On Acoustic Covert Channels Between Air-Gapped Systems

Brent Carrara(✉) and Carlisle Adams

School of Electrical Engineering and Computer Science,
University of Ottawa, Ottawa, ON, Canada
{bcarr092,cadams}@uottawa.ca

Abstract. In this work, we study the ability for malware to leak sensitive information from an air-gapped high-security system to systems on a low-security network, using ultrasonic and audible audio covert channels in two different environments: an open-concept office and a closed-door office. Our results show that malware installed on unmodified commodity hardware can leak data from an air-gapped system using the ultrasonic frequency range from 20 kHz to 20.5 kHz at a rate of 140 bps and at a rate of 6.7 kbps using the audible spectrum from 500 Hz to 18 kHz. Additionally, we show that data can be communicated using ultrasonic communication at distances up to 11 m with bit rates over 230 bps and a bit error rate of 2 %. Given our results, our attacks are able to leak captured keystrokes in real-time using ultrasonic signals and, using audible signals when nobody is present in the environment - the **overnight attack**, both keystrokes and recorded audio.

Keywords: Malware communication · Audio communication · Ultrasonic · Jumping air-gaps · Out-of-band covert channels

1 Introduction

Physically separating computers on a high-security network from computers on a low-security network is a security practice that is commonly employed in high and medium security environments to protect sensitive information from unauthorized access [12,22]. This security practice prevents malware on the low-security network from attacking computers on the high-security network by means of traditional virus and worm network propagation vectors (e.g. e-mail, network shares, web servers, and web clients). The physical separation between the two networks is commonly referred to as an **air-gap**, a term that predates wireless networks, because of the literal physical air-gap separation between the two networks.

Air-gapped networks are employed in many industrial and government settings, including military [12], financial [12], nuclear power [22], and aviation [27] environments. In addition to being used in high-security environments, the practice of using air-gapped machines is recommended by security researchers (see Schneier [23], for example) and even Osama bin Laden was reported to have used

F. Cuppens et al. (Eds.): FPS 2014, LNCS 8930, pp. 3–16, 2015.
DOI: 10.1007/978-3-319-17040-4_1

one to protect his communications while in seclusion [6]. Given Schneier's setup in [23], once the air-gapped machine is configured and disconnected from the network there is no risk that a machine connected to the Internet (low-security machine) could access the air-gapped system via the network. Even so, there are still a number of ways that malware could be installed on the air-gapped system including a malicious insider [19], trojan horse executable [25], malicious payload in a trusted file format [24], or a malicious actor in the supply chain [4]. Once the malware is installed on the air-gapped system, the software requires a covert channel to propagate data to and from the low-security network. A covert channel can be characterized as a channel that has a low probability of intercept (LPI) and when evaluating a covert channel, the adversary and the adversary's capabilities must be assumed so that the effectiveness of the adversary's ability to detect the presence of the covert channel can be evaluated.

In our study, we analyzed the ability for malware introduced onto an air-gapped system, using any of the methods described above, to communicate sensitive information to a remote system on a low-security network using an *out-of-band covert channel*, the audio channel, in real-world settings. We studied both ultrasonic audio communication in the range of 20 kHz to 22 kHz (20 kHz is understood to be the cutoff frequency that can be heard by a young person, which decreases with age [9]) and audible audio communication in the range of 0 Hz to 20.5 kHz. Our results demonstrate that, in general, captured keystrokes, encryption key material (e.g. private keys, shared keys), credentials (e.g. passwords), documents, and even recorded audio can effectively be leaked from compromised air-gapped systems. Furthermore, our attack requires no hardware modifications to the air-gapped system or any system on the low-security network. Our attack requires only that the air-gapped system be equipped with a speaker and that a system on the low-security network be equipped with a microphone. Additionally, we consider our victim as unaware and unassuming of our attack and therefore measure the covertness of our channel by a human's natural ability to hear the communication.

In our study, we analyzed the effectiveness of our attack, i.e. the malware's ability to leak data using audio communications, in two traditional office environments: a closed-door office containing a single desk and multiple computers (approximately 3 m × 3 m × 2.8 m in dimension) and an open-concept office containing four desks, each holding multiple computers (approximately 4.27 m × 4.27 m × 2.8 m in dimension). Furthermore, we evaluated the ability for malware to covertly communicate within both of these environments in two different scenarios: the malware on the air-gapped system leaks data from the air-gapped machine when humans are present in the room (i.e. during regular business hours) and the malware leaks data from the compromised air-gapped machine after hours when no humans are present (i.e. after regular business hours), an attack we call the **overnight attack**.

As a result of our study, we make the following novel contributions:

1. We demonstrate that, in general, unmodified commodity hardware from major hardware vendors is capable of the following:

   (a) communication using the ultrasonic spectrum (i.e. 20 kHz to 20.5 kHz) at bit rates greater than 140 bps and bit error rates (BERs) below 10 %.
   (b) communication using the audible spectrum (i.e. 0 Hz to 20 kHz) at bit rates greater than 6.7 kbps and BERs below 15 %.
2. We introduce the concept of the **overnight attack** and demonstrate that, given our achievable bit and BERs, the threat of the overnight attack challenges the traditional threat model that is based on the assumption that covert audio communication is strictly low-bandwidth [7,8].
3. We demonstrate that covert ultrasonic audio communication is an effective mechanism for leaking data, including captured keystrokes in real-time, from air-gapped systems in real-world environments (e.g. open-concept and closed-door offices) under real-world settings (e.g. people talking and a radio playing nearby).
4. We demonstrate that ultrasonic communication can leak data from compromised systems at distances up to 11 m at bit rates of over 230 bps and a BER of 2 %.

Our paper is organized as follows. In Sect. 2, we review the related literature. In Sect. 3, we introduce the acoustic channel, its effects on audio communication, and the physical environments we studied in this work. In Sect. 4, we present our experiments in detail as well as our results. Furthermore, in Sect. 5, we present protection mechanisms that can be employed to detect our attack and, finally, in Sect. 6, we conclude.

2 Related Work

Utilizing audio signals for the purpose of covert communication and leaking information from an air-gapped system has previously been discussed in [7,8,19]. In [7], Hanspach, et al., built a proof-of-concept network using five identical Lenovo laptops to demonstrate that audio communication between the computers can be achieved in the near-ultrasonic range from 17 kHz to 20 kHz. Their research demonstrates that frequency-hopping spread spectrum (FHSS) with 48 sub-channels can be used effectively to establish a covert channel capable of transmitting data at a rate of 20 bps up to a distance of 19.7 m. While novel, their research demonstrates extremely low bit rates and their paper provides no in-depth BER analysis. Furthermore, their study only examines the ability of five identical model laptops, Lenovo T400, running Debian Linux to communicate with one another. In contrast to their study, we demonstrate our attack on multiple systems from various manufacturers running both Mac OS X and Windows. Furthermore, we demonstrate that orthogonal frequency-division multiplexing (OFDM) is a much more appropriate modulation scheme given the nature of the audio channel and that as a result of using OFDM we are able to effectively demonstrate bit rates over 25x faster than those previously reported using the near-ultrasonic and ultrasonic bandwidths. In [8], Hanspach, et al., documented the ability to communicate using the same Lenovo T400 model laptops, over the ultrasonic range from 20.5 kHz to 21.5 kHz at a speed of 20 bps up to a range of 8.2 m. In our study, we found that, in general, commodity hardware has difficulty reliably communicating above 20.5 kHz, but that bit rates above 200 bps can be achieved, in general, in the 20 kHz to 20.5 kHz bandwidth. In [7,8], Hanspach, et al., also proposed filtering out near-ultrasonic and ultrasonic frequencies from all audio being sent to the speakers as a defensive mechanism. This technique would be ineffective against our **overnight attack**. In [19], O'Malley, et al., established a covert channel between a MacBook Pro and a Lenovo Tablet using Frequency Shift Keying (FSK) in the ultrasonic range between 20 kHz and 23 kHz using an external speaker. The bit rates they achieved using external hardware, in the ultrasonic range, are comparable to ours; however, in our attack no additional hardware is required.

The use of audio has also been researched as an alternative to traditional wireless communication (e.g. infrared (IR), Bluetooth, radio frequency (RF), Wi-Fi, and near-field communications (NFC)); however, due to the relatively low bandwidth available in the channel when compared to IR and RF as well as the negative impacts of audio on humans and animals, this alternative solution has been primarily only studied in academic circles [5,10]. In [5], Gerasimov, et al., studied the use of audio communication over-the-air for the purposes of device-to-device communication and were able to achieve bit rates of 3.4 kbps, about half the bit rate we were able to achieve using the spectrum from 0 Hz to 20 kHz. In [13–15], researchers examined the ability to communicate using audio signals that are pleasant to humans to exchange pre-authorization information required for wireless networks as well as uniform resource locators (URLs). The researchers synthesized audio signals using frequencies from musical scales, chords, and lullabies as well as from fictional characters (e.g. R2D2 from Star Wars) and insects. Similarly, in [3], Domingues, et al., studied the ability to communicate using audio signals that sound like musical instruments (e.g. piano, clarinet, and bells). In [16,17], Madhavapeddy, et al., examined audio communication as an alternative to Bluetooth wireless communication through the use of dual-tone multi-frequency (DTMF) signalling, on-off keying and melodic sounds. Of relevance to our study, Madhavapeddy, et al., studied ultrasonic communication between two laptops with third-party speakers. Lastly, in [18], Nandakumar, et al., experimented with audio communication as an alternative to NFC.

3 Acoustic Channel

In this section, we take an in-depth look at the characteristics of the acoustic channel and demonstrate, by way of measurement, that the over-the-air channel causes multipath delays and has a non-ideal frequency response.

3.1 Environments Studied and System Requirements

The main goal of our study was to challenge the existing threat model posed by unauthorized audio communication between air-gapped systems. Previous researchers have demonstrated the ability to use audio signals to bridge the

Table 1. Machines and parameters of our study

ID	Make	Model	Distance, ∠ in closed environment	Distance, ∠ in open environment
Audio1	Lenovo	IdeaPad S10	1.19 m, 50 °	3.89 m, 4 °
Audio2	Lenovo	ThinkPad X120e	1.55 m, 10 °	N/A, N/A
Audio3	Dell	Precision T3500	1.88 m, 33 °	3.68 m, 21 °
Audio4	HP	HP Mini	1.50 m, 38 °	1.47 m, 90 °
Audio5	Acer	Aspire One	1.91 m, 355 °	3.33 m, 335 °
Audio6	Alienware	M15X	2.36 m, 330 °	1.45 m, 305 °
Audio7	Sony	Vaio	2.11 m, 340 °	3.84 m, 47 °
Audio8	Apple	MacBook Pro	N/A, N/A	N/A, N/A

air-gap; however, they have only demonstrated very limited bit rates on very specific hardware in constrained environments [7,8,19]. Our study shows that not only is audio communication possible using unmodified commodity hardware but that the achievable bit rates in both the ultrasonic and audible ranges are well above what has previously been reported. Our work demonstrates that audio provides an effective covert channel to communicate low volumes of data, such as captured keystrokes and cryptographic key material, as well as more substantial data streams, such as captured audio.

In our study, we placed a desktop with a USB headset as well as seven laptops in two real-world office environments. In all our experiments, the MacBook Pro, **Audio8**, was used as the air-gapped system and the remaining seven systems were connected to the low-security network. The distances and angles between **Audio8** and the other machines can be seen in Table 1 for both our environments, where an angle of 0° can be taken to be directly in front of **Audio8** and positive offset angles are measured clockwise. The systems in our study were all configured with Windows 7, aside from the MacBook Pro, which was configured with OS X 10.9. Additionally, none of the machines in our study had any hardware added or modified with the exception of the Dell desktop, **Audio3**, which had a USB headset with a microphone and speaker added because it had none built in. Unfortunately, our Lenovo ThinkPad, **Audio2**, suffered from a hardware failure following our closed-door office tests and was unavailable for testing during our open-concept office tests.

The detailed hardware and configuration requirements for the systems in our study are summarized here:

- **Air-Gapped System:**
  1. An audio speaker
  2. No network connections to the low-security machines
  3. Configured according to the steps outlined in [23]
- **Low-Security System:**
  1. A microphone

Given the uniform manner in which the systems were configured we feel that our study's environmental model closely mimics a real-world corporate office where the same version of Windows is installed on every machine. It should be noted that, for the sake of brevity, we only show detailed bit rate and BER results with the MacBook Pro as the air-gapped system using uni-directional communication. In Sect. 4, we provide the combined microphone and speaker frequency responses of each of the other machines' speakers and the MacBook Pro's microphone in order to assure the reader that all the machines tested are capable of transmitting audio signals in the audible and ultrasonic ranges in addition to receiving them.

3.2 Measured Channel Characteristics in Our Environments

To properly engineer our attack, we measured the noise and the reverberations caused by objects in the environment as well as the frequency response of the microphones and speakers built into the systems used in our study. In this section, we present the experiments we used to measure each of these quantities.

It has previously been reported that audible background noise drops off exponentially with frequency [26]. For the purposes of our study, we noted that the noise levels in our environment, in general, are quite low. Particularly at frequencies above 5 kHz and especially so for frequencies above 10 kHz. This is due to a combination of factors: equipment in our environment is generating very little audible background noise at frequencies above 5 kHz and the frequency response of the microphones in the systems we studied are not as sensitive to frequencies above 10 kHz as they are to frequencies between 0 Hz and 10 kHz. Furthermore, ultrasonic communication is not subject to the same degree of background noise that audible communication is subject to. This manifests itself in our results through the fact that the error rates for ultrasonic communication are lower than the error rates for audible communication.

We quantified the *multipath spread* of the channel in our environments by performing two experiments. In the first, we transmitted a 100 ms signal from **Audio8** consisting of pure-tone frequencies between 5 kHz and 10 kHz at 500 Hz intervals and, in the second, we transmitted a 100 ms signal from **Audio8** consisting of pure-tone frequencies between 20 kHz and 22 kHz also at 500 Hz intervals. To measure the *multipath spread* of the channel, we took the two received waveforms, filtered out all frequencies outside of the passbands (e.g. 5 kHz to 10 kHz and 20 kHz to 22 kHz, respectively), then cross-correlated them with the original signal according to the algorithm in [20]. We then plotted the normalized magnitude of the cross-correlated signals over time to determine the *multipath spread* for two dB thresholds, -10 dB and -15 dB. Our experiment showed that copies of our original signal were received for 175 ms at the -10 dB threshold after the line-of-sight signal was received and 200 ms at the -15 dB threshold in the closed-door environment using audible signals. The observed reverberations varied from system to system due to their location in the room. Given the results of our experiments in the closed office setting, we performed experiments using delays of 125 ms for our tests using ultrasonic frequencies and

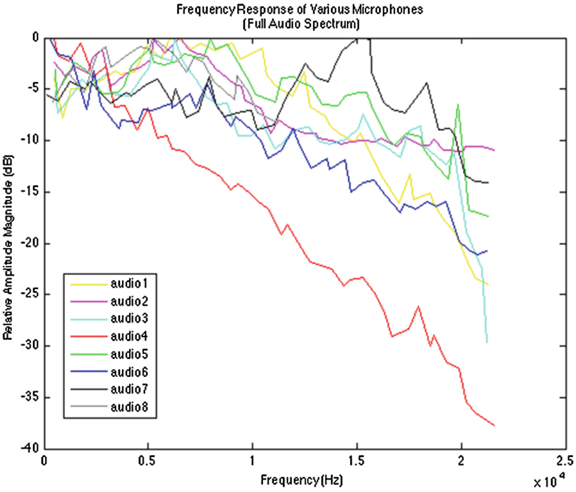

Fig. 1. The combined normalized frequency response of the audio channel. From the diagram it is clear that, over the $W = 0\,\text{Hz}$ to 22.05 kHz bandwidth, the channel does not have an ideal frequency response.

delays of 175 ms and 225 ms for our tests using audible frequencies. For our tests in the open-concept office environment, we performed experiments using 80 ms for our ultrasonic tests and 125 ms and 150 ms for our audible tests as there were much fewer echoes to deal with in the open-concept environment.

The frequency responses of the systems are shown in Fig. 1. All quantities are in dB and are normalized to the highest amplitude received by each of the speakers for any frequency component. An ideal frequency response would show as a horizontal line at the 0 dB level across all frequencies. From the figure it can be seen that the audio channel used in our study does not have an ideal frequency response. Furthermore, the majority of the frequency responses demonstrate a near-ideal response up to around 5 or 6 kHz. This is to be expected as the human voice consists primarily of frequencies between 300 Hz and 3.4 kHz [2] and it would be expected that the manufacturers of the microphones would design a device that has an ideal frequency response in this range. Given the non-ideal frequency response of the audio channel, we used OFDM as our modulation scheme. We divided the available channel bandwidth, W, into a number of equal-bandwidth sub-channels, N, such that each of the sub-channels has an ideal frequency response. We then modulated symbols using FSK on each sub-channel.

4 Experiments and Results

We performed a number of experiments in both the open-concept and closed-door office environments to determine the achievable bit rates and corresponding BERs using audio communication in the ultrasonic and audible frequency ranges. We measured the achievable bit rates to better model the threat posed by malware using an audio covert channel to leak sensitive data from an air-gapped machine. By quantifying the achievable bit rates we are able to determine the type of sensitive data that can be leaked by malware in real-world scenarios. We tested our attack in real-world environments (e.g. open-concept office, closed-door office) under real-world conditions (e.g. radio playing, people talking) in real-world scenarios (e.g. when humans are present, when they are not). In this section, we present the results of the following experiments:

1. In both the open-concept and closed-door office environments, we varied the channel bandwidth, W, the delay or *settle time*, k, and the number of sub-channels, N, to determine the maximum bit rate at which we could communicate with the lowest BER (note that increasing W, increasing N, and decreasing k all increase our effective bit rate).
2. Given the optimal bandwidth, *settle time*, and number of sub-channels, we attempted to communicate in the ultrasonic spectrum while a clock radio was playing a local radio station in the closed-door office environment.
3. Given the optimal bandwidth, *settle time*, and number of sub-channels, we attempted to communicate in the ultrasonic spectrum while conversations were taking place in the closed-door office environment.
4. We used **Audio7** to determine the maximum distance that we could communicate over using our ultrasonic attack.
5. Lastly, we tested **Audio8**'s ability to receive ultrasonic communication from each of the other machines in our study.

In Fig. 2, we show the BER for our ultrasonic tests using the $W = 20\,\text{kHz}$ to $20.5\,\text{kHz}$ bandwidth. From the figure it can be seen that as we increased the number of channels in the ultrasonic experiments, our BER increased dramatically.

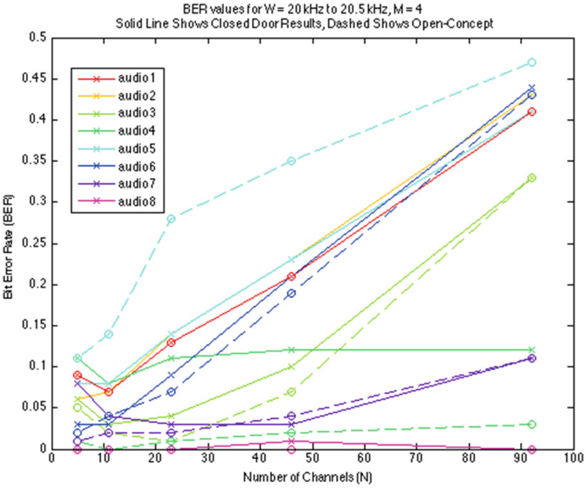

Fig. 2. BER for our ultrasonic tests in both the open-concept and closed-door environments.

Table 2. Average time (s) to leak popular document types

Document type	Average size (kb) per page [1]	Leak time using **overnight attack**
Microsoft Word	15	3.27
Microsoft Excel	6	1.31
Microsoft PowerPoint	57	12.42
Portable Document Format	100	21.79
Text	1.5	0.33
Email	10	2.18
Tagged Image File Format	65	14.16

In the closed-door office environment we achieved a BER of approximately 10 % or less for all machines with $N = 5$ and $N = 11$ and in the open-concept environment we also achieved a BER of 10 % with $N = 5$. With $N = 11$ all machines had a BER below 10 % except our Acer machine, **Audio5**, in the open-concept environment. This was due to the increased distance between **Audio8** and **Audio5** in our two environments, i.e. 1.91 m in the closed-door environment versus 3.33 m in the open-concept environment. The corresponding bit rates for $N = 5$ and $N = 11$ were 140 bps and 189 bps respectively. The results of our tests using the $W = 20$ kHz to 21 kHz bandwidth produced BERs above 20 % and 25 % for all machines in the closed-door and open-concept environments, respectively. This leads us to conclude that receiving ultrasonic signals above 20.5 kHz is not generally well supported in commodity hardware at the distances we tested.

Although not shown, we were able to communicate using the combined near-ultrasonic and ultrasonic ranges (i.e. 18 kHz and above) using $N = 17$ sub-channels for an effective bit rate of over 500 bps and BERs of 10 % and 15 % in the closed-door office and the open-concept office environments, respectively. As a comparison, previous researchers were only able to achieve a bit rate of 20 bps with a BER of 0 % using the same bandwidth. Furthermore, in our tests using the audible spectrum (i.e. $W = 0$ Hz to 20.5 kHz), we achieved BERs below 10 % and 15 % for the closed-door and open-concept environments respectively, with the exception again being **Audio5**, which achieved a BER of 19 % in the open-concept tests. We used $N = 812$ to achieve these results, which gave us a bit rate of over 6.7 kbps. We were also able to transmit data at a rate over 8.7 kbps using $W = 500$ Hz to 20.5 kHz and $N = 1857$ to achieve BERs below 25 % and 30 % for all machines in the closed-door and open-concept office experiments, respectively.

In order to reduce our error rates to acceptable levels, an $[n, k, d]$ block code (where n is the block length, k is the message length, and d is the distance), capable of correcting $\lfloor \frac{d-1}{2} \rfloor$ random errors, such as $[n, k, n - k + 1]$ Reed-Solomon Codes [21], could be used. To correct up to 10 % and 15 % bit errors using Reed-Solomon codes, an overheard of approximately 20 % and 30 %

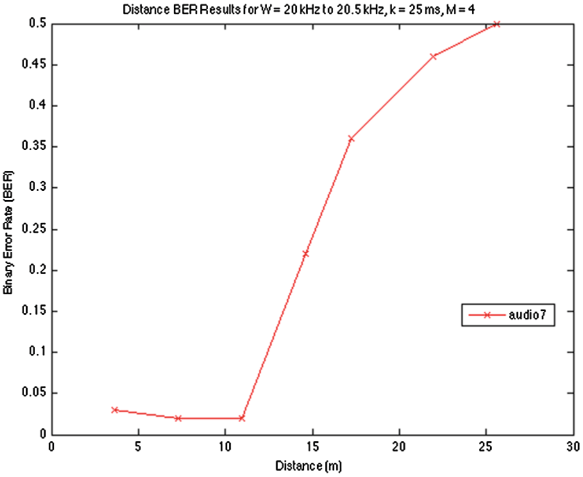

Fig. 3. BER for our distance experiment. We tested **Audio8**'s ability to communicate with **Audio7** over increasing distances. The parameters for the distance tests were $W = 20\,\text{kHz}$ to $20.5\,\text{kHz}$, $k = 25$ ms, $M = 4$, and $N = 46$.

is respectively required. Given our results, our bit rates after error correcting would then be reduced to 112 bps and 4.7 kbps using the ultrasonic and audible frequency ranges, respectively. To put these data rates into perspective, an individual typing 7-bit ASCII text at 80 words per minute and an average word length of 5.1 characters would produce data at an average rate of 47.6 bps. Similarly, voice can be streamed using the LPC-10 codec at 2.4 kbps [11]. Given these bit rates, our **overnight attack** can effectively leak both keystrokes and recorded audio data. Additionally, Table 2 provides average document sizes in kilobits (kb) and the amount of time that would be required to leak the document using our **overnight attack**. Lastly, both our attacks are effectively able to leak cryptographic key material such as a 256-bit shared key within seconds.

Using the parameters $W = 20\,\text{kHz}$ to $20.5\,\text{kHz}$, $k = 25$ ms, $M = 4$, and $N = 11$, we tested our ultrasonic attack with both a clock radio playing in the background as well as conversations taking place in the room while data was being leaked. We saw our resulting BER only increase marginally. This was due to the fact that while the human voice is predominantly composed of frequencies below 4 kHz there are still some near-ultrasonic and ultrasonic frequencies present which interfered with our data symbols. We noted that the majority of the energy in the clock radio and the conversations was in the 0 Hz to 15 kHz bandwidth and, in general, did not adversely affect our ultrasonic communication. Additionally, in order for our test to be performed in true ultrasonic fashion our pilot signal needed to be set to an ultrasonic frequency. In our experiments we used a pure-tone pilot signal at 20 kHz. Although this did work in most cases, some of the equipment, namely **Audio4**, **Audio5**, and **Audio6** had issues reliably picking up a 20 kHz pilot tone.

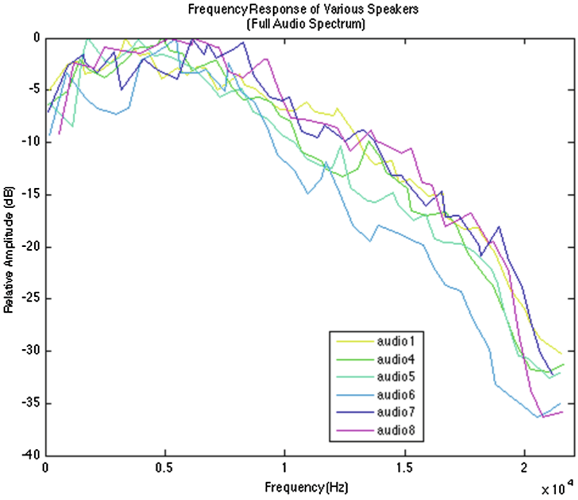

Fig. 4. The combined normalized frequency response of the audio channel. Each curve represents the frequency response of the channel with each of the machines in the study as the transmitter and **Audio8** as the receiver. Although the received signal strength in the ultrasonic range from 20 kHz to 20.5 kHz is relatively low, ultrasonic frequencies are still present.

Additionally, we performed a distance experiment to determine the maximum distance over which we could transmit ultrasonic signals as a means to compare our results to previous work. To determine our maximum transmission range we set up **Audio8** to transmit to our Sony laptop, **Audio7**, over increasing distances in 3.66 m (12') increments from 3.66 m to 25.60 m (84'). The results of our experiment are shown in Fig. 3. Our attack using ultrasonic communication is able to effectively communicate up to distances of 11 m with BERs under 2 % and up to a distance of 15 m with BERs around 20 %. With $W = 20$ kHz to 20.5 kHz, $k = 25$ ms and $N = 46$, we are able to communicate at over 230 bps. With error correcting we are therefore able to communicate at a distance of 15 m with an effective bit rate of 138 bps. As a comparison, the researchers in [8] were only able to communicate 20 bps up to a maximum distance of 8.2 m with a BER of 0 %.

Lastly, we performed an experiment to measure **Audio8**'s ability to receive ultrasonic communication from all the other machines in our study. In our experiment we placed each of the other systems, with the exception of **Audio2** and **Audio3**, 1 m in front of **Audio8** and played the same broadband white noise sample that we used to determine the frequency response of the channel in our previous experiment. We then plotted the combined frequency response of the other systems' speakers and **Audio8**'s microphone. The results are shown in Fig. 4. From the figure it is clear that the frequency response of the channel as observed by the microphone in **Audio8** is similar to the frequency response curves shown in Fig. 1. For the sake of brevity, we present this as evidence that the other systems in our study are capable of transmitting ultrasonic signals.

The results of our experiments show the following:

1. In general, malware can leak sensitive data using ultrasonic communication to systems on a low-security network at a bit rate of 140 bps (112 bps after error correction) with all machines observing a pre-error corrected BER of 10 % or under.
2. In general, malware can leak sensitive data using our **overnight attack** to systems on a low-security network at a bit rate of 6.7 kbps (4.7 kbps after error correction) with all machines observing a pre-error corrected BER of 15 % or under (the exception being **Audio5**, which experienced a BER of 19 %, at least 4 % higher than all other machines).
3. Our ultrasonic attack is not affected by either a clock radio playing music in the environment or conversations taking place while the communication is taking place.
4. Our ultrasonic attack is capable of leaking sensitive information at distances up to 11 m and bit rates over 230 bps. Furthermore, ultrasonic communication can be used to leak sensitive information at a distance of 15 m and a bit rate of 138 bps.
5. In general, all the systems in our study can produce ultrasonic signals in the $W = 20\,\text{kHz}$ to $22\,\text{kHz}$ range.

5 Protection Mechanisms

Our attack can be eliminated by physically removing any speakers from the high security machine. If removal is not possible, disabling the speaker in software is the next best option; however, disabling the speaker in software could be overwritten by malware installed on the machine and could therefore provide a false sense of assurance to users of high security machines that our attack is not possible. The most effective way to detect our attack is to passively monitor the audio channel for abnormally high peaks of energy, especially in areas of the spectrum where there are typically none (e.g. the ultrasonic frequency range above 20 kHz). An effective way to detect our attack would be to periodically perform the fast Fourier transform (FFT) of the ambient audio in the environment and examine the resulting frequency components for abnormally high peaks. By calculating the FFT and comparing the resulting energy levels in a given bandwidth to some baseline or by checking if the resulting energy is above a certain threshold, our attack could be detected. With this method however, the distance from the transmitter is of utmost importance. If the device scanning the ultrasonic spectrum is too far from the transmitter, these approaches will not work. Furthermore, while we have discussed this countermeasure for our ultrasonic attack, this methodology works similarly for our **overnight attack** with the spectrum analysis simply shifted down into the audible range. Lastly, we feel it is imperative that this analysis be done by a specialized hardware device that is less likely to be attacked by malware that could render it ineffective.

6 Conclusion

In this work, we studied the ability for malware to leak sensitive information to low-security machines using ultrasonic and audible audio communication. We measured the achievable bit rate and BERs when transmitting signals in the 20 kHz to 20.5 kHz range as well as the 0 Hz to 20.5 kHz range using commodity hardware from a number of major laptop and desktop vendors. Our study showed that, in general, data can be communicated using ultrasonic communication at a bit rate of 140 bps with BERs below 10 %. Furthermore, we showed that malware can leak information when nobody is around to hear it, using an attack we call the **overnight attack**, at a rate of 6.7 kbps with a BER below 15 %. Additionally, we showed that our ultrasonic attack is not affected by ambient conversations taking place at the same time as the communication, nor by a radio playing a local radio station. Lastly, we showed that data can be communicated using ultrasonic communication across distances up to 11 m and at bit rates over 230 bps and a BER of 2 %. Given the achievable bit rates, our attack is able to leak captured keystrokes in real-time using the ultrasonic attack and both keystrokes and recorded audio using the **overnight attack**. A passive detection strategy was also presented.

References

1. File sizes and types (2014). http://help.netdocuments.com/file-sizes/
2. Baken, R.J., Orlikoff, R.F.: Clinical Measurement of Speech and Voice. Cengage Learning, Clifton Park (2000)
3. Domingues, N., Lacerda, J., Aguiar, P.M., Lopes, C.V.: Aerial communications using piano, clarinet, and bells. In: 2002 IEEE Workshop on Multimedia Signal Processing, pp. 460–463. IEEE (2002)
4. Ellison, R.J., Goodenough, J.B., Weinstock, C.B., Woody, C.: Evaluating and mitigating software supply chain security risks. Technical report, DTIC Document (2010)
5. Gerasimov, V., Bender, W.: Things that talk: using sound for device-to-device and device-to-human communication. IBM Syst. J. **39**(3.4), 530–546 (2000)
6. Goldman, A., Apuzzo, M.: How bin Laden emailed without being detected (2011). http://www.nbcnews.com/id/43011358/
7. Hanspach, M., Goetz, M.: On covert acoustical mesh networks in air. J. Commun. **8**(11), 758–767 (2013)
8. Hanspach, M., Goetz, M.: Recent developments in covert acoustical communications. In: Sicherheit, pp. 243–254 (2014)
9. Kinsler, L.E., Frey, A.R., Coppens, A.B., Sanders, J.V.: Fundamentals of Acoustics, 4th edn., p. 560. Wiley-VCH, December 1999. ISBN: 0-471-84789-5
10. Landström, U.: Noise and fatigue in working environments. Environ. Int. **16**(4), 471–476 (1990)
11. Lee, K.S., Cox, R.V.: A very low bit rate speech coder based on a recognition/synthesis paradigm. IEEE Trans. Speech Audio Process. **9**(5), 482–491 (2001)
12. Lindqvist, U., Jonsson, E.: A map of security risks associated with using COTS. Computer **31**(6), 60–66 (1998)

13. Lopes, C.V., Aguiar, P.M.: Aerial acoustic communications. In: 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics, pp. 219–222. IEEE (2001)
14. Lopes, C.V., Aguiar, P.M.: Acoustic modems for ubiquitous computing. IEEE Pervasive Comput. **2**(3), 62–71 (2003)
15. Lopes, C.V., Aguiar, P.M.: Alternatives to speech in low bit rate communication systems. arXiv preprint. arXiv:1010.3951 (2010)
16. Madhavapeddy, A., Scott, D., Sharp, R.: Context-aware computing with sound. In: Dey, A.K., Schmidt, A., McCarthy, J.F. (eds.) UbiComp 2003. LNCS, vol. 2864, pp. 315–332. Springer, Heidelberg (2003)
17. Madhavapeddy, A., Sharp, R., Scott, D., Tse, A.: Audio networking: the forgotten wireless technology. IEEE Pervasive Comput. **4**(3), 55–60 (2005)
18. Nandakumar, R., Chintalapudi, K.K., Padmanabhan, V., Venkatesan, R.: Dhwani: secure peer-to-peer acoustic NFC. In: Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM, pp. 63–74. ACM (2013)
19. O'Malley, S.J., Choo, K.K.R.: Bridging the air gap: inaudible data exfiltration by insiders. In: 20th Americas Conference on Information Systems (AMCIS 2014), pp. 7–10 (2014)
20. Proakis, J.G.: Digital Communications. McGraw-Hill, New York (2008)
21. Reed, I.S., Solomon, G.: Polynomial codes over certain finite fields. J. Soc. Ind. Appl. Math. **8**(2), 300–304 (1960)
22. Sanger, D.E.: Obama order sped up wave of cyberattacks against Iran. The New York Times 1, 2012 (2012)
23. Schneier, B.: Air Gaps (2013). http://aiweb.techfak.uni-bielefeld.de/content/bworld-robot-control-software://www.schneier.com/blog/archives/2013/10/air_gaps.html?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+feedburner%2FbDnSB+(Schneier+on+Security)
24. Stallings, W.: Network Security Essentials: Applications and Standards. Pearson Education, India (2007)
25. Szor, P.: The Art of Computer Virus Research and Defense. Pearson Education, Indianapolis (2005)
26. Tempest, W.: The Noise Handbook. Academic Press, New York (1985)
27. Zetter, K.: FAA: Boeings new 787 may be vulnerable to hacker attack (2008). http://www.wired.com/politics/security/news/2008/01/dreamlinersecurity