Eavesdropping: Network Mediated Performance in Social Space

Jack Stockholm

Live performance has migrated to social networks providing an opportunity for interaction between musician and audience 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 between audience members and with the artist to provide new means of social engagement. The interactions available in 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 connectivity of internet devices provides for a pairing of network connectedness and physical social spaces, allowing for overlap of both interactive methods [2]. This article explores the network audio project Eavesdropping, an internet-based audio system I designed to mix networked performance with live social interaction in a public space.

Eavesdropping is an internet-based, audio composition and performance system designed for public spaces where several computer users are gathered, such as a café. Anyone can initiate a performance by visiting the website, selecting a composition, and typing in the name of the location. The initiator then announces to café customers that an audio art performance is beginning and that anyone willing to participate should visit the website and join the listed location. At this point the server starts sending different audio samples to each participant's computer based on the composition. Participants hear a combination of the audio being played from their own computer and from all other participants' computers in the room. Compositions are designed to capitalize on personal association with the sounds emanating from one's own computer. When a unique or interesting sound originates from one computer in the room, the other participants become aware of the person whose computer made the sound. This project highlights the intentionality and exhibitionism of bringing private actions into the public sphere by increasing shared experiences in the environment. Visitors to a café, like the audience in a music venue, seek the passive awareness of others to achieve a sense of connectedness born from shared experience in proximity [3].

Figure 1. Eavesdropping participants interacting in a café.

This project utilizes a prototype system designed for sonification of the moods of a distributed group of people in a networked environment. The motivation is to develop an audio system which will raise awareness between individual, networked participants in the same physical space, to increase connectedness and to facilitate interaction, by using audio samples to portray participants' moods. Eavesdropping utilizes this system as an art installation allowing musicians to create compositions as a sequence of layered moods in a localized, multi-channel environment. The prototype web 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.

There are many sonic art projects which have used interconnected musical networks for a variety of different roles [4]. Many early projects focused on multiple musicians in different locations collaboratively performing via network connections [5, 6]. In these 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, but provides audience connectivity to the host server which acts as a conductor for the audio each 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 [7].

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 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 localized environment [8]. Barbosa's *Public Sound Objects* provides users with a visual representation of sound objects on a screen which can be manipulated to affect the pitch, reverberation and amplitude, during their synthesis and playback in a public installation [7].

The Eavesdropping project 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 Conductor to adjust the audio based on the number of participants. The composer uploads a variety of audio 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].

Uploading and Encoding Audio

The Eavesdropping project has three primary components: an audio composition environment, an audio selection engine, and a performance interface. The system was developed as an ASP.NET application in C# and is online and accessible via the project website (www.oddible.com/cafe).

Composers with special access upload mp3 files via the composition interface and encode 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 identify on a linear numeric scale whether the 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 when there are a large number of participants or selecting dense, full samples when the number of participants is few. Another variable, *Harmelodic*, represents a range that runs from harmonic background sounds to lead melodies, allowing the system some control in variety of play style. Moods are identified via a two variable mood matrix which ranges from angry to exultant along an outer mood axis and from depressed to happy along an inner mood axis (Figure 1) [9].

These values are initially input by the musician uploading the files, though these characteristics introduce some arbitrary mappings which may not maintain consistent meaning amongst composers or audience. In order to improve the validity of these mappings, a future version of the system is in development with a reinforcement learning system built into the performance interface to fine tune these representations.

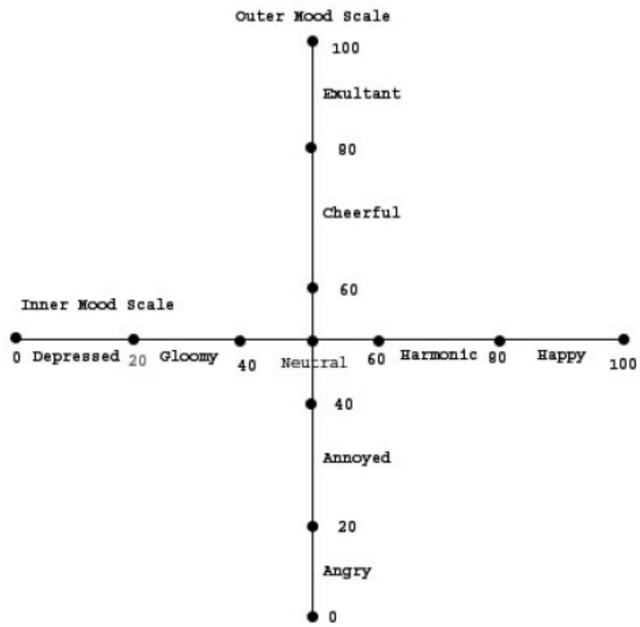

Figure 2. Mood Matrix dèfined by Eladhari, et Al.

One variable which has been specifically omitted is key or pitch information. The addition of pitch-based variables introduces an undesirable focus on key-based composition rather than the original intent of a more abstract meta-composition environment based on moods and density patterns.

Musicians uploading files should be aware of some guidelines in designing audio for this system. First, the composer will have 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. Musicians should therefore design audio that aligns to a complimentary set of keys, or a set of audio that explores a 12-tone range and which appeals to the generative possibilities of the system. Lastly, enough audio files should be included to cover a wide variety of potential characteristics. Compositions could be designed for a specific set of audio, but ideally audio sets may be swapped into different compositions as if skinning the composition with the musical style of the audio set.

Composition

The composition environment provides a piano roll-style interface for arrangement. 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, InnerMood*. The web-based interface presents a timeline in which composers can layer several compositional segments to shape the composition (Figure 2). 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, all participants in the room visit the Eavesdropping website and join the performance at their location. The audio selection engine then picks audio files based on the compositional elements and sends one to each participant. Subsequent audio selection takes 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. The participants' web browsers will continue to request new samples until the composition is complete.

Conductors

The centerpiece of the audio selection engine is a decision module called a 'Conductor.' 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. When a participant requests a file, the conductor uses its specific algorithm to compare the files which 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 which conductor works best with their composition and audio.

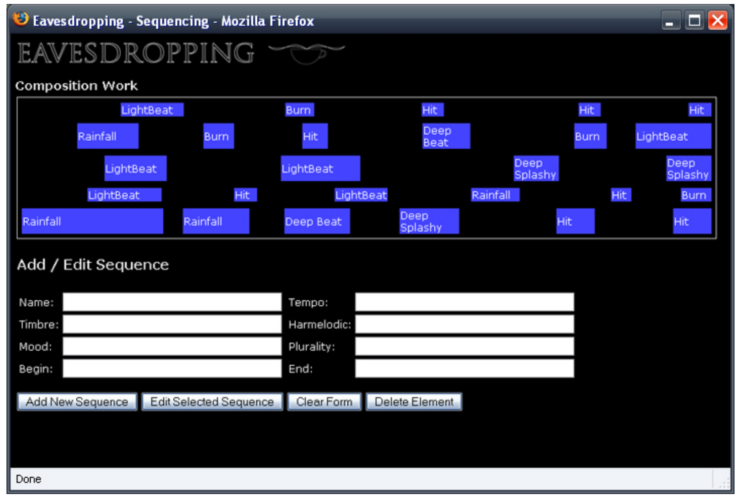

Figure 3. Eavesdropping Composition Interface.

This article discusses two conductors, defined as simple and complex. Each conductor has two basic functions, selecting a compositional element to play and selecting an audio file which matches that compositional element. First the conductor must evaluate the audio files that are currently playing in comparison to the set of elements that has been defined by the composer in the piano roll at that location on the timeline. When a conductor selects a compositional element it increments the play count for that element so it knows which 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 which 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, 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 which will keep the composition on track, it then searches the audio file database for a file which most closely matches these characteristics. This is the second primary function of the conductor module. The simple conductor assembles 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', which represents the lowest deviation from the target element, gets 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.

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 the compositional elements provided by the composer do not allow enough variety for 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 which 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, however an issue of user agency manifest both immediately during the performance as well as in user frustration expressed during subsequent question and answer sessions. During the performance, participants engaged in all sorts of actions which 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 an audience is accustomed to passive listening, once their laptops are performing for the rest of the room, users want agency. A version of Eavesdropping is in development which allows users to directly input their own moods into the system to be performed in the room, providing an immediate venue for individual expression without creating significant distraction from the performance. User studies of both musicians and audience are planned for summer 2008.

This research is partially funded by a commission from Rhizome.org with additional resources provided by Simon Fraser University. You can view Eavesdropping online at http://www.oddible.com/cafe.

References

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

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

3. T. Erickson and W. A. Kellogg, "Social translucence: an approach to designing systems that support social processes," *ACM Trans.Comput.-Hum.Interact.,* Vol. 7, No. 1, 59-83 (2000).

4. Weinberg,G. The aesthetics, history and future challenges of interconnected music networks. In *Proceedings of the International Computer Music Association Conference* (Goteborg, Sweden, Sept. 2002). 2002, 349-356.

5. Bischoff, J., Gold, R. and Horton, J. Music for an interactive network of microcomputers. *Computer Music Journal 2, 3* (Dec, 1978), 24-29.

6. Gresham-Lancaster, S. The aesthetics and history of the hub: The effects of changing technology on network computer music. *Leonardo Music Journal 8 (*1998), 39-44.

7. Barbosa, A. Displaced soundscapes: A survey of network systems for music and sonic art creation. *Leonardo Music Journal 13 (*2003), 53-59.

8. Brown, C. Talking drum: A local area network music installation. *Leonardo Music Journal 9 (*1999), 23-28.

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

10. Ibrahim,A. A. A. and Hunt,A., A general HCI framework of sonification applications. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility (*Oct. 23-25, 2006). ACM Press, Portland, Oregon, USA, 2006, 221-222.

11. Gaver,W. W., What in the world do we hear? an ecological approach to auditory event perception. *Ecological Psychology 5,1 (*1993). 1-29.