Various threat models to circumvent air-gapped systems for preventing network attack

Eunchong Lee, Hyunsoo Kim, and Ji Won Yoon

Center for Information Security Technologies (CIST)
Korea University, Republic of Korea
{gr4ce,aitch25,jiwon_yoon}@korea.ac.kr

Abstract. In order to prevent incidents related with information leakage, many enterprises and organizations have installed an air-gapped system. The system is used for separating their own network from a public network such as the Internet. However, researchers have demonstrated possibilities that the air-gapped system can be inactivated by attackers, especially about their advanced attacks with various covert channels. In this paper, we analyzed how much the information could be leaked via the covert channel. We conducted experiments about data communication between a speaker and a microphone which are regarded as a conventional acoustic covert channel. At the same time, we also had expanded the attack scenario into an environment without any microphone. That is, we tested whether the critical information could be leaked and transferred via two loud-speakers as a limited environment where the air-gapped system. Finally, it is shown that the speaker based covert network can be effectively expanded to centrally controlled embedded loudspeakers which have not been considered in a conventional acoustic covert channel.

Keywords: Air-gap malware, Malware communication, Acoustic covert channel communication

1 Introduction

There have been several cyber-terror incidents throughout the world. For instance, the centrifuge of nuclear facilities in Iran was damaged by Stuxnet in 2010 [1], and computer network at Nonghyup, Korean bank, was paralyzed in 2011 due to hacking attack [2]. The computer network of a major media corporate was also stopped in the Republic of Korea [3] in March 20, 2013. Furthermore, the critical information was leaked from Korea hydro & nuclear power company in 2014 [3]. From this history, governments and enterprises have applied network separation to prevent information leakage and to protect their systems from such threats.

Especially, Korean government and national intelligence service (NIS) have built the air-gapped system with network separation to avoid leakage of important information [4]. In February, 2012, business operators were required to

set network separation in their company if they have over one million personal information data or if their yearly sale is over 10 billion in Korean won[5].

By the effort of the government and the enterprises for resisting damages and reducing risks from leaking important information and maximizing their benefit, a network separation has been utilized in many organization such as Ministry of National Defense[6], financial computer system[6], nuclear power[7] and aviation[8]. In addition, the system has been applied to banks, public corporations, courts, and factories in order to build security system from 2009 to 2014. Several solutions exist to build an effective air-gap environment including Server Based Computing (SBC) and physical separation [9].

Although the government struggles to protect their important information by separating a network, there have been various threats to incapacitate the air-gapped system. Researchers have studied about this, they could derive meaningful results by introducing some methods which circumvent the air-gap security system. In this paper, a covert channel using voice grade sound is focused on, and various threat models via the covert channel are dealing with. Especially, we have demonstrated the possibility of communication between not only a speaker and a microphone but also a speaker to a speaker through this work. Manipulating Windows registry provides us to change the role of the speaker from an output device to an input device, we were also able to experiment in a data communication environment with a speaker to speaker.

However, Prior works mostly focused on the necessity of air-gap separation to prevent the leakage of important information. To make the system effectively, people must be aware of the fact that air-gap separation can be neutralized and the possibility of such detour around according to the solution. Throughout this research, various experiments were practiced by reproducing the air-gap separation circumvention attack via covert channels between a speaker and a microphone. In addition, we show the possibility to neutralize air-gap system and to gain information by only using speakers.

2 Related work

In 2013, Hanspach et al. used a 5-bit address system for data communication instead of TCP/IP stack in order to reduce an unnecessary overhead[10]. Through this, they demonstrated the possibility of communicating via a speaker and a microphone within $19.7m$ at a speed of 20bit/s. This network stack adapts an emulation system into the underwater acoustic network, which is composed of four parts, called APP, NET, EC, and PHY. The applied layer uses generic underwater application language (GUWAL), from FWG/FKIE to build appropriate communication for low bandwidth [14], and has 16-byte data frame, 2-byte header, 2-byte CRC checksum. The link layer uses 16-bit error verification code and ACS modem based on frequency hopping spread spectrum (FHSS), to build communication with strong resistance of error. In 2014, Guri et al. used audio frequency-shift keying and dual-tone multiple frequency modulation systems to build the physical layer and increased performance by applying DSP such as

adaptive noise filtering, equalization, and so on[11]. They could show that the communication between a computer and a cell-phone is possible by using FM signal emitted from video display unit within $1-7$m, at a speed of $13-60$bit/s. The research proved that we can make various secret channels for a covert channel. In addition, Guri et al. built the covert channel between two computers with a temperature sensor and transferred information within $0-40$cm, at a speed of $1-8$ bit/s. Despite the limitation of performance [12], the possibility of information leakage remains noticeable. Additionally, Carrara et al. reveals that utilizing OFDM communication mechanism between a speaker and a microphone is possible [13]. Their approach performed at a speed of 6.7k byte/s for overnight attack and 230 byte/s for the ultrasonic attack.

Generally, speakers are used as an output device. Nevertheless, Lee et al. (2013) show that the loud-speaker can be used as a microphone by simply modifying the operating system (OS) setting [15] and proved that the speaker can provide a certain level of signal to the attacker within 30 cm. There are a lot of similarities between the speaker and the microphone. Especially, they both have a diaphragm which is the key element for utilizing the speaker as an input device. The diaphragm is an important component for input and output of sound. It interacts with the robin and the magnetic field generated from the permanent magnet and creates a vibration from sound pressure. Whenever the speaker used as an input device, it may have different current direction and different trembling vibration due to the difference of material, yet can still play the role of a microphone [15].

Frequency hopping spread spectrum (FHSS) is one of the communication methods used in initial stage of a wireless network through 802.11 physical layer [16]. FHSS is widely used in war industry due to its resistance against jamming. This method sends the same signal through multiple frequencies to increase the reliability of signal transmission. FHSS has 23 channels and transfers data by random hopping sequence throughout the whole channels. Before the signal randomly hops to other channel, it scans a channel whether it has noise or interference of electric wave in order to find stable channel that helps to secure the data while it is transmitting or receiving [17]. In FHSS, whenever signal hops to other channel, it follows the rule of hopping code. It stays very short period of time in a frequency band and hops to another frequency band by the rule of hopping code and keep repeating this process [17].

Orthogonal frequency division multiplexing (OFDM) which is announced after FHSS and DSSS[1], is still actively developed and studied due to its adequacy for high speed data transmission, both through wired and wireless channels [18, 19]. The OFDM has high efficiency of frequency utilization because of multiple orthogonal carrier waves. The process of modulation and demodulation of those carrier waves in the transmitting and receiving stage can be realized at a

[1] direct-sequence spread spectrum (DSSS) is a spread spectrum modulation technique. Spread spectrum systems are such that they transmit the message bearing signals using a bandwidth that is in excess of the bandwidth that is actually needed by the message signal.

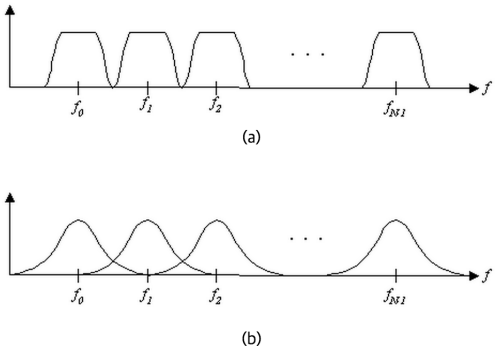

Fig. 1. Frequency Spectrum in OFDM

great rate. Since, the method suits for high-speed data transmission, it is chosen as a standard in IEEE 802.11a, high speed wireless LAN of HIPERLAN/2, Broadband Wireless Access of IEEE 802.16, Digital Audio Broadcasting, Digital Terrestrial Television Broadcasting, ADSL, and VDSL.

In the OFDM, the transmitting cycle increases in accordance with the a number of carrier waves increases. In 1966, the theory to send multiple carrier waves concurrently without interference between channels was introduced, which is shown above in Fig.1(a) [20]. Later, Orthogonal multiplexing QAM[2] was proposed to prevent interference between channels, which is shown above in Fig.1(b).

3 Threat model

In this paper, we have classified three different threat models for air-gapped system.

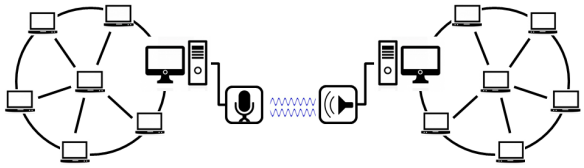

Fig. 2. Covert network between speaker and Microphone (Case 1)

[2] Quadrature amplitude modulation (QAM) is both an analog and a digital modulation scheme. It conveys two analog message signals, or two digital bit streams, by changing (modulating) the amplitudes of two carrier waves, using the amplitude-shift keying (ASK) digital modulation scheme or amplitude modulation (AM) analog modulation scheme.

– **Covert network between speaker and microphone (Case 1)**: For the first threat model, let us assume that there are two different devices A and B which are disconnected by the separated network. B in separated networks has been infected with malware. First, a device A (the sender) and a device B (receiver) in the air-gapped system have speakers and microphones respectively. Now, A can transfer the data to B using A's speaker and B' microphone. This is a basic and traditional model to disable the air-gapped system as shown in Fig. 2. This model is well-studied as described in section 2 but we also tested this for the comparison.

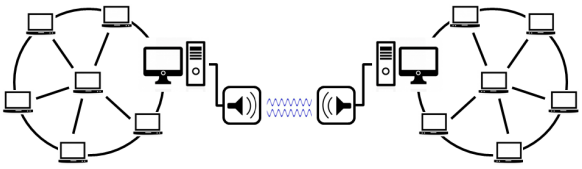

Fig. 3. Covert network between speaker and speaker (Case 2)

– **Covert network between speaker and speaker (Case 2)**: The case 1 is effective and well-studied, but its usage is rather limited. In many situations, targeted system (B) does not equip any microphones physically or the targeted system removes the microphones to avoid the case 1's attack. In this case, case 1 cannot be used since B cannot receive the data from B via the covert channel. For this case, we have made an extension of Case 1 with replacement of the microphone by a speaker for the receiving device. That is, the speaker can be used both for sending and receiving devices. The attacker can use this model since it provides more flexible attack by handling only speakers.

– **Covert network between the speaker and centrally controlled, embedded speaker (Case 3)**: Note that the microphones and speakers in case 1 and case 2 are connected to nearby terminals or devices. However, if we use speakers as a receiver, threat model can be simply expanded as well. For example, speakers can be separately installed and connected to other speakers over a shared line in the building and they are controlled by the central control center. In this case, there is no physically connected nearby terminals or devices. However, hackers can transfer the data to the system using shared loud-speakers which are equipped on the wall. In this case, the target terminal does not have to be nearby to the sender. For example, the attacker can transmit the data using the loudspeaker in toilet of the building if the speaker on toilet wall is controlled by central server and the central server is connected to the separated local network as shown in Fig. 4.

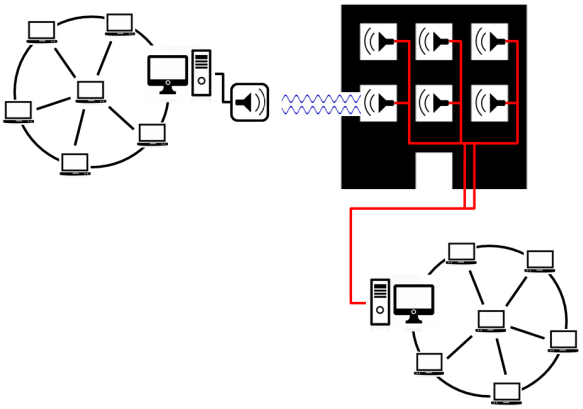

Fig. 4. Covert network between speaker and centrally control embed speaker (Case 3)

4 Technical background

This research proposes *Speaker Based Covert Network* (SBCN), a communication network based on speakers through the acoustic band. Furthermore, we compare SBCN with existing speaker and microphone based covert network to demonstrate the possibility to communicate with only speakers. For the communication method, OFDM [13] is used to communicate at the speed of 6.7kbit/s through acoustic frequency band and FHSS [10] is used to communicate at the speed of 20bit/s when the environment has not any noises. In addition, hamming code is applied to our method for minimizing the error and cyclic redundancy check (CRC) and verifying received data.

4.1 Collecting sound signal with redefining terminal pin number

According to the paper of Lee et al. (2013) [15], loud-speakers can be used as an input device, such as a microphone. To use a speaker as a microphone, modifying terminal setting about each definition is necessary. Especially in Windows, by changing Windows registry, terminal setting for the loud-speaker can be modified simply as a terminal for the microphone. In this scenario, a malicious user can utilize the speaker for the covert network communication after modifying terminal setting of Windows registry.

Table 1 shows the description of pin number and its function. This setting is used to conduct an experiment on Windows 7. Pin numbers “10” to “1b” denote the two front side and six rear sides I/O terminals.

Fig 5 shows the Windows registry modification process in order to use the speaker as a microphone. Each pin registry key has a designated hexadecimal value according to its function, and the sound device will act differently following the value changes. Table 2 matches each hexadecimal value and its function. In this experiment, hexadecimal code “04 00 00 00” is modified to “01 00 00 00”

Pin #	Function	Pin #	Function
Pin10	Line-in(rear blue)	Pin11	Mic-in(rear pink)
Pin14	Front speakers(rear green)	Pin15	Rear Speakers(rear black)
Pin16	Center / Sub-woofer(rear orange)	Pin17	Side Speakers(rear gray)
Pin19	Front Mic-In(rear pink)	Pin1b	Front Headphone(rear green)

Table 1. ALC882 codec based sound card pin assignment

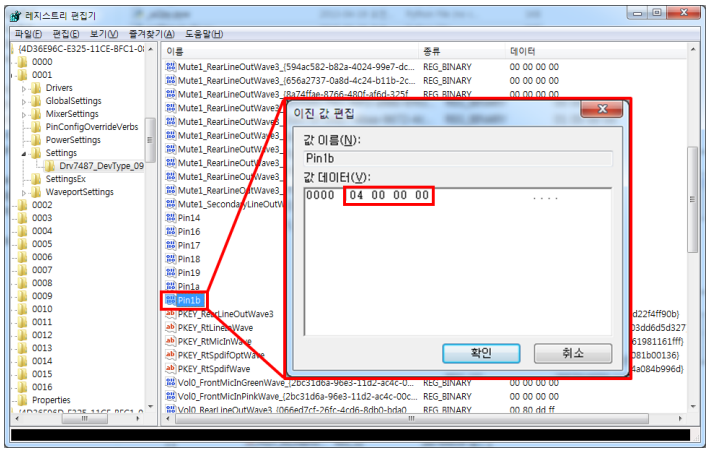

Fig. 5. Process for modifying Windows registry to use speaker

for loud-speaker communication. A malicious code with a permission of administrator can expose normal users at the threat by simply modifying this value. Thus, attackers may be able to control sound devices and force the computer to communicate regardless of air-gap separation.

Banary	Function	Banary	Function
00 00 00 00	Line-in	01 00 00 00	Mic-in
02 00 00 00	Headphones	03 00 00 00	Front speakers 2nd Output
04 00 00 00	Front speakers	05 00 00 00	Center / Sub-woofer
06 00 00 00	Side speakers		

Table 2. ALC882 Codec based sound card binary value assignment

4.2 OFDM communication at acoustic band

OFDM communication is conventionally used at a frequency over 1GHz for high-speed communication. However, it this paper, we used the frequency band range

of 3000-20000Hz, which are that of a speaker. The frequency band is divided into multiple bands of 1kHz. In addition, 2QAM, 4QAM and 16QAM modulation methods are applied. OFDM becomes more sensitive to noise as carrier waves in the frequency band increase, resulting in higher error rate. Thus, this paper focuses on finding the adequate balance between error and speed rate.

4.3 FHSS communication at acoustic band

FHSS was developed in the United States to satisfy the need for a new modulation method which is less influenced by jamming. FHSS is also used in 802.11 and CDMA communication in accordance with this advantage. Therefore, FHSS is applied to SBCN for preventing noise affection which is a major problem in speaker based communication. The frequency band is divided into a series of 3kHz, and acoustic bands are sorted into six channels of 3kHz, 6kHz, 9kHz, 12kHz, 15kHz, and 18kHz.

4.4 Verification of transmitting error and recovery code

When an error occurs during communication, retransmission is generally requested. However, forward error correction (FEC) is applied in situations where retransmissions are impossible, such as simplex transmission. In FEC, the sender encodes the message in a redundant way by using an error-correcting code. This redundancy allows the receiver to detect a limited number of errors and often attempt to correct them. In this work, we minimize errors that occur during communication by using hamming error correction code [21], which is generally used for FEC to verify and recover an error. Finally, transmitted data is verified by CRC [22].

5 Experiment design

5.1 Environment and system configuration

The major objective of this study is to establish a connection between an unapproved network and an external network through sound communication between speakers in an air-gapped system, measure the distance and the bit per second for communication through the relevant convert channel, and check whether there exists a possibility of private data leakage in an internal network, separated from the outside.

In order to verify the possibility of leakage in various environments, the maximum communication distance, the transmission success rate in different communication environments, the data transmission rate per second and the leakage time of each data capacity were experimented.

Lee et al. (2013) paper [15] employed an earphone and a headphone, which are passive speakers to be easily found around, to prove eavesdropping in a speaker-based environment. In this study, a passive speaker, mentioned above,

was used for the SBCN communication experiment, and a room-type experiment environment that was similar to an office, was constructed at a distance of 0.1 to 7m.

Prior to this experiment, performance of OFDM and FHSS communication was compared and analyzed. First, the OFDM algorithm was applied to SBCN communication, however it was not appropriate to use due to strong noise and interference of the speaker. The OFDM communication experiment was conducted using speaker and microphone, and FHSS, which is resistant to noise, was applied to only speaker communication.

6 Experiment and results

6.1 The Maximum communication distance of each frequency in only speaker environment and with microphone environment

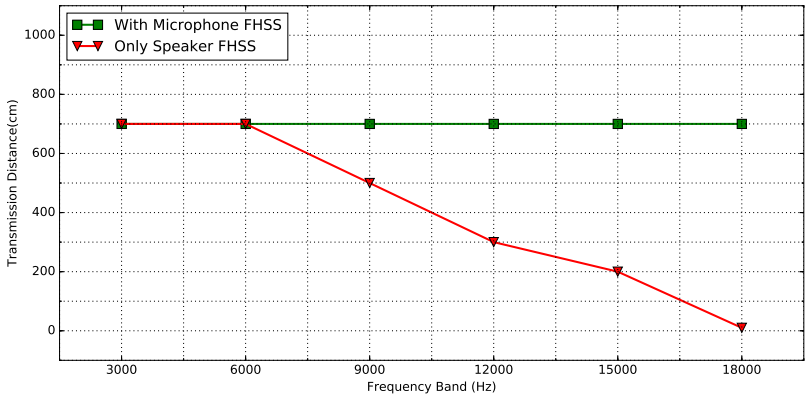

Fig. 6. Maximum Transmission Distance depend on Frequency Band

As shown in Fig. 6, the maximum communication distance experiment was performed at six frequencies ranging from 3kHz to 21kHz (3kHz-6kHz, 6kHz-9kHz, 9kHz-12kHz, 12kHz-15kHz, and 18kHz-21kHz) in the environment with only speaker and with microphone.

In the environment with microphone, communication at all the frequencies was available at a distance of 7m, which is the maximum distance of the experiment. However in the environment with only speaker, communication was available at a long distance, as the frequency became lower, and the communication distance was rapidly shortened at higher frequencies. A frequency of 3kHz to 6kHz had communication with no error at a distance of 7m, which is

the maximum distance of the experiment. In the experiment with an inaudible frequency of 18kHz to 21kHz, communication was available only at a short distance of 10cm. Though there might be slight differences depending on the performance of the speaker, in the inaudible domain of the only speaker environment, communication was available only under limited conditions of the short distance.

6.2 Comparison of the transmission success rate of each frequency in the environment with only speaker and with microphone

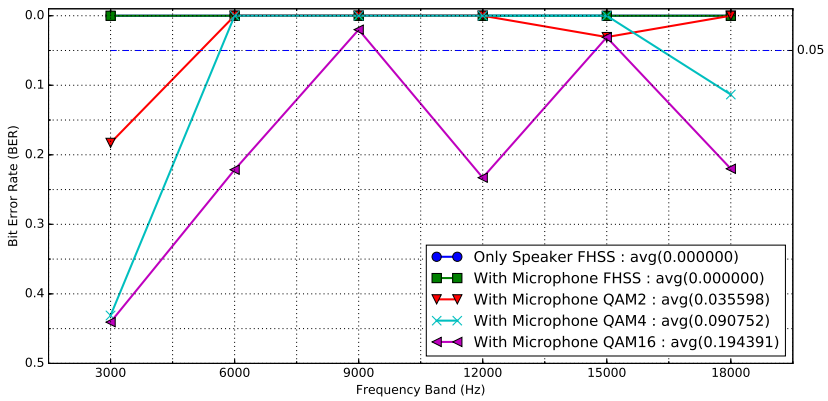

Fig. 7. Bit Error Rate on Frequency Band

As displayed in Fig. 7, the only speaker environment and the environment with microphone both showed a perfect transmission success rate in FHSS communication. In the only speaker environment of OFDM communication, transmission was impossible due to strong noise, and in the with microphone environment, generally, the success rate of data transmission fell, as carrier waves increased. The success rate declined overall at a frequency of 3kHz, used in human voice, and 2QAM and 4QAM showed smooth communication at all the frequencies except 3kHz, but 16QAM showed a markedly low performance at a frequency of 6kHz, 12kHz, and 18kHz.

6.3 Comparison of data transmission of each frequency in only speaker environment and with microphone environment

As indicated in Fig. 8, FHSS communication had a speed of 8bit/s, which was slower than OFDM communication. 2QAM of OFDM showed stable communication with a high data success rate like FHSS, but its transmission speed was

slower than that of 4QAM and 16QAM. Performance of 16QAM can fall to zero at specific frequencies (3kHz and 18kHz in the experiment) depending on the signal conditions, since this is sensitive to noise, compared to 2QAM and 4QAM, but in a nice signal environment, communication was available at a speed of 3.2kbit/s, which more than doubled the speed of 2QAM and 4QAM.

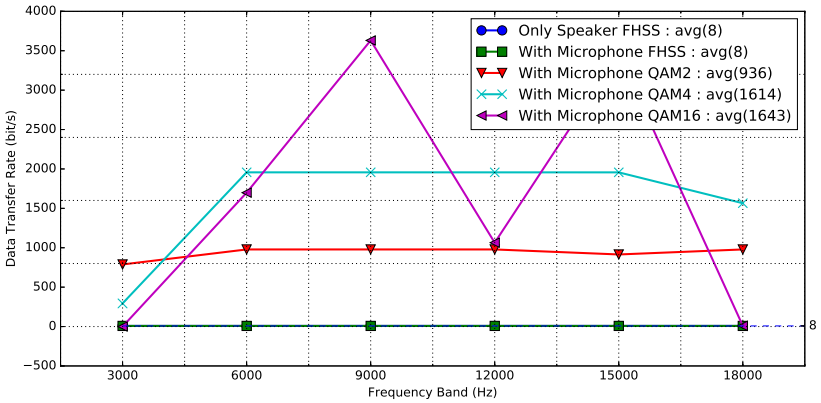

Fig. 8. Transfer Rate on Frequency Band

7 Discussion and conclusion

7.1 Discussion

The best way to defend the attack suggested in this paper is to physically remove the speaker. Alternatively, the speaker should be deactivated from the operating system. However, in this case, a malicious code may infect the system and reactivate the speaker. If the speaker cannot be removed, the most effective way to determine such attacks is to detect the presence of abnormal signals. Methods for detecting abnormal signals have already developed. If these methods are applied to the air-gap system, the system is able to protect attacks mentioned by this paper. Particularly, the covert channel which we built could be perceived by the technique using abnormal signal detection with fast Fourier transform(FFT). The covert channel uses a high-frequency sound for communication that is higher than voice grade. Let me show the example with Fig. 9. Basically, a range of voice grade is from $300Hz - 4kHz$. Therefore, we can assume that the sound from $300Hz - 4kHz$ would be diffused in the office during business hours. Consequently, if we surveil frequency bands which are higher than voice grade and check whether there are abnormal patterns of the signal, we can perceive abnormal signals and prevent the attack.

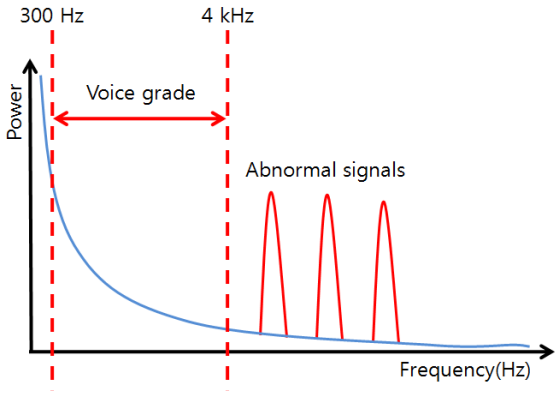

Fig. 9. Perceiving abnormal signals

7.2 Conclusion

To conclude, we have demonstrated various acoustic covert communications given three threat models in this paper: from a speaker to a microphone (case 1), from a speaker to another speaker (case 2) and from a speaker to a centrally controlled embedding loudspeaker on the wall(case 3). The maximum communication speed are 3.2 kbit/s when both the speaker and the microphone are attached while it is 8 bit/s when only the speakers are used. We focused on the various possibilities to circumvent air-gap system in different environments rather than their communication speed. With regard to mentioned threat models, we insist that air-gap system is not a completely secure method and at the same time, it also implies that an attacker is able to exfiltrate significant information from the separated network. Consequently, in order to protect the important information in a private network, system managers have to consider of this kind of threat models as well as network separation.

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (NRF-2013R1A1A1012797)

References

1. N. Falliere, L. O. Murchu, and E. Chien. W32. stuxnet dossier. *White paper, Symantec Corp., Security Response*, 5, 2011.
2. J. H. Eom. Cyber Defense Strategy for Information Superiority in Cyberspace. *Journal of Security Engineering*, 9(5), 2012.
3. J. H. Eom. Management Plan of Cyber Reserve Forces for Cyber Warfare. *Journal of Security Engineering*, 12(2), 2015.

4. Y. Jeong and K.-D. Nam. An Investigation of Network Separation Solution for Government Network. volume , pages 1125–1126. Korea Institute Of Communication Sciences, 2011.
5. W. J. Shim. A Study on Considerations for Effective Network Partition. *Soonsil University, Information Security 2015. 2*, 2015.
6. U. Lindqvist and E. Jonsson. A map of security risks associated with using COTS. *Computer*, 31(6):60–66, 1998.
7. D. E. Sanger. Obama order sped up wave of cyberattacks against iran. *The New York Times*, 1:2012, 2012.
8. K. Zetter. Faa: Boeing's new 787 may be vulnerable to hacker attack. *Wired (Jan 8) http://www. wired. com/politics/security/news/2008/01/dreamliner security.(accessed May 2009)*, 2008.
9. E. Lee and K. Kim. A Study on data protection based air-gap separate. *REVIEW OF KIISC*, 20:39–46, 2010.
10. M. Hanspach and M. Goetz. On covert acoustical mesh networks in air. *arXiv preprint arXiv:1406.1213*, 2014.
11. M. Guri, G. Kedma, A. Kachlon, and Y. Elovici. AirHopper: Bridging the air-gap between isolated networks and mobile phones using radio frequencies. In *Malicious and Unwanted Software: The Americas (MALWARE), 2014 9th International Conference on*, pages 58–67. IEEE, 2014.
12. M. Guri, M. Monitz, Y. Mirski, and Y. Elovici. Bitwhisper: Covert signaling channel between air-gapped computers using thermal manipulations. *arXiv preprint arXiv:1503.07919*, 2015.
13. B. Carrara and C. Adams. On Acoustic Covert Channels Between Air-Gapped Systems. In *Foundations and Practice of Security*, pages 3–16. Springer, 2014.
14. I. Nissen and M. Goetz. Generic under water application language (GUWAL)-Specification of tactical instant messaging in underwater networks,. *Research Department for Underwater Acoustics and Marine Geophysics*, 2012.
15. S. J. Lee, Y. M. Ha, H. J. Jo, and J. W. Yoon. The danger and vulnerability of eavesdropping by using loud-speakers. *Journal of the Korea Institute of Information Security and Cryptology*, 23:1157–1167, 2013.
16. B. P. Crow, I. Widjaja, J. G. Kim, and P. T. Sakai. IEEE 802.11 wireless local area networks. *Communications Magazine, IEEE*, 35(9):116–126, 1997.
17. R. C. Dixon. *Spread Spectrum Systems: With Commercial Applications*. John Wiley & Sons, Inc., New York, NY, USA, 3rd edition, 1994.
18. J. AC Bingham. Multicarrier modulation for data transmission: An idea whose time has come. *Communications Magazine, IEEE*, 28(5):5–14, 1990.
19. H. Sari, G. Karam, and I. Jeanclaude. Transmission techniques for digital terrestrial TV broadcasting. *IEEE communications magazine*, 33(2):100–109, 1995.
20. R. W. Chang. Synthesis of Band-Limited Orthogonal Signals for Multichannel Data Transmission. *Bell System Technical Journal*, 45(10):1775–1796, 1966.
21. R. W. Hamming. Error detecting and error correcting codes. *Bell System technical journal*, 29(2):147–160, 1950.
22. G. Castagnoli, S. Brauer, and M. Herrmann. Optimization of cyclic redundancy-check codes with 24 and 32 parity bits. *Communications, IEEE Transactions on*, 41(6):883–892, 1993.