Journal of Communications and Information Networks
Vol.1, No.4, Dec. 2016
DOI: 10.11959/j.issn.2096-1081.2016.055

Review paper

Securing acoustics-based short-range communication systems: an overview

CHEN Si[1], QIN Zhan[2], XING Guoliang[3], REN Kui[2]

1. West Chester University of Pennsylvania, West Chester PA 19383, USA
2. State University of New York at Buffalo, Buffalo NY 14260, USA
3. Michigan State University, East Lansing MI 48824, USA

Abstract: Mobile devices such as smartphones and tablets have continued to grow in recent years. Nowadays, people rely on these ubiquitous smart devices and carry them everywhere in their daily lives. Acoustic signal, as a simple and prevalent transmitting vector for end-to-end communication, shows unique characteristics comparing with another popular communication method, i.e., optical signal, especially on the applications performed over smart devices. Acoustic signal does not require line-of-sight when transmission, the computational power of most smart devices are sufficient to modulate/demodulate acoustic signal using software acoustic modem only, which can be easily deployed on current off-the-shelf smart devices. Therefore, many acoustics-based short range communication systems have been developed and are used in sensitive applications such as building access control and mobile payment system. However, past work shows that an acoustic eavesdropper snooping on the communication between a transmitter and its legitimate receiver can easily break their communication protocol and decode the transmitted information. To solve this problem, many solutions have been proposed to protect the acoustic signal against eavesdroppers. In this overview, we explore the designs of existing solutions, the corresponding implementations, and their methodologies to protect acoustic signal communication. For each dependable and secure acoustics-based short range communication system, we present the major technical hurdles to be overcome, the state-of-the-art, and also offer a vision of the future research issues on this promising technology.

Key words: acoustic, eavesdropper, short-range communication, ICA, BSS, randomize channel, motion source, usability

Citation: CHEN S, QIN Z, XING G L, et al. Securing acoustics-based short-range communication systems: an overview[J]. Journal of communications and information networks, 2016, 1(4): 44-51.

1 Introduction

Recent advancement of smartphones and tablet computing devices has witnessed the increasing popularity of short-range communication in many mobile applications and services. For instance, NFC (Near Field Communication) enables a low-power radio communication between two NFC-enabled devices by a simple touch, 1D/2D barcode enables millions of transactions which are finished over barcode-based payment services for retail customers. Comparing with these two prevalent short-range communication methods, acoustic short-range communication is a new and promising communication technique for smart devices, and has been used in a variety of smart device applications,

Manuscript received Sept. 6, 2016; accepted Nov.22, 2016
This work is supported in part by the U.S. National Science Foundation (No. CNS-1421903)

including mobile payment systems, building access control[1], and mobile data exchange[2,3]. Some of the most well-known applications in people's daily lives include Alipay, Soundpays and AasaanPay.

The unique properties of acoustic communication provide a number of highly desirable features for smart devices as well as clearly defined security strength. First of all, the transmission of acoustic signal does not require line-of-sight, which offers much higher usability than the barcode based communication systems. Secondly, the computational power of most smart devices are sufficient to modulate/demodulate acoustic signals using a software acoustic modem. Therefore, such acoustic communication systems can be easily deployed on most of the off-the-shelf smart device platforms. Unlike NFC chips, it is safe to assume that all current smartphones are readily equipped with a speaker and microphone as required by the functionality of phones. Thirdly, sound wave has inherent localization in the air medium, and it fades quickly when travels in distance. As a coin has two sides, this feature naturally enhances the data confidentiality of acoustic communication systems against eavesdropping. Finally, when the carrier frequency of a smartphone acoustic communication system lies within audible bandwidth, it is easy to detect jamming like DoS (Denial-of-Service) attacks and locate the adversaries by human ears. Due to these features of acoustic signal communication, it enables short-range communication designs with ultra-low cost, high compatibility, efficiency, and usability. Meanwhile, after carefully reviewing the state-of-the-art systems, we find that most of them typically adopt weak encryption protocols, or lack encryption altogether, leaving them widely exposed to security threats[4].

The main disadvantage of acoustic signal short-range communication is the vulnerability to eavesdropping attacks. In practice, the adversary can easily conduct eavesdropping as a passive attack method by using specialized microphones[5]. For example, an adversary can hide one or several remotely controlled wireless microphone(s) to record the communication between the acoustic signal transmitter and the receiver. In most cases, eavesdropping attacks can be addressed with more sophisticated encryption protocols, namely (elliptic curve) Diffie-Hellman key exchange protocol. However, such an approach would introduce too many cryptography-based procedures into the system, which are computationally expensive and power-consuming, and may severely compromise the communication efficiency.

To tackle this problem, recently, researchers utilize the physical layer of wireless communication to protect acoustic signal communications without utilizing those "expensive" crypto tools. Among these state-of-the-art methods, acoustic friendly jamming[6] is one of the most popular and novel secure communication schemes, which is under fast development[7]. The basic idea of acoustic friendly jamming is to let the receiver transmit a random jamming signal (artificial noise) while the transmitter is transmitting the data signal. This method keeps the confidentiality of communication as the jamming signal and the data signal arrives at the attacker at the same time (a mixture signal). Therefore, the attacker cannot separate and demodulate the data signal correctly. However, since the jammer is also the one who receives the messages, he can easily remove the jamming signal from the received mixture signal and decode the message. In both Refs.[2] and [3], the authors use the random white Gaussian noise as jamming signal and PSK (Phase-Shift Keying) or FSK (Frequency-Shift Keying)[8] to modulate the data. They both prove that the proposed acoustic self-jamming schemes are immune to eavesdropping attack, DoS attack, and BSS (Blind Source Separation) attack.

2 System model and assumptions

In the acoustics-based short-range communication system, we assume that both transmitter and receiver are the off-the-shelf smartphones, no priori knowledge of any secret information have been shared before the communication. The acoustics-based system uses the existing microphones and speakers on smartphones to enable short range communication, thus, eliminating the need for specialized NFC hardware. Same as in conventional NFC where communication through magnetic coupling is confined to a short range, acoustics-based short range communication systems also require the communication range to be confined within a short range (few cm) as shown in Fig.1. We model the channel between the smartphones as a LoS (Line-of-Sight) channel. A key advantage of acoustics-based short range communication system is that it is a purely software-based solution that can be implemented on legacy phones, as long as they have a speaker and a microphone. Unlike conventional NFC, which does not incorporate any security at the physical or MAC layers since the short range of communication is in it presumed to offer a degree of protection, acoustic based short-range communication system provides security at the physical layer using a novel friendly-jamming technique. The security thus obtained is information-theoretic.

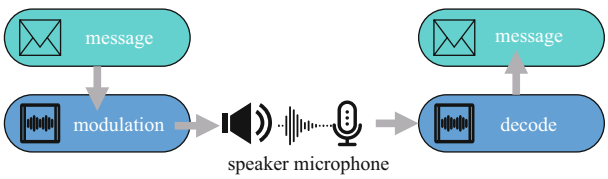


Figure 1 System model for acoustics-based short-range communication system.

In the acoustics-based short-range communication system, we mainly consider a passive eavesdropper whose objective is to obtain the exchanged data during the acoustic transmissions. The threat model is based on the scenario where adversary may deploy multiple sensors (hidden microphones) at any fixed location in priori to the acoustic communication. A more detailed threat model will be discussed at the following section.

3 Secure acoustics-based short-range communications

3.1 A brief view of secure acoustics-based short-range communications

Acoustic communication has received continuous and extensive attention and it has been widely used in many underwater wireless communication systems. To secure these acoustics-based short-range communication systems, researchers are utilizing the physical layer of wireless communication to design novel ways to protect communications without using any prior shared key. We take a glance of those state-of-the-art friendly-jamming systems first.

Dhwani: The first acoustics-based short-range communication system, called Dhwani, uses the special random white Gaussian noise as jamming signal, which consists of parts of several one-tie random white Gaussian noises. The main components of Dhwani include an ingress filter, an OFDM (Orthogonal Frequency Division Multiplexing) based software module, and a self-jamming module. It applies PSK (Phase-Shift Keying) as the digital modulation way to module the data.

PriWhisper: As a parallel and independent work, another acoustics-based short range communication system called PriWhisper was proposed in Ref.[2]. PriWhisper also uses the random white Gaussian noise as jamming signal. Different from Dhwani, which focuses on more implementation aspects, the authors aim to provide rigorous security analysis of friendly-jamming technology in acoustics-based

short-range communication systems. In their model, they well designs the generated noise to cover the frequencies selected by FSK (Frequency-Shift Keying) modulation scheme.

3.2 Threat model and design challenges

Prior work has considered the problem of eavesdropping over acoustic emanations as a side channel. The system security is analyzed in the standard LoS channel model. Both PriWhisper and Dhwani are expected to provide secure communication in the presence of either single or multiple passive eavesdropper(s). A typical scenario is that the eavesdropper places one or more remotely controlled wireless microphone(s) near a user's workspace in priori and records the acoustic signal during the transmission. In particular, multiple-sensor eavesdroppers may try to separate the data signal from its recorded mixture signals. Fig.2 illustrates a typical attack scenario where the attacker utilizes multiple microphones for eavesdropping (R_1, R_2). Note that s_1 and s_2 are two original signals (data signal and jamming signal), x_1 and x_2 are two received mixed signals.

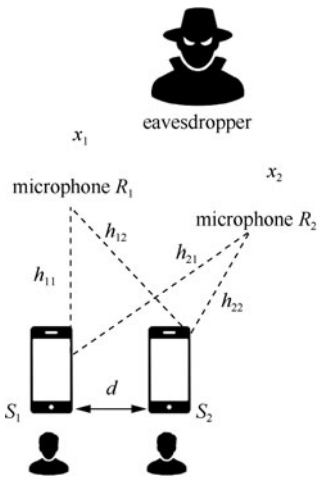


Figure 2 System workflow for PriWhisper

To fully separate s_1 and s_2 from x_1 and x_2, the adversary continuously launching a separation attack. In this attack, the adversary tries to estimate the data signal and jamming signal by using BSS techniques. FD-ICA (Frequency Domain Independent Component Analysis) is one of the most famous algorithms. Upon success, the adversary can separate the transmitted data from the mixed signal.

BSS is a technique for estimating original source signals using only observed mixtures. It has a wide range of applications including high-quality telecommunication system and robust speech recognition. FD-ICA is one of the most popular blind signal segmentation techniques. Assume that the original data signal and jamming signal are $s_i(t)$ (i=1,⋯,N), the signals which are observed by the eavesdropper using microphone j are $x_j(t)$(j=1,⋯,M), and the separated signals are $y_k(t)$(k=1,⋯,N), the BSS model can be described by the following equations:

$$x_j(t) = \sum_{i=1}^{N} (h_{ij} * s_i)(t),$$

$$y_k(t) = \sum_{j=1}^{N} (\omega_{kj} * x_j)(t),$$

where h_{ji} represents the coefficient from source i to the eavesdropper's microphone j, ω_{kj} is the coefficient for the FIR (Finite Impulse Response) filter, and $*$ denotes the convolution operator. Instead of applying an ordinary ICA algorithm in the time domain to solve the BSS problem, we can first use a STDFT (Short-Time Discrete Fourier Transform) to convert the time domain signal into frequency domain, and then apply ICA in the frequency domain. The model is approximated as:

$$\boldsymbol{X}(\omega, n) = \boldsymbol{H}(\omega)\boldsymbol{S}(\omega, n),$$

where, ω is the angular frequency, and n denotes the frame index of the recorded audio. $\boldsymbol{S}(\omega, n) = [\boldsymbol{S}_1(\omega, n),\cdots,\boldsymbol{S}_N(\omega, n)]^T$ is the source signal in frequency bin ω, $\boldsymbol{X}(\omega, n)=[\boldsymbol{X}_1(\omega, n),\cdots,\boldsymbol{X}_M(\omega, n)]^T$ denotes the mixed signals. The separating process can be represented by the following equation:

$$\boldsymbol{Y}(\omega, n) = \boldsymbol{W}(\omega)\boldsymbol{X}(\omega, n),$$

where $\boldsymbol{Y}(\omega, n) = [\boldsymbol{Y}_1(\omega, n), \cdots, \boldsymbol{Y}_N(\omega, n)]^T$ denotes the estimated data and jamming signal, and $\boldsymbol{W}(\omega)$ represents the separating matrix. The goal of FD-ICA is to determine the $\boldsymbol{W}(\omega)$ so that $\boldsymbol{Y}_i(\omega, n)$ and $\boldsymbol{Y}_j(\omega, n)$ become mutually independent. Fig.3 shows the whole process.

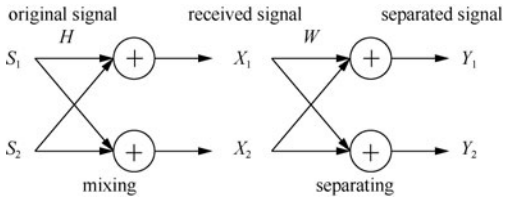


Figure 3 Model of BSS system

3.3 System architecture

The general architecture of self-jamming communication system is shown in Fig.4. Both PriWhisper and Dhwani are designed to enable key-less secure acoustic short-range communication in smartphone-smartphone and smartphone-terminal scenarios. The distance between two devices should be in a few cm. At the beginning of the transmission, the receiver sends jamming noise and the transmitter sends data signal simultaneously. Once the receiver gets the mixture signal, the receiver removes the jamming noise with the help of its own knowledge. However, due to the phase distortion, the receiver can only remove parts of the jamming noise. To protect the confidentiality of the message, the duration time of the jamming noise should always cover the duration time of the message signal.

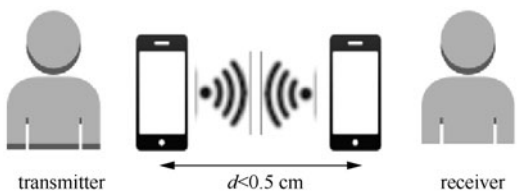


Figure 4 Acoustic self-Jamming communication Architecture

The workflows of PriWhisper are shown in Fig.5. For PriWhisper, the transmitter first broadcast a start signal to inform the receiver to prepare for receiving messages. Then the transmitter starts detecting jamming signal. The receiver, once detect the starting signal, begins recording audio signals. In the next step, the receiver plays a synchronization sound for itself to mark the beginning spot of the jamming signal. Once finished, the receiver immediately plays the jamming signal. The power of the jamming signal is the maximum power that the receiver can reach. Once the transmitter detects the jamming signal, it begins to transmit messages. Finally, the receiver removes the jamming signal from its received mixture signal and decode it to get the message.

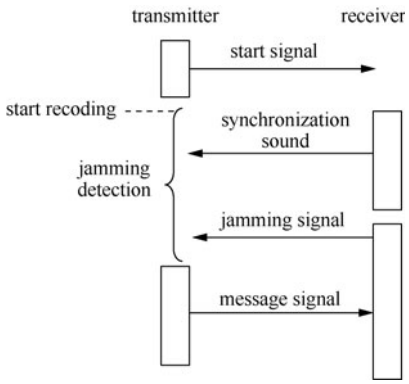


Figure 5 System workflow for PriWhisper

4 Future research issues

The practical use of these acoustics-based short range communication systems in a real-world scenario relies on the system usability, data through-put and the security strength provided by the friendly jamming technique. Further research study needs to be done in the following areas.

4.1 Non-invasive design for acoustics-based communication

To make a non-invasive design of acoustics-based

communication, the transmission is made in the ultrasonic or near ultrasonic frequency range that humans cannot hear. The carrier frequency of the smartphone usually has a wider range. However, as most off-the-shelf smartphone speakers are only capable of producing sound with a 44.1 KHz sample rate, based on Nyquist-Shannon sampling theorem, this allows us to use a maximum frequency of about 22 KHz for transmitting ultrasonic sound on smartphones. The ultrasonic range for most humans is above 19.5 KHz. Motivated by this observation, it is possible to have a non-invasive design of acoustics-based communication system using the non-voice frequency band, achieving even higher security guarantee.

To implement non-invasive design into these existing systems, new and novel mechanisms of jamming signal generation needed to be designed. Recently, researchers have proposed the perceiving power ratio threshold in Ref.[9]. According to Ref.[9], human can perceive specific audio from background wide-band noise if SNR (Signal-to-Noise-Ratio) of single-frequency audio is above13 dB. While for multi-frequency audio, SNR should be larger than 0 dB. One solution is to adopt spread spectrum encoding where each bit of message is transmitted as random noise that has been filtered to be only in the ultrasonic spectrum 19.5 KHz to 22 KHz. Modulated noise rather than signal is chosen so that it blends more discreetly into the background. Using the entire spectrum also allows for a low power but high SNR signal. To defend against activate adversaries using ultrasonic frequency, a novel physical layer integrity checking scheme need to be further designed.

4.2 Enforced proximity communications

Existing acoustic friendly jamming system is assumed that the transmitter and the receiver are both in a fixed location. This makes the system vulnerable to multi-eavesdropping attack if the distance d between the transmitter and the receiver is larger than a threshold h_d (usually 0.5 cm). To enforce proximity communications, researchers rely on two main techniques: distance estimation and bounding protocol and contextual co-presence approach. Distance estimation and bounding protocol aims to cryptographically bound the distance between transmitter and receiver by measuring the response time. However, exiting distance estimation protocols[10-12] and distance bounding protocols[13] either can only work in special environments or require specialized hardware. Contextual co-presence approach[10-12,14], on the other hand, comparing the ambient information (e.g., RSSI level, GPS, etc.) sensed by transmitter and receiver to enforce proximity. These contextual co-presence approach also suffer from a few limitations. First, they aim to determine relative distance (e.g. which device is closer) instead of absolute distance between transmitter and receiver. Second, these approaches are insecure because attackers can modify the ambience around the transmitter and receiver. Recently, a system named Dolphin[15] has been proposed to restrict distance. Dolphin utilizes the fast decay property of acoustic signals to ensure the distance and uses full-duplex communication to defend against eavesdropping attacks. As a future research direction, it is important and necessary to investigate the effectiveness of these acoustic near filed assertion system, under a very powerful attacker equipped with multiple microphones, by conducting both analytical evaluation and experiments.

4.3 Channel randomization

Another solution for improving system security is to randomize the acoustic channels from the receiver to the eavesdropper to prevent accurately decoding the message. To implement channel randomization in

practice, one possible solution is to leverage recent results in under water acoustic communication[16]. Ref.[16] shows that due to multi-path effect, even small motion of the transmitter and receiver can create large variation in the acoustic channel. Thus, we could place the transmitter and receiver both on a rotating frame, and randomly change the relative location of both transmitter and receiver to randomize the signal. This creates fast varying acoustic channels with a random distribution. It also provides the channel diversity of an acoustic transmitter with a huge number of possible locations, which renders an eavesdropper unable to separate and decode.

In fact, the BSS algorithm running by the eavesdropper is highly directional and the solution of the FD-ICA is as the same way as an adaptive beamformer. Because of this special characteristic, FD-ICA is robust as regards a moving transmitter. But it may fail to decode when we add a moving receiver since the moving receiver add randomizing to the acoustic channels which makes the eavesdropper impossible to form a spatial null towards the jamming signal. Thus, to defend against multiple passive adversaries, the use of channel randomization techniques may help effectively protect today's widely used commercial acoustic systems from eavesdroppers without degrading usability.

4.4 Performance enhancement

For performance enhancement design, it has been observed that the data rate, maximum transmission range and robustness of the system is closely related to the frequency modulation scheme and error correction scheme. To further improve the system performance and dependability, it is possible to explore advanced modulation methods (e.g. multiple frequencies amplitude modulation) and repetition error correction scheme. On top of this, a systematic study of system SNR, indoor and outdoor multi-path effect and frequency amplitude modulation methods is also very critical to evaluate the robustness of the acoustics-based short communication system. Moreover, other practical countermeasures and security enhanced schemes may also need to be designed to improve the overall system security.

5 Concluding remarks

In this article we presented an overview of existing acoustics-based short range communication systems, specifically, we picked two state-of-the-art acoustic friendly jamming systems, Dhwani and PriWhisper. They both provide near field communication functionalities and enable stronger security guarantees but require less strict hardware support. However, as we pointed out in the discussion section (Section 4), there is still much room to improve the usability and practicality of these acoustics-based systems in terms of security strength, channel security, transmission data rate and so on. Further, we demonstrated that the non-invasive design, enforced proximity techniques and acoustic channel randomization technique can be combined with many existing security primitives, which opens doors to the designing of a variety of safe acoustic short-range communication systems. We also believe that these acoustics-based short range communication systems will be widely adopting in our daily life after we successfully overcome these technical challenges.

References

[1] CHEN S, LI M, QIN Z, et al. AcousAuth: an acoustic-based mobile application for user authentication[C]//IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2014: 215-216).

[2] ZHANG B, ZHAN Q, CHEN S, et al. Enabling keyless secure acoustic communication for smartphones[J]. IEEE internet of things journal. 2014, 1(1): 33-45.

[3] NANDAKUMAR R, CHINTALAPUDI K K, PADMANABHAN V, et al. Dhwani: secure peer-to-peer acoustic NFC[C]//ACM SIGCOMM Computer Communication Review, 2013, 43(4): 63-74.

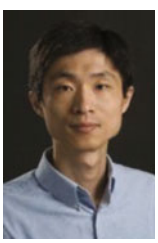
[4] WEI T, WANG S, ZHOU A, et al. Acoustic eavesdropping through

wireless vibrome-try[C]//The 21st Annual International Conference on Mobile Computing and Networking, 2015: 130-141.
[5] KORTVEDT H, MJOLSNES S. Eavesdropping near field communication[C]//The Norwegian Information Security Conference (NISK) 2009: 57.
[6] GOEL S, NEGI R. Guaranteeing secrecy using artificial noise[J]. IEEE transactions on wireless communications, 2008, 7(6): 2180-9.
[7] KUMAR S, UMA M K, KUMAR K. Peer-to-peer acoustic near field communication[J]. IOSR journal of electronics and communication engineering.
[8] SKLAR B. Digital communications[M]. Upper Saddle River: Prentice Hall, 2001.
[9] STUART J R. Noise: methods for estimating detectability and threshold[J]. Journal of the audio engineering society, 1994, 42(3): 124-40.
[10] MA D, SAXENA N, XIANG T, et al. Location-aware and safer cards: enhancing RFID security and privacy via location sensing[J]. IEEE transactions on dependable and secure computing, 2013,10(2): 57-69.
[11] SCHRMANN D, SIGG S. Secure communication based on ambient audio[J]. IEEE transactions on mobile computing, 2013, 12(2): 358-70.
[12] HALEVI T, MA D, SAXENA N, et al. Secure proximity detection for NFC devices based on ambient sensor data in computer security C ESORICS 2012[M]. Berlin: Springer Berlin Heidelberg, 2012: 379-396.
[13] BRANDS S, CHAUM D. Distance-bounding protocols[C]// Workshop on the Theory and Application of of Cryptographic Techniques, 1993, 344-359.
[14] KRUMM J, HINCKLEY K. The nearme wireless proximity server[C]//InInternational Conference on Ubiquitous Computing, 2004: 283-300.
[15] LI L, XUE G, ZHAO X. The Power of Whispering: Near Field Assertions via Acoustic Communications[C]//The 10th ACM y on Information, Computer and Communications Security, 2015: 627-632.
[16] STOJANOVIC M, BEAUJEAN P P. Acoustic communication[R]. DOI 10.1109/ACCESS.2016.2552538, IEEE Access, 2016.

About the authors

CHEN Si received the Ph.D. degree in computer science from State University of New York at Buffalo. He is a member of the Ubiquitous Security and Privacy Research Laboratory (UbiSeC Lab). He is now an assistant professor at West Chester University of Pennsylvania. His research interests include cyber-physical system security and mobile crowd-sensing system. (Email: schen@wcupa.edu)

QIN Zhan is currently working toward his Ph.D. degree at the Ubiquitous Security and Privacy Research Laboratory (UbiSeC) in the Computer Science and Engineering Department of the State University of New York at Buffalo. His research interests are in the areas of cloud computing and security, with focus on differential privacy data collection and publication, cybersecurity in smart grid. He is a student member of the IEEE, IEEE COMSOC, and ACM. (Email: zhanqin@buffalo.edu)



XING Guoliang is an associate professor of Computer Science and Engineering Department at Michigan State University. He earned his D.Sc. (2006) and M.S. (2003) in computer science from Washington University in St. Louis. He received a B.S. degree in electrical engineering and M.S. degree in computer science from Xi'an Jiaotong University in 1998 and 2001, respectively. Prior to joining MSU, he was an assistant professor of computer science at the City University of Hong Kong. His research interests include wireless sensor networks, mobile computing, and networked embedded systems. (Email: glxing@cse.msu.edu)

REN Kui [corresponding author] is a professor of computer science and engineering and the director of UbiSeC Lab at State University of New York at Buffalo (UB). He received his Ph.D. degree from Worcester Polytechnic Institute. Kui's current research interest spans cloud & outsourcing security, wireless & wearable systems security, and mobile sensing & crowdsourcing. His research has been supported by NSF, DoE, AFRL, MSR, and Amazon. He received UB Exceptional Scholar Award for Sustained Achievement in 2016, UB SEAS Senior Researcher of the Year Award in 2015, Sigma Xi/IIT Research Excellence Award in 2012, and NSF CAREER Award in 2011. He has published 170 peer-review journal and conference papers and received several Best Paper Awards including IEEE ICNP 2011. He currently serves as an associate editor for IEEE Transactions on Dependable and Secure Computing, IEEE Transactions on Mobile Computing, IEEE Wireless Communications, IEEE Internet of Things Journal, and IEEE Transactions on Smart Grid. Kui is a fellow of IEEE, a distinguished lecturer of IEEE, a member of ACM, and a past board member of Internet Privacy Task Force, State of Illinois. (Email: kuiren@buffalo.edu)