CAPTURING THE SOUNDS OF AN URBAN GREENSPACE

Ewan Klein [1, *], Simon Chapple [2], Joachim Fainberg [1], Cat Magill [1], Martin Parker [3], Charles Raab [4,] Jonathan Silvertown [5]

[1] School of Informatics, University of Edinburgh, Edinburgh, UK – {Ewan.Klein, J.Fainberg, C.Magill}@ed.ac.uk
[2] Information Services Group, University of Edinburgh, Edinburgh, UK –Simon.Chapple@ed.ac.uk
[3] Edinburgh College of Art, University of Edinburgh, Edinburgh, UK – Martin.Parker@ed.ac.uk
[4] School of Social and Political Sciences, University of Edinburgh, Edinburgh, UK – Charles.Raab@ed.ac.uk
[5] School of Biological Sciences, University of Edinburgh, Edinburgh, UK – Jonathan.Silvertown@ed.ac.uk

KEY WORDS: urban soundscape, biodiversity, Internet of Things, privacy, citizen engagement, sonic art

ABSTRACT:

Acoustic data can be a source of important information about events and the environment in modern cities. To date, much of the focus has been on monitoring noise pollution, but the urban soundscape contains a rich variety of signals about both human and natural phenomena. We describe the CitySounds project, which has installed enclosed sensor kits at several locations across a heavily used urban greenspace in the city of Edinburgh. The acoustic monitoring components regularly capture short clips in real-time of both ultrasonic and audible noises, for example encompassing bats, birds and other wildlife, traffic, and human. The sounds are complemented by collecting other data from sensors, such as temperature and relative humidity. To ensure privacy and compliance with relevant legislation, robust methods render completely unintelligible any traces of voice or conversation that may incidentally be overheard by the sensors. We have adopted a variety of methods to encourage community engagement with the audio data and to communicate the richness of urban soundscapes to a general audience.

1. INTRODUCTION

The concept of the 'smart city' has been subject to controversy and conflicting visions, especially over the relative primacy of technology-driven versus citizen-centric agendas (Hemment and Townsend, 2013; Hollands, 2015; Kitchin, 2014; McFarlane and Söderström, 2017). Nevertheless, the importance of environmental sensing tends to be a common thread running through these different models and has been boosted further by the spread of Internet of Things (IoT) sensor networks (Alvear et al., 2018; Gabrys, 2014; Jin et al., 2014; Perera et al., 2014). Sensors are used for a wide range of purposes in urban contexts, including air quality, footfall and traffic intensity, weather conditions and energy consumption. Although capture of sounds has figured in a number of city sensing initiatives, the focus has typically been on noise pollution.

Yet sounds can tell us much about what is happening in the world: about birds in the garden, people celebrating a night out on the town, fire engines rushing to an emergency. In the CitySounds project, we have set ourselves the goal of capturing a comprehensive 'picture' of the urban soundscape. Initially, our focus is on the sounds that can be detected in the Meadows, a much-used city-centre greenspace in Edinburgh. Longer term, we intend to expand to a wider variety of urban contexts, and indeed to reach out into the surrounding region, much of which is rural and sparsely populated.

The project, which is still at an early stage, involves a collaboration between university researchers (with expertise in data science, ecology, privacy policy, participatory design and digital sound composition) and members of third sector organisations and community groups; see section 4.1. In large part due to the engagement of these latter groups, the initial phase of work is mainly concerned with audio indicators of biodiversity. However, the Meadows park is bounded by one major road, and is itself an important thoroughfare for pedestrians and cyclists. It hosts annual fairgrounds and marquee-based festival events, as well many sporting facilities. Consequently, as well as capturing biophony — sounds created by birds, mammals, insects and other biological organisms, we are also interested in anthrophony — sounds created by human activities, including music, traffic, sirens, and so on; finally, geophony — sounds from the movement of wind and (to a lesser extent) water —also plays a role (Pijanowski et al., 2011b, 2011a).

Sounds lend themselves to both scientific study and artistic creation, and both of these have a part to play within our project. We have completed a 'sprint' to install a small array of networked sensor kits — Audio Capture Devices (ACDs) — and are in the early stages of a 12-month phase of 24/7 data collection; cf. section 3. In section 5 we highlight the technical and organizational measures taken by the project to support transparency and governance, with the goal of inculcating and sustaining public trust, complementing the support and participation stemming from the engagement activities described in section 4.

The audio samples that we will eventually collect through a year-long natural cycle will constitute a unique and substantial soundscape dataset. With further work, this large body of audio data could provide valuable training input for machine-learning.

Our project includes the following distinctive characteristics:
- Most previous work has focussed either on audible or ultrasonic frequencies; by contrast, CitySounds cover the whole gamut of audio signals from 20Hz up to 96kHz.
- We use a novel design of sensor kit based on low-cost commodity hardware.
- We have developed a highly secure architecture which protects data at source, in transit and in storage, combined with a robust set of measures for ensuring that any published data preserves privacy.

2. RELATED WORK

Research and practice involving acoustic data in urban contexts has many contributing strands. In this section, we survey some of the major components that provide a context for our research.

2.1 Soundscapes

The notion of *soundscape* can be traced back to (Schafer, 1994), where it refers to any acoustic environment that is treated as a field of study. Schafer's main concern was with noise, which "results when man does not listen carefully. Noises are the sounds we have learned to ignore" (Schafer, 1994, p. 4). He argued that noise and noise pollution should be considered a factor in human and planetary well-being, and that the best approach to combating it was a positive one, by identifying sounds that "we want to preserve, encourage, multiply" (Schafer, 1994, p. 4). Ingold (2007) has argued against the use of the term 'soundscape' on the grounds that sound is not something that can be 'scaped', or regarded from a distance. Nevertheless, the term is extremely useful both in connecting up disciplines where landscape is involved and in describing sonic environments.

2.2 Noise, Health and Wellbeing

The negative effects of noise pollution on health have been well documented. For example, chronic exposure to low-level noise triggers an increase in stress hormones which in turn can interfere with sleep, activity and cardiovascular health (Babisch, 2003, 2002); persistent chronic noise exposure increases the risk of cardiometabolic diseases such as hypertension, type 2 diabetes and stroke (Münzel et al., 2018); animal and human studies suggest that exposure to noise is also associated with cognitive impairment and with emotional response in humans, including annoyance and irritation (Münzel et al., 2018); systematic studies indicate that exposure to aircraft and road traffic noise is related to adverse birth outcomes (Nieuwenhuijsen et al., 2017).

The research of the World Soundscape Project[1] led to some of the first governmental noise policies in the world. The EC Environmental Noise Directive 2002/49/EC (END, 2002) deals with noise from road, rail, air traffic and industry, focussing on the impact of such noise on individuals and requires member states to publish noise maps that will inform citizens of acoustic levels, measured in decibels. The maps are created by collecting data from static sensors and computing averages over the course of a year; the noise maps are updated once every five years. Often, the noise levels are represented as gradients on a heat map, as shown for example on the Amsterdam Noise Map 2017.[2] However, with the wider availability of technology for automatic online noise measurement, it has become more feasible to carry out real-time monitoring (Alsina-Pagès et al., 2016), the results of which can be displayed interactively, such as on the Dublin City Environmental Noise website.[3] The Glasgow 3D Sound Map[4] goes a step further in allowing the user to view street panoramas on a web portal while listening to sounds recorded at that location.

Münzel et al. (2018) speculate that noise and air pollution may combine to act synergistically, although there is a lack of detailed evidence to support this. Interestingly, there is research to show that traffic noise correlates with air pollution from black carbon and particulate matter (Dekoninck et al., 2013, 2015), though Apparicio et al. (2016) have found little correlation in cyclists' exposure to air pollution and noise. Given that commercial low-cost air quality sensors are still relatively inaccurate (Castell et al., 2017), at least at the concentration levels found in most European cities, measuring sound as a complementary data source may offer practical advantages.

2.3 Soundscape Ecology and Biodiversity

Acoustic monitoring of the natural environment has a relatively long history. SOSUS (Sound Surveillance System) was a project initiated during the cold war and involved underwater listening systems positioned around the Pacific and Atlantic oceans. SOSUS has been replaced by a number of other surveillance systems but many of the microphone arrays are still active and are used in Marine science research, in particular by the Pacific Marine Environmental Laboratory.[5] This group is gathering long-term data to assist in evaluating human impact on ocean wild life.

Many animals use sound for communication within social groups, during mating, feeding, for navigation and territory holding. By eavesdropping on animal sounds using passive acoustic recording (Browning et al., 2017), ecologists and conservationists can continuously monitor a wide range of ecological and behavioural parameters, including the presence and abundance of different species and activity patterns through the day and around the year. The technology has to date been used to monitor bats, birds, frogs, crickets, marine mammals, elephants and even some fish.

Recording devices are becoming cheap enough (e.g., Hill et al., 2018) to be deployed in large numbers and for use in citizen science (Newson et al., 2017). Since vast quantities of audio data can be generated by such surveys, it is becoming increasingly essential to automate the classification of recorded sounds by the species that made them (Stowell et al., 2016). The emerging field of sound scene analysis is developing computational techniques for analysing audio data captured in various environments, and involves detecting sound events and the context within which those events occur (Benetos et al., 2018). Automated classifiers using deep learning algorithms can now identify bat species in noisy, real world recordings with high precision (Mac Aodha et al., 2018).

Within increasingly congested and polluted cities, greenspaces provide multiple benefits, both in terms of ecosystem services and the mental and physical health of residents (Braubach et al., 2017; Wolch et al., 2014). Moreover, the quality of urban greenspaces and human well-being are closely linked (Niemelä, 2014), hence using acoustic data to monitor biodiversity in urban parks can form an important part of a larger strategy for managing public land.

2.4 Soundscape Composition

'Soundscape composition' is a genre of music production where recordings taken from the world are reformed and restructured into pieces of music. Notable pioneers include Luc Ferrari, Barry Truax and Hildegard Westerkamp. There have been subsequent developments in the worlds of fine art (Janet Cardiff), installation art (Jem Finer), concert music (Francisco Lopez) and from film and television field recording (Bernie Krauss, Chris Watson).

A well-established soundscape project called *Locus Sonus,* led by a dedicated community of sonic artists, maintains a series of live microphones across the world that stream local soundscapes to a

---

[1] https://www.sfu.ca/sonic-studio/WSP
[2] https://maps.amsterdam.nl/geluid/
[3] http://dublincitynoise.sonitussystems.com
[4] http://www.glasgow3dsoundmap.co.uk/soundmap.html
[5] https://www.pmel.noaa.gov/acoustics/

dedicated server.[6] The project leverages computer networks to create sonic connections on a global scale. The purpose of making accessible all of these real-time streams of sound is not to create an archive — the sounds are not recorded at source — but rather to open up new ways of listening to the world. Locus Sonus has become prominent for its role in hosting the annual International Dawn Chorus Day in early May.[7]

An important initiative that connects soundscape, gardens and community engagement is the site-responsive sonic artwork *Sounding out Spaces* (Hayes and Stein, 2018), which ran in the Arizona Desert in 2017. Using embedded computers such as the Bela Platform[8] and Raspberry Pi, the project analyses audio and biotic phenomena and creates a living soundscape for the garden in real-time. Young people from the area around the garden were invited to design sounds and sound-making machines that added to the existing soundscape.

2.5   Privacy, Trust and Data Literacy in the Smart City

The use of ubiquitous sensing and massive data collection to drive technical solutions for smart cities has raised widespread concern about the risks to privacy, and there has been a corresponding effort to develop measures for protecting privacy and engendering trust (Hancke et al., 2012; Martinez-Balleste et al., 2013; van Zoonen, 2016). Trust is a highly complex concept (Gambetta, 1988; Misztal, 1966; Sztompka, 1999) and is multidimensional in IoT applications, especially where people are unfamiliar with complex, networked systems (Harwood and Garry, 2017).

One ingredient for building higher levels of trust is engaging citizens in initiatives that foster greater understanding and agency and facilitate co-production of the smart city (de Waal and Dignum, 2017; Degbelo et al., 2016; Hemment et al., 2016; Twidale et al., 2013; Veeckman et al., 2017). The advent of digital tools based on connected, low-cost sensors is empowering ordinary people to increasingly participate in the collection and interpretation of data from urban environments (Alvear et al., 2018; Balestrini et al., 2017); participatory methods for measuring noise pollution in public spaces are discussed in (D'Hondt et al., 2014; Woods et al., 2018).

3.   DATA COLLECTION AND ANALYSIS

3.1   Technical Framework

We collect data by means of custom-designed Audio Capture Devices (ACDs). Within the ACD box, there are two main components: a Raspberry Pi Zero W running a variant of Linux (Raspbian), and an Ultramic-192 microphone connected to the Raspberry Pi. Since we did not want to be constrained to gain access to mains power, we decided to opt for batteries as a power source, connected to a Raspberry Pi. A 30,000mAH battery currently lasts around a week or longer on a full charge, depending on how cold the weather is — battery performance deteriorates as the temperature drops.

The ACD enclosures have been designed to look like a large wooden weather-proofed "bird box" and to be placed in trees at a height of around 2.5–3m from the ground. For safety and security, they are secured by locked wire cable, and additionally held in place by bungee cord. The fastening system and box are designed specifically not to cause damage to the trees themselves,

as they do not require screws or other fastenings to be embedded in or fixed to the trees. When the devices are *in situ*, we need to visit them periodically to exchange the battery. The box design enables secured easy access for this task.

Figure 1. ACD box installed in tree

The six ACDs operate as a collective continuous sound recording system. Each one captures a 10-second sample of audio per minute (cf. section 5.2), interleaved with one another in sequence so that a full 60 seconds of sound per minute is captured across all the ACDs. Immediately after an ACD completes its 10-second sample recording, it transfers the encrypted sound file over a dedicated, isolated and encrypted WiFi network we have created specifically for this project. The use of a wholly separate WiFi network for this project helps ensure the security and performance of the main institutional networks are not impacted nor at risk.

In addition to the six ACDs, we have installed a Libelium-based sensor kit that measures temperature, atmospheric pressure and relative humidity.   This data is transferred over a separate LoRaWAN infrastructure established prior to CitySounds.[9] Sound frequency propagation in air is fundamentally affected by temperature, relative humidity and atmospheric pressure and so it is important to track these metrics for correct energy calculation of sound sources captured at the detector. The same metrics will enhance interpretation of weather-based sound events. Certain sound profiles, particularly those produced by wildlife, may also be correlated with environmental conditions.
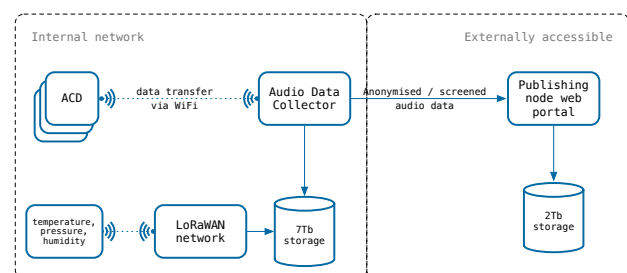

Figure 2. System architecture for collecting audio data

We are using two Virtual Machines, one to act as data collector and the other to host the publishing node. Figure *2* illustrates the system architecture which integrates the sensors with the Data Collector and Publishing node.

3.2   Detecting Sound Events

CitySounds is still a work-in-progress, and so far, we have only captured a few months' worth of data. The audio data files are

---

[6] http://locusonus.org/locustream/
[7] http://soundtent.org/

[8] http://bela.io
[9] http://iot.ed.ac.uk/iot-services/

large in size and long in duration. 10-second samples at 192kHz are 3.5MB in size and each ACD generates 1.26GB every 6 hours. For preliminary exploration, we have developed an application that allows us to extract the spectral energy across each 10-second segment and save these aggregated snapshots of spectra alongside one another. It is possible to 'scrub' through the spectral content quickly but still use our ears to pick out areas of interest very quickly. Using this approach, we have been able to quickly identify the presence of bats, birds, people walking, builder activity, ambulance sirens and buskers.

Another approach we have explored involves concatenating the files into 6-hour chunks and generating long-duration spectrograms (Towsey et al., 2018) that give an overview of the period in question.[10] Whole days and different ACDs can quickly be compared while trends, anomalies and other features (such as sounds in the ultrasonic band) noticed at a glance. Researchers can then home in on specific files and generate further close-up spectrograms of particular areas of interest.

4. ENGAGEMENT & PUBLIC COMMUNICATION

4.1 Partners and Stakeholders

A key partner for this project was Edinburgh Living Landscape, a consortium of organisations focussing on biodiversity in urban environments.[11] Their joint policy objective is the search for more efficient and reliable ways to use new and existing data to inform land management at the ecosystem-scale for the benefit of people, wildlife and the economy. We are also working closely with FOMBL, a citizen-based organisation which coordinates a community garden on the Meadows and a programme of volunteer-led biodiversity projects.[12]

4.2 Engagement Workshops

Overall coordination of community engagement within the project was managed by Edinburgh Living Lab.[13] A key part of our approach to this project was to involve the most relevant stakeholders from the outset. This included soliciting input from core partners to the initial project proposal; inviting a broad spectrum of people onto our management team; and inviting an even wider group to an initial co-design workshop, which focussed on building engagement and impact around the project.

One of the first aims of our engagement workshops was to develop a shared understanding of the technology that was being used, what it could accomplish, how it might be applied to monitor biodiversity in different contexts, and what the avenues for broader engagement were. We also wanted to use the workshops to better understand how CitySounds data could contribute to community interest in the environment, particularly with respect to biodiversity in the city. We set three goals:

1. Develop a grounded understanding of *what* people want to know about their environment (the places and spaces around them) in the city and *how* they want to know about it, e.g., by collecting data, accessing data collected by others, viewing visualisations.
2. Demonstrate how new technologies and new methods of data collection and analysis can support people in

finding out about their environment and exploring what skills can make this possible.
3. Demonstrate the value for the city (and eventually for the region) of supporting people in finding out about their environment through technology and data.

While it is not always easy to get a community gardener, a data scientist, a municipal biodiversity officer and a sound designer talking to each other, it worked amazingly well in this project. We formed new relationships and are continuing to build on them.

The project helped us to reflect on what we mean by 'community' and 'citizen' and on the potential role of new ideas and opportunities at the interface of technology and city issues. We aimed to reach biodiversity enthusiasts with new monitoring technology, data and communication methods; technology enthusiasts with sound recording devices and biodiversity data; sound art enthusiasts with audio and biodiversity data; but also to engage people using or interacting with the Meadows who have little or no experience of biodiversity monitoring, sound recording devices or audio data analysis and presentation. All of these people are in some way community members and citizens of Edinburgh, but the latter group was our ideal target and was, unsurprisingly, the most difficult to identify, reach out to, and draw in.

4.3 Sonic Art Installation

In order to communicate the value of soundscape data to the broader public, over and above the biodiversity community, we created *Sonikebana* (Parker, 2018): a sonic art installation, or more accurately, a long-form soundscape composition that is transformed and shaped by visitor-listeners who are invited to move portable loudspeakers around the space.[14] Each speaker box contains a Raspberry Pi 3B+, an Adafruit LSM303 compass, a 20-watt amplifier, some loudspeaker cones usually found in car sound systems and a 30,000 mAh 5v battery. The speakers were built into standard-sized plywood packing cases. Holes were drilled in the bottom of the boxes where the speaker cones fired downwards. This meant that although sounds felt as though they 'belonged' to a specific box, they spread along the floor, making the speaker behave partially omni-directionally. We commissioned a graphic artist to design and then laser-etch images of abstracted local flora. Placing this on the lids of the boxes meant that there was a visual reward for getting closer to the boxes and pushing them around.

In order to reduce ground loop effects caused by the same battery powering the computer and the amplifier, a ground loop eliminator was placed on the audio cable connecting the amp and computer. The Raspberry Pi computer ran a Python script (Macleod, 2018) that passed compass data across to Pure Data,[15] a computer music programming language, using a protocol called Open Sound Control.[16]

The sound quality of the recordings from the ACDs was noisier than we expected and at the point of the installation opening (April 2018), wind and the dreadful weather of early 2018 hampered many of the sounds we collected. Undesirable noise presents less of a problem for the biodiversity aspects of the project as we are looking for moments of change and difference across a very wide audio spectrum and can factor out many of

---

[10] For an example spectrogram, see
https://citysounds.eu/2018/07/exploring-the-data/
[11] https://edinburghlivinglandscape.org.uk
[12] http://www.fombl.org.uk

[13] http://edinburghlivinglab.org
[14] Cf. http://www.tinpark.com/2018/04/sonikebana-v1/
[15] http://puredata.info
[16] http://opensoundcontrol.org/

these irritants. For the sonic art-piece we also needed distinctive and clear sounds from the site so that audio processes, when used, wouldn't simply contribute more noise. Consequently, the installation used professionally gathered field recordings as well as sounds from the ACDs.

Recordings from various locations and times of day were grouped together and a different group of sounds placed inside each speaker. Data from the compass was mapped to change which recording was being played and to select which parameters would modulate the sounds. This guaranteed that every time a speaker was rolled along the floor, new sounds would emerge.

Much of the sound processing was relatively light in terms of the demands made on the Raspberry Pi's CPU. These broadly involved changing the pitch of playback and how many layers might play at one time via slowly changing delay times and filters. However, real-time analysis of harmonic spectra was used to generate synthesised chords and distinctive melodies that were made dynamically from slowed-down bird song. This lent a particularly musical feel to the sound world.

5. PRIVACY AND SECURITY

5.1 Overview

It is important that the use of microphones is not, and is not perceived as, an audio surveillance system aimed at gathering information about human speech, conduct and interactions with harmful effects on individual privacy, dignity, and freedoms. Providing this assurance is a necessary requirement for a project that relies on public trust and collaboration, and that embraces the spirit of responsible research. From the outset, therefore, the project was designed to include governance and ethics mechanisms as well as oversight procedures that were strongly shaped by persons who are not members of the project team.

We are not able to completely eliminate the risk that some of the audio data we capture would contain traces of voice from passers-by. If the content of such spoken utterances is intelligible, it might contain information that would enable the identification of specific individuals and would therefore constitute personal data. Although we believe there is only a small risk of the ACDs recording intelligible speech, much of the value of the audio data that is captured will come from sharing it for further research and community engagement. Consequently, we have implemented a two-fold approach to ensure that none of the published audio data infringes on individual privacy. We believe that this meets the requirements of 'data protection by design' as set out in Article 25 of the General Data Protection Regulation (GDPR, 2016).

The first strand of our approach allows manual filtering of the audio data to select samples which contain no speech at all and thus require no further processing to ensure privacy. Two members of the project team independently review each such sample by listening to it and viewing the sample's frequency spectrogram. The second strand of our approach is adopted when manual filtering fails to conclusively confirm that voice traces are absent, or it is not practical to use labour-intensive manual filtering at the required scale. In these cases, prior to any publication we will apply a voice-scrambling process to the audio sample that renders speech unintelligible. The voice scrambling algorithm is described in section 5.3.

Once we had established these mechanisms for addressing privacy concerns, we were able to produce a Privacy Impact

Assessment (Data Protection Impact Assessment, in the terminology of Article 35 of the GDPR (GDPR, 2016) that explains the way in which the processing of any personal data captured by the project is minimised and any effects mitigated, so that the project is demonstrably within the scope of legal requirements.[17] In addition, the ethical requirements that underlie the project are aimed at ensuring that possible adverse impacts that might not be sufficiently comprehended by the principles of data protection – for example, discriminatory effects – are anticipated and remedied. Among the procedures are a process for handling enquiries and complaints from members of the public, and an independent ethics and governance component that oversees the project.

5.2 Secure Data Collection

As mentioned earlier, in each specific ACD location only a 10-second sound sample is recorded, then a gap of 50 seconds before the next 10-second sample is recorded. This localised, discontinuous recording is advantageous from the perspective of privacy, meaning that only intermittent snippets of conversation could be captured by the system, even if individuals conversing remain standing next to one of the ACDs for an extended period of time.

The ACDs do not persist or cache any sound recordings on their local filesystem (stored on an SD card); they are held purely in volatile memory until transferred successfully (or transfer fails), at which point they are immediately deleted from the device. Thus, if someone were to tamper with the device, they would not have access to any sound recordings that have been made by that device.

The ACDs are in constant communication with the Data Collector (cf. Figure 2) and if an ACD fails to deliver a new audio file within a 60-second window of time, this automatically generates an alert to take remedial action, which includes immediately disabling the ACD account on the server to which it delivers data while the issue with the ACD in question is investigated further.

5.3 Voice Scrambling

In order to render speech unintelligible, various approaches have been explored that come under the general heading of 'voice scrambling' (Jayant et al., 1981). One possible approach would have been to use a Voice Activity Detector (VAD) (Ramírez et al., 2007); this would allow us to distinguish between speech and non-speech sounds, and to drop the signal whenever a voice is found to be present. However, we decided that it would be too challenging to create a VAD that would perform with high precision and recall on data in which the speech source may be distant. A VAD would probably also struggle to differentiate human voice from biotic sounds (e.g., bird calls) that occur in the same frequency range.

Instead, we developed an algorithm based on "sound shredding" which affects the signal at all times, whether or not voice is present. The basic principle in sound shredding is to first isolate the frequency range in which speech sounds occur; then to chop up the signal into equal-sized chunks called "grains"; and then to randomly shuffle the grains so that contiguous sounds are split up and dispersed. Techniques involving some variant of sound shredding seem to be the dominant choice in the literature and have been adopted by similar urban soundscape projects. Our

---

[17] https://citysounds.eu/privacy/

algorithm closely follows the approach proposed by (Park et al., 2014) which in turn shows a significant overlap with related literature (e.g., Joliat et al., 2013; Kumar et al., 2015; Schmandt and Vallejo, 2003). In more detail, the algorithm proceeds as follows:

1. The speech audio band is extracted with a 3rd-order Butterworth bandpass filter.
2. The bandpassed signal is segmented into individual "grains" (windows) with 50% overlap.
3. Grains in close proximity are randomly shuffled — the shuffling window is shifted forwards at each time step.
4. A random set of grains are individually time-reversed.
5. A smoothed window (e.g., Hann) is applied to each grain.
6. The grains are recombined with overlap-add (OLA).
7. Finally, we reapply the bandpass filter (removing unexpected artefacts) before adding back the other bands.

Critical to the success of this approach are settings for the following parameters (default values in parentheses):

- speech band frequencies (80Hz to 7kHz);
- Butterworth filter order (3);
- overlap during segmentation (50%);
- size of grains (200ms);
- size and hop of shuffling group windows (3 and 1);
- probability of reversing a grain (0.3);
- smoothing window function (Hann).

The frequencies of the bandpass filter are crucial: as the sound shredding algorithm is only applied to this part of the signal, any speech left in the remaining bands would be easily extracted. If the grain size is too large it would make the speech more intelligible and easier to stitch back together in the right order; however, making the grains too small would make it very hard to recognise biotic sounds in the filtered frequencies. Smoothing by overlap and reversal of some grains makes it possible to retain a fairly large grain size without sacrificing effectiveness. Similarly, for the moving window within which grains are shuffled, there is a trade-off between how easy it would be to reverse the algorithm and the extent over which interesting sounds may be temporally distributed. An example of the effect of the algorithm is shown in Figure 3.
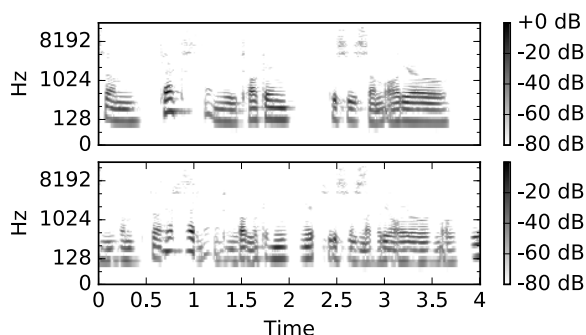

Figure 3. Power spectra for original signal (top), and sound-shredded signal (bottom).

The main drawback of the algorithm is that it treats everything in the speech band as equal. It would be interesting to look at source separation or noise removal techniques which would allow us to keep any background noise constant.

We have carried out an indicative perceptual test with one external expert in order to estimate the residual intelligibility (Jayant et al., 1981) of speech that has been scrambled using our approach, and this confirmed the effectiveness of the sound shredding algorithm. In order to maximally 'stress' the algorithm, the pre-scrambled speech input was recorded in good conditions, without background noise.

6. CONCLUSIONS AND FUTURE WORK

We have described a cyber-physical system that is able to capture wide-spectrum audio data 24/7. The project as described here is still at an early stage, and so far only has only captured approximately three months' worth of data in what is intended to be a full 12 month cycle. We believe that the outcome will constitute a unique and substantial soundscape dataset. This in turn will allow us to develop robust audio classifiers capable of recognising different categories of events within a noisy city soundscape. Such machine learning models would be capable of automatically detecting the presence of birds, bats and other wildlife, and potentially distinguishing individual species on the basis of their sound signatures. Collecting data over this 12-month period will also enable us to demonstrate how biodiversity varies throughout the seasons.

The general public's understanding and acceptance of smart city developments cannot be taken for granted, and projects must demonstrate their trustworthiness in a climate of growing suspicion about data collection and use, especially through systems that appear unusual and opaque. We see citizen engagement and co-production as crucial ingredients in developing trust. As we further develop our city IoT network, we will continue outreach activities in order to build relationships with people and groups interested in biodiversity monitoring. We will explore new methods of using audio data to improve the ways that we interact with and value greenspace and the natural environment in the city. Finally, we will be looking at ways that we can connect with and support community initiatives that are already underway and have strong and active groups around them.

REFERENCES

Alsina-Pagès, R.M., Hernandez-Jayo, U., Alías, F., Angulo, I., 2016. Design of a Mobile Low-Cost Sensor Network Using Urban Buses for Real-Time Ubiquitous Noise Monitoring. Sensors 17, 57. https://doi.org/10.3390/s17010057

Alvear, O., Calafate, C.T., Cano, J.-C., Manzoni, P., 2018. Crowdsensing in Smart Cities: Overview, Platforms, and Environment Sensing Issues. Sensors 18, 460. https://doi.org/10.3390/s18020460

Apparicio, P., Carrier, M., Gelb, J., Séguin, A.-M., Kingham, S., 2016. Cyclists' exposure to air pollution and road traffic noise in central city neighbourhoods of Montreal. J. Transp. Geogr. 57, 63–69. https://doi.org/10.1016/j.jtrangeo.2016.09.014

Babisch, W., 2003. Stress hormones in the research on cardiovascular effects of noise. Noise Health 5, 1–11.

Babisch, W., 2002. The noise/stress concept, risk assessment and research needs. Noise Health 4, 1.

Balestrini, M., Rogers, Y., Hassan, C., Creus, J., King, M., Marshall, P., 2017. A City in Common: A Framework to Orchestrate Large-scale Citizen Engagement Around Urban Issues, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17. ACM, New York, NY, USA, pp. 2282–2294. https://doi.org/10.1145/3025453.3025915

Benetos, E., Stowell, D., Plumbley, M.D., 2018. Approaches to Complex Sound Scene Analysis, in: Computational Analysis of Sound Scenes and Events. Springer, Cham, pp. 215–242. https://doi.org/10.1007/978-3-319-63450-0_8

Braubach, M., Egorov, A., Mudu, P., Wolf, T., Thompson, C.W., Martuzzi, M., 2017. Effects of Urban Green Space on Environmental Health, Equity and Resilience, in: Nature-Based Solutions to Climate Change Adaptation in Urban Areas, Theory and Practice of Urban Sustainability Transitions. Springer, Cham, pp. 187–205. https://doi.org/10.1007/978-3-319-56091-5_11

Browning, E., Gibb, R., Glover-Kapfer, P., Jones, K.E., 2017. Passive acoustic monitoring in ecology and conservation (No. 1 (2)), WWF Conservation Technology Series. WWF, Woking, United Kingdom.

Castell, N., Dauge, F.R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., Broday, D., Bartonova, A., 2017. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? Environ. Int. 99, 293–302. https://doi.org/10.1016/j.envint.2016.12.007

de Waal, M., Dignum, M., 2017. The citizen in the smart city. How the smart city could transform citizenship. It - Inf. Technol. 59. https://doi.org/10.1515/itit-2017-0012

Degbelo, A., Granell, C., Trilles, S., Bhattacharya, D., Casteleyn, S., Kray, C., 2016. Opening up Smart Cities: Citizen-Centric Challenges and Opportunities from GIScience. ISPRS Int. J. Geo-Inf. 5, 16. https://doi.org/10.3390/ijgi5020016

Dekoninck, L., Botteldooren, D., Int Panis, L., 2013. An instantaneous spatiotemporal model to predict a bicyclist's Black Carbon exposure based on mobile noise measurements. Atmos. Environ. 79, 623–631. https://doi.org/10.1016/j.atmosenv.2013.06.054

Dekoninck, L., Botteldooren, D., Panis, L.I., Hankey, S., Jain, G., S, K., Marshall, J., 2015. Applicability of a noise-based model to estimate in-traffic exposure to black carbon and particle number concentrations in different cultures. Environ. Int. 74, 89–98. https://doi.org/10.1016/j.envint.2014.10.002

D'Hondt, E., Zaman, J., Philips, E., Boix, E.G., De Meuter, W., 2014. Orchestration support for participatory sensing campaigns. ACM Press, pp. 727–738. https://doi.org/10.1145/2632048.2632105

END, 2002. Environmental Noise Directive (END). 2002/49/EC of the European parliament and the Council of 25 June 2002 relating to the assessment and management of environmental noise. Off. J. Eur. Communities L 189.

Gabrys, J., 2014. Programming Environments: Environmentality and Citizen Sensing in the Smart City. Environ. Plan. Soc. Space 32, 30–48. https://doi.org/10.1068/d16812

Gambetta, D. (Ed.), 1988. Trust: Making and Breaking Cooperative Relations. Basil Blackwell, Oxford.

GDPR, 2016. {Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)}. Off. J. Eur. Union L119, 1–88.

Hancke, G.P., Silva, B. de C. e, Hancke, J., 2012. The Role of Advanced Sensing in Smart Cities. Sensors 13, 393–425. https://doi.org/10.3390/s130100393

Harwood, T., Garry, T., 2017. Internet of Things: understanding trust in techno-service systems. J. Serv. Manag. 28, 442–475. https://doi.org/10.1108/JOSM-11-2016-0299

Hayes, L., Stein, J., 2018. Desert and sonic ecosystems: Incorporating environmental factors within site-responsive sonic art. Appl. Sci. Switz. 8, 111. https://doi.org/10.3390/app8010111

Hemment, D., Townsend, A. (Eds.), 2013. Smart Citizens. Future Everything, Manchester.

Hemment, D., Woods, M., Appadoo, V., Bui, L., 2016. Community Key Performance Indicators (Community KPIs) for the IoT and Smart Cities: A Collaborative Framework for Project Assessment (FutureEverything Report).

Hill, A.P., Prince, P., Piña Covarrubias, E., Doncaster, C.P., Snaddon, J.L., Rogers, A., 2018. AudioMoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. Methods Ecol. Evol. https://doi.org/10.1111/2041-210X.12955

Hollands, R.G., 2015. Critical interventions into the corporate smart city. Camb. J. Reg. Econ. Soc. 8, 61–77. https://doi.org/10.1093/cjres/rsu011

Ingold, T., 2007. Against soundscape. Autumn Leaves Sound Environ. Artist. Pract. 10–13.

Jayant, N., McDermott, B., Christensen, S., Quinn, A., 1981. A comparison of four methods for analog speech privacy. IEEE Trans. Commun. 29, 18–23.

Jin, J., Gubbi, J., Marusic, S., Palaniswami, M., 2014. An Information Framework for Creating a Smart City Through Internet of Things. IEEE Internet Things J. 1, 112–121. https://doi.org/10.1109/JIOT.2013.2296516

Joliat, N., Mayton, B., Paradiso, J.A., 2013. Spatialized anonymous audio for browsing sensor networks via virtual worlds, in: Proceedings of the 2013 International Conference on Auditory Display (ICAD 2013). Georgia Institute of Technology, pp. 67–75.

Kitchin, R., 2014. The real-time city? Big data and smart urbanism. GeoJournal 79, 1–14. https://doi.org/10.1007/s10708-013-9516-8

Kumar, S., Nguyen, L.T., Zeng, M., Liu, K., Zhang, J., 2015. Sound Shredding: Privacy Preserved Audio Sensing, in: Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications. ACM, pp. 135–140.

Mac Aodha, O., Gibb, R., Barlow, K.E., Browning, E., Firman, M., Freeman, R., Harder, B., Kinsey, L., Mead, G.R., Newson, S.E., Pandourski, I., Parsons, S., Russ, J., Szodoray-Paradi, A., Szodoray-Paradi, F., Tilova, E., Girolami, M., Brostow, G., Jones, K.E., 2018. Bat detective—Deep learning tools for bat acoustic signal detection. PLOS Comput. Biol. 14, e1005995. https://doi.org/10.1371/journal.pcbi.1005995

Macleod, C., 2018. Code to read the Adafruit LSM303 sensor and send the readings to Pure Data [WWW Document]. Gist. URL https://gist.github.com/notexactlyawe/c34e4edbe64c6cca46ec9e8994de994d (accessed 4.26.18).

Martinez-Balleste, A., Perez-martinez, P.A., Solanas, A., 2013. The pursuit of citizens' privacy: a privacy-aware smart city is possible. IEEE Commun. Mag. 51, 136–141. https://doi.org/10.1109/MCOM.2013.6525606

McFarlane, C., Söderström, O., 2017. On alternative smart cities: From a technology-intensive to a knowledge-intensive smart urbanism. City 312–328. https://doi.org/10.1080/13604813.2017.1327166

Misztal, B.A., 1966. Trust in Modern Societies: The Search for the Bases of Social Order. Polity Press, Cambridge, Mass.

Münzel, T., Sørensen, M., Schmidt, F., Schmidt, E., Steven, S., Kröller-Schön, S., Daiber, A., 2018. The Adverse Effects of Environmental Noise Exposure on Oxidative Stress and Cardiovascular Risk. Antioxid. Redox Signal. 28, 873–908. https://doi.org/10.1089/ars.2017.7118

Newson, S.E., Evans, H.E., Gillings, S., Jarrett, D., Raynor, R., Wilson, M.W., 2017. Large-scale citizen science improves assessment of risk posed by wind farms to bats in southern Scotland. Biol. Conserv. 215, 61–71. https://doi.org/10.1016/j.biocon.2017.09.004

Niemelä, J., 2014. Ecology of urban green spaces: The way forward in answering major research questions. Landsc. Urban Plan. 125, 298–303. https://doi.org/10.1016/j.landurbplan.2013.07.014

Nieuwenhuijsen, M.J., Ristovska, G., Dadvand, P., 2017. WHO Environmental Noise Guidelines for the European Region: A Systematic Review on Environmental Noise and Adverse Birth Outcomes. Int. J. Environ. Res. Public. Health 14. https://doi.org/10.3390/ijerph14101252

Park, T.H., Turner, J., Musick, M., Lee, J.H., Jacoby, C., Mydlarz, C., Salamon, J., 2014. Sensing Urban Soundscapes., in: EDBT/ICDT Workshops. Citeseer, pp. 375–382.

Parker, M., 2018. Sonikebana, version 1.0 documentation example. Sonikebana v1.0. http://dx.doi.org/10.7488/ds/2351

Perera, C., Zaslavsky, A., Christen, P., Georgakopoulos, D., 2014. Sensing as a service model for smart cities supported by Internet of Things. Trans. Emerg. Telecommun. Technol. 25, 81–93. https://doi.org/10.1002/ett.2704

Pijanowski, B.C., Farina, A., Gage, S.H., Dumyahn, S.L., Krause, B.L., 2011a. What is soundscape ecology? An

introduction and overview of an emerging new science. Landsc. Ecol. 26, 1213–1232.

Pijanowski, B.C., Villanueva-Rivera, L.J., Dumyahn, S.L., Farina, A., Krause, B.L., Napoletano, B.M., Gage, S.H., Pieretti, N., 2011b. Soundscape ecology: the science of sound in the landscape. BioScience 61, 203–216.

Ramírez, J., Górriz, J.M., Segura, J.C., 2007. Voice Activity Detection. Fundamentals and Speech Recognition System Robustness, in: Robust Speech Recognition and Understanding. Vienna, Austria, pp. 1--22.

Schafer, R.M., 1994. The Soundscape: Our Sonic Environment and the Tuning of the World. Destiny Books, Rochester, Vermont.

Schmandt, C., Vallejo, G., 2003. "ListenIn" to domestic environments from remote locations, in: Proceedings of the 2003 International Conference on Auditory Display (ICAD 2003). Georgia Institute of Technology, Boston, MA.

Stowell, D., Wood, M., Stylianou, Y., Glotin, H., 2016. Bird detection in audio: A survey and a challenge, in: 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP). Presented at the 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP), pp. 1–6. https://doi.org/10.1109/MLSP.2016.7738875

Sztompka, P., 1999. Trust: A Sociological Theory. Cambridge University Press, Cambridge.

Towsey, M., Znidersic, E., Broken-Brow, J., Indraswari, K., Watson, D.M., Phillips, Y., Truskinger, A., Roe, P., 2018. Long-duration, false-colour spectrograms for detecting species in large audio data-sets. J. Ecoacoustics 2, IUSWUI. https://doi.org/10.22261/JEA.IUSWUI

Twidale, M.B., Blake, C., Gant, J., 2013. Towards a Data Literate Citizenry. IConference 2013 Proc. 247–257. https://doi.org/10.9776/13189

van Zoonen, L., 2016. Privacy concerns in smart cities. Gov. Inf. Q., Open and Smart Governments: Strategies, Tools, and Experiences 33, 472–480. https://doi.org/10.1016/j.giq.2016.06.004

Veeckman, C.M., McCrory, G., Walravens, N., 2017. Data Literacy for Greater Civic Participation in Smart Cities - Visualising Open Mobility Data, in: Proceedings of the 21th Conference of the Environmental and Sustainability Management Accounting Network (EMAN). Liège.

Wolch, J.R., Byrne, J., Newell, J.P., 2014. Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough.' Landsc. Urban Plan. 125, 234–244. https://doi.org/10.1016/j.landurbplan.2014.01.017

Woods, M., Balestrini, M., Bejtullahu, S., Bocconi, S., Boerwinkel, G., Boonstra, M., Boschman, D.-S., Camprodon, G., Coulson, S., Diez, T., Fazey, I., Hemment, D., Horn, C. van den, Ilazi, T., Jansen-Dings, I., Kresin, F., McQuillan, D., Nascimento, S., Pareschi, E., Polvora, A., Salaj, R., Scott, M., Seiz, G., 2018. Citizen Sensing: A Toolkit. Making Sense. https://doi.org/10.20933/100001112