Mobile Music Making

Atau Tanaka
Sony Computer Science Laboratory Paris
6 rue Amyot
F-75005 Paris FRANCE
+33-1-4408-0512
tanaka@csl.sony.fr

ABSTRACT
We present a system for collaborative musical creation on mobile wireless networks. The work extends on simple peer-to-peer file sharing systems towards ad-hoc mobility and streaming. It extends upon music listening from a passive act to a proactive, participative activity. The system consists of a network based interactive music engine and a portable rendering player. It serves as a platform for experiments on studying the sense of agency in collaborative creative process, and requirements for fostering musical satisfaction in remote collaboration.

Keywords
Mobile music, wireless ad-hoc networks, peer-to-peer, Creative Commons.

1. INTRODUCTION
Music is by nature a social art. Whether in songwriting, recording, or performing, collaborative acts are at the heart of the musical creative process, independent of technology. Networks infrastructures, meanwhile, have fostered the growth of cooperative computing and groupware.

Use of sensors on mobile devices have been explored for collaborative computing [1], remote monitoring [2], and for interfacing the real world to the World Wide Web [3]. PDA's with wireless network connections have been used to indicate presence and exchange of music playlists [4]. We build upon this work to propose a system for collaborative musical creation on mobile wireless networks. The work extends peer-to-peer dynamic to go beyond simple file sharing to streaming and ad-hoc mobility. It transforms music listening from a passive act to a proactive, participative activity. The system consists of a network based interactive music engine and a portable rendering player. It serves as a platform for experiments on studying the sense of agency in collaborative creative process, and defining requirements for fostering musical satisfaction in remote collaboration.

The system consists of three main components: 1) a handheld device that is the main input/output hardware 2) a generative music engine, and 3) network services that manage authentication and media delivery. These three components work over ad-hoc 802.11b wireless networks allowing multiple socially defined groups of friends to collectively create music, independent of any pre-existing network infrastructure.

2. MOBILE TERMINAL
The handheld device is a personal digital assistant (PDA) modified to be a mobile terminal device for music. It functions as an input device for user actions, as a graphical display interface, and audio rendering destination. It is network enabled using a standard 802.11b (Wi-fi) network card, permitting input, interface, and musical output to be synchronized in real-time with similar peers.

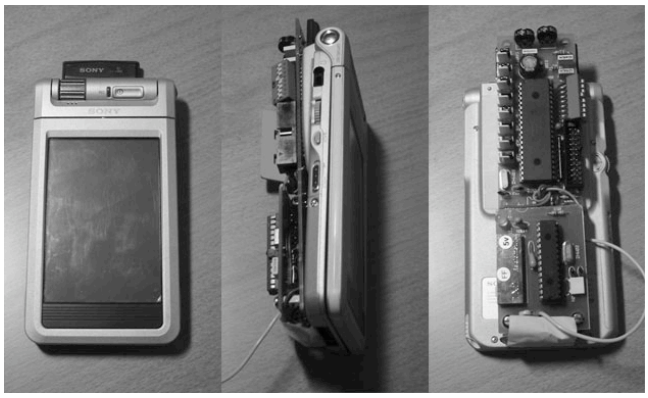

Figure 1. A PDA as mobile terminal with sensor subsystem

User input to the mobile terminal takes places via the touchscreen as well as by a sensor sub-system. The touchscreen captures *voluntary* input to the graphical user interface (GUI). System functions such as user log-in, as well as musical functions like slider manipulation take place over this channel. The sensor sub-system is a data-acquisition board operating in conjunction with the PDA, and allows *involuntary* input to the system. Force sensing resistors (FSR) capture grip pressure, while accelerometers sense gesture and motion in three-dimensional space. This channel allows expressive information more typically associated with musical instruments to be captured and sent up to the music generation engine.

The graphical user interface gives visual feedback to the user. The first screen displays the available group of "friends" within range of the wireless network. Once the user has joined a session, musical controls and network status are displayed.

The audio rendering engine is a network audio streaming client, capable of invoking multiple channels of MP3 format audio streams from the music generation engine. These multiple musical sources can be mixed by the onscreen sliders or by gestural input.

3. USER MANAGEMENT AND TRUST
The building of music making groups is handled by a suite of network services that manage acquaintance and access. It is based on a social metaphor of friendship and introduction, where security in the form of trust is a natural outcome [5]. A typical scenario for bootstrapping a music group is as follows.

A user's mobile device detects the presence of other users through Service Discovery protocols [6]. Once other potential partners are detected in range, the task turns to sorting the detected users by acquaintance level. Access permissions are managed by a social model – a "friend" is at a high trust level, and therefore has full access privelages. A "friend of a friend" is a less well known entity, slightly lower in trust, so may have some limitations in access permissions. A "friend of a friend of a friend" is again lower in trust, and might not have access.

We use public key encryption techniques to manage the acquaintanceship network. When two members become friends, they exchange certificates respectively signed by the other's public key. When one user goes to a new unknown members introducing himself as "a friend of a friend", the fact that the two users each have the public key of their common friend enables them to securely verify each other's identities.

These trust relationships are managed in a decentralized fashion, with no central authority or database. Each user keeps a record of his own social interactions. Parsing the list of keys that each user has in a given space allows the system to send to each user's display a representation of the acquaintanceship network. As there is no centralized database, the system is robust against attack. Even if a malicious user collected all the public keys he could, he would not be able to reconstruct the network of who knows who. In this way, our system is distinct from existing socially driven reputation systems [7]. Our solution is particularly suited for the spontaneous nature of group musical collaboration.

4. TRUST AND PERMISSIONS

By using the trust propagation properties of our social acquaintance system, we maintain the idea of making music with a group of friends. As the Internet is an open infrastructure open not only to attack but also to eavesdropping, the issue of permissions was fundamental. However, there was a desire to keep the process of bootstrapping a musical group as spontaneous and natural as possible. Instead of distributing passwords or managing accounts, each user simply discovers whom around him he knows. The theorem, "six degrees of separation" stipulates that with six levels of acquaintanceship, a path can be created between also any pair of people [8]. The interest of our system was not to see whom we know distantly, but to play music with who we know closely, while giving some access to those we know less well. We thus limited the number of acquaintance levels in our system to four, and defined the following access permissions:

Level 1: play music together

Level 2: listen to friends playing. Access to individual tracks of the participants, able to see the input of each member

Level 3: listen to result of friends playing. No access to individual tracks, no visual display

Level 4: no access

5. MUSIC GENERATION

The music generation engine is built using the MaxMSP real time music environment. Instead of the classic configuration of taking user input from the mouse/keyboard or MIDI controller to synthesize music to be output to a soundcard and speaker system, the input/output to the engine are the network. Gestural input from the group of connected users arrives via XML or OSC messages [9]. The engine reconciles the multiple control inputs to generate several parallel music channels that are streamed up to an Icecast type streaming audio server [10].

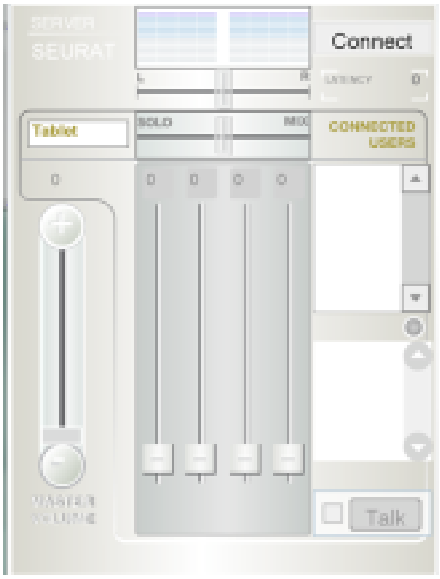

Figure 2. Onscreen graphical interface

Music pieces are conceived as open forms with elemental modules. Modules that make up a piece include rhythms, fragments of sequences, and samples. These elements are processed by time and frequency domain signal processing and combined to make a single musical flow. Treatments could include time stretching to reconcile differing tempos of the modules, filtering, or time domain re-ordering. The modules are submixed and assigned to different "tracks" that represent the intervention of individual connected users. The master mix represents the sum total of the different tracks.

The time re-ordering of elements in the music is applied at the rhythmic level and also at the structural level of a song. The low level re-sequencing allows user actions to intuitively create new rhythms and melodies based on existing elements in the music. The high level re-ordering allows the top level structure of the music to be *malleable*, making total song length flexible to match the corresponding social activity that drives the progress of the music.

These techniques are applied to standard popular songs and assume a constant meter and tempo. The temporal reconstitution uses techniques from sampler looping, allowing synchronization and assuring musical rhythm. At the same time, a second order temporal organization takes the music out of the repetitiveness of simple looping systems and permits an evolving high level structure for the song.

This approach is applied as a technique for creating new musical content for the system, but can also be used to *repurpose* existing songs, or *legacy* musical content.

The music is sent both in multitrack and mixed form up to the streaming audio server. They are accessible by the clients independently at different mountpoints of the streaming server. The ability of the listener to distinguish his own input within the total resulting music is a crucial element in musical agency discussed below.

6. COLLABORATIVE COMPOSING

Collaborative authoring systems exist in computer science and have been studied in sociological situations [11]. Music at first seems by default a collaborative process. But the compositional act remains a solitary activity. The famous music writing partnerships in the history of music have primarily been collaboration at the level of lyrics and music. There have been attempts to make network based musical groupware tools [12]. In most cases, music is created at each endpoint and uploaded for synchronization and reconciliation. Open source programming philosophy has been applied to cultural asset creation in the form of the Creative Commons [13]. In keeping the programming paradigm, there is a notion of versioning and incremental editing.

In the work presented here, a community of users contribute to the creation of a single musical stream in real time. It is not a merging of separate musical output, nor is it an incremental process of review, comment, and improvement. The dynamic we seek to create is a live musical interaction. In this sense, we seek to recreate as much as possible the qualities of music making itself.

7. AGENCY AND SATISFACTION

The network orientation of the system poses tangible problems in attaining this goal to recreate the dynamic of "real" music making,. Inasmuch as the participants are contributing to the evolution of a single piece of music in real time, the latency effects of network data transmission are inevitable. The question here is not whether latency can be eliminated, but in what ways it can be tuned to maintain musical satisfaction.

At the same time, remote collaboration allows users to span geographical separation. It is not that the distance is collapsed to zero, but that the network enables long distance collaboration. The sense of distance, then, needs to be apparent to the user. If distance somehow is not encoded in the participation of each member, not only is the appreciation of remote collaboration lost, but it becomes difficult to distinguish which musical part is contributed by whom [14].

These two needs, one of immediacy and the other of representing distance, seem to be diametrically opposed. One serves to provide the user a sense of *agency* for his own contribution to the music. The other serves to distinguish and give a *sense* to the remote partners' input. To attain musical satisfaction, these two needs must be met. This points out needs for the further development of the current rendering engine.

8. CONCLUSION

We have presented a network based hardware software system for group music making. The system exploits ad-hoc wireless networks and mobility to allow a community of users to participate in the real-time creation of a single piece of music. Socially defined permissions define levels of access to the creative process and resulting music. This extends existing notions of creative commons towards a dynamic real time activity. It seeks to encourage music listening not as a passive act of consumption, instead proposing a proactive participative activity. This is facilitated by detection of voluntary and involuntary gestures of the user on the mobile terminal. Tenets of musical instrument practice are applied to lay users. This raises issues of agency, representation, and musical satisfaction that will be addressed in future studies.

9. ACKNOWLEDGMENTS

This work was carried out with the participation of interns Guillaume Valadon and Laurent Violette.

10. REFERENCES

[1] Hinckley , K., "Synchronous Gestures for Multiple Users and Computers". *ACM UIST Symposium on User Interface Software & Technology*. 2003, 149-158.

[2] Milanovic N., et al. "Bluetooth Ad-hoc Sensor Network". *Proc IX Telecommunications Forum TELFOR 2001,* Belgrade, 2001.

[3] Kindberg, T., et. al., "People, Places, Things: Web Presence for the Real World". *HP Laboratories, Technical Report,* HPL-2000-16, Feb 2000. (http://www.cooltown.hp.com).

[4] Bassoli, A., Cullinan, C., Moore, J., Agamanolis, S. "TunA: A Mobile Music Experience to Foster Local Interactions", *Proceedings UBICOMP*. 2003.

[5] Tanaka, A., Valadon, G. *Social Mechanisms for Bootstrapping Trust in Wireless Networks.* Sony CSL Internal Report 2004-01. 2004.

[6] Rosenberg, J. et al. *SIP: Session Initiatation Protocol*. Internet Engineering Task Force RFC 3261, May 2002.

[7] Yu, B., Singh, M. P. "A Social Mechanism of Reputation Management in Electronic Communities". *Proc. 4th Intnl Workshop on Cooperative Information Agents*. 2000.

[8] Barabasi, A. *Linked: The New Science of Networks*. Addison-Wesley, Reading MA, 2003.

[9] Wright, M., Freed, A. "Open Sound Control: A New Protocol for Communicating with Sound Synthesizers". *Proc International Computer Music Conference*. 1997.

[10] Icecast streaming audio server. http://www.icecast.org

[11] Bentley, R., et al., "Basic Support for Cooperative Work on the World Wide Web". In *International Journal of Human-Computer Studies: Special issue on Innovative Applications of the World Wide Web*, 1997, 827-846.

[12] Jorda, S., Wust, O. "Architectural Overview of a System for Collaborative Music Composition over the Web". *Proc International Computer Music Conference*. 2001.

[13] Creative Commons. http://www.creativecommons.org

[14] Tanaka, A. "Musical implications of media and network infrastructures". *Proc H2PTM'01: Hypertextes Hypermédias*, Hermes Science Publications, Paris. 2001, 241-250.