SONDI: Audio-based Device Discovery and Pairing for Smart Environments

Hannu Kukka
University of Oulu
Department of Computer Science and Engineering
Erkki Koiso-Kanttilankatu 3
FIN-90570 Oulu, Finland
hannu.kukka@ee.oulu.fi

Pauli Marjakangas
University of Oulu
Department of Computer Science and Engineering
Erkki Koiso-Kanttilankatu 3
FIN-90570 Oulu, Finland
pauli.marjakangas@student.oulu.fi

NordiCHI '14, Oct 26-30 2014, Helsinki, Finland
ACM 978-1-4503-2542-4/14/10.
http://dx.doi.org/10.1145/2639189.2670275

Abstract
In this paper we propose a system called SONDI that uses high frequency audio signals (called audio signatures) to pair mobile devices with fixed devices in smart environments. The system allows users to discover interaction possibilities in the environment they might have otherwise missed, through unobtrusive and non-audible signals sent from fixed devices. Benefits of SONDI include fast discovery times (<1.8 seconds), effortless interaction from the user, and high availability as SONDI does not require any additional hardware on the users' mobile devices.

Author Keywords
Device discovery; device pairing; audio; interaction;

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction
We present a novel system for device discovery and pairing using so-called *audio signatures*. The system, called SONDI, allows mobile clients to serendipitously encounter fixed smart devices in the environment and proactively propose pairing to users. The basic premise of the system is as follows: A fixed device broadcasts a

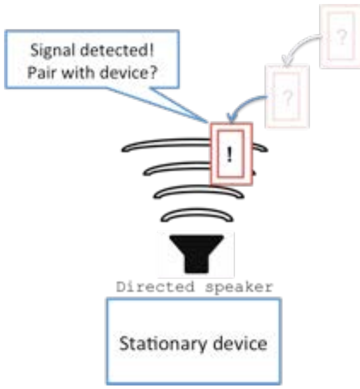

Figure 1. SONDI Concept

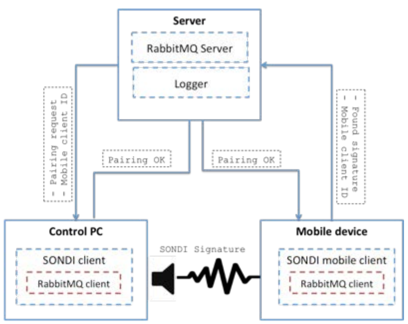

Figure 2. Technical components

unique audio signature to its immediate vicinity on a frequency that is inaudible to the human ear using a directed speaker. The mobile client listens to these signatures, and when a signature is detected, we can determine that the user is *a)* close enough to a device for pairing; and, since we're using a directed speaker, *b)* the user is in front of the device, as opposed to being *e.g.* behind it. The position of the user is important in cases such as public displays, which naturally are only useable if the user can see the screen. Such positioning is difficult to realize with undirected signals such as Bluetooth. After the mobile client recognizes a signature, the user is notified via tactile or audio feedback that a device supporting pairing is close by, and the user can then decide whether or not s/he wants to go ahead with the pairing (figure 1).

Related Work

Device pairing methods can roughly be divided into proactive and reactive methods. In the former, the pairing procedure is initiated by an entity other than the user's personal mobile device. As an example, Bluetooth hotspots can be used to continuously scan the environment for devices, and push information to mobile clients when they are in range. This approach has been used for *e.g.* Bluetooth-based advertisement [1, 6], or topical information delivery [7]. Reactive pairing approaches, on the other hand, require the user the initiate pairing and subsequent data exchange by e.g. scanning a visual code (for example [2]), or performing some other physical activity such as "scratching" a printed barcode [4].

A difficulty with reactive approaches, such as visual codes, is that they require quite many steps from the user. The user has to first become aware that s/he is in the vicinity of a device that supports pairing; become motivated enough to approach the device and initiate the procedure; oftentimes launch a mobile application such as a barcode reader; and scan the presented visual code, after which pairing can be initiated. Previous research has found that people may get confused about how to interact with a visual code, and for example try to click on the code instead of scanning it [8]. Proactive approaches, such as Bluetooth can suffer from long latency in device discovery, and users may also get confused as they have no way of knowing when they are within range of the Bluetooth transmitter [7].

Audio has previously been used for several prototypes and concepts in both mobile and ubiquitous computing research. Smith *et al.* [10] used ultrasonic pulses coupled with wireless RF signals for indoor positioning of moving mobile devices. Similarly, the Active Bat location system presented by Harter *et al.* [5] utilized ultrasonic "chirps" and small sensor tags carried by the users to track their location in indoor spaces. Peng *et al.* [9] used acoustic "beeps" for measuring the distance between two mobile devices. Goodrich *et al.* [3] used the audio channel to securely exchange public keys between devices where one device broadcasts it's public key encoded as audio, and the other decodes and interprets the audio to locate the key. Audio has also been used to encode information in Digital Audio Watermarking (DAW) (e.g. [11]). DAW systems are similar to SONDI because in both, a digital signal is hidden and subsequently retrieved using *e.g.* a mobile phone. Difference is, in DAW the signal is hidden in another piece of audio, whereas SONDI does not have

SONDI Components

The SONDI ID Generator is stand-alone program implemented used to create unique audio signatures for the system.

The SONDI Client is responsible for both broadcasting the audio signature, and responding to incoming pairing requests from the SONDI Server.

The SONDI Mobile Client listens to audio signatures, and once a signature is found, contacts the SONDI client in question. If the device in question is available the mobile client notifies the user and prompts pairing

The SONDI Server handles pairing requests. After a mobile client has recognized and decoded a SONDI signature, the mobile client contacts the SONDI server with it's own ID and the ID received from the signature.

to hide the signal except to make it inaudible to the human ear, while still keeping it distinguishable from the ambient background noise.

Technical Implementation

The SONDI system comprises several related components, as shown in figure 2. The main components are the SONDI ID generator, responsible for generating unique audio signatures; the SONDI Mobile Client, which listens to the audio signatures and presents the main entry point to the system; the SONDI client, which broadcasts the audio signatures and listens to incoming pairing requests from mobile clients; and the server, which manages addressing, messaging, and logging between the various other components of the system. These components are described in the sidebar. SONDI uses an asynchronous messaging broker to route messages to and from different clients. This allows for a more robust and fault-tolerant implementation, where clients can enter and leave the system without requiring extensive manual reconfigurations or addressing.

SONDI uses audio signatures that are created by segmenting a sound wave with the duration of 354 milliseconds into 13 frequency bands. The 19kHz frequency band is allocated as pilot signal used to correct for the Doppler effect, and the rest of the bandwidth is allocated to data, altogether 12 frequencies with 80Hz gaps between each frequency. Each frequency accounts for one bit, where an active frequency indicates a logical 1 and absence of frequency implies logical 0. The signature files are created using a sampling rate of 44.1 kHz (mono), which matches the typical recording capabilities of mobile devices. SONDI uses Hamming (4 12) coding for linear error-correcting by segmenting the 12 total bits into 8 data bits and 4 parity bits used for error correcting.

In the current implementation, the SONDI client broadcasts back-to-back signature waves, each lasting 6 seconds. When in active scanning mode, the mobile client continuously runs so-called recording cycles, each lasting 0.9 seconds. During each cycle, the mobile client records one sample lasting 0.354 seconds from the signal, and the rest of the cycle is used to perform signature detection on the captured sample and, in case a signature is captured, initiate a second recording cycle for confirmation. If no signature is captured, a new recording cycle is initiated. Maximum total detection time of 1.8 seconds was implemented because SONDI is designed for mobile devices that move about it a given space; therefore, given the average human walking speed of about 5 km/h or 1.39 m/s, we need to detect a signature when a person either directly approaches the audio source, or is passing it at a given angle.

Results from our initial testing show that SONDI performs well under normal office conditions, and also when pre-recorded ambient traffic noise is introduced. SONDI works reliably up to distances of 5 meters from the audio source (likely further, but we limited measurements to 5m). Figure 3 shows signature recognition percentages from 0.5m to 5m in half meter intervals using two different smartphones (Samsung Galaxy Nexus and S4 Mini). From the figure we can see that the larger, more high-end Galaxy Nexus works more reliably across the distance range, but the differences between the two devices mostly fall within 1%. Figure 4 shows the effect of added ambient traffic

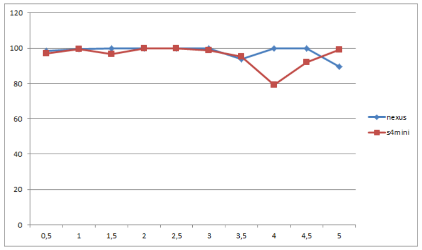

Figure 3. Signature detection rates

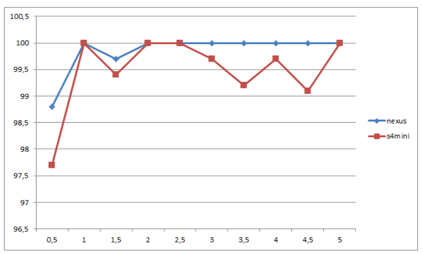

Figure 4. Signature detection with added ambient noise

noise used to simulate interference. Again, both phones detect signatures at around 100%, although the S4 mini had a slight performance drop at the distance of 4 meters which requires further investigation.

SONDI Demonstration

We will demonstrate the capabilities of SONDI by deploying an interactive public display with a control PC and a speaker in the demo session. We will have several mobile phones running the SONDI mobile client available for visitors. As users carrying SONDI-enabled devices roam around the display, the devices will utilize haptic feedback (vibration) to notify the users every time a signature is recognized; the users may experiment on the range of detection, on the effect of movement, or with how various common materials such as fabric effect the signature recognition when the phone is carried in *e.g.* a pocket or a purse. The display will be used to demonstrate the multi-user capabilities and pairing process of SONDI.

References

1. Aalto, L., Göthlin, N., Korhonen, J., & Ojala, T. (2004). Bluetooth and WAP push based location-aware mobile advertising system. *Proc. MobiSys '04, 49-58*
2. Alt, F., Kubitza, T., Bial, D., Zaidan, F., Ortel, M., Zurmaar, B., ... & Schmidt, A. (2011). Digifieds: insights into deploying digital public notice areas in the wild. *Proc. MUM '11 165-174*
3. Goodrich, M. T., Sirivianos, M., Solis, J., Soriente, C., Tsudik, G., & Uzun, E. (2009). Using audio in secure device pairing. *IJSN, 4(1), 57-68*
4. Harrison, C., Xiao, R., & Hudson, S. (2012). Acoustic barcodes: passive, durable and inexpensive notched identification tags. *Proc. UIST '12, 563-568*
5. Harter, A., Hopper, A., Steggles, P., Ward, A., & Webster, P. (2002). The anatomy of a context-aware application. *Wireless Networks, 8(2/3), 187-197.*
6. Kukka, H., Kruger, F., & Ojala, T. (2009). BlueInfo: Open architecture for deploying web services in WPAN hotspots. *Proc. ICWS '09, 984-991*
7. Memarovic, N., Langheinrich, M., Cheverst, K., Taylor, N. and Alt, F. (2013) P-LAYERS – A layered framework addressing the multi-faceted issues facing community-supporting public display deployments. *ACM Trans. Comput.-Hum. Interact. 20, 3, Article 17*
8. Peng, C., Shen, G., Zhang, Y., Li, Y., & Tan, K. (2007). Beepbeep: a high accuracy acoustic ranging system using cots mobile devices. *Proc. SenSys '07, 1-14*
9. Sharifi, M., Payne, T., & David, E. (2006). Public display advertising based on Bluetooth device presence. *Proc. MIRW '06*
10. Smith, A., Balakrishnan, H., Goraczko, M., & Priyantha, N. (2004). Tracking moving devices with the cricket location system. *Proc. MobiSys '04, 190-202*
11. Löytynoja, M., Keskinarkaus, A., Cvejic, N., & Seppänen, T. (2008). Watermark-enabled value added services to broadcast audio. *Proc. DEST '08, 388-396*