Mobile Hearing Diagnostics

by

Benjamin Yu

Supervisor: Parham Aarabi

April 2013

Abstract

The purpose of this research was to design a robust mobile play-based diagnostic application for iOS. The app consisted of four different game modes that utilized pure-tone audiometry and threshold equalizing noise to determine the user's minimum audible threshold. The threshold is then compared to 11 different audiogram shapes using mean square error to determine the pathological cause for the hearing deterioration.

Table of Contents

Acknowledgements

Foremost, I would like to thank my thesis supervisor Professor Aarabi for his continued support and mentorship. Your expertise and guidance was invaluable, and I couldn't have completed my thesis without it. I would also like to thank my parents for their years of unconditional love and support and my brother for showing me the courage of the human spirit.

List of Symbols, Tables & Figures

Introduction

The proliferation of mobile devices and commoditization of sensor technologies through the past decade has created a new healthcare movement of self-quantification. Biometric information, such as heart rate, blood sugar, and sleeping patterns can now be analyzed at time scales that were thought to be unattainable. This technology could potentially usher in a new paradigm shift in health care delivery and patient-doctor relationships.

One area that has not been thoroughly explored with mobile diagnostics is audiology and the ability to quantify one's hearing ability. Hearing loss is the most prevalent congenital sensory deficit in the world [1]. In 2004, over 275 million people globally had moderate-to-profound hearing impairment [1]. Nearly 15% of school-aged children had hearing deficits at low and high frequencies. The problem is further compounded where impoverished countries do not have audiology or otolaryngology services available [2]. A mobile audiometric diagnostic solution could potentially tap into an extremely large patient population, along with becoming a new diagnostic tool for doctors.

The gold standard of audiometric diagnostics is pure tone audiometry (PTA) [3]. This is a procedure where an audiometrist plays pure tones of varying frequency levels and volume increments, until a response is elicited from the patient [3]. This establishes a minimum threshold of hearing, creating a base measure of the patient's hearing ability. The fidelity of this procedure is limited primarily due to the small number of trials that can be performed within a single testing session. In one study, an average bilateral test took 7.7 minutes to establish an audiogram of only 30 recordings [2]. PTA is also extremely prone to patient and psychometric bias. Differential diagnosis using audiogram shapes is also clinically subjective, which is more prone to misdiagnosis. An accessible procedure with higher data fidelity could drastically improve this procedure's diagnostic ability.

Proposed Research

Figure 1 - Sones, an iOS hearing diagnostic game

My research will be exploring this gap through a play-based iOS mobile application (app), Sones, that will evaluate a user's hearing ability. Sones will feature a collection of games that will test a user's hearing in an entertaining & informal setting. Gamification of audiometric testing will provide more extensive and robust diagnostic data than traditional methods. By testing on a mobile platform, we also gain the ability to diagnose patients who don't have access to basic audiology services. Users will also be restricted to non-standard insert headphones, allowing us to explore the viability of automated testing within a non-ideal sound environment. My research will also explore the viability of mobile diagnostics, and the possibility of altering current health care paradigms through continuous data acquisition and personal quantification.

Sones will consist of a progression of games of increasing complexity that will train and stress the user's hearing abilities. The classification will be done algorithmically through simple mean square error comparisons. My research will be part of a larger research project, with one team developing an Android implementation of a similar application, and another hearing test application on iOS that will emulate traditional clinical audiometric tests. Having Sones run on both Android and iOS environments will help us reach a larger target audience, and will help qualify each platform's ability to support remote audiometric testing. The results from the separate clinical test emulator will allow us to measure the viability of these new techniques.

Literature Review

Pure tone audiometry (PTA) is used worldwide as the primary mode of hearing disorder assessment. PTA is any test that tries to establish the minimum threshold of hearing, through the use of pure tones played within a controlled anechoic environment. Typically pure sine tones of varying frequencies are played via calibrated headphones. The patient then indicates at what volume he/she can hear the sound. Various strategies are used, but common procedures involve playing tones starting at about 125 Hz and increasing by octaves, half-octaves, or third-octaves to about 8000 Hz. Hearing tests of right and left ears are generally done independently. The results of such tests are summarized in audiograms. Audiograms are then compared to an established normal threshold of hearing, and the physician can make a diagnosis.

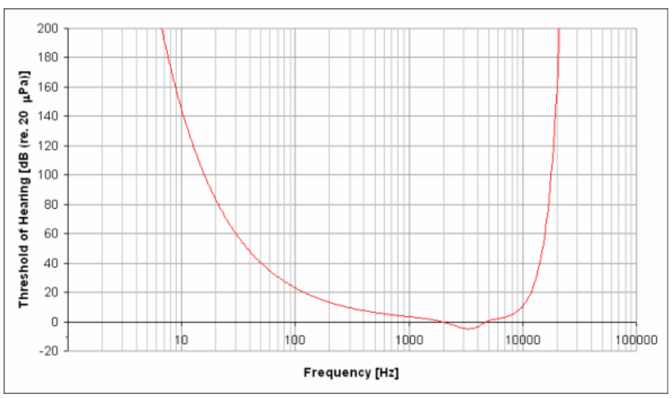

Figure 2 - Minimum audible threshold as established by Terhardt [4]

Human threshold for hearing was established empirically by Terhardt to be close to [4]:

$$T(f) = 3.64(f/1000)^{-0.8} - 6.5e^{-0.6(f/1000-3.3)^2} + 10^{-3}(f/1000)^4$$

Human hearing is the most sensitive around 3.3kHz and sensitivity at frequencies above 10kHz rapidly decreases [5]. Deviations from the ideal model should indicate some form of hearing loss.

Classical Methods of Minimum Threshold Acquisition

In pure-tone audiometry, there are a variety of testing methodologies for determining the minimum audible threshold. Sones will employ many of the same techniques in accessing hearing ability. The most common method is the method of limits, where patients are presented with stimuli that are uniformly distributed over a range of volumes and frequencies. They are then presented to the patient in a random order, and the patient indicates whether the sound was heard. A more intelligent variation is using adaptive methods, where the first sequence of tones starts at a level that the listener did not hear. We then predict the threshold on an online basis so we can reduce the volume range we need to search.

Another common method used in practice is where the patient is given control of the stimulus level, and is instructed to adjust it until they cannot hear the tone. An example is Bekesy tracking where the listener pushes a button as long as he can hear the tone and lets go when he stops hearing the tone, so the level goes up and down around threshold. In Bekesy tracking the frequency of the tone changes during the course of the test so that thresholds can be estimated at many frequencies [6].

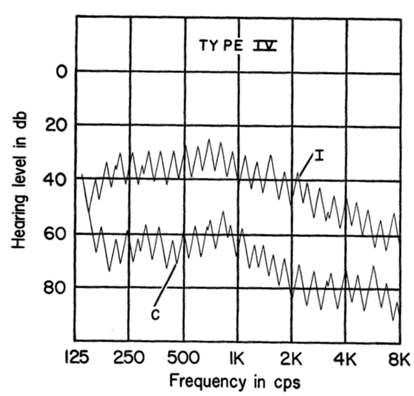

Figure 3 - Audiogram of Bekesy Tracking. Midpoint of peaks can establish threshold [6]

Two-alternative Forced Choice

A common issue with audiometric testing is user response bias, where the user guesses whether the sound is present, even if they don't hear anything. One way to circumvent this is to present two alternatives signals for the user to choose from. On interval one, what the listener thinks happened is affected by both his sensation and his response bias. The same is true in the second interval. When the subject is asked to choose which interval contains the signal, the bias between both choices cancel each other out, effectively removing response bias from the test.

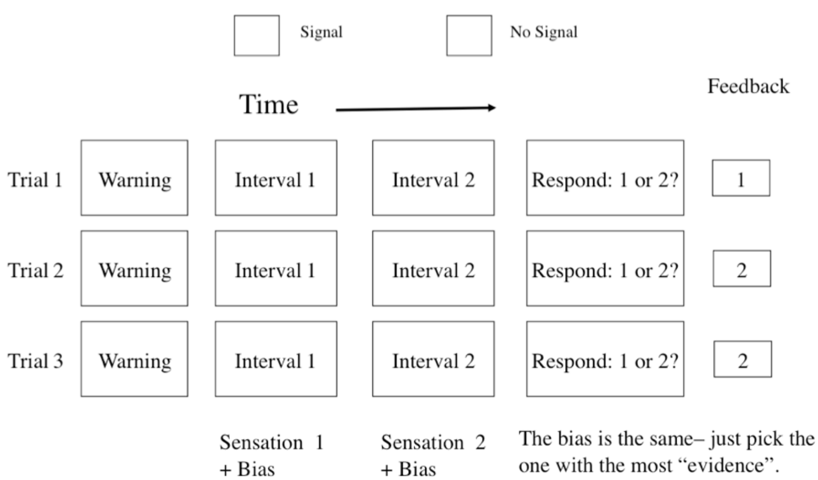

Figure 4 - Forced Choice Procedure that eliminates choice bias

Neural Tuning Curves

Dead regions, areas where outer hair cells are damaged or dead, cannot be clearly found via PTA audiograms [7]. Typically, the neurons innervating the dead region cannot react to vibration at their characteristic frequency [7]. If the basilar membrane vibration is large enough, neurons tuned to different characteristic frequencies such as those adjacent to the dead region, will be stimulated due to the spread of excitation, and a response from the patient at the test frequency will be obtained [7].

This is referred to as off-place listening, and is also known as off- frequency listening [7]. This will lead to a false threshold being found. Thus, it appears a person has better hearing than they actually do, resulting in a dead region being missed. Therefore, using PTA alone, it is impossible to identify the extent of a dead region [7]. Since Sones will essentially employ PTA for its hearing analysis, it is important that we take the limitations of PTA into account.

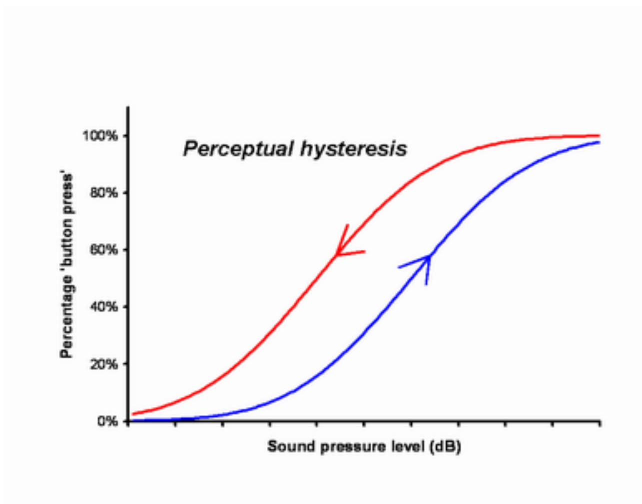

Figure 5 – Perception Hysteresis Effect

Another major issue encountered by previous methodologies is the variability of the patient's perception of the stimuli. It can be shown that the absolute threshold of hearing is actually a time-dependent feature. This theory is based on speculation that stimulus is somewhat connected to memory, and is accumulated until some level of audibility. Plomp and Bouman suggest that the detection system approximates a low pass 'leaky' integrator. This psychophysical phenomenon is conceptualized as a hysteresis effect. Sones will also have to take these psychophysical effects into account.

Psychophysical Tuning Curves

Psychophysical Tuning Curves are a relatively new method for determining minimum audible thresholds that attempts to account for psychophysical affects [7]. To measure a

PTC, the sinusoidal signal is fixed in frequency and presented at a fixed (usually low) sensation level (about 10 dB SL) [7]. A narrowband noise is usually used as the masker. For each of several masker centre frequencies, the level of the masker required just to mask the signal is determined [7]. The masker attempts to mask the off-frequency excitations to OHC's that are adjacent to the target frequency range [7].

Threshold Equalizing Noise

A new quicker way that was explored by researches is using a threshold- equalising noise (TEN) [7-8]. TEN is a broadband noise that is spectrally shaped so that the masked threshold is approximately the same for all frequencies in the range of 250 Hz to 10 kHz, for people with normal hearing. It is essentially follow the same concept at PTC's where the masking noise will cancel out the off-frequency effect from dead regions, making it possible to identify such dead ranges. The advantage of this technique is that the masking signal covers the entire frequency spectrum, reducing the computation needed to create a unique mask for each test frequency. This appears to be a technique that can be easily implemented in Sones.

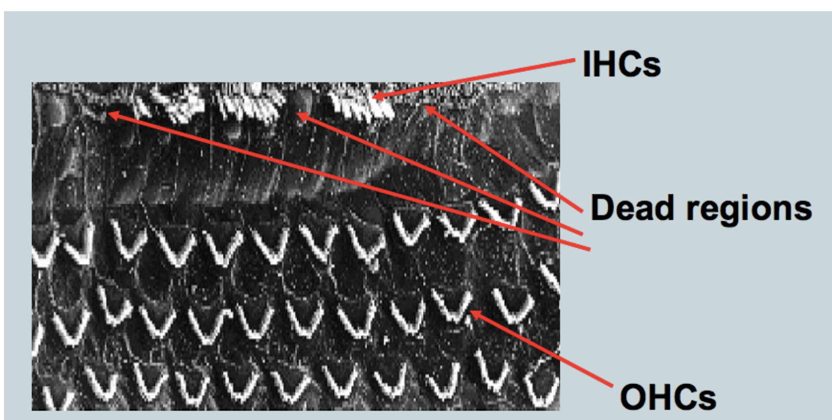

Figure 6 - Regions where OHC's are damaged [7]

Auditory Filter Models

To better understand the underlying causes for hearing disorders, and to better treat patients, it would be extremely beneficial to be able to map diagnostic data through a system model of the human auditory system. Furthermore, if the performance of individual listeners could be predicted through auditory modeling, this would be particularly useful to help design the best compensation strategy for individuals.

One modeling strategy is to use information from the auditory nerve fibers (e.g., average rate, synchrony, and nonlinear phase information) [9]. They typically predict performance that is 1 to 2 orders of magnitude better than human performance [9]. Other types of auditory models are inspired by neurophysiological findings and make simplifying assumptions about each auditory processing stage [9]. One common method is to use a temporal window model where volume accumulates over time [9].

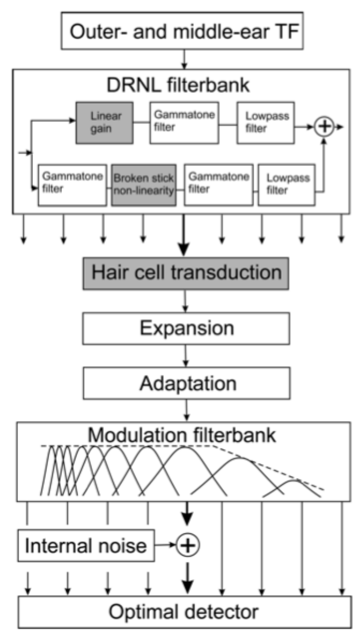

Figure 7 - CASP Model for hearing

Jepsen et al. modeled the auditory pathway using the CASP model, which is shown in Figure 3 [9]. The processing stages comprise outer and middle-ear filters, nonlinear BM processing, IHC transduction, expansion, adaptation, and a modulation filterbank [9]. Finally, the model includes an optimal detector designed to deal with n-interval alternative forced choice paradigms [9]. To simulate consequences of sensorineural hearing loss, the stages associated with hair cell loss were modified. The changes were thus included in the DRNL filterbank and the IHC transduction stage [9]. Sones could employ a similar classification model to assess hearing ability.

Automation and Differential Diagnosis

There have been several studies over the past decade exploring the possibility of applying machine learning techniques to categorize audiogram shapes, and better define the pathological underpinnings of such patterns.

K-means Clustering has been used in several studies as a basic classification technique with quite remarkable results. Lee et al. identified 11 basic shapes that highly correlated with existing clinical diagnosis practices, using K-means over audiogram data over 1633 patient data sets [10]. Ciletti & Flamme studied patient data over 5000 patient records categorized by gender and living location versus urban and rural locations [11]. They established that males generally have 20 different audiometric configurations, while women were limited to 8 [11]. Tyler et al. also established 4 sub-groups in relation to degrees of tinnitus, using audiogram and subjective reports from 528 patients [12].

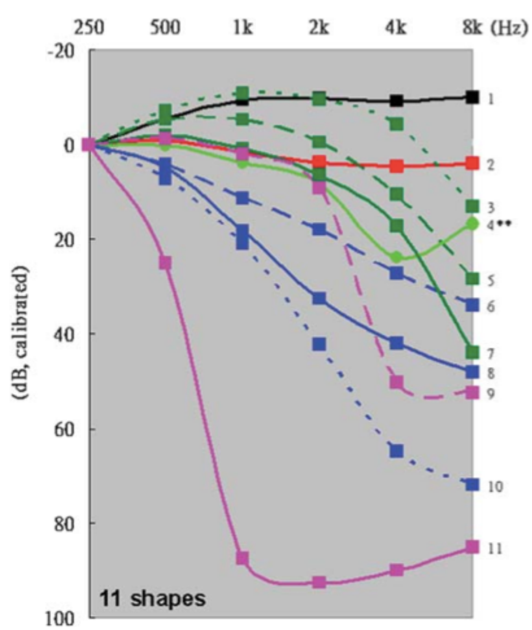

Figure 8 – Audiogram shapes from K-means Clustering [10]

More complex supervised techniques have been used to attempt to train diagnostic classification models. Muhammad et al. conducted a study over 180,000 audiogram datasets to identify contributing factors for using inner ear hearing aids versus behind the ear hearing aids [13]. Using a combination of PCA, Chi sq. testing and logistic regression analysis, they were able to classify hearing aid usage with success ranging from 0.79-0.87 [13]. The resulting model created a decision tree that could aid or supplement an audiologist's diagnosis. Nouraei et al. created an artificial neural network as a screening classifier for patients with vestibular schwannoma, a slow growing auditory nerve tumor [14]. Ozdamar et al. in 1990, derived a Bayesian estimation model known as CAST which generated a probability table of 11 audiogram shapes, which would help audiologists determine which class the patient would fall into during the diagnosis procedure [15]. Again, Sones could employ a similar classification model to assess hearing ability. These methods appear to be more viable due to their simplicity in implementation, and their ability to be modified in an online setting.

Visual Reinforcement and Play-based Audiometry

Visual reinforcement and play-based techniques are commonly employed in clinical settings to diagnose younger children who lack the attention span or mental facilities for common audiometric tests [16]. By the developmental age of six months, a child can make a head turn towards a sound source (localization) and may also be suitable for operant conditioning [16]. Operant conditioning is a technique in which the child is made aware that a specific response (localization of an auditory stimulus) would result in a visual reward (visual reinforcement). Play-based audiometry follows a similar vein, but instead of a visual reward, game elements serve as the reward mechanism.

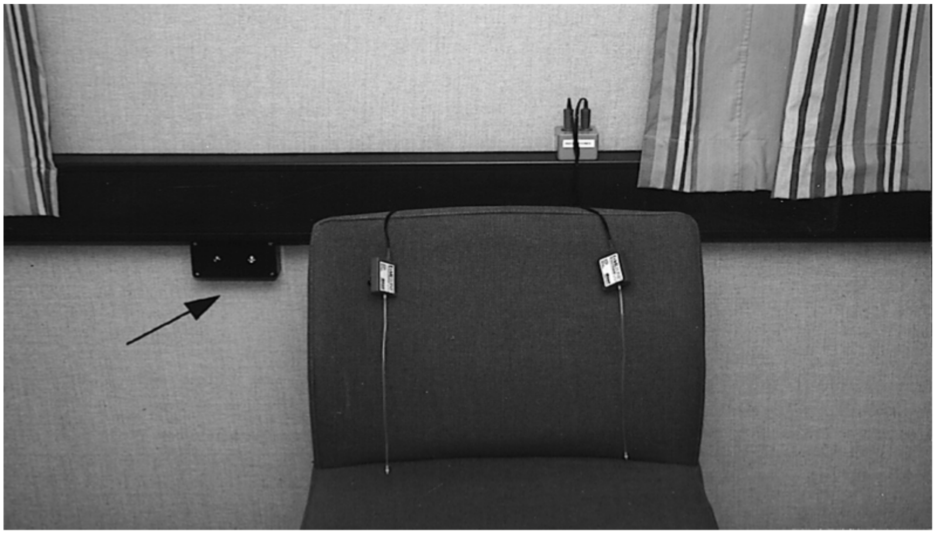

Figure 9 - Clinical visual reinforcement audiometry setup. Lights on either side of chair elicits a response from the patient

Baldwin et al. performed a cross-validation study on visual reinforcement, OAE's and tympanometry measurements [17]. Only 5-10% deviation was found in results [17]. Current play-based and visual reinforcement techniques employ technologies such as Flash-based games, physical toys, DVDs etc... [18]. It is important to assess the viability of these techniques, since the game elements within Sones will essentially be employing similar techniques.

Design Overview

Sones ultimate goal was to replicate the same procedures followed in pure-tone audiometric tests, while incorporating gamification elements that would make the procedure more interactive and appealing to the end user. In order to design such a product, we defined a set of overarching requirements to guide our design process and methodology.

Requirements

First and foremost, Sones had to be able to diagnose hearing problems using games, without compromising the core aspects of the diagnostic procedures or the game elements themselves. The game still had to be fun, while still maintaining the integrity of the individual diagnostic elements. This posed a challenge as it could take many design iterations where we could combine the two together, but still have a fun and engaging gaming experience.

The environments Sones would be used in would vary significantly. For most scenarios, Sones would be used in loud noisy environments which would impair the diagnostic fidelity of the app. Our design would have to account for the variability in the environment to produce reliable results.

The headphones and earbuds used with Sones also posed a serious issue. Due to differences in their builds, internal impedance, and the driving capabilities of the devices, the headphones would have large differences in their dynamic range during the diagnostic procedures. In order to account for this, Sones would require some calibration mechanism to evaluate the volume characteristics of each headset and to make sure our measurements are precise.

The hardware available on each device also posed restrictions on the user interface and the gamification elements we could introduce. Screen size and graphic capabilities varied in each device, so the design would have to ensure that the same user experience would be achieved, regardless of the hardware platform.

In order to reach a larger target audience, App Store approval and publication was a priority. This would require adherence to Apple Guidelines, such as restrictions in usable APIs, software practices, and human interface standards.

A secondary priority I had was to maintain code portability, to minimize the overhead for porting Sones to other platforms. It is also an essential part of good software development, as it ensures a modular design and improves development and debugging efficiency.

Design Conceptualization

Figure 10 - Sones Main Menu

Sones consisted of a collection of games that accessed the user's hearing ability through a variety of methodologies. By having several modes of assessment, we would be able to provide a more interactive and varied experience, as compared to a single clinical PTA procedure. This also allows us to isolate specific elements and techniques, so that the effectiveness of each procedure could be assessed independently within the context of this study. This could potentially also allow us to collect more data from the user, and provide a much more robust framework of audiometric assessment since we have an ensemble of results for each user, rather than just one.

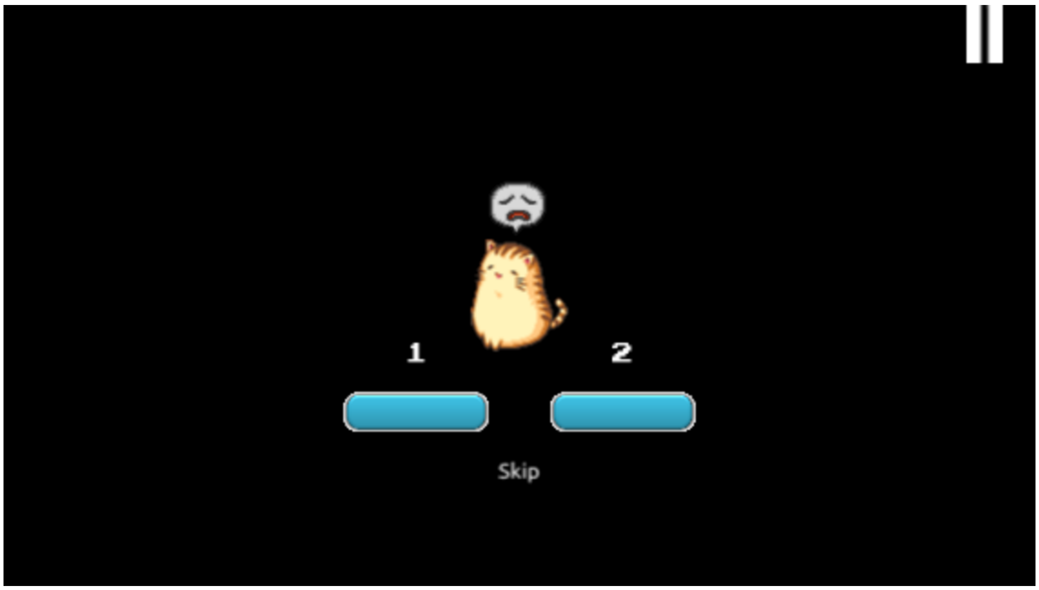

Figure 11 - Counting Game

The first game we designed was *Counting Game*, where the player listened to one or two pure tones in succession. The player would then indicate how many tones he/she had actually heard by pressing the corresponding button, as shown in Figure 11. The volume of the second tone would progressively change towards finding the minimum threshold. This volume threshold would be found through a binary search based on the player's input. Playing single tones acted as catch trials to account for player bias, and perception hysteresis, since the player would be unaware if the second tone was just extremely quiet, or not played at all.

Counting Game was designed such that the game complexity would be extremely simple. This removed extraneous factors such as reaction time, or memory from the diagnostic procedure. It was our hope that this would improve the diagnostic fidelity of Sones, and would more closely replicate traditional diagnostic procedures.

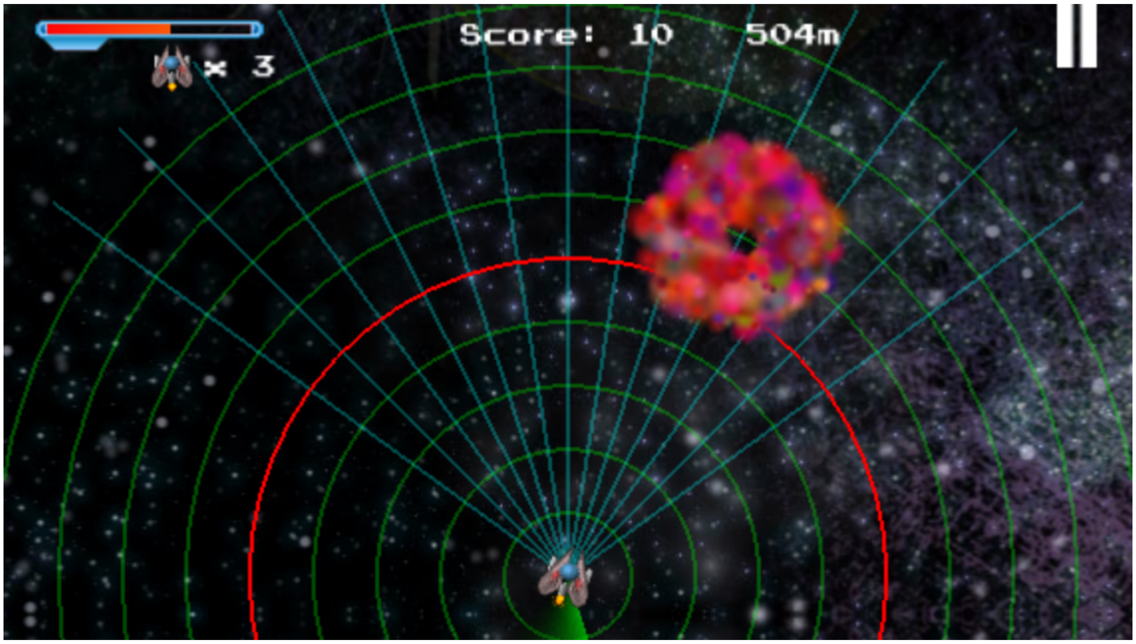

Figure 12 - Target Practice

In order to introduce more complexity and interesting gameplay, we developed another game mode, *Target Practice*. In this game, the user tries to shoot down invisible enemies that are attacking your spaceship. Pure tones of varying frequencies are played on the left and right channels to indicate that an enemy is approaching. As the enemy gets closer, the volume of the pure-tone also increases. This game mechanics allows to user to detect the location of the enemy on screen by identifying the particular frequency and volume of the pure-tone, and associating that with its distance and radial direction. As the game progresses, the enemies remain invisible for longer durations of time. The user will also have to attempt to differentiate the tones more accurately, was the number of attempts/sec will also be reduce as he/she progresses through the levels. The final goal is to progress through as many levels as possible and try to attain the highest possible score.

The philosophy behind *Target Practice* was to implement the simplest gameplay elements possible that could measure the player's hearing ability, while still posing some sort of skill challenge to the player. We were already unsure on the viability of adding gaming elements to audiometric testing, so following Occam's Razor and implementing the simplest solution seemed ideal. The trade-off is that 'fun' is usually highly correlated with complexity and difficulty. We attempted to implement a gradual difficulty curve that increased the speed of the incoming enemies and reduced the time between your shots. This would keep the player entertained and help mitigate boredom.

Figure 13 - Moving Invaders

During initial internal testing, we realized that there was no new content being introduced to the user with *Target Practice*, and the difficulty didn't seem to scale very well to encourage continued play. One possible solution we explored was to emulate already popular game genres, and add more complexity through additional game elements. In the game mode *Moving Invaders*, we tried to create an arcade bullet hell game. The player can move his ship around using the device's accelerometer while collecting coins and avoid enemy attacks. The hearing aspect of the game involves dodging lasers that come in on the left or right side of the screen. A warning tone is played on the corresponding headphone channel, before the laser is appears. For example, if the laser was going to be shot on the right side, as shown in Figure 13, then a warning tone would be played on your right headphone. The warning tone volumes change based on the player's performance, trying to search for their minimum audible threshold. Similar to the *Counting Game*, we utilized a simple binary search procedure to converge on the volume threshold.

We also implemented a maximum likelihood approach to search for the threshold level by fitting sigmoid functions.

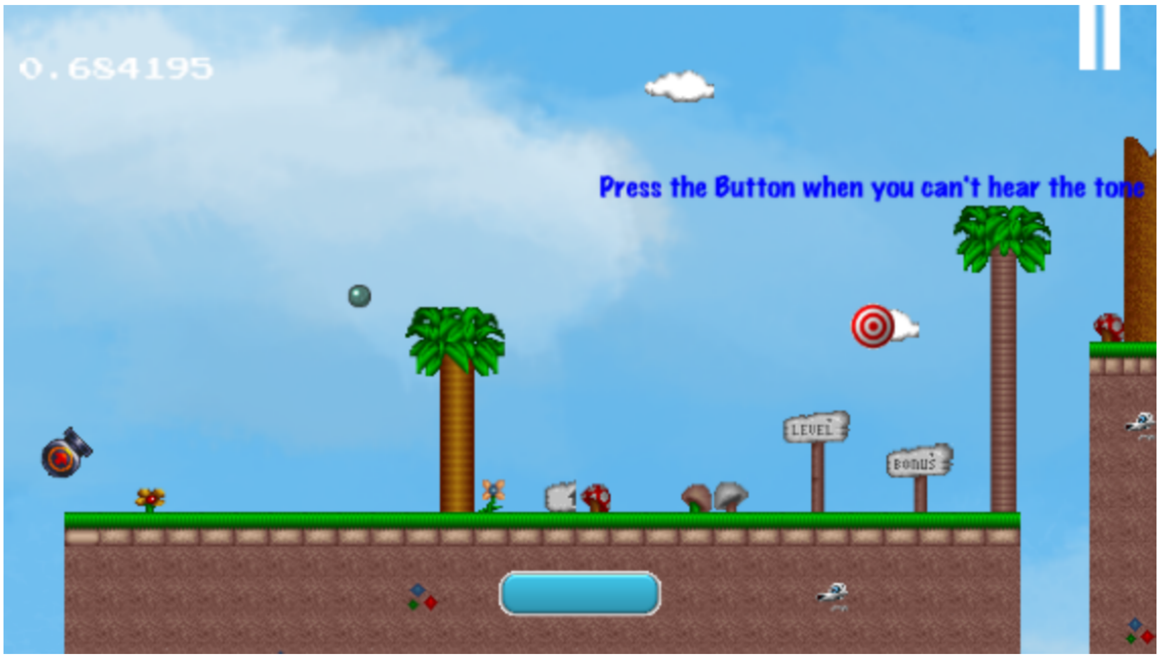

Figure 14 - Cannon Launch

As noted in our Literature Review, pure-tone audiometry is limited in its diagnostic fidelity. All our previous game modes essentially utilized PTA techniques, which meant that their diagnostic power was also somewhat limited. To remedy this, we created third game mode called *Cannon Launch*. The idea behind this game was to emulate threshold equalizing noise (TEN) testing. The game would be similar to *Angry Birds* where the player will control the angle of a projectile, and will try to knock down buildings, and other structures. The power/strength of the projectile will correspond to how closely a masking noise completely covers a pure-tone being played. The masking noise was provided by a 10 second sample of white noise. It was the hope that its flat power spectral density would mitigate frequency-shifting effects as discussed in the Literature Review. Theoretically, this should enable us to measure dead zones more effectively than using

PTA techniques.

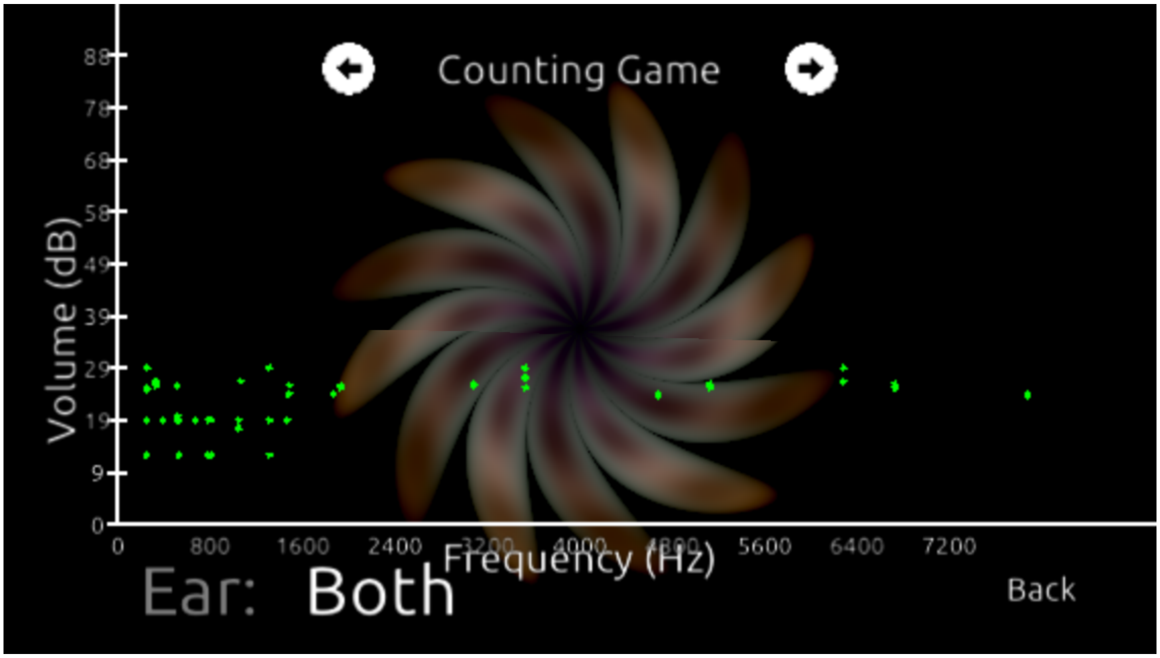

Figure 15 - Audiogram Result View

All collected results can be viewed in Sones in the form of audiograms. Sones would fit the data to audiogram shapes based on their mean square error, and then based on the audiogram shape, determine a diagnosis of your hearing condition. The actual diagnosis is not presented currently due to ethical considerations.

In order to collect data from the user to properly conduct a user study, player performance and other metrics were stored locally on the iOS Device. The database file can be retrieved through iTunes, as shown in Figure 16. The database can then be opened by any relational database view that supports SQLite. This will allow easy access to all the collected information

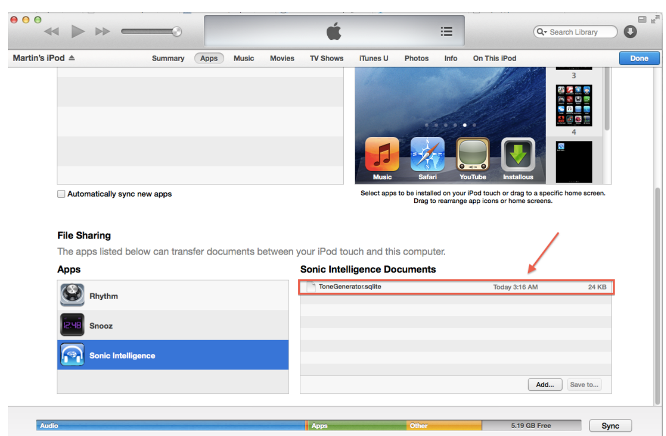

Figure 16 - SQL Database accessed on iTunes

when conducting user studies in a one-

Remote data acquisition was also possible with Sones. We currently have the functionality to send all collected user metrics remotely to a website (http://pacific-springs-8731.herokuapp.com/). All user identification is anonymous, with only a Universally Unique identifier (UUID) to differentiate between each patient. Sones communicates with the website through HTTP requests, which send POST requests with the user data packed as JSON objects. The website then provides a web interface, as shown in Figure 11, which allows us to view all collected audiogram data. UofT Ethics Board has not approved remote acquisition of patient data due to security concerns, so it has been disabled in the current release version.

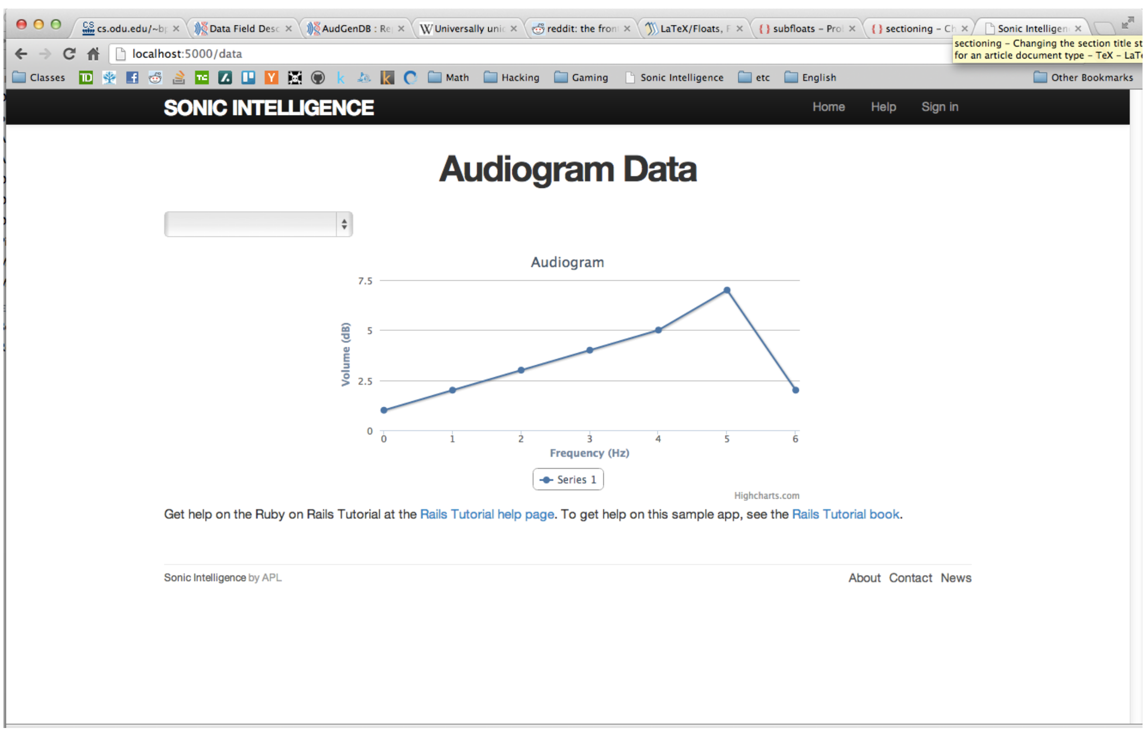

Figure 17 - Sones Website

Technical Implementation

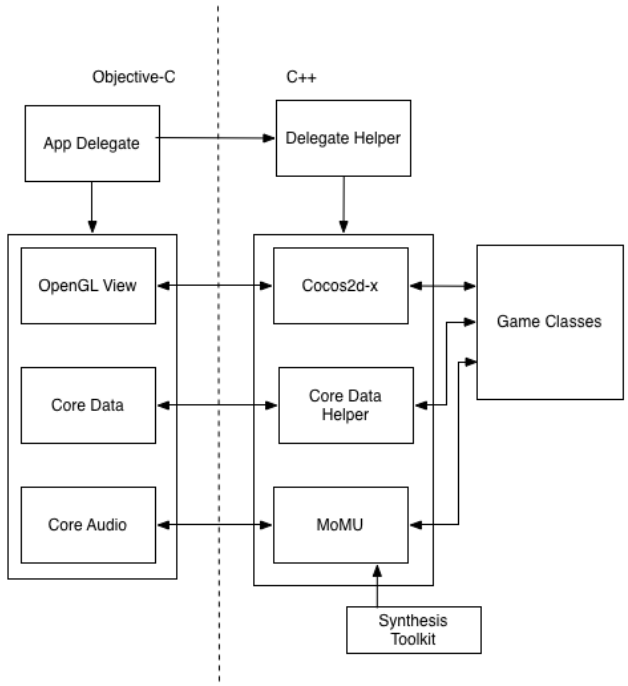

Figure 18 - System Level Overview of Sones

The figure above shows a high level system overview of Sones and it's technical implementation. The primary components are Cocos2d-x, an open source 2D game engine, and MoMu, the Mobile Music Toolkit, an audio synthesis engine. Both are implemented in C++, making them platform independent. This was essential to ensure that Sones could be easily ported to other systems such as Android or Windows 8.

Cocos2d-x is the port of Cocos2D, a popular 2D game engine for iOS. It supports OpenGL ES and has essential features for game development such as scene management, texture atlas support, and integration with Box2D and Chipmunk physics engines. It was also developed with multi-platform in mind, so porting between Android and iOS was possible. It also natively supported multi-resolution modes so developing for different device resolutions was trivial.

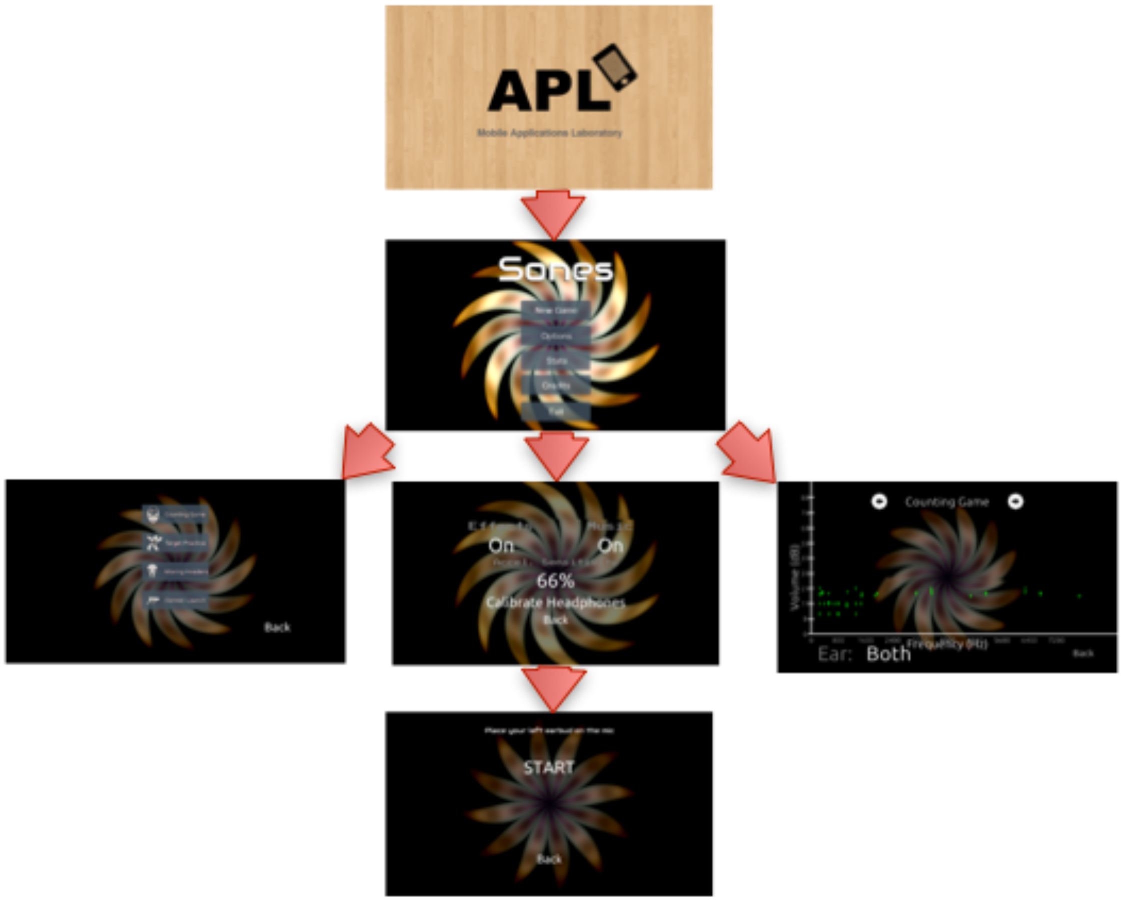

Figure 19 - Cocos2d-x Scene Flow

Cocos2d-x organizes the game with a scene graph, as shown in Figure 11. A director object handles the transitions between each scene and is responsible for rendering each scene to the OpenGL ES view. Each scene then consists of different layers, which are transparent drawable areas that can be stacked on top of one another to render a scene. Each game mode is encapsulated within its own scene. The source code for each is listed in the Appendices. Cocos2d-x also supports integration with the Box2D physics engine, and tile mapping, which was used in the *Moving Invaders* and *Cannon Launch* games. Cocos2d-x also provides scheduling support, which allows for simple asynchronous event-based programming.

MoMu is an audio library created by the Stanford Music, Computing, and Design Research

Group specifically for mobile platforms. This library was invaluable because it brings in the Synthesis Toolkit, which allows us to synthesize sounds other than pure-tones. It allows us to generate saw-tooth signals, white noise, and custom signal envelopes. The library was also developed for mobile platforms, so the synthesis functions were developed with low computation overhead in mind.

Audiogram and calibration data was stored locally using the iOS Core Data library. Core Data is essentially just a light SQLite wrapper for iOS. I developed a class around Core Data in C++, so the game itself would be independent from the storage components of the app. This would make cross platform development easier, when Sones gets ported to Android or Windows 8.

Remote data acquisition was possible using libCurl, which allows me to send HTTP requests to a remote server with the collected data. Data was packaged into JSON objects, which could be easily parsed by the webserver. The server backend was developed using Ruby on Rails, which also renders a webpage that allows a doctor to access patient audiograms remotely. The audiograms were rendered using Highcharts, a Javascript graphing library.

Testing Methodology

We followed the following methodology to evaluate the viability of Sones:

1. Collected diagnostic results from Sones from at least 20 participants. At least 2 participants should be in each 10-year age bracket and half of participants should have diagnosed hearing disorder.
2. Obtained participant's clinical audiogram where possible. If unobtainable, obtain audiogram using MLP toolbox (Grassi 2009)
3. Compared Sones 's audiogram with participant's audiogram noting mean square error and shape differences

4. Compared Sones's audiogram with Terhardt's ideal hearing model [4] or similar models for minimum hearing threshold for patients without a diagnosed hearing disability.
5. Analyzed the Sones's classification model with ANOVA and/or ROC analysis
6. Gauged usability by user ratings and feedback to questions in user study.

Results

Headphone/Earbud Calibration

Before any we could assess the viability of the diagnostic techniques we employed, we looked at the dynamic range of typical headphones and earbuds, and more importantly the fidelity of control available through the Core Audio API. On the iOS Core Audio, we are only given access to a gain control that scaled from 0.0 – 1.0, with the precision of a 32-byte float. We accessed the range of three different types of headphones/earbuds playing a 1kHz pure tone on the 2nd generation iPad, as shown below:

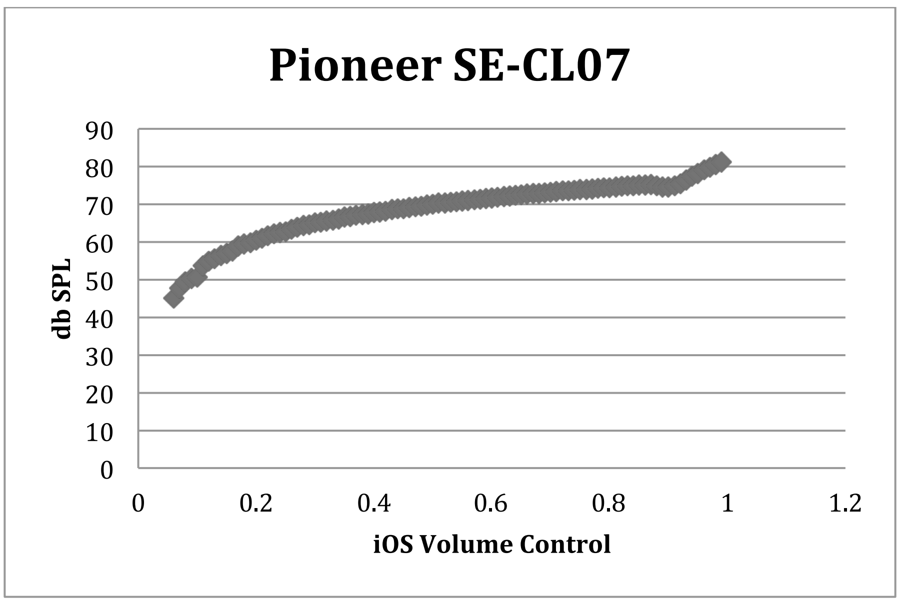

Figure 20 - Pioneer SE-CL07 Dynamic Range

The Pioneer earbuds scaled logarithmically, with a slight peak near the maximum volume levels allowed. The marked sensitivity is 100dB. Other tested earbuds showed a similar logarithmic range and are listed in Appendix A.

Classification Performance

We conducted several user studies to assess the viability of Sones under normal usage conditions. We then looked at the compiled audiograms, and their mean sq. error relative to typical audiogram shapes as established by Lee et al. [10]. The shape classification was determined by the lowest mean square error. These preliminary results look at Sones classification performance with the game *Target Practice*:

Table 1 - Mean Sq. Error and Classification Results

	User #1	User #2	User #3	User #4	User #5	User #6	User #7
Mean Sq. Error	247	500	2192	2040	1630	460	4170
Class	1	1	6	10	6	6	6

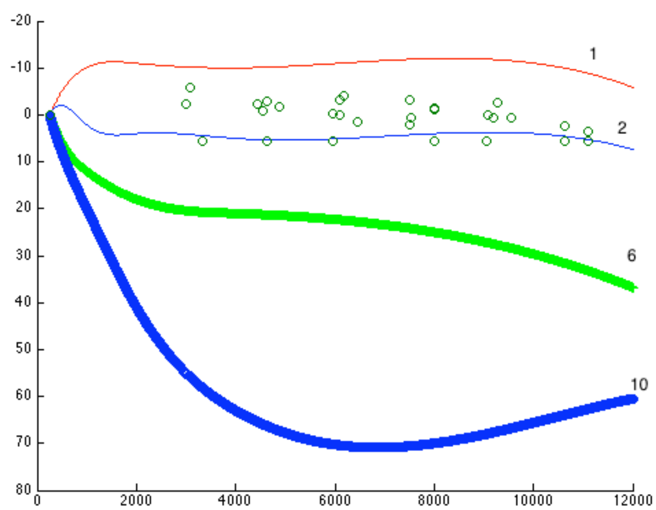

Figure 21 - User #1 Performance vs. Audiogram Shapes

Table 2 - Audiogram Shapes

Index	Audiogram Shape
1	Rising
2	Flat
3	Peaked 8kHz
4	4 kHz dip
5	8 kHz dip
6	Mild sloping
7	Severe 8kHz dip
8	Sloping
9	Abrupt Loss
10	Severe Sloping
11	Profound abrupt loss

Our user study was conducted on patients the ranging from 16-50 years of age. There were no prior records of hearing impairments or significant deterioration. Considering shapes of 1 & 6 are typical of normal hearing ability, our small trial had a classification rate of 85.7%.

User Feedback

Feedback was recorded after each trial about their whole diagnostic experience and comments on the app in general. Most comments focused on the following key issues:

1. Tutorial Instructions should be more informative
2. In *Target Practice* it was hard to associate the direction the enemy was traveling with the pure-tone it was emitting
3. Pure-tones aren't very pleasant to hear, especially at higher volumes

Discussion

From the headphone calibration data, we have established that the Core Audio API does map to the full dynamic range of most headsets and earphones. The resolution of the volume control is also sufficient for our diagnostic purposes. Since the volume mapping isn't linear, most minimum thresholds would usually be within the range of 0.01-0.3. The fidelity of the possible dynamic ranges on the iPhone/iPad will depend on the headset being used. The headphone jack typically only supplies 30mW of power per channel. Thus, if the output impendence of the headset is too high, the possible volume range will be lowered.

The user study we performed only surveyed 7 users thus far, so it's hard to say if the results we obtained are indicative of Sones' actual viability. The largest gap was that the tests were performed only on users without any hearing disorders. The test group was also extremely small. A sample size of around 20-40 users would have been desirable. The results do look promising so far, with 6/7 classifications seemingly making sense with their current hearing pathology. In the future, we should cross validate our results by conducting actual audiometric testing, such as using the MLP Matlab toolbox, to measure actual clinical audiograms.

From the user feedback, it appears that the game concepts we implemented in Sones are not immediately intuitive. Specifically, it was unintuitive having different tones map to a specific radial direction. A more intuitive approach would be to lose the mapping all together and spatialize the sound played into 3D space. It is also apparent that more work can be done with improving the usability of the app in general to help mitigate the steep learning curve. One solution would be to just run more user studies, and iterate our interface based on the feedback we receive.

From a design perspective, Sones has fulfilled most of its design requirements that we established earlier. We have successfully meshed together several diagnostic techniques with game elements, creating an extremely interactive diagnostic tool on a mobile device. Sones is also capable of syncing the diagnostic results to a remote server, allowing for audiologists to provide virtual consultations. We also created a platform that can properly measure the dynamic range of headphones and calibrate its own playback based on those measurements.

Conclusion

From our preliminary user study, we established that Sones does provide a viable and robust mobile hearing diagnostics platform. The sensors on our mobile devices are capable of performing complex medical tasks. In the future, with more complex peripherals, it maybe possible to have an entire suite of diagnostic tools on your phone, vastly changing our whole conceptualization of healthcare and our perception of the standard patient/doctor relationship.

In the future, we would like to develop a more precise and robust classification model. The model will be built on collected user data, and possibly external sources. AudGenDB looks to be a promising source, as it has 57,278 readily accessible and anonymous pediatric audiograms. A simple k-means clustering model should probably suffice, although other methods maybe explored. We would also like to explore other diagnostic techniques, such as exploring the correlation between sound localization ability and tinnitus.

References

. [1] W. H. Organization, “Deafness and hearing impairment,” Febru- ary 2012.

. [2] S. D. W., M. S, M. S, M. H, and T. S, “Hearing assessment re- liability, accuracy, and efficiency of automated audiometry,” TELEMEDICINE and e-HEALTH VOL. 16 NO. 5, 2010.

. [3] B. S. of Audiology, Pure tone air and bone conduction thresh- old audiometry with and without masking and determination of uncomfortable loudness levels. British Society of Audiology, 2004.

. [4] E. Terhardt, “Calcuating virtual pitch,” Hearing Res., 1979.

. [5] L. D. Fielder and E. M. Benjamin, “Subwoofer performance for accurate reproduction of music,” ournal of the Audio Engineering Society, Volume 36, Number 6, 1988.

. [6] J. Jerger, “Hearing tests in otologic diagnosis,” ASHA, 1962.

. [7] P. McKinnon, Dead regions in the cochlea: What are they and why do they matter? Siemens, 2006.

. [8] V. Summers, M. R. Molis, H. Musch, B. E. Walden, R. K. Surr, and M. T. Cord, “Identifying dead regions in the cochlea: Psy- chophysical tuning curves and tone detection in threshold- equalizing noise,” Ear & Hearing, 2003.

. [9] M. L. Jepsen and T. Dau, “Characterizing auditory processing and perception in individual listeners with sensorineural hearing loss,” J.Acoust.Soc.Am, 2011.

. [10] C.-Y. Lee, J.-H. Hwang, S.-J. Hou, and T.-C. Liu, “Using cluster analysis to classify audiogram shapes,” International Journal of Audiology, 2010.

. [11] L. Ciletti and G. A. Flamme, “Prevalence of hearing impairment by gender adn audiometric configuration: Results from the na- tional health and nutrition examination surey adn the keokuk county rural health study,” Journal of the American Academy of Audiology, 2008.

. [12] R. Tyler, C. Coelho, P. Tao, H. Ji, W. Noble, A. Gehringer, and S. Gogel, “Identifying tinnitus subgroups with cluster analysis,” American Journal of

Audiology, 2008.

. [13] M. N. Anwar and M. P. Oakes, “Data mining of audiology pa- tient records: factors influencing the choice of hearing aid type,” BMC Medical Informatics and Decision Making, 2011.

. [14] S. Nouraei, Q. Huys, P. Chatrath, J. Powles, and J. Harcourt, “Screening patients with sensorineural hearing loss for vestibular schwannoma using a bayesian classifier,” Clinical Otolaryngology, 2007.

. [15] O. Ozdamar, R. E. Eilers, E. Miskiel, and J. Widen, “Classifica- tion of audiograms by sequential testing using dynamic bayesian procedure,” Acoustical Society of America, 1990.

. [16] N. SE and O. SO, “Validation of play-conditioned audiometry in a clinical setting,” Scand Audiol, 1997.

. [17] S. M. Baldwin, B. J. Gajewski, and J. E. Widen, “An evalua- tion of the cross-check principle using visual reinforcement au- diometry, otoacoustic emissions and tympanometry,” J Am Acad Audiol, 2010.

. [18] A. S, “Visual reinforcement audiometry: An adobe flash based approach,” Journal of Visual Communication in Medicine, 1997.

Appendix A – Headphone/Earbud Dynamic Ranges

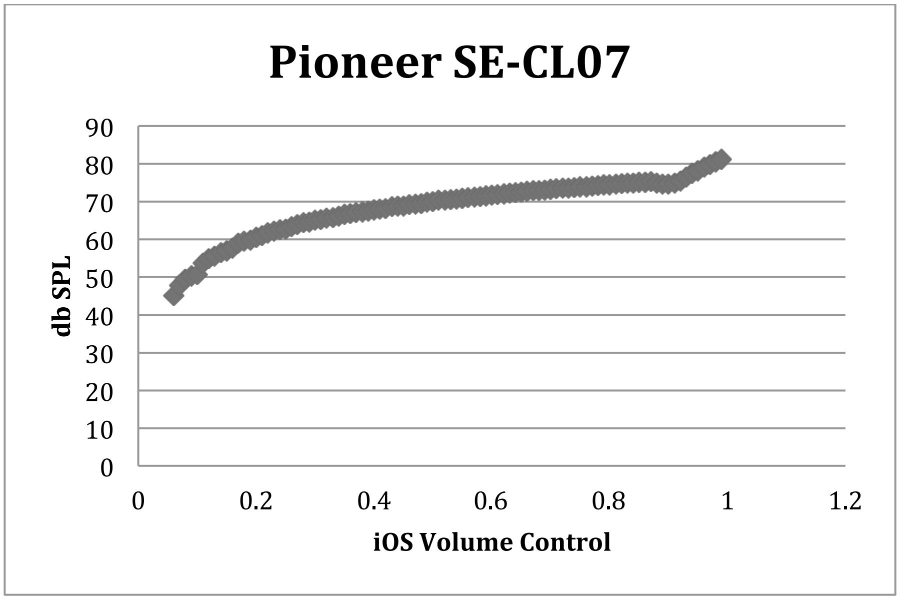

Figure 22 - Pioneer SE-CL07 Dynamic Range

The Pioneer earbuds scaled logarithmically, with a slight peak near the maximum volume levels allowed. The marked sensitivity is 100dB.

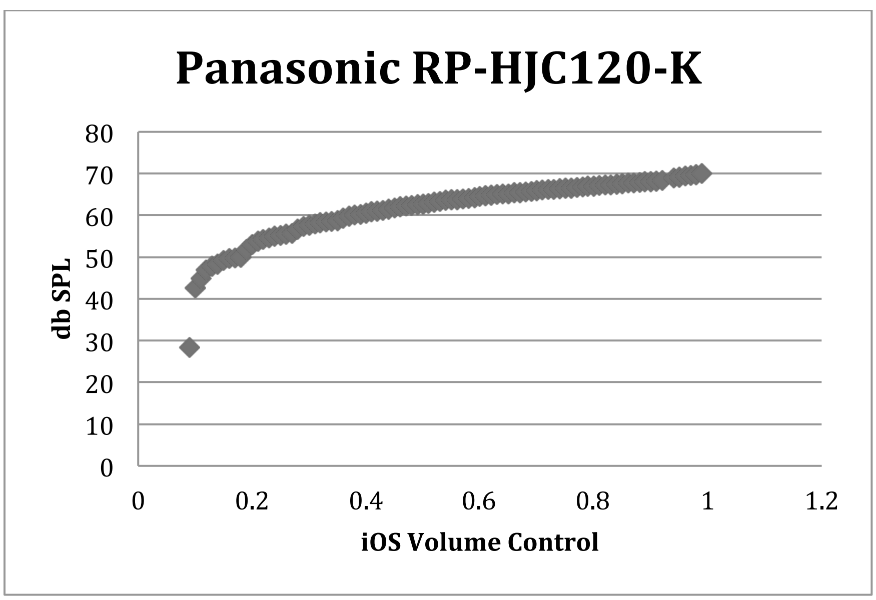

Figure 23 - Panasonic RP-HJC120-K Dynamic Range

The Panasonic earbuds also scaled logarithmically, without the peak. The marked sensitivity is 96dB.

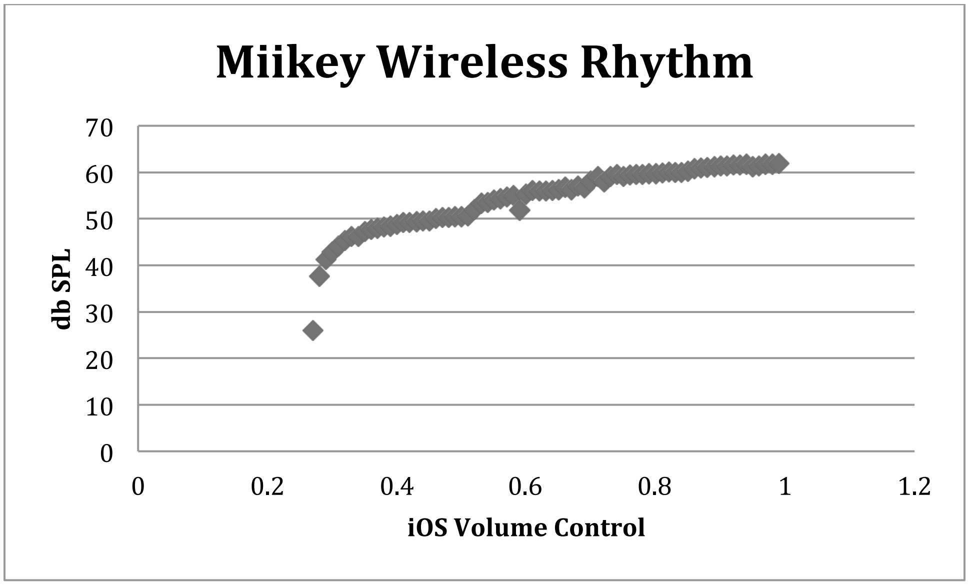

Figure 24 - Miikey Wireless Rhythm Dynamic Range

The Miikey Rhythm headsets also scaled logarithmically, with a bit more perturbation compared to the earbuds.

Appendix B – User Study Data

User #1	ZATTEMPTS	ZCHANNEL	ZFINAL_TIME	Z_ENT	ZFREQ	ZINIT_TIME	ZVOLUME
1	1	0	4098160640	1	8003	0	24.5
2	1	1	6912	1	8003	6656	17.30850029
3	1	0	7168	1	8003	6656	17.57861137
4	2	1	8192	1	6453	6912	20.22405624
5	1	0	7680	1	253	7424	18.88272285
6	1	0	12544	1	9052	12288	24.5
7	7	1	13568	1	5952	7680	24.5
8	2	0	10240	1	7502	7936	20.74430656
9	1	0	8960	1	5952	8704	18.48893738
10	1	1	7680	1	7502	7424	15.73930645
11	1	0	12032	1	10626	11776	21.16067696
12	3	0	11264	1	7526	8192	19.47432709
13	10	1	17152	1	10626	13056	24.5
14	3	1	9472	1	4426	0	16.57908821
15	1	1	10752	1	9076	10496	18.97146988
16	1	1	12800	1	9203	12544	19.36003494
17	2	0	12288	1	6103	11008	18.90225601
18	2	1	10496	1	6103	8192	15.75114536
19	2	1	10752	1	3003	9472	16.62670135
20	3	1	12032	1	4553	8960	18.08447838
21	1	1	11776	1	9286	11520	16.14225388
22	1	1	9728	1	3086	9472	13.08052444
23	1	1	10752	1	6186	10496	14.59756184
24	7	1	18432	1	4636	10752	24.5
25	1	0	11520	1	4636	11264	15.86966038
26	1	1	13056	1	11093	12800	22.45724106
27	1	1	14080	1	11093	14080	24.23987389
28	10	0	14592	1	3343	7424	24.5
29	1	1	9984	1	4893	9728	16.94346237
30	1	1	11264	1	9543	11264	19.51440048

Z_User #2

#2	ZATTEMPTS	ZCHANNEL	ZFINAL_TIME	ZFREQ	ZINIT_TIME	ZVOLUME
1	1	0	4098982144	8010	0.000861581	6.125
2	1	0	10496	260	10240	6.125
3	6	1	12800	8010	7680	6.125
4	3	1	11008	3360	8448	6.125
5	4	0	11008	3360	8192	6.125
6	1	0	13056	11238	12800	6.125
7	1	0	9472	5038	9216	4.768900871
8	5	1	14080	3488	8960	6.125
9	1	1	10240	3488	9984	5.260752678
10	3	1	13056	8138	10752	6.125
11	1	0	12288	8591	12032	5.340962887
12	1	0	11264	2391	11008	4.91937542
13	2	0	11264	5491	9984	4.924444675
14	6	1	14848	7041	10496	6.125
15	1	1	12800	8591	12800	5.616518021
16	2	1	12032	9313	1792	4.603342056
17	1	1	12544	1563	12544	4.854453087
18	1	1	12800	7763	12800	4.974452972
19	3	0	12800	6213	9984	4.934453011
20	1	0	12288	3113	12032	4.666675568
21	1	1	16384	9947	16384	5.589143753
22	1	0	13312	2197	13312	4.50469923
23	1	1	14080	3747	13824	4.754575729
24	1	0	13312	5297	13056	4.517538548
25	1	0	15616	8397	15616	5.324451923
26	1	1	2048	4401	1792	0.91871345
27	1	0	9216	9051	8960	3.911571026
28	1	1	9984	2851	9728	4.272355556
29	2	0	12544	1301	11008	5.375298023
30	3	0	13056	2851	10240	5.600616932

User #3	Z_PK	Z_OPT	ZATTEMPTS	ZFINAL_TIME	ZFREQ	ZINIT_TIME	ZCHANNEL	ZVOLUME
1	1	1	5	12320	5348	8864	1	98
1	2	1	3	11904	2248	7968	1	98
1	3	1	2	<null>	8586	1.23E-25	0	0.000222222
1	4	1	2	12640	8586	32	1	98
1	5	1	5	11040	5486	6592	1	98
1	6	1	2	12000	7036	11072	0	98
1	7	1	4	11616	7036	9152	0	98
1	8	1	11	9920	8853	128	0	7.300926685
1	9	1	5	9184	2653	480	0	82.39444733
1	10	1	6	14016	7303	288	0	98
1	11	1	3	8128	4203	288	0	66.27574921
1	12	1	8	14720	7303	3392	1	98
1	13	1	7	9888	8005	3648	1	6.664286137
1	14	1	1	15936	9555	15840	0	98
1	15	1	8	15680	6455	9760	1	98
1	16	1	2	16896	9555	128	0	98
1	17	1	7	12512	1805	64	0	87.60905457
1	18	1	5	4064	7658	128	1	5.404583454
1	19	1	15	9312	9208	128	1	10.09805584
1	20	1	2	9280	1458	8992	1	66.84486389
1	21	1	14	15904	6108	9344	0	97.42236328
1	22	1	2	8768	3008	8512	1	53.83124924
1	23	1	24	19744	8244	2176	0	4.993703842
1	24	1	3	16192	494	128	0	93.0451889
1	25	1	2	2080	6694	1184	0	11.44111156
1	26	1	12	17984	5144	10496	1	97.91222382
1	27	1	2	11584	2044	11232	0	63.14678955
1	28	1	7	18176	10322	160	1	5.975464344
1	29	1	1	9888	5672	9696	1	73.53644562

User # 4	Z_ENT	Z_OPT	ZATTEMPTS	ZCHANNEL	ZFINAL_TIME	ZFREQ	ZINIT_TIME	ZVOLUME
30	1	1	1	0	9568	2312	9312	93.98066711
31	1	1	3	0	12384	3862	7808	98
32	1	1	1	0	7392	762	7072	72.24910736
33	1	1	3	1	13408	2312	8192	98
34	1	1	4	1	13024	2312	8448	98
35	1	1	4	1	14912	10250	192	7.206111908
36	1	1	2	1	7712	2500	128	70.06852722
37	1	1	6	1	13696	7150	7520	98
38	1	1	3	1	14304	10250	11872	98
39	1	1	11	0	16064	7150	7232	98
40	1	1	2	1	13184	8710	128	6.420476437
41	1	1	0	1	480	8710	128	9.752698898
42	1	1	5	0	13440	4060	8384	94.2731781
43	1	1	4	0	14144	7160	9984	98
44	1	1	1	1	10048	960	9792	70.21730804
45	1	1	2	0	15424	9718	14016	5.191249847
46	1	1	7	1	18048	6618	9376	98

User #5	ZATTEMPTS	ZCHANNEL	ZFINAL_TIME	ZINIT_TIME	ZVOLUME	ZFREQ
1	3	1	8384	5856	82.14778137	5408
2	1	0	7232	6880	70.76999664	2308
3	1	1	6624	6304	65.08110809	2308
4	2	0	8768	7072	85.78866577	2308
5	3	0	47360	44608	98	758
6	1	1	10720	10464	7.016667366	8558
7	1	1	7072	6720	64.75907898	808
8	1	0	7104	6752	58.12166977	2358
9	1	0	6912	6560	56.32019043	808
10	1	1	5312	4928	43.33055878	7008
11	1	1	10816	10560	6.826826096	11134
12	11	1	14784	5184	98	4934
13	4	0	15328	11040	98	11134
14	2	0	7200	5472	50.46873474	4934
15	1	1	5792	5376	40.47254181	4934
16	1	1	10528	10272	6.044583321	8481
17	2	0	7072	128	49.42250061	731
18	3	1	9824	6336	60.0890274	3831
19	15	0	12128	6112	74.31124878	3831
20	2	1	8704	7200	53.40458298	3831
21	2	0	12352	10880	5.056913376	9701
22	1	1	7648	7328	46.64913559	1951
23	2	1	10720	9152	58.21641922	8151
24	1	1	9760	9472	53.15962982	1951
25	2	0	7520	64	41.0233345	1951
26	1	1	11008	10752	6.612830639	11688
27	1	0	5312	4992	42.94286346	3938
28	4	1	12032	7840	82.45941925	8588

User #6	ZATTEMPTS	ZCHANNEL	ZFINAL_TIME	ZFREQ	ZINIT_TIME	ZVOLUME
1	1	0	4098160640	8003	0	24.5
2	1	1	6912	8003	6656	17.30850029
3	1	0	7168	8003	6656	17.57861137
4	2	1	8192	6453	6912	20.22405624
5	1	0	7680	253	7424	18.88272285
6	1	0	12544	9052	12288	24.5
8	2	0	10240	7502	7936	20.74430656
9	1	0	8960	5952	8704	18.48893738
10	1	1	7680	7502	7424	15.73930645
11	1	0	12032	10626	11776	21.16067696
12	3	0	11264	7526	8192	19.47432709
14	3	1	9472	4426	0	16.57908821
15	1	1	10752	9076	10496	18.97146988
16	1	1	12800	9203	12544	19.36003494
17	2	0	12288	6103	11008	18.90225601
18	2	1	10496	6103	8192	15.75114536
19	2	1	10752	3003	9472	16.62670135
20	3	1	12032	4553	8960	18.08447838
21	1	1	11776	9286	11520	16.14225388
22	1	1	9728	3086	9472	13.08052444
23	1	1	10752	6186	10496	14.59756184
25	1	0	11520	4636	11264	15.86966038
26	1	1	13056	11093	12800	22.45724106
27	1	1	14080	11093	14080	24.23987389
29	1	1	9984	4893	9728	16.94346237
30	1	1	11264	9543	11264	19.51440048

User #7	ZFREQ	ZVOLUME
1	1764	86.68820953
2	1764	98
3	1260	79.14172363
4	1512	58.24376297
5	1260	98
6	1764	98
7	291	42.95729446
8	2619	97.91118622
9	291	98
10	2037	91.71920013
11	582	93.07369995
12	2619	98
13	291	91.91269684
14	2619	86.10771179
15	2910	98
16	271	98
17	2439	94.62169647
18	542	95.20219421
19	2168	98
20	1355	98
21	1897	98
22	813	94.81519318
23	1355	98
24	271	90.55820465
25	1355	91.71920013
26	1897	89.20370483
27	2439	98
28	542	92.88019562
29	1626	91.91269684
30	1626	98
31	2710	87.07521057
32	542	88.62320709
33	271	84.17271423
34	813	87.26870728
35	1084	96.94368744
36	1084	98
37	283	98
38	2547	78.7547226