Eavesdropping: Network Mediated Performance in Social Space

Jack Stockholm

ABSTRACT

The author describes an Internet-based audio composition and diffusion system, *Eavesdropping* (2007–2008), designed for public spaces where several computer users are gathered, such as cafés. Compositions are created from abstract mood objects rather than musical structures. A composer uploads a set of audio files to represent the different moods in the composition. During a performance, a server-based Conductor selects audio files from this set to be played at each participant's laptop based on the composition, the number of participants in the room and the time they joined the performance. This project aims to enhance awareness of and connectedness among individual members of an audience at a generative musical performance by encouraging shared experiences.

Live performance has migrated to social networks, providing an opportunity for interaction between musician and audience that was formerly unavailable in broadcast media. While recorded music and broadcast media removed the musician from the performance venue and often isolated audience members from each other, Internet-based audio arts are redefining the interaction among audience members, as well as with the artist, and providing new means of social engagement. The interactions available with Internet audio reflect many of the social interactions of the concert hall, but the specifics of the medium also offer a variety of new and global articulations [1]. Additionally, the ubiquitous nature of the Internet and the connectivity of Internet devices provide for a pairing of network connectedness and physical social spaces, allowing for overlap of both interactive methods [2].

Walkthrough and Motivations

Eavesdropping is an Internet-based audio composition and performance system I designed for public spaces where several computer users are gathered, such as cafés. Anyone can initiate a performance by visiting the web site, selecting a composition and typing in the name of their location. Compositions have been defined by composers in advance; their audio files have been uploaded already and are ready to be sequenced by the system in real-time. Once a performance is started, the initiator announces to café customers that an audio art performance is beginning and that anyone willing to participate should visit the web site and join the listed location. At this point the server will select and send different audio files to each participant's computer following the composition.

Participants hear a combination of the audio from their own computers and from all other participants' computers in the room. When a unique or interesting sound originates from a computer in the room, the other participants become aware of the person whose computer made the sound. Compositions are designed to capitalize on this personal association between a user and his or her computer by utilizing music that will focus participants' attention on each other. Visitors to a café, like an audience in a music venue, seek the passive awareness of others to achieve a sense of connectedness born from experience shared in proximity [3]. Though people in a café may be focused on their own work, they have chosen a public space in which to do it. This project highlights the voyeurism and exhibitionism of bringing private actions into the public sphere by increasing the exposure of participants to each other via diffusion of audio from their private computers.

Jack Stockholm (musician, researcher, educator), School of Interactive Arts and Technology, Simon Fraser University Surrey, 250–13450 102 Avenue, Surrey, BC V3T 0A3 Canada. E-mail: <jack_stockholm@sfu.ca>. Web site: <www.oddible.com>.

Figure 1 shows *Eavesdropping* participants at different tables interacting during a performance.

Eavesdropping is the first release of a platform designed for the sonification of the moods of a distributed group of people in a networked environment. The motivation is to develop an audio system that will raise awareness among individual, networked participants in the same physical space, to increase connectedness and to facilitate interaction by using audio samples to portray participants' moods. This initial implementation utilizes the platform as an art installation. Instead of capturing mood data from participants, it offers composers a tool to create compositions based on a sequence of moods rather than using formal musical figures. In this prototype, a web-based server selects audio files to match the mood-based composition by evaluating the number of participants, the time the participants joined the performance and the network latency in the environment. Each performance of a composition will differ due to these factors.

Background

There are many sonic art projects that have used interconnected musical networks in a variety of different roles [4].

Fig. 1. *Eavesdropping* participants interacting at Prado Café in Vancouver. (Photo © Jack Stockholm)

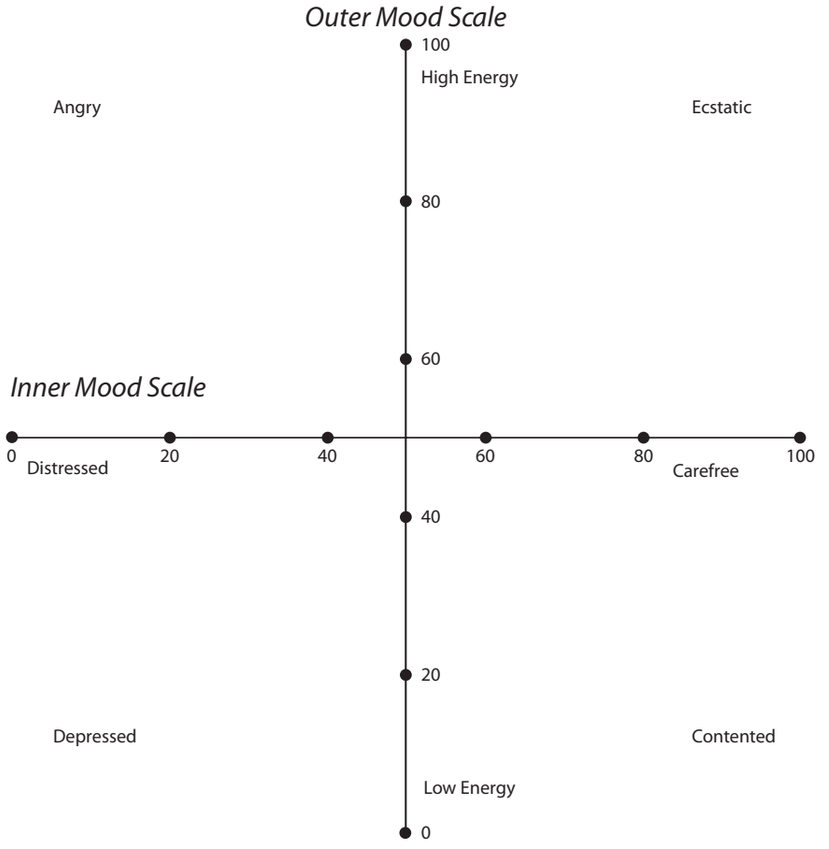


Fig. 2. Mood Matrix derived from the Mood Matrix by Eladhari et al. [13] and the Mood Model by Thayer [14]. (© Jack Stockholm)

Many early projects focused on multiple musicians in different locations collaboratively performing via network connections [5,6]. In these kinds of systems, the goal is to achieve a low-latency means to communicate the actions of the various musicians over the network. In *Eavesdropping*, the network does not act as a means of communication between performers, rather it provides audience connectivity to the host server, which acts as a conductor of the audio each participant will play. Perfect timing of the audio presented from each participant's computer is not a goal of *Eavesdropping*. In the social acoustic ecology of physical spaces, people make sounds at irregular moments; their lack of synchronicity often elicits a variety of interesting interpretations. Computational and network delay has been increasingly tapped as a performance element due to its high variability.

A key focus of *Eavesdropping* is that the networked system is designed for compositions to be performed in a localized environment. Similar projects offer an instrument-based approach to localized network performance by allowing musicians and audience alike to perform together in a collaborative sound space. In Chris Brown's *Talking Drum*, a server-based Conductor monitors input from microphones and generates collaborative audio to be performed via four speakers in a shared physical space [7]. Barbosa's *Public Sound Objects* provides users with a visual representation of sound objects on a screen that can be manipulated to affect the pitch, reverberation and amplitude, during their synthesis and playback in a public installation [8].

Composition and Performance

Eavesdropping offers a composition environment abstracted from formal musical representation by allowing composers to create compositions of moods. These moods are arranged in a piano-roll interface and are assigned to the various participants in the room. This research attempts to resolve issues raised in prior multi-user, collaborative mood ecologies [9] by applying a situation-aware, server-based software algorithm I call the "Conductor" to adjust the audio based on the number of participants. The composer uploads a variety of audio files to represent the moods in the composition, and the Conductor chooses which files to play to address issues of sound density, stream segregation and acoustic ecology [10,11].

Eavesdropping has three primary components for (1) uploading and tagging of audio files, (2) composing and (3) conducting a performance. The system was developed as an ASP.NET application in C# and is online and accessible via the project web site [12].

Uploading and Tagging Audio

A composer uploads MP3 files via the composition interface and encodes them with a variety of formal and abstract characteristics. These characteristics were developed to give the system basic information to associate files to moods as well as to address specific issues in layering multiple audio files. For instance, a variable named *Plurality* was created to specify on a linear numeric scale whether a file is dense or sparse, indicating whether it would be better played alone or with other samples. This allows the system to adapt to the number of participants by selecting minimal samples in cases where there are a large number of participants or selecting dense, full samples when the number of participants is small. Another variable, *Harmelodic*, represents a range that runs from harmonic background sounds to lead melodies, allowing the system a variety of play styles. Moods are identified via a two-variable mood matrix that ranges from low energy to high energy along an outer mood axis and from distressed to carefree along an inner mood axis (Fig. 2) [13,14].

A composer uploading files should be aware of some guidelines in designing audio for this system. First, the composer has no control over the timing of playback and therefore will not be able to specifically align beats. Second, the lack of pitch information in the representational data creates an environment where files in various keys could be combined. A composer should therefore design audio that aligns to a complimentary set of keys or a set of audio that explores a 12-tone range and that appeals to the generative possibilities of the system. Lastly, sufficient audio files should be included to cover a wide variety of potential characteristics, allowing the system to generate a variety of different performances.

Composition

The composition environment provides a piano roll–style interface for arrange-

ment (see Fig. 3). The piano roll segments are not associated with the audio files themselves but represent compositional elements encoded with characteristics similar to the audio files *Plurality, Harmelodic, Tempo, Timbre, OuterMood* and *InnerMood*. The web-based interface presents a timeline with which a composer can layer several compositional segments to shape the composition. For instance, the composer can indicate a long harmonic section to be played with several shorter melodic segments of alternating moods.

Once a performance is initiated, willing participants can visit the *Eavesdropping* web site and join the performance at their locations. The audio selection engine then picks audio files based on the compositional elements and sends one to each participant. Subsequent audio selections take into account the audio currently playing in the room. Participants who join the performance after it has already begun are added to the system and factored into the arrangement. A participant's web browser will continue to request new samples until the composition is complete.

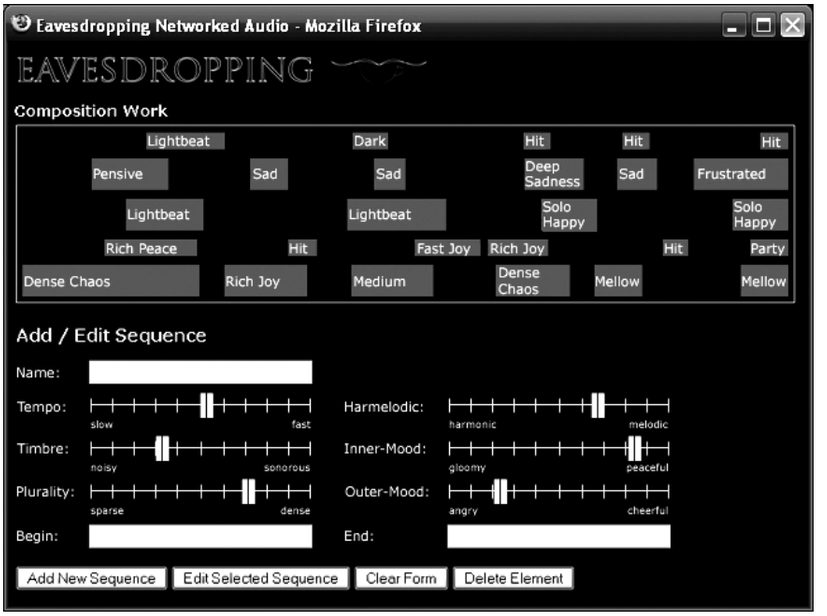


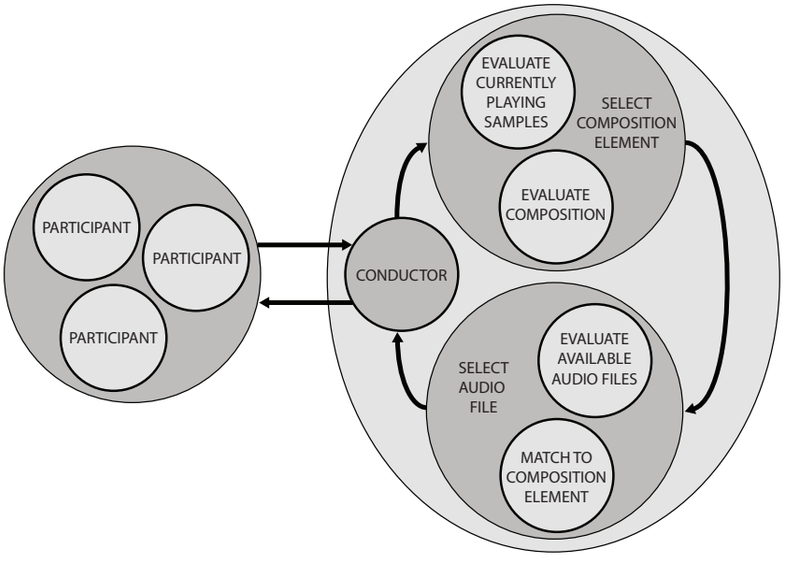
Fig. 3. *Eavesdropping* Composition Interface. (© Jack Stockholm)

Conductors

The Conductor is a decision module that is the centerpiece of the audio selection engine. Conductors are plug-ins to the system, and each has its own variations on how it analyzes the participant information and the composition to determine which samples to play. The use of Conductors allows experimentation with the decision-making logic without re-coding of the server system. During the development process several Conductors were created for the system in order to explore different audio selection methods. When a participant requests a file, the Conductor uses its specific algorithm to compare the files that are currently playing with the set of compositional elements assigned by the composer for that specific moment. Only one Conductor is active during a performance, and the composer can select the Conductor that will work best with his/her composition and audio. Figure 4 shows the basic file-selection process performed when a participant requests a new file.

In the following description, one simple and one complex Conductor are discussed. Each Conductor has two basic functions when a participant requests audio to play: selecting a compositional element to play and selecting an audio file that matches that compositional element. First the Conductor compares the audio files that are currently playing to the set of compositional elements in the piano roll on the timeline. It is frequently the case that there are more or fewer participants than compositional elements. The goal is to ensure that each compositional element will be represented in the audio played and that no compositional element will be overrepresented. When a Conductor selects a compositional element, the Conductor increments the play count for that element so it knows which elements have been selected in prior operations. The simple Conductor merely selects the oldest element with the lowest play count. The complex Conductor evaluates the individual characteristics to select an element. For instance, it evaluates for the current density of the composition by adding up the *Plurality* values of all the compositional elements at the current location in the piano roll as a target density. It then selects an element that, when added to the *Plurality* of all the audio files currently playing in the room, is closest to this target density. In choosing a compositional element,

Fig. 4. Conductor file selection process. (© Jack Stockholm)



the complex Conductor weighs certain characteristics over others by sequentially evaluating characteristics and narrowing a candidate set of potential elements in each operation.

Once the Conductor has determined a set of characteristics that will keep the composition on track, it then searches the audio-file database for a file that most closely matches these characteristics. This is the second primary function of the Conductor module. The simple Conductor performs a nearest-neighbor search by assembling a "select value" from the sum of the differences between characteristics of the target element and those of each of the audio files. The sample with the lowest select value that represents the lowest deviation from the target element is sent to the participant. The complex Conductor selects a best match for each characteristic, then evaluates the percentage difference of all the characteristics for the selected samples. The sample with the lowest sum of percentage difference is chosen and sent to the participant.

Performance and Conclusions

Overall, the engine has performed as expected, with audio delivered to various-sized groups, clearly shaped by the composition in mood, structure and density. In some instances it has been observed that the engine strays from the intended path when there are not enough compositional elements provided by the composer to allow the system to adjust to large numbers of participants. One possible solution would allow the Conductor the flexibility to create its own target element by defining characteristics that better aligned currently playing files to composed elements rather than being limited to the compositional elements provided by the composer.

Initial performances with *Eavesdropping* were successful at initiating conversations between disparate people in the performance environment. During the performance some issues did arise with participants demanding agency, the desire to act as more than a passive audience. Participants engaged in all sorts of actions that clearly expressed their intent to be involved in the performance. Many people turned their laptops around to face the other participants as if to be heard. Others opened multiple browser windows to the system so that their computers were playing multiple sessions. Still others opened music players on their machines and contributed outside sources of audio to the mix. Despite the fact that audiences are accustomed to passive listening, once participants' laptops are performing for the rest of the room, users want agency. A version of *Eavesdropping* is in development that allows users to input their own moods directly into the system, thereby influencing the composition to be performed in the room, providing an immediate venue for individual expression without creating significant distraction from the performance. User studies of both composers and audience using questionnaires and participant observation are being performed during Summer 2008 and will be published as part of my Master's Thesis work from the School of Interactive Arts and Technology at Simon Fraser University.

Acknowledgment

This research is partially funded by a commission from Rhizome.org with additional resources provided by Simon Fraser University.

References

1. A.J. Bongers, "Interaction in Multimedia Art," *Knowledge-Based Systems* **13**, No. 7–8, 479–485 (2000).

2. Jane McGonigal, "Super Gaming: Ubiquitous Play and Performance for Massively Scaled Community," *Modern Drama* **48**, No. 3, 471–491 (2005).

3. T. Erickson and W.A. Kellog, "Social Translucence: An Approach to Designing Systems that Support Social Processes," *ACM Trans. Comput. Hum. Interact.* **7**, No. 1, 59–83 (2000).

4. G. Weinberg, "The Aesthetics, History and Future Challenges of Interconnected Music Networks," *International Computer Music Association Conference,* Göteborg, Sweden, September, 2002 (2002) pp. 349–356.

5. J. Bischoff, R. Gold and J. Horton, "Music for an Interactive Network of Microcomputers," *Computer Music Journal* **2**, No. 3 (December 1978) pp. 24–29.

6. S. Gresham-Lancaster, "The Aesthetics and History of the Hub: The Effects of Changing Technology on Network Computer Music," *Leonardo Music Journal* **8** (1998) pp. 39–44.

7. C. Brown, "*Talking Drum*: A Local Area Network Music Installation," *Leonardo Music Journal* **9** (1999) pp. 23–28.

8. A. Barbosa, "Displaced Soundscapes: A Survey of Network Systems for Music and Sonic Art Creation," *Leonardo Music Journal* **13** (2003) pp. 53–59.

9. Mirjam Eladhari, Rik Nieuwdorp and Mikael Fridenfalk, *The Soundtrack of your Mind: Mind Music—Adaptive Audio for Game Characters* (Hollywood, CA, and New York, NY: ACM Press, 2006).

10. A.A.A. Ibrahim and A. Hunt, "A General HCI Framework of Sonification Applications," *ACM SIGACCESS Conference on Computers and Accessibility,* 23–25 October 2006 (Portland, OR: ACM Press, 2006) pp. 221–222.

11. W.W. Gaver, "What in the World Do We Hear? An Ecological Approach to Auditory Event Perception," *Ecological Psychology* **5**, No. 1 (1993) pp. 1–29.

12. See <www.oddible.com/cafe>.

13. See Eladhari et al. [9].

14. R.E. Thayer, *The Biopsychology of Mood and Arousal* (Oxford Univ. Press, 1989).

Jack Stockholm is an instructor and graduate student at Simon Fraser University's School of Interactive Arts and Technology. His research focuses on creativity in auditory representation for networked collaboration. He is a musician and Net artist who explores improvisation in audio interaction.