MoNoDeC: the Mobile Node Controller for audience-involved sound diffusion

Nick Hwang
University of Wisconsin-Whitewater
hwangn@uww.edu

Anthony T. Marasco
Michigan State University
tmarasco125@gmail.com

ABSTRACT

We present MoNoDeC, a multichannel audio system designed for immersive audio experiences by leveraging audience mobile phones and Internet of Things (IoT)-driven speakers. This network-based system uses audience mobile devices and speakers as point-source audio and dynamic audience participation through their mobile displays, allowing users to interact with and influence performances. MoNoDeC enables designers to craft interactive audio environments where audience members can modify musical elements, contribute to a collective canvas, or change instrument timbres. This paper highlights MoNoDeC's ability to foster engaging, immersive, and interactive musical experiences beyond traditional loudspeaker systems, showcasing its potential for interactive audience-based performances and sound diffusion.

1. INTRODUCTION

MoNoDeC, short for "Mobile Node Controller," is a multichannel audio system that utilizes audience mobile phones and IoT-driven speakers as point sources in immersive audio experiences. MoNoDeC is a network-based system currently with two client use cases: audience mobile phones and Raspberry Pi-driven audio devices, allowing for interactive audience-based performances, mobile and connected audio sound diffusion system [1]. Audience phones allow for dynamic interfaces for participatory elements during performance.

Audience participants use their mobile phones to register their location within a customizable audience space (rows, cloud, or free-form), turning their devices into point sources during performances or installations. Designers can choose for mobile users to move about during a performance or experience.

A performance controller manages the state, audio output, and interface data delivery to participants during the performance stage. The controller selects performance parameters and shapes the sound from the mobile devices through mapped touch gestures. Audience members interact with the mobile interface, affecting the experience by modifying musical aspects, contributing to a collective canvas, or altering instrument timbres.

MoNoDeC engages the audience in immersive and interactive experiences and performances unattainable with traditional loudspeaker setups. This paper explores the design, functionality, and capabilities of MoNoDeC as a sound diffusion system and shows examples of the dynamic interfaces it can present to mobile users during performance.

2. PRIOR SYSTEMS & HISTORICAL CONTEXT

Techniques for assembling speaker arrays from audience-provided or readymade devices have been present since the mid-20th century. As radios became more ubiquitous and portable, composers experimented within the limits of this technology to write works around these media players turned instruments. Composers writing new works for radios explored differing approaches to culling sound from the airwaves [2]. Through the advent of the Web Audio API and mobile devices, the evolving social expectation that audience members will be equipped with a device capable of receiving audio and/or control data over the internet has allowed composers to push the boundaries of complexity and collaboration experienced in works for distributed musical networks.

The past decade has seen the rise of frameworks aimed at streamlining the creation of musical networks for installations and live performances. Leveraging the built-in sensors and audio processing systems in mobile phones and single board computers [3], these frameworks allow artists to use mobile devices as both gestural controllers for shaping musical content in performance and portable speakers integrated into the concert hall. Prior research shows a collection of client-server systems similar to MoNoDeC, some developed for conducting specific studies centered on audience participation [4] and others designed to aid in creating new distributed music works. The Soundworks framework [5, 6] allows creators to develop networked musical works featuring bidirectional audio and data transmission between client devices. Both SoundSling [7] and Nü Soundworks [8] combine audience member web browser interfaces with a controller made in the Max programming environment to explore audio diffusion in 1-to-*n* and *n*-to-*n* layouts respectively. Focusing on collaborative improvisation with networked mobile devices, the AuRal framework [9] and the Daisyphone software [10] use a shared music composition workspace and relative geolocation to turn data transmission and local audio generation into communal music-making experiences.

MoNoDeC differs from preexisting systems in a handful of ways. The current version of MoNoDeC's user-facing software elements is web-based allowing its user interfaces to provide dynamic visual feedback for the mobile user that many existing systems do not. MoNoDeC aims to help newcomers develop and deploy projects quickly by including starting templates, preset speaker node layouts, and the ability to save frequently used settings. Additionally, the option to add customizable IoT speaker nodes into a setup is simplified by providing firmware that works to connect embedded computers to the MoNoDeC server.

3. MONODEC

The main purpose of MoNoDeC is to use audience members' mobile phones as point-source sound diffusion speakers. MoNoDeC is also capable of using custom-made IoT devices that we call autonomous hubs as point sources and standalone digital instruments as well. Additionally, MoNoDeC is capable of presenting dynamic interfaces to audiences members which may include providing performance affecting options or digital instruments localized on their own device. MoNoDeC can be used independently as an immersive audio system with sound diffusion or in conjunction with existing paradigms such as loudspeaker orchestras or ambisonic arrays.

MoNoDeC is built with existing communication software created by the authors (Collab-Hub)[1] that uses a web-based server-client framework allowing for two-way communication between web-pages, apps, and IoT Devices. MoNoDeC's interfaces for the controller and audience members are served through web pages or apps, and the autonomous hubs connect directly to the server. Dynamic interfaces for controller and audience provide active participation options during performance. The autonomous hubs are intended to be stationary. Collab-Hub enables the communication pathways of MoNoDeC. For ease of connectivity and setup, currently MoNoDeC relies on the Wide Area Network (WiFi internet) to communicate with the server and audience members' phones, although alternative setups could use a localized version of the MoNoDeC / Collab-Hub server to reduce possible latency.

3.1 Visual Rendering

The controller and audience mobile phone interfaces are rendered with HTML, CSS, and JavaScript. HTML5 Canvas objects render the interactable layouts and instrument control elements with P5.js.[2]. The state of MoNoDeC's data and phases are stored on the remote server.

3.2 Phases

MoNoDeC has three phases which are managed by the controller: layout and design, registration, and performance:

3.2.1 Layout and Design Phase

Layout and Design Phase allows the controller to make decisions on the layout of the available audience spaces. These audience spaces represent predetermined seating for the performance, where during the Registration Phase audience members select their seats. Visual landmarks like the Controllers table / mixing booth, stage, or easily noticeable physical object in the performance help orient the audience to the graphical layout. These graphical landmarks can be added during this Layout and Design Phase. The controller may select from premade or saved layouts.

[1] https://www.collab-hub.io/
[2] https://p5.js

This phase can occur anytime before a performance; the layouts are saved on the server and can be recalled.

3.2.2 Registration Phase

Registration phase occurs prior to performance and once the audience is seated. During this phase audience members' interfaces display the phase name and the interactable interface displays the selectable seat options on their mobile device. Audience members are instructed to select their seat. Upon selection, their seat graphic will change colors and new text will instruct audience members to await the next phase.

The controller's interface displays the entire layout with seats with colors indicating unregistered (blue) and registered colors (red). During this phase the controller is unable to change the layout, but is able to return to Layout and Design Phase. If the controller returns to the previous phase, all audience member's selections reset. The controller is able to advance to the Performance Phase once audience members have finished selecting their seats.

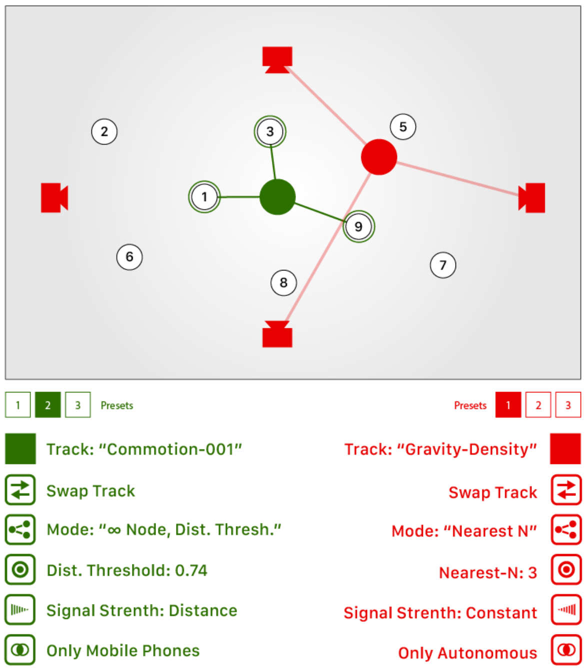

Figure 1. Controller interface during *Performance Phase*.

3.2.3 Performance Phase

The Performance Phase begins when the controller starts a new performance. All connected devices (mobile phones and autonomous hubs) have registered their positions and interfaces remind audience members to unmute and turn their volumes up.

All devices, mobile phones, and autonomous hubs, are considered 'nodes' in the 'performance space'.[3]

[3] Due to current web-based interaction and iOS/Android programming limitations, audiences must unmute their devices. One side effect to this for MoNoDeC is notifications, alerts, and reminders from such devices may potentially sound during performances. At the time of this writing, the authors have no programming solution, but a workaround is to ask the audience to unmute and turn off all notifications while in performance.

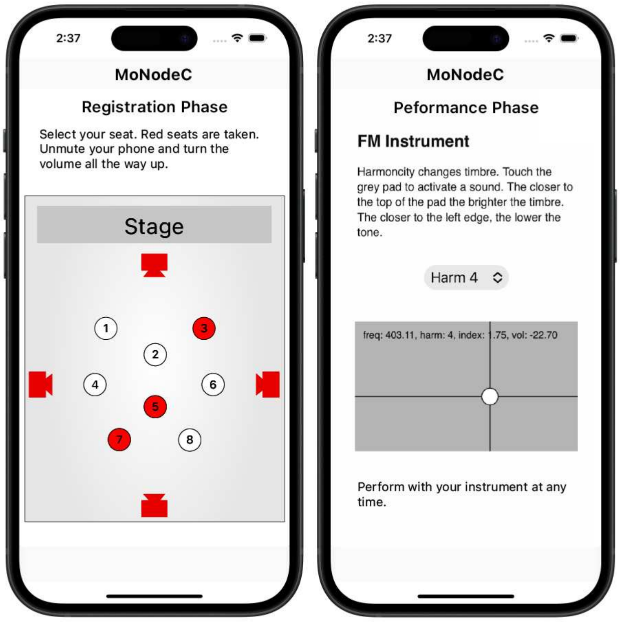

Figure 2. Audience Mobile Interfaces. Left: Registration. Right: A Performance Interface

.

3.3 Sound Control During Performance

The controller is able to select and change control parameters during performance, including main audio material, nearest neighbor distance threshold, number of nearest neighbors, signal strength, devices to include.

Main Audio Material - the controller can choose one sound file for diffusion throughout the performance space. The controller can change the sound file anytime.

Number of Nearest Neighbors - A nearest neighbor algorithm is employed to identify the closest node(s) (Mobile phone or autonomous hubs). The number of nearest neighbors is the maximum number of the possible neighbor nodes. This parameter is used in conjunction with the Nearest Neighbor Max Distance.

Nearest Neighbor Max Distance - This distance is the maximum distance to calculate nearest neighbors from the current touch location in the layout. The algorithm sorts nodes by shortest distance first then number of neighbors.

Signal Strength - This setting can be either 'Distance' or 'Constant' and refers to the amplitude value sent to nodes based on their Nearest Neighbor evaluation. If the parameter is 'Distance', a node within the Nearest Neighbor grouping will receive a ratio of their distance from the controller's touch divided by the Nearest Neighbor Max Distance. Any excluded node would receive a signal strength value of zero. Signal Strength value will not exceed 1. Signal Strength value is the Node Distance divided by the Nearest Neighbor Max Distance. If the parameter is set to 'Constant', all included Nearest Neighbor Nodes will receive a value of 1.

MoNoDeC's approach to diffusion is different than techniques practiced with conventional loudspeaker orchestras [11]: Rather than using multiple channels of audio routed to specific speakers, MoNoDeC uses two audio files routed to all phones with amplitude touch gestures creating the sense that the sound is following that touch path.

4. USE OF DIGITAL INSTRUMENTS

MoNoDeC makes use of digital instruments as one part of its audio synthesis with mobile phones and autonomous hubs. Prior versions used the WebAudio API for audio synthesis on mobile phones with Tone.JS.[4] The current versions on the mobile and autonomous hubs use different exports of RNBO patches. RNBO is an extension of the Max patching environment designed to export software creations to other frameworks.[5] RNBO patch export destinations include Web, C++, and others. These digital instruments exported from RNBO have differing designs because of their expected level of interaction: the hardware instruments will be discussed in the next section. The mobile phone instruments which are controlled by the audience members have simplified controls such as the frequency and modulation index of an Frequency Modulation instrument are mapped to an XY-pad touch area. A single finger's touch and gesture would initiate the onset of sound and continuous control while that finger is down. Part of the digital instrument that the controller commands is obscured from the audience member. This part of the instrument has sound file playback controls and simple oscillators. The controller can select a sound file that will play on the mobile phones. The amplitude of the sound file is controlled by touch gesture through a mouse or finger touch mentioned in *Sound Control During Performance*. These instruments do not stream nor receive streaming audio data. Pre-selected audio files are uploaded on the server prior to the performance, and the digital instruments load those files when performance begins.

During a performance, the controller commands the sound on mobile phones including sound file playback, oscillators and their envelopment, and the amplitude of all sound from MoNoDeC.

The controller changes the interface on the participating mobile phones including instrument controllers or choice-based options.

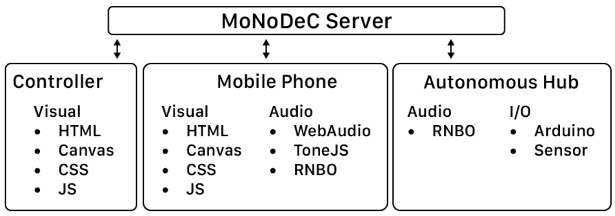

Figure 3. Server-Client Model for MoNoDec

5. HARDWARE

Extending the distributed/networked musical performance device set beyond mobile devices and laptops, MoNoDeC users can incorporate our autonomous hub devices into their network topology. Autonomous hubs are capable of playing simultaneous roles in a MoNoDeC project: each device can serve as a point source in the wireless speaker array, a self-performing and remote-controllable computer music

[4] https://tonejs.github.io/
[5] https://cycling74.com/products/rnbo

instrument, and a sensor-laden interface for communicating with other MoNoDeC clients. The current version of our autonomous hubs are built using a Raspberry Pi Zero W2, a stereo DAC shield with onboard amplifiers, and two 3-watt speakers. This setup can be expanded to incorporate additional shields that add connections for analog and digital sensors or add multichannel audio IO. A custom server runs in tandem with the RNBO runner or web browser on each hub and handles OSC message transmission between the audio engine, sensors, and the main MoNoDeC server.

With the ability to host either a webpage or an instance of a RNBO virtual instrument, autonomous hubs provide MoNoDeC users with a wide variety of options for featuring IoT hardware alongside mobile devices in their creative works. The same audio engine running on participants' phones can be run on each hub, allowing a one-to-one comparison between the sonic contributions of both devices. With the addition of proximity or motion sensors, physical interaction with each hub can allow audience participants to generate control data along the network. The physical computing Hub hardware is compact and lightweight, making it easy to hide or enclose inside fabricated housings, podiums, sculptures, and more. The flexible form factor of the autonomous hubs coupled with their connectivity and audio processing ability provide an alternative to the ubiquitous shape and user experience of consumer mobile electronics.

6. FUTURE WORK

Areas for future development include enhancements to MoNoDeC's software and rendering, digital instrument, and hardware capabilities. The authors plan to develop interfaces into native applications for tablets and mobile devices. MoNoDeC currently does not stream audio, and the authors intend to add that capability. To reduce possible network latency, the authors intend to release a version of the server software capable of running on a single-board computer. A variety of compositional examples, digital instruments, and more audience choices will expand the performance capabilities available to users.

7. CONCLUSIONS

MoNoDeC melds mobile music computing, networked music performance, sound spatialization techniques, and the development of software and hardware systems. By integrating audience mobile phones and IoT-driven speakers into immersive audio experiences, MoNoDeC combines notions of electronic music performers and audience, creating a dynamic and interactive soundscape that extends beyond traditional performance paradigms.

8. REFERENCES

[1] J. Harrison, "Sound, space, sculpture: some thoughts on the 'what', 'how' and 'why' of sound diffusion," *Organised Sound*, vol. 3, no. 2, pp. 117–127, Aug. 1998.

[2] B. Taylor, "A History of the Audience as a Speaker Array," in *Proceedings of the International Conference on New Interfaces for Musical Expression*. Copenhagen, Denmark: Aalborg University Copenhagen, 2017, pp. 481–486. [Online]. Available: http://www.nime.org/proceedings/2017/nime2017_paper0091.pdf

[3] A. Tanaka, "Mobile Music Making," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Hamamatsu, Japan, 2004, pp. 154–156. [Online]. Available: http://www.nime.org/proceedings/2004/nime2004_154.pdf

[4] O. Hödl, C. Bartmann, F. Kayali, C. Löw, and P. Purgathofer, "Large-scale audience participation in live music using smartphones," *J. New Music Res.*, vol. 49, no. 2, pp. 192–207, feb 2020.

[5] S. Robaszkiewicz and N. Schnell, "Soundworks - A Playground for Artists and Developers to Create Collaborative Mobile Web Performances," in *Proceedings of the International Web Audio Conference*, ser. WAC '15, S. Goldszmidt, N. Schnell, V. Saiz, and B. Matuszewski, Eds. Paris, France: IRCAM, January 2015.

[6] B. Matuszewski, "Soundworks - A Framework for Networked Music Systems on the Web - State of Affairs and New Developments," in *Proceedings of the International Web Audio Conference*, ser. WAC '19, A. Xambó, S. R. Martín, and G. Roma, Eds. Trondheim, Norway: NTNU, December 2019, pp. 65–70.

[7] A. T. Marasco and J. Allison, "SoundSling: A Framework for Using Creative Motion Data to Pan Audio Across a Mobile Device Speaker Array," in *Proceedings of the International Web Audio Conference*, ser. WAC '18, J. Monschke, C. Guttandin, N. Schnell, T. Jenkinson, and J. Schaedler, Eds. Berlin, Germany: TU Berlin, September 2018.

[8] D. Poirier-Quinot, B. Matuszewski, N. Schnell, and O. Warusfel, "Nü Soundworks: Using spectators smartphones as a distributed network of speakers and sensors during live performances." in *Proceedings of the International Web Audio Conference*, ser. WAC '17, F. Thalmann and S. Ewert, Eds. London, United Kingdom: Queen Mary University of London, August 2017.

[9] J. Allison and C. Dell, "AuRal: A Mobile Interactive System for Geo-Locative Audio Synthesis," in *Proceedings of the International Conference on New Interfaces for Musical Expression*. Ann Arbor, Michigan: University of Michigan, 2012. [Online]. Available: http://www.nime.org/proceedings/2012/nime2012_301.pdf

[10] N. Bryan-Kinns and P. G. Healey, "Daisyphone: Support for Remote Music Collaboration," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Hamamatsu, Japan, 2004, pp. 27–30. [Online]. Available: http://www.nime.org/proceedings/2004/nime2004_027.pdf

[11] L. Austin, "Sound diffusion in composition and performance: An interview with Denis Smalley," *Computer Music Journal*, vol. 24, no. 2, pp. 10–21, Jun. 2000.