

EXPLORING SONIC PARAMETER MAPPING FOR NETWORK DATA STRUCTURES

Brian Hansen, Leya Breanna Baltaxe-Admony, Sri Kurniawan, and Angus G. Forbes

Computational Media Department
University of California, Santa Cruz
Santa Cruz, California, USA

{brmhans, bbaltaxe, skurnia, angus}@ucsc.edu

ABSTRACT

In this paper, we explore how sonic features can be used to represent network data structures that define relationships between elements. Representations of networks are pervasive in contemporary life (social networks, route planning, etc), and network analysis is an increasingly important aspect of data science (data mining, biological modeling, deep learning, etc). We present our initial findings on the ability of users to understand, decipher, and recreate sound representations to support primary network tasks, such as counting the number of elements in a network, identifying connections between nodes, determining the relative weight of connections between nodes, and recognizing which category an element belongs to. The results of an initial exploratory study (n=6) indicate that users are able to conceptualize mappings between sounds and visual network features, but that when asked to produce a visual representation of sounds users tend to generate outputs that closely resemble familiar musical notation. A more in-depth pilot study (n=26) more specifically examined which sonic parameters (melody, harmony, timbre, rhythm, dynamics) map most effectively to network features (node count, node classification, connectivity, edge weight). Our results indicate that users can conceptualize relationships between sound features and network features, and can create or use mappings between the aural and visual domains.

1. INTRODUCTION

A network data structure is an arrangement of data into interconnected groupings of information (nodes) according to relationships between groupings (edges). Network data structures are an integral component of our daily lives. Social networks facilitate personal and professional communication. The internet, a network of linked documents, is a ubiquitous utility used in nearly every facet of contemporary life. Transportation networks, such as subway maps, are used by millions of people who rely on these networks for their daily commute. Similarly, in data science, the “hairball”—a densely connected network—has become the dominant icon for the information age, describing the need for analysts to invent new methods to untangle the complex relationships between data points [1].

Sonification has the potential to play a key role in illuminating relationships present in network data structures. In social

networks, a sonic queue could signify a group of friends or social connections with common interests and be assigned a unique *earcon* [2]. With respect to website hierarchies, there is often a lack of topological orientation present when an individual visits a web page. A sonification could give the user a sense of place within the website topology, facilitating more accurate and relevant navigational decisions. Sonification has already been used to help passengers interpret navigational systems. Japanese composer Minoru Mukaiya has composed over 100 unique jingles for different train stations throughout Tokyo which are played each time a train leaves the station. Each jingle acts as an earcon, conveying a range of information. For instance, a crescendo and rising pitch in the Shibuya station departure song represents the train’s uphill journey to the next platform. The melodies themselves are strongly mnemonic and reinforce the passenger’s awareness of their location. Moreover, the jingle for each station along a route can be concatenated to form a coherent song, providing confirmation to a passenger where they are headed and at what point of the journey they are in [3].

Despite existing examples of and speculative uses for network sonification, there is a lack of research on the ability of users to create useful mappings between network elements and sonic parameters. In this paper, we present an initial investigation into how network data structures could be effectively sonified. Our contributions include: *a*) a characterization of the challenges unique to network sonification (Sec. 3); *b*) a formulation of initial hypotheses that incorporate these challenges (Sec. 4); *c*) the results of two qualitative pilot studies (Secs. 4.1 and 4.2) that assess user interpretation of sonic parameters mapped to network features; and *d*), a delineation of themes extracted from user responses that identify representational elements that many users expect, and that point to potentially useful avenues for future study (Sec. 4.2.3). Our results indicate that users can create an effective mental model of a sonified network structure. Furthermore, we find that users can make meaningful associations between network features and sonic parameters.

2. BACKGROUND AND RELATED WORK

Network data structures are commonly represented by node-link diagrams consisting of nodes and edges, where the nodes are positioned either to facilitate readability, or based on a particular grouping criteria, and the edges depict a connective relationship between the nodes [4]. This approach was popularized by Jacob Moreno as early as 1932, in which he formalized graphical characteristics to represent actors and relationships in social networks [5]. Visual characteristics used in Moreno’s drawings include arrows



This work is licensed under Creative Commons Attribution Non Commercial 4.0 International License. The full terms of the License are available at <http://creativecommons.org/licenses/by-nc/4.0>

to indicate the direction of nodal connections, colors for multiple layers of nodes and nodal connections, shapes of nodes to communicate characteristics of social actors, and variations in the nodal location to emphasize structural features of the data.

Network graphs are often multivariate in nature, and mapping parameters to particular sounds has the potential to enhance the interpretation and analysis of network features. Prior work in sonification has shown success in this regard, as the properties of audio allow the presentation of multiple dimensions without information overload for users. This is demonstrated in one of the earliest studies of sonification by Pollack and Ficks [6], who evaluate mappings of multidimensional data onto sound. They measure information transmitted to subjects as the sum of the number of bits in each correctly identified dimensional level and find that multidimensional audio displays outperform uni-dimensional displays. Yeung [7] presents a parameter mapping between seven sonic parameters and seven chemical variables, utilizing two pitches, loudness, damping, direction, duration, and rest (silence) to represent the structure of a chemical. Upon hearing the sonification, classification occurs with a 90% success rate before training, and increases to 98% after training.

3. CHALLENGES

Given that sound has shown to be successful in representing multivariate data, our goal is to determine sonic parameter mappings appropriate for network data structures. The relational aspects among nodes and edges pose a unique challenge for our sonification. Specifically, this includes representing nodal connectivity, location, and orientation. In addition, it can be advantageous to have an overall impression of the network architecture, allowing for the acquisition and recall of the structural topology.

The size of the network structure is also of important consideration. A very large network with numerous nodes and a highly complex edge topology may yield a sonification that is too acoustically saturated for interpretation. In this case, the sound may be best utilized to convey the gist of the overall data structure [8]. Alternatively, if examining a network structure that is either very small or at a highly localized level, a sonification may saliently communicate the low level details present in the data.

3.1. Connectivity

In representing network connectivity, fundamentally we need to determine the best sonic representation for a connection between two nodes. This could be accomplished numerous ways including via musical texture, articulation, or the production of sound effects. For example, if two nodes are represented as sequential tones separated by time, then the melodic transition between them could represent the presence or absence of a connection. A legato or glissando articulation between them would indicate a nodal connection, while the presence of silence a detachment. Alternatively, two connected tones could sound simultaneously producing a harmonic texture. Further, a sound effect could be introduced between the tones signifying their connection. For example, a synthesized Doppler effect could indicate not only that tones are connected but also give the impression of the connective direction.

The problem of representing nodal connectivity becomes increasingly difficult when considering the addition of numerous nodes and the combinatorial possibilities for their connections. Connections may be unidirectional or bidirectional. Simultaneous

connections may exist among the nodes as one-to-many or many-to-one mappings. Groups of nodes can be chained together generating a higher order of connectivity. The ensemble of possibilities has the potential for yielding a highly saturated sonic display that could be very difficult to interpret. Thus, a sonification of network data structures must carefully consider how the multiple types of connections are displayed.

3.2. Location and Orientation

Nodal location and orientation plays an important role in depicting nodal relationships and an overall network architecture. This poses a unique challenge for data sonification because nodes are commonly represented in a geospatial context, with their location and relative orientation displayed in two or three dimensional space, and exclusive of a relationship to time. These facets pose a challenge for our sonification because the perception of sonic events depends primarily on the temporal domain, where the majority of sonifications involve the presentation of data as a sequence of sonic events over time.

Although the vast majority of sonifications involve time based representations, there has been successful exploration using sound to communicate spatial data. For example, Flowers, Buhman, and Turnage used the dimensions of frequency and time to display 2D scatter plots of data [9]. In addition, Alty and Rigas devised the tool AudioGraph that paired notes to represent geospatial data points. In their representation, timbre indicated a particular axis and frequency the distance along that axis [10, 11].

Further, in the area of cartography, Schito and Farikant utilize parameter mapping sonification to represent digital elevation models, where the sonic parameter of pitch was shown to be the most successful in accurately interpreting sonic displays. Krygier proposes a set of nine “sound variables”— including sound source location, loudness, pitch, register, timbre, duration, rate of change, order, and envelope— that could be used to represent spatial data [12]. Krygier paralleled his approach to the semi-otic system for graphics previously established by the cartographer Bertin [13, 14].

Questions remain as to how effective such a geospatial sonification can be. In comparison between sonic and visual mappings, visual representations of spatial data are much more accurate. In a visual map, data points can be precisely plotted on an x, y coordinate plane and their location can be clearly comprehended. Compared to source localization of sound, a listener’s notion of the sound object’s position is much less accurate [15].

As a possible solution, the temporal characteristics of sound could be mapped to geospatial metrics. For example duration, measured in a unit of time such as beats per minute, could represent the distance to a sonically positioned object. In addition, meter, which is often conceived as a one-dimensional grid, can be used to quantify distance. Rhythm could also be used to quantify distance and could be expanded to multiple dimensions via the notion of polyrhythms, where this could help represent distances in a multi-dimensional space.

Although a network data structure is often times presented spatially, the spatial relationships presented may not exist in reality. For example, social network visualizations may show nodes of individuals and their interconnectivity organized into groupings on a two dimensional space. In reality, the orientation of the individuals (that is, the physical location of the individuals in the real world) has no relationship to nodal orientation displayed on

a graph. The importance of the spatial representation is unique to data visualization, in that it acts as a means to communicate data groupings. When it comes to sound, groupings can be represented in a similar fashion, perhaps by assigning a unique timbre or register to the data points.

Regardless of the approach, it is important for us to distinguish which features of our sonification require geospatial precision and which do not. If certain features can not be precisely represented via sound, then the question may be to what extent can the information be conveyed? For example, given the geospatial placement of a sound source, we may not be able to determine its precise distance from us, but we can at least perceive that it is either close or far. Additionally, given the placement of multiple sources, we may at a minimum be able to determine which source is relatively closest or furthest. If we must accept that geospatial relationships cannot be perceived accurately, then we may be required to de-emphasize this feature in the sonification.

3.3. Topological Impression (Acquisition and Recall)

It is important to perceive the overall impression of a network topology. The minutiae of relationships within network structures are distinct and complex. Taken in aggregate, the impression of the overall structure is an identifying principle that signals the general relationships among its more detailed components [16]. As such, a higher level impression of the structure enhances our ability for knowledge acquisition and recall, allowing us to more efficiently identify, categorize, and compare network data sets.

Musical structures, such as a melodic phrase, are similar in that a lower level of multivariate sonic information is encapsulated in the higher level structure. When a person hums a tune, they are referencing an abundance of parametric data with an elaborately organized collection of pitches, durations, onsets, rhythms, articulative effects, and dynamics. All of these features are efficiently encoded within the musical structure which can be commonly recalled to a high degree of accuracy.

We look to take advantage of this facet via a successful parameter mapping, where upon transforming a network data set into a higher level musical structure, the network features will be more efficiently acquired and recalled. To accomplish this, we looked to utilize the musical genre of a *jingle*. Most commonly employed in advertising, jingles exhibit strong mnemonic qualities, facilitating learning and recall [17]. Yalch presented experiments with jingles where it was shown that they are highly effective in low-exposure advertising (low frequency of slogan occurrence), thus demonstrating the strong recall qualities of the genre [18]. Jingles are “catchy,” consisting of simple musical phrases that are easily sung and can be recalled with a high degree of accuracy. The compositional makeup of a jingle supports recall through its frequent use of the pentatonic scale (the most universally utilized scale) and the 4/4 time signature (the most common musical meter), its brevity (usually no more than two bars in length with a highly limited set of melodic notes), and by being registrally within a nominal singable range.

4. USER STUDIES

In this section, we describe two pilot studies that each explore parameter mapping between jingles and small networks. To address the challenges in determining which musical parameters are best

sued to represent elements of network data structures, we consider the following questions related to nodes, edges, and recall:

Nodal content

- How many nodes are present in the representation?
- How does the representation convey nodal position?
- How does the representation convey qualitative aspects of the nodes?

Edge content

- How many edges are present in the representation?
- How does the representation convey connectivity between edges and nodes?
- How does the representation convey qualitative aspects of the edges?

Mnemonic strength of the representation

- How easily and accurately can knowledge of the representation be obtained and recalled?

Based on our survey of the literature and our own experience conducting research in information visualization and data sonification, we formulate the following hypotheses:

Hypothesis 1 (H1): Subjects can conceptualize a general representation of a network graph upon hearing a musical example.

Justification: A mental construct is formed when a person is presented with a sonic stimulus. Musical structures, being sonic stimuli, consist of sonic elements that are grouped into specific relationships. As network data structures exhibit similar features, they can be conceptualized by a musical representation.

Hypothesis 2 (H2): Subjects can conceive specific and meaningful correspondence between musical and graphical features of a network structure with only minimal exposure to the musical excerpt.

Justification: Musical phrases, in particular jingles, have demonstrated a strong mnemonic quality allowing a person to acquire and recall structural details to a high degree of accuracy. With the musical structure of a jingle internalized, subjects can recall and continuously reference the structure, allowing them to formulate meaningful correspondence with a network graph. Due to the mnemonic quality of a jingle, this can be accomplished with minimal exposure to the musical source.

Hypothesis 3 (H3): When presented with a mapping between a musical example and network graph, subjects can identify correspondence between musical and graphical features.

Justification: Sonification has a demonstrable history of success with respect to parameter mapping. In such cases, subjects are able to associate musical and graphical parameters.

Hypothesis 4 (H4): Subjects have a preconceived notion of how musical parameters correspond to features of a network graph.

Justification: Musical features are described in a qualitative and quantitative fashion. This implies musical perception consists of associative structural descriptors. As such descriptors exist prior

to knowledge of a parameter mapping to network data structures, subjects will have a preconceived notion of how musical features ought to correspond to features of a network graph.

Hypothesis 5 (H5): Subjects can consistently justify their association of musical examples with network graphs when mapping musical parameters to network structural features.

Justification: Effective parameter mapping sonifications have demonstrated a high degree of accuracy and consistency when sonic parameters intuitively map to particular data elements. The presence of highly salient parameter mappings generates consistency, allowing mappings to be preserved over numerous domains.

To validate these hypotheses, we conducted two pilot studies to obtain qualitative feedback on the relationship between musical and network structure representations. For each study, we composed a set of jingles containing musical elements with the potential to correspond to features of a small, highly localized network structure. Each jingle emphasized a specific musical parameter, such as pitch, melody, harmony, rhythm, timbre, or dynamics. For example, one jingle highlighted dynamics, containing notes that are either quiet or loud. Other jingles presented harmonic vs. monophonic content, while yet others contained obvious drastic changes in timbre among melodic tones. As we describe below, each pilot study included a range of tasks in which subjects were presented with jingles and asked to make correspondences between the musical elements and network features. Their feedback gives us insight into sonic parameter mapping and provides initial empirical evidence about which feature correspondences are meaningful.

4.1. First Pilot Study

Our first pilot study aimed to discover if individuals could conceptualize networks as sound and to explore an individual's mental model of a sonified network. We were also curious to see if individuals could intuitively understand the mapping between pre-composed jingle-network pairs. The study consisted of three tasks and was conducted as follows:

4.1.1. Participants

This pilot study was conducted with six individuals - five males and one female. Four were graduate students in music, one was a graduate student in computer science, and one was a professor of music. All participants had a strong background in both music and technology, but had varying experience working with abstract network representations. An introduction to networks was presented in order to ensure that everyone was familiar with fundamental network concepts.

4.1.2. Methodology

To begin, participants were shown examples of common network visual representations (an organization chart, a bus map, and an abstract node-link network diagram), preparing them to participate in three tasks. A detailed explanation of network features, including core concepts such as connectivity, edge weight, and node value, was given between Task 1 and Task 2 so that the immediate inclination of participants in Task 1 would not be influenced. A post-study survey was conducted to obtain participants' musical

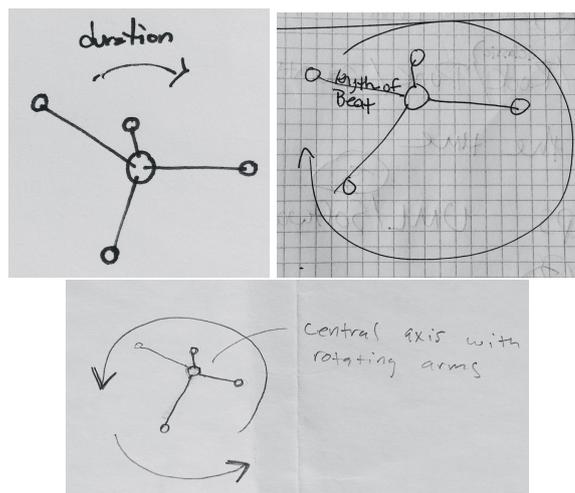


Figure 1: This figure shows examples from selected participants' feature mapping for a graph and jingle pair from Pilot Study 1, T3.

experience and familiarity with networks.

Task 1 (T1) - Drawing a Network:

The first task contained two subtasks (Subtask 1 and Subtask 2). In each, we played a jingle for our subjects. The jingle presented for this task exhibited common characteristics of the genre (see Sec. Topological Impressions): it was limited to 2 measures of 4/4 time at a tempo of 120 beats per minute, totaling 4 seconds in duration. The texture of the jingle was purely monophonic, comprised of an easily singable melody confined to a nominal register, with a total number of notes limited to 9 or less. This jingle placed particular emphasis on dynamic contrast among the notes.

T1-Subtask 1 (S1) - Memory: Upon hearing the jingle, subjects were then instructed to draw a picture of a network structure that resembled what they heard and justify their reasoning. In this task, we aimed to test knowledge acquisition and memory, and thus asked them to listen to the jingle only once before drawing a graph.

T1-Subtask 2 (S2) - Detail: In the second task, they were instructed to listen to the jingle from Task 1 as many times as they wished and revise their original drawing as desired. This task was designed to see if individuals could conceptualize and generate a meaningful mapping of sonic parameters to visual features of a network.

Task 2 (T2) - Drawing Network Features:

The second task was focused on the number of nodes present and their connectivity. After an explanation of core network concepts, we asked the subjects to draw another network graph based on a new jingle they had not yet heard. The jingle played in this task was similar to that employed in T1, but placed particular emphasis on timbral contrast. We emphasized that they should consider the number of nodes and their connectivity in the drawing and in their justification. This task aimed to uncover sonic inclinations for particular network features.

Task 3 (T3) - Identifying Correspondences between Network Features and Sonic Parameters:

In the third task, we presented a precomposed pair— a graphical representation of a simple network structure and a corresponding jingle which we composed to “match” it. The outer nodes of the graph, shown in Figure 1 were sonically represented by each pitch in a clockwise manner with the arc length corresponding to the duration of each note. This was done to establish the position and orientation of each node in the graph. In this representation, the central node acted as a connectivity hub for the other nodes and was sonically treated as a tonal center (tonic). Nodes that appeared above this hub sounded higher in register, while nodes below sounded lower in register. Nodal distance from the hub was portrayed by pitch, where a higher or lower pitch relative to the tonic, indicated distance extremity. Without revealing this chosen mapping, we asked the subjects to label the network concepts on the provided graph with what they felt was represented by musical concepts from the audio. This task aimed to discover whether individuals could interpret a preconceived mapping without guidance, and uncover any points of friction in our mapping.

4.1.3. Results

In this study we observed individuals create their own graphs corresponding to provided music passages and interpret existing music/graph pairs. This shows us that these individuals were capable of conceptualizing a mapping between sound and network representations, confirming H1.

In T1, 5 out of 6 participants were able to generate some graph after hearing the jingle once. 3 participants were able to provide noticeably more detail to their graphs in T1-S2 with many listens than in T1-S1 with few. Some participants listened all the way through the jingle multiple times before drawing, while others paused throughout playback to draw. Each participant did seem to value the allowance to go back through and listen again, and willingly took time to play the jingle until they felt comfortable with their interpretation. This refutes H2.

In T1 and T2 we found that participants had a tendency to draw networks that closely resemble musical notation. This was unexpected, but reveals the power of a known structure on an individual’s mental model. As shown in the example in Fig. 2, almost all participants map pitches as nodes occurring over time with a height corresponding to pitch. One participant had a hard time breaking from this musical structure at all. This may be a result of the subjects’ musical background, but music representations such as sheet music and MIDI scrollers are commonplace, and this inclination might be present in anyone who has seen them. Responses to T2 indicate that subjects were thinking about specific parameter mappings. However, the act of drawing graphs for two graph concepts was not an effective way of quantifying the parameter mappings.

For T3, 5 of 6 individuals intuitively mapped both the pitch duration to arc length. 5 of 6 mapped note order to circular node position, but only 3 of 6 correctly identified the mapping to be clockwise (see Fig. 1). However, no participant was able to grasp how pitch corresponded to cardinal node position. Thus H3 is partially supported. This is a promising finding, but indicates that more work is required to identify a set of meaningful mappings. Another interesting finding from T3 is individual’s curiosity about the mappings. After the study was completed, participants engaged in a discussion of what they thought the mappings were, and con-

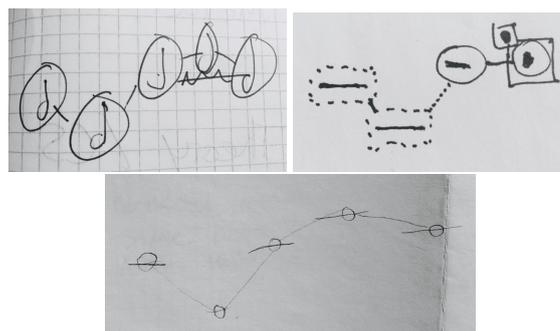


Figure 2: This figure shows selected examples of the graphs which participants drew to correspond with the jingle played in Pilot Study 1, T2. These selected graphs closely parallel music notation, where notes are “plotted” pitch vs. time.

tinued inquiry with the experimenter even after the study ended. Some feedback we received was that one participant could have created drawings and mapped features better if they had been told what the mappings were to begin with. From these interactions it seems likely that once mappings have been identified, participants would be able to interpret the sonic data better when the mappings are provided.

Overall, we concluded that indeed correspondences can be made between musical parameters and network features. However, at the conclusion of this first pilot study, it was not clear to us which parameter mappings were the most effective. This led us to develop a second pilot study that aimed to determine more precisely what constitutes a salient parameter mapping (H4 and H5).

4.2. Second Pilot Study

The second pilot study aimed to find a specific mapping of musical features to network features. As in Pilot Study 1, we again asked users to listen and respond to jingles. Pilot Study 2 consisted of three tasks, followed by a general survey. The study was conducted as follows:

4.2.1. Participants

The study was carried out with a convenience sample of 26 undergraduate participants with a range of musical and computational experience. All participants were students in a course offered within our university’s engineering department. This group included 7 females, 17 males and 2 non-binary individuals. 22 of 26 participants have had past experience playing an instrument or singing, with 9 of the 22 consider themselves to be currently practicing musicians. The university majors of this group were (including two students with double majors): Computer Science (9), Art & Design: Games and Playable Media (9), Cognitive Science (6), Computational Media (2), Technology and Information Management (1), and Theater (1).

4.2.2. Methodology

To begin, a presentation on network and music foundations was given to help solidify participant’s understanding of each by

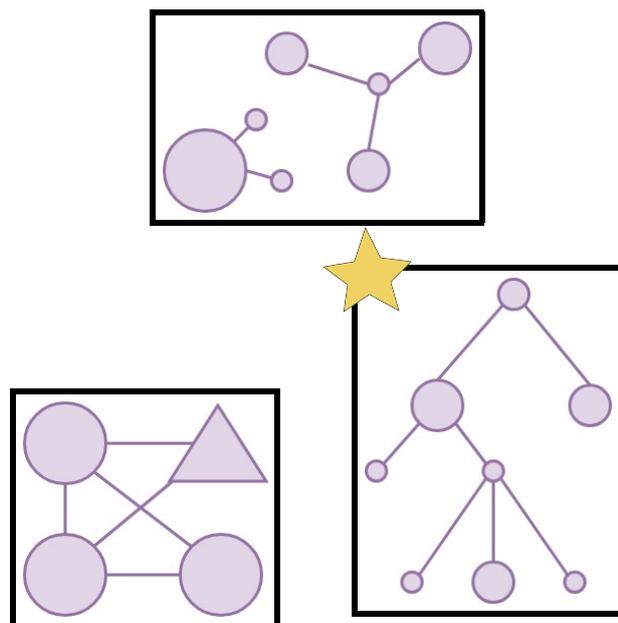


Figure 3: This figure shows examples of the networks used for Pilot Study 2, T2. The network which was most frequently selected to match the jingle played in T2 is starred.

playing different variations of sounds on a speaker in a classroom setting. Key terms for each space were defined during this presentation. Following the introductory material, participants were asked to participate in two tasks.

T1 - Mapping Sonic Parameters to Network Features:

The first task presented the subjects with a range of different network features, including: the number of nodes, how the nodes are connected, strength of the connectivity, and nodal shape. We then asked the subjects to choose which musical feature best represented given network features and instructed them to justify their answer. They were limited to select from the musical features of pitch, melody, harmony, rhythm, timbre, and volume and were confined to choose only one musical feature to represent each correspondence in the feature mapping.

T2 and T3 - Matching a Network Diagram to a Jingle:

In each of tasks two and three, subjects were presented with three unique graphical representations of a network structure. They were then played a jingle and asked to select which of the three graphs most closely matched the musical excerpt they heard. They were also asked to justify their choice. For this study, each jingle was conceived of and composed independently of any concept of a visual representation. This was done to avoid introducing bias in parameter mapping. The jingles were composed similarly to those of Pilot Study 1. However, the jingle in T2 placed particular emphasis on harmonic dyads integrated into the melody, and the jingle in T3 placed particular emphasis on a timbral effect placed on specific melodic notes.

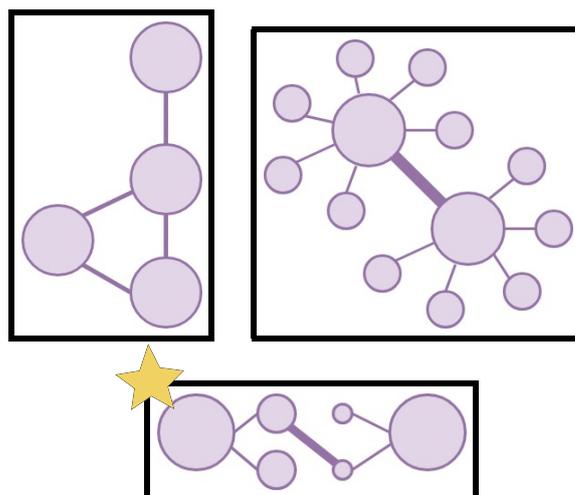


Figure 4: This figure shows examples of the networks used for Pilot Study 2, T3. The network which was most frequently selected to match the jingle played in T3 is starred.

4.2.3. Results

The responses of the 26 individuals were recorded and analyzed using grounded theory [19], where three researchers individually identified emerging themes from the collected data. We report the number of occurrences of each unique response for all short answer prompts which shows general trends for parameter mapping. No further statistical analysis has been conducted due to the small sample size.

We found quantitatively from T1 that subjects have the strongest preconceived notion of mapping for the features connectivity, weight, and shape, which supports H4. However, users were not as effective at mapping the number of nodes. Thus H4 is partially supported. The top scoring mappings for each of the four features is as follows: 21 of 26 participants chose timbre to represent node shape, 15 of 26 participants chose volume to represent edge weight, 13 of 26 participants chose melody to represent node connectivity, and 9 out of 26 participants chose rhythm to represent the number of nodes present. (The remaining values can be seen in Table 1.)

We found from T2 and T3 that although participants were decisive about what musical features should represent graph concepts, those choices did not necessarily remain consistent. We expected individuals to strongly hold to their preconceived notions as exhibited in T1. However, there were many inconsistencies between individuals conceptual choice in T1 and their actions when relating graphs to music in T2 and T3. Further, even when our subjects articulated a reasonable rationale for mapping a parameter to a network feature, those mappings would change when presented with different sounds. Also, in some cases multiple musical parameters were cited as justification for a singular graph, and in other cases a singular musical parameter was cited as justification for multiple graphs. An example of this variability can be seen under the “Note Count” theme, where Participant 18 chooses two possible mappings for note count. Thus, H5 was not fully supported, as each user’s parameter mapping may not be consistent across varying contexts. This indicates a possible level of adapt-

	# of Nodes	Connectivity	Weight	Shape
Duration	4	1	4	0
Rhythm	9	3	0	0
Pitch	6	1	0	4
Harmony	2	7	3	0
Melody	2	13	1	1
Volume	2	0	15	0
Timbre	0	1	1	21

Table 1: This table shows the number of votes for particular musical feature mappings recorded for T1 of Pilot Study 2, including: the number of nodes, whether or not nodes were connected to one another, the edge weight of these connections, and the shape of the nodes.

ability of a user’s mental model when relating graphs to audio data. Although our original expectation was not met, the implication that users have the ability to flexibly intuit new mappings appropriate for particular contexts can serve as an advantage, as it allows for a broader range of parameter mapping choices when designing a network sonification.

We were also able to identify several themes in participants justifications which are independent of their graph selection in T2 and T3. Those themes are linearity, character, and note count. Selected quotes and analysis of each theme are detailed below.

Linearity:

A desire for network sonifications to have a notion of “start” and “finish” that aligns to the time dimension of the audio was the strongest theme. This trend was noticeable in Pilot Study 1, where many of the network drawings resembled MIDI musical notation—a well known representation for displaying pitches over time. In addition, this concept was present in participants’ answers, regardless of their selection in T2 and T3. This theme also emerged as the most selected node-link diagram for both tasks, as seen in Figs. 3 and 4. In each case the most frequently selected diagram was the most linear of the group, portraying no cycles and exhibiting an obvious location for the “start” and “finish” for each group. Quotes from subjects related to this theme include:

“Notes ... lasted the longest at the start and the end. The picture ... when viewed left to right, has larger circles at the start and the end.”

- Participant 12 referencing Fig. 4, bottom.

“[When viewed] from the top down the number of nodes will match what is playing [over time].”

- Participant 1 referencing Fig. 3, bottom right.

Character:

Most individuals chose to focus on specific features of the given networks and audio. However, some individuals chose to focus on the overall “feeling” or “character” of the music. Participants created a narrative for the network diagram or related it to known objects. Aside from choosing a specific feature mapping, it may be important to consider how the sonification supports the overall character of the network. Quotes from subjects related to this theme include:

“[The graph] look[s] like those instruments that have beads attached to them that bounce off the drum ... this [matches] the plentiful, trailing, low volume notes. ...”

- Participant 18 referencing Fig. 4, top right.

“[The] graph is [like a] father and son playing catch. The music gives off that vibe too: peaceful, calm, happy”

- Participant 21 referencing Fig. 4, bottom.

Note Count:

Participants exhibited a tendency to try to match the number of notes in a jingle to a graph feature. The number of notes seems to be a salient feature that individuals want to find specific meaning for. Note count was not present as a musical feature in the mapping step and needs to be explored further as a potentially powerful feature in network sonification. Two examples of note counting follow:

“... There were 5 notes/chords played which might correspond to number of connections”

- Participant 20 referencing the jingle in Pilot Study 2, T3.

“There were 5 notes played and this graph has the closest number of nodes. There are also 5 links in the graph. ... one link could [be represented by] one note”

- Participant 18 referencing the jingle in Pilot Study 2, T3.

“The beginning sounds like a harmony [of 2 notes] which is held and then changes to one note, converging again to 2 notes. ...”

- Participant 11 referencing the jingle in Pilot Study 2, T2.

5. CONCLUSIONS

Our initial results show that individuals are able to conceptualize a mapping between audio and network representations (H1). Due to our convenience sample, all findings from these studies cannot be said to be universal—our sample size was small and all participants were affiliated with the university. However, we believe that these initial findings can be validated in future studies of network sonification, as well as to inspire the development of network sonification tools.

From our first pilot study, we found that individuals’ mental models were influenced by their previous exposure to various music notation (plotted pitch vs. time). We did not find that jingles were memorable enough for individuals to fully grasp a network from only one listening (H2). In this study, we also found that individuals could effectively reason about the mappings between a precomposed network and jingle pair, but that they were not able to confidently identify all feature mappings (H3).

In our second pilot study, we found that individuals could justify feature mappings of unrelated networks (H3), and that they did have some preconceived notions of what network and musical features should correspond to one another (H4). These notions were however somewhat flexible as shown by inconsistency between some of their mappings (H5). More specifically, Pilot Study 2 showed that node connectivity, edge weight, and node shape correspond most popularly to melody, volume, and timbre respectively. Three important themes with respect to participants’ mapping justifications were also identified: linearity, character, and note count.

Both studies show that users can create associations between musical and network structures. The results showed attempts made by subjects to formulate specific relationships between the musical and graphical details. For example, in T3 of Pilot Study 2, subjects who selected the second graph primarily did so because they associated the sound effect present with the graphical image. Comments described the sound as “rippling” or “twinkling,” and provided a precise description of how this quality represented the nodal structure in the graph. Although this is a promising indication, more work needs to be done with a larger sample size and more rigorous analysis to conclusively identify parametric relationships at a more detailed level. We plan to devise and test a more robust set of parameter mappings and confirm their validity in a future study.

Although a sonification can serve as an aid to the visual representation of a network data structure, ideally the sonic representation would stand on its own. If fundamental aspects of the network could be represented in the audio domain, this could free the visual domain to encode other data elements, or could lead to network displays that are more readily accessible by visually impaired populations. Results from Pilot Study 2 appear to support such segregation, as the preconceived notions that subjects had about a parameter mapping were not reflected by their actual associations selected. Thus, it appears that a person’s mental model of a sonic representation is flexible and may be liberated from a pre-existing visual anchor. In the future, we plan to conduct studies to further explore relationships between the purely conceptual elements of a network structure and sonic parameters.

Materials and data from the conducted pilot studies are attainable via the project repository located at <https://github.com/CreativeCodingLab/SonifyingNetworksData/>.

6. REFERENCES

- [1] A. D. Lander, “The edges of understanding,” *BMC biology*, vol. 8, no. 1, p. 40, 2010.
- [2] D. Oswald, “Non-speech audio-semiotics: A review and revision of auditory icon and earcon theory,” in *Proc. ICAD*, 2012.
- [3] Hermesauto, “The man behind Japan’s train departure melodies,” Apr 2018. [Online]. Available: <https://www.straitstimes.com/asia/east-asia/the-man-behind-japans-train-departure-melodies>
- [4] T. von Landesberger, A. Kuijper, T. Schreck, J. Kohlhammer, J. J. van Wijk, J.-D. Fekete, and D. W. Fellner, “Visual analysis of large graphs,” *Computer Graphics Forum*, vol. 30, no. 6, pp. 1719–1749, 2011.
- [5] L. C. Freeman, “Social network visualization, methods of,” *Computational Complexity: Theory, Techniques, and Applications*, pp. 2981–2998, 2012.
- [6] I. Pollack and L. Ficks, “Information of elementary multidimensional auditory displays,” *The Journal of the Acoustical Society of America*, vol. 26, no. 2, pp. 155–158, 1954.
- [7] E. S. Yeung, “Pattern recognition by audio representation of multivariate analytical data,” *Analytical Chemistry*, vol. 52, no. 7, pp. 1120–1123, 1980.
- [8] T. Hermann, A. Hunt, and J. G. Neuhoff, *The sonification handbook*.
- [9] J. H. Flowers, D. C. Buhman, and K. D. Turnage, “Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples,” *Human Factors*, vol. 39, no. 3, pp. 341–351, 1997.
- [10] J. L. Alty and D. I. Rigas, “Communicating graphical information to blind users using music: the role of context,” 1998.
- [11] M. Britnell, “Seeking a reference frame for cartographic sonification,” in *Proc. ICAD*, 2018.
- [12] J. B. Krygier, “Sound and geographic visualization,” in *Modern cartography series*. Elsevier, 1994, vol. 2, pp. 149–166.
- [13] J. Bertin, *Semiology of graphics: diagrams, networks, maps*. University of Wisconsin Press, 1983.
- [14] J. Schito and S. I. Fabrikant, “Exploring maps by sounds: using parameter mapping sonification to make digital elevation models audible,” *International Journal of Geographical Information Science*, vol. 32, no. 5, pp. 874–906, 2018.
- [15] T. Nasir and J. C. Roberts, “Sonification of spatial data,” in *Proc. ICAD*, 2007.
- [16] B. Shneiderman, “The eyes have it: A task by data type taxonomy for information visualizations,” in *The Craft of Information Visualization*. Elsevier, 2003, pp. 364–371.
- [17] M. Alexomanolaki, C. Loveday, and C. Kennett, “Music and memory in advertising: Music as a device of implicit learning and recall,” *Music, Sound, and the Moving Image*, vol. 1, no. 1, pp. 51–71, 2007.
- [18] R. F. Yalch, “Memory in a jingle jungle: Music as a mnemonic device in communicating advertising slogans,” *Journal of Applied Psychology*, vol. 76, no. 2, p. 268, 1991.
- [19] A. Strauss and J. Corbin, “Grounded theory methodology,” *Handbook of qualitative research*, vol. 17, pp. 273–85, 1994.