January 2015

Birds of a Feather (Les Oiseaux de Même Plumage): Dynamic Soundscapes using Real-time Manipulation of Locally Relevant Birdsongs

William Walker
Mozilla Corporation, bwalker@mozilla.com

Brian Belet
San Jose State University, brian.belet@sjsu.edu

Birds of a Feather (Les Oiseaux de Même Plumage): Dynamic Soundscapes using Real-time Manipulation of Locally Relevant Birdsongs

William Walker, PhD
Senior Manager
Mozilla Corporation
Mountain View, California, USA
(01) 650 903 0800
bwalker@mozilla.com, softwarewalker.com

Brian Belet, DMA
Professor of Music
San Jose State University
San Jose, California, USA
(01) 408 924 4632
brian.belet@sjsu.edu, beletmusic.com

ABSTRACT
This paper and live audio demonstration explores the capabilities of using Web Audio API as a digital audio workstation (DAW) to manipulate sounds from massive server-side databases. Sonic source material comes from a database of birdsongs recorded worldwide by volunteer recordists at xeno-canto.org. Sounds from xeno-canto are chosen to match recent, nearby bird sightings submitted by volunteer birders at eBird. The result is a virtual soundscape derived from the sounds of birds currently present in the user's geographical region.

Our client-server architecture delegates database queries and archival storage to the server, leaving the client to concentrate on the aesthetic context of sound modification and manipulation. Engineering issues include separation of client versus server concerns and mashups of crowdsourced databases. Aesthetic issues include which tasks are automated server-side, which are user-controlled client-side, and why. Social issues include single user versus multiple user paradigms, artistic soundscape composition versus commercial applications (e.g., games with evolving sound tracks) using public domain sound sources, music as foreground art versus background audio content, and the larger role of sound and music in current society. Audio results will be demonstrated as each topic is addressed.

All the source code for this project is free available under the MIT License at [https://github.com/wfwalker/loco-xeno-canto]. A live demo is at [http://birdwalker.com:9090/quartet.html]

Categories and Subject Descriptors
C.3. [**Special-purpose and Application-based Systems**]: Real-time and embedded systems; D.2.6. [**Software Engineering**]: Programming Environments – *interactive environments*; G.3. [**Probability and Statistics**]: Random number generation; H.1.2. [**Models and Principles**]: User/Machine Systems – *human factors*; H.3.3. [**Information Search and Retrieval**]: Information Search and Retrieval – *information filtering, retrieval models, selection process*; H.3.5. [**Information Search and Retrieval**] Online Information Services – *web-based services*; H.5.2. [**Information Interfaces and Presentation**]: User Interfaces – *auditory (non-speech) feedback, user-centered design*; H.5.5. [**Information Interfaces and Presentation**]: Sound and Music Computing – *signal analysis, synthesis, and processing*; J.5. [**Arts and Humanities]:** Music; K.4.3. [**Computers and Society**]: Organizational Impacts – *computer-supported collaborative work*.

General Terms
Performance, Design, Experimentation, Human Factors

Keywords
AJAX, JavaScript, Web Audio API, client-server architectures, location-based services, crowdsourcing and mashups, real-time audio processing and performance.

1. INTRODUCTION
The authors are both outdoor enthusiasts, and as such the natural world of fauna and flora influence various aspects of their 'indoor' professional work. Walker's passion for birding takes him around the world to observe and photograph birds in their habitats. [http://birdwalker.com/] He is a keen participant in citizen science projects for birding, including Cornell's eBird as well as the Audubon Christmas Bird Count. Belet has produced field recordings (and the inherent unpredictability that is part of this audio capture process) as source material for soundscape compositions, most recently for *Sea Lion Mix*, 2009. [http://www.beletmusic.com/Compositions.html]

Both authors are active musicians who have collaborated formally and informally on a variety of projects over the past thirty years. Their collaboration started at the University of Illinois in the 1980's where a unique blend of computer audio art, science, and engineering found a home in the CERL Sound Group. Walker worked with composer Sal Martirano creating a next generation of his seminal SalMar Construction using MIDI devices and a real-time Smalltalk-80 engine. This led to Walker's ImprovisationBuilder framework, which was the basis for a joint composition with Belet for two Disklavier instruments, each

controlled by a version of the software that affected the other performer's piano (*Cross-Town Traffic*, 1996). A paper describing this collaborative project was published and presented at the 1996 International Computer Music Conference. [Walker and Belet, 1996]

Both authors continue to work in computing and composition: Walker leads a team of partner engineers at Mozilla focused on delivering great web experiences on mobile devices [http://marketplace.firefox.com]; Belet utilizes Smalltalk-80 within Symbolic Sound Corporation's Kyma digital audio workstation for his composing and performing platform. [http://www.symbolicsound.com]

2. WEB BROWSER AS DIGITAL AUDIO WORKSTATION (DAW)

2.1. Historical Context and Current Status

The open web is now becoming a compelling platform for audio experiences, thanks to the convergence of several trends: (1) competition among Apple, Google, and Mozilla to achieve the best possible performance for web standards like CSS3, JavaScript, and HTML5 means that developers can rely on unprecedented client-side capabilities that work in all modern browsers; (2) the new Web Audio API is becoming a de facto standard in modern browsers even as it moves toward formal standardization at the W3C; and (3) the rise of service-oriented architectures for the web has led to a proliferation of REST API's that can be easily combined to create novel experiences.

2.2. Specific Browser Aspects Used For This Project

This project relies on the modern web browser's ability to determine the user's current geolocation (more precisely, the location of the physical device running the browser), to access cloud-based API's using JavaScript and AJAX, to retrieve and decode compressed audio files using AJAX and Web Audio API, and to mix and playback sound files in real time using Web Audio API.

3. AUDIO SOURCE FILES

eBird is a citizen science project in which 150,000 participants worldwide have created a database of over 140,000,000 bird sightings. New sightings are uploaded hourly, creating a dynamic, ongoing picture of bird migration and activity. These observations inform research projects and policy decisions at many levels. We use eBird's REST API [https://confluence.cornell.edu/display/CLOISAPI/eBird+API+1.1] to determine which bird species are being seen recently near the user's location.

Xeno-canto.com is a collection of over 190,000 birdsong recordings made by 2,000 different recordists. The recordings document over 9,000 bird species from all over the world. New recordings are added weekly. We use xeno-canto's REST API [http://www.xeno-canto.org/article/153] to retrieve recordings for the bird species included in recent eBird reports for the user's location.

4. SERVER-SIDE PROCESSING

4.1. Field Recordings of Birds

When users launch "Birds of a Feather," four field recordings of birds are retrieved, filtered by the user's geographic location, as described above. Each sound file is loaded into its own audio player, and file playback begins with an auto-loop function. Each sound file is assigned a distinct random location in the stereo field. The output of all four audio players passes through a mixer for a combined stereo audio output (see Figure 1).

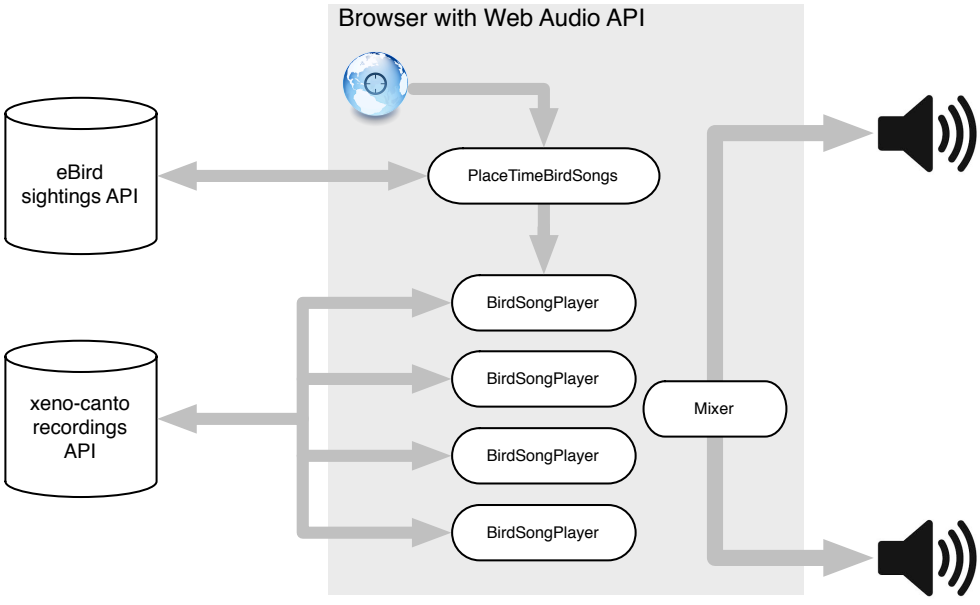

Figure 1. Architecture of "Birds of a Feather" client-side Web Audio API components configured from server-side database queries.

Durations of xeno-canto recordings vary widely, from a few seconds to a few minutes. Complexity quickly arises even before the user interacts with the application. There are four simultaneous audio loops that largely share a similar frequency range and timbral spectrum. Looping multiple recordings with different durations creates ever-changing juxtapositions of the soundscape. The resultant polyphonic texture is reminiscent of *Ars Nova* isorhythmic motets. (This late Medieval period culminated with the music of Guillaume de Machaut [d. 1377], whose art songs used conflicting yet interlinking texts, often in different languages simultaneously, with a high degree of rhythmic complexity and independence. These songs were sung by solo voices of the same type, with similar ranges and timbre.) This invites dual (and duel) aural processing by the listener, as there is an attempt to retain the sonic reality of each unique birdsong while also exploring the artistic abstraction of the new multi-layered complex.

4.2. User Interface

Our goal is to provide all users with a compelling audio experience, even those who don't engage with the user interface. Upon entering the piece, relevant sounds will begin to play as soon as they are chosen, downloaded, and decoded.

We have designed ways for an interested user to produce discernible but unpredictable changes in the soundscape. We have used the shuffle icon, familiar to most smartphone users, to convey that the controls we provide will introduce chance into the soundscape.

The current user interface is intentionally simple, based on Twitter's Bootstrap framework. Bootstrap allows the user interface to expand and collapse in response to laptop, tablet, and smartphone screen sizes and shapes. This 'no frills' clarity is useful for the current development and testing phase. A different interface design may be implemented at a later date, once the initial development issues are settled. User options are clearly enclosed within colored buttons (each with a shuffle icon), while the individual sound files and their relevant data are listed in plain text and graphic data fields, organized within panels (see Figure 2).

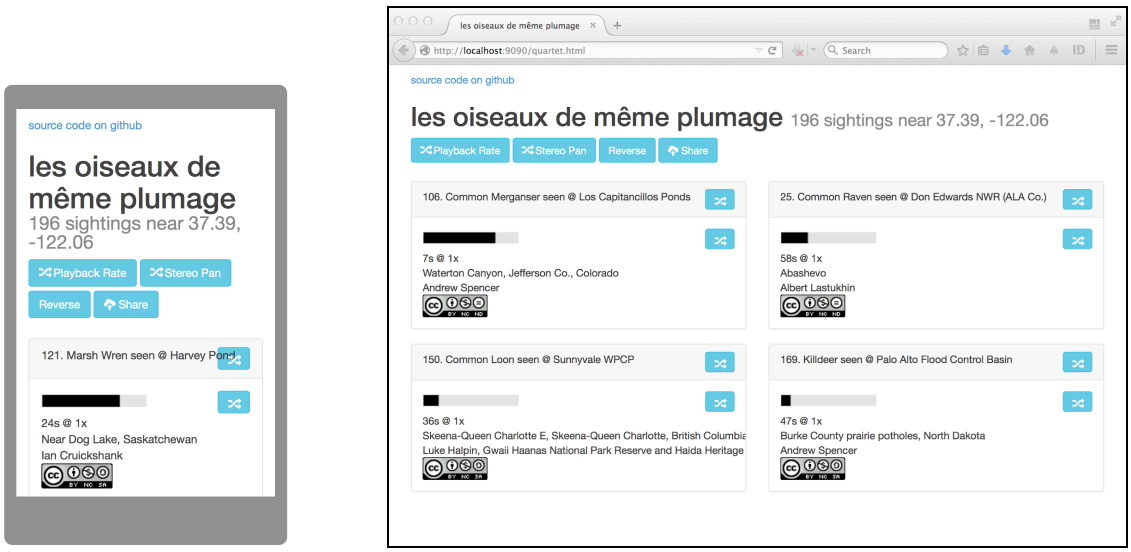

Figure 2. "Birds of a Feather" user interface for smartphones (left) and laptops (right).

User controls include modifying the sound files' playback rate, stereo location, direction of file playback, capturing the state of the system for sharing with others, and selecting new sound files (with two options). Details regarding these options are described in the next section. Current states for the four birdsong files are listed within two interior data fields, each with its own shuffle icon. The upper field lists the bird species and where it was seen in the user's geographic region. The lower field displays an active linear VU meter that displays the signal level of the sound file as playback proceeds, the duration of the sound file and its current playback speed (where 1x equals the original recorded speed), the location where the sound file was recorded, the name of the person who recorded the birdsong, and the license terms.

5. CLIENT-SIDE OPTIONS

The user's geographic location is gathered client-side when the piece loads in the browser, but is used server-side to search millions of bird sightings. Once the four sound files are selected and playback is initiated, the user has options to manipulate the performance soundscape through several processes. Playback Rate generates random playback rates for all four sound files as they are playing in real time. With time and frequency domains linked via the file playback rate, lower frequencies result in time stretched durations. As real birdsongs are mostly high frequency and short duration events, these permutations create rich, unexpected sounds. In addition, a wide range of sonic artifacts, often imperceptible at regular speed, are brought to the aural foreground when the playback rate is quite slow. Combined with

wind and weather, nearby traffic, other bird and animal sounds (including human sounds, often produced by the recordist), these artifacts take on new and artistically significant roles that are of primary interest to the sound designer and composer.

Stereo Pan reassigns each sound file to a new random location in the stereo field, which changes the perceived context of the combined soundscape. The stereo field is derived from Web Audio API's PannerNode, which incorporates a full 3D model of the listener and sound source [https://developer.mozilla.org/en-US/docs/Web/API/PannerNode]

Reverse toggles the playback direction for all four sound files together between forward and reverse. The results of this retrograde permutation are not always obvious at or near the original playback rate, as many landbird call notes or chirps display a near symmetrical amplitude envelope (rapid attack, very little sustain, and rapid decay). However, when the sound file is dramatically slowed down, any differences in the attack and decay envelope slopes become more evident. Recordings of gulls, owls, and waterfowl sound noticeably different in retrograde compared to their original direction.

Share captures the current state of all four sound players, plus the list of eBird sightings from which they were configured. The user is then given a unique URL that can be used to recreate this same soundscape later.

New birdsong files can be selected at two different levels. The shuffle icon at the end of the bird species (first) line selects an entirely new bird from recent eBird sightings, with a corresponding new set of xeno-canto recordings. In contrast, the second shuffle icon, following the audio file data line (to the right of the VU meter), selects a new audio recording of the same (current) bird species. The first option produces a dramatic audio change by default. The second option, while more subtle on first consideration, still produces new sonic results, since a new sound file has its own unique duration (for overall juxtaposition among the four sound files) as well as the specific variant of the birdsong, the recording location and setting, including any artifacts that are present in the recording. In both cases, selecting a new birdsong retains the current playback rate.

6. MUSIC AESTHETICS AND COMPOSITION

Musique concréte is one of the two primary historical branches of electro-acoustic music, the French approach that we trace to Pierre Schaeffer in the 1940s. Aesthetically for Schaeffer "the question was to collect concrete sounds, wherever they came from, and to abstract the musical values they were potentially containing." [Reydellet, 1996] These concrete, or real, sounds originated from both the traditional musical and environmental realms. All sounds and all manipulations of these sounds were fair game for compositional exploration. Modern *musique concréte* includes all aspects of sampling (from whatever original sound source, including birdsongs) and subsequent digital processing of these samples, whether studio rendering or real-time performance, and is therefore the historical precedent and practical basis for this project.

The artistic (and social) interest of *musique concréte*, then and now, is to start with a sound source from our real world, something that is recognizable as tangible and concrete. Birdsongs are recognizable to most as originating from birds, even when the listener is not an expert in the field. Our library of field recordings include unexpected (and uncontrollable) levels of audio data on several levels. The sound files are recorded by different people in different locations, using different levels of recording equipment and displaying different levels of recording skill. Sonic artifacts can and do appear, including wind, traffic, other birds, and human voices. We view these added sounds as serendipitous occurrences, neither positive nor negative. Rather, they are just part of the available sonic landscape the user can explore. Adopting the American experimental aesthetic championed by John Cage, we accept for this project that ''[A]nything … may happen. A 'mistake' is beside the point, for once anything happens it authentically is." [Cage, 1961]

7. SOCIAL IMPLICATIONS AND ISSUES

7.1. Direct Implications of "Birds of a Feather"

"Birds of a Feather" supports hands-on modification of sounds that originate in the real physical world, thereby venturing into the realm of aural fantasy. This encourages serious play as a user motivation and activity. We believe this approach has other applications, including, but not limited to, music composition in the *musique concréte* tradition, sound exploration for its own sake, sound design, real-time performance, evolving soundtracks for games, films, and public installations, and the pedagogical study of sound parameters. The unique strength of this piece is its ease of use (by means of the ubiquitous web browser) and its emphasis on social democracy on many levels.

This experimental approach to music composition (and, by extension, to sound design in general), which we consider central to this project, was articulated by the American composer and theorist James Tenney: "Curiosity … is the strongest motivating force of almost everything I do…. I often think about the pieces that I compose: I write them because I want to know what they are going to sound like." [Belet, 1987]

Our particular collaboration presents a working model that offers benefits to both authors' primary disciplines. A comprehensive musical grounding (historical precedents and contemporary composition practices) provides a practical aesthetic context for the project. Questions such as "Why use these source sounds?" and "Which manipulations of the sounds are musically interesting?" are enlightened from this perspective. In return, the deterministic nature of writing code creates the need for clarity and precision (in both conception and execution) that demands focus of those working in the abstract and flexible realm of music creation.

This soundscape project is only possible thanks to the thousands of field recordings created by hundreds of bird enthusiasts around the world, and then shared under Creative Commons via xeno-canto. We honor their contributions by displaying their names and Creative Commons license icons, highlighting the power of open, collaborative communities.

Social issues for further consideration include single versus multiple users. Would multiple users have the same four birdsongs, or their own unique collection? Would one user's manipulation of the sound files affect only their own collection, or also interfere with another users? These issues lead to larger questions of sound in society (Is it background or foreground?), society's levels of concern for others in a shared (albeit physically displaced) space, and the value of sound as a carrier of real and abstracted information.

7.2. Larger Issues of Acoustic Ecology

There can be no doubt that ambient noise levels in our society have increased since the Industrial Revolution, particularly in urban and industrial settings. The rather ubiquitous distribution and use of personal (with public leakage) audio devices, whether iPhones with screaming ear buds or booming automobile sound systems in our current time period have added a high degree of ambient noise pollution to our daily lives. The field of acoustic ecology has emerged to study this issue from scientific (e.g., The Acoustic Ecology Institute [www.acousticecology.org]) and artistic (e.g., Electronic Music Foundation's 'Ear to the Earth' program [www.eartotheearth.org]) perspectives.

Urban structural surfaces (which reflect and also distort high frequencies) and increased overall noise levels (which mask low frequencies) have also affected birds. One recent study found that bird species that sing in urban areas are shifting their songs upward in pitch [Dowling, 2012]. The rise in the minimum frequency is clearly an adaptation to rise above the low frequency masking caused by urban noise.

8. SUMMARY

8.1. What "Birds of a Feather" Currently Does

As this project is in its first-phase development stage, we are focused on making an initial restricted set of audio processes that are technically stable and aurally interesting. This permits easier error checking on both the coding and user's perception levels. Current server-side processing includes retrieving the four initial birdsong sound files, assigning stereo field locations for each file, and starting their playback at 1x speed with an auto loop function. Current client-side processing options include shuffling the playback rate, stereo location, and playback direction for all four sound files as a group action (with individual parameter values for each sound file). The user can shuffle individual birdsongs on two levels: selecting a new bird species or selecting a new audio recording for the current species. In addition, the current state of the four sound files, as well as the list of eBird sightings from which they were configured, can be captured for later recall and sharing.

8.2. 'Wish List' for Future Work

Additional client-side processing algorithms that we plan to explore include filtering, granulation, delay (reverb and echo), and nonlinear modifications (e.g., phase vocoding, ring modulation, et al). Each of these processes can be explored on a continuum from gentle (sound modification) to extreme (resynthesis). Adding a shuffle option for the start and end points within a sound file is another option to explore. This process can obscure the context of the original when these points generate very small file segments, and these segments can occur in reverse (when the start value is higher than the end value) as well as forward direction.

It is interesting to note that once implemented, this app can work with any collection of easily available sound files. In this regard there are limitless sonic possibilities for the application.

9. ACKNOWLEDGMENTS

The authors would like to thank Soledad Penadés and Samuel Goldszmidt for information and encouragement; with additional thanks to all the citizen scientists contributing to eBird and xeno-canto.

10. REFERENCES

[1] Belet, B. 1987. An Interview with James Tenney. *Perspect. New Music*. Vol. 25, Nos. 1 & 2, Winter – Summer 1987, 459-66.

[2] Cage, J. 1961. Composition: To Describe the Process of Composition Used in *Music of Changes* and *Imaginary Landscape No. 4*. *Silence* (Middletown, CT: Wesleyan University Press) 57-59. https://archive.org/details/silencelecturesw1961cage.

[3] Dowling, J.L., Luther, D.A., and Marra, P.P. 2012. Comparative effects of urban development and anthropogenic noise on bird songs. *Behav. Ecol.,* Vol. 23, No. 1, 201-09. http://beheco.oxfordjournals.org/content/23/1/201.

[4] Reydellet, Jean de. 1996. Pierre Schaeffer, 1910-1995: The Founder of 'Musique Concrete.' *Comput. Music J.*. Vol. 20, No. 2, Summer 1996, 10-11. http://www.jstor.org/stable/3681324.

[5] Walker, W., and Belet, B. 1996. Applying *ImprovisationBuilder* to Interactive Composition with MIDI Piano. In *Proceedings of the 1996 International Computer Music Conference* (Hong Kong, China, August 19-24). International Computer Music Association, San Francisco, CA, 386-89.