SoniTalk Protocol Specification

Working Draft v0.1.1

Introduction

The ultrasonic frequency band represents a novel and so far hardly used channel for the communication of different devices, such as mobile phones, computers, TVs, and personal assistants. Ultrasonic communication is a promising technology since it requires only a standard loudspeaker and a microphone (as built into our phones) for communication. SoniTalk is an open near-ultrasonic data transmission protocol ("data over audio"). This is a first draft of the specification of the protocol for data transmission. This specification targets the lowest layer (physical layer) of the protocol stack as well as the second layer (data link layer).

Physical layer

The goal of the physical layer is to encode individual bits in the form of ultrasonic signals. Different techniques, such as amplitude shift keying, frequency shift keying and phase shift keying exist for this purpose.
The theoretically possible ultrasonic frequency band for SoniTalk ranges from 18 to 22kHz. This band is frequently corrupted by noise, which requires a number of counter measures to assure a robust signal transmission. Especially the temporally varying characteristics of the channel makes the transmission of messages over longer time-spans likely to be corrupted, i.e. that parts of the message get lost. SoniTalk tries to circumvent this source of error by a design that uses parallel signal transmission on several frequency bands. This means that the target frequency band of N to N+B Hertz is divided into C carrier frequencies. The intended encoding is a simple amplitude shift keying, i.e. each of the C carrier frequencies is either on (for logical 1) or off (for logical 0). Each such message block has a duration of D milliseconds (usually in the range of 10-30ms). For illustration, each message can be thought of as a spectral barcode that encodes a part (a word) of the message.

Several such message blocks (barcodes) together make up an entire message. A second means to make the SoniTalk protocol robust to noise is the choice of a fixed message length. This means that a message always consists of M message blocks that are sent in sequence. The use of a fixed message length makes it easier at the decoder side to find the start and end of a message as well as to locate the boundaries between the individual message blocks.

The current SoniTalk specification represented by this draft uses the following parameters

Current specification:
N = 19000Hz
B = 1000Hz
C = 64 frequencies
D = 30ms
M = 10 blocks

Each message encodes 64x10=640bits. The maximum data rate that can be achieved with these parameters is: 21.333 bits/second.

Data Link Layer:

To enable a robust decoding of the message even in the case of corruptions of individual bits by noise, SoniTalk includes different mechanisms for error correction. Error correction is first employed on the individual blocks of a message, i.e. a subset c of the C carrier frequencies is used for redundant coding of the signals. With this first error correction, small transmission errors in the individual blocks can be corrected and corrupted blocks (in case there are too many errors) can be detected.

A second level of robustness is introduced by a temporal error correction scheme which goes over the individual message blocks. For this purpose a subset m of the M message blocks is used to encode error detecting and correcting information for each carrier frequency individually. This scheme should ensure that even in case a message block could not be transmitted successfully, its information can still be reconstructed from the m redundantly coded message blocks (parity blocks).

Current specification:
c = 20 frequencies
m = 4 blocks

By using redundant encoding as specified above, an effective data rate of 8.800 bits/second is possible.

Remarks:

The parameters in this draft specification may change based on empirical tests and theoretic findings in the process of further development and implementation of the SoniTalk specification.

License