Characterizing Privacy Risks in Healthcare IoT Systems

Shuai Li[1], Alessio Baiocco[2], and Shouhuai Xu[1]*

[1] Department of Computer Science
University of Colorado Colorado Springs
Colorado Springs, CO USA 80918
{sli,sxu}@uccs.edu
[2] Department of Information Security and Communication Technology
Norwegian University of Science and Technology NTNU
Gjøvik, Norway
alessio.baiocco@ntnu.no

Abstract. The advancement in Internet of Things (IoT) technology has been having a huge societal and economic impact, effectively changing the paradigms in doing business including the healthcare industry. While citizens now can enjoy the convenience brought up by healthcare IoT systems, such as wearable healthcare IoT devices, the privacy risks incurred by these systems and devices are not well understood, let alone adequately addressed. In this paper we systematically characterize the privacy risks in healthcare IoT systems, by considering a range of privacy attack vectors such as those that can be imposed by healthcare IoT device fingerprinting and semi-honest Internet Service Providers. Then, we leverage these characteristics to guide us in exploring countermeasures for mitigating privacy risks in healthcare IoT systems. We hope the present study will serve as a baseline for designing a systematic solution to protect citizen's privacy in healthcare IoT systems.

Keywords: Healthcare IoT Systems · Healthcare IoT Devices · Privacy · Data Breach · Privacy Risk · Cyber Risk

1 Introduction

Internet of Things (IoT) devices have become popular recent years, including smart watch and smart home ecosystem (e.g., smart thermostat, WiFi plug). Correspondingly, the healthcare industry has leveraged IoT devices to improve their services, leading to the notion of *healthcare IoT systems*, which have been widely employed for healthcare functions such as automated and remote patient monitoring, glucose monitoring, smart inhaler, and health data collection [21,30, 43,64]. This leads to a large amount of data that can not only be used by doctors for purposes such as diagnosing and immediate attention to health issues, but also be used for healthcare research purposes. It is predicted that the global

[*] Corresponding author

market size of healthcare IoT devices will rapidly grow from US$291.2 billion in 2023 to US$861.3 billion in 2030 [1].

With the fast-growing popularity of healthcare IoT systems, we must adequately understand, characterize, and mitigate the potential security and privacy risks pertinent to healthcare IoT systems. For example, healthcare IoT devices are widely known to have unaddressed vulnerability surface that makes them susceptible to many attack vectors. This imposes a big risk because healthcare data often, if not always, include sensitive personal information, such as one's vital signs (e.g., pulse rate, body temperature) and medical problems (e.g., blood sugar level, blood pressures). These sensitive healthcare data pose a serious threat to citizens' privacy and potential unfair social welfare (e.g., a healthcare insurance company may refuse to sell insurance to a person when the company knows what kinds of medical problems from which the person is suffering).

To the best of our knowledge, the privacy problem in the context of healthcare IoT systems has not received the due amount of attention, meaning much research remains to be done. In this paper, we conduct a systematic study on characterizing the privacy risks in healthcare IoT systems.

Our Contributions. This paper makes two contributions. First, we characterize the privacy risks that can occur to healthcare IoT systems, through the following perspectives: (i) attack vectors via healthcare IoT device fingerprinting, (ii) attack vectors associated with Internet communications despite the employment of standard countermeasures such as cryptography-protected communications (e.g., Transport Layer Security or TLS, VPN-protected communications) and anonymous communication channels (e.g., Tor), (iii) attack vectors that are applicable to IoT data collection, (iv) attack vectors that may be waged by a curious or semi-honest Internet Service Provider (ISP), and (v) attack vectors that may be waged against healthcare service provider servers. Second, we leverage the resulting characteristics to guide our exploration of potential countermeasures to protect healthcare privacy against those attack vectors. This exploration would pave a way for future studies on designing systematic and practical solutions to harden healthcare privacy.

Related Works. Healthcare IoT techniques have shown great potential in providing high-quality healthcare services to citizens as evidenced by the following studies. Liu *et al.* [43] proposed an IoT-based heart ECG monitoring system that can detect cardiac abnormality in real time. Wu *et al.* [64] proposed to integrate ECG sensors into a T-shirt and use a bio-potential chip to collect quality ECG data. Istepanian *et al.* [30] reported a non-invasive IoT glucometer to monitor the glucose in real time. Fu and Liu [21] designed a non-invasive tissue oximeter to measure the blood oxygen saturation level, along with heart rate and pulse parameters.

However, the security and privacy issues in healthcare IoT systems have received much less attention and there are not so many studies. Tang *et al.* [62] designed a privacy-preserving healthcare data aggregation scheme that can achieve secure data collection from multiple sources and provide fair incentives for contributing patients. Fang *et al.* [19] proposed an anomaly detection scheme to detect healthcare IoT devices that have been compromised by attackers. Li *et al.* [37] adopted consortium blockchains to allow patients to manage, share, trade their medical records securely. In addition, there are early studies on analyzing risks associated with IoT systems in the healthcare domain [3] and there are some studies on exploring countermeasures [36, 58, 63]. For example, the Open Web Application Security Project (OWASP) [58] has identified some threats and vulnerabilities in healthcare IoT systems, including the lack of authorization, the insufficient authentication associated with the pertinent Internet communications, the insecure web interface of the healthcare service provider servers, the lack of transport layer encryption, the insecure network service, the insecure cloud interface, the inadequate security configuration, and the insecure mobile interface. Despite these studies, there is no systematic understanding on the privacy risks in healthcare IoT systems. For example, even [3] does not present a system model that would be comparable to what we will propose in this paper, which means that their analysis of risks would not be applicable to our setting. The present study aims at a systematic characterization of privacy risks. For example, we investigate privacy risks despite the employment of countermeasures that could have prevented some vulnerabilities discussed in [3] (e.g., employing sufficient authentication and transport layer cryptographic mechanisms).

Paper Organization. The rest of the paper is organized as follows. Section 2 presents the system model of healthcare IoT systems. Section 3 characterizes the privacy risks associated with healthcare IoT systems. Section 4 explores countermeasures to mitigate the privacy risks. Section 5 discusses the limitations of the present study. Section 6 concludes the paper with several exciting future research directions.

2 System Model

Figure 1 describes the system model of healthcare IoT systems, which will severe as the basis for discussing privacy risks in healthcare IoT systems and countermeasures for mitigating these privacy risks. We consider a range of healthcare IoT devices, including: emergency button (for emergency care), heart rate monitor, glucose monitor, sleep tracker, and blood sugar monitor. These devices are assumed to be connected to a gateway or a smartphone. The connection can be via WiFi, ZigBee, or Bluetooth Low Energy (BTLE).

The gateway is responsible for collecting data from those devices and communicating healthcare data to some healthcare servers by some healthcare service providers. There are multiple servers because the IoT devices are manufactured by different companies and each service provider has its own server to collect and analyze the data collected from its customers or patients. The analysis results are assumed to be returned back to a customer or patient via an App provided by the service provider, and the App runs in the gateway (i.e., smartphone). The communications between the gateway and the healthcare servers are based on the Internet, likely facilitated by some Internet Service Provider (ISP).

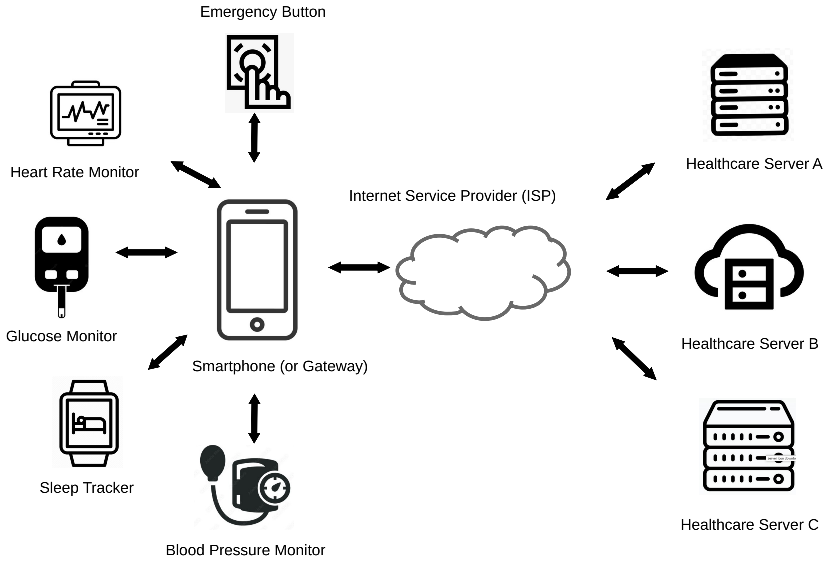

Fig. 1. System model of healthcare IoT systems

To protect the privacy of customers or patients, the healthcare servers often adopt some "standard" security and privacy mechanisms. For example, they would employ the TLS protocol or the Virtual Private Network (VPN) technology to protect the communication between the gateway and the server in question. Moreover, this gateway-server communication may even be protected by some anonymous communication mechanisms, such as the Tor technology. At the application layer, a customer or patient may have created some pseudonym rather than using real name in the communication (e.g., for authenticating to a server), and a server may not know the mapping between the pseudonym and the real name of the customer or patient.

The research question we ask is: even if the aforementioned security and privacy mechanisms have been employed in healthcare IoT systems, is the privacy of customers or patients adequately protected? In what follows we characterize the privacy risks in healthcare IoT systems despite the employment of the standard security and privacy mechanisms mentioned above.

3 Privacy Risks in Healthcare IoT Systems

Due to the sensitive nature of the healthcare data, privacy protection in healthcare IoT systems is a necessity and required by law. In what follows we discuss six categories of privacy risks with respect to the system model presented above.

3.1 Privacy Risks Incurred by Communications between Healthcare IoT Devices and the Gateway (Smartphone)

Healthcare IoT devices connect to the gateway via wireless communications, which provide the attackers the chance to eavesdrop the connection. Encrypting the traffic of healthcare IoT devices to the gateway can mitigate this risk, but encryption alone is far from sufficient because studies have showed that wireless devices could be fingerprinted by exploiting the identifiable features at the physical layer, the Medium Access Control (MAC) layer, and the network layer [69].

First, the imperfection in the manufacture process of wireless transmitters can cause varying wireless communication features (e.g., clock skew, frequency offset, and phase offset) [24]. As a consequence, healthcare IoT devices manufactured by the same vendor would share some common wireless features, which can be exploited by an attacker to recognize what IoT devices are being used and who are their vendors.

Second, some details of the MAC layer protocol are unspecified in the pertinent standard. As a consequence, the concrete implementation is largely left to the vendors and the discrepancy between these implementations can be exploited to infer the vendor to which a healthcare IoT device belongs [12].

Third, an attacker could exploit network-layer features to fingerprint healthcare IoT devices effectively [59]. This is because different kinds of IoT devices communicate with the gateway in different ways (e.g., the frequency of communication, the number of packets that are communicated). More specifically, network-layer features such as the number of packets, the packet size, and the direction of the packets can be exploited by an attacker to distinguish different kinds of healthcare IoT devices [38]. An even more concrete example is that the heart rate monitor may incur more frequent updates than the glucose monitor does, and this discrepancy can be exploited to tell these devices apart. This privacy risk cannot be prevented by encrypting the traffic because cryptosystems cannot hide the frequency of communications.

3.2 Privacy Risks Incurred by Internet Communications Despite Standard Countermeasures

Pertinent to the system model described in Figure 1, there are privacy risks associated with the communications between the gateway and the healthcare service provider servers (or clouds), which would be partly or completely based on Internet. To characterize these risks, we consider three scenarios: the communications are protected by TLS; the communications are protected by VPN; the communications are protected by anonymous communication techniques such as Tor. We do not consider the scenario where none of these mechanisms (i.e., TLS, VPN, Tor) is employed because in which case a passive attacker can breach a user's privacy completely by eavesdropping the channel owing to the fact that the personal medical data is sent in plaintext.

Privacy Risks Despite TLS-protected Communications. In this scenario, the smartphone that acts as a gateway of the BWAN (Body Wide Area Network), which consists of the smartphone and the healthcare IoT devices (including their associated sensors), is expected to communicate securely with the healthcare service provider servers using the well-established TLS protocol. The TLS protocol provides, among other things, authentication, confidentiality, and data integrity services between two parties by establishing an authenticated private channel via an appropriate key-exchange procedure. That is, TLS allows mutual authentication between users' smartphones / gateways and the remote medical servers, encrypts data to assure it confidentiality, and assures that the data (encrypted or not) is not manipulated during transmission.

However, the employment of TLS might give a false sense of privacy protection, at least for the following four reasons. First, earlier versions of TLS (i.e., version 1.1 and version 1.2) are known to be vulnerable, explaining why TLS version 1.3 has been proposed. However, many computers or devices in the real world (smartphone and/or the server in this context) have not employed the most updated protocol [31, 60], meaning that some of them would be susceptible to known attacks. Second, even if TLS version 1.3 is used, the authenticity of public key certificate of the gateway and/or the server represents another source of potential vulnerability even if these certificates are issued by trusted certification authority and their integrity is assured. One root cause of this phenomenon is that there are systems-based attacks against cryptosystems that may not be detected immediately, and these delays undermine the trustworthiness of cryptographic services owing to the presence of (for example) compromised cryptographic signing keys and/or functions [80]. Third, there are vulnerabilities that can affect the assurance offered by TLS, such as the CRIME attack [31] and the BREACH attack [22, 31, 49], which exploit the cookie mechanism by brute-forcing them. Fourth, TLS requires a granular configuration (application to application), which offers a great configuration flexibility but, if not properly managed, opens doors to serious implementation vulnerabilities that can be exploited by attackers [56].

Even if the preceding risks associated with TLS are carefully prevented, TLS does not provide any means to assure anonymous communications between a gateway and a healthcare service provider server. As a consequence, an attacker passively eavesdropping the Internet could easily figure out which user (via the user's smartphone) is communicating with which healthcare service provider, which could breach privacy. For example, if a user's smartphone often communicates with a cancer care service provider, this communication alone would expose that the user has the kind of cancer in question.

Privacy Risks Despite VPN-protected Communications. The VPN technology aims to create a secure point-to-point connection over the Internet (insecure network) between two networks or devices (e.g., the gateway and a healthcare service provider in the context of the present paper) [2]. By creating a point-to-point connection, the VPN technology encapsulates IP packets to prevent attackers from sniffing the network traffic. It can also prevent ISPs from spying on the network traffic of its users. In general, VPN can guarantee data integrity, confidentiality and authenticity of network communications.

Still, privacy risks can emerge despite the employment of the VPN technology for a gateway to access a healthcare service provider server. For example, when the VPN service is outsourced to an external service provider that handles the activities of its users, the VPN service provider can be a threat to privacy. Second, the cryptographic protocol that is used by a VPN may have vulnerabilities, such as those associated with the key exchange protocols whose cryptographic properties are extremely delicate. Such vulnerabilities could cause the compromise of data confidentiality, and thus privacy of the users. Third, the environment in which VPN is used can pose as a threat. This can be demonstrated by the LocalNet attack [81]. Putting this attack into the context of the present paper, a compromised router between a gateway and a healthcare service provider server can, despite the use of VPN, provide the client with incorrect network settings (e.g., routing tables), which represent public IP addresses of interest to the attackers as part of the local network. As a consequence, the data communicated in a VPN channel falls under the exclusions and bypasses the VPN tunnel [81].

Even if the preceding risks associated with VPNs are carefully prevented, VPN, similar to TLS, does not provide any means to assuring anonymous communication between a gateway and a healthcare service provider server. As a consequence, an attacker passively eavesdropping the Internet can still breach a user's privacy by monitoring communication patterns (e.g., what disease a user may be suffering as shown in the case of TLS).

Privacy Risks Despite Tor-protected Communications. Tor [13,38,55] is one implementation of the concept of Onion Routing. It is a network of virtual tunnels that aim to assure privacy and anonymity of its users by preventing passive attackers from tracking the traffic generated by, in the context of this paper, a gateway and a healthcare service provider server. While using cryptosystems to protect the confidentiality and integrity of the payload, Tor could give a false sense of privacy that it can hide the source-destination communication relationship because Tor is also vulnerable to attacks [13]. As one example, the *Autonomous System* eavesdropping attack [4] can be waged by the ISP used by a smartphone (i.e., gateway in this paper) to allow the autonomous system that deals with the re-routing of packets, to be present in the access relay and exit relay of the smartphone. As a consequence, the attacker can carry out a correlation attack between the incoming and outgoing traffic to breach privacy. As another example, the *exit node* eavesdropping attack [33] can be waged by its ISP to intercept the traffic from an exit Tor node. Since the exit relay traffic may not be encrypted by the Tor user and Tor doesn't encrypt the exit relay traffic by itself, the ISP may successfully eavesdrop the exit relay traffic.

3.3 Privacy Risks in Healthcare IoT Data Collection

Healthcare IoT devices leverage a smartphone as the gateway to communicate with a healthcare service provider server largely because of their constrained power and computational capability. Unfortunately, the gateway (smartphone) is subject to cyber attacks. For example, a smartphone often runs many apps, and a malicious app may be able to break its sandbox via privilege-escalation to control the smartphone and the other apps (e.g., healthcare apps). Moreover, the apps may share the smartphone's cache in the CPU, making cache side-channel attacks possible. For example, a malicious app that uses the same cache with a healthcare app can perform a side-channel attack (e.g., PRIME+PROBE [50]) to learn how the cache has been affected by the healthcare app, effectively inferring the activities of the healthcare app and even its cryptographic private keys. In addition, studies (e.g., [65]) even showed that an attacker can exploit inaudible attacks to compromise a smartphone (e.g., sending unauthorized text messages, triggering malicious downloads, changing the WiFi settings, or performing context-aware voice recording). As a consequence, the attacker can exploit these attacks to breach private healthcare data.

3.4 Privacy Risks Incurred by Curious Internet Service Providers

It is important to highlight a particular privacy threat vector that can be waged by ISPs, including the scenarios that an ISP itself is compromised and then abused to breach the privacy of its users. Healthcare IoT device users typically depend on ISPs to connect to healthcare care service provider servers. However, an ISP can be curious (i.e., semi-honest) to learn its users' usage of healthcare IoT devices. In what follows we discuss why standard techniques for privacy protection are not adequate.

- ***TLS***. TLS can protect the communications between users' smartphones / gateways and the healthcare service provider servers, but not the communications between the healthcare IoT devices and the associated smartphone / gateway in most cases (owing to the computational and communication overhead incurred by TLS). Although employing TLS can prevent an ISP from learning the application-layer data, it cannot prevent the ISP from learning "who is communicating with whom" or "which IP address is communicating with which other IP address." This can breach the privacy of the users. For example, if a user frequently communicates with a healthcare service provider offering cancer therapy, the ISP can infer that the user might be suffering from a cancer.
- ***VPN***. The healthcare IoT gateway may use the VPN technology to communicate with a healthcare service provider server. In this case, the gateway would incur Internet communication traffic with the VPN server offered by the healthcare service provider, which still exposes with which healthcare service provider a user is communicating. Moreover, an ISP could exploit the traffic correlation technique to figure out the destination of the traffic originating from the gateway [32].
- ***Tor***. As discussed above, Tor [13] can hide the communication relationship between a user and a healthcare service provider server to some extent. However, studies [38, 59] have showed that traffic features, such as packet count and packet direction, could still leak information about the source

and destination of the connection. In particular, privacy risk can be incurred when the Tor routers belong to a single ISP.

The preceding analysis shows that standard techniques cannot adequately prevent ISPs from breaching the privacy of users in healthcare IoT systems.

3.5 Privacy Risks Incurred by Attacks against Healthcare Service Provider Servers

The healthcare service provider servers can become the target of cyberattacks. Recent years have witnessed some large-scale data leakage from healthcare servers, impacting millions of patients [20,34,67]. The leaked information include patient name, home address, date of birth, and appointment information. One approach to mitigating the privacy risks incurred by attacks against the healthcare service provider servers is to make these servers store personal information as little as possible. However, this is not practical for at least two reasons: (i) the healthcare service providers would treat the user's data as their assets; (ii) keeping all data pertinent to a user would enable a better healthcare service to the user because a complete medical history is a critical source of information when a user gets a serious disease.

3.6 Privacy Risks Despite the Use of Application-Layer Pseudonyms

One may suggest to use application-layer pseudonyms to alleviate privacy risks against a malicious or compromised healthcare service provider server and the other attacks mentioned above. For example, a user may use a pseudonym rather than personal identifier to index their healthcare data. However, this technique is often vulnerable to the re-identification attack because it cannot guarantee complete anonymity (e.g., the data may retain some "fingerprints" or "linkability" that can be exploited to recover the identity of a data owner [9, 54]). These deanonymization attacks may be waged by a malicious healthcare service provider, or by an attacker that compromised a healthcare service provider server.

To be more specific, we note that patient re-identification has been reported by correlating medical data saved in data servers with patient discharge logs [82]. It is intuitive that the greater the medical details of the pathology categorized with codes in the databases, the greater the possibility that this data will be re-identified. Moreover, the greater the side information made public by healthcare facilities, the greater the ability to deanonymize patient data by correlating them (e.g., information about hospitalizations and medical conditions, including prescription data, medical mailing lists , employers, debtors, friends, and family).

Note that making a user to have multiple pseudonyms is no good idea because it will reduce the usefulness of the healthcare data as a medical doctor cannot see the complete medical history of a patient and the statistical analysis conducted by a medical researcher would not obtain high-quality results.

4 Exploring Countermeasures for Enhancing Privacy in Healthcare IoT Systems

As analyzed above, privacy risks impose challenges that demand new solutions. In this section we explore four kinds of countermeasures that would be needed in order to mitigate those privacy risks.

4.1 Traffic Feature Obfuscation between IoT Devices and Gateway

In order to defeat the healthcare IoT device fingerprinting attacks discussed above, we propose obfuscating the identifiable features at the physical layer, the MAC layer, and the network layer.

Fingerprint obfuscation at the physical layer. The imperfection in the manufacture process of wireless transmitters is inevitable, which makes fingerprinting possible. One approach to alleviating the risk would be to make the physical layer's features device-specific rather than vendor-specific. The intuition is that if the physical-layer features of a healthcare IoT device are unique to the device and do not leak any information about its vendor, it would prevent an attacker from inferring the vendor. What remains to be investigated include: (i) the unique features could be exploited to unambiguously identify a device, which may have another kind of privacy implication because unique fingerprint may serve as a unique identifier of the device; (ii) how the manufacturing process may be “randomized” to achieve fingerprint obfuscation; and (iii) the required distance to the patient that makes the fingerprinting attacks feasible.

Fingerprint obfuscation at the MAC layer. MAC-layer features are incurred by vendor-specific implementations of the protocols when dealing with the unspecified details of MAC-layer protocols. To prevent the vendor-specific features at the MAC layer, a vendor should adopt more common implementation shared with other vendors, instead of crafting its own. To achieve this, the healthcare IoT industry should standardize the unspecified MAC layer details so as to preserve the healthcare IoT users' privacy at the MAC layer.

Fingerprint obfuscation at the network layer. When a healthcare IoT device user is in a public space, an attacker can monitor its wireless traffic to learn the IoT devices that are being used. The network-layer features include packet size, packet count, packet directions, and burst pattern. To defeat such attacks, we propose obfuscating the network traffic features. In what follows we propose two preliminary schemes for this purpose.

- ***Random Dummy Packet Injection.*** When a healthcare IoT device communicates with the gateway (smartphone), both sides need to inject dummy packets randomly so as to obfuscate traffic features such as packet counts and packet direction. In order to prevent interference to the existing IoT data processing services, an IoT device should send its real packets with no delay, while injecting dummy packets to the gateway with a pre-determined probability p_i. Similarly, the gateway sends its real packets to the IoT device as usual, while sending dummy packets with a probability p_g. To better protect user privacy, we prefer larger p_i and p_g. However, the IoT devices are usually energy-constrained, which means that p_i should not be too large; otherwise, the IoT device's power would be drained quickly. Thus, we need to make a trade-off when determining probability p_i. On the contrary, the gateway has no such constraints, and we can choose a larger p_g to achieve better privacy protection.

- ***Random Packet Padding.*** If the packets sent by the healthcare IoT devices and the gateway have different sizes, we propose padding the packets in order to prevent the packet size from leaking any information to the attacker. When an IoT device or the gateway is about to send a packet with size s_c, the sender should pad the packet to the size $s_c + d$ with a probability p_c, where d is a random integer between 0 and $s_{max} - s_c$ and s_{max} is the maximum size a packet is allowed to have.

Note that the general idea of traffic padding has been proposed in other context [35], but hasn't been systematically studied in medical IoT settings. Open research questions include: How can we determine the padding parameters for medical IoT devices? Moreover, these techniques should be used together with cryptographic mechanisms for encrypting the content of device-gateway communication; otherwise, an attacker could recognize the dummy packets. For this purpose, light-weight cryptosystems, including both confidentiality and integrity protection mechanisms, should be used.

4.2 Privacy Enhancement for Healthcare IoT Device Data Collection

We propose taking countermeasures to defeat an attacker from learning healthcare data from the gateway or smartphone. In response to a malicious app taking control of the smartphone or gateway, we advocate adopting the Trusted Execution Environment (TEE) and implement the healthcare apps to leverage TEE [18]. This assures that even if a process with a higher privilege has been hijacked by a malicious app, the attacker still cannot compromise the data of the healthcare app because the data is encrypted in the memory space. Note that it is only when the healthcare app gets executed, the data is decrypted in the TEE. In addition, we need to address the cache side-channel attacks by run-time diversification or cache partitioning [42], so that a malicious app and healthcare app do not share a common cache resource.

4.3 Thwarting a Curious ISP

We propose addressing the privacy risks that can be incurred by a curious ISP by hiding the true destination of the gateway's Internet traffic. TLS protocol cannot conceal the true destination's IP address, while VPN service is subject to the traffic correlation attack mentioned above. Therefore, we propose adopting the onion routing [55] or the TOT implementation to conceal the true destination of the gateway's traffic, but in a more sophisticated way than the standard use described in Section 3.4, by using onion routers that belong to different ISPs. Recall that Tor contains thousands of volunteer nodes or onion routers. We propose that the gateway (smartphone) should choose at least three onion routers as its *entry guard*, *middle relay*, and *exit relay*. A gateway should have its traffic to go through the chosen onion routers sequentially before reaching the healthcare service provider server. The gateway exhibits traffic destined to the Tor onion routers such that a curious ISP has no idea about the true destination. It's also much harder to conduct traffic correlation attacks on Tor onion routers as long as the three onion routers belong to different ISPs (i.e., no single ISP can monitor these three onion routers).

Note that recent studies [38, 59] find that the traffic features of Tor networks can be informative, and that an advanced attacker can harvest these traffic features to tell which website a Tor user has visited. In our context, such an attack can be exploited by a curious ISP to learn which healthcare service provider server the gateway has connected to, which compromises the user's privacy. In order to mitigate this attack, we propose that the gateway should initiate deliberate web browsing activities when connecting to the healthcare service provider server, and both the web browsing traffic and IoT traffic use the same set of onion routers. The goal is to mix the two kinds of traffic together so that the traffic features of healthcare IoT devices will be obfuscated by the web browsing traffic.

4.4 Data Privacy Protection for Healthcare Provider Server

Healthcare service provider servers can become the subject of the cyber attacks discussed above. Although the stored medical records in the these servers may be anonymized, the records are still subject to re-identification attacks. To protect data privacy, we propose the following scheme. First, the data should be encrypted using some appropriate homomorphic encryption schemes that allow the desired operations over ciphertexts corresponding to some applications (e.g., statistical analysis). Second, the private key for decrypting the data and the ciphertext resulting from homomorphic operations over the ciphertexts should be protected in a secure environment, such as TEE such that the private key cannot be compromised (assuming the attacker cannot exploit any side-channel). However, one must recognize that even if the private key cannot be compromised, data privacy can still be at risk. This is because the cryptographic function corresponding to the private key could be compromised without compromising the private key. In theory this is known as *oracle access* to the cryptographic function (e.g., decryption or digital signing). In practice, this threat has inspired many studies to mitigate the problem, such as [11, 14, 15, 28, 52, 72, 80].

5 Limitations

The present study has three limitations, which need to be addressed in future studies. First, the system model we considered, which focused on data collection, can be extended to accommodate other emerging components, such as edge computers with which the gateway may communicate with. Second, the extended system model may introduce new threat vectors, meaning that the threat model may need to be revised correspondingly. Third, we proposes several countermeasures to address the privacy risks but without experimental evaluation. The high-level design we proposed need to be elaborated, refined, and evaluated.

6 Conclusion and Future Research Directions

We have presented a characterization of privacy risks associated with healthcare IoT systems via a range of privacy attack vectors. We have explored countermeasures to mitigate these privacy risks. We hope this study will inspire many future endeavors on adequately assuring privacy in healthcare IoT systems.

Open problems for future studies are abundant. In addition to addressing the limitations mentioned above, we highlight the following.

First, how should we design a systematic architecture to assure privacy in healthcare IoT systems? This architecture should be holistic in the sense of encompassing all the layers, including the application layer and communication layer, because the preceding discussion suggests that privacy breaching can be achieved by exploiting information gathered at multiple layers.

Second, how should we design a systematic set of privacy-protection mechanisms to mitigate privacy risks in healthcare IoT systems? Similarly, the mechanisms should consider multiple layers and the inference attacks that may be waged by attackers. This is nontrivial because the current research are often geared towards point solutions. For example, cryptographic multiparty computation [23], an elegant mechanism for protecting data privacy when multiple participants need to conduct some joint computational tasks over the union of their data, does not prevent an eavesdropper to infer which participants are conducting such activities with which other participants. This means that additional mechanisms are needed in order to prevent the eavesdropped from making such inferences.

Third, how should we quantify privacy risks in healthcare IoT systems? How should we quantify the effectiveness or capabilities of each privacy mechanisms? What privacy metrics are needed? Although there have been some very nice and useful privacy metrics, such as differential privacy [17], we observe they are geared towards the application layer if not a particular kind of applications. As mentioned above, privacy risks can be incurred by exploiting information collected at different layers (e.g., application layer and communication layer). This highlights that privacy is an emergent property, which suggests that holistic privacy cannot be achieved by composing building-block or point solutions each of which achieves certain privacy assurances in their respective models [70].

Fourth, the preceding discussion suggests that privacy in the healthcare sector (and perhaps in a broader context) should be treated holistically. This is reminiscent of the notion of *cybersecurity dynamics* [71, 76, 77], which intends to model, quantify, and analyze cybersecurity from a holistic perspective because cybersecurity also exhibit emergent behavior [8, 53, 70, 78]. This prompts us to envision the notion of *privacy dynamics*, which intuitively means the following: the degree of privacy breached by the adversary evolves with time, and there could be a threshold of tolerable privacy reach above which the privacy in question is completely breached. This means that privacy breaching would exhibit the *phase transition* phenomenon, which has been exhibited by theoretical cybersecurity dynamics studies [10, 25, 26, 39, 41, 45, 51, 66, 68, 73–75, 79, 83, 84] and data-driven cybersecurity studies [6,7,27,61]. For example, it would be very interesting to know whether the rich phenomena exhibited by cybersecurity dynamics would be exhibited by privacy dynamics as well, such as: global convergence [84] and global attractivity [26] for preventive and reactive cyber defense dynamics, and chaos for active cyber defense dynamics under certain circumstances [83]. The privacy implications of these phenomena would also deserve investigation. Moreover, we would need to define privacy metrics to accommodate the emergent properties, reminiscent of studies on defining cybersecurity metrics to measuring cybersecurity from a holistic perspective [5, 8, 16, 46, 53, 78].

Fifth. going beyond the healthcare sector, it is important to realize that cyber attackers are interested in compromising healthcare data (e.g., via cyber social engineering attacks [44, 47, 48, 57]) not only for the purpose of breaching privacy, but also for garnering patients' information and then exploiting the breached data to wage further attacks (e.g., blackmailing) [29, 40].

Acknowledgement. We thank the anonymous reviewers for their useful comments. Shouhuai Xu is supported in part by NSF Grant #2115134 and Colorado State Bill 18-086. This research work is also a contribution to the International Alliance for Strengthening Cybersecurity and Privacy in Healthcare (CybAlliance, Project no. 337316).

References

1. Internet of things in healthcare market size report, 2030. *https://www.grandviewresearch.com/industry-analysis/internet-of-things-iot-healthcare-market.*
2. H. Abbas, N. Emmanuel, M. Amjad, T. Yaqoob, M. Atiquzzaman, Z. Iqbal, N. Shafqat, W. Shahid, A. Tanveer, and U.r Ashfaq. Security assessment and evaluation of vpns: A comprehensive survey. 55(13s), jul 2023.
3. Nasser Abouzakhar, Andy Jones, and Olga Angelopoulou. Internet of things security: A review of risks and threats to healthcare sector. In *2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pages 373–378, 06 2017.
4. M. Akhoondi, C. Yu, and H. Madhyastha. Lastor: A low-latency as-aware tor client. In *2012 IEEE Symposium on Security and Privacy*, pages 476–490, 2012.
5. John Charlton, Pang Du, and Shouhuai Xu. A new method for inferring ground-truth labels and malware detector effectiveness metrics. In *Science of Cyber Security - Third International Conference (SciSec'2021)*, volume 13005 of *Lecture Notes in Computer Science*, pages 77–92. Springer, 2021.

6. H. Chen, J. Cho, and S. Xu. Quantifying the security effectiveness of firewalls and dmzs. In *Proc. HoTSoS'2018*, pages 9:1–9:11, 2018.
7. Huashan Chen, Hasan Cam, and Shouhuai Xu. Quantifying cybersecurity effectiveness of dynamic network diversity. *IEEE Transactions on Dependable and Secure Computing*, 2021.
8. Jin-Hee Cho, Shouhuai Xu, Patrick M. Hurley, Matthew Mackay, Trevor Benjamin, and Mark Beaumont. Stram: Measuring the trustworthiness of computer-based systems. *ACM Comput. Surv.*, 51(6):128:1–128:47, 2019.
9. A. Creţu, F. Monti, S. Marrone, X. Dong, M. Bronstein, and Y. Montjoye. Interaction data are identifiable even across long periods of time. *Nature Communications*, 13, 01 2022.
10. G. Da, M. Xu, and S. Xu. A new approach to modeling and analyzing security of networked systems. In *Proc. HotSoS'14*, pages 6:1–6:12, 2014.
11. W. Dai, P. Parker, H. Jin, and S. Xu. Enhancing data trustworthiness via assured digital signing. *IEEE TDSC*, 9(6):838–851, 2012.
12. L. Desmond, C. Yuan, T. Pheng, and R. Lee. Identifying unique devices through wireless fingerprinting. In *Proc. ACM WiSec*, pages 46–55, 2008.
13. R. Dingledine, N. Mathewson, and P. Syverson. Tor: The second-generation onion router. In *Proc. Usenix Security*, volume 4, pages 303–320, 2004.
14. Y. Dodis, J. Katz, S. Xu, and M. Yung. Key-insulated public key cryptosystems. In *Proc. EUROCRYPT'2002*, volume 2332 of *LNCS*, pages 65–82, 2002.
15. Y. Dodis, J. Katz, S. Xu, and M. Yung. Strong key-insulated signature schemes. In *Public Key Cryptography (PKC'03)*, pages 130–144, 2003.
16. P. Du, Z. Sun, H. Chen, J. H. Cho, and S. Xu. Statistical estimation of malware detection metrics in the absence of ground truth. *IEEE T-IFS*, 13(12):2965–2980, 2018.
17. Cynthia Dwork. Differential privacy. In *Proc. Automata, Languages and Programming, 33rd International Colloquium (ICALP'06)*, pages 1–12.
18. Jan-Erik Ekberg, Kari Kostiainen, and N. Asokan. Trusted execution environments on mobile devices. In *Proc. ACM CCS'2013*, page 1497–1498, 2013.
19. L. Fang, Y. Li, Z. Liu, C. Yin, M. Li, and Z. Cao. A practical model based on anomaly detection for protecting medical iot control services against external attacks. *IEEE Trans. on Industrial Informatics*, 17(6):4260–4269, 2020.
20. Zijian Fang, Maochao Xu, Shouhuai Xu, and Taizhong Hu. A framework for predicting data breach risk: Leveraging dependence to cope with sparsity. *IEEE Trans. Inf. Forensics Secur.*, 16:2186–2201, 2021.
21. Yu Fu and Jian Liu. System design for wearable blood oxygen saturation and pulse measurement device. *Procedia manufacturing*, 3:1187–1194, 2015.
22. Yoel Gluck, Neal Harris, and Angelo Prado. Breach: reviving the crime attack. *Unpublished manuscript*, 2013.
23. O. Goldreich, S. Micali, and A. Wigderson. How to play any mental game or a completeness theorem for protocols with honest majority. In *Proc. 19th ACM Symp. on Theory of Computing*, pages 218–229. ACM, 1987.
24. Jeyanthi Hall, Michel Barbeau, Evangelos Kranakis, et al. Enhancing intrusion detection in wireless networks using radio frequency fingerprinting. *Communications, internet, and information technology*, 1, 2004.
25. Y. Han, W. Lu, and S. Xu. Characterizing the power of moving target defense via cyber epidemic dynamics. In *HotSoS*, pages 1–12, 2014.
26. Y. Han, W. Lu, and S. Xu. Preventive and reactive cyber defense dynamics with ergodic time-dependent parameters is globally attractive. *IEEE TNSE*, 8(3):2517–2532, 2021.

27. Richard E. Harang and Alexander Kott. Burstiness of intrusion detection process: Empirical evidence and a modeling approach. *IEEE Trans. Inf. Forensics Secur.*, 12(10):2348–2359, 2017.
28. K. Harrison and S. Xu. Protecting cryptographic keys from memory disclosures. In *IEEE/IFIP DSN'07*, pages 137–143, 2007.
29. Mohammad Hijji and Gulzar Alam. A multivocal literature review on growing social engineering based cyber-attacks/threats during the covid-19 pandemic: challenges and prospective solutions. *Ieee Access*, 9:7152–7169, 2021.
30. R. Istepanian, S. Hu, N. Philip, and A. Sungoor. The potential of internet of m-health things "m-iot" for non-invasive glucose level sensing. In *2011 IEEE Conference of Engineering in Medicine and Biology Society*, pages 5264–5266, 2011.
31. Oleksandr Ivanov, Victor Ruzhentsev, and Roman Oliynykov. Comparison of modern network attacks on tls protocol. In *2018 IEEE International Conference Problems of Infocommunications. Science and Technology*, pages 565–570, 2018.
32. Aaron Johnson, Chris Wacek, Rob Jansen, Micah Sherr, and Paul Syverson. Users get routed: Traffic correlation on tor by realistic adversaries. In *Proc. ACM CCS'2013*, page 337–348, 2013.
33. Torbjörn Jonsson and Gustaf Edeby. Collecting and analyzing tor exit node traffic. MS Thesis, Blekinge Institute of Technology, 2021.
34. The HIPPA Journal. Healthcare data breach statistics. https://www.hipaajournal.com/healthcare-data-breach-statistics/.
35. M. Juárez, M. Imani, M. Perry, C. Dıaz, and M. Wright. Wtf-pad: toward an efficient website fingerprinting defense for tor. In *Proc. ESORICS*, 2016.
36. Ronald L Krutz and Russell Dean Vines. Cloud security: A comprehensive guide to secure cloud computing wiley publishing. *Inc. Indianapolis, Indiana*, 2010.
37. Chaoyang Li, Mianxiong Dong, Jian Li, Gang Xu, Xiubo Chen, and Kaoru Ota. Healthchain: Secure emrs management and trading in distributed healthcare service system. *IEEE Internet of Things Journal*, 8(9):7192–7202, 2021.
38. S. Li, H. Guo, and N. Hopper. Measuring information leakage in website fingerprinting attacks and defenses. In *Proc. ACM CCS'2018*, page 1977–1992, 2018.
39. X. Li, P. Parker, and S. Xu. A stochastic model for quantitative security analyses of networked systems. *IEEE TDSC*, 8(1):28–43, 2011.
40. Tian Lin, Daniel E Capecci, Donovan M Ellis, Harold A Rocha, Sandeep Dommaraju, Daniela S Oliveira, and Natalie C Ebner. Susceptibility to spear-phishing emails: Effects of internet user demographics and email content. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 26(5):1–28, 2019.
41. Z. Lin, W. Lu, and S. Xu. Unified preventive and reactive cyber defense dynamics is still globally convergent. *IEEE/ACM ToN*, 27(3):1098–1111, 2019.
42. Fangfei Liu, Yuval Yarom, Qian Ge, Gernot Heiser, and Ruby B. Lee. Last-level cache side-channel attacks are practical. In *2015 IEEE Symposium on Security and Privacy*, pages 605–622, 2015.
43. Ming-Liang Liu, Liu Tao, and Zhou Yan. Internet of things-based electrocardiogram monitoring system. *Chinese Patent*, 102(764):118, 2012.
44. Theodore Longtchi, Rosana Montañez Rodriguez, Laith Al-Shawaf, Adham Atyabi, and Shouhuai Xu. Sok: Why have defenses against social engineering attacks achieved limited success? *CoRR*, 2022.
45. Wenlian Lu, Shouhuai Xu, and Xinlei Yi. Optimizing active cyber defense. In *International Conference on Decision and Game Theory for Security*, pages 206–225. Springer, 2013.

46. J. Mireles, E. Ficke, J. Cho, P. Hurley, and S. Xu. Metrics towards measuring cyber agility. *IEEE Transactions on Information Forensics and Security*, 14(12):3217–3232, 2019.
47. Rosana Montañez, Adham Atyabi, and Shouhuai Xu. *Book Chapter in "Cybersecurity and Cognitive Science"*, chapter Social Engineering Attacks and Defenses in the Physical World vs. Cyberspace: A Contrast Study. Elsevier, 2022.
48. Rosana Montañez, Edward Golob, and Shouhuai Xu. Human cognition through the lens of social engineering cyberattacks. *Frontiers in Psychology*, 11:1755, 2020.
49. Christopher Ng. Ssl-tls security flaws: the breach and logjam attacks. NTU Technical Report, 2021.
50. Dag Arne Osvik, Adi Shamir, and Eran Tromer. Cache attacks and countermeasures: The case of aes. In *Proc. CT-RSA'2006*, page 1–20, 2006.
51. Keith Paarporn and Shouhuai Xu. Analysis of contagion dynamics with active cyber defenders. *CoRR*, 2023.
52. T. Paul Parker and Shouhuai Xu. A method for safekeeping cryptographic keys from memory disclosure attacks. In *1st International Conference Trusted Systems (INTRUST'2009)*, volume 6163 of *LNCS*, pages 39–59, 2009.
53. Marcus Pendleton, Richard Garcia-Lebron, Jin-Hee Cho, and Shouhuai Xu. A survey on systems security metrics. *ACM Comput. Surv.*, 49(4):62:1–62:35, December 2016.
54. Jovan Powar and Alastair R Beresford. Sok: Managing risks of linkage attacks on data privacy. *Proceedings on Privacy Enhancing Technologies*, 2:97–116, 2023.
55. M.G. Reed, P.F. Syverson, and D.M. Goldschlag. Anonymous connections and onion routing. *IEEE J. Selected Areas in Communications*, 16(4):482–494, 1998.
56. I. Ristic. *Bulletproof SSL and TLS: Understanding and Deploying SSL/TLS and PKI to Secure Servers and Web Applications.* 2013.
57. Rosana Montanez Rodriguez and Shouhuai Xu. Cyber social engineering kill chain. In *Proceedings of International Conference on Science of Cyber Security (SciSec'2022)*, 2022.
58. J. Shahid, R. Ahmad, A. Kiani, T. Ahmad, S. Saeed, and A Almuhaideb. Data protection and privacy of the internet of healthcare things (iohts). *Applied Sciences*, 12(4):1927, 2022.
59. Payap Sirinam, Nate Mathews, Mohammad Saidur Rahman, and Matthew Wright. Triplet fingerprinting: More practical and portable website fingerprinting with n-shot learning. In *Proc. 2019 ACM CCS*, page 1131–1148, 2019.
60. Preeti Sirohi, Amit Agarwal, and Sapna Tyagi. A comprehensive study on security attacks on ssl/tls protocol. In *2016 2nd international conference on next generation computing technologies (NGCT)*, pages 893–898. IEEE, 2016.
61. Z. Sun, M. Xu, K. Schweitzer, R. Bateman, A. Kott, and S. Xu. Cyber attacks against enterprise networks: Characterization, modeling and forecasting. In *Proc. of SciSec'2023*, 2023.
62. Wenjuan Tang, Ju Ren, Kun Deng, and Yaoxue Zhang. Secure data aggregation of lightweight e-healthcare iot devices with fair incentives. *IEEE Internet of Things Journal*, 6(5):8714–8726, 2019.
63. Vic JR Winkler. *Securing the Cloud: Cloud computer Security techniques and tactics.* Elsevier, 2011.
64. Taiyang Wu, Jean-Michel Redouté, and Mehmet Yuce. A wearable, low-power, real-time ecg monitor for smart t-shirt and iot healthcare applications. In *Advances in Body Area Networks I: Post-Conference Proceedings of BodyNets 2017*, pages 165–173. Springer, 2019.

65. Qi Xia, Qian Chen, and Shouhuai Xu. Near-ultrasound inaudible trojan (nuit): Exploiting your speaker to attack your microphone. In Joseph A. Calandrino and Carmela Troncoso, editors, *Proc. Usenix Security*, 2023.
66. M. Xu, G. Da, and S. Xu. Cyber epidemic models with dependences. *Internet Mathematics*, 11(1):62–92, 2015.
67. M. Xu, K. Schweitzer, R. Bateman, and S. Xu. Modeling and predicting cyber hacking breaches. *IEEE Trans. Inf. Forensics Secur.*, 13(11):2856–2871, 2018.
68. M. Xu and S. Xu. An extended stochastic model for quantitative security analysis of networked systems. *Internet Mathematics*, 8(3):288–320, 2012.
69. Q. Xu, R. Zheng, W. Saad, and Z. Han. Device fingerprinting in wireless networks: Challenges and opportunities. *IEEE Communications Surveys & Tutorials*, 18(1):94–104, 2016.
70. S. Xu. Emergent behavior in cybersecurity. In *Proc. HotSoS*, pages 13:1–13:2, 2014.
71. S. Xu. The cybersecurity dynamics way of thinking and landscape (invited paper). In *ACM Workshop on Moving Target Defense*, 2020.
72. S. Xu, X. Li, T. Parker, and X. Wang. Exploiting trust-based social networks for distributed protection of sensitive data. *IEEE T-IFS*, 6(1):39–52, 2011.
73. S. Xu, W. Lu, and L. Xu. Push- and pull-based epidemic spreading in networks: Thresholds and deeper insights. *ACM TAAS*, 7(3), 2012.
74. S. Xu, W. Lu, L. Xu, and Z. Zhan. Adaptive epidemic dynamics in networks: Thresholds and control. *ACM TAAS*, 8(4), 2014.
75. S. Xu, W. Lu, and Z. Zhan. A stochastic model of multivirus dynamics. *IEEE Transactions on Dependable and Secure Computing*, 9(1):30–45, 2012.
76. Shouhuai Xu. Cybersecurity dynamics. In *Proc. Symposium on the Science of Security (HotSoS'14)*, pages 14:1–14:2, 2014.
77. Shouhuai Xu. Cybersecurity dynamics: A foundation for the science of cybersecurity. In *Proactive and Dynamic Network Defense*, volume 74, pages 1–31. Springer, 2019.
78. Shouhuai Xu. Sarr: A cybersecurity metrics and quantification framework. In *Third International Conference on Science of Cyber Security (SciSec'2021)*, pages 3–17, 2021.
79. Shouhuai Xu, Wenlian Lu, and Hualun Li. A stochastic model of active cyber defense dynamics. *Internet Mathematics*, 11(1):23–61, 2015.
80. Shouhuai Xu and Moti Yung. Expecting the unexpected: Towards robust credential infrastructure. In *International Conference on Financial Cryptography and Data Security (FC)*, pages 201–221. Springer, 2009.
81. N. Xue, Y. Malla, Z. Xia, C. Pöpper, and M. Vanhoef. Bypassing tunnels: Leaking {VPN} client traffic by abusing routing tables. In *Proc. Usenix Security*, pages 5719–5736, 2023.
82. Ji Su Yoo, Alexandra Thaler, Latanya Sweeney, and Jinyan Zang. Risks to patient privacy: a re-identification of patients in maine and vermont statewide hospital data. *J Technol Sci*, 2018100901:1–62, 2018.
83. R. Zheng, W. Lu, and S. Xu. Active cyber defense dynamics exhibiting rich phenomena. In *Proc. HotSoS*, 2015.
84. R. Zheng, W. Lu, and S. Xu. Preventive and reactive cyber defense dynamics is globally stable. *IEEE TNSE*, 5(2):156–170, 2018.