Acoustic Data Communication with Mobile Devices

Victor I. Khashchanskiy
First Hop Ltd., Finland

Andrei L. Kustov
First Hop Ltd., Finland

INTRODUCTION

One of the applications of m-commerce is mobile authorization, that is, rights distribution to mobile users by sending authorization data (a token) to the mobile devices. For example, a supermarket can distribute personalized discount coupon tokens to its customers via SMS. The token can be a symbol string that the customers will present while paying for the goods at the cash desk. The example can be elaborated further—using location information from the mobile operator, the coupons can only be sent to, for example, those customers who are in close vicinity of the mall on Saturday (this will of course require customers to allow disclosing their location).

In the example above, the token is used through its manual presentation. However, most interesting is the case when the service is released automatically, without a need for a human operator validating the token and releasing a service to the customer; for example, a vending machine at the automatic gas station must work automatically to be commercially viable.

To succeed, this approach requires a convenient and uniform way of delivering authorization information to the point of service—it is obvious that an average user will only have enough patience for very simple operations. And this presents a problem.

There are basically only three available local (i.e., short-range) wireless interfaces (LWI): WLAN, IR, and Bluetooth, which do not cover the whole range of mobile devices. WLAN has not gained popularity yet, while IR is gradually disappearing. Bluetooth is the most frequently used of them, but still it is not available in all phones.

For every particular device it is possible to send a token out using some combination of LWI and presentation technology, but there is no common and easy-to-use combination. This is a threshold for the development of services.

Taking a deeper look at the mobile devices, we can find one more non-standard simplex LWI, which is present in all devices—acoustical, where the transmitter is a phone ringer. Token presentation through acoustic interface along with general solution of token delivery via SIM Toolkit technology (see 3GPP TS, 1999) was presented by Khashchanskiy and Kustov (2001). However, mobile operators have not taken SIM Toolkit into any serious use, and the only alternative way of delivering sound tokens into the phone-ringing tone customization technology was not available for a broad range of devices at the time the aforementioned paper was published.

Quite unexpectedly, recent development of mobile phone technologies gives a chance for sound tokens to become a better solution for the aforementioned problem, compared with other LWI. Namely, it can be stated that every contemporary mobile device supports either remote customization of ringing tones, or MMS, and in the majority of cases, even both, thus facilitating sound token receiving over the air.

Most phone models can playback a received token with only a few button-clicks. Thus, a sound token-based solution meets the set criteria better than any other LWI. Token delivery works the same way for virtually all phones, and token presentation is simple.

In this article we study the sound token solution practical implementation in detail. First, we select optimal modulation, encoding, and recognition algorithm, and we estimate data rate. Then we present results of experimental verification.

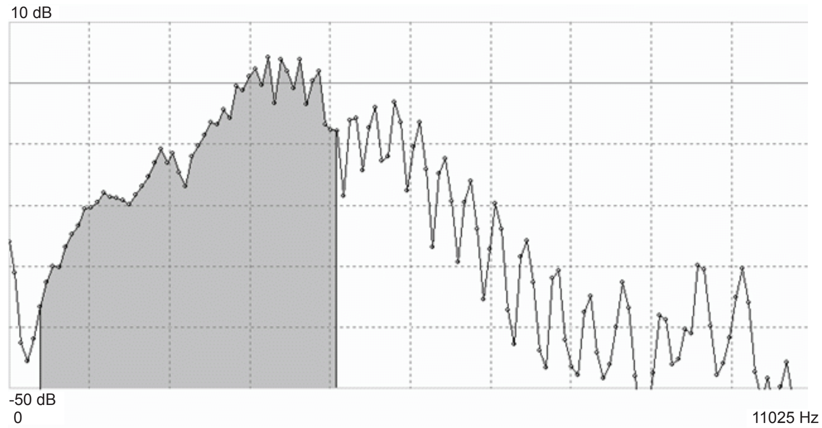
ACOUSTIC DATA CHANNEL

We consider the channel being as follows. The transmitter is a handset ringer; information is encoded as a sequence of sine wave pulses, each with specified frequency and amplitude.

Multimedia message sounds and most ringing tones are delivered as sequences of events in MIDI (musical instrument digital interface) format. A basic pair of MIDI events (note on and note off) defines amplitude, frequency, duration of a note, and the instrument that plays this note. MIDI events can be used to produce information-bearing sound pulses with specified frequency and amplitude.

Widely used support of polyphonic MIDI sequences allows playback of several notes simultaneously. Nonetheless, this has been proved worthless because in order to get

Figure 1. Frequency response measured with test MIDI sequence in hold-max mode



distinguished, these notes have to belong to different non-overlapping frequency ranges. Then the bit rate that can be achieved would be the same as if wider frequency range was allocated for a single note.

The receiver is a microphone; its analog sound signal is digitized and information is decoded from the digital signal by recognition algorithm, based on fast fourier transform (FFT) technique. FFT is, in our opinion, a reasonable trade-off between efficiency and simplicity.

We investigated acoustics properties of mobile devices. After preliminary comparison of a few mobile phone models, we found that ringer quality is of approximately the same level. All handsets have a high level of harmonic distortions and poor frequency response. The results shown in Figures 1 and 2 are obtained for a mid-class mobile phone SonyEricsson T630 and are close to average.

MIDI-based sound synthesis technology applies limitations on pulse magnitude, frequency, and duration. At the same time, ringer frequency response is not linear and the level of harmonic distortions is very high. Figure 1 shows frequency response measured with a sweeping tone or, to be precise, a tone leaping from one musical note to another. To obtain this, the phone played a MIDI sequence of non-overlapping in-time notes that covered a frequency range from 263 to 4200 Hz (gray area).

The frequency response varies over a 40 dB range, reaching its maximum for frequencies from approximately 2.5 to 4 kHz. Moreover, spectral components stretch up to 11 KHz, which is caused by harmonic distortions. This is illustrated also by Figure 2.

Horizontal axis is time; overall duration of the test sequence is 15 seconds. Vertical axis is sound frequency, which is in range from 0 to 11025 Hz. Brightness is proportional to sound relative spectral density; its dynamic range is 60 dB, from black to white.

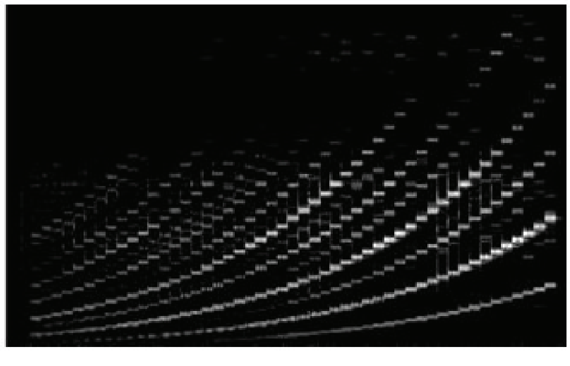
We also found that frequency of the same note may differ in different handsets. Nevertheless, the ratio of note frequencies (musical intervals) remains correct, otherwise melodies would sound wrong.

For a simplex channel with such poor parameters, as reliable a data encoding method as possible is to be used. Frequency shift keying (FSK) is known as the most reliable method which finds its application in channels with poor signal-to-noise ratio (SNR) and non-linear frequency response.

It is not possible to negotiate transfer rate or clock frequency, as it is usually done in modem protocols because acoustic channel is simplex. To make the channel as adaptive as possible, we have chosen to use differential FSK (DFSK), as it requires no predefined clock frequency. Instead, frequency leaps from one pulse to another provide the channel clocking. The difference between frequencies of consecutive pulses determines the encoded value.

Once encoding scheme is selected, let us estimate possible transfer rate before we can find the balance between data

Figure 2. A spectrogram of the test MIDI sequence



Related Content

High-Performance Reconfigurable Computing
Mário Pereira Vestias (2019). *Advanced Methodologies and Technologies in Network Architecture, Mobile Computing, and Data Analytics (pp. 731-744).*
www.irma-international.org/chapter/high-performance-reconfigurable-computing/214656

Large Imagery on Small Screens: Novel Technology for Device Adaptation in Mobile Services
René Rosenbaum (2013). *Mobile Services Industries, Technologies, and Applications in the Global Economy (pp. 134-152).*
www.irma-international.org/chapter/large-imagery-small-screens/68656

Modulation Recognition of Digital Multimedia Signal Based on Data Feature Selection
Hui Wang, Li Li Guoand Yun Lin (2017). *International Journal of Mobile Computing and Multimedia Communications (pp. 90-111).*
www.irma-international.org/article/modulation-recognition-of-digital-multimedia-signal-based-on-data-feature-selection/188626

Sampling and Reconstructing User Experience
Panos Markopoulosand Vassilis-Javed Khan (2011). *International Journal of Handheld Computing Research (pp. 53-72).*
www.irma-international.org/article/sampling-reconstructing-user-experience/55891

Snapshot Assessment of Asia Pacific BWA Business Scenario
C. Chin Wong (2007). *Encyclopedia of Mobile Computing and Commerce (pp. 906-911).*
www.irma-international.org/chapter/snapshot-assessment-asia-pacific-bwa/17194