Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.23, NO.6, Dec. 2013 ISSN 2288-2715(Online)
http://dx.doi.org/10.13089/JKIISC.2013.23.6.1157

스피커를 이용한 도청 위험에 대한 연구

이 승 준,† 하 영 목, 조 현 주, 윤 지 원‡
고려대학교

The danger and vulnerability of eavesdropping by using loud-speakers

Seung Joon Lee,† Young Mok Ha, Hyun Ju Jo, Ji Won Yoon‡
Center for Information Security Technologies (CIST), Korea University, Korea

요 약

최근 정보통신기술 발달과 함께 야기되는 불법적 도청으로 인한 기업의 정보유출 및 개인의 사생활 침해문제는 현대사회에서 중요한 문제로 인식되고 있다. 특히 스피커는 소리를 내는 장치임에도 불구하고, 공격자의 의도에 따라 마이크의 역할로서 작용 되어 도청 및 감청의 도구로 이용될 수 있다. 일반적으로 스피커를 도청장치로 이용할 수 있다는 인지가 적다는 점과 기존의 도청장치 탐지 장비로 쉽게 탐지하기 어렵다는 점은 도청도구로서 스피커를 더욱 위협적이게 한다. 이러한 점에서 악의를 갖는 공격자나 해커에 의해서 악용될 소지가 있다. 따라서 스피커를 이용한 도청 피해를 최소화하기 위해서는, 스피커를 이용한 도청 위험성의 인지 및 예상되는 도청 시나리오에 대한 연구가 요구된다. 본 논문에서는 스피커를 이용한 도청방법과 실험을 통해 스피커가 음성수집 도구로 이용할 수 있다는 취약점과 그 위험성을 보이고자 한다.

ABSTRACT

The development of electronic devices has recently led to many problems such as personal information rape and leakage of business information. Conventional loud-speakers have been generally used to output devices. It can be, however, operated as a micro-phone which was abused as a means for eavesdropping since the speaker and microphone have basically the equivalent structure. Most importantly, the general peoples are not aware of the approaching danger about using speaker as microphone. And, traditional eavesdropping detection equipment does not check the attack. In this paper, we demonstrate that there is a serious danger and vulnerability in using loud-speakers since they can be used as eavesdropping devices.

Keywords: Eavesdropping, Tapping, Bugging, Speaker

I. 서 론

정보통신기술의 발달로 생활이 점점 편해지고 있지만 그와 더불어 정보 유출로 인한 위험 또한 증가하고 있다. 최근 사회 문제로 거론되는 IT 기기 기반 정보 유출은 다양한 과학기술의 발달과 함께 점차 확대되고 있다. 공격자들은 불법적으로 정보를 입수하는 고전적 방법 중 하나인 음성 도청을 정보통신기술과 결합하여 개인의 정보와 산업정보를 탈취하기 위해 사용하고 있다.

사회가 도청을 사회적 문제로 생각하기 시작한 것은 어제 오늘 일이 아니다. 일례로, 지난 2011년 4월 발생한 사상 초유의 농협 해킹 사고의 도청 사례를 들 수 있다. 범인들은 문제의 노트북에 악성코드와 함께 백도어(Backdoor) 도청 프로그램을 설치하였다. 후에 그들은 업체 보안담당자의 언행을 감시함으로써 공격대상의 IP와 루트계정 비밀번호를 습득 및 공격하였고, 이는 사회적 문제를 일으켰다.[1]

또한 올해 4월, 스마트기기가 마이크를 기본으로 내장하고 있는 점을 이용하여 불법 어플리케이션 설치를 통해 동의 없이 상대방의 음성신호를 불법 도청하는

접수일(2013년 8월 14일), 수정일(2013년 10월 4일), 게재확정일(2013년 10월 8일)
† 주저자, seungjoonlee@korea.ac.kr
‡ 교신저자, jiwon_yoon@korea.ac.kr(Corresponding author)

사건이 있었다.[2] 이밖에 노트북 해킹을 이용한 음성 도청을 통해 개인정보를 유출시키는 것은 물론 기업의 회의를 도청하는 등의 사건이 있었고, 이러한 도청 문제는 사회적 문제로 더욱 대두되고 있는 추세이다.

저자는 본 논문에서 일반적으로 도청 도구로 생각하는 마이크가 아닌, 음원 출력 장치인 스피커를 이용한 도청 방법을 제안하고자 한다. 스피커는 일반적으로 소리를 내는 장치이지만, 간단한 조작을 통해 공격자의 의도에 따라 마이크로 작동할 수 있다.[3] 기존 도청 장비들은 추가적인 장비의 설치를 통해 마이크를 통한 도청은 복잡하거나 값비싼 비용이 드는 조작없이 사용자가 쉽게 도청 장치로 이용할 수 있다. 제안하는 공격 방법은 주변에서 흔히 찾을 수 있는 스피커를 이용하며, 방출하는 에너지가 스피커와 비슷하기 때문에 기존의 도청 탐지 장비로 탐지가 불가능하다. 일반적으로 마이크는 경계하지만 스피커를 경계하지는 않는다는 심리를 고려한다면 스피커는 기존 도청 장치의 좋은 대안이라고 할 수 있다.

본 논문에서는 스피커를 이용한 도청 예상 시나리오 및 도청 기술을 소개하고, 음성신호처리 기술을 적용함으로써 도청 방지 기법이 적용된 상황에서 스피커를 이용한 도청이 공격 방법으로 충분한 이용될 수 있다는 점을 보이고자 한다. 논문 II 장에서는 제안하는 연구를 이해하는데 필요한 기초적 내용을 기술하였다. 스피커와 마이크의 구조, 그리고 스피커로부터 얻어낸 음성 신호를 이용한 공격의 효과를 높이기 위해 사용된 은닉음원분리및(Blind Source Separation, BSS) 독립 성분 분석(Independent Component Analysis, ICA)에 대한 간단한 설명을 기술하였다. Ⅲ,Ⅳ장에서는 예상되는 스피크 도청 시나리오 및 기술에 대하여 서술하였다. 음성 신호 분석을 이용한 키보드 해킹에 있어서 음원 신호 수집을 위해 스피커가 음원 입력 단자 역할을 충분히 할 수 있음을 제안하였다. V장에서는 스피커를 이용한 도청 방법에 대한 실험 결과를 기술한 후 VI장 결론과 함께 이 논문을 마무리 한다.

II. 관련연구

2.1 스피커와 마이크의 구조

그림 1은 일반적인 마이크의 간단한 구조를 보여준다. 음성 신호가 마이크 내부 얇은 판막에 음압을 가하면 진동이 일어나 보빈(Bobig)이 자극(자석)사이

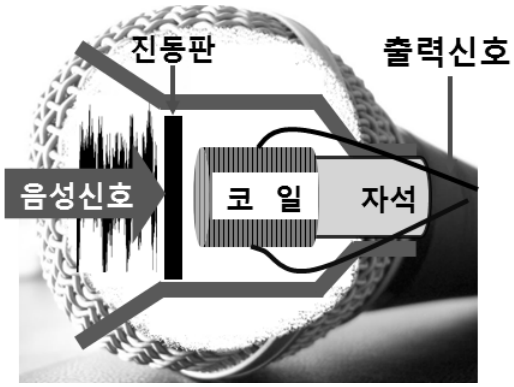

Fig.1. A moving-coil microphone structure

를 전후 왕복 운동한다. 마이크는 음압을 일으킨 음성 신호를 전기적 신호로 바꿔 해당 음성 신호를 수집한다. 스피커 역시 마이크와 그 구조가 유사하다 [4]. 주위에서 가장 흔하게 찾아 볼 수 있는 스피커는 일반적으로 그림 2의 무빙코일(Moving coil)구조의 스피커다. 특수한 형태의 스피커로는 정전형(Electrostatic) 스피커와 리본(Ribbon) 스피커, 그리고 흔히 부저라고 불리는 압전형(Piezoelectric) 스피커 등이 있다. 본 논문에서 사용한 스피커는 가장 보편적으로 사용되는 무빙코일 구조를 가진다. 무빙코일 스피커의 구조는 그림 2와 같다. 스피커는 마이크에서 볼 수 없는 부품을 가지고 있는데, 진동판 가운데 부분에 있는 먼지 커버라는 뚜껑이 바로 그것이다.

스피커와 마이크의 형태는 그림 1과 그림 2에서 볼 수 있듯 유사한 점이 많다. 그 중 공격자가 음원 입력 단자로서 스피커를 사용할 수 있게 하는 중요 요소는 진동판이다. 진동판은 영구자석에서 나오는 자기장과 보빈의 상호작용에 의해 직접 소리를 재생하거나 받아

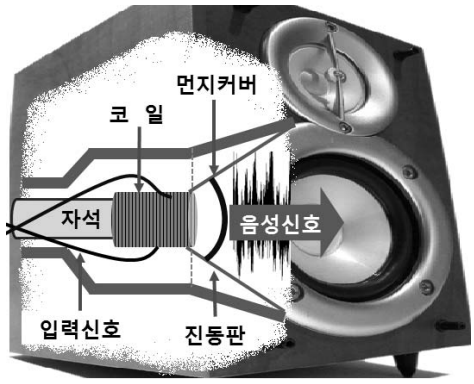

Fig.2. A moving-coil speakerphone structure

들이는 역할을 하는 중요한 부품이다. 이 진동판은 스피커와 마이크에서 공통적으로 살펴볼 수 있는데, 앞에서 마이크의 구조를 설명 하였듯이 음압으로 인한 떨림을 만드는 중요한 역할을 한다. 같은 스피커라도 진동판막이 더 민감하게 반응 할수록, 마이크의 형태와 유사 할수록 음원 입력 장치 역할을 효과적으로 한다.

이와 같이 스피커는 구성 부품 및 구조가 유사하며 음원 입력 장치 역할을 할 수 있는 조건을 가지고 있다. 스피커를 음원 입력 장치로 사용할 경우 전류의 방향과 재질에 따른 진동판막의 떨림이 다를 뿐 마이크와 같은 역할을 한다. 도청용 스피커는 음성신호 출력에 최적화한 재질 및 형상을 지니는 것이 좋고, 설계할 때 지향성 및 분할진동의 발생요소를 고려하는 것이 좋다.

2.2 마이크로서 역할을 가능케 하는 스피커의 조건

스피커는 구조적으로 마이크와 유사하기 때문에 입력 단에 직접적으로 연결하여 마이크 역할을 할 수 있다. 하지만 스피커를 마이크로서 효과적으로 활용하기 위해선 스피커의 자체의 성향을 좌우하는 진동판의 소재별 특성과 노화 상태, 스피커의 구조를 고려해야한다.

진동판의 특성을 결정짓는 여러 요소들 중 가장 고려해야할 요소들은 두께, 질량, 탄성률(modulus), 내부손실(Internal Loss)이다. 내부손실률은 어떤 소재에 에너지가 입력되었을 때 입력된 에너지의 일부가 전달되지 않고 내부적으로 에너지가 손실되는 정도를 말한다. 두께가 얇고 가벼울수록 진동운동이 원활하며 내부손실률이 낮아지며, 에너지가 잘 전달돼 코일의 움직임을 쉽게 만들어낸다. 즉, 내부손실률이 낮을수록 스피커는 신호 입력 단자의 역할을 충분히 할 수 있다는 것이다.

진동판의 소재는 스피커의 내부 손실률을 결정하는 요인이다. 천연펄프, 수지, 메탈계를 소재로 하는 진동판의 경우 스피커의 내부 손실률이 높고, 합성섬유 진동판의 경우 내부 손실률이 낮다. 소재에 따른 진동판의 이러한 특성은 특정 스피커에 의한 공격 효과가 좋지 않을 수도 있다는 것을 말한다. 때문에 안정적인 공격을 위해선 추가적인 하드웨어 조작이 필요하며, 그림 2와 같이 캐패시터, 오디오 트랜스포머 및 정교하게 설계된 증폭회로의 설치가 그것이다.

완제품의 스피커가 기본 스피커 유닛만으로 구성되어 있는지, 증폭 내장형의 스피커인지 파악하는 것 은 스피커를 이용한 도청 가능성 유무를 파악하는데 중요한 요소 중 하나이다. 일반적으로 완제품 스피커는 크게 두 종류로 분류 될 수 있는데 첫째는 스피커 형태가 내부 증폭회로가 없는 형태인 수동 스피커이며, 나머지 하나는 내부 증폭회로가 존재하는 형태를 가지는 능동 스피커이다. 능동 스피커의 경우 앰프사용으로 인해 전류가 출력 단으로만 흐르게 되어 있기에, 단선 후 직접 연결을 이루지 않는 한 도청용도로 신호를 받아들이는 것이 불가능하다. 수동형 스피커의 경우 주변에서 쉽게 찾아 볼 수 있는데 PC에 연결되어 있는 이어폰 헤드폰이 우리 주변에서 가장 쉽게 찾아볼 수 있는 수동형 스피커이다. 해당 스피커는 입력 단으로 연결하면 바로 마이크용도로 사용 할 수 있다. 이 밖에 주택에서 사용되어지는 방송용 장치(Public Address System)에 있는, 공동 신호선과 연결한 스피커 또한 그 예 중 하나이다.

2.3 은닉 음원 분리 (Blind Source Separation)

여러 지점에서 송출하는 신호를 혼합한 혼합신호를 마이크와 같은 장치로 받아들일 때, 각 송출 신호들이 어떻게 혼합되었는지 정보를 모르는 경우가 많다. 이 때 공격자는 혼합에 대한 정보가 전혀 없는 상태에서 혼합되기 전의 신호를 분리해야 한다는 과제를 안게 되며, 문제를 풀기 위해 음원분리 기법을 사용해야 한다. 이 때 사용하는 기술을 블라인드 음원분리(Blind Source Separation)이라한다. 저자는 도청방지 기법 중 하나인 사운드 마스킹 기법이 적용된 환경을 가정하고 문제를 풀기 위해 간단한 블라인드 음원 분리 기법을 사용하였다.

(1)은 m 개의 음원에서 나온 신호가 서로 혼합되어 n 개의 혼합된 신호를 형성한다는 것을 나타낸다.

$$S_n = \left(s_n^{(1)}, s_n^{(2)}, s_n^{(3)}, \ldots, s_n^{(m)}\right) \quad (1)$$

(2)는 음원 신호와 소스가 각각 2개라고 했을 때, 신호 혼합과 관련된 소스의 특성 행렬을 나타낸다. (3)은 혼합 신호의 성분과 음원 간 선형 관계를 나타낸다. 음원과 혼합 신호가 이러한 선형 관계를 가지고 있을 때, 공격자는 혼합된 신호로부터 음원 신호를 특성 행렬의 역행렬(A^{-1})을 통해 알아낼 수 있다.

$$A = \begin{bmatrix} a_{11} a_{12} \\ a_{21} a_{22} \end{bmatrix} \quad (2)$$

$$x_1 = a_{11}s^{(1)} + a_{12}s^{(2)} + \epsilon_1$$
$$x_2 = a_{21}s^{(1)} + a_{22}s^{(2)} + \epsilon_2 \quad (3)$$

하지만, 공격자가 관측할 수 있는 신호는 x뿐이며, 문제를 풀기 위해 공격자는 S의 각 요소들이 독립적일 것이라는 가정을 할 수 있다. 해당 가정과 함께 공격자는 각 신호의 독립성을 최대한 높이는 방향으로 행렬 A^{-1}를 추정할 수 있으며 관측된 신호 x_1, x_2로부터 음원 신호 S 를 추정 할 수 있다. 이 방법을 독립 성분 분석 (Independent Component Analysis) 라고 한다.(

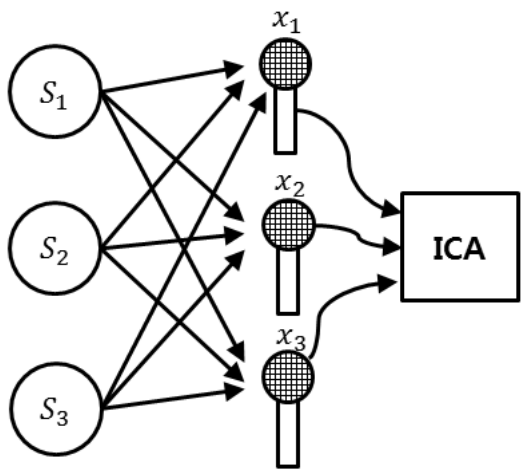

Fig.3. Diagram of ICA

2.4 독립 성분 분석 (Independent Component Analysis)

앞서 언급했던 바와 같이, 본 논문에서는 서로 독립적인 여러 신호가 섞인 상태에서 각 신호를 추정하는 독립 성분 분석 알고리즘을 구현하였다. 중심극한정리 (Central Limit Theorem)[6]에 의하면 서로 독립 신호들을 하나의 신호로 혼합할 때 그 신호들의 개수가 많을수록 혼합 신호의 분포는 점점 더 가우스 함수 분포에 가까운 형태를 가진다. 이에 따라 스피커나 마이크로폰을 통해 흘러나오는 신호들을 혼합한 신호는 원래 신호들보다 정규 분포를 따른다고 가정할 수 있으며, 이러한 특성은 여러 개의 소스를 사용하여 혼합한 신호를 분석하는데 독립 성분 분석 방법이 적절할 것이라는 기대를 준다. [8]

독립 성분 분석 기법에 대한 자세한 내용은 본 논문의 연구주제에 벗어남으로 자세한 기술적인 내용은 참고문헌을 통하여 참조하였다.

2.5 마이크를 이용한 키 누름 소리 탐지

2005년 네이쳐 뉴스에 타이핑 소리 분석을 이용한 키보드 해킹에 관한 연구 내용이 소개 된 적 있다. 해당 연구를 진행한 캘리포니아 버클리대학교 연구진은 10$ 정도의 마이크를 이용하여 컴퓨터 키보드 타이핑 소리를 수집 및 분석하였다. 연구진은 96%의 정확도로 해당 키보드를 통해 무슨 내용을 입력하고 있는지 하는 지 알아낼 수 있는 소프트웨어를 발표하였고, 보안전문가들에게 비상한 관심을 불러일으켰다[9].

키 타격음을 이용한 입력된 문자의 재구성 작업은 키보드 키를 눌렀을 때와 눌린 키가 원상태로 돌아올 때의 에너지 레벨에 초점을 둔다. 기존 연구에서는 각 단계의 에너지 임계값을 설정하여 그 값에 따라 키를 감지하는 방식의 방법론이 제시되었다.[10][11]

그림 4는 각 단계별 임계값에 대한 예시이다. 해당 논문에서는 타격 음을 수집하기 위해 사용된 $10 정도의 PC 마이크, 사운드 카드의 종류 및 주의해야 할 사항 등에 대해 기술되어 있다. 하지만 실제 실험 및 적용에 있어 환경에 따라 중요하게 작용할 수 있는 마이크 증폭 및 입력 레벨 조건은 기술되어있지 않다.

본 논문에서는 마이크의 증폭 기능을 사용하는 상황을 가정하여 기존에 제안된 공격모델이 공격 도구로 마이크를 사용할 때보다 효과적일 수 있음을 실험 결과를 통해 보인다.

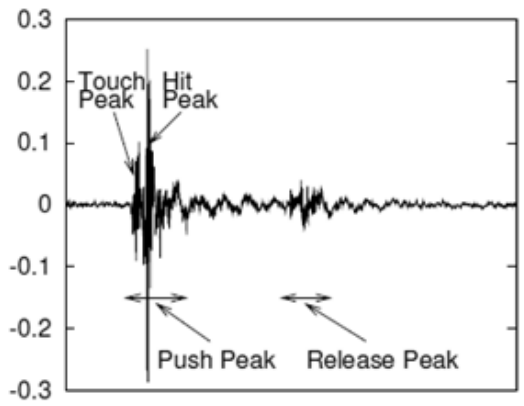

Fig.4. Acoustic signal of keystroke

III. 운영체제에 기반 한 전자기기의 스피커를 이용한 음성도청

3.1 스피커 단자 재 정의를 이용한 음성신호 수집

부분의 소리와 관련되어 입출력 역할을 하는 사운드관련 주변 장치들은 입출력포트들의 역할 핸들링 소프트웨어 조작을 통해 재정의가 가능하다. 스피커 단자를 입력단자로 변경시키면 해당 스피커는 마이크로의 역할을 수행하게 되고, 공격자는 이를 이용해 스피커로 음성 정보를 수집하는 것이 가능하다.

Table 1. ALC882 Codec based sound card pin assignment

핀 번호	기능
Pin14	라인 인 (뒤쪽 파랑)
Pin15	뒤쪽 스피커단자(뒤쪽 초록)
Pin16	뒤쪽 마이크단자(뒤쪽 분홍)
Pin18	5채널용 스피커 단자(뒤쪽 검정)
Pin19	5채널용 스피커 단자(뒤쪽 자주)
Pin1a	마이크단자(앞쪽 분홍)
Pin1b	헤드폰 스피커단자(앞쪽 초록)

표 1은 본 연구에서 수행한 환경인 윈도우7 운영체제에서 사운드장치의 입출력단자 핀 번호와 그 값을 보여준다. 표에서 상단 핀 '14'번 부터 핀 '1b'은 컴퓨터 후면패널 5개의 입출력 단자와 전면 패널 2개의단자의 번호를 의미한다. 다음 표는 본 연구에서 사용된 'ALC882 '코덱기반 사운드장치에 기본으로 할당된 각 핀의 기능이다.[12]

그림 5에서 볼 수 있듯 각 핀 번호는 고유의 기능에 따라 지정된 이진 값이 설정되어 있다. 해당 값의 수정은 사운드 장치의 기능을 바꿀 수 있다. 악성코드가 관리자 권한을 획득한다면, 소프트웨어로 정의되어 있는 핀의 기능을 레지스트리 값 수정만으로도 간단하게 스피커를 마이크로 사용할 수 있다. 즉, 악성코드를 설치하여 스피커가 입력 장치의 기능을 하도록 핀의 값을 재 할당 한다면, 소프트웨어 접근만으로도 스피커를 도청 장치로 이용할 수 있다.

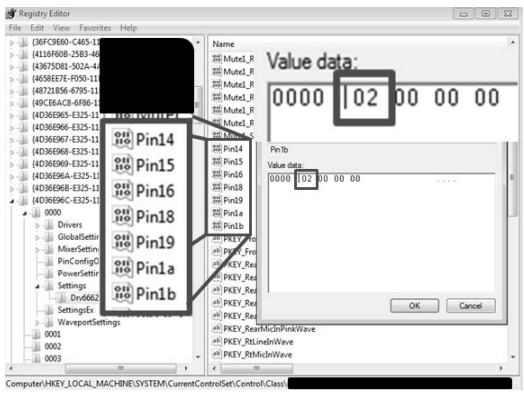

Fig.5. Redefinition of sound card pin assignment using system registry modification

저자는 출력포트를 입력포트로 재 정의하고 스피커로부터 보다 원만한 소리를 수집하기 위해 사운드장치 내부 증폭을 임의로 변경할 수 있는 소프트웨어 설계를 완료하였다. 또한, 스피커의 도청 장치로서 충분한 성능을 낸다는 점을 부각하기 위해 아래 실험 결과 부분에서 음성신호 후처리를 통한 결과를 기술하였다.

IV. 장내방송설비(Public Address System) 종단 스피커를 이용한 도청

4.1 PA시스템 종단 스피커를 이용한 실내 도청 가능성

그림 6은 본 연구에서 수행한 스피커 다수로 연결된 공동신호선 도청실험 환경을 도식화한 것이다. 특히 이러한 실험 환경은 장내방송설비의 공동 신호선에서 쉽게 찾을 수 있으며 이는 다수의 스피커를 연결되어 각 방의 천정이나 벽면에 설치되어 있는 환경과 유사하다. 특히 이러한 스피커는 일반적으로 패시브(Passive) 형태이기 때문에 마이크와 같은 역할이 가능하다. 하지만 장내방송설비 스피커가 도청장치로

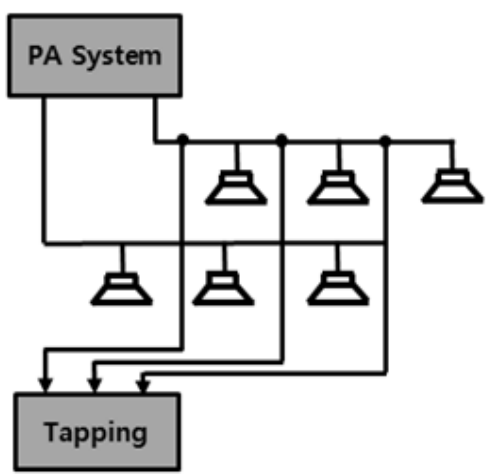

Fig.6. Schematic diagram of PA system tapping

서 마이크와 같이 충분한 전기신호를 수신하기에는 세 가지 제약이 따른다.

첫째, 장내방송설비에서 사용되는 일반적 스피커의 진동판은 다소 출력이 높은 신호를 송출하기 위해 두께가 두꺼운 천연펄프 재질이며 콘(cone)형이 일반적이다.[13] 따라서 진동판의 재질에 따른 내부손실률이 높고 형태적인 이유 때문에 받아들이는 신호가 미약하다. 둘째, 그림 6와 같이 하나의 공동 신호 선에 다수의 스피커가 직렬과 추가적 병렬로 연결되어 있다.[14][16] 셋째, 공동 신호선에 흐르는 고전압 신호를 수신하기 위해 임피던스 매칭용 트랜스포머가 각 스피커 직전에 연결되어 있음으로 스피커로 수집되는 전기 신호가 미약하다는 점이다.[15][16] 하지만 위에서 기술한 방송설비시스템에 연결되어 있는 스피커가 마이크특성을 가지기에는 좋지 않다는 점, 방송설비의 구조적인 이유에 따른 제약 조건은 증폭장비로 어느 정도 해결이 가능하다. 이때 스피커로부터 받아들인 음성신호가 충분히 앰프에 전달하여 주기 위해서는 스피커에서 증폭기로의 출력 임피던스가 증폭기 자체의 입력 임피던스보다 적어도 5배에서 10배 이상 작아야한다. 이처럼 작은 출력 임피던스를 큰 입력 임프던스에 연결하는 것을 브릿징(bridging)이라 하는데, [17] 본 연구에서는 높은 입력 임피던스를 갖는 앰프를 이용함에 따라 임피던스 브리징을 시켜줌으로서 미세한 전류 값의 변화를 잡아 주어 도청장치로의 역할을 가능 하게 할 수 있었다.

그림 7은 스피커로부터 받은 미약한 신호를 증폭시켜주는 간단한 증폭회로를 보여준다. 본 연구에서는 PA시스템에 높은 신호 증폭 장비를 추가하여 마이크와 같은 음성신호수집 도구로 스피커를 이용하는 것이 가능하다는 것을 확인하였다. 다수의 스피커를 직렬로 연결하는 방식은 한 도선에 연결한 모든 스피커로부터 수집한 신호로 혼합신호를 만들 수 있게 해준다. 위와 같은 공격 형태에서 원치 않는 배경잡음 및 시스템 자체의 특성이 야기한 잡음으로 인해 청음 시 음성신호의 명료성이 저하 될 수 있다. 때문에 공격의 실효성을 높이기 위해서는 SNR(신호대잡음비)를 높이기 위한 후처리 기술이 요구될 수 있다.

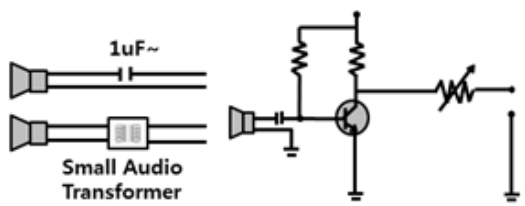

Fig.7. Simple amplifier circuit diagram for speaker as microphone

V. 실험결과

일반적으로 스피커는 이어폰, 음악 감상용 스피커를 비롯하여 거의 모든 제품들이 하나의 유닛으로 구성되어 있는 것이 아닌 쌍으로 구성되어 있다. 또한 공동 신호선에서와 같이 스피커가 배열로 연결되어 있는 경우도 볼 수 있다. 따라서 본 논문에서는 위와 같은 이유로 여러 개의 스피커로부터 소리를 수집할 수 있다면, 추가적인 음성신호처리를 이용하여 신호품질을 높이는 것이 가능하다는 점에 착안하여 연구를 수행하였다. 이러한 실험은 스피커를 이용한 도청의 충분한 가능성을 보여준다.

추가적으로 제안하는 기술이 도청 방지 기술이 적용된 상황에서도 사용될 수 있다는 것을 보이기 위해 도청 방지 기법이 적용된 환경을 가정하였다. 본 실험에서는 도청 목표 음성을 물소리, 백색잡음, 사이렌 소리 등을 잡음이 조성된 환경에 노출시켜 시행하였다. 해당 환경을 조성하기 위해 본 논문에서는 사운드 마스킹을 사용하였다. 이는 도청방지에 주로 사용되는 기술이다. 사운드 마스킹은 청자로 하여금 듣고자 하는 소리, 특히 사람의 말소리를 인식하지 못하게 하는 것으로, 사람이 상대적으로 큰 소리나 주파수가 낮은 소리를 선호하여 인식하는 것을 이용한다. 사운드 마스킹은 도청방지 기법이 적용된 환경에서 제안하는 공격 방법을 이용할 수 있다는 것을 보이기 위해 사용되었다.

본 실험에서는 도청 방지 기법이 적용된 환경에서 음성신호를 분리하여 보다 좋은 결과를 얻기 위해 신호처리 기법인 독립 성분 분석(ICA)을 이용하였다. 또한, 본 논문에서는 스피커를 이용한 도청이 마이크를 이용한 도청보다 효율적일 수 있음을 보이기 위해 키보드 타격소리를 이용한 키 해킹에 제안하는 기술을 적용해 보았다. 키보드 해킹에서 스피커를 이용한 키보드 타격 소리 수집은 마이크를 이용한 수집보다 더 효율적일 수 있다는 것을 실험결과를 통해 보였다.

본 공격모델의 평가를 위해 ALC882 코덱 기반 메인보드 내장형 사운드카드 환경에서 일반적으로 사용되는 $10 미만의 스피커 모듈을 이용하였다. 도청 실험에 사용된 목표 음성 신호(이하 원음)는 30대 일반인 여성 화자로부터 한국어로 발음된 짧은 단어로 이

루어졌으며, 1부터 10까지 숫자에 대한 해당 발음을 약 30cm 이내 근거리에서 녹음하여 사용하였다. 사운드 마스킹 신호의 경우 백색잡음과 사이렌소리를 매틀랩(Matlab)을 이용하여 직접 생성하였으며, 흐르는 물소리는 공용 수도꼭지로부터 얻었다. 본 논문에서는 같은 공간 안에 대역잡음을 줄 수 있는 사운드 마스킹 신호와 여성의 음성신호를 동시에 재생하여 혼합한 음원(이하 혼합 음원)을 대상으로 실험하였다.

도청을 위한 도구로써 스피커가 충분한 역할을 수행할 수 있는지를 알아보기 위해 스피커에 혼합 음원을 입력 하였다. 사람의 소리를 좀 더 잘 알아들을 수 있도록 입력받은 혼합 음원을 원음과 잡음으로 분리하였으며, 이를 위해 블라인드 음원 분리(Blind Source Separation) 기술 중 하나인 독립 성분 분석(Independent Component Analysis)기술을 이용하였다. 본 논문은 스피커의 도청장치로써 역할 가능여부를 위한 것으로 독립 성분 분석에 대한 기술적인 논의는 주제에 벗어나며, 이에 대한 자세한 설명은 참고문헌 [8]에 기술되어있다. 독립 성분 분석 기법을 이용하기 위해 독립 성분 분석에 관한 활발한 연구를 진행 중인 헬싱키 대학 독립 성분 분석 연구팀에서 무료로 제공하는 매틀랩(Matlab) 소프트웨어를 이용하였다. 스피커로 도청한 음성 신호의 음질 평가를 위해 해당 신호에 대한 신호 대 잡음 비율(Signal to Noise Ratio)을 측정하여 비교하였다. 신호 대 잡을 비율이 설명하지 못하는 부분에 평가하기 위해 주관적 실제 청음 평가 또한 평가방법으로 이용되었다.

5.1 실험에서 사용된 잡음신호 분석

그림 8은 잡음으로 사용한 물소리, 백색잡음, 사이렌소리와 같은 스윕잡음에 대한 스펙트로그램 (Spectrogram)이다. 백색잡음은 주파수에 대한 전력밀도의 스펙트럼이 거의 일정한 잡음을 말하며, 스윕잡음은 일정시간동안 주파수를 연속적으로 변화시키는 잡음을 일컫는다. 잡음의 스펙트로그램은 각 잡음이 서로 다른 잡음 환경 및 도청방지 기법이 적용되었음을 시각적으로 보여주기 위해 사용되었다. 스펙트로그램은 소리 파형의 시각화 기법으로서, 파형과 스펙트럼의 특징이 조합된 시각도구이다. 스펙트로그램은 시간축과 주파수축의 변화에 따른 진폭차를 색의 농담으로 나타낸다. 다시 말하면 어두운 색에 가까워질수록 진폭이 큰 것을 의미하며, 밝은 색상으로 가까워질수록 진폭이 작은 것을 의미한다.

백색잡음(B)은 넓은 주파수 범위에서 거의 일정한 주파수 스펙트럼을 가지는 신호로 특정한 청각패턴을 갖지 않는다. 반면 수도꼭지의 물소리(A)는 신호가 백색잡음과 같이 주파수 성분이 고루 분포되어 있지 않고 고주파 영역일수록 신호의 세기가 커지며, 백색잡음과 다른 스펙토그램을 보인다. 사이렌소리(C)는 주파수 영역에서 보았을 때 톱니형태를 가지며 가청 주파수 대역에서 스윕 잡음을 가지는 것을 볼 수 있다. 사이렌소리(C)는 주파수 영역에서 보았을 때 톱니형태를 가지며 가청 주파수 대역에서 일정 시간 동안 주파수를 연속적으로 변화시켜주는 스윕(Sweep) 형태의 잡음을 가지는 것을 볼 수 있다.

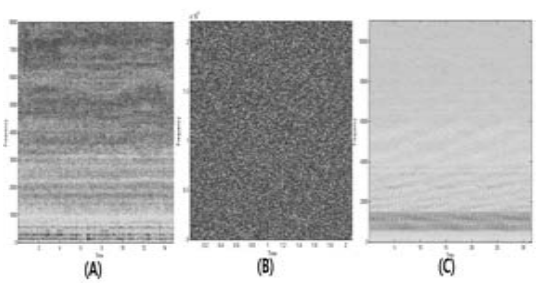

Fig.8. Spectrogram of waterfall sound(A), whitenoise(B), siren sound(C)

5.2 잡음환경에서 스피커를 이용한 음성수집 및 신호분리

그림 9는 두 스피커를 이용해 얻은 혼합 신호를 나

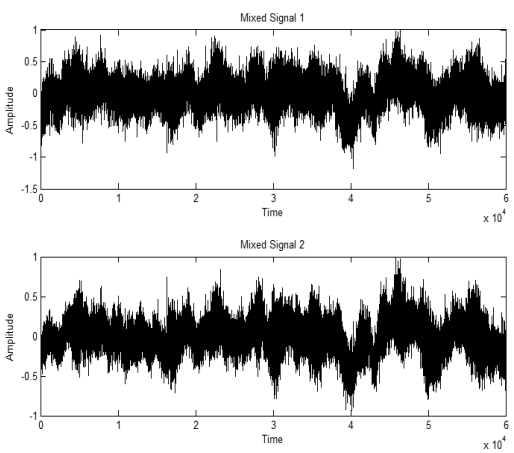

Fig.9. Two observed acoustic signals at sound-masking sound added environment

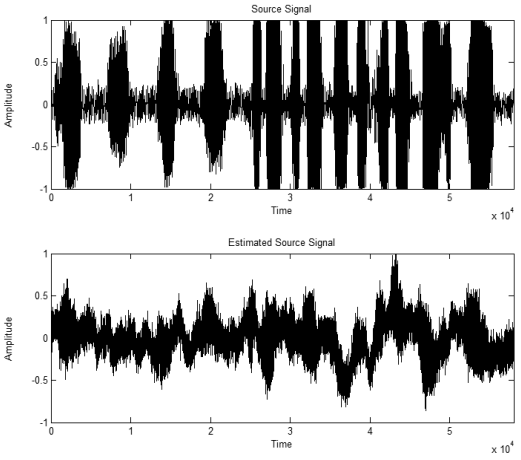

Fig.10. Original voice signal (top) and mixed signal with waterfall sound (bottom)

타낸다. 두 그림은 동일한 조건 하에 스피커들의 거리 및 녹음되는 공간을 다르게 하여 얻은 혼합신호들을 나타낸다. 각 스피커에서는 잡음과 원음이 흘러나오도록 설정하였으며, 해당 잡음(물소리)과 음원에 대한 특별한 평균 전력비 조건을 가하지 않고 혼합 음원을 얻었다. 두 혼합 신호는 약간 다른 형태를 보이는 것을 볼 수 있으며, 이것은 해당 변수들에 따라 혼합 신호의 형태가 달라질 수가 있다는 것을 보여준다. 그림 10의 아래 그림은 혼합 음원에서 잡음을 분리한 신호를 보여주며, 위 그림은 사운드 마스킹 잡음이 섞이지 않은 원음을 보여준다. 표 2는 독립 성분 분석을 이용하였을 때 목표 음원을 더욱 잘 인지 할 수 있을 것임을 수치로 보여주며, 적절한 신호처리 기법이 적용되었을 때 스피커 도청의 성능이 향상될 수 있음을 보여준다. 독립 성분 분석 도구를 이용한 분리는 분리한 잡음신호와 음성신호의 위치에 변화를 줄 수 있다. 이때 분리한 음성신호와 잡음의 주파수 영역이 역이 서로 겹치게 되면 분리하고자 하는 신호의 심각한 손상을 야기할 가능성이 있고, 어느 신호가 얻고자 하는 신호인지 알 수 없게 할 수도 있다. 이러한 한계는 기계의 음성신호인식 분야에서 치명적일 수 있지만, 본 연구의 목적은 음성 도청이기 때문에 실제 청취를 함으로써 그 한계를 극복할 수 있다고 가정하였다. 본 실험에서는 도청 도구로 스피커가 어느 정도 성능을 낼 수 있는지를 평가하기 위해 스피커로 받아들인 입력신호에 대해 객관적 평가를 하였다. 제안하는 기술은 도청 공격이 목적이기 때문에 청자로 하여금 얻은 신호로부터 목표 음성 신호를 어느 정도 수준으로 인지할 수 있는지 객관적, 주관적으로 평가하는 것이 중요하다. 본 실험에서는 도청 도구로 스피커가 어느 정도 성능을 낼 수 있는지를 평가하기 위해 스피커로 받아들인 입력신호에 대해 객관적 비교 평가를 하였다. 저자는 아날로그 신호처리에서 신호 질 평가의 대표적 방법인 신호 대 잡음비 (Signal to Noise Ratio, SNR) 평가 방법을 사용하였다. 해당평가는 신호전력과 잡음신호 전력의 상대적 비율 값을 제시함으로써 신호의 명료도가 어느 정도인가를 확인시켜주는 수치를 제공 해준다. 다음 식 (4)은 신호 대 잡음비 식을 나타낸다.

$$SNR_{db} = 10\log_{10}\left(\frac{V_0^2}{V_1^2}\right) = 20\log_{10}(V_0/V_1) \quad (4)$$

V_0는 스피커로부터 얻은 음원을, V_1은 원음을 의미한다. 이 값은 스피커로부터 얻은 음원과 원음이 어느 정도의 차이가 있는지를 수치로 알려줄 수 있으며 객관적인 지표로 사용될 수 있다.

그림 11의 위 2 그림은 원음과 백색 잡음이 혼합된 혼합 음원을 스피커로 받아들였을 때의 파형을 나타낸다. 아래는 혼합 음원으로부터 백색잡음을 분리하여 얻은 음성 신호를 보여준다. 표 2의 수치로 볼 때 백색 잡음을 분리하여 얻은 신호는 수치상 신호 대 잡음비가 7.747dB로 물소리(3.953dB), 스윕 잡음(5.411dB)의 경우보다 원음으로부터 거리가 먼 신호를 얻는다. 하지만 이것은 백색잡음 특성상 전 대역에 잡음 에너지가 분포되어 있기 때문에 일어나는 현상이며, 실제 청음 평가 및 스펙트럼 관찰에서는 상당부분

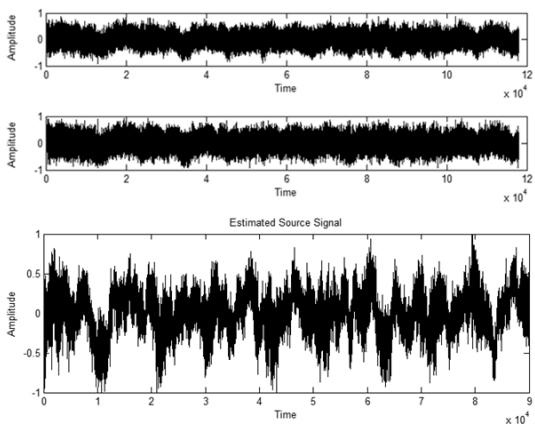

Fig.11. Mixed signal with white-noise (top) and separated voice signal (bottom)

Table 2. SNR(Signal to Noise Ratio) values of mixed and separated signals

잡음 종류	혼합신호	분리된 신호
물소리	32.668 dB	3.953 dB
백색잡음	10.952 dB	7.747 dB
스윕잡음(사이렌)	12.274 dB	5.411 dB

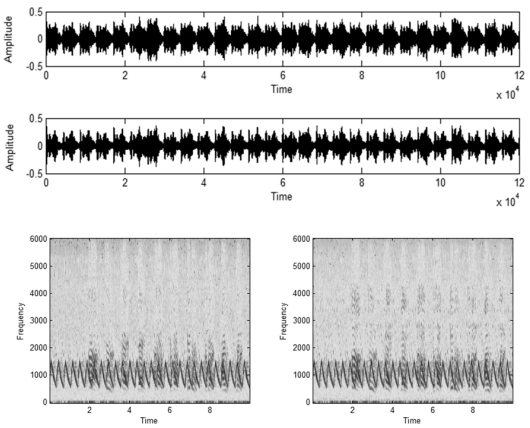

Fig.12. Mixed signal with siren noise (top) and their spectrogram (bottom)

분리되었다. 이는 본 실험의 성능 평가를 위한 방법으로 신호 대 잡음비의 설명 한계를 나타내며 실제 청음 등 주관적, 객관적인 평가를 더한 복합적 성능 평가의 필요성을 알려준다. 그림 12는 사이렌과 같은 가청주파수대역 스윕 잡음을 이용한 혼합 잡음 신호를 보여준다. 그림 13은 각각의 관찰된 신호의 사이렌과의 혼합신호 분리된 음성신호의 시간영역 그래프 및 스펙트로그램이다.

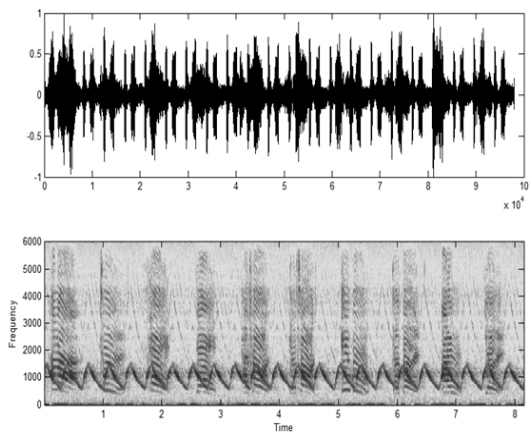

Fig.13. Separated signal from mixed signal with siren noise and their spectrogram

위의 실험결과들은 스피커가 일반적으로 하나의 유닛을 사용하는 것이 아닌 쌍으로서 사용되어지기 때문에, 추가적인 신호처리 기술로 스피커가 도청 장치로의 역할을 충분히 할 수 있다는 것을 보여준다. 대부분의 사람들이 스피커가 음성 신호를 받아들여 공격자에게 일정 수준 이상의 신호를 제공할 수 있다는 사실을 인지하지 못하고 있다는 사실은 공격자로 하여금 공격 방법으로써 스피커 도청을 좀 더 고려하게끔 만들어준다. 다음은 공격 도구로 마이크를 사용했을 때 보다 스피커를 이용했을 때 보다 효과적인 공격이 가능 하다는 것을 보이는 실험에 관한 것이다.

VI. 결 론

일반적으로 사람들은 스피커를 도청 장치로 인식하지 않는다. 또한 스피커는 우리 주변에 쉽게 찾을 수 있으며, 기존의 어떠한 도청 탐지 장비로도 도청 스피커가 그 기능을 하고 있는지 알 수 없다. 본 논문에서는 스피커 도청이 일반 스피커와 운영체제를 비교적 간단하게 조작 하여 여러 시나리오에서 공격이 가능하다는 것을 제시하였다. 더하여, 본 논문은 스피커를 이용한 도청 공격을 간단한 사운드 마스킹이 적용된 환경과 키보드 해킹에 적용해 보았고, 해당 실험을 통해 스피커 도청 공격이 가능함을 보였으며, 특정 환경에서의 키보드 해킹에 스피커가 마이크보다 효과적일 수 있다는 것 또한 보였다. 향후 연구에서는 본 논문에서 소개한 공격모델이외에 가능한 많은 시나리오에 대한 연구를 발전시키고, 스피커 구성 재질에 따른 집음 능력 실험을 추가적으로 수행할 예정이다. 또한 스피커도청에 대한 명확한 방지책에 대하여 논의 할 것이다.

Acknowledgement

본 연구는 "고려대학교 특별연구비에 의하여 수행되었음"(Supported by a Korea University Grant)

References

[1] "'Borderless' surrender by DPRK administration office sixth cyber attack," Chosun Daily, 2013.4.10., http://biz.chosun.com

/site/data/html_dir/2013/04/10/2013041001593.html
〔2〕 "Smartphone application 'Spy phone' first exposure," Dong-A Daily, 2013. 4. 5, http://dkbnews.donga.com/3/all/20130405/54219969/3
〔3〕 W. Zemlin, "Speech and Hearing Science," 4th Ed., Boston:Allyn & Bacon, 1998.
〔4〕 Georgia State University, http://hyperphysics.phy-astr.gsu.edu/hbase/sound/soucon.html
〔5〕 T. Yamamoto and T. Tsukagoshi, "New materials for loudspeaker diaphragms and cones-An overview," J. Acoust. Soc. Am. Volume 69, Issue S1, 1981.
〔6〕 P. Comon and C. Jutten, Handbook of Blind Source Separation, 1th Ed., Academic Press(Elsevier), Feb 2010.
〔7〕 C. M. Grinstead, "Introduction to Probability," 2th Ed., American Mathematical Society, pp.325-364, 1997.
〔8〕 J. V.Stone, "Independent Component Analysis : A tutorial Introduction," 1th Ed., The MIT Press Cambridge, Massachusetts, 2004 Londeon, England, Feb. 2010.
〔9〕 Nature News Snoopers can hear what you type, http://www.nature.com/news/2005/050922/full/news050919-9.htm
〔10〕 D. Asonov AND R. AgrawalR. "Keyboard Acoustic Emanations", In Proceedings of the IEEE Symposium on Security and Privacy (2004), pp. 3-11.
〔11〕 L. Zhuang and F. Zhou, "Keyboard Acoustic Emanations", In Proceedings of the IEEE Symposium on Security and Privacy (2004), pp. 373-382.
〔12〕 Realtek ALC882 Series Datasheet, oct, 2005
〔13〕 Korea Electronics Association, "Speaker industrial trend", http://www.eiak.org/electronic_info/data/Speaker_2008.pdf
〔14〕 Bosch Audio, Bosch Communications Systems 2011 Bosch Public Address Product Catalogue
〔15〕 Daniels, Drew. "Notes on 70-Volt and Distributed System Presentation", db, March/April 1988.
〔16〕 Don Davis and Eugene Patronis, "Sound System Engineering", 2nd Ed., Indianapolis, Howard W. Sams Co,, 1987, pp. 85-87, 402-405.
〔17〕 Sedra and Smith, Microelectronic Circuits, 6th Ed., Oxford University Press, USA, 2011.

〈저자소개〉

이 승 준 (Seung Joon Lee) 학생회원
2013년 2월: 한림대학교 컴퓨터공학과 졸업
2013년 9월~현재: 고려대학교 정보보호대학원 석박사 통합 과정
〈관심분야〉 정보보호, 하드웨어 보안, 임베디드 하드웨어, 디지털통신

하 영 목 (Young Mok Ha) 학생회원
2012년 2월: 고려대학교 정보통신대학 졸업
2012년 9월~현재: 고려대학교 정보보호대학원 석사과정
〈관심분야〉 정보보호, 빅데이터 보안, 통계신호처리, 전파통신

조 현 주 (Hyun Ju Jo) 학생회원
2012년 8월: 국민대학교 경영정보학과 졸업
2013년 3월~현재: 고려대학교 정보보호대학원 석사과정
〈관심분야〉 정보보호, 스태가노그라피, 소셜네트워크 보안, 모바일 보안

윤 지 원 (Ji Won Yoon) 정회원
2003년 2월: 성균관대학교 정보공학사 졸업
2005년 2월: University of Edinburgh, 정보학과 석사 졸업
2008년 11월: University of Cambridge 전자공학과 박사 졸업
2008년 2월~2009년 5월: University of Oxford, 로봇연구소 박사후과정
2009년 5월~2011년 5월: University of Dublin 통계학과 연구원 및 강사
2011년 7월~2012년 8월: IBM 연구소 정규 연구원
2012년 9월~현재: 고려대학교 정보보호대학원 조교수
〈관심분야〉 신호정보처리, 응용통계, 빅데이타 분석 기술, 도감청 탐지기술