambient noises can both be considered noise. In a quiet, echo-free environment, the communication range can be extended.

Use Cases

SRSM is designed to reduce human efforts to create a reliable data channel intended for data transfer. The primary scenario for SRSM is sending data to the sound-equipped devices, which cannot be accessed by any other non-destructive means.

Because SRSM is a software-only solution, OEM and end manufacturers do not have to modify the hardware to gain the extra functionality for their products.
Several scenarios in which SRSM is a reliable and the only possible solution are outlined below.

IoT: Configuring smart devices to perform a variety of complicated tasks: communicating with other devices, listening to environmental sounds, checking sensors, and sending data to the database. With SRSM the user can change parameters without accessing either the device itself or the database. For example, changing the heater’s temperature threshold.

Proximity detection: Because of the short-range nature of SRSM it can be used as a door opener and/or presence indicator. Data in the payload may indicate the user’s ID and provide some other information. Our tests show that the generated signal cannot be reliably recorded from long distances, and thus in this scenario, the user's ID cannot be spoofed.

Two-way acoustic communication: There are a variety of two-way communication scenarios where two devices talk to each other. However, unlike with RF communication, with SRSM the dialogue cannot be reliably recorded. This brings, for example, key exchange to a higher level of security.

Communication in RF-restricted environment: In some cases, RF communication is not possible (in some hospitals rooms, for example). SRSM overcomes this limitation by bringing reliable data transfer to these places.

Toy and robot control: These days, toys and robots can talk and listen to each other. With two-way SRSM they can exchange the information with their own mysterious and entertaining language.