A Formalised Approach to Designing Sonification Systems for Network-Security Monitoring

Louise Axon, Jason R. C. Nurse, Michael Goldsmith, Sadie Creese

Department of Computer Science, University of Oxford,
Parks Road, Oxford, UK
Email: {louise.axon, jason.nurse, michael.goldsmith, sadie.creese}@cs.ox.ac.uk

***Abstract*—Sonification systems, in which data are represented through sound, have the potential to be useful in a number of network-security monitoring applications in Security Operations Centres (SOCs). Security analysts working in SOCs generally monitor networks using a combination of anomaly-detection techniques, Intrusion Detection Systems and data presented in visual and text-based forms. In the last two decades significant progress has been made in developing novel sonification systems to further support network-monitoring tasks, but many of these systems have not been sufficiently validated, and there is a lack of uptake in SOCs. Furthermore, little guidance exists on design requirements for the sonification of network data. In this paper, we identify the key role that sonification, if implemented correctly, could play in addressing shortcomings of traditional network-monitoring methods. Based on a review of prior research, we propose an approach to developing sonification systems for network monitoring. This approach involves the formalisation of a model for designing sonifications in this space; identification of sonification design aesthetics suitable for real-time network monitoring; and system refinement and validation through comprehensive user testing. As an initial step in this system development, we present a formalised model for designing sonifications for network-security monitoring. The application of this model is demonstrated through our development of prototype sonification systems for two different use-cases within network-security monitoring.**

Keywords–Sonification; Network Security; Anomaly Detection; Network Monitoring; Formalised Model; Situational Awareness.

I. INTRODUCTION

The cybersecurity of enterprises crucially depends on the monitoring capabilities of the Security Operations Centres (SOCs) operating on their behalf, aiming to maintain network and systems security; in particular, their ability to detect and respond to cyber-attack. Organisations today are frequently the target of cyber-attacks, the nature of which varies widely from ransomware to denial-of-service (DoS) attacks to the exfiltration of sensitive data by insiders, for example. These attacks can be highly damaging both financially, and in terms of the reputation of the organisation. In the face of a constantly evolving set of threats and attack vectors, and changing business operations, there is a constant requirement for effective monitoring tools in SOCs to both automatically and semi-automatically detect attacks.

One of the key challenges that SOCs face in monitoring large networks is the huge volume of data and metadata that can be present on the network. This consists of both the data created by the day-to-day operations of the enterprise, and the data created by security tools. For real-time monitoring, tools that present this data in a form that can be processed in negligible time are essential [1]. Intrusion Detection Systems (IDSs) and visualisations are general examples of classes of tools that are widely used to convey information pertaining to network security in a form that can be easily understood by analysts. The detection algorithms that usually underlie such tools have certain limitations, and can produce false-positive and false-negative results [2,3]. Detecting attacks, and recognising which risks must be prioritised over other attacks and malign activities is difficult, and the degree of inaccuracy in detection systems can make it even more so.

Sonification can provide a potential solution to the challenges of network-security monitoring in SOCs. Sonification is the presentation of data in an audio (generally non-speech) form. Over the last two decades, the incorporation of sonification systems into the monitoring activity of SOCs has been considered [1]. A range of systems has been proposed in which sonified data are presented to support security analysts in their network-monitoring tasks. Some prior work has provided strong evidence of the role sonification could play in improving SOC monitoring capabilities. It has already been shown, for example, that using sonification techniques enables users to detect false-positives from IDSs more quickly [4]. However, the use of sonification systems in this context has not been sufficiently validated, and there is a lack of uptake in SOCs. Sonification has not yet been used operationally in SOCs to our knowledge. Based on the current state of the art, there are clear needs for further research and testing to validate the usefulness of sonification for efficient network monitoring, and to develop appropriate and effective sonifications to enhance network-monitoring capabilities.

This paper is an extension of a survey paper by Axon et al. [1]. In that paper, the major developments over the last two decades in sonification and multimodal systems for network monitoring were reviewed, with particular focus on approaches to design and user testing. That article also contributed a research agenda for advancing the field. This agenda included comprehensive user testing to assess the extent to which, and ways in which, sonification techniques can be useful for network-monitoring tasks in SOCs; the development of aesthetic sonifications appropriate for use in continuous network-monitoring tasks; and the formalisation of an approach to sonifying network-security data. In this paper, we extend that work by proposing an approach to designing sonification systems for network-security monitoring, and presenting a formalised sonification model as part of that approach. We illustrate the application of the model by using it to design two different sonification-system prototypes.

The remainder of this paper is structured in six sections: in Section II, we present traditional approaches to network

monitoring and detail their shortcomings. Section III presents a review of prior work in using sonification for network monitoring, and highlights outstanding challenges in the field. In Section IV, we propose an approach to developing sonification systems for real-time network monitoring. We present our initial work in a part of this approach – the formalisation of a sonification design model – in Section V. In Section VI we apply this model to develop prototype sonification systems for two different use-cases within network-security monitoring. We conclude in Section VII, and indicate directions for future work.

II. Traditional Approaches to Network-Security Monitoring

Network-security monitoring is generally conducted by security analysts, who observe activity on the network – usually using a variety of tools – in order to detect security breaches. According to the UK government's Cyber Security Breaches Survey for companies across the UK, published in May 2016, two-thirds (65%) of large organisations reported that they had detected a security breach in the last twelve months, with the most costly single breach experienced by an organisation during that time purported to have cost £3 million [5]. In the face of such frequent and potentially costly breaches, network-monitoring and attack-detection capabilities are of extremely high importance.

A variety of tools are used in network monitoring: IDSs, Intrusion Prevention Systems (IPSs), visualisations, textual presentations, and firewalls are some of the tools with which analysts conduct their monitoring tasks. The subject of our research is primarily detection, rather than prevention capabilities. We therefore focus on IDSs and anomaly-detection techniques. We also describe the data-presentation methods generally used to convey network-security monitoring information to security analysts – security visualisation tools, and text-based interfaces.

Network monitoring is largely based on alerts given by IDSs. Many IDSs have been based on Denning's model [6]. In general, there are two types of IDS. Anomaly-based IDSs monitor network traffic, and compare it against an established baseline (based on bandwidth, protocols, ports, devices, and connections that are "normal"). Signature-based IDSs, on the other hand, compare packets monitored on the network against a database of signatures or attributes from known malicious threats [2]. Leading SOCs typically craft their own signatures, defined by analysts in the form of rules. Recent advances automate the collection and analysis of data from a range of sources such as logs and IDS alerts using novel Machine Learning and Data Mining approaches.

Anomaly-detection techniques describe methods for the detection of changes in systems that may indicate the presence of threat, and so be of interest from a monitoring perspective. In contrast with signature- or rule-based detection, which relies on comparison with known attack signatures, in anomaly detection, the state of the network is monitored and compared with a "normal" baseline. Anomalous activity is that which exceeds an acceptable threshold difference from this baseline. Anomaly detection often informs the output of IDSs and visualisations. There are several reports reflecting on the state of the art in anomaly-detection techniques [2, 7, 8]. In general, we can divide anomaly-detection methods into three categories [2, 9]: detection methods based on Statistics, in which values are compared against a defined acceptable range for deviation [10, 11]; detection methods based on Knowledge Systems, in which the current activity of the system is compared against a rule-based "normal" activity [12]; and detection methods based on Machine Learning, automated methods in which systems learn about activities and detect whether these are anomalous through supervised or unsupervised learning [7, 13].

Data-presentation techniques convey network-security monitoring information to security analysts. Command-line interfaces are commonly used mediums for presenting the output of network-monitoring appliances such as IDSs and network firewalls. Security visualisations are another widely-used class of tool that convey the output of automated detection tools, and may also present information about the raw network data. While some security-visualisation systems are very basic, there are a number of recent surveys of the state of the art in visualising complex network data. Zhang et al. [14] and Etoty et al. [15] present reviews as of 2012 and 2014 respectively, reporting research into improving graphical-layout and user-interaction techniques [16, 17]. Visualisations generally work by mapping network-data parameters to visual parameters, such that analysts can observe the changes in the visualisation presented and from this deduce changes in, and information about, the network. The design of effective visualisation involves identifying mappings that represent the data in a way that can be understood by security analysts, in SOCs for example, without inducing cognitive overload, and can clearly convey information pertaining to the security of the network.

There are certain drawbacks to current approaches to the monitoring and analysis of security data. Existing automated techniques can be unreliable or inaccurate. Signature-based IDSs may suffer from poorly-defined signatures, and are limited to detecting only those attacks for which signatures are known. The algorithms underlying anomaly-detection techniques using Statistics or Machine Learning also produce false-positives and false-negatives [2, 3]. There is, therefore, a requirement to identify improved anomaly-detection methods. Alongside ongoing research into improving the accuracy of automated detection methods, one avenue that has been researched in security-visualisation work is the detection of anomalies by humans observing aspects of the network data [18].

Given the potential inaccuracy of the alerts produced by the automated detection-system used, it is important that the human analyst has situational awareness and an understanding of the network state, in order that he can interpret alerts and accurately decide their validity; this is one of the key roles of data-presentation techniques. A shortcoming of existing text-based and visualisation-based network-monitoring systems is the requirement that operators dedicate their full attention to the display in order to ensure that no information is missed – for real-time monitoring especially – which can restrict their ability to perform other tasks. Furthermore, the number of visual dimensions and properties onto which data can be mapped is limited [19], and the presentation of large amounts of information visually may put strain on the visual capacity of security analysts.

III. Network Monitoring Using Sonification

Based on the shortcomings we identify in existing monitoring techniques, we believe that sonification may have the

potential to improve monitoring capabilities in SOCs, in a number of ways. While many promising advances have been made recently in novel data-analytics approaches in particular, we highlight that automated network-monitoring systems do not always produce reliable outputs. Presenting network-monitoring information as a continuous sonification could improve analysts' awareness of the network-security state, aiding their interpretation of the alerts given by automated systems. Such awareness could also enable analysts to detect patterns, recognise anomalous activity and prioritise risks differently from the way their systems do, acting as a human anomaly-detector of sorts.

Sonification could also offer a solution to the shortcomings of current data-presentation techniques – in particular, text-based presentation and security visualisations – as an extra interface that requires humans to use their sense of hearing rather than vision. It is important to design representations of large volumes of network data that are as easy as possible for analysts to use, understand and act on. A potential advantage of using sonification in this context is that sound can be presented for peripheral listening. This means that, if designed correctly, sonification could enable analysts to monitor the network-security state as a non-primary task, whilst performing other main tasks. Furthermore, using sound offers another set of dimensions in addition to visual dimensions onto which data can be mapped. The addition of sonification to existing visualisation-based or text-based data presentation approaches could provide a useable method of monitoring highly complex, multivariate network data.

A. Sonification: a Background

Sonification is the presentation of data in an audio (generally non-speech) form. It is used in numerous fields, such as financial markets, medicine (Electroencephalography (EEG) monitoring [20], image analysis [21]) and astronomy. User testing has validated that the presentation of sonified data can improve certain capabilities in a number of applications: improved accuracy in monitoring the movement of volatile market indices by financial traders [22], and improved capabilities for exploratory analysis of EEG data [23], for example.

A variety of techniques and guidelines have been developed for the design and implementation of sonification [24–27]. Throughout sonification literature there are three main approaches recognised: earcons/event-based sonification (discrete sounds representing a defined event), parameter-mapping sonification (PMSon – in which changes in some data dimensions are represented by changes in acoustic dimensions), and model-based sonification (in which the user interacts with a model and receives some acoustic response derived from the data).

The current state of the art in sonification for network and server monitoring is summarised by Rinderle-Ma et al. [19], who identify systems for the sonification of computer-security data, in various stages of maturity. It is concluded that there is a lack of formal user and usability testing, even in those systems that are already fully developed [28–30]. Our survey work differs from that of Rinderle-Ma et al.: while that survey gives an overview of the design approaches taken in some existing sonification systems, our survey provides much greater detail on the sonification design of existing systems in terms of sonification techniques, sound mapping types, the network data and attack types represented and the network-monitoring scope. Furthermore, in this paper we propose an approach to designing and testing the utility of sonification systems for network monitoring, and we go on to actually report on the implementation of that research vision, namely our work on the development of a formalised model for designing sonifications for anomaly-based network monitoring.

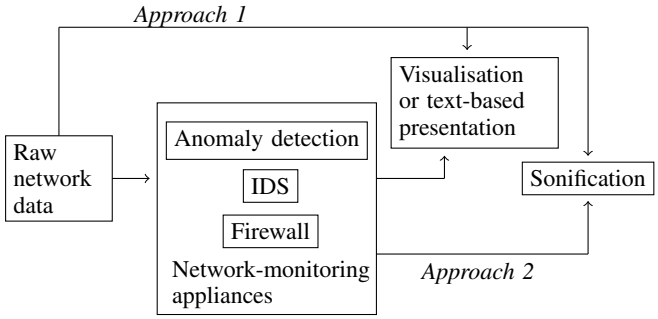

Figure 1. A summary of the existing relationship between traditional monitoring techniques and their potential relationship with sonification systems in SOCs.

Figure 1 shows the existing relationship between raw data, anomaly-detection techniques, network-monitoring appliances such as IDSs, and data-presentation techniques, and the position we envisage sonification might take in this setup. The figure shows two approaches to sonifying network-security monitoring data. In *Approach 1*, the raw network data is represented in the sonification – perhaps with some scaling or sampling methods applied. In *Approach 2*, the network data is not sonified in its raw form but is subject to some automated detection procedure prior to sonification. This either means that the output of some network-monitoring appliance – an IDS, for example – is sonified, or that there is some detection algorithm involved in the sonification method itself prior to the rendering of the data as sound.

TABLE I. EXAMPLES OF TYPES OF RAW NETWORK DATA

Data Type	**Description**
Packet header	The header information for individual packets on the network (including timestamp, source/destination IP/port, packet size, for example) from a network packet capture
Netflow	Data on collected network flows – sequences of packets sent over the same connection (including timestamp, flow duration, source/destination IP/port, for example)
Machine Logs	Data recorded at individual machines on the network. For example, network packets received and sent; processes running; central processing unit (CPU) usage

In Table I, we clarify the meaning of "raw network data" as it is used in Figure 1, by illustrating examples of types of data. The list is not exhaustive, but gives some indication of the network data to which we refer. These data are examples of the raw network data that is sonified directly in *Approach 1* of Figure 1.

B. Applications of Sonification to Network Monitoring

PEEP, a "network auralizer" for monitoring networks with sound, is presented in [28]. PEEP is designed to enable system administrators to detect network anomalies – both in security and general performance – by comparing sounds

with the sound of the "normally functioning" network. The focus of PEEP is on the use of "natural" sounds – birdsong, for example – in sonifying network events. Recordings are mapped to network conditions (excessive traffic and email spam, for instance), and are played back to reflect these conditions. Abnormal events are presented through a change in the "natural" sounds. PEEP represents both network events (when an event occurs it is represented by a single natural sound) and network state (state is represented through sounds played continuously, which change when there is a change in some aspects of the state, such as average network load). There is no experimental validation of the performance of PEEP and its usefulness for monitoring networks, but the authors report the ability to hear common network problems such as excessive traffic using the sonification.

The Stetho network sonification system is given in [31]. Stetho sonifies network events by reading the output of the Linux tcpdump command, checking for matches using regular expressions, and generating corresponding Musical Instrument Digital Interface (MIDI) events, with the aim that the system creates sounds that are "comfortable as music". The aim of Stetho is to convey the status of network traffic, without a specific focus on anomaly detection. The research includes an experimental evaluation of the Stetho system – users' ability to interpret the traffic load from the sounds generated by Stetho is examined. The experiment shows that this monitoring information can be recognised by users from the sounds created by Stetho; however, only four users (subjects familiar with network administration) are involved in the evaluation experiment.

Network Monitoring with Sound (NeMoS) is a network sonification system in which the user defines network events, and the system then associates these events with MIDI tracks [32]. The system is designed to allow monitoring of different parts of a potentially large network system at once, with a single musical flow representing the whole state of the part of the system the system manager is interested in. The focus is not on network security but on monitoring network performance in general; printer status and system load, for example, can be represented through two different sound channels.

More recently, Ballora et al. look to create a soundscape representation of network state which aids anomaly detection by assigning sounds to signal certain types and levels of network activity such as unusual port requests [33] ("soundscape" definition given by Schafer [34]). The concept is a system capable of combining multiple network parameters through data fusion to create this soundscape. The fusion approach is based on the JDL Data Fusion Process Model [35], with characteristics of the data assigned to multiple parameters of the sound. The authors aim, firstly, to map anomalous events to sound and, secondly, to represent the Internet Protocol (IP) space as a soundscape in which patterns can emerge for experienced listeners. No user testing is carried out to establish the usefulness of the system for anomaly-detection tasks. However, the authors report being able to hear patterns associated with distributed denial-of-service (DDoS) and port-scanning attacks (see Table III).

Vickers et al. sonify meta properties of network traffic data [36] as a countryside soundscape. In that system, the log returns of successive values of network traffic properties (number of packets received and sent, number of bytes received and sent) are used to modulate the amplitude, pan, phase or spectral characteristics of four sound channels, including the sound of a running stream and rain. The aim of the system is to alert the system administrator to abnormal network behaviour with regard to both performance and security; it is suggested, for example, that a DDoS attack might be recognisable by the system's representation of an increase in certain types of traffic. There is, however, no evaluation of users' ability to recognise such information using the system. Vickers et al. then extend that work to further explore the potential for using sonification for network situational awareness [37]. For this context, i.e., continuous monitoring for network situational awareness – it is argued that solutions based on soundscape have an advantage over other sonification designs, and that there is a need for sonifications that are not annoying or fatiguing and that complement the user's existing sonic environment.

A soundscape approach is also adopted in the InteNtion system [29] for network sonification. Here, network traffic analysis output is converted to MIDI and sent to synthesisers for dynamic mixing; the output is a soundscape composed by the network activity generally rather than the detection of suspicious activity specifically. It is argued that the system could be used to help administrators detect attacks; however this is not validated through user testing. DeButts is a student project available online in which network data is sonified with the aim of aiding security analysts to detect anomalous incidents in network access logs [38].

García-Ruiz et al. investigate the application of sonification as a teaching and learning tool for network intrusion detection [39, 40]. This work includes an exploratory piece in which information is gathered regarding the subjects' preferred auditory representations of attacks. Sonification prototypes are given for the mapping of log-registered attacks into sound. The first uses animal sounds – auditory icons – for five different types of attack ("guess", "rcp", "rsh", "rlogin", "port-scan"); the second uses piano notes at five different frequencies as earcons to represent the five types of attack. Informal testing was carried out for these two prototypes, and suggested that the earcons were more easily identifiable, while the subjects could recall the attack types more easily using the auditory icons. While this is a useful start to comparing approaches to sonification design for network data, the mappings tested are limited, and further research is required into mappings involving other sound and data types.

Systems have been proposed to sonify the output of existing IDSs, and to act as additions to the function of these systems. Gopinath's thesis uses JListen to sonify a range of events in Snort Network Intrusion Detection System (a widely used open-source network IDS for UNIX derivatives and Windows) to signal malicious attacks [4]. The aim is to explore the usefulness of sonification in improving the *accuracy* of IDS alert interpretation by users; usability studies indicate that sonification may increase user awareness in intrusion detection. Experiments are carried out to test three hypotheses on the usability and efficacy of sonifying Snort. The findings are: musical knowledge has no significant effect on the ability of subjects to use the system to find intrusions; sonification decreases the time taken to detect false positives; immediate monitoring of hosts is possible with a sonified system. As noted by Rinderle-Ma et al. [19], however, the comparison is somewhat biased since the control group without auditory

TABLE II. REVIEW OF APPROACHES TO AND USER TESTING IN EXISTING SONIFICATION SYSTEMS FOR NETWORK MONITORING, ORDERED BY YEAR.

Author	Year	Sonification approach description	User testing	Number of participants	Nature of participants	Network data type mapped	Sound type	Sonification technique	Monitoring scope	Evaluates for utility security monitoring?	Multimodal
Gilfix [28]	2000	"Natural" sounds mapped to network conditions	✗			Raw data (network packet logs)	Natural (wildlife and nature) sounds	PMSon	Anomaly detection: conditions such as high traffic load and email spam are mapped to sound	✗	✗
Varner [41]	2002	Multimodal system: visualisation conveys status of network nodes; sonification conveys additional details on network nodes selected by the user	✗			Not specified	Not specified	Not specified	Network attack detection	✗	✓
Kimoto [31]	2002	Maps parameters of sound to raw network data	✓	4	Subjects familiar with network administration	Raw data (Linux tcpdump output)	Musical	PMSon	General network activity and network anomaly detection	✓	✓
Malandrino [32]	2003	Associates MIDI tracks to user-defined network events	✗			Raw data (printer status, server CPU, file server logs, network packet logs)	Musical	Event-based	Network performance	✗	✗
Gopinath [4]	2004	Instrument and pitch mapped to IDS alert type	✓	20	Computer Science students and staff	IDS alerts (Snort)	Real-world and musical	PMSon	Intrusion detection: IDS logs sonified to aid users monitoring intruders and vulnerable hosts	✓	✗
Papadopoulos [42]	2004	Combines network events rendered as spatial audio with 3D stereoscopic visuals to form a multimodal representation of network information. Sounds are created in response to changes in data patterns using Gaussian Mixture Modelling	✗			Raw data (incoming network flows)	Real-world and musical	PMSon	Anomaly detection: network data presented for pattern recognition	✗	✓
Qi [43]	2007	Maps traffic pattern (classified, queued and scheduled) to audio; bytes and packet rate are mapped to frequency and intensity of audio respectively	✗			Raw data (network packet logs)	Musical	PMSon	Network attack detection (DoS, port scanning)	✗	✓
El Seoud [40]	2008	Auditory icons (non-instrumental) and earcons (instrumental) mapped to attack type	✓	29	Telematics engineering students	Marked attacks from network log	Real-world and musical	Event-based	Network attack detection	✗	✗
Brown [44]	2009	Proposed system maps raw network traffic to sound to convey information on network status; current system maps properties of traffic classified as disruptive by an IDS to properties of piano notes	✗			Raw data (network packet logs) and IDS output	Musical	PMSon	Network anomaly detection (increase in traffic; HTTP error messages; number of TCP handshakes)	✗	✓
Ballora [33]	2011	Parameter-mapping soundscape for overall IP space; obvious sound signals for certain levels of activity	✗			Raw data (network packet logs)	Musical	PMSon	Anomaly detection: anomalous incidents sonified, and network state presented to human to enable pattern recognition	✗	✗
Giot [29]	2012	MIDI messages mapped to data output by SharpPCap library network traffic analysis; MIDI messages mixed to produce a soundscape	✗			Raw data (network packet logs)	Musical	PMSon	General network activity and attack detection	✗	✗
deButts [38]	2014	Maps distinct notification tones to anomalous network events; visualises network traffic activity (multimodal)	✗			Raw data (access logs)	Musical (single tones)	Event-based	Anomaly detection: defined anomalous incidents mapped to sounds	✗	✓
Vickers [36]	2014	Parameters of each sound generator (voice) mapped to the log return values for the network's self-organised criticality	✗			Raw data (network packet logs)	Natural	PMSon	Network performance and attack detection	✗	✗
Worrall [30]	2015	Multimodal system for real-time sonification of large-scale network data. Maps data parameters and events to sound; parameter-mapping sonification approach using melodic pitch structures to reduce fatigue	✗			Raw data (sampled network packet traffic)	Musical	PMSon	General network activity	✗	✓
Mancuso [45]	2015	Multimodal system for representing data on military networks, in which each source and destination IP is mapped to an instrument and pitch, and the loudness is increased when a packet size threshold is exceeded	✓	30	Local population and air force base personnel	Raw data (network packet logs)	Musical	PMSon	Network anomaly detection (packet size threshold, source and destination IPs sonified)	✓	✓

support had to conduct the tasks by reading log files, without access to the visualisation-based tools to which the group tested with auditory support had access.

Multimodal systems, that combine visualisation and sonification for network monitoring, have also been explored. Varner and Knight present such a system in [41]. Visualisation is used to convey the status of network nodes; sonification then conveys additional details on network nodes selected by the user. This multimodal approach is useful because it combines advantages of the two modalities – the spatial nature of visualisation, and the temporal nature of sonification – to produce an effective and usable system. García et al. describe the benefits and pitfalls of using multimodal human-computer interfaces for the forensic analysis of network logs for attacks. A sonification method is proposed for IDSs as part of a multimodal interface, to enable analysts to cope with the large amounts of information contained in network logs. The sonification design approach is not detailed, and the system is not tested with users.

The CyberSeer [42] system uses sound to aid the presentation of network-security information with the aim of improving network-monitoring capability. Sound is used as an additional variable to data-visualisation techniques to produce an audio-visual display that conveys information about network traffic log data and IDS events. The requirement for user testing to establish the most effective audio mappings is recognised, but no testing is carried out. García-Ruiz et al. describe the benefits and pitfalls of using multimodal human-computer interfaces for analysing intrusion detection [46]. A sonification method is proposed for IDSs as part of a multimodal interface, to enable analysts to cope with the large amounts of information contained in network logs.

Qi et al. present another multimodal system for detecting intrusions and attacks on networks in [43]; distinctive sounds are generated for a set of attack scenarios consisting of DoS and port scanning. The authors stipulate that the sounds generated could enable humans to recognise and distinguish between the two types of attack; however, user testing is needed to validate this conclusion and investigate the extent to which this approach is effective. A similar approach is adopted by Brown et al. [44]: the bit-rates and packet-rates of a delay queue are sonified in a system for intrusion detection.

NetSon [30] is a system for real time sonification and visualisation of network traffic, with a focus on large-scale organisations. In this work, there are no user studies, but the system is being used at Fraunhofer IIS, a research institution, who provide a live web stream of their installation [47]. Microsoft have a multimodal system, *Specimen Box*, for real-time retrospective detection and analysis of botnet activity. It has not yet been presented in a scientific publication, but a description and videos of the functioning system are presented online [48]. The system has not been subject to formal evaluation, but is used in operations at the Microsoft Cybercrime Centre.

Mancuso et al. conducted user testing to assess the usefulness of sonification of network data for military cyber operations [45]. Participants were tasked with detecting target packets matching specific signatures (see Table III), using either a visual display (a visual interface that emulated network packet analysis software such as Wireshark) only, or both visual and sonified displays. The aim of the testing was to assess the extent to which sonification can improve the performance and manage the workload of, and decrease the stress felt by, users conducting cyber-monitoring operations on military networks. The testing results show that the use of sonifications in the task did not improve participants' performance, workload or stress. However, only one method of sonifying the data was tested, in which each possible source and destination IP address was represented by a different instrument and note, and the loudness increased if a threshold packet size was exceeded. The results do not, therefore, show that using sonification does not improve performance, stress and workload in this context, but demonstrate only that this particular method of sonifying the data is ineffective.

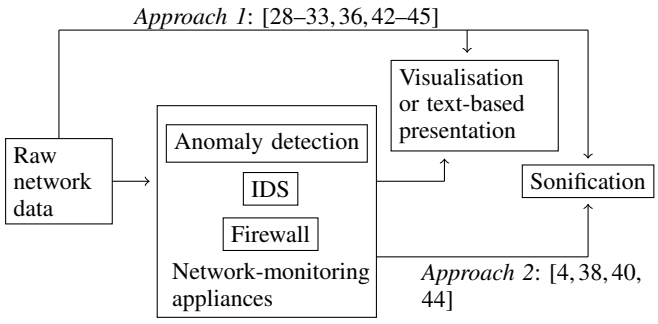

Figure 2. A summary of the data types used in previous network data sonification approaches.

In Figure 2 we show the approaches taken in previous work to designing network-data sonifications, in terms of the type of network data sonified. In this figure, we position the existing sonification systems surveyed onto the monitoring tool relationships diagram presented in Figure 1. Previously-proposed sonifications of network-security data can be divided into two sets: those that take *Approach 1* (in which the sonification system takes as input some raw network data, with scaling functions applied such that the sonification is a representation of the raw network data itself), and those that take *Approach 2* (in which the systems sonifies the output of some network monitoring tool such as an IDS, or sonifies the output of some inbuilt anomaly detection technique).

In Table II, we summarise the sonification design techniques used and user testing carried out in prior work. In Table III we examine in greater detail those existing sonification systems developed for enabling attack detection by sonifying raw network data specifically (*Approach 1* represented in Figure 2). For each system, we present the types of attacks targeted, and the network data features represented in the sonification. We summarise the reported effectiveness of these systems for "hearing" cyber attacks.

In summary, some prior work shows that sonification systems have promising potential to enable network-security monitoring capabilities. Previously-designed sonification systems have been reported to produce sonic patterns from which it is possible to "hear" cyber attacks [28, 33, 36, 43]. In particular, it is reported that DoS attacks and port-scanning attacks can be heard in previous systems sonifying raw network data. User testing has shown that other sonification design attempts were not useful for network-security monitoring tasks [45]; however, the sonification designs and applications tested in this work were limited, and this result is not comprehensive

TABLE III. ATTACK DETECTION AND NETWORK DATA FEATURE REPRESENTATION IN PREVIOUS SONIFICATION SYSTEMS.

Author	Network data features sonified	Can attacks be "heard"?	Attacks targeted
Gilfix [28]	Incoming and outgoing mail; average traffic load; number of concurrent users; bad DNS queries; telnetd traffic; others unspecified	Not assessed, but authors report ability to "easily detect common network problems such as high load, excessive traffic, and email spam"	Not specified
Varner [41]	Not specified	Not assessed	
Papadopoulos [42]	Packet rate; others not specified	Not assessed	
Qi [43]	Packet rate; byte rate	No experimental assessment, but authors report that the system produced sounds "notably" different enough that distinguishing between DoS and port scanning attacks is "relatively easy", while no sounds were produced under "normal" traffic conditions	DoS; port scanning
Brown [44]	Prolonged increase in traffic volume; number of TCP handshakes in progress; number of HTTP error messages	Not assessed	
Ballora [33]	Source IP address; destination IP address; frequency of packets in ongoing socket connections; packet rate; requests to unusual ports; geographic location of sender (suggested but not implemented)	Not assessed, but authors report finding "that patterns associated with intrusion attempts such as port scans and denials of service are readily audible"	Dataset used contains DoS and port-scanning attacks
Giot [29]	Packet size; Time-to-Live (TTL) of packet; bandpass of network; source IP address; destination IP address; protocol (type of service); number of useless packets (e.g. TCP ACK packet)	Not assessed	
Vickers [36]	Data sonified are log returns of successive instances of the following values: number of bytes sent; number of packets sent; number of bytes received; number of packets received	Changes in soundscape not noticeable under "normal" network conditions; noticeable change occurs when log returns large (large log return for number of packets received might indicate DDoS, for example)	Not specified
Mancuso [45]	Source IP address (of packet); destination IP address (of packet); packet size	Use of sonification alongside the visual interface did not improve participants' performance in detecting "target packets" compared with their performance using the visual interface alone	Not specific attacks – target packet characterised by "signatures": network transmissions originating from either of two particular source IP addresses, directed to either of two destination IP addresses, using either of two protocols, with packet size 500 bytes or more

enough to suggest that further research in this area is futile. It is clear that variations in sonification design approach may affect the usefulness of the system for network-security monitoring, and as such further research is required into appropriate sonification designs for the context.

C. Outstanding Challenges

Table II presents a summary of the sonification systems previously developed for network monitoring (solutions for which full systems or prototypes have been developed). From this, we have identified the key areas in which research is lacking: formalisation of a model for designing sonification systems for network monitoring, identification of data requirements, investigation of appropriate sonification aesthetics, and validation of the utility of the approach through user testing.

In general, a weakness in the articles is the amount of user testing carried out with the intended users – security analysts. Table II shows that little user testing has been carried out, and of that which has, little has specifically targeted security analysts – it is possible that some of the Air Force Base personnel who participated in the user testing by Mancuso et al. were security analysts, but this is not made clear in that paper [45]. Table II shows also that there has been little (and no comprehensive) evaluation of the usefulness of existing sonification systems for network anomaly detection. Gopinath evaluates the usefulness of a sonification with a focus on aiding users in monitoring the output of IDSs [4]. Mancuso et al. evaluate the effectiveness of their sonification system in enabling users to detect packets matching specific signatures, but test only one sonification design. There is therefore a clear need to assess and compare the use of a number of sonification designs for network anomaly detection. Extensive user testing is required to validate the usefulness of the approach and of proposed systems, and to refine the sonification design.

The systems listed vary in the data they represent. Some map raw network data to sound, some map the output of IDSs, while some aim to map attacks to sounds. However, there is no comparison of the efficacy of these approaches, or of the usefulness of sonic representations of different attack types. Identification of the network-data sources and features that should be sonified in order to represent network attacks is needed. The sonification design approaches used (event-based, parameter-mapping, and soundscape-based) also vary, as do the sound types (natural sounds, sounds that are musically informed) but there is as yet no comprehensive investigation into, or comparison of, the usefulness of these methods. Based on this, we propose that comparative research into the sonification aesthetics most appropriate for use in network monitoring is crucial, in order to inform sonification design. We further identify a requirement for the development of a formalised approach to designing sonifications in this field, to underpin developments and enable comparison. Next, we outline our proposed approach to sonification development and testing, with which we aim to address these issues.

IV. Proposed Approach

We propose an approach to developing sonification systems in this space. The approach involves formalising a model for designing sonifications for network monitoring, identifying the network data representation requirements, investigating appropriate design aesthetics for the context, and assessing the utility of the developed systems through comprehensive user testing. We believe that these elements combine to form a solution to the problem of designing and testing the utility of sonification systems for network-security monitoring.

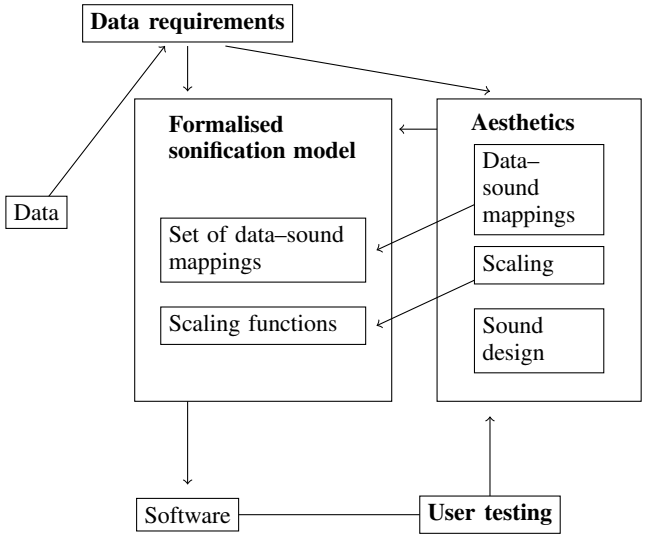

Figure 3. Proposed approach to designing sonification systems for network-security monitoring.

Figure 3 shows the parts of our approach, and their relationship with each other. The formalised network data sonification model takes as input aesthetic requirements and data requirements, and incorporates the results of iterative user testing. We now detail the research questions to be answered for each part of the approach.

A. Requirement for a Formalised Sonification Model

To enable us to architect and experiment with sonifications in a flexible way, we need an underpinning sonification model. This should enable us to utilise heterogeneous sonifications alongside each other in order to compare performance. No such model currently exists, and we therefore propose the development of a formalised model specifically for developing sonification systems for network monitoring. This model should describe a grammar for the representation of network data through parameter-mapping sonification that enables incorporation of and experimentation with appropriate design aesthetics, techniques of musical composition, and the science of auditory perception. It is important that the model encompasses prior art, and enables comparison with previous approaches to designing sonifications in this space.

A model for designing sonifications for use in the network-monitoring context should tailor aspects of sonification design such as cross-field interference to produce sonifications that are appropriate for network-monitoring tasks. A simple example is a simultaneous change in two network parameters: a statistically significant increase in traffic load, and messages received from an IP address that is known to be malicious (these two changes would generally be found by the statistical anomaly-based IDSs and signature-based IDSs, respectively, described in Section II). This could be the result of a DoS attack, and the sonification system should therefore attract the attention of the analyst. Cross-field interference could be leveraged through the choice of data-sound mappings used in this case (with a mapping to higher pitch and increased tempo – two sound parameters which interact such that each appears more increased that it really is – for the two data parameters respectively) to ensure that the attack is highlighted by the sonification.

In order to prevent sonification designs from causing listener fatigue, we propose that a rule-based approach to aesthetic sound generation may be appropriate. In particular, a sonification model should be non-prescriptive in terms of musical genre, and be applied to a variety of genres of music to generate a set of different-sounding sonifications of the same network data. We hypothesise that with this approach, users could be allowed to move between a set of musical genres at choice, each of which would sonify the network data according to the same grammar, and this could reduce the fatigue caused by the sounds. Below, we give the key questions to be addressed in building this model.

- **Which are the requirements specific to the network-security monitoring context for the mapping of data to sound?** In general, huge quantities of multivariate and highly complex data move through organisational computer networks. It is important that the model enables the sonification designer to reason about the parts of the data to be sonified, the key information about these parts that must be conveyed to the sonification user, and the most appropriate method of representing this information through sound. For example, an important task in SOCs is monitoring the security state of sensitive servers on the network – this could be those servers containing databases of customer records. Devising methods of mapping required information about selected aspects of the network to sound will be a key part of the model development process.
- **What are the inputs and outputs of the sonification model?** The sonification model should take as input both the data requirements for the representation, and the aesthetics derived: appropriate data-sound mappings and sound design, and methods of scaling the data to the sound domain. The model should provide a method for mapping the required data to sound, following the aesthetic requirements input. The model should itself then produce the input to some sonification software. Adaptability of the model according to differing aesthetic requirements is important, particularly as we aim to compare multiple aesthetic approaches, and refine the aesthetic requirement specification through iterative user testing.
- **How can we verify that the sonification model is capable of addressing prior art approaches?** In order to enable comparison of new sonification system designs with the approaches taken in prior work, it is important that prior approaches can be replicated through use of the sonification model developed. We

can verify the correctness of the model for this task by verifying that it has some representation of each relevant prior sonification approach.

B. Data Requirements

The data requirements include firstly the data sources used, since these produce different data types. For example packet capture header data might be represented – a different data type to machine log data (including machine CPU, for example) or file access log data. The data requirements must also include the data features addressed. These are the properties of the data that we choose to sonify, and may be low-level properties (such as a representation of source IP address from which each packet is received) or may be attack-detection features (such as packet rate thresholds against which data are compared).

The data requirements depend to a large part on the use-case. In developing sonification systems for anomaly detection by humans, data requirements should be derived from information about all data sources and features that enable network anomaly detection, and through which attacks are conveyed. On the other hand, for use in a multimodal system, which conveys part of the network data sonically while other data is conveyed visually, the sonification data requirements would depend on which data had been selected to convey visually, and which using the sonification. As another example, if the aim of sonification was to enable analysts to monitor network security as a non-primary task, the data requirements should be informed by the data sources that analysts may be frequently required to monitor while simultaneously conducting other tasks – these sources might include IDS alert logs, or the logs of critical servers on the network, for example.

- **Which data sources should be included in developing sonifications for network-security monitoring purposes?** It is important to identify those sources for which a sonified representation might add value in network monitoring; these might be raw network data sources such as packet captures, Netflow or Domain Name System (DNS) logs, or the sources might be monitoring systems such as IDSs or network firewalls. Buchanan et al. categorised the potential data sources used by security analysts in answering a number of different analytical questions (for example, in searching for the activities associated with a particular suspicious IP address) [49]. We hypothesise that raw network packet capture data is most suitable for network attack detection, because this constitutes a full representation of traffic on the network. However, it would be valuable to identify the network data sources security analysts consider most useful for network attack detection, and the methods by which those sources are currently monitored. For example, the information output of multiple such data sources are often integrated in Security Information and Event Management (SIEM) tools for monitoring by analysts.
- **Which data properties or features should be sonified to enable network anomaly detection by analysts?** In order to identify the data properties to be represented, attack characterisation can be used to extract the ways in which classes of network attacks (flooding attacks, for example) manifest in the network data sources selected for representation in the sonification. Some prior work identifies network data features for network anomaly detection, and for the detection of particular classes of threat such as Advanced Persistent Threats (APTs) and Botnets [50–52]. Some of this work involves interviews with security analysts to identify the properties of data analysts search for in network security monitoring to enable attack detection [49, 53]. The findings from attack characterisation and prior work can be bolstered through interviews with security analysts, to gather their views on the importance of particular network data features for network attack detection.

C. Sonification Aesthetics

While there has been some work in aesthetic sonification, as reported in Section III, it has not been heavily applied in the context of network monitoring. Prior work indicates that sonification aesthetic impacts on its effectiveness. In an experiment comparing sonifications of guidance systems, for example, it was shown that sonification strategies based on pitch and tempo enabled higher precision than strategies based on loudness and brightness [54]. It was also shown in [55] that particular sonification designs resulted in better participant performance in identifying features of Surface Electromyography data for a range of different tasks involved.

The aesthetics of the design are an important factor in producing sonifications that are suitable for continuous presentation in this context. In particular, the sounds should be unfatiguing [37, 56] and, if intended for use in non-primary task monitoring, should achieve a balance in which they are unobtrusive to the performance of other tasks while drawing sufficient attention when necessary to be suitable for SOC monitoring. While there are other techniques that may be useful, we propose an approach to this design that draws on techniques and theories of musical composition. We can draw on work in aesthetic sonification by Vickers [56], and on work in musification, i.e., the design of sonifications that are musical. Some key questions to be answered regarding sonification aesthetics for network-security monitoring are described below.

1) **Which are the most appropriate mappings from network-security data to sound?** Prior work has indicated preferred mappings from data to sound in certain contexts; for mapping physical quantities such as speed and size, for example [57]. Useful parallels can be drawn between these previous experiments and the network-monitoring context, and hypotheses can thus be made about appropriate data-sound mappings. However, it is important to perform a context-specific assessment of these mappings, in terms of their ability to convey the required network-monitoring information in a way that users can comprehend. Associations formed through the previous experiences of users may affect the ease with which they can use certain mappings; for example, based on prior work we might expect a mapping from packet rate to tempo of music to be intuitive. We propose that user experiments should be carried out as part of the sonification design process, to establish which mappings from data to sound are most appropriate. The results of these user experiments will form an input to the sets of data-sound mappings used in the sonification model, as shown in Figure 3.
2) **Which sonification aesthetics are suitable for use in**

network-security monitoring in SOCs? Comparison of a range of musical aesthetics (for example, a comparison between Classical Music and Jazz Music), should be carried out to identify those most suitable for the context. In particular, aesthetics that are unfatiguing, unobtrusive to other monitoring work, and able to attract the required level of attention from analysts, should be chosen. It may be that a suitable approach is to enable analysts to choose between a selection of musical aesthetics at will. It is important to assess the extent to which musical experience affects the ability of security analysts to use musically-informed sonification systems in network-monitoring tasks. The effect of users' musical experience on their ability to understand and make use of the sonification systems design will require investigation. Here, musical experience refers to the level of prior theoretical and aural musical training attained by the user. For this SOC monitoring context, analysts' use of the systems should not be impaired by a lack of musical experience.

3) **What granularity of network-security monitoring information can we represent usefully using sonification?** Given the huge volumes of network data observed on organisational networks, and the high speed of packet traffic on these networks, it should be assumed that some scaling or aggregation will be required in the sonic representation of certain data sources. The amount of information that can be conveyed through sound should be identified. This is both in the sense that sonification software is actually capable of rendering the information, and that humans can usefully interpret the information presented and hear the network data required for anomaly detection, i.e. that the sound is not overwhelming. Methods for producing network data inputs that can be usefully rendered as sound, such as aggregating packets over time intervals, or scaling quantities such as packet rate, should then be experimented with. Sampling packets is another possible approach; for example, Worrall uses sampled network packets as the input to his network data sonification using the *Sflow* tool, which takes packets from the traffic stream at a known sampling rate [30]. Comparative testing would be valuable at this stage to assess the levels of granularity of data sonification at which network anomalies can be heard. The results of this assessment of appropriate data granularity will form an input to the scaling functions of the sonification model, as shown in Figure 3.

Besides aesthetics, aspects of human perception must influence the design: the prior associations sounds may hold for users and the way in which this affects interpretation; the effect of musical experience on perception. It is important that the design takes into account factors in perception such as cross-field interference (in which different dimensions of sound – pitch and tempo, for example – interact in a way that affects perception of either) and does not induce cognitive overload for the user.

D. Comprehensive User Studies

As well as addressing sonification-aesthetic requirements through iterative user testing, we need to conduct user experiments to investigate the utility of sonification systems for network-security monitoring tasks. Section III indicates that of the proposals made for sonification systems for network monitoring, very few have conducted any user testing. None have conducted testing to the extent required for an appropriate understanding of the use of these systems and their suitability for actual deployment in security monitoring situations. As such, we identify a requirement for significantly more in-context user testing of sonifications for network-monitoring tasks, carried out with security analysts in SOCs, to inform the design and investigate the advantages and disadvantages of the approach. It is important that sonification systems are tested in the SOC environment, in order to investigate how well they incorporate with the particular characteristics of SOCs – a variety of systems running simultaneously, collaborative working practice, high levels of attention required from workers.

We will conduct user testing to investigate the hypothesis that sonification can improve the network-monitoring capabilities of security analysts. This hypothesis is proposed in light of prior work in other fields in which it is proven that certain capabilities can be improved by the presentation of sonified data, as outlined in Section III, and of the limited experimental evidence that shows that sonification can be useful for tasks involving network data specifically [4, 31].

For the validation of sonification as a solution to improving network-monitoring capabilities, there are certain key research questions that need to be answered through user testing.

1) **To what extent, and in what ways, can the use of sonification improve the monitoring capabilities of security analysts in a SOC environment?** User testing is required to establish the extent to which sonification can aid security analysts in their network-monitoring tasks. We theorise that there may be a number of use-cases for sonification of network data in SOCs. For example, investigation is needed to establish whether the presentation of network data through sonification can enable analysts to "hear" patterns and anomalies in the data, and in this way detect anomalies more accurately than systems in any cases. Given the strong human capability for pattern recognition in audio representations [56, 58, 59], and for contextualising information, it is plausible that a system that presents patterns in network data may enable the analyst to detect anomalies with greater accuracy than traditional rule-based systems. User testing should also establish whether presenting sonified network data can enable analysts to monitor the network as a non-primary task, maintaining awareness of the network-security state while carrying out other exploratory or incident-handling tasks. Finally, we propose user testing to assess whether multimodal systems, which fuse visualisations and sonifications of complex data – which might usually be presented visually across multiple monitors, for example – can aid analysts in their network-monitoring tasks.
2) **Are there certain types of attack and threat that sonify more effectively than others, and what implications does this have for the design of sonification systems for network monitoring?** It may be the case that certain types of attacks are better-represented through sonification than others, and that some attacks sound anomalous in a way that is particularly easy for analysts to use while others do not sonify well. Findings on this subject should inform sonification system design by distinguishing the attacks and threats in relation to which sonification performs best, and the areas in which the technique therefore

has the potential to be most effective.

3) **How does the performance of the developed sonification tools in enabling network attack detection compare with the performance of other network attack detection tools?** The performance of users in network attack detection using sonification alone, and using network monitoring setups incorporating sonification, should be compared to their performance using visualisation and text-based interfaces. Users' performances in detecting network attacks using the sonification should also be compared with the performance of automated systems such as IDSs. It is important to compare the attack detection performance (in terms of accuracy and efficiency) of humans using the sonification to that of automated systems, for particular classes of network attack.

Answers to these questions will provide a greater understanding of the role sonification can play in improving monitoring capabilities in SOCs, the limits of the approach, and the extent to which it can be reliable as a monitoring technique. In conducting this testing, we expect to draw from existing research on conducting user studies in general, and in a security context [60, 61].

V. Formalised Model for the Sonification of Network Security Data

In this section, we expand on our proposal in Section IV by presenting a formalised approach for the musical parameter-mapping sonification of network-security data. In particular, we focus on our formalised sonification model (as introduced in Section IV.A). We first identify requirements for sonifying network-security data, and from these requirements, construct a model for developing sonifications for network-security monitoring uses.

Some work in formalising the sonification of data has been presented previously. For parameter-mapping sonification, a formalised representation of the sonification mapping function is given by Hermann [23]. That representation was the basis of the parameter-mapping sonification model that we developed for network-security monitoring. In Hermann's representation, the parameter-mapping function $\mathbf{g} : \mathbb{R}^d \to \mathbb{R}^m$ describes the mapping from a d-dimensional dataset $\langle x_1, ..., x_d \rangle \in \mathbb{R}^d$ to an m-dimensional vector of acoustic attributes which are parameters of the signal generator. The q-channel sound signal $s(t)$ is computed as a function $\mathbf{f} : \mathbb{R}^{m+1} \to \mathbb{R}^q$ of the parameter-mapping function $\mathbf{g}$ applied to the dataset, and time t:

$$s(t) = \sum_{i=1}^{d} \mathbf{f}(\mathbf{g}(\mathbf{x}_i), t).$$

In developing our model, we draw on de Campo's Sonification Design Space Map (SDSM), which describes the questions to be addressed in any sonification design process [62]. The map presents, as axes, three key questions for reasoning about data aspects in sonification design. We also use the work of Hermann [23]; in particular, we extend Hermann's parameter-mapping sonification formalisation, by addressing the design questions indicated by the SDSM.

A. Requirements of the Model

In what follows, we describe the use of the SDSM design questions to extract requirements for the model. We present each question, then consider context-specific answers. We thus identify requirements particular to sonification for network-security monitoring.

- *Question 1: How many data points are required for patterns to emerge?*
  The presentation of network data at a range of different resolutions may be required for different monitoring applications – see Subsection IV.B:

  Requirement 1: the model should enable any number of data points to be represented.

- *Question 2: What properties of data dimensions should be represented?*
  The properties of data dimensions represented should be those through which indicators of attacks are shown. These may vary based on the network type and the source of the monitoring information:

  Requirement 2: the model should enable the inclusion of appropriate data dimensions for individual designs.

  Furthermore, these dimensions may be continuous (for example, packet rate), or discrete (for example, direction of packet flow – incoming/outgoing). Appropriate mapping of both continuous and discrete data dimensions should be enabled in order to prevent unnecessary loss of resolution in the data representation (for example, there would be a loss of resolution in a representation in which data with continuous values, such as packet rate, was mapped to a sound with a small number of discrete values, such as type of instrument):

  Requirement 3: the model should provide a systematic method of mapping continuous and discrete data dimensions to continuous and discrete sound dimensions.

- *Question 3: How many sound streams should be present in the design?*
  This depends on the network type, use-case and monitoring information source, but in general network data is multivariate, with many network elements, data sources, and packet flows that require monitoring. We require a method of communicating which of these streams is represented by particular sounds: we need to represent information about a number of different channels of the network data. This means, we need to know what is happening, and to which parts of the data:

  Requirement 4: the model should allow the inclusion of appropriate sound channels for individual designs, and provide a method for systematically identifying the channels and the dimensions required in the representation.

The formalised model should also meet certain other requirements, based on the observations that were made in Section IV. These can be summarised as follows:

- We argued that sonification aesthetics, and mappings, require testing for the context in which they are used. The model should therefore facilitate the insertion of those data-sound mappings selected, according to experimental results and user preferences:

  Requirement 5: the model should not prescribe data-sound mappings.

- We also argued that the problem of listening fatigue may be reduced, if users can select their own music

and change it at will. Furthermore, experimentation with different musical aesthetics is required to determine those most suitable for the SOC environment. Therefore:

Requirement 6: the model should not prescribe musical genre, and should allow for choice in its selection.

B. Formalised Sonification Model

In Tables IV and V we present a formalised sonification model for designing musical parameter-mapping sonifications for use in network-security monitoring, developed to meet the requirements identified.

To construct the model, we divided Hermann's formalisation for the parameter-mapping sonification of a dataset [23] into individual mapping functions for *data channels* (corresponding to the channels identified in *Requirement 4*), *continuous data dimensions* and *discrete data dimensions* (corresponding to the dimensions identified in *Requirements 2* and 3). In Table IV, we define these *data channels* and *data dimensions*. Our approach is well-suited to this particular context because it allows us to reason about the channels of information to be presented for each particular use-case. Moreover, we can systematically identify continuous and discrete data, and their most appropriate mappings to sound. At the end of this section, we discuss how the model we develop meets the requirements identified.

The model comprises *components* (individual parts of the data and the sound to be mapped, which we present in Table IV), and *relations* (by which *components* are associated with one another, which we present in Table V). The *relations* are described by *mapping functions*.

A sonification is described by the tuple of its *components* and *relations* (the meaning of each of these is explained in Tables IV and V):

$$\langle CD_R, DD_R, VD_R, Rel_c, Rel_{d\alpha}, Rel_{d\beta}, Rel_v \rangle.$$

The *relations* presented in Table V are described by the *channel-mapping function* (which describes the *channel relation* Rel_c) and the *dimension-mapping function* (which describes both the *dimension relation* Rel_d and the *value relation* Rel_v). We also treat *sound dimensions ds* as functions of *sound channels cs*, which have values in the tuple of *sound values* of each *sound dimension*, *vs*.

The *channel-mapping function* $\boldsymbol{\psi}$:$\mathbb{R}^n \rightarrow \mathbb{R}^m$ describes the mapping from a tuple of *n data channels* $CD = \langle cd_1, ..., cd_n \rangle$ to a tuple of *m sound channels* $CS = \langle cs_1, ..., cs_m \rangle$. The *q*-dimensional sound signal $s(t)$ is computed as the sum over *m sound channels cs* of the *dimension-mapping function* $\boldsymbol{\Gamma}$: $\mathbb{R}^{m+1} \rightarrow \mathbb{R}^q$,

$$s(t) = \sum_{i=1}^{m} \boldsymbol{\Gamma}_i(cs_i, t),$$

where cs_i is the output of the *channel-mapping function* $\boldsymbol{\psi}$: $\mathbb{R}^n \rightarrow \mathbb{R}^m$ applied to the data channel cd_j and time t:

$$cs_i = \langle \boldsymbol{\psi}_i(cd_j, t) | j \in \{1, ..., n\} \rangle,$$

and $\boldsymbol{\Gamma}_i$ is the tuple of *dimension-mapping functions* $\boldsymbol{\gamma}_{ik}$, which are applied to the z data dimensions dd_{ik} of the data channels cd_j that were mapped by $\boldsymbol{\psi}_i$ to sound channel cs_i, and time t. The functions $\boldsymbol{\gamma}_{ik}$ describe the x continuous dimension mappings $\boldsymbol{\gamma}\alpha_1, ..., \boldsymbol{\gamma}\alpha_x$, and the y discrete dimension mappings $\boldsymbol{\gamma}\beta_1, ..., \boldsymbol{\gamma}\beta_y$, for each sound channel cs_i:

$$\boldsymbol{\Gamma}_i = \langle \boldsymbol{\gamma}_{i1}, ..., \boldsymbol{\gamma}_{iz} \rangle = \langle \boldsymbol{\gamma}\alpha_{i1}, ..., \boldsymbol{\gamma}\alpha_{ix}, \boldsymbol{\gamma}\beta_{i1}, ..., \boldsymbol{\gamma}\beta_{iy} \rangle.$$

We now explain how this model meets the requirements we identified. Since the *sound channels* and *sound dimensions* are left as an abstraction, *Requirements 5 and 6* are met. *Requirement 1* is also met through the use of abstract functions to describe the mappings themselves, meaning the resolution of the data presentation (number of data points presented) can be addressed through the choice of a function appropriate to any particular use of the model.

Requirement 4 is addressed by the division of the parameter-mapping into channels and dimensions; the *channel mapping function* addresses *Requirement 4*, while the *dimension mapping function* addresses *Requirement 2*. *Requirement 3* is met by the division of the *dimension mapping function* into a continuous and a discrete mapping function.

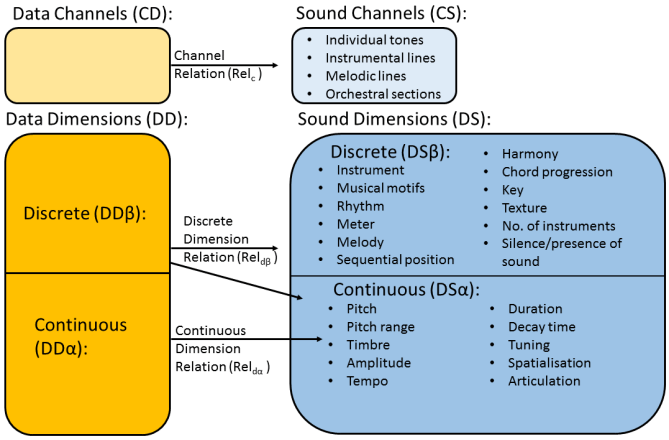

Figure 4. Data Sound Mappings Space of the Model

In Figure 4, we illustrate the space of data-sound parameter-mappings produced by the model. This shows the mappings from the sets of *data channels* and *data dimensions* (continuous and discrete) to possible *sound channels* and *sound dimensions*. We devised the list of *sound channels* and *sound dimensions* by drawing on sonification design literature such as a survey by Dubus et al. of sonification mappings used in prior work [57]; many of the items presented in Figure 4 are further described in that work. We also considered aspects of musical composition in creating these lists, which are not necessarily exhaustive, and can be added to.

VI. Application of the Model to Facilitate Prototype Design

To illustrate the application of the model, this section shows how we used it to design two prototype sonifications of network packet capture data, aimed at two different potential use-cases of sonification within network-security monitoring. We begin by presenting the two use-cases we considered. This is followed by an outline of the network attack characterisation that we used to derive the attack indicators to be represented for a defined network-monitoring scope. We demonstrate how the formalised model was applied, using these attack indicators, to generate prototype sonification systems for the two use-cases, and we describe the implementation of the prototypes.

TABLE IV. DESCRIPTION AND FORMAL NOTATION OF MODEL COMPONENTS

Component	Description	Formal Notation
Data channels	Parts of the network-security monitoring information, about which information should be presented, e.g., individual packets, IDS alerts, sensitive IP addresses on the network	The tuple CD of *data channels* cd
Data dimensions	Types of information we can present about *data channels*, e.g., amount of activity (at network IPs, for example), protocol used (in packet transmission), CPU usage (of network machines). These can have continuous or discrete values	The tuple DD of *data dimensions* dd. The tuple of data dimensions DD is the concatenation $DD\alpha^\frown DD\beta$ of the tuple $DD\alpha$ of *continuous data dimensions* $dd\alpha$, and the tuple $DD\beta$ of *discrete data dimensions* $dd\beta$
Data values	The values *data dimensions* can take. These can be continuous or discrete, e.g. a continuous scale from low to high (for packet rate, for example); discrete names (of protocols)	The tuple VD_{dd} of *data values* vd_{dd} of the data dimension dd
Sound channels	Streams of sound which we can vary sonically, e.g., individual note events, or separate melodic/instrumental lines	The tuple CS of *sound channels* cs
Sound dimensions	Types of sonic variations we can make to sound channels, e.g. varying the tempo or loudness at which they are presented, or the harmonic structure they follow. These can have continuous or discrete values	The tuple DS of *sound dimensions* ds. The tuple of sound dimensions DS is the concatenation $DS\alpha^\frown DS\beta$ of the tuple $DS\alpha$ of *continuous sound dimensions* $ds\alpha$, and the tuple $DS\beta$ of *discrete sound dimensions* $ds\beta$
Sound values	The values sound dimensions can take. These can be continuous or discrete, e.g. a continuous scale from slow to fast (tempo); discrete names of instruments	The tuple VS_{ds} of *sound values* vs_{ds} of the sound dimension ds

TABLE V. DESCRIPTION AND FORMAL NOTATION OF MODEL RELATIONS

Relation	Description	Formal Notation
Channel relation	*Data channels* are mapped to *sound channels*	*Channel relation* Rel_c: $CD \leftrightarrow CS$ is a total relation between the tuple of *data channels* and the tuple of *sound channels*
Dimension relation	*Data dimensions* are mapped to *sound dimensions* (which can be discrete or continuous) • *Continuous dimension relation*, in which *continuous data dimensions* are mapped to *continuous sound dimensions* • *Discrete dimension relation*, in which *discrete data dimensions* are mapped to *continuous* or *discrete sound dimensions*	*Dimension relation* Rel_d: $DD \leftrightarrow DS$ is a total relation between the tuple of *data dimensions* and the tuple of *sound dimensions* • *Continuous dimension relation* $Rel_{d\alpha}$: $DD\alpha \leftrightarrow DS\alpha$ is a total relation between the tuple of *continuous data dimensions* and the tuple of *continuous sound dimensions* • *Discrete dimension relation* $Rel_{d\beta}$: $DD\beta \leftrightarrow DS\beta$ is a total relation between the tuple of *discrete data dimensions* and the tuple of *discrete sound dimensions*
Value relation	Values of data dimensions are mapped to values of sound dimensions	For each *data dimension* dd, mapped to *sound dimension* ds, *value relation* Rel_{vdd}: $VD_{dd} \leftrightarrow VS_{ds}$ is a total relation between the tuple of *data values* of dd and the tuple of *sound values* of ds

Finally, we show how the model can be used to capture prior-art approaches to the sonification of network data.

A. Use-Cases

In Section II we highlighted potential advantages of using sonification for network monitoring. Here, we extend that discussion to create two different use-cases for sonification for network monitoring in SOCs. The first case focuses on enabling anomaly detection by security analysts deliberately listening to low-level network data, while the second case focuses on enabling peripheral monitoring of network-security information by security analysts as a non-primary task.

The two use-cases have different design requirements, since they target different modes of monitoring. Vickers differentiates between modes of auditory monitoring [56]. We associate *Use-Case 1* (as described below) with Vickers' description of the direct monitoring mode, in which the user deliberately listens to an audio interface as their main focus of attention, aiming to extract information or identify salient characteristics. *Use-Case 2* is associated with Vickers' peripheral monitoring mode, in which the user focuses attention on another primary task, while indirectly monitoring required information relating to another non-primary task, which is presented through a peripheral auditory display.

Use-Case 1: high-granularity sonification of network data to enable attack detection through pattern recognition by human security analysts: Humans have used sound in the past to detect anomalies with very high levels of resolution; an example is human sonar operators, who classify underwater sources by listening to the sound they make [63, 64]. Furthermore, sonification systems have been successfully designed for pattern recognition [58], and anomaly detection [59], for example, in prior work involving complex datasets.

The motivation for this use-case is that, as described in Section II, automated systems such as IDSs do not always detect attacks effectively or accurately, producing false-positives and false-negatives [2,3]. Presentation of data to humans in a visual form, using security visualisations, can enable detection of malicious network activity that is undetected by automated systems. Given the human ability for pattern recognition through listening, it should not be assumed that vision is the most effective medium for this in all cases without first comparing performances using vision and hearing experimentally [65].

To enable anomaly detection by humans, we aim to represent low-level network data with the highest granularity and resolution of information possible, such that patterns in the data may emerge naturally.

Use-Case 2: high-level sonification of network data for monitoring aspects of the network-security state as a non-

primary task: Analysts are required to carry out multiple tasks while monitoring the network for security breaches, maintaining an awareness of the security of the network [37]. This may mean, for example, continuing to monitor real-time network or IDS logs, while exploring or handling a potential security incident [66]. The aim of sonification in this use-case is to represent sonically the information that analysts need to maintain an appropriate awareness of the network-security state, in such a way that the information can be effectively monitored as a non-primary task. To produce a sonification suitable for use in peripheral monitoring tasks, we aim to present a higher level of information than in *Use-Case 1*: summaries of the data to enable comprehension of network-security state, rather than perception of anomalies and deviations from the normal.

Vickers argues that visual monitoring methods are not well suited to situations in which users are required to focus attention on a primary task, while monitoring other information directly, because of the demands this places on visual attention [56]. He summarises why sonification is well suited to monitoring peripheral information: "...the human auditory system does not need a directional fix on a sound source in order to perceive its presence". Experimental work has shown that sonification is an effective method of presenting information for monitoring as a secondary task. Hildebrandt et al. compared participants' performances in monitoring a simulated production process as a secondary task in three conditions – visual only, visual with auditory alerts, and visual with continuous sonification – while solving simple arithmetic problems as a primary task [67]. The results showed that participants performed significantly better in the secondary monitoring task using the continuous sonification than in the visual, or auditory alert, conditions. Furthermore, secondary monitoring using the continuous sonification had no significant effect on participants' performances in the primary task.

B. Data Requirements: Network Attack Characterisation

Despite the differences in the levels of information required for the two prototypes, the underlying data requirements are the same: for both cases, we require network data to be represented such that attacks are signalled by the sonification. We therefore used the same attack characterisation as the basis for both prototypes, enabling us to identify the network data that should be monitored to indicate the attacks within a defined network and monitoring information sources scope. We varied the treatment of the resulting data requirements in the application of the formalised model, taking into consideration the purpose of the prototype, and the required resolution of data presentation. In particular, for Prototype 1, we aimed to represent all attack indicators derived, while in Prototype 2 we focused on representing one particular attack indicator derived – the traffic rate at destination IP addresses on the monitored network, such that the amount of traffic received at each of these IPs may be monitored as a non-primary task.

We characterised the data requirements for representing indicators of attacks that can be detected within a network-monitoring scope; the scope was defined as follows:

1) The network is a local area network (LAN).
2) The network data monitored is packet header information, excluding packet contents.
3) Network data is monitored in real-time only (we, therefore, excluded aspects such as supply chain attacks on hardware components during manufacture and transportation).

With this scope in mind, we considered attacks in the Mitre Common Attack Pattern Enumeration and Classification (CAPEC) list (https://capec.mitre.org). This is a comprehensive listing and classification of computer attacks for use by, amongst others, security analysts. From this list, we selected attacks that fell within the defined scope. Excluding packet contents especially enabled us to narrow the scope of attacks considered initially to a list of around twenty types of attack, including reconnaissance such as port scanning, and threat realisation such as service flooding. This is because many of the attacks listed in CAPEC could not be detected by monitoring only packet header information without packet contents.

We characterised the attacks in terms of the way they are indicated through the network data monitored, i.e. packet header information. After completing this work, we were able to produce a summary of the data features needed to capture indicators of the attacks within the network monitoring scope. The data features we selected are defined as follows.

- *Packets*: the flow of packets into, out of or within the network.
    - *Rate*: the amount of traffic.
    - *Direction*: The direction in which network traffic is moving (entering network, leaving network, moving within network).
    - *Size*: the byte count of a packet.
    - *Protocol*: the protocol with which traffic is associated.
        - *Rate*: the amount of traffic transmitted using a particular protocol.
    - *Source IP*: the IP from which packets are sent, within or outside the network.
        - *Rate*: the amount of traffic associated with a source IP address.
        - *Range*: the number of source IP addresses as which traffic is observed.
    - *Destination IP/port*: the IP and port to which packets are sent, within or outside the network.
        - *Rate*: the amount of traffic associated with a destination IP address or port.
        - *Range*: the number of destination IP addresses or ports at which traffic is observed.

The derived data features are shown in Table VI. In the table, the leftmost three columns display the data features, while the rightmost three columns show the characterisation of three examples of attacks (TCP SYN scan, data exfiltration and DDoS) in terms of these features. The data features entered in each column are characteristics of those in the preceding column. For example, *rate* (third column) is a characteristic of *source IP*, (second column), which is itself a characteristic of *packet* (first column). The attack characterisation columns show how we used the data features to characterise three different network attacks. For example, given a data exfiltration attack, the data features listed in the second attack characterisation column of Table VI are required.

TABLE VI. NETWORK ATTACK CHARACTERISATION EXAMPLES AND DERIVED DATA PRESENTATION REQUIREMENTS

Required Data			**Attack characterisation**			
Features	**Features (Characteristics of First Column)**	**Features (Characteristics of Second Column)**	**Attack type:**	**TCP SYN scan**	**Data exfiltration**	**DDoS**
			Attack description:	TCP protocol, SYN packets sent to a range of destination ports on a host	Data exfiltrated from network to external address	Network is flooded by a high wide of traffic sent from multiple hosts
Packet						
	Rate					High
	Direction			Inbound	Outbound	Inbound
	Size					
	Protocol				FTP	
		Rate			High	
	Source IP			Single IP outside network	Single IP inside network	Multiple IPs outside network
		Rate		High	High	High
		Range				Wide
	Destination IP			Single IP inside network	Single IP outside network	One or more IPs inside network
		Rate		High	High	High
		Range				
	Destination port			Ports on single host IP inside network		
		Rate				
		Range		Wide (all ports targeted – scan)		

C. Prototype 1: Low-Level Network Data Sonification for Anomaly Detection by Humans

The aim of Prototype 1 is to sonically represent network data through which an attack might be signalled with as high a resolution as possible, in order to enable anomaly detection through emerging sound patterns. We show how we applied the model in the design of the sonification by considering appropriate data channels, dimensions and values. We develop a prototype design, and highlight challenges in the implementation.

Here we seek to present prototype designs. The purpose is not to develop final system designs, but to illustrate the use of the sonification model for designing sonifications for particular use-cases within network-security monitoring, and to demonstrate how the application of the model can be varied depending on the use for which the sonification is intended.

1) Applying the Sonification Model: We derived the *data channels*, *data dimensions* and *data values* for the prototype using the data requirements presented in Table IV. In this case, in order to achieve the highest possible resolution in the sonification of these data requirements, we aimed to present, as closely as possible, each packet captured, and to represent as much information about each packet as possible as dimensions of the packet channel. We therefore let the entries in the first column (a single entry: packets) of Table VI be the tuple of *data channels*, and all entries in the second column be the tuple of *data dimensions*. The entries in the third column are then conveyed naturally, through the mapping of the selected data channels and dimensions (for example, range of source IPs – third column – does not have a defined mapping, but is presented through the cumulation of the presentation of the source IP dimension for each individual packet).

For this prototype, the sonification is described by the tuple $\langle CD_R, DD_R, VD_R, Rel_c, Rel_{d\alpha}, Rel_{d\beta}, Rel_v \rangle$:

- $CD_R = \langle cd_{R1} \rangle = \langle packets \rangle$
- $DD_R = DD\alpha_R^\frown DD\beta_R = \langle dd\alpha_{R1}, dd\alpha_{R2}, dd\alpha_{R3}, dd\alpha_{R4}, dd\alpha_{R5} \rangle^\frown \langle d\beta_{dR1}, d\beta_{dR2} \rangle = \langle$Source IP, Destination IP, Destination Port, Rate, Size$\rangle^\frown \langle$Direction, Protocol$\rangle$
- VD_{dR}=$\langle vd_{dR1}, vd_{dR2}, vd_{dR3}, vd_{dR4}, vd_{dR5}, vd_{dR6}, vd_{dR7} \rangle = \langle \{1,...,2^{32}\}, \{1,...,2^{32}\}, \{1,...,2^{16}\}, \{$low,normal,high$\}, \{$small,normal,large$\}, \{$incoming,outgoing,internal$\},$(the protocols present in the dataset)$\rangle$
- Rel_c is described by the function $\boldsymbol{\psi}_i : \mathbb{R}^1 \rightarrow \mathbb{R}^m$, $cs_i = \boldsymbol{\psi}_i(cd_1)$
- Rel_d and Rel_v are described by the function $\boldsymbol{\Gamma} : \mathbb{R}^{m+1} \rightarrow \mathbb{R}^q$, $\boldsymbol{\Gamma}_i$=$\langle \boldsymbol{\gamma}\alpha_{i1}, ..., \boldsymbol{\gamma}\alpha_{ix}, \boldsymbol{\gamma}\beta_{i1}, ..., \boldsymbol{\gamma}\beta_{iy} \rangle$ $\forall i \in \{1,...,m\}$

We assume that the IP version is IPV4 in the above description of source and destination IP values. Although source and destination ports and IPs are not technically continuous they have such a high number of possible values (2^{32} for IPV4) that we treat them as such. In describing some data values, we used a notion of "normal". This is left as an abstraction in the model, and describes some *expectation* for the observed behaviour of the data dimensions. We discuss how this normal abstraction might be implemented in sonification designs in Section VI.E.

To simplify the design process, we describe data values for rate and size as discrete points of interest (for example,

low, *high*, *narrow*, *wide*). This description does not exclude the possibility of mapping continuously in the representation, but allows indication of the polarity required in dimension mapping. In sonification, polarity is the direction of the mapping from data to sound. For example, *positive* mapping polarity from the data dimension *rate* to the sound dimension *amplitude* would be described:

- rate: high → amplitude: loud;
- rate: low → amplitude: soft.

In Figure 5, we present the sonification mapping space introduced in Figure 4, applied to Prototype 1. This shows the *data channels*, *continuous data dimensions* and *discrete data dimensions* with all possible mappings to *sound channels* and *sound dimensions*.

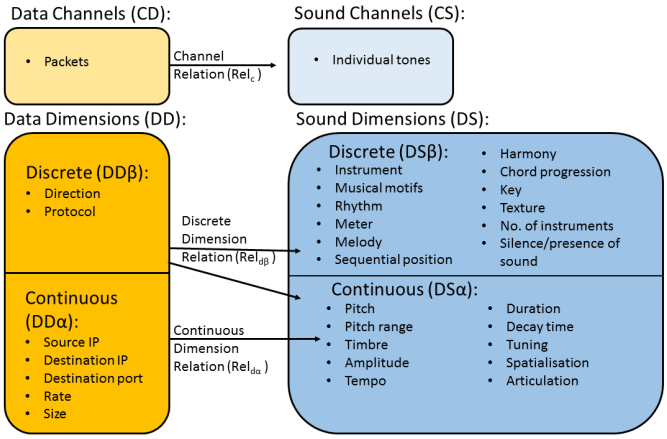

Figure 5. Data Sound Mappings Space: Prototype 1

To determine the mappings from data to sound for the prototype, we selected *sound channels*, *continuous sound dimensions* and *discrete sound dimensions* from the sets *CS*, *DS*α and *DS*β respectively. We have not yet carried out testing of appropriate mappings for the data dimensions, as described in Section IV; this is left to future work. Instead, we made predictions about appropriate mappings at this stage, drawing on a survey of mappings used in the sonification of physical quantities in prior sonification work [57]. In that paper, prior work in which physical quantities are sonified is surveyed, and it is noted whether data-sound mappings were: assessed as good; assessed as poor; implemented but not assessed; or not implemented but mentioned as future work. We applied those assessed as good for quantities we considered representative of our data dimensions (for example, for the data dimension *rate*, we considered the physical quantities velocity, activity and event rate from [57]). From this, we derived the following information, which was then incorporated into the prototype design.

- **Rate.** Good mappings described for *velocity*: pitch, brightness, tempo, rhythmic duration. Good mappings described for *activity*: tempo, rhythmic duration [57].
- **Size.** Bad mappings described for *size*: pitch, tempo [57].

Applied to our sound mappings space, this generated the following rules.

- rate → pitch, tempo, rhythmic duration
- size NOT → pitch, tempo.

We thus arrived at the following set of *relations* for the prototype design.

- *Data channels*:
  - Packet → individual tone ($cs_1 = \boldsymbol{\psi}_1(cd_1,t)$)
- *Data dimensions (continuous)*:
  - Rate → tempo (positive polarity) ($ds_{11} = \boldsymbol{\gamma}\alpha_1(dd_{11},t)$)
  - Destination IP → spatialisation (pan from left to right headphone) ($ds\alpha_{12} = \boldsymbol{\gamma}\alpha_2(dd\alpha_{12},t)$)
  - Source IP → pitch ($ds\alpha_{13} = \boldsymbol{\gamma}\alpha_3(dd\alpha_{13},t)$)
  - Destination port → articulation ($ds\alpha_{14} = \boldsymbol{\gamma}\alpha_4(dd\alpha_{14},t)$)
  - Size → amplitude (positive polarity) ($ds\alpha_{15} = \boldsymbol{\gamma}\alpha_5(dd\alpha_{15},t)$)
- *Data dimensions (discrete)*:
  - Protocol → instrument ($ds\beta_{11} = \boldsymbol{\gamma}\beta_1(dd\beta_{11},t)$)
  - Direction → register ($ds\beta_{12} = \boldsymbol{\gamma}\beta_2(dd\beta_{12},t)$)

Figure 6 shows the prototype design developed from these *relations*. In this sonification, each packet observed triggers an individual note event; these events shown as musical notes in Figure 6. The above dimension mappings are represented: the sonification maps data dimensions to the sound dimensions (including instrument, for example) of each note. The sound is panned on a continuous scale between left and right, corresponding to the continuous destination IP dimension. The rate of traffic at each destination IP is represented by the tempo of the notes played at that pan location; source IPs map continuously to frequency such that source IP range is represented by the range of frequencies played. As shown in Figure 6, destination ports map to articulation on a continuous scale. The instrument by which each note is played represents the protocol in which the packet is transmitted, and the direction of traffic is conveyed by differentiating between low, medium and high registers of music.

D. Prototype 2: High-Level Network-Data Sonification for Monitoring Network-Security Information as a Non-Primary Task

The aim of Prototype 2 is to enable security analysts to monitor network data for indicators of attacks as a non-primary task. The sonification must represent aspects of the network data that might signal an attack, in a way that is unobtrusive usually, but draws analysts' attention to aspects of the data when required (when a potential attack indicator arises). The use-case is different to an alert system: the goal here is to be informative about which data has changed, and how it has changed.

Our design approach for this use-case is to sonify a subset of the data through which attacks are indicated – the traffic rate at the destination IP addresses on the network. The rationale for this approach is that to be suitable for peripheral monitoring, the sonification should be uncomplicated to understand. We therefore elect not to sonify all indicators derived in our network-attack characterisation, as in Prototype 1, but to produce a simpler representation of a subset of these indicators. Our network-attack characterisation showed that high traffic rates at particular destination IP addresses on the network were frequently indicators of attacks (see Table VI).

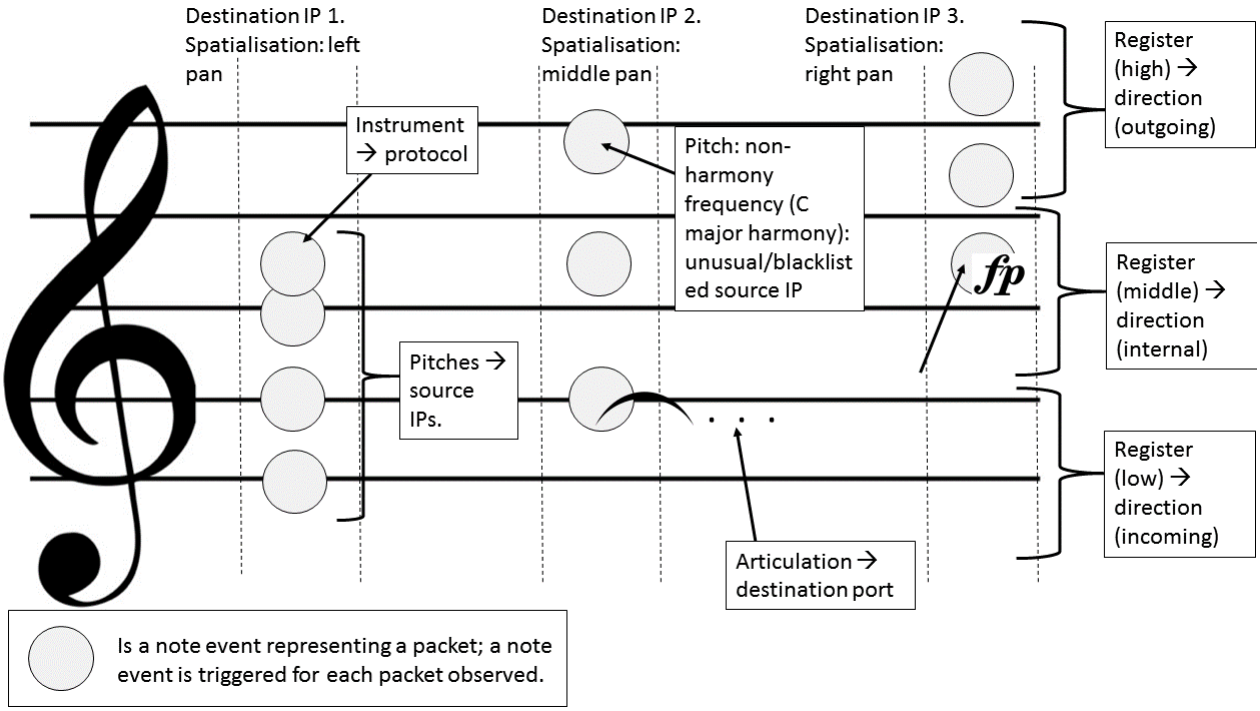

Figure 6. Prototype Diagram: Prototype 1

Monitoring the amount of activity at sensitive machines on the network such as those on which databases containing sensitive information are stored is important, and we selected this as the aim of this peripheral-monitoring sonification prototype. In the remainder of this section we show how this difference in approach influences the application of the model and leads to differing prototype designs.

1) Applying the Sonification Model: We derived the *data channels*, *data dimensions* and *data values* for the prototype by considering the data requirement: present the traffic rate at destination IP addresses on the network. Given the purpose of the sonification is to be suitable for use in peripheral monitoring, we aim to present sonified information such that data changes judged significant (in this case, large increases in traffic rate at any destination IP represented) draw attention, and the sonification is otherwise unobtrusive. We let 10 individual destination IP addresses be the data channels, and the packet rate be the data dimension.

For this prototype, the sonification is described by the tuple $\langle CD_R, DD_R, VD_R, Rel_c, Rel_{d\alpha}, Rel_{d\beta}, Rel_v \rangle$:

- $CD_R = \langle cd_{R1} \rangle = \langle$Destination IP addresses$\rangle$
- $DD_R = DD\alpha_R \frown DD\beta_R = \langle dd\alpha_{R1}, dd\alpha_{R2}, dd\alpha_{R3} \rangle \frown \langle d\beta_{dR1} \rangle = \langle$Rate$\rangle$
- $VD_{dR} = \langle vd_{dR1} \rangle = \langle \{$low,normal,high$\} \rangle$
- Rel_c is described by the functions $\boldsymbol{\psi}_i : \mathbb{R}^{10} \to \mathbb{R}^m$, $cs_i = \langle \boldsymbol{\psi}_i(cd_j) | j \in \{1,...,n\} \rangle$, where n is the number of network destination IP addresses represented
- Rel_d and Rel_v are described by the function $\boldsymbol{\Gamma} : \mathbb{R}^{m+1} \to \mathbb{R}^q$, $\boldsymbol{\Gamma}_i = \langle \boldsymbol{\gamma}\alpha_{i1}, ..., \boldsymbol{\gamma}\alpha_{ix}, \boldsymbol{\gamma}\beta_{i1}, ..., \boldsymbol{\gamma}\beta_{iy} \rangle$ $\forall i \in \{1,...,m\}$

The notes on the prescription of a "normal" value, and representations of polarity, following the presentation of the sonification for Prototype 1, hold for this case also: the "normal" packet rate for each IP could be prescribed by a human, set as an average calculated statistically, or calculated using Machine Learning.

In Figure 7, we present the sonification mapping space introduced in Figure 4, applied to Prototype 2. This shows the *data channels* and *continuous data dimensions* (there are no *discrete data dimensions* in this case) with all possible mappings to *sound channels* and *sound dimensions*.

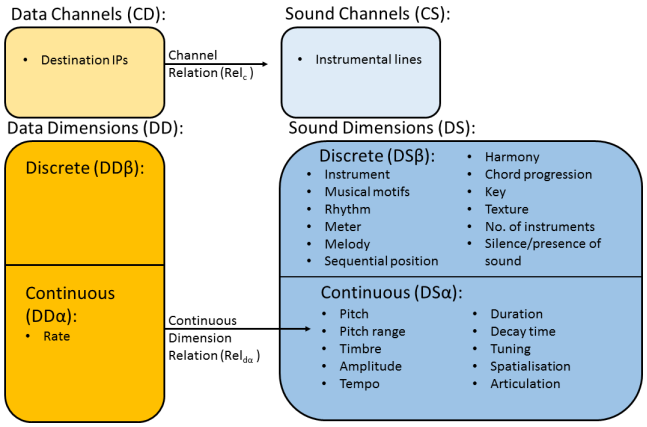

Figure 7. Data Sound Mappings Space: Prototype 2

As described for Prototype 1, we drew on prior work [57] to select from the set of *sound channels* and *continuous sound dimensions*. The **relations** we arrived at are as follows.

- *Data channels*
  - 10 destination IP addresses → 10

instrumental lines
$(cs_i = \boldsymbol{\psi}_i(cd_i,t) \forall i \in \{1,...,10\})$

- *Data dimensions (continuous)*:
  - Rate $\rightarrow$ tempo (positive polarity)
    $(ds\alpha_{i1} = \boldsymbol{\gamma}\alpha_1(dd\alpha_{i1},t) \forall i \in \{1,...,10\})$

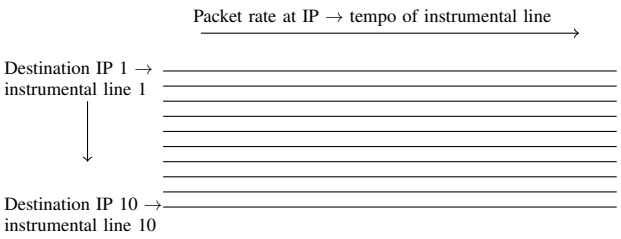

Figure 8. Prototype Diagram: Prototype 2

Figure 8 shows the prototype design developed from these relations. In this sonification, the individual instrumental lines that form the musical piece each present information about an individual destination IP address on the network (in the figure, we present an example in which 10 destination IP addresses are monitored). The lines each follow the base tempo of the musical piece when the packet rate at the destination IP addresses they represent is at its "normal" value. When the packet rate at an individual destination IP address exceeds its "normal" value, note repetition is introduced in the corresponding instrumental line, and the speed of note repetition is scaled to convey the size of the increase in packet rate. As such, a destination IP with a high traffic rate is represented in the sonification as an instrumental line with fast note repetition.

E. Implementation of Prototypes

We implemented both prototypes and used them to sonify the Centre for Applied Internet Data Analysis (CAIDA) "DDoS Attack 2007" dataset [68]. We describe our processes for, and the challenges that arose during, implementation of this dataset. The dataset contains a DDoS attack in which a large flood of incoming traffic is observed, sent from a wide range of source IP addresses to destination IP addresses on the network. We reflect on the sounds produced by the two sonifications of this dataset; in particular, the sounds produced at the time when the flooding begins compared with the sounds prior to the flooding.

We implemented the prototypes by reading the dataset in Python, and parsing the data values according to the mapping functions presented in Tables VII and VIII. These parsed values were then rendered as sound using Supercollider (http://supercollider.github.io/), a platform for audio programming and synthesis frequently used in prior sonification work. The sound rendering was controlled by Open Sound Control (OSC) messages sent from Python to Supercollider.

Although we have not yet conducted user testing for these prototypes, our initial assessment from listening to the sonifications ourselves is that there is a significant change in sound in both prototypes at the time that the dataset shows flooding from multiple source IPs. We invite the reader to listen to audio clips of each of the two network-security monitoring prototypes running on this dataset (https://soundcloud.com/user-71482294).

We encountered some challenges during the implementation phase; in the following, we reflect on possible solutions to these challenges, and hence identify directions for future development. The most significant challenge in the implementation of Prototype 1 arose as a result of the sheer number of packets logged in the dataset, and the small times between their arrival. Because of this, it was challenging to implement the *channel relation* $\boldsymbol{\psi}_1$ – to render each packet observed as individual notes without overloading the sound engine, or creating sounds too complicated to be of use to human listeners.

We sampled randomly every 1 in 10 packets in the dataset to address this challenge; however, as future work it is important to investigate the most appropriate methods of aggregation, sampling and scaling. For example, a solution might be to aggregate the packets sent in each individual connection (between the same two IP addresses and ports, and using the same service) over time intervals (for example, every 0.1 seconds), and represent the aggregation over this time interval with a single note, whose amplitude varies depending on the number of packets aggregated in this time. This would be a potential way of addressing the problem of packet rates too fast to sonify, without losing the granularity of information provided by the representation of each individual packet. Grond and Berger write that sonification mapping functions may sometimes be linear, but other forms may be more suitable depending on the data [56]. Scaling exponentially, or using methods such as step-change analysis or Fourier Transforms, are examples of avenues worth exploring. Establishing the resolution with which we can represent each of the listed *data channels* and *data dimensions* will be a key part of the development and testing process.

The destination IP representation approach of Prototype 2 may become challenging on large networks. The aim of the prototype is to represent monitoring information in a relatively simple fashion suitable for peripheral monitoring. However, the number of instrumental lines required to represent the many destination IP addresses on a large organisational network would likely introduce complexity to the sonification and make extracting information about individual IP addresses difficult for the user. It is important to investigate experimentally how much information we can represent – in this case, how many destination IP addresses we can represent information about simultaneously in a way that is useful, and whether this can match the monitoring requirements of SOCs in large organisations.

F. Addressing Prior-Art Approaches Using the Sonification Model

We describe the use of our formalised sonification model in representing previously-published sonification system designs. In particular, we verify that our model can address all previous systems (those in which the sonification design is specified completely) that use a musical parameter-mapping sonification method to represent raw network data (these aspects of the systems are presented in Table II) [29, 33, 43–45]. Other relevant systems which use a musical parameter-mapping approach to represent raw network data are presented in [30, 31, 42], but the sonification designs for these works are not specified in enough detail to include.

In Table IX we present the relevant pre-existing sonification systems in terms of the *data channels* and *data dimensions*, and the *sound channels* and *sound dimensions* of our model. In Table X we present the *channel relations* and *dimension relations* for each prior sonification approach

TABLE VII. IMPLEMENTATION OF PROTOTYPE 1

Relation Addressed	**Description of Implementation**	**Mapping Function**
Channel relation: $cs_1 = \boldsymbol{\psi}_1(cd_1, t)$	Individual packets observed are mapped to individual tones	The function $\boldsymbol{\psi}_1$ can be described: for the p^{th} packet cd_{1p} observed at time t, play a single tone cs_{1p} at time t
Dimension relation: $ds\alpha_{12} = \boldsymbol{\gamma}\alpha_2(dd\alpha_{12}, t)$	Destination IP is mapped to spatialisation (pan from left to right headphone). Here, the possible destination IP addresses take values in the range $[0, 2^{32}]$, we converted destination IP addresses to values in this range using a function such that IP address 0.0.0.0 $\rightarrow$ 0, and 0.0.0.1 $\rightarrow$ 1. The pan value varies continuously in the range $[-1, 1]$	The function $\boldsymbol{\gamma}\alpha_2$ can be described: for a note cs_{1p} played at time t, and IP conversion function *IPVal*, the pan value is $ds\alpha_{12p} = \frac{IPVal(dd\alpha_{12p})\times 2}{2^{32}} - 1$
Dimension relation: $ds\alpha_{13} = \boldsymbol{\gamma}\alpha_3(dd\alpha_{13}, t)$	Source IP is mapped to pitch. Here, the possible source IP addresses take values in the range $[0, 2^{32}]$ and the frequencies vary in the chosen range $[261.63, 2093]$. Frequency 261.63Hz corresponds to C4 – middle C – while frequency 2093Hz corresponds to C7, three octaves higher. We also use a hotlisting method: the top 50 source IPs we expect to observe are mapped to harmonic tones (the notes of a C major 7^{th} chord), while source IPs outside this hotlist are mapped on a continuous scale to frequencies in the selected range	For a source IP hotlist tuple H_s, and tuple M_n of musical notes $\langle C, E, G, B\rangle$, the function $\boldsymbol{\gamma}\alpha_3$ can be described: for note cs_{1p} at time t, and IP conversion function *IPVal*, the pitch value is $dd\alpha_{13p} \in H_s \implies ds\alpha_{13p} \in M_n$, $dd\alpha_{13p} \notin H_s \implies ds\alpha_{13p} = \frac{IPVal(dd\alpha_{13p})\times(2093-261.63)}{2^{32}} + 261.63$
Dimension relation: $ds\alpha_{14} = \boldsymbol{\gamma}\alpha_4(dd\alpha_{14}, t)$	Destination port is mapped to articulation. Here, the possible destination ports take values in the range $[0, 2^{16}]$, and the articulation takes values in the range $[0.1, 1]$. Many packets observed in this dataset did not have destination port values; in these cases we set the sound articulation value to be 0.5 in Supercollider.	The function $\boldsymbol{\gamma}\alpha_4$ can be described: for a note cs_{1p} played at time t, the articulation value is $ds\alpha_{14p} = \frac{dd\alpha_{14p}\times 0.9}{2^{16}} + 0.1$
Dimension relation: $ds\alpha_{15} = \boldsymbol{\gamma}\alpha_5(dd\alpha_{15}, t)$	Size is mapped to amplitude (positive polarity). Here, for the dataset we implemented the average packet size was 60 bytes, while occasional packet sizes were very large. We mapped the size values of the packets to the amplitude values of the sound using a logarithmic function, in which the average packet size, 60, mapped to an amplitude value we judged "comfortable" – the amplitude value 1 in Supercollider.	The function $\boldsymbol{\gamma}\alpha_5$ can be described: for a note cs_{1p} played at time t, the amplitude value is $ds\alpha_{15p} = \frac{1}{2}(log_{10}(\frac{dd\alpha_{15p}}{60} \times 100))$
Dimension relation: $ds\beta_{11} = \boldsymbol{\gamma}\beta_1(dd\beta_{11}, t)$	Protocol is mapped to instrument. Here, a hotlisting method is used again. The two protocols most frequently seen in this dataset are mapped onto two different instruments; the remaining protocols are mapped to another instrument. For this dataset, the tuple of hotlisted protocols is: $H_p = \langle$ICMP, TCP$\rangle$, and the tuple of instruments selected was: $M_i = \langle$strings, saxophone, piano$\rangle$	The function $\boldsymbol{\gamma}\beta_1$ can be described: for a note cs_{1p} played at time t, the instrument value is $dd\beta_{11p} \in H_p \implies ds\beta \in \langle M_{i1}, M_{i2}\rangle$, $dd\beta_{11p} \notin H_p \implies ds\beta = M_{i3}$

TABLE VIII. IMPLEMENTATION OF PROTOTYPE 2

Relation Addressed	**Description of Implementation**	**Mapping Function**
Channel relation: $cs_i = \boldsymbol{\psi}_i(cd_i, t) \forall i \in \{1, ..., 10\}$	Destination IP addresses within a hotlist of 10 addresses $H_d = \langle dst_1, ..., dst_{10}\rangle$ are mapped to 10 musical lines in the tuple $M = \langle m_1, ..., m_{10}\rangle$	The function $\boldsymbol{\psi}_i$ can be described: at any time t, play all musical lines $m_i \in M$
Dimension relation: $ds\alpha_{11} = \boldsymbol{\gamma}\alpha_1(dd\alpha_{i1}, t) \forall i \in \{1, ..., 10\}$	Rate is mapped to tempo (positive polarity), scaled such that the average rate for a particular destination IP is mapped to the base tempo of the music. The rate is measured by aggregating the number of packets observed at each IP per second, and comparing this with the average number to derive the tempo for the corresponding second of music	The function $\boldsymbol{\gamma}\alpha_1$ can be described: for a musical instrumental line $m_i \in M$ played at time t, where the average rate for the corresponding destination IP address dst_i is $avrate_i$ and the base tempo of the music is $avtempo$, the tempo value is $ds\alpha_{i1} = \frac{dd\alpha_{i1}}{avrate_i} \times avtempo$

TABLE IX. APPLYING THE FORMALISATION TO CAPTURE PREVIOUS MUSICAL PARAMETER-MAPPING SYSTEMS FOR THE SONIFICATION OF RAW NETWORK DATA: COMPONENTS

Author	**Data Channels**	**Data Dimensions**	**Sound Channels**	**Sound Dimensions**
Qi [43] Mapping 1:	Traffic queue 16 (cd_1)	**Continuous**: byte rate ($dd\alpha_{11}$); packet rate ($dd\alpha_{12}$)	Piano notes (cs_1)	**Continuous**: frequency ($ds\alpha_{11}$); amplitude ($ds\alpha_{12}$)
Qi [43] Mapping 2:	Traffic queues 1–16 ($cd_1, ..., cd_{16}$)	**Continuous**: byte rate ($dd\alpha_{11}$); packet rate ($dd\alpha_{12}$)	16 groups of piano notes ($cs_1, ..., cs_{16}$)	**Continuous**: frequency ($ds\alpha_{11}$); amplitude ($ds\alpha_{12}$)
Brown [44]	Network traffic (cd_1)	**Continuous**: packet rate ($dd\alpha_{11}$; number of TCP handshakes ($dd\alpha_{12}$); number of HTTP error messages ($dd\alpha_{13}$)	Existing musical piece (cs_1)	**Continuous**: number of sharp notes ($ds\alpha_{11}$); pitch ($ds\alpha_{12}$); rhythm ($ds\alpha_{13}$)
Ballora [33]	Socket exchanges (cd_1); requests to unusual ports (cd_2); traffic in 5 different monitoring locations (within 2 subnets; between subnets; external traffic going to each subnet) (cd_3)	**Continuous**: source IP ($dd\alpha_{11}$); destination IP($dd\alpha_{12}$); frequency of packets in ongoing socket connections ($dd\alpha_{13}$); traffic rate ($dd\alpha_{34}$) **Discrete**: port number ($dd\beta_{21}$)	An individual strike of a gong (cs_1); humming sound (cs_2); 5 distinct whooshing sounds (cs_3)	**Continuous**: rumble's timbre ($ds\alpha_{11}$); sizzle's timbre ($ds\alpha_{12}$); stereo pan position ($ds\alpha_{13}$); force of strike ($ds\alpha_{14}$); timbre (of humming sound) ($ds\alpha_{25}$); amplitude (of whooshing sound) ($ds\alpha_{36}$)
Giot [29]	Packets (cd_1); useless packets (e.g. ACK packets) (cd_2)	**Continuous**: packet size ($dd\alpha_{11}$); time-to-live (TTL) ($dd\alpha_{12}$); rate/bandpass ($dd\alpha_{13}$); number of useless packets ($dd\alpha_{21}$) **Discrete**: Protocol ($dd\beta_{11}$)	Individual note events (MIDI) (cs_1); noise (cs_2)	**Continuous**: frequency ($ds\alpha_{11}$); note duration ($ds\alpha_{12}$); bandpass of resonant filter ($ds\alpha_{13}$); amount of noise ($ds\alpha_{24}$) **Discrete**: sound synthesiser ($ds\beta_{11}$);
Mancuso [45]	Individual packets (cd_1)	**Continuous**: source IP ($dd\alpha_{11}$); destination IP ($dd\alpha_{12}$) **Discrete**: packet size ($dd\beta_{11}$)	String note (cs_1); wind note (cs_2)	**Continuous**: pitch ($ds\alpha_{11}$, $ds\alpha_{21}$); amplitude ($ds\alpha_{12}, ds\alpha_{22}$)

TABLE X. APPLYING THE FORMALISATION TO CAPTURE PREVIOUS MUSICAL PARAMETER-MAPPING SYSTEMS FOR THE SONIFICATION OF RAW NETWORK DATA: RELATIONS

Author	Channel Relations	Dimension Relations
Qi [43] Mapping 1:	Single traffic queue $\rightarrow$ all piano notes ($cs_1 = \psi(cd_1)$)	Byte rate $\rightarrow$ frequency ($ds\alpha_{11} = \gamma\alpha_1(dd\alpha_{11}, t)$); packet rate $\rightarrow$ amplitude ($ds\alpha_{12} = \gamma\alpha_2(dd\alpha_{12}, t)$)
Qi [43] Mapping 2:	Traffic queue $i \rightarrow$ piano notes group i ($cs_i = \psi(cd_i) \forall i \in \{1, ..., 16\}$)	Byte rate $\rightarrow$ frequency ($ds\alpha_{i1} = \gamma\alpha_1(dd\alpha_{i1}, t) \forall i \in \{1, ..., 16\}$); packet rate $\rightarrow$ amplitude ($ds\alpha_{i2} = \gamma\alpha_2(dd\alpha_{i2}, t) \forall i \in \{1, ..., 16\}$)
Brown [44]	Network traffic $\rightarrow$ existing musical piece ($cs_1 = \psi(cd_1)$)	Traffic rate $\rightarrow$ number of sharp notes ($ds\alpha_{11} = \gamma\alpha_1(dd\alpha_{11}, t)$); number of TCP handshakes $\rightarrow$ pitch ($ds\alpha_{12} = \gamma\alpha_2(dd\alpha_{12}, t)$); number of HTTP error messages $\rightarrow$ rhythm ($ds\alpha_{13} = \gamma\alpha_2(dd\alpha_{13}, t)$)
Ballora [33]	Socket exchange $\rightarrow$ individual strike of gong ($cs_1 = \psi(cd_1)$); request to unusual port $\rightarrow$ humming sound; traffic in five different monitoring locations $\rightarrow$ five distinct whooshing sounds	Source IP $\rightarrow$ gong rumble's timbre ($ds\alpha_{11} = \gamma\alpha_1(dd\alpha_{11}, t)$); destination IP $\rightarrow$ gong sizzle's timbre ($ds\alpha_{12} = \gamma\alpha_2(dd\alpha_{12}, t)$); source IP, destination IP $\rightarrow$ stereo pan position ($ds\alpha_{13} = \gamma\alpha_3(dd\alpha_{11}, dd\alpha_{12}, t)$); frequency of packets $\rightarrow$ force of strike ($ds\alpha_{14} = \gamma\alpha_4(dd\alpha_{13})$); port number $\rightarrow$ timbre of humming sound ($ds\alpha_{25} = \gamma\beta_1(dd\beta_{21})$); traffic rate $\rightarrow$ amplitude of whooshing sound ($ds\alpha_{36} = \gamma\alpha_4(dd\alpha_{34})$)
Giot [29]	Packets $\rightarrow$ individual note events ($cs_1 = \psi(cd_1)$); useless packets $\rightarrow$ noise ($cs_2 = \psi(cd_2)$)	Packet size $\rightarrow$ frequency ($ds\alpha_{11} = \gamma\alpha_1(dd\alpha_{11}, t)$; TTL $\rightarrow$ note duration ($ds\alpha_{12} = \gamma\alpha_2(dd\alpha_{12})$); rate $\rightarrow$ bandpass of resonant filter ($ds\alpha_{13} = \gamma\alpha_3(dd\alpha_{13})$); protocol $\rightarrow$ sound synthesiser ($ds\beta_{11} = \gamma\beta_1(dd\beta_{11})$); number of useless packets $\rightarrow$ amount of noise ($ds\alpha_{24} = \gamma\alpha_4(dd\alpha_{21})$)
Mancuso [45]	Individual packets $\rightarrow$ string note, wind note ($cs_1 = \psi(cd_1)$, $cs_2 = \psi(cd_1)$)	Source IP $\rightarrow$ pitch of string note ($ds\alpha_{11} = \gamma\alpha_1(dd\alpha_{11}, t)$); destination IP $\rightarrow$ pitch of wind note ($ds\alpha_{21} = \gamma\alpha_2(dd\alpha_{12}, t)$); packet size $\rightarrow$ amplitude (of string note and wind note) ($ds\alpha_{12} = \gamma\beta_1(dd\beta_{11})$), ($ds\alpha_{22} = \gamma\beta_1(dd\beta_{11})$)

addressed. This shows that the systems addressed can all be represented in our model, which allows for comparative testing of newly-developed sonification systems against pre-existing approaches.

VII. Conclusion and Future Work

We conclude that there is a growing requirement for the validation of sonification as a means of improving certain monitoring capabilities in SOCs. The current state of the art provides evidence of the potential of sonification in advancing network-security monitoring capabilities. Systems proposed and in use have been shown to be as effective as, or more effective than, other network monitoring techniques insofar as a limited amount of testing has been performed [19].

As future work, we intend to perform proof-of-concept experiments for the sonification prototypes. For Prototype 1, we will sonify a number of network packet capture datasets containing instances of network attacks using the prototype, and assess whether "patterns" appear, or deviations from the "normal" sound of the sonification are heard, at the time of the attacks. For Prototype 2, we will assess experimentally whether the sonification conveys the packet rate at individual destination IP addresses on the network, in a way suitable for monitoring as a non-primary task.

As described in Section IV, a key stage in the sonification development is experimental identification of appropriate aesthetics: intuitive mappings from data to sound, for example. We have applied mappings in both presented prototypes based on our own intuition, and relevant aspects of prior work [57]. A direction for future work is conducting design experiments to determine the optimal mapping aesthetics, and incorporating these mappings into the formalised sonification model to generate final system designs. To assess the effectiveness of our sonification model and aesthetic approach, we need to contrast our approach with pre-existing approaches to parameter-mapping sonification for network-security monitoring [28, 33, 36, 43, 45], by comparing their performance in highlighting network attacks.

During the presentation of prototypes, we highlighted our use of a "normal" in describing the values of certain data dimensions. A challenge in the implementation of the prototypes lies in determining appropriate meanings of this "normal", which is left as an abstraction in the model. The normal might in practice be defined, or calculated using Statistics or Machine Learning for a particular network. The normal could also be defined not by the system itself, but discerned by the humans using the system, based on what they expect to be, or have become accustomed to, hearing. The former approach is likely to be more appropriate for enabling the peripheral monitoring capability targeted in *Use-Case 2*, while the latter (in which humans learn to "hear" some normal) may apply to *Use-Case 1*, given the aim to enable humans to detect anomalies.

Alternative methods of extracting the data requirements for network-attack detection should be explored. The attack characterisation approach taken here could be extended, and validated, through security analysts' input on their real network-data monitoring requirements. This should explore both how analysts detect anomalies indicating attacks through network data, and which aspects of network data they may realistically be required to monitor as a non-primary task (for addressing *Use-Case 2* in particular).

Also left to future work is the exploration of the potential interactions between sonification and visualisation, and of how multimodal system designs can be leveraged for the context. In Prototype 1, for example, we envisage that, while sonification is used here for the *perception* of anomalous events on the network – the recognition by humans that "something is wrong" – visualisation could complement the system by enabling *comprehension* of the nature of the events perceived, directed by the sonification. Similarly, Prototype 2 could be complemented by a visualisation that conveys exactly which destination IP address has experienced an increase in packet rate, following the event that the listening analyst's attention is drawn to some change in the sonification.

Further work should be carried out, as highlighted in Section IV, in user testing of the system, in order to assess whether users (in particular, the intended users: security analysts) can hear the patterns generated in the sonification at the time of the attacks. We intend to research the potential for sonification to match, or improve on, the performance of

existing monitoring systems in the SOC environment such as security visualisations and IDSs. At this stage, usability aspects such as integration of sonification into the SOC environment should also be addressed.

REFERENCES

[1] L. Axon, S. Creese, M. Goldsmith, and J. R. C. Nurse, "Reflecting on the use of sonification for network monitoring," in Proceedings of the International Conference on Emerging Security Information, Systems and Technologies (IARIA), 2016, pp. 254–261.

[2] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Maciá-Fernández, and E. Vázquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges," computers & security, vol. 28, no. 1, 2009, pp. 18–28.

[3] S. Axelsson, "The base-rate fallacy and its implications for the difficulty of intrusion detection," in Proceedings of the 6th ACM Conference on Computer and Communications Security. ACM, 1999, pp. 1–7.

[4] M. Gopinath, "Auralization of intrusion detection system using Jlisten," Development, vol. 22, 2004, p. 3.

[5] Klahr, Rebecca and Amili, Sophie and Shah, Jayesh Navin and Button, Mark and Wang, Victoria, "Cyber Security Breaches Survey 2016," 2016.

[6] D. Denning, "An intrusion-detection model," IEEE Transactions on Software Engineering, no. 2, 1987, pp. 222–232.

[7] A. Lazarevic, L. Ertöz, V. Kumar, A. Ozgur, and J. Srivastava, "A comparative study of anomaly detection schemes in network intrusion detection." in Proceedings of SIAM International Conference on Data Mining, 2003, pp. 25–36.

[8] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," ACM Computing Surveys (CSUR), vol. 41, no. 3, 2009, p. 15.

[9] V. Kumar, J. Srivastava, and A. Lazarevic, Managing cyber threats: issues, approaches, and challenges. Springer Science & Business Media, 2006, vol. 5.

[10] D. E. Denning and P. G. Neumann, "Requirements and model for ides—a real-time intrusion detection expert system," Document A005, SRI International, vol. 333, 1985.

[11] N. Ye, S. Emran, Q. Chen, and S. Vilbert, "Multivariate statistical analysis of audit trails for host-based intrusion detection," IEEE Transactions on Computers, vol. 51, no. 7, 2002, pp. 810–820.

[12] D. Anderson, T. F. Lunt, H. Javitz, A. Tamaru, and A. Valdes, Detecting unusual program behavior using the statistical component of the Next-generation Intrusion Detection Expert System (NIDES). SRI International, Computer Science Laboratory, 1995.

[13] C. Tsai, Y. Hsu, C. Lin, and W. Lin, "Intrusion detection by machine learning: A review," Expert Systems with Applications, vol. 36, no. 10, 2009, pp. 11 994–12 000.

[14] Y. Zhang, Y. Xiao, M. Chen, J. Zhang, and H. Deng, "A survey of security visualization for computer network logs," Security and Communication Networks, vol. 5, no. 4, 2012, pp. 404–421.

[15] R. E. Etoty and R. F. Erbacher, "A survey of visualization tools assessed for anomaly-based intrusion detection analysis," DTIC Document, Tech. Rep., 2014.

[16] R. F. Erbacher, K. L. Walker, and D. A. Frincke, "Intrusion and misuse detection in large-scale systems," Computer Graphics and Applications, IEEE, vol. 22, no. 1, 2002, pp. 38–47.

[17] B. Shneiderman, "Dynamic queries for visual information seeking," IEEE Software, vol. 11, no. 6, 1994, pp. 70–77.

[18] J. Nicholls, D. Peters, A. Slawinski, T. Spoor, S. Vicol, J. Happa, M. Goldsmith, and S. Creese, "Netvis: a visualization tool enabling multiple perspectives of network traffic data," 2013.

[19] S. Rinderle-Ma and T. Hildebrandt, "Server sounds and network noises," in Cognitive Infocommunications (CogInfoCom), 2015 6th IEEE International Conference on. IEEE, 2015, pp. 45–50.

[20] Z. Halim, R. Baig, and S. Bashir, "Sonification: a novel approach towards data mining," in Proceedings of the International Conference on Emerging Technologies, 2006. IEEE, 2006, pp. 548–553.

[21] T. Hinterberger and G. Baier, "Parametric orchestral sonification of EEG in real time," IEEE MultiMedia, no. 2, 2005, pp. 70–79.

[22] P. Janata and E. Childs, "Marketbuzz: Sonification of real-time financial data," in Proceedings of the International Conference on Auditory Display, 2004.

[23] T. Hermann, "Sonification for Exploratory Data Analysis," Ph.D. dissertation, 2002, Bielefeld University.

[24] G. Kramer, Auditory display: Sonification, audification, and auditory interfaces. Perseus Publishing, 1993.

[25] A. de Campo, "Toward a data sonification design space map," in Proceedings of the International Conference on Auditory Display, 2007, pp. 342–347.

[26] S. Barrass and C. Frauenberger, "A communal map of design in auditory display," in Proceedings of the International Conference on Auditory Display, 2009, pp. 1–10.

[27] S. Barrass et al., "Auditory information design," Made available in DSpace on 2011-01-04T02: 37: 33Z (GMT), 1997.

[28] M. Gilfix and A. Couch, "Peep (the network auralizer): Monitoring your network with sound." in Proceedings of the Large Installation System Administration Conference, 2000, pp. 109–117.

[29] R. Giot and Y. Courbe, "Intention–interactive network sonification," in Proceedings of the International Conference on Auditory Display. Georgia Institute of Technology, 2012, pp. 235–236.

[30] D. Worrall, "Realtime sonification and visualisation of network metadata," in Proceedings of the International Conference on Auditory Display, 2015, pp. 337–339.

[31] M. Kimoto and H. Ohno, "Design and implementation of stetho—network sonification system," in Proceedings of the International Computer Music Conference, 2002, pp. 273–279.

[32] D. Malandrino, D. Mea, A. Negro, G. Palmieri, and V. Scarano, "Nemos: Network monitoring with sound," in Proceedings of the International Conference on Auditory Display, 2003, pp. 251–254.

[33] M. Ballora, N. Giacobe, and D. Hall, "Songs of cyberspace: an update on sonifications of network traffic to support situational awareness," in SPIE Defense, Security, and Sensing. International Society for Optics and Photonics, 2011, pp. 80 640P–80 640P.

[34] R. Schafer, The soundscape: Our sonic environment and the tuning of the world. Inner Traditions/Bear & Co, 1993.

[35] O. Kessler et al., "[functional description of the data fusion process]," 1991, Office of Naval Technology Naval Air Development Center, Warminster PA.

[36] P. Vickers, C. Laing, and T. Fairfax, "Sonification of a network's self-organized criticality," arXiv preprint arXiv:1407.4705, 2014.

[37] P. Vickers, C. Laing, M. Debashi, and T. Fairfax, "Sonification aesthetics and listening for network situational awareness," in Proceedings of the Conference on Sonification of Health and Environmental Data, 2014.

[38] B. deButts, "Network access log visualization & sonification," Master's thesis, Tufts University, Medford, MA, US, 2014.

[39] M. Garcia-Ruiz, M. Vargas Martin, B. Kapralos, J. Tashiro, and R. Acosta-Diaz, "Best practices for applying sonification to support teaching and learning of network intrusion detection," in Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications, 2010, pp. 752–757.

[40] S. El Seoud, M. Garcia-Ruiz, A. Edwards, R. Aquino-Santos, and M. Martin, "Auditory display as a tool for teaching network intrusion detection," International Journal of Emerging Technologies in Learning (iJET), vol. 3, no. 2, 2008, pp. 59–62.

[41] P. Varner and J. Knight, "Monitoring and visualization of emergent behavior in large scale intrusion tolerant distributed systems," Technical report, Pennsylvania State University, 2002.

[42] C. Papadopoulos, C. Kyriakakis, A. Sawchuk, and X. He, "Cyberseer: 3d audio-visual immersion for network security and management," in Proceedings of the ACM workshop on Visualization and data mining for computer security. ACM, 2004, pp. 90–98.

[43] L. Qi, M. Martin, B. Kapralos, M. Green, and M. García-Ruiz, "Toward sound-assisted intrusion detection systems," in On the Move to Meaningful Internet Systems 2007: CoopIS, DOA, ODBASE, GADA, and IS. Springer, 2007, pp. 1634–1645.

[44] A. Brown, M. Martin, B. Kapralos, M. Green, and M. Garcia-Ruiz, "Poster: Towards music-assisted intrusion detection," 2009, poster presented at IEEE Workshop on Statistical Signal Processing.

[45] V. F. Mancuso et al., "Augmenting cyber defender performance and workload through sonified displays," Procedia Manufacturing, vol. 3, 2015, pp. 5214–5221.

[46] M. García-Ruiz, M. Martin, and M. Green, "Towards a multimodal human-computer interface to analyze intrusion detection in computer networks," in Proceedings of the First Human-Computer Interaction Workshop (MexIHC), Puebla, Mexico, 2006.

[47] "Fraunhofer IIS Netson," 2016, URL: http://www.iis.fraunhofer.de/en/muv/2015/netson.html [accessed: 24/02/2017].

[48] "Specimen Box, The Office for Creative Research," 2014, URL: http://ocr.nyc/user-focused-tools/2014/06/01/specimen-box/ [accessed: 24/02/2017].

[49] L. Buchanan, A. D'Amico, and D. Kirkpatrick, "Mixed method approach to identify analytic questions to be visualized for military cyber incident handlers," in Visualization for Cyber Security (VizSec), 2016 IEEE Symposium on. IEEE, 2016, pp. 1–8.

[50] T.-F. Yen, A. Oprea, K. Onarlioglu, T. Leetham, W. Robertson, A. Juels, and E. Kirda, "Beehive: Large-scale log analysis for detecting suspicious activity in enterprise networks," in Proceedings of the 29th Annual Computer Security Applications Conference. ACM, 2013, pp. 199–208.

[51] D. Zhao, I. Traore, B. Sayed, W. Lu, S. Saad, A. Ghorbani, and D. Garant, "Botnet detection based on traffic behavior analysis and flow intervals," Computers & Security, vol. 39, 2013, pp. 2–16.

[52] D. Acarali, M. Rajarajan, N. Komninos, and I. Herwono, "Survey of approaches and features for the identification of http-based botnet traffic," Journal of Network and Computer Applications, vol. 76, 2016, pp. 1–15.

[53] A. D'Amico, K. Whitley, D. Tesone, B. O'Brien, and E. Roth, "Achieving cyber defense situational awareness: A cognitive task analysis of information assurance analysts," in Proceedings of the human factors and ergonomics society annual meeting, vol. 49, no. 3. SAGE Publications, 2005, pp. 229–233.

[54] G. Parseihian, C. Gondre, M. Aramaki, S. Ystad, and R. K. Martinet, "Comparison and evaluation of sonification strategies for guidance tasks," IEEE Transactions on Multimedia, vol. 18, no. 4, 2016, pp. 674–686.

[55] S. C. Peres, D. Verona, T. Nisar, and P. Ritchey, "Towards a systematic approach to real-time sonification design for surface electromyography," Displays, 2016.

[56] T. Hermann, A. Hunt, and J. Neuhoff, The sonification handbook. Logos Verlag Berlin, GE, 2011.

[57] G. Dubus and R. Bresin, "A systematic review of mapping strategies for the sonification of physical quantities," PloS one, vol. 8, no. 12, 2013, p. e82491.

[58] E. Yeung, "Pattern recognition by audio representation of multivariate analytical data," Analytical Chemistry, vol. 52, no. 7, 1980, pp. 1120–1123.

[59] M. Ballora, R. Cole, H. Kruesi, H. Greene, G. Monahan, and D. Hall, "Use of sonification in the detection of anomalous events," in SPIE Defense, Security, and Sensing. International Society for Optics and Photonics, 2012, pp. 84 070S–84 070S.

[60] J. Rubin and D. Chisnell, Handbook of usability testing: how to plan, design and conduct effective tests. John Wiley & Sons, 2008.

[61] J. R. C. Nurse, S. Creese, M. Goldsmith, and K. Lamberts, "Guidelines for usable cybersecurity: Past and present," in Proceedings of the Third International Workshop on Cyberspace Safety and Security (CSS). IEEE, 2011, pp. 21–26.

[62] A. de Campo, "A data sonification design space map," in Proc. of the 2nd International Workshop on Interactive Sonification, York, UK, 2007.

[63] F. Briolle, "Detection and classification of the audiophonic sonar signal: perspectives of space simulation under headphones," Undersea Defense Technology, 1991.

[64] R. Mill and G. Brown, "Auditory-based time-frequency representations and feature extraction techniques for sonar processing," Speech and Hearing Research Group, Sheffield, England, 2005.

[65] M. Ballora and D. Hall, "Do you see what I hear: experiments in multichannel sound and 3D visualization for network monitoring?" in SPIE Defense, Security, and Sensing. International Society for Optics and Photonics, 2010, pp. 77 090J–77 090J.

[66] A. D'Amico and K. Whitley, "The real work of computer network defense analysts," in VizSEC 2007. Springer, 2008, pp. 19–37.

[67] T. Hildebrandt, T. Hermann, and S. Rinderle-Ma, "Continuous sonification enhances adequacy of interactions in peripheral process monitoring," International Journal of Human-Computer Studies, 2016.

[68] "The CAIDA UCSD DDoS Attack 2007 dataset," URL: https://www.caida.org/data/passive/ddos-20070804_dataset.xml [accessed 24/02/2017].