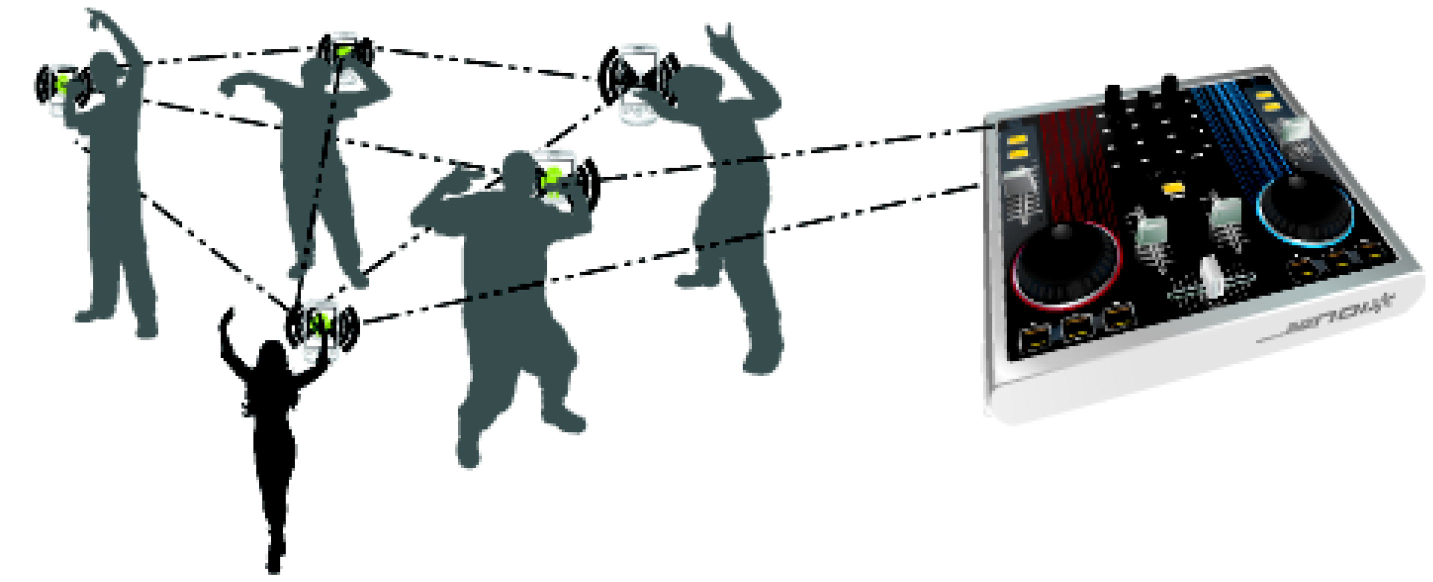
A framework for distributed audio smartphone games

K. Collins[1], P.J. Taillon[1], B. Kapralos[2], J.-M. Trivi[3]
[1]CCAT University of Waterloo, [2]University Ontario Inst. of Technology, [3]Google, Inc.



Introduction

The explosive growth of smartphone sales in recent years has almost no comparison in the sector of consumer electronics. Sales in 2010 almost doubled those of 2009 year, with Android-based units showing a startling 1,580% growth in the last quarter of 2010, as compared to the same period the previous year [2].

Given such a proliferation of mobile computing power, there is much need for further exploration in the areas of distributed gaming, multimedia, networking, and audio for mobile devices. This project explores the use of smartphone technology as computing units for ad hoc gaming support. The underlying environment assumes a close-proximity setting, such as a classroom, conference room, or social setting (e.g. an activity area on the order of 15m²).

Project Overview

The work presented here is the first of three applications encompassing our research programme. The objective is explore the potential for networking smartphones for distributed audio-based games such as:

1. audio playback using smartphones of synchronized tracks for social gaming applications, providing opportunity for interactivity between participants;
2. exploration of the role smartphones can play interacting with a smart-table in gaming;
3. using smartphones to explore educational games relating to acoustics, using sound field effects created by physically moving the phones in a monitored environment.

For example, we envisage a music-based game where streamed audio data (MIDI) of musical sequences is sent as individual, separate music tracks to a series of players who have a synchronized music sequencer on their phone. Players can re-mix the track in real-time, playing a kind of musical "telephone" game where they change a note or beat and send their selected sequence on to the next person. Eventually the sequence returns to them, altered from its original but still fitting with the other tracks that are simultaneously being altered.
This phase of activity details the system framework for confederating mobile phones for distributed audio games.

Design Considerations

1. Where possible, the framework should not rely on specialized hardware, beyond what is packaged in any off-the-shelf desktop/laptop and smartphone.
2. The support software on the phone itself must present modest processing requirements, and the data storage should leave a minimum footprint.
3. The system must impose low power demands on the phone battery.

Network subsystems

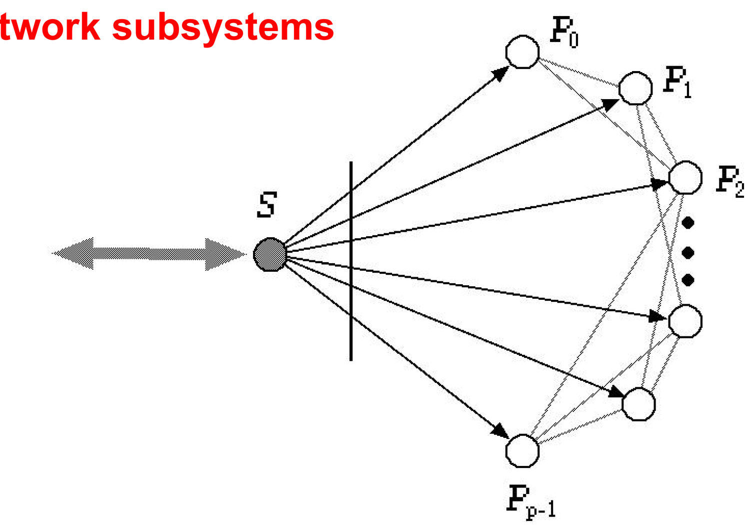


Figure 1. A central server, *S*, broadcasts audio stream data via high-bandwidth WiFi. A small network between phones nodes, P_i, supports the exchange of control data via low-bandwidth, low-power Bluetooth.

System Architecture

Taking into account bandwidth constraints outlined in [1], the system framework consists of two networking subsystems (see Figure 1):

- media streaming from a server to the phones over WiFi;
- a Bluetooth control network established between the phones, that enables initial position determination; information exchange; and network teardown.

A WiFi/Bluetooth card on the server acts as a local hotspot, making it discoverable by the phones. A main consideration with audio streaming is the control and synchronization of the transmission and playback. We use the Real Time Streaming Protocol (RTSP), which establishes and controls one or more time-synchronized streams of continuous audio data [3]. The server operates in a manner oblivious to the number of phones participating, i.e. it simply broadcasts the streams over WiFi. The phone nodes implement an RTSP app client that produces the sound. See Figure 2. Additional phones can be recruited to support the channel playback based on location information, e.g. see [4].

Phases of organization

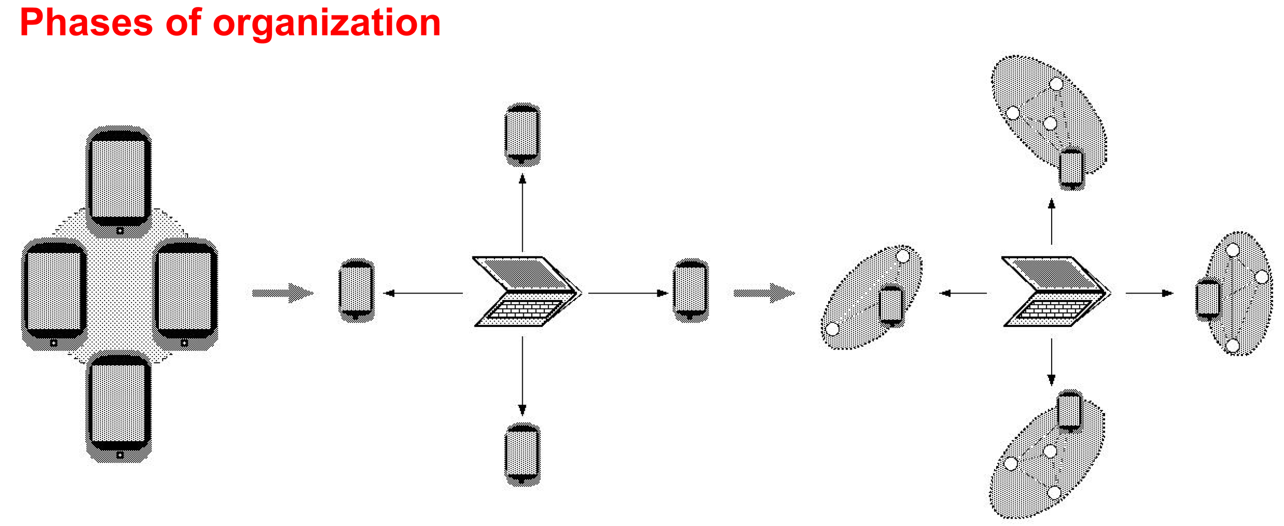
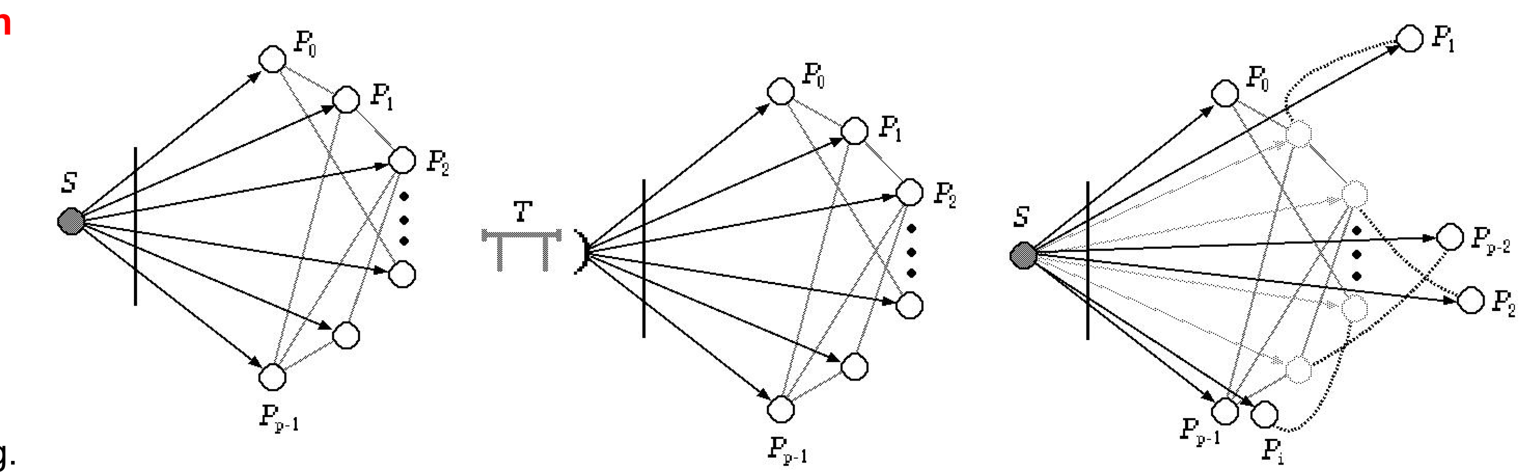


Figure 2.
Phase 1: Master nodes are designated.
Phase 2: Master nodes are assigned audio channels based on positions relative to the server.
Phase 3: Additional phones can be recruited into piconets to support channel playback.

Future project development based on framework

Figure 3. The framework provides infrastructure support for subsequent projects in the programme:
i. **current framework**: business and multimedia applications;
ii. audio-based game using smart-table;
iii. social application and education gaming.



Experimental Performance

Sound Field Quality

The sound playback on each phone is limited by the quality of the speakers. Loudspeakers on phones are generally designed to reproduce the mid-range frequencies (i.e. human voice) at the expense of having a wider frequency range. Given the size of the speaker in smartphones, the reproduction of lower range (bass) frequencies is particularly problematic. Volume is likewise limited by the speaker size, however with add-on portable speakers like those by Livespeakr offer a potential solution to both of these problems (to some extent). Further experiments with different compression technologies and file sizes will follow the present prototype.

Audio Synchronization and Phasing Issues

The synchronization of audio files across an array of phones poses interesting challenges. While as mentioned our use of the RTSP protocol has enabled the temporal synchronization of files, depending on the positioning of the smartphones, a variety of interactions between the sound playing through the phones may occur. We anticipate that users will not remain in a static location, and therefore there may be sound wave constructive or destructive interference, phasing effects, and potentially even Doppler shift depending on the velocity and location of the phones and listeners. These problems have not been a consideration for the present prototype, but the authours acknowledge that further exploration with simultaneous microphone capture on the phones may help to alleviate some of the potential acoustical issues.

Acknowledgments

The authours wish to acknowledge the generous research sponsorship of Google, Inc.

Portions of this research were submitted to the Wireless Personal Multimedia Communications Symposium (WPMC'11).

Conclusions and Future Work

A prototype system framework has been developed that demonstrates the feasibility and utility of the paradigm. As the sound quality from the speakers has significant impact of the user experience, future experiments will consider incorporating small low-cost USB speakers to augment the phones. Subsequent research phases will focus on exploring more complex audio games, with mobile phones as key components (see Figure 3):

- peripheral computing devices such as smart-tables could be used to drive an educational or game application where smartphones provide support for the audio component, e.g. see [5];
- phones could be used to playback streamed audio based on their location in a room relative to objects, other phones, etc.

Literature

[1] AES, "Best Practices in Network Audio," *J. of Audio Engineering Society*, Vol.57, No.9, pp.729--741, 2009.
[2] Gartner Inc., "Worldwide mobile phone sales grew 17 per cent in first quarter 2010," May 2010. [On-line]: Available: http://www.gartner.com/it/page.jsp?id=1372013.
[3] Internet Engineering Task Force (IETF), "Real Time Streaming Protocol (RTSP)," Apr. 1998. [On-line]: Available: http://www.ietf.org/rfc/rfc2326.txt.
[4] X. Li, "Collaborative localization with received-signal strength in wireless sensor networks," *IEEE Trans. on Vehicular Tech*., Vol.56, No.6, pp.3807--3817, 2007.
[5] A. S Shirazi, et al., "Poker surface: Combining a multi-touch table and mobile phones in interactive card games," in *Proc. of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2009*, Bonn, Germany, Sept. 2009.

For further information

Contact *p.taillon@uwaterloo.ca* or *collinsk@uwaterloo.ca*.

More information on this and related projects can be obtained at *http://www.gamessound.com/index.htm*.