A Survey on Security and Privacy Issues in Modern Healthcare Systems: Attacks and Defenses

AKM IQTIDAR NEWAZ, AMIT KUMAR SIKDER, MOHAMMAD ASHIQUR RAHMAN, and A. SELCUK ULUAGAC, Florida International University

Recent advancements in computing systems and wireless communications have made healthcare systems more efficient than before. Modern healthcare devices can monitor and manage different health conditions of patients automatically without any manual intervention from medical professionals. Additionally, the use of implantable medical devices, body area networks, and Internet of Things technologies in healthcare systems improve the overall patient monitoring and treatment process. However, these systems are complex in software and hardware, and optimizing between security, privacy, and treatment is crucial for healthcare systems because any security or privacy violation can lead to severe effects on patients’ treatments and overall health conditions. Indeed, the healthcare domain is increasingly facing security challenges and threats due to numerous design flaws and the lack of proper security measures in healthcare devices and applications. In this article, we explore various security and privacy threats to healthcare systems and discuss the consequences of these threats. We present a detailed survey of different potential attacks and discuss their impacts. Furthermore, we review the existing security measures proposed for healthcare systems and discuss their limitations. Finally, we conclude the article with future research directions toward securing healthcare systems against common vulnerabilities.

CCS Concepts: • General and reference → Surveys and overviews; • Security and privacy → Systems security; • Applied computing → Health care information systems;

Additional Key Words and Phrases: Medical device, health IoT, healthcare system, security and privacy

ACM Reference format:

1 INTRODUCTION
In recent years, the healthcare domain has experienced myriad advancements in terms of new technologies and treatment methods. Modern healthcare systems have changed the lives of patients and medical professionals in many respects. Today, different healthcare applications have been embedded in consumer devices to remotely collect physiological information of a patient and provide automatic treatment. For instance, smartwatches

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Association for Computing Machinery.
2637-8051/2021/06-ART27 $15.00
https://doi.org/10.1145/3453176

can monitor different body mechanisms like heart rate, as well as rhythm via electrocardiogram (ECG); smartphones can track physical activities and sleep apnea; and an implanted glucose monitor can automatically control the sugar level by injecting insulin into a patient. Moreover, the development of low-power wearable biosensors [1], implantable medical devices (IMDs) [2], ultra-low-power body area networks (BANs) [3, 4], Internet of Things (IoT) technologies [5–7], and numerous lightweight communication protocols [8] have helped to develop small-scale sense-actuate healthcare devices that can collect and send different physiological values (blood pressure, heart rate, etc.) from a patient to medical professionals remotely and instantly to provide better treatments. Indeed, the increasing popularity and diverse utilities of modern healthcare systems have made the healthcare industry grow at a massive rate. The global medical device market is forecasted to grow at a compound annual growth rate of 4.5% from 2018 to 2023 and is expected to reach $409.5 billion by 2025 [9].

In the ecosystem of medical devices and applications, modern technologies, such as implantable and wearable medical devices (IWMDs), biosensors, and BANs, have certainly enhanced overall healthcare systems for patients and medical professionals. However, these smarter and advanced healthcare systems are “more” complex in software and hardware. Although the adaptation of new technologies in the healthcare domain is at an early stage, several software and hardware defects have already been found, which can lead to possible malicious attacks [10]. The open source development platforms and continuous connectivity pave the way for the attackers to exploit the security and privacy in healthcare systems. In recent years, several healthcare security issues have been reported both in the media and the academic community. A story became popular in media that doctors disabled the wireless connectivity of a former U.S. vice president’s pacemaker to protect it from being hacked [11]. Adding to this story, researchers demonstrated several cyber attacks on commercial IMDs, including attack scenarios of remotely disabling and reprogramming the therapies performed by an implantable cardiac defibrillator (ICD) [12–14]. Moreover, healthcare/medical devices are remotely exploitable through the communication media [15] (Wi-Fi, Bluetooth Low Energy (BLE), Zigbee, Z-Wave, etc.), and attackers can easily eavesdrop on the communication channel to access the transmitted data [12]. The lack of standard practice, the need for timely security patches, and the push from the government to keep devices and applications secure exacerbate this situation. Because of the catastrophic health consequences, any security issue concerning healthcare systems should be addressed aggressively and proactively. Unfortunately, there is no comprehensive security solution available in the industry and research community to mitigate the emerging cyber attacks on healthcare systems. Researchers have proposed a few countermeasures (privacy-preserving communication protocols, encrypted databases, etc.) that cannot address the overall attack surface in healthcare systems. Therefore, the security and privacy in healthcare systems require an immediate attention of the security research community, medical device industry, and regulatory bodies [2].

Contributions. The aim of this survey is to provide a comprehensive overview of the security and privacy trends and emerging threats to healthcare systems to facilitate the understanding of the pressing security and privacy challenges. The contributions of this work are as follows:

- First, we provide a detailed overview of a typical healthcare system and discuss its components.
- Second, we explore different security and privacy goals for healthcare systems and discuss potential adversarial models.
- Third, we present a detailed taxonomy of the existing attacks in the healthcare domain by analyzing them as reported by the research community and industry. We also discuss the impact of these attacks based on common vulnerability scoring metrics.
- Fourth, we summarize the existing solutions that have been proposed to mitigate these attacks and identify the challenges faced by the research community to ensure security and privacy in healthcare systems.
- Finally, we articulate several open challenges and future research directions toward solving the security and privacy issues in healthcare systems.
Table 1. Comparison Among Our Survey and Existing Surveys

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rushanan et al. [16]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Ellouze et al. [17]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Zhang et al. [2]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Altawy and Youssef [18]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Rathore et al. [19]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Camara et al. [20]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Kim et al. [21]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Alemdar and Ersoy [22]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>David and Jeyachandran [23]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Patil and Seshadri [24]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Qayyum et al. [25]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Sametinger et al. [26]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Pantelopoulos and Bourbakis [1]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Razaque et al. [5]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Pramanik et al. [27]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Abouelmehdi et al. [28]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Habibzadeh and Soyata [29]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Islam et al. [30]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Kruse et al. [31]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Yaqoob et al. [32]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Nasiri et al. [33]</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Our survey</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

○ = No information provided; ◁ = partial information provided; ● = complete information provided.

Organization. The remainder of this article is organized as follows. We provide an overview of existing literature surveys on healthcare systems in Section 2. We present the architecture of a typical healthcare system and discuss its different components in Section 3. In the following section, we provide security and privacy goals for healthcare systems. In Section 5, we present a detailed taxonomy of existing security and privacy attacks on healthcare systems and summarize the impacts of these attacks based on common vulnerability metrics. In Section 6, we discuss existing approaches that have been proposed to secure healthcare systems by researchers. Limitations of current security solutions, requirements to form a secured healthcare system, and corresponding challenges are discussed in Section 7. Finally, we conclude the article in Section 8.

2 RELATED WORK

In recent years, several surveys have been conducted to review existing security and privacy attacks on healthcare systems. However, these works either focus on specific attacks or security solutions for specific devices without considering overall security and privacy issues in healthcare systems. In this section, we summarize these surveys and discuss their differences from our work.

Existing surveys. Existing surveys are mostly focused on security and privacy problems, major vulnerabilities, and solutions related to the privacy and safety issues of IMDs [2, 16–21]. Among these surveys, Rushanan et al. [16] extensively reviewed security and privacy problems corresponding to telemetry interfaces and software programs, security frameworks, and standard practices that aimed at improving the security of IMDs. Alemdar and Ersoy [22] evaluated the current research activities and issues that need to be addressed to enhance remote health monitoring. David and Jeyachandran [23] presented a survey on wireless medical sensor networks (WM-SNs), cryptographic approaches to preserve health data, and the trade-off between security and reliability of...
WMSNs. Sametinger et al. [26] reported the critical issues that were being faced to ensure the security of medical devices and provided an illustrative example. Several other useful surveys have focused on applications of big data in the modern healthcare systems [27, 28]. Patil and Seshadri [24] studied the state-of-the-art security and privacy issues in big data as applied to the healthcare industry. Qayyum et al. [25] presented an overview of various applications in healthcare that leveraged machine learning (ML) techniques from a security and privacy point of view and reported associated challenges. A comprehensive review of existing research and development on wearable biosensor systems for health monitoring was presented by Pantelopoulos and Bourbakis [1]. Razaque et al. [5] presented a flow of information in the healthcare domain with a particular focus on IoT connection. Habibzadeh and Soyata [29] studied the emerging trends in healthcare applications and discussed potential threats, vulnerabilities, and consequences of cyber attacks in healthcare systems. Kruse et al. [31] presented a survey on cybersecurity challenges in healthcare systems. Yaqoob et al. [32] demonstrated possible attack vectors and security vulnerabilities, and applicable attacks were demonstrated for networked medical devices. Researchers also analyzed IoT security and privacy features from the healthcare perspective. Islam et al. [30] surveyed advances in IoT-based healthcare technologies, network architectures, and industrial trends in IoT-based healthcare solutions. An overview of the features and concepts related to security requirements for IoT in a healthcare system was provided by Nasiri et al. [33].

**Differences from existing surveys.** The main differences between our work and existing surveys can be articulated as follows. First, although most current surveys are focused on the security and privacy of IMDs and IWMDs, our survey focuses on the overall healthcare system, which covers end-to-end components, including medical devices, sensors, networking/communication, and healthcare providers. Second, we provide a formal architecture on healthcare systems and identify its major components to outline security and privacy needs. Third, we categorize the existing security and privacy attacks on healthcare systems and use the common vulnerability scoring system to measure the impact of the attacks. Fourth, although the existing surveys are hardly focused on the limitations of current security solutions, our work identifies the limitations and discusses them. Fifth, our survey provides categorical directions for researchers to explore mitigation measures against common security vulnerabilities in healthcare systems. We present a comparison among the existing surveys and our survey in Table 1.

### 3 BACKGROUND AND DEFINITIONS

In this section, we provide a detailed overview of different components of healthcare systems to understand the significance of security and privacy needs in the healthcare domain. A healthcare system usually comprises one or more medical devices that are equipped with different sensors to collect patients’ vitals and makes autonomous decisions to provide enhanced treatments. The overall architecture of a healthcare system is shown in Figure 1. We identify five major components that are typically important to perform general functionalities of a healthcare system. These five components are as follows: medical device, sensor, networking, data processing, and healthcare provider.

**Medical device.** Any device, instrument, appliance, or apparatus that is intended for one or more medical purposes, such as diagnosis, monitoring, treatment, and alleviation, is called a medical device. According to the U.S. Food and Drug Administration (FDA), medical devices range from simple tongue depressors to complex programmable ICDs. The FDA provides classification standards for medical devices based on the potential risk of causing harm to patients in case of device malfunction or malicious attacks. Devices with a minimal level of risk and a minimum level of regulatory control, like elastic bandages and dental floss, belong to Class I medical devices. Pregnancy testing kits, powered wheelchairs, and so on are Class II devices, which are more complicated and riskier than Class I devices and require stringent regulatory controls. Class III devices, such as implantable pacemakers and breast implants, possess the highest risk and complexity, and they require highly stringent regulatory controls. In addition, the European Commission provides several other classification standards for medical devices based on non-invasive, invasive, and active therapeutic properties [34]:
Fig. 1. Overview of an example healthcare system.

(1) Non-invasive devices: Non-invasive devices are intended to use for body-liquid collection in such a way that return flow back to the human body is unlikely (e.g., urine collection bottles). In addition, this type of device only contacts the patient’s skin and is intended for channeling or storing blood, body liquids or tissues, liquids, or gases for eventual infusion (e.g., antistatic tubing for anesthesia, syringes for infusion pumps).

(2) Invasive devices: These types of devices are introduced into the body either through a break in the skin or an opening in the body. Invasive devices can be further categorized into four groups, namely transient use, short-term use, long-term use, and connected to an active medical device. Surgically invasive transient use medical devices (<60 minutes) are precisely controlled, directly contacted with the central nervous system, and reusable. Surgically invasive devices for short-term use (>60 minutes, <30 days) can be directly contacted with the central nervous system to precisely monitor, diagnose, or control the heart central circulatory system. Surgically invasive long-term use and active medical devices (>30 days) can be placed in the mouth, or have direct contact with the heart or central circulatory system to administer medicines.

(3) Active therapeutic devices: Active therapeutic devices are intended to administer or exchange data, whether used alone or in combination with other medical devices, to deliver to or remove medicines from the body. Examples of such devices include muscle stimulators, hearing aids, and therapeutic X-ray sources.

Sensor. In the healthcare domain, sensors are used to monitor and measure the patient’s vitals. Different physiological sensors (blood sugar sensor, heart rate sensor) are used as a trigger to automate different functionalities (diagnosis, monitoring, etc.) of healthcare systems. We divide sensors into the following three categories:

(1) Physiological sensors: These sensors measure physiological signals (e.g., ECG, electromyography) and features to give an overall estimation of the patient’s health condition at any given time.

(2) Biological sensors: These sensors integrate the biological elements in a human body with the physico-chemical transducer to produce an electric signal. Glucose and alcohol sensors are examples of this kind.

(3) Environmental sensors: These sensors can sense different environmental parameters to understand any change in the proximity of a patient. For example, an accelerometer and gyroscope in a smartwatch can detect a patient’s movement to measure motion and sleep data.

Networking. Networking components are concerned with how different medical devices and sensors communicate with each other, as well as with other components of a healthcare system. As Figure 1 illustrates, the transmission of measured data in a healthcare system needs to be performed primarily for two different purposes: (1) transferring the physiological signal from the sensors or devices to the system’s central node and (2) sending the aggregated measurements from the central node to or from a health server or healthcare professional. Transmission of data for the short range can be handled by wired or wireless channels. However, wired communication may hinder the patient’s mobility and comfort [35]. Autonomous sensor nodes can follow a primary star topology network to form a BAN for transmitting data to the central node of the BAN [36].
The most commonly used wireless communication standards in BANs are IEEE 802.15.1 (Bluetooth) and 802.15.4 (Zigbee), which are a part of the 802.15 working group for the wireless personal area network (WPAN). Bluetooth is an industry specification for short-range radiofrequency (RF)-based connectivity between portable and fixed devices. It is a low-power, low-cost RF standard, operating in the unlicensed 2.4-GHz spectrum [37]. It uses a frequency hopping technique (FHSS) over 79 channels in the industrial, scientific, and medical (ISM) band to combat interference and supports up to 3 mbps in the enhanced data rate mode with a maximum transmission distance of 100 m. The Zigbee standard also targets low-cost, low data-rate solutions with high battery life. It operates in 16 channels in the 2.4-GHz ISM band (250 kbps, OQPSK modulation), in 10 channels in the 915-MHz band (40 kbps, BPSK modulation), and in one channel in the 868-MHz band (20 kbps, BPSK modulation) [38]. Alternative technologies for short-range intra-BAN communication include infrared data association (IrDA), ultra-wideband (UWB), and medical implant communication service (MICS). UWB is a low-cost communication protocol for the short-range exchange of data over infrared light. MICS is an ultra-low-power, unlicensed, mobile radio service for transmitting low-rate data in support of diagnostic or therapeutic functions associated with medical devices. It uses the 402-405-MHz frequency band, with 300-kHz channels [39].

For long-range communication between a healthcare system and a health server or a healthcare provider, there is a wide variety of available wireless technologies (e.g., WLAN, GSM, GPRS, UMTS, WiMAX, LoRa), which can offer broad coverage and ubiquitous network access. Moreover, future advances in 5G mobile communication systems are expected to guarantee worldwide seamless access to the Internet at much higher data rates, providing the ability to collect data from remote medical devices in real time. More recently, with the advent of Z-Wave and BLE, more devices are expected to be in the market using these low-power communication protocols.

**Data processing.** The data processing component collects data from devices and sensors to produce meaningful information. A central data processing unit is shown in Figure 1, which communicates with the medical device and sensor components via communication and control modules. It has a data processing unit, along with a local database to save initial data about the patient. Its alarm generator informs the patient if there is any anomaly. It uses a wireless transmission module to make a connection with the health server and healthcare provider.

**Healthcare provider.** Health servers and healthcare professionals are elements of the healthcare provider component. They communicate with the data processing component through a wireless transmission module. The health server saves healthcare data in the cloud. Healthcare professionals can access this data to treat patients remotely or physically.

4 SECURITY AND PRIVACY NEEDS IN EXISTING HEALTHCARE SYSTEMS

To discuss the security and privacy issues in healthcare systems, we refer to Figure 2 as the use case scenario, which is a complex multidisciplinary and integrated healthcare system. In this section, we first present security and privacy requirements, and then we review the corresponding security and privacy goals.

4.1 Security and Privacy Requirements

Figure 2 presents the general security and privacy goals of a healthcare system. Here, a patient carries several invasive and non-invasive medical devices that are placed on or around the patient’s body to monitor constantly various vital signs of the body (e.g., ECG signal, pulse, blood pressure) and important environmental parameters (e.g., ambient temperature and humidity). The sensor readings and patient profiles are together called *patient-related data* that is collected and transmitted to other devices like smartphones and computers. These devices can perform further data processing, aggregation, or distributed storage. The patient-related data can also be sent to a central healthcare server for permanent records and to healthcare professionals and the hospital for continuous monitoring of the patient’s physical condition. In summary, the overall architecture of a personal healthcare system is divided into three tiers. Tier 1 consists of medical devices, including invasive and non-invasive devices, and Tier 2 has personal devices like smartphones and computers. Health servers and healthcare professionals form the third tier.
For ensuring security, authentication is required in Tiers 1, 2, and 3. Before transmitting any patient-related data to a personal device or from a personal device to a health server, users must be strictly authorized at each tier. Device information and data should only be accessible by authorized healthcare professionals and should not be modifiable by unauthorized users. Confidentiality and integrity should also be ensured at Tiers 2 and 3. As medical devices perform various sensitive operations and handle personal data, this information should be kept confidential in an access log. Medical devices should be reachable all the time, as unavailability of the device data may impact the treatment of the patient. Non-repudiation and availability should be maintained in all three tiers.

Furthermore, to achieve privacy goals, one needs to maintain device anonymity, which means that only the patients and authorized users should know what medical devices a specific patient is carrying. For transmitting data from Tiers 1 to 2 and 2 to 3, patients and doctors’ real information should not be disclosed to maintain data anonymity. Communication between the patients and healthcare professionals/hospitals via the healthcare system should be untraceable for adversaries to achieve communication anonymity and unlinkability.

4.2 Security and Privacy Goals

We identify the following security goals for healthcare systems based on the previous discussion. To maintain these goals, the following properties should be considered throughout the life cycle of a healthcare system:

1. **Authentication**: Strong authentication is a fundamental component for securing healthcare systems where one needs to consider the environment setup, single-factor vs. multi-factor authentication [40], grace periods, and emergency situations. Most of the current networked medical devices (e.g., IMDs, IWMDs) have weak password authentication schemes where password files are stored on the local hard drive [41, 42]. As a result, an attacker with privileges can delete/modify the files or install new software on the device.

   a. **Environment considerations**: Different healthcare settings have different architectures, and one needs to select the appropriate authentication mechanism accordingly. For example, proximity cards may be suitable and easily accessible for regular patient interaction, but not for authenticating operating rooms.

   b. **Single vs. multi-factor authentication**: Before the period of access control evaluation, it is important to consider where to employ single or multi-factor authentication. For example, one-factor authentication may be sufficient for blood pressure or temperature readings, but accessing data from healthcare servers may require multi-factor authentication.

   c. **Continuous authentication**: Healthcare professionals need to use user credentials repeatedly while accessing the patients’ data. One existing solution is establishing a grace period after building the initial trust. However, the grace period may lead to malicious scenarios, as any unauthorized personnel can access the device within the grace period. One possible solution is continuous authentication [43]
that can be achieved by implementing different methods, such as wearable-assisted or sensor-assisted authentication.

(d) **Emergency considerations**: One needs to consider multiple scenarios for accessing medical devices in the event that one method is not available. For example, a medical device may be considered only to transmit data after authentication but may still be allowed to access in any emergency scenario related to the patient’s health concerns [44].

(2) **Confidentiality**: Device information, system configuration, and healthcare data should be accessible only by authorized personnel or entities. These entities needed to be authenticated before accessing any healthcare-related confidential information. However, it is possible to eavesdrop on existing healthcare devices, such as an insulin pump communication channel, and get patient data and device-related information [14].

(3) **Integrity**: The data, device information, and system configuration should not be modifiable by unauthorized users’ devices or applications. For instance, if there is no integrity-checking mechanism in medical devices, data can be altered, and medical devices might accept malicious inputs, which can lead to severe attacks such as a code injection attack [45]. Current fitness tracker devices (e.g., Fitbit Charge, Garmin Vivosmart) lack the integrity check mechanism for firmware updates [46, 47].

(4) **Non-repudiation**: A healthcare system performs different operations, and this information is usually kept secret in an access log. Any modification in this log should be traced and monitored and only performed by verified users. The attackers might want to delete these logs to cover their traces. For instance, the Fitbit smartwatch keeps daily logs in clear-text files where an unauthorized user can change the log file to reverse engineer the communication protocol without keeping any trace [48]. There are many resource-limited medical devices where there is no log in the systems, and attackers might try to gain access to the system without leaving any footprints.

(5) **Availability**: The service provided by the healthcare system should always be available to authorized users for accessing device systems, as well as patients’ data in normal or emergency situations. For instance, an implementation flaw was found in an ICD, which does not let the device go into sleep mode when a communication session ends [49]. Such a flaw can be exploited to trigger denial-of-service (DoS) attacks, thus making the device unavailable.

In addition to satisfying the preceding security goals, one should ensure necessary privacy requirements in healthcare systems. In this work, we consider the privacy goals based on device, data, and communication anonymity properties [50]:

(1) **Device anonymity**: This means that the identity of a medical device is unknown to the system so that unauthorized entities should not be able to determine the existence of the device type, specific device ID, and traditional identifiers such as IP and MAC addresses.

(2) **Data anonymity**: The goal of data anonymity is to prevent unauthorized users from identifying a user and the user’s sensitive data. In lieu of any obvious identifiers, patients and doctors could employ pseudonyms or other temporary identifiers for protection [51].

(3) **Communication anonymity**: Unauthorized entities should not be able to identify the connection between the user and the system. Effective necessary mechanisms, such as collision-resistant pseudonyms [52] should be used to ensure anonymous communication.

(4) **Unlinkability**: An attacker who tracks the data transactions between the sender and the receiver should not be able to establish a relationship between data and sender.

5 **ATTACK MODEL AND EXISTING SECURITY AND PRIVACY ATTACKS ON HEALTHCARE DEVICES AND APPLICATIONS**

As existing healthcare devices and applications fail to meet the security and privacy requirements (as discussed in Section 4), attackers can exploit different components of healthcare systems to perform malicious activities.
In this section, we explain attack goals considering the capabilities of an attacker and attack methodologies to perform different attacks on healthcare systems. Moreover, we discuss various attacks on different components of healthcare systems (e.g., sensor, device, network) and summarize how attackers can compromise the security and privacy of targeted healthcare systems. We formally categorize existing security and privacy attacks on healthcare systems reported by the research community and developers and explain the attack methods and effects in detail.

**Attacker goals.** An attacker can target a medical device to perform various malicious activities, including communication interception, data modification, and device or data unavailability. We categorize attack goals in the following categories based on the attack impacts on the healthcare devices and patients:

1. **Hardware modification:** An attacker tries to tamper with the device hardware architecture so that he/she can insert malicious hardware Trojan during chip manufacturing time.
2. **Unavailability:** An attacker seeks to use malicious written programs to perform different attacks (e.g., malware, ransomware) and make the device and data unavailable (may be until a ransom is paid).
3. **Communication delay:** An attacker attempts to connect with the healthcare device using an unauthorized programmer device (e.g., smartphone, personal computer) and forces the device to continue communication with an unauthorized programmer.
4. **Data sniffing:** An attacker tries to capture the communication of healthcare devices for collecting sensitive information such as the patient’s vital signs and device information.
5. **Data modification:** An attacker attempts to modify the patient’s vital signs by breaking into the device or intercepting and modifying the communication packet between the healthcare and the programmer device.
6. **Information leakage:** An attacker tries to retrieve confidential and sensitive information from healthcare systems. For instance, he/she can extract secret cryptographic keys, device power consumption, and personal information (e.g., bank card, PINs) using several methods, such as statistical analysis, electromagnetic interference (EMI) radiation, and malicious software applications.

**Attacker capabilities.** We consider the following capabilities for an attacker to successfully implement different attacks on healthcare systems:

1. An attacker has physical and/or remote access to the healthcare systems’ environment.
2. An attacker has the knowledge of which communication standard and protocol are used by the healthcare devices to establish communication with the programmer device.
3. An attacker can access communication channels using third-party devices (e.g., sniffer, off-the-shelf hardware and software).
4. An attacker can use a programmer device to impersonate herself/himself as a patient to collect sensitive information from the healthcare device.

**Attack types.** Depending upon their goals, capabilities, and relationship to the system, adversaries in healthcare systems can be categorized as follows:

1. **Passive adversary:** An adversary of this kind can eavesdrop on communication channels, including side channels or unintentional communication channels, without interrupting the communication. It is a direct threat to confidentiality and authentication for an insecure communication channel. By reading messages only, she/he may determine whether a person carries any medical device or not, find out the device model, capture telemetry data, and disclose private information. Recently, the value of personal health information in underground markets has been rising. If no authentication mechanism is enforced, then the adversary can obtain private information such as the surgery type and social security number related to the patient.
2. **Active adversary:** Such an adversary can interrupt the communication channel and read, modify, and inject data. The adversary can be capable of capturing messages exchanged over the radio channel.
Table 2. List of Existing Attacks to Healthcare Devices and Applications

<table>
<thead>
<tr>
<th>Attack</th>
<th>Attack Type</th>
<th>Target Medical Device</th>
<th>Target Component</th>
<th>Vulnerability Metrics††</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hardware</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malware</td>
<td>Active therapeutic</td>
<td>Device, data,</td>
<td>Active Low</td>
<td>✓ ✓ I, A</td>
<td>[58–60]</td>
</tr>
<tr>
<td>Ransomware</td>
<td>Active therapeutic</td>
<td>Data, healthcare</td>
<td>Active Low</td>
<td>✓ ✓ I, A</td>
<td>[61–65]</td>
</tr>
<tr>
<td>Outdated operating systems</td>
<td>Active therapeutic</td>
<td>Device, data,</td>
<td>Passive High</td>
<td>✓ ✓ I, A</td>
<td>[66], [42]</td>
</tr>
<tr>
<td>Electroencephalography</td>
<td>Non-invasive</td>
<td>Device</td>
<td>Passive Low ✓</td>
<td>✓ C</td>
<td>[67]</td>
</tr>
<tr>
<td>Counterfeit firmware update</td>
<td>Invasive, non-invasive</td>
<td>Device, healthcare</td>
<td>Passive High ✓</td>
<td>✓ ✓ I, A</td>
<td>[44], [46, 47, 68–73]</td>
</tr>
<tr>
<td><strong>System level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak authentication schemes</td>
<td>Invasive, non-invasive</td>
<td>Device, data,</td>
<td>Passive High ✓</td>
<td>✓ ✓ C, I</td>
<td>[41, 42], [24, 48, 74–86]</td>
</tr>
<tr>
<td>Privilege escalation</td>
<td>Invasive</td>
<td>Device, data</td>
<td>Passive Low ✓ ✓</td>
<td>✓ ✓ I, A</td>
<td>[87]</td>
</tr>
<tr>
<td>Sensor spoofing</td>
<td>Invasive</td>
<td>Sensor</td>
<td>Active High ✓</td>
<td>✓ ✓ A</td>
<td>[91]</td>
</tr>
<tr>
<td>Differential power analysis</td>
<td>Non-invasive</td>
<td>Device</td>
<td>Passive High ✓</td>
<td>✓ ✓ I, A</td>
<td>[92]</td>
</tr>
<tr>
<td>Eavesdropping</td>
<td>Invasive, non-invasive</td>
<td>Network</td>
<td>Passive Low ✓</td>
<td>✓ ✓ C, I</td>
<td>[14], [93–106]</td>
</tr>
<tr>
<td>Replay</td>
<td>Non-invasive</td>
<td>Network</td>
<td>Active Low ✓</td>
<td>✓ ✓ C, I</td>
<td>[94], [107], [74]</td>
</tr>
<tr>
<td>Impersonation</td>
<td>Non-invasive</td>
<td>Network</td>
<td>Passive High ✓</td>
<td>✓ ✓ I</td>
<td>[14], [108]</td>
</tr>
<tr>
<td><strong>Communication channel</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DoS</td>
<td>Invasive, non-invasive</td>
<td>Network</td>
<td>Active Low ✓</td>
<td>✓ ✓ A</td>
<td>[112], [49, 109–112], [63], [78]</td>
</tr>
<tr>
<td>Multiple input and multiple output</td>
<td>Invasive</td>
<td>Device</td>
<td>Passive High ✓</td>
<td>✓ ✓ C</td>
<td>[113]</td>
</tr>
<tr>
<td>Man-in-the-middle</td>
<td>Invasive Devices,</td>
<td>Network</td>
<td>Active High ✓</td>
<td>✓ ✓ C, I</td>
<td>[102], [114–119]</td>
</tr>
<tr>
<td>Battery depletion</td>
<td>Invasive, non-invasive</td>
<td>Device</td>
<td>Active Low ✓</td>
<td>✓ ✓ A</td>
<td>[49, 109], [120–122]</td>
</tr>
</tbody>
</table>

†† Impacted security: Integrity (I), availability (A), confidentiality (C).

Vulnerability metrics: Attack approach (AA), attack complexity (AC), privilege requirement (PR), user cooperation (UC).

Responding attacks may involve a sequence of interceptions, modifications, interruptions, and generations of extra messages to achieve different goals. Moreover, an active attacker can impersonate a programmer medical device (e.g., smartphone, personal computer), which is a third-party device used in IMDs. It can request confidential information, reprogram the medical device, cause a shock to the patient, or drain the battery of the medical device. An adversary may track a patient (e.g., patient’s location, diagnosis, blackmail-worthy material) so that he can cause physical or psychological harm.

It is worth noting that it is not essential for an attacker to be close to the healthcare devices to conduct an attack. An adversary can be an external or internal entity with respect to the system. The adversary can also be a manufacturer, a patient, a physician, or even a hospital administrator.

To understand the effect of the attacks on real-life healthcare systems, we introduce a vulnerability metrics based on a widely accepted measure of the common vulnerability scoring system (CVSS) to quantify the impact of these attacks [53]. We consider the following vulnerability metrics to illustrate the severity of different attacks on healthcare system:
A Survey on Security and Privacy Issues in Modern Healthcare Systems: Attacks and Defenses

(1) **Attack approach (AA):** This reflects how an attacker exploits a healthcare system to perform malicious activities. Based on the attack approaches, it can be categorized as follows: active attack and passive attack. Passive attacks refer to an attack that performs malicious activities in a healthcare system without obstructing the normal operation of the system, whereas active attacks obstruct the normal operation of a healthcare system to perform malicious activities.

(2) **Attack complexity (AC):** This metric specifies the amount of information an adversary needs to perform an attack on a healthcare system. An attacker needs partial (e.g., device model, used communication protocol) or full information (e.g., network structure, encryption type, etc.) of healthcare devices to perform an attack. For instance, a **man-in-the-middle (MITM)** attack needs physical access (high complexity) to the network, whereas a replay attack can be performed by capturing the communication packet passively (low complexity) and sending the same packet repeatedly.

(3) **Privilege requirement (PR):** To perform an attack, the attacker needs certain privileges or access to the healthcare system. We use the required privilege of an attacker to the system to explain the impact of the attack. For example, a communication medium attack such as packet sniffing does not need any access to perform malicious activities, whereas an impersonation attack requires access to the healthcare system.

(4) **User cooperation (UC):** An attack may require human interaction other than the attacker to exploit the vulnerability successfully. For instance, to install malware, user interaction is needed in the healthcare system.

In the following sections, we group the existing attacks according to their relevant attack surfaces and provide an explicit categorization of the attacks. We also present a detailed summary in Table 2.

### 5.1 Hardware Attacks

Hardware attacks refer to an exploitable weakness in a device hardware that can be used to gain physical or remote access to the device to perform malicious activities. An attacker may know or gain access to the internal hardware architecture of the device and insert **hardware Trojans (HTs)** during chip manufacturing that can lead to data corruption, causing serious harm to the medical devices [54]. Indeed, HTs have emerged as a major security concern for **integrated circuits (ICs)** as most ICs are manufactured in outsourced fabrication facilities. Third-party vendors can include unverified intellectual property cores that act as HTs and can be used to perform malicious activities, including leaking information from the medical devices. HTs can be classified based on **physical attributes** (e.g., chip layout, activation) and **action characteristics** (e.g., logic functions, chip activities) [123]. Physical attributes describe the Trojans that can be injected through the addition or deletion of transistors or gates in the chip manufacturing stage. Action characteristics refer to an HT where it changes the chip’s function by adding or bypassing existing logic.

The FDA has released numerous reports on changing a patient’s health data by modifying medical device hardware [55, 56]. In recent work, researchers presented an HT attack on the bacillus Calmette-Guerin (BCG) scale [57]. They injected a malicious payload that modifies the logic of an XOR gate on the input bus. It is a relatively less severe HT attack and cannot be detected if spread out among thousands of gates in the embedded system of the healthcare device.

### 5.2 Software Attacks

Software attacks refer to maliciously written programs to deliberately impact healthcare devices, associated computers, or servers. The use of embedded and customizable software in a healthcare environment is increasing rapidly, which certainly improves the patients’ treatment and monitoring. However, there is no satisfactory security measure to verify the functionalities and authenticity of medical software. As a result, healthcare systems are facing various threats related to software and apps, such as malware, ransomware, outdated **operating systems (OSs)**, counterfeit firmware update, and **electroencephalography (EEG)** attacks.
Malware. Any software or application that is written with malicious intent is called malware. A healthcare device infected with malware can stray away from its normal functionalities, such as slowing or shutting down a device. For instance, Conficker, a relatively old malware, was recently detected on 104 devices, including X-ray machines, mammography, and a gamma camera for nuclear medicine at the James A. Haley Veteran’s Hospital in Tampa, Florida [58]. This malware affected the Microsoft Windows OS from a thumb drive because the network drivers were not patched with the MS08-067 patch from Microsoft. Hence, a remote and unauthenticated attacker could execute arbitrary code on the vulnerable system. In January 2010, a Veterans Affairs catheterization laboratory in New Jersey was temporarily closed due to a malware infection into the computer systems [59]. As a result, patients were unable to get any medical services from that hospital. Affected devices were X-ray machines and lab equipment manufactured by reputed companies. Moreover, malware like Kwampirs can introduce instability into healthcare systems by triggering equipment malfunction or delay in accessing information [60].

Ransomware. Ransomware is a unique subset of malware that limits or blocks users’ access by locking the system and data unless a ransom is paid. In May 2017, around 50 hospitals in the United Kingdom were directly affected, and many hospitals preemptively shut down their computer systems due to ransomware. It caused considerable disruption, such as affecting care delivery, compromising patient safety, and potentially eroding trust [61]. The ransomware encrypted and blocked the patient’s data and threatened to publish or delete them unless a ransom was paid. In 2016, a ransomware shut down the network of the Hollywood Presbyterian Medical Center in Los Angeles, California, for 10 days. It prevented its staff from accessing medical records or using medical equipment until the hospital paid a ransom of $17,000 [62]. Freedom of information requests in the United Kingdom found that in 2015–2016, up to half of the national health service (NHS) trusts were hit by ransomware [62]. In addition, two U.S.-based health centers (Hancock Health and Erie County Medical Center) were hit by SamSam ransomware and ended up meeting the ransom demand. In all of these incidents, on average, it would take 12 days to restore limited system access and 6 weeks to restore full access to the system [63].

Recently, a new ransomware called Zeppelin was reported in healthcare companies across Europe, the United States, and Canada [64]. Zeppelin is a Delphi-based highly configurable ransomware that could be deployed as an .exe, a .dll, or wrapped in a PowerShell loader. This ransomware employed a standard combination of symmetric and asymmetric encryption with randomly generated keys for each file (advanced encryption standard (AES)-256 in cipher-block chaining (CBC) mode).

Outdated OSs. Outdated OSs poses severe threats to healthcare devices as newfound bugs are not addressed in the older versions of the OS by the vendor. As a result, attackers can exploit the existing bugs of the OSs by simply injecting malicious code snippets or software. According to the Duo security research team, 70% of healthcare devices in North America and Europe were still running on outdated Windows 7 OS at the end of 2020, although Microsoft stopped releasing any patches for Windows 7 [66]. As an example, the WannaCry ransomware attacks were launched against unpatched healthcare devices, where information technology professionals neglected to download the OS update on time. A group of researchers conducted a vulnerability assessment in a radiology department, where the majority of networked medical equipment (e.g., medical ventilator, X-ray machine, anesthesia machine) were running on an old and insecure version of OSs (e.g., Windows Vista, Windows XP) [42]. These OSs were running unprotected, insecure, and vulnerable applications that had no firewall or protection against malware.

Counterfeit firmware update. Counterfeit firmware in medical devices introduces numerous threats to healthcare devices, as an attacker can gain access to the devices and manipulate the applications using fake copies of firmware. Counterfeit firmware is produced and distributed in such a way that it appears to be authentic. Hanna et al. [41] analyzed an automated external defibrillator (AED) (Cardiac Science G3 Plus 9390A) and found four vulnerabilities, including a software update mechanism that accepts counterfeit firmware. An attacker could replace the firmware of that specific AED with custom firmware designed to exploit the AEDUpdate package to perform buffer overflow. Rios and Butts [68] showed that it is possible to update unverified firmware of a home
monitoring device that is connected to an ICD. As there was no digitally signed firmware within the ICD ecosystem, the unverified firmware allowed performance of an MITM attack on the ICD. Ricke [69] reverse engineered a wearable healthcare device (Withings Activite fitness tracker) to identify and reconstruct the header structure within a firmware update. As the authentication scheme of this device only computes a checksum over the actual content of the image, it is possible to create a fake copy of the firmware by alternating the checksum in the header field. In other studies [46, 47, 70–73], fitness tracker devices’ application codes and firmware were reverse engineered to extract and manipulate fitness-related data. These devices lacked authentication, encryption, and integrity check mechanisms for firmware updates.

EEG attacks. In an EEG-based attack, an attacker is a malicious third-party application developer who is using an EEG-based brain-computer interface (BCI) device. The main goal of this device is to learn secret and private information about the user. An attacker developed a malicious software called brain spyware that was integrated into a BCI device to detect private information of the user [67]. Moreover, an attacker can specially design the videos and images that can be shown to the user to maximize the information leakage from the BCI device during the time of the attacks. It has been shown that the captured EEG signal can reveal personal information (e.g., bank cards, PINs, where one lives). In other studies [98, 99], researchers used brain spyware to extract not only private information about users’ memories and prejudices but also their possible neurological disorders.

5.3 System-Level Attacks

System-level attacks directly focus on system-level vulnerabilities such as memory modules, system applications, and design flaws in a healthcare system. Attackers can exploit these vulnerabilities to gain unauthorized control and access to sensitive data. There are two major types of system-level attack that can be performed on a healthcare system. These are weak authentication scheme exploitations and privilege escalation attacks on healthcare devices.

Weak authentication scheme exploitations. Authentication is a process where one needs to prove his/her identity to an application or system to access a service. Weak authentication describes a scenario where the strength of the authentication mechanism is relatively weak compared to the value of the assets. In a recent study, researchers investigated weak password-based authentication in healthcare devices focusing on external and internal defibrillators. According to the study, an AED using MDLink software has a weak password authentication scheme where the password file is stored on the local hard drive [41]. As a result, anyone with privileges can delete or change the password file and install any new software on the machine. Furthermore, researchers reverse engineered the MDLink authentication mechanism and wrote a small utility to change or recover a user’s password.

In another work, Xiao et al. [74] used a malicious BCI app to steal the patient’s EEG data by exploiting the standard software development kit (SDK), as the calling application programming interfaces have no authentication schemes [74]. In 2017, a group of researchers identified a hard-coded authentication system in the Medfusion 4000 Wireless Syringe Infusion Pump from Smiths Medical that allowed FTP server connections without any verification [75]. In a recent article, researchers reported an EEG-based application that allows execution of malicious code on the EEG device. When a client requests an EEG file, this application exploits the path requested by the client using the buffer overflow to remotely access the EEG device and make the device unavailable [83].

Most networked medical devices pose weak authentication schemes during the time of reading or writing data from these devices. Moses and Korah [42] conducted a study on onsite networked medical devices in a radiology department to identify vulnerabilities by using a port scanner and a network vulnerability scanner. This study reported that around 85% of the networked medical devices allowed unauthorized users to read or write data from a portable USB storage medium. Moreover, a CD or DVD drive in 17 out of 31 networked items allowed unauthorized users to copy or upload data from the equipment. Researchers from CyberMDX studied the improper authentication vulnerabilities in GE Aestiva and Aespire Anesthesia devices [77]. In this work, researchers used serial devices to connect to a TCP/IP server via an unsecured terminal that allowed remote access to modify...
device configuration and disable alarms. Security researchers of Alfonso Powers and Bradley Shubin studied connected cardiology devices made by Change Healthcare [79]. They have reported insecure file permission in the default installation that might allow an attacker with local system access to execute unauthorized arbitrary code. Security researchers of Philips reported improper authentication and missing sensitive data encryption vulnerabilities in medical image management systems that could enable an attacker to see usernames, passwords, and personal data [81]. In some cases, weak authentication in the image management allowed direct access to the memory locations to execute arbitrary code, alter the intended control flow, or cause the system to crash. Mahler et al. [82] reported authentication flaws in medical imaging devices such as magnetic resonance imaging (MRI) or computed tomography (CT) machines. The scan configuration file inside a host controller PC of imaging device used a default username and password that allowed an attacker to manipulate the file to change the CT’s behavior or control the entire CT operation [82].

Several Medtronic ICDs and associated equipment use a conexus RF protocol that does not have any authentication or authorization [80]. This allows a nearby attacker to access any implantable cardiac devices with a radio turned on. An attacker can inject, replay, modify, and/or intercept data within the telemetry communication using unauthorized access. Rahman et al. [48] reported that the Windows supported Fitbit application stored its daily logs containing data requests, responses, and social network data in clear-text files without any authentication scheme. Researchers reverse engineered the communication protocol (ANT) of Fitbit and demonstrated active and passive attacks using the off-the-shelf software module.

Researchers of CyberMDX also discovered two security vulnerabilities in the firmware and web management of Alaris Gateway Workstations (AGWs) that were used to provide mounting, power, and communication support to infusion pumps [76]. AGWs were vulnerable to an exploit where an attacker could remotely exploit firmware files, which required no special privileges to execute. Moreover, an attacker could manipulate gateway communication with connected infusion pumps. As the web management system did not require any credentials or passwords, attackers could easily connect to a workstation using an IP address and monitor the infusion pump’s status, event logs, an so on. McMahon et al. [78] showed that an earlier version of Dropbear SSH Server used by several IMDs provided memory access to any users without any proper authentication. An attacker could get local access to the process memory by simply running a trace with a -v option that might disclose sensitive information of the patient held on the database. D’Orazio and Choo [84] reverse engineered a recent version of the iOS healthcare app that stored its patient information in an encrypted database (VaultDataModel.sqlite) without providing any authentication mechanism. Researchers examined the database with the SQLite DB browser and found sensitive patient information such as the patient’s name, healthcare card number, and healthcare services cost. Researchers of vpnMentor discovered an improperly configured MongoDB healthcare database of a prescription medication, which was left open and exposed to allow access by anyone on the Internet [85]. The data included the information (e.g., patients’ names, addresses, phone numbers, email addresses) of more than 78,000 patients. In another study, security researchers of UpGuard reported an exposed Amazon S3 bucket containing a healthcare vendor’s spreadsheet with a default username and password that appeared to have potentially sensitive files such as business data, medical data, and insurance data [86].

Privilege escalation attacks. A privilege escalation attack takes advantage of bugs, design flaws, or configuration failures in an OS or application to access healthcare devices and data that usually require exclusive permission or authorization. Privilege escalation attacks can be launched by rogue users (e.g., patients or physicians) who have access to healthcare systems and perform malicious activities, such as calibration failures and data modification. Yan et al. [87] introduced two types of privilege escalation attacks (pressure-based attacks and time-based attacks) to IMDs, which can mislead the diagnostic process by altering collected and stored data after bypassing the initial access control mechanism. In pressure-based attacks, the attacker can change the pressure value of the sensors connected to the IMDs to report misdiagnosis of the patient. The attacker postpones the pressure data of some pressure sensors for certain time slots in time-based attacks.
5.4 Side-Channel Attacks

Side-channel attacks aim at extracting sensitive data (ongoing task, encryption method, etc.) from a healthcare system/device by analyzing physical parameters without interrupting the ongoing task. Examples of physical parameters include how the circuit works, what data it is processing, and when a victim’s device is being used, among others. There are three major types of side-channel attacks on healthcare systems. These are EMI, sensor spoofing, and differential power analysis (DPA) attacks on medical devices.

**EMI attacks.** EMI attacks are performed by measuring the electromagnetic (EM) radiation emitted from a device and performing signal analysis to infer sensitive information. Kune et al. [88] showed that analog sensors used in medical devices (e.g., infusion pumps, ICDs) are sensitive to EMI and can provide an unchecked entry point into the medical devices [88]. Here, an attacker can inject EMI signals in an ICD’s sensing unit to alter the sensor readings and trick the medical devices to prevent data communication. Additionally, several prior works reported that EMI could cause device malfunction in pacemakers and ICDs [89, 90].

**Sensor spoofing attacks.** In a sensor spoofing attack, an adversary alters the physical environment in a way so that a medical system behaves abruptly. Park et al. [91] introduced a sensor spoofing attack against the infrared drop (ID) sensor embedded in the infusion pump. An ID sensor has a linear property of input-output stimuli, which can be manipulated to non-linear behavior by exceeding the upper bound operating region of the infusion pump. Researchers showed that an attacker could inject an external power signal to a targeted sensor to block the sensor response to environmental changes, which results in over-infusion or under-infusion to the patient. Over-infusion allowed the infusion pump to infuse about 333% of the fluid as compared to the normal operation, whereas under-infusion infused approximately 45% less than the normal operation.

**DPA attacks.** DPA attacks use different analysis techniques (statistical, error correction, etc.) to infer sensitive information from power consumption data. Zhang et al. [92] introduced a DPA attack that can extract secret keys from extremely noisy channels in a heart rate monitor using a symmetric cipher. Here, the heart rate monitor uses AES encryption to encrypt the measured heart rates before transmitting to an end device (hub or storage). An attacker can recover the secret key used in the encryption scheme by analyzing the current consumption rate while measuring the heart rate of the patients. If the same key is used in the same model of all heart rate monitor devices, an attacker can publicize the inferred secret key and thus make the cryptographic protection ineffective for a large number of devices.

5.5 Attacks via Communication Channel

Wireless communication is used for the connectivity among healthcare devices for remote monitoring, diagnosis, treatment, and emergency support. For healthcare devices, attacks through communication channels have become a major concern, as attackers can perform various attacks including eavesdropping, replay, impersonation, DoS, multiple input and multiple output (MIMO), MITM, and battery depletion attacks to compromise the integrity of the device operation.

**Eavesdropping.** Eavesdropping refers to an attack where an adversary tries to steal information over a communication medium by taking advantage of the unsecured communication channels. Several eavesdropping attacks on medical devices (e.g., Medtronic and OneTouch Ping insulin pumps) have been reported, which captures the cleartext communication to capture sensitive patient data such as blood glucose results and insulin dosage [93, 94]. Li et al. [14] demonstrated an eavesdropping attack in an insulin pump by using off-the-shelf hardware and a software radio platform. As the communication channel does not use any authentication, researchers showed that attackers could capture glucose level, device type, device PIN, and medical condition of the patient by eavesdropping on the communication channel. A group of researchers was able to capture enough sensitive data from Withings Blood Pressure Monitors’ network traffic to determine the time and frequency of blood pressure testing on a patient [105]. As the information sections of all queries and responses in Withings devices are transmitted
in clear-text format, an attacker can easily monitor and capture network traffic to eavesdrop on sensitive data including the device ID, device type, and patient’s readings.

Wearable healthcare devices, such as smartwatches and fitness bands, are also vulnerable to passive eavesdropping attacks, as attackers can capture network traffic and sensitive data by simply using a sniffer module. Cusack et al. [96] used a BLE sniffer to capture communication packets of four wearable devices using BLE 4.0 and 4.2 (e.g., Fitbit Charge HR, Samsung Gear3) and performed packet analysis to extract sensitive information [96]. The captured packets were uploaded to Wireshark for further analysis, and researchers found that sensitive information, such as the long-term key to the BLE pairing process, sender, and receiver MAC address, and communication messages were transmitted as plain text. Traffic from two wearable smartwatches (TW64 and Mambo HR) was remotely sniffed and analyzed using TI SmartRF and BLETestTool [100]. Attackers could remotely control these two smartwatches, such as making them vibrate for a long time, possibly by keep sending fake command messages. These two devices had no technical security protection mechanism at all. Fawaz et al. [102] performed a passive eavesdropping attack on BLE-enabled healthcare devices by sniffing the communication over advertisement channels. Further analysis of these captured packets revealed that BLE-enabled healthcare devices use a fixed Bluetooth address for long periods, making the address randomization process ineffective. Hence, an attacker could sniff and capture sensitive health data of a patient without any interruption for a long period. Furthermore, the authors recovered the original Bluetooth signal from the jammed signal using a MIMO receiver that contains detailed information of the patient’s vitals. Lofty and Hale [104] captured the network traffic between a smartwatch and smartphone and showed that it is possible to convert the HEX-encoded data to human-readable data using reverse-engineering techniques. A passive attack was accomplished to sniff the internal LAN on an infusion pump, which was integrated into the information technology networks [95]. In this work, researchers found an open port in the infusion pump unit where the default password setting was not changed. Moreover, the information on the correct login was not monitored, and the communication was unencrypted.

Kim et al. [101] showed that it is possible to infer the encryption key by capturing the vibration of a smartphone while transmitting to an IWMD [101]. This vibration of the smartphone also leaked an audible acoustic signal that was captured using a microphone. The recorded acoustic signal was highly correlated to the vibration waveform that could effectively block the transmission of the encryption key. Halevi and Saxena [103] reported that auxiliary audio channels could be breached by close-range eavesdropping. Here, researchers eavesdropped on the IMD key pairing process and detected the initial sequence of the secret key using a signal processing algorithm. Moreover, they extracted spectrum features from each consecutive bit and used these features as input to ML algorithms for classifying each bit value. Li et al. [106] showed that the communication between prosthetic limb application and neural implant devices could be eavesdropped on to capture brain neural signals, decompose raw signals, and obtain users’ private information. Furthermore, an attacker can get control of prosthetic limbs of patients and give dangerous movement to patients without being in the close proximity of the victim. 

**Replay attacks.** A replay attack is a form of attack in which an adversary intercepts the data transmissions and fraudulently retransmits it to misdirect the receiver. For instance, One Touch Ping insulin pumps and blood glucose meters do not use any sequence numbers or timestamps, which allows attackers to capture transmissions and replays them later to perform an insulin bolus without specialized knowledge [94]. Radcliffe [107] showed that a continuous glucose monitoring device (CGM) without any timestamp or other protection methods in network packets could be exploited by a replay attack. This attack led to an unusual insulin dosage to the patient resulting in a hypoglycemic condition. In another work, Xiao et al. [74] showed that software-defined radio (SDR) waves emitted from EEG devices can be recorded to replay in an RF dongle to recover the patient’s EEG signals maliciously [74].

**Impersonation attacks.** In an impersonation attack, an adversary successfully disguises as a valid user in the communication system to gain access to the victim’s sensitive information and take advantage of the clear-text
communication between healthcare devices. Li et al. [14] showed that an unencrypted communication between a glucose monitoring device and the insulin delivery system could be sniffed, and by applying reverse-engineering methods, it is possible to discover the device PIN [14]. Furthermore, this PIN can be used to authenticate a patient maliciously to perform an impersonation attack. In another work, researchers introduced a hijacking attack using a smartphone application and its corresponding Medical IoT (MIoT) devices (e.g., pulse oximeter, glucometer) [108]. MIoT devices can store offline readings when the user’s smartphone application is not available to upload the results in the smartphone interface. A hijacker with stolen user credentials can open an account from another smartphone and retrieve all of the offline readings that can be verified using digital forensics techniques.

**DoS attacks.** In DoS attacks, the attackers usually make a healthcare device or system unavailable temporarily or permanently to legitimate users by sending excessive and unnecessary service requests. For example, an ICD can remain in the standby mode for 5 minutes after activation even though there is no active communication session. This wait time can be exploited by initiating false communication sessions and keep the ICD in standby mode for longer times [49]. Ransford et al. [109] reported a crash attack to cardiovascular implantable electronic devices (CIEDs) in which attackers send undisclosed radio traffic to disrupt the radio connectivity of the CIED, causing the device to stop working [109]. A group of researchers reverse engineered the communication protocol of a battery-powered ICD to communicate with an unauthenticated device that posed a potential DoS risk to the ICD [12]. An exploitable DoS vulnerability was identified in the use of a return value in an EEG-based software applications program [83]. As a consequence, a specially crafted network packet could cause an out of bounds read to trigger this vulnerability.

Communication protocols of healthcare devices are also vulnerable to DoS attacks. A team of cybersecurity researchers reported SweynTooth, a repository of 12 security vulnerabilities, affecting thousands of BLE-enabled smart medical devices [110]. This repository includes a DoS attack where an attacker in radio range performs a buffer overflow by manipulating the link-layer length field. This attack triggers a deadlock state when a device received a packet with a clear link layer ID, primarily leading to an OS attack. Wang et al. [112] reported a data overflow vulnerability in the medical image-based communication standard called digital imaging and communications in medicine (DICOM). Researchers developed a DICOM vulnerability framework and found that when the content of the received image file was greater than 7,080 lines, archiving and communication systems refused to respond to any request from the server.

Medical web servers are also being targeted by the attackers to perform DoS attacks by sending numerous fake requests. In 2014, one of the largest children’s hospitals in the United States was the target of a distributed DoS attack by flooding the website with numerous fake requests over a 7-day period [111]. In consequence, the hospital’s website was unreachable and day-to-day operations at the hospital were slowed down. In another work, McMahon et al. [78] exploited an outdated version of hypertext preprocessor (PHP < 4.4.5) to perform DoS attacks, as well as remote code executions.

**MIMO attacks.** MIMO refers to a setting where multiple antennas are used by the transmitter to transmit a wireless message to a receiver with multiple antennas. In the MIMO attack, attackers try to recover signals sent by the transmitter in the presence of a friendly jammer, without any collaboration with the jammer or transmitter. An attacker can recover confidential messages from distances even when the friendly jammer and the data source are few centimeters apart, and the attacker is several meters away. Friendly jamming is often used to protect the confidentiality of the communicated data, which also enables message authentication and access control. Researchers showed that MIMO attacks are still possible with two receiving antennas from a range up to 3 m [113].

**MITM attacks.** An MITM attack occurs when communication between different components of healthcare systems is monitored and modified by unauthorized users. This attack can be used to inject malicious codes to a healthcare device or server, intercept sensitive information like protected health information, expose
confidential information, and modify trusted information. Researchers introduced distance hijacking attack to intercept an ongoing communication in healthcare systems [114]. Here, two medical devices, the prover and the verifier, are considered in the distance bounding protocol, where the verifier establishes physical proximity with the prover. The authors considered various adversarial capabilities for falsifying physical abilities to the prover to create a false or rogue prover that can intercept the communication and establish a new communication channel with the authorized verifier. In another study, a Bluetooth-enabled pulse oximeter was used to perform an MITM attack [115]. In this attack, attackers jammed the Bluetooth device to break the existing connection to pair the device with an access point (AP). Hei et al. [116] presented a MITM attack where an attacker compromised the wireless communication between an insulin pump and a USB device. As the communication in the insulin pump was unencrypted, the attacker could perform signal acute overdose and chronic overdose attacks. Signal acute overdose issued a one-time overdose to the patient, whereas chronic overdose issued extra portions of medication to the patient over a long period of time. Marin et al. [97] demonstrated an MITM attack by intercepting the communication between the ICD and the programmer device. Additionally, the researchers reverse engineered the communication protocol of the ICD to show that these proprietary protocols do not provide any security during communication. Newaz et al. [119] presented an MITM attack to establish a proxy connection between the personal medical device and the smartphone to capture the BLE communication packets [119]. Hence, all communication packets between the personal medical device and the smartphone were rerouted through a publicly available BLE MITM framework. An attacker could observe the BLE packets and extract sensitive information including device information, payload, and so on forwarded by the personal medical device.

Fawaz et al. [102] demonstrated an MITM attack against BLE-enabled healthcare devices that accept connection from unauthorized programmer devices. The unauthorized BLE pairing allows an active attacker to inject malicious traffic into any BLE channel at any given point in time without crossing the bounds of BLE specifications. Hence, an attacker can obtain an inventory of the patient’s device and learn the patient’s health condition, preferences, habits, an so on. Chauhan et al. [117] used an MITM proxy to capture and decrypt the network traffic generated by smartwatch apps. Researchers inspected the captured traffic and found personal information about the user, such as location, app credentials, health data (e.g., heart rate, water intake), and user activities as a result of the unencrypted communication between the device and app. Paolletti et al. [118] presented a formal approach to perform effective and stealthy reprogramming attacks on ICDs. Researchers focused on the ICD software that implements a discrimination algorithm along with multiple discrimination criteria (discriminators) for the detection and classification of arrhythmia episodes based on the analysis of intracardiac signals features. In this attack, an attacker tried to change the discrimination features that might alter the device’s parameter to induce misclassification and inappropriate or missed therapy to the patient. For performing this attack in real life, an attacker needs to know the ICD model of the victim so that it can select the appropriate discrimination algorithm. In addition to this work, an attacker can send discovery signals to the device to know the ICD model [12].

Battery depletion attacks. A battery depletion attack is a forced authentication attack where an attacker tries to connect with an IMD to perform multiple authentications and drain the battery of the device. Raymond et al. [120] presented a denial-of-sleep attack that prevents the medical device from activating power-down mode in case of a failed authentication attempt to exhaust the battery life. Security researchers of MedSec studied St. Jude Medical Merline’s CIED and reported a battery drain attack that reduced the CIED operating time cycle [109]. In a recent report, an implementation flaw in an ICD was reported where the ICD did not go to the sleep mode even after ending an active communication session [49]. This flaw can trigger a DoS attack and drains the battery of the ICD. Hei and Du [121] presented a battery depletion attack on an IMD by exploiting the wireless communication between the IMD and programmer device [121]. As the programmer device needs to authenticate itself to the IMDs, an unauthorized programmer device can send several authentication requests to consume a considerable amount of battery life of the IMD. Researchers reported unsecured authentication in earlier versions of wearable
devices, such as Google Glass, which allows root access to the attackers [122]. Using this root access, attackers can establish a connection with the wearable devices and pass certain commands to recognize the users’ face, record video footage, and record the voice of the user.

5.6 Summary of Existing Attacks
We categorize 84 attacks on healthcare systems reported by research communities and developers in five different categories. These publicly known attacks have been selected from books, research articles, and other published online materials such as the databases of PubMed (Medline), IEEE Digital Library, ACM Digital Library, Elsevier ScienceDirect, Directory of Open Access Journals (DOAJ), Cybersecurity and Infrastructure Security Agency (CISA), and Semantic Scholar. We specifically consider the research articles and news that addressed cybersecurity issues related to software, hardware, and network vulnerabilities, attacks, possible solutions, and architectures from the healthcare domain perspective. The key terms used in the search criteria were based on cybersecurity, healthcare device attacks, medical field, cyber attacks, architectures, threats, and vulnerabilities. Additionally, in this section, we explain the attack methods and discuss the impacts of these existing attacks based on different metrics (i.e., attack approach, impacted security, targeted medical devices, targeted components, and types of attack). We summarize our findings about these attacks in Figure 3.

**Attack approach.** Based on the attack approach in a healthcare system, the attacks can be categorized as active or passive. Active attacks (e.g., DoS attack, MITM attack) try to change the healthcare system resources or affect the system’s operations, whereas passive attacks (e.g., eavesdropping, weak authentication scheme exploitations) read or make use of the information from the healthcare system resource. From Figure 3(a), one can observe that passive attacks have been reported more (55.6%) in healthcare than active attacks (44.4%).
Impacted security. After successful exploitation, confidentiality (C), integrity (I), and/or availability (A) of a healthcare system are impacted by the existing attacks. Active attacks (e.g., malware, ransomware) mostly affect the integrity and availability of a healthcare system. On the contrary, passive attacks (e.g., eavesdropping, MIMO) jeopardize the confidentiality of a healthcare system. As Figure 3(b) shows, the integrity of the systems is impacted the most (44.0%) due to reported attacks in healthcare.

Targeted medical devices. Based on the existing attacks in a healthcare system, as Figure 3(c) presents, non-invasive medical devices are the most targeted by the attackers. Most of the non-invasive devices (e.g., smartwatches, BCI devices) do not use encryption and authentication mechanisms during their communication with the programmer but rather perform clear-text data transmission. As a consequence, an attacker can perform active attacks (e.g., DoS attack, replay attack) and passive attacks (e.g., eavesdropping, impersonation) on these devices. Attacks on active therapeutic devices are much lower compared to the attacks on invasive devices.

Targeted components. Among the five components in a healthcare system (explained in Section 3), medical devices and data are the most affected due to attacks (Figure 3(d)). As the patients’ data and information produced by healthcare devices can be used for blackmail and extortion [124], attackers target medical devices and data the most. Besides these components, network and healthcare provider attacks are also high in number. Sensor-level attacks are the minimum in number compared to the other targeted components.

Types of attacks. Based on the current attacks, as Figure 3(e) presents, communication channel attacks are the most common attack on healthcare systems. Healthcare devices use various wireless communication protocols (e.g., Wi-Fi, Bluetooth, Zigbee) to check the patient’s status remotely, which at the same time makes the patient vulnerable to the attacks. Software attacks are emerging as more medical devices are getting smart and support third-party applications. Hardware attacks occur because third-party vendors develop most of the ICs. Although small in number, hardware attacks are still a concern for the patients, as they are increasing day by day.

6 CURRENT SECURITY AND PRIVACY SOLUTIONS

In this section, we discuss the security measures on healthcare systems that have been proposed by research communities and developers to defend against the attacks presented in Section 5. Most of the security measures directly address the trade-offs in healthcare devices, whereas others propose several countermeasures against specific types of attack. Several of these solutions are widely used or considered as standards. In most cases, the primary concern of a security measure is how to deal with the emergency when there is risk to a patient’s safety. In the following sections, we discuss different existing security measures for healthcare systems by categorizing them into five broad categories. A summary of these security and privacy solutions for healthcare systems is presented in Tables 3 and 4.

6.1 Hardware-Centric Solutions

Hardware-centric solutions are designed to protect healthcare devices from hardware-level attacks like HTs. The solutions are broadly divided into several categories, including HT triggering and physical unclonable function (PUF), physical separation, and online HT detection methods.

HT triggering and PUF. This type of HT detection depends on testing the design after the chip fabrication. Wu et al. [125] proposed the golden die method to detect HTs with significant footprints differentiated from the base standard. However, cleverly inserted HTs may not be easily triggered by this approach, as the testing mechanism may not know the presence of the HT and its location in the chip. Fracq and Frick [126] presented a functional verification method that depends on checking the functionality of the hardware by monitoring the output and checking for expected behavior. Compared to the aforementioned methods, the PUF method is useful for healthcare applications, as it derives a secret from the physical characteristics of the IC instead of storing secrets in digital memory [127].
<table>
<thead>
<tr>
<th>Solution Type</th>
<th>Defense Mechanism</th>
<th>Attack Type</th>
<th>Summary</th>
<th>Limitations</th>
<th>References</th>
</tr>
</thead>
</table>
| Hardware centric | HT triggering and PUF | Hardware, communication channel | • Testing the design right after the chip fabrication step.  
• Detect HTs with significant footprints like a golden die method.  
• PUF method derives a secret from the physical characteristics of the IC. | • Cleverly inserted HTs may not be easily triggered.  
• Does not guarantee the correctness of the design at runtime.  
• Offline full functionality test is inefficient and time consuming. | [125–127] |
| | Physical separation | Software | • Executes critical security applications on isolated hardware that is free of observation and interference through direct physical access. | • Does not hide the data processing, which may reveal patient information.  
• Vulnerable to DoS attack. | [128, 129] |
| | Online HT detection method | Hardware | • Identify HTs by checking underlying hardware functionality.  
• Architecture is divided into two-chip generating signatures deep in the hardware and later checks it during data processing and transmission. | Relied on the digital logic modules that minimally impact the performance of the system. | [57] |
| Software centric | Secure execution environment | Software, system level | • Secure virtual machines provide a secure network interface, secure storage, and secure execution environment. | Management environment can still be a compromised OS.  
Performance penalties are based on execution-specific domain operations as well as several benchmarks. | [130] |
| | Static analysis | Software | • Provides almost complete coverage of the code and helps to detect potentially fatal errors.  
• May not easily be detected through conventional testing methods (e.g., CodeSonar). | It identifies many bugs in the software, but static analysis tools are not a replacement for testing. | [131] |
| | Runtime analysis | Software | A dynamic binary instrumentation-based malware detection framework.  
• Can trace the untrusted program during execution time in a virtualized testing environment.  
• With a large number of input values and with extensive security policies, it can detect malware behavior. | Depends on the accuracy of the security policies used.  
Along with the number of observed paths, in particular, observed malicious paths, in the testing environment. | [132] |
Table 3. Continued

<table>
<thead>
<tr>
<th>Solution Type</th>
<th>Defense Mechanism</th>
<th>Attack Type</th>
<th>Summary</th>
<th>Limitations</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formal verification</td>
<td>Software</td>
<td></td>
<td>• Functional specification of the medical device are expressed in input/output sequences</td>
<td>• Current software verification tools have been written in a high-level programming language.</td>
<td>[133–135]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Then translated into assert verifiable property so that medical device software accepts it.</td>
<td>• Not suitable for highly platform specific and low-level programs.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• A model checker is used to feed transformed code and valid assertions.</td>
<td>• System-level properties need to be verified with the real-world medical device interfaces.</td>
<td></td>
</tr>
<tr>
<td>Encryption</td>
<td>Communication channel</td>
<td></td>
<td>• PRESENT and KATAN lightweight hardware-oriented block ciphers are as small as 1000–1500 gate equivalent for a 64-bit block size encryption.</td>
<td>• Low-power symmetric ciphers still may increase the energy consumption, which will shorten the battery life.</td>
<td>[2], [136], [137], [138–152]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Stream mode utilizes output feedback to obtain a scalable stream cipher.</td>
<td>• If the patient is unconscious, then key distribution needs to be done without the patient’s intervention for timely treatment.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Encompression is the combination of compressive sensing, encryption, and integrity checking.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Symmetrical encryption algorithms are the distribution of shared key between devices.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Machine learning</td>
<td>Communication channel</td>
<td></td>
<td>• A decision tree algorithm was used to detect malicious attacks in healthcare devices.</td>
<td>• Energy consumption is an issue for resource-limited healthcare devices.</td>
<td>[153–156]</td>
</tr>
<tr>
<td>based approaches</td>
<td></td>
<td></td>
<td>• Support vector machine, trees, and ensemble algorithms were used to authenticate users for healthcare devices.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• An ML-based security framework was proposed to detect malicious activities in a healthcare system.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access control mechanisms</td>
<td>System level</td>
<td></td>
<td>• An identity-based access control proposed to protect private health data in a cloud-assist health monitoring system.</td>
<td>• Most access control mechanisms focused only on the healthcare device access authentication.</td>
<td>[12, 14], [157–164]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Attribute-based access policy has been employed in many research works to control access to medical data, BAN, and cloud system.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Proximity-based access control is based on the distance between the programmer and the healthcare device.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solution Type</td>
<td>Defense Mechanism</td>
<td>Attack Type</td>
<td>Summary</td>
<td>Limitations</td>
<td>References</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>---------</td>
<td>-------------</td>
<td>------------</td>
</tr>
</tbody>
</table>
| Data protection | Blockchain-based approaches | Communication channel | • A hybrid approach that combined advantages of the private and public key, blockchain, and many other lightweight cryptographic primitives to develop a patient-centric access control for electronic medical records.  
• Researchers introduced a reliable data communication and storage with more advanced and lightweight cryptographic techniques like the ARX encryption scheme. | • Storage is an issue for the resource-constrained healthcare devices. | [165–168] |
| Current analysis | Hardware | | • Static CMOS gates are subject to leakage current in the idle mode.  
• Needs to measure the current from multiple power pins. | • Unable to detect small HTs because of power and timing variations. | [169–171] |
| Delay variation and characterization techniques | Hardware | | • High-precision, low-overhead embedded test structure (REBEL used to detect anomalies in HTs.  
• Capable to deliver high-resolution measurement of path delays.  
• Linear regression used to classify delay behavior  
• Apply on a large number of unobservable internal combinatorial register-to-register paths. | • Timing-based HT is a backtracking-based algorithm used to identify reconvergent paths in the circuit that unable to catch the delay variations caused by HT.  
• Additional power and area overheads may not acceptable for medical devices. | [172, 173] [174] |
| Battery-constraint mitigation | Communication channel | | • BAN protocols can be used to mitigate this problem.  
• The node will not wake up without any messages that are outside of negotiated time intervals.  
• Data-independent power consumption as circuit-level solutions can be used here.  
• A security protocol called IMD-fence can be used here. | • Jamming-based protection is provided by many external security devices, but it is not always effective because it can cause battery depletion. | [12], [175–179] |
| Side-channel analysis | | | | |
| Power consumption analysis | Software | | • Measures the power consumption of traditional programs and known malware.  
• Compares them with the power signatures to find any anomaly in the power consumption.  
• Any aberration in power consumption can be detected as anomaly behavior in case of medical devices. | • No consistent base set of known good behaviors on a PC.  
• A system like WattsUp-Doc would likely to raise false alarms because of an inconsistent or inaccurate internal model. | [180] |
<table>
<thead>
<tr>
<th>Solution Type</th>
<th>Defense Mechanism</th>
<th>Attack Type</th>
<th>Summary</th>
<th>Limitations</th>
<th>References</th>
</tr>
</thead>
</table>
| Shielding and filtering | Side channel | Side channel | • The exterior of a healthcare device was covered with a conducting surface.  
• Attackers were forced to transmit a $10^4$ times more powerful signal to have the same effect as before.  
• Faraday cage is another countermeasure to block EM radiation. | • Very difficult to defeat EM analysis, except if the circuit and its countermeasures are overlapped. | [88], [181] |
| Masking | Side channel | Side channel | • Intermediate values of a cryptographic computation are randomized by masking.  
• A key masking method can be used against DPA.  
• A band-pass filter or a current-flattening circuit can be added to the cryptosystem to suppress information leakage. | • Very difficult to defeat EM analysis because energy overhead is a big concern for healthcare devices. | [182–185] |
| Biometrics | Communication channel | Communication channel | • ECG signal asserts the time between heartbeats, or interpulse interval.  
• Creates a high level of randomness and can be measured from anywhere on the body.  
• Temporal and morphological alterations of ECG measurements could be detected using arterial blood pressure signal. | • If any physiological signals staying within the human body are incorrect, both the security and privacy of schemes may be affected. | [186, 187], [188–197] |
| Trust-management framework | Out-of-band (OOB) authentication channel | Communication channel | • Use audio, visual, or haptic channels for authentication.  
• Visual OOB authentication, such as ultraviolet or visible tattoos to record permanent implantable medical devices keys. | • Auxiliary channels can be breached by close-range eavesdropping.  
• Visual OOB suitable for the emergency but will be a problem for key revocation and usability concern. | [12], [136, 137, 198, 199] |
| Close-range Communication channel | Communication channel | Communication channel | • Near-field communication (NFC) and radiofrequency identification (RFID)-based channel can be utilized here.  
• Distance bounding communication between the medical device and external devices.  
• Access is granted only if the devices are within a safe range. | • A successful RFID eavesdropping attack is possible at a distance of a few meters with off-the-shelf antenna kits | [200–205], [114] |
Table 4. (Continued)

<table>
<thead>
<tr>
<th>Solution Type</th>
<th>Defense Mechanism</th>
<th>Attack Type</th>
<th>Summary</th>
<th>Limitations</th>
<th>References</th>
</tr>
</thead>
</table>
| External device | Communication channel | • Ensure radio security without any modification to the IWMD itself.  
• Patients’ ECG signals used in IMDGuard to extract keys explicitly.  
• Physical characteristics such as received signal strength indicator, time of arrival, differential time of arrival, and angle of arrival are used to detect anomalies. | • Jamming-based protection not always effective. | | [113], [163, 206–208] |

**Physical separation.** This is a hardware-centric design solution where security applications run on separate hardware from the device’s main architecture. The trusted platform module is proposed as a physical separation method where the cryptographic keys for a specific host computer are stored in a separate module to use in IWMDs. As a separate module introduces overhead, a software-based trusted platform module can be used to increase the effectiveness in healthcare devices [128]. Sorber et al. [129] proposed Plug-n-Trust, which is a MicroSD card that provides a trusted computing platform on a smartphone. Plug-n-Trust encrypts all medical data transmitted from the devices and can only be decrypted in the card for further analysis in a verified application programming interface. However, it does not hide data processing, which may reveal the patient’s information and can be vulnerable to DoS attacks.

**Online HT detection method.** The online HT detection method refers to identify HTs at runtime. Wehbe et al. [57] proposed an online method to identify HTs by checking underlying hardware functionality at runtime. Here, researchers divided the whole architecture in two-chip, generating signatures deep in the hardware and later checked it during data processing and transmission. This technique relies on the digital logic modules that minimally impact the performance of the system.

### 6.2 Software-Centric Solutions

Software-centric solutions are designed to protect healthcare devices from software-level and system-level attacks (e.g., malware, ransomware, weak authentication scheme exploitations, counterfeit firmware). The software-centric solutions can be categorized as follows: secure execution environment, static analysis, runtime monitoring, and formal verification.

**Secure execution environment.** A secure execution environment is a safe execution space for executing the code that ensures security to the code and loaded data. For ensuring the safety of healthcare applications, one needs to run these applications on a secure execution environment to protect from a compromised OS. In this sense, **secure virtual machines (VMs)** can be a viable solution that provides a network interface, storage, and execution environment. The main advantage of using a secure VM is that only security-critical healthcare applications will be running on the VM, making them isolated from other applications that may be compromised. However, the management environment can still be a compromised OS [130].

**Static analysis.** Although a secure execution environment can protect healthcare applications from a compromised system, it cannot defend them if the healthcare application itself is a malware. Static analysis techniques can be used to characterize the execution behavior of a program and find program flaws by analyzing the source code. By using symbolic execution techniques to explore the execution paths of the software, static analysis de-
tects potentially fatal errors that may not be easily detected through conventional testing methods. Jetley et al. [131] proposed CodeSonar, a static analysis tool to automatically detect buffer overrun, initialized variables and null pointer dereference in healthcare apps. However, static analysis depends on the availability of the source code, which is not openly available for healthcare applications.

Runtime analysis. Runtime analysis can be used to detect unintended behavior of a healthcare app at runtime. Aaraj et al. [132] proposed a dynamic binary instrumentation (DBI)-based malware detection framework that observes the execution of unknown programs, models safe/unsafe behavior with respect to specified security policies, and ensures that the program does not deviate from safe behavior [132]. This framework uses two virtual execution environments: testing and real environment. In the testing environment, DBI collects specific information in the form of execution traces to construct a hybrid model for representing dynamic control and data invariants. This hybrid model along with a behavioral model generated from security polices is used to define the malware behavior in the DBI framework. In the test environment, a program is analyzed based on the generated model in the test environment to detect malicious behavior of a program in healthcare devices. An unknown program is only moved into the real environment to monitor its execution if the behavior pattern is matched with the allowed model generated in the test environment.

Formal verification. Formal verification methods are proposed in several prior works to develop reliable medical device systems [133–135]. Formal methods are well-formed statements in mathematical logic, and formal verification provides strict deduction of that logic. As a result, the entire state space of a system can be examined to establish a security property for all possible input. Formal verification can be used to verify whether or not medical devices are free from vulnerabilities. An example of a formal verification approach is described by Li et al. [133], where medical device software is the first subject to the source transformation to address the semantic gaps. Properties based on the functional specification of the medical device are expressed in input/output sequences and then translated into a verifiable property to make it acceptable for the medical device software. After that, a model checker is used to feed transformed code, and valid assertions, which verifies the code against the statements and reports whether the code is acceptable or an anomaly is detected. Current software verification tools are written in a high-level programming language, which is not suitable for platform-specific and low-level programs of medical devices. Researchers proposed a semiformal verification approach that was a combination of dynamic and static verification to cover the state space exhaustively [134]. Pre and post-market analyses were carried out based on formal verification techniques to support the process of reviewing healthcare software in the work of Jetley et al. [135].

6.3 Data Protection Mechanisms for Communication and System-Level Attacks

Data protection mechanisms ensure the safety of important medical information from unintended corruption, compromise, or loss regardless of its physical or logical location. Although the most common solution for data protection is encryption, selecting an appropriate cryptographic algorithm for resource-constraint healthcare devices is important. In this regard, ML and blockchain-based approaches can be alternatives to encryption that can protect health data from tampering and detect attacks in healthcare systems. To defend against system-level attacks, an access control mechanism is a good solution to protect healthcare data from unauthorized users.

Encryption. For transmitting sensitive data, especially over a wireless channel, encryption is a fundamental technique to secure the communication channel. Although there are many encryption techniques to follow, high energy consumption and implementation costs of encryption are big issues for resource-constraint medical devices [148, 149]. As symmetric encryption mechanisms usually consume less power than asymmetric encryption mechanisms, it is more practical to use them for resource-constrained platforms like medical devices [139]. In this domain, PRESENT and KATAN, two lightweight hardware-oriented block ciphers [138, 140] were developed with 64-bit block size and an 80-bit key. Additionally, a low-power block-cipher-based security protocol was
A Survey on Security and Privacy Issues in Modern Healthcare Systems: Attacks and Defenses

proposed by Beck et al. [150], which offers two operational modes: stream and session mode. For short message transmission, the stream mode utilizes output feedback (OFB) to obtain a scalable stream cipher that enables strict duty cycling for energy. The session mode utilizes the CBC and a challenge-response scheme to provide advanced security to health data.

Symmetric encryption algorithms depend on the distribution of shared keys between devices where healthcare devices often have to communicate with previously unknown devices. If the patient is unconscious, then key distribution needs to be done without the patient’s intervention for timely treatment. The secret key can be imprinted on a card or on wearable devices where it can be hidden secretly [136] or printed directly onto the patients’ skin using ultraviolet-ink micro-pigmentation that is visible only under ultraviolet light [137]. Even if the secret key is shared secretly or pre-loaded, it needs to be changed periodically, and new keys should be generated with a secure agreement protocol, high randomness, and minimum energy overhead. For instance, Received signal strength indicator is the symmetrical property of the wireless channel between two devices that can generate a symmetrical key from the communication [144–147]. These techniques are useful not only for the initial key sharing setup but also for renewing the key periodically to prevent eavesdropping.

However, even low-power symmetric ciphers may increase power consumption, which will shorten battery life. To address this issue, different compression techniques can be used before applying encryption to reduce the power consumption and transmission cost [2]. Compressive sensing is well suited since compression can be realized with a very low computational and energy footprint [141]. Zhang et al. [151] presented Encompression, a combination of compressive sensing, encryption, and integrity checking, which utilizes the sparsity of sensor data for reducing power consumption in medical devices [151]. Researchers showed that Encompression can reduce consumption by 78% compared to traditional encryption and integrity checking with a reasonable compression ratio of 6 to 10x. The AES [142] and the secure hash algorithm (SHA) [143] are used for data integrity and confidentiality where a hash algorithm is used on original data. As a result, imposters cannot generate encrypted data without knowing the AES secret key. Bu et al. [152] presented a MAC-then-encrypt (MtE) security mechanism combined with AEC-CBC mode to protect IMDS from communication-based attacks. As current IMDS are equipped with AES using 128- or 192-bit encryption keys, the IMD’s information, the timestamp, and the authentication signature can be wrapped all under 128 bits or 192 bits depending on the demand. As a result, MtE adds no extra transmission overhead in IMDS equipped with 128- or 192-bit AES-CBC mode.

**ML-based approaches.** ML is a data analytic technique that provides healthcare systems the ability to learn from data and perform specific tasks such as anomaly detection and behavior analysis from experience without being explicitly programmed. ML algorithms have been explored widely by the research community to detect attacks on healthcare systems [156]. Saeedi [153] demonstrated how a decision tree algorithm can be used to detect malicious attacks in healthcare devices. Here, the author used the normal behavior of healthcare devices as ground truth, and any deviation from the normal behavior was identified as an attack. Vhaduri Poellabauer [154] used several physiological and behavioral parameters such as calorie burn, average step counts, and minute heart rate as features of support vector machine to detect unauthorized access to a healthcare device and its captured data. In a recent work, Newaz et al. [155] proposed HealthGuard, a novel ML-based security framework to detect malicious activities in a connected healthcare system. HealthGuard collects the vital signs of different healthcare devices and uses several ML algorithms to correlate the changes in body functions of the patient to distinguish between benign and malicious activities. Although ML algorithms can identify anomalous behavior in a healthcare system, implementing an ML-based solution on medical devices can consume more energy, which is an issue for these resource-constrained devices.

**Blockchain-based approaches.** A blockchain is a distributed system that maintains a continuously growing list of data records and keeps these records safe from tampering. Blockchain-based security frameworks are widely used by researchers to protect healthcare data from unauthorized entities. Chen et al. [165] proposed a blockchain-based storage scheme for medical data to ensure safe data storage and sharing. Researchers also introduced a
service framework for sharing medical records to describe the process of personal medical data management in some applications. A novel hybrid approach was introduced by Dwivedi et al. [166] that combines several cryptographic primitives (e.g., private key, public key, blockchain) to develop a patient-centric access control for electronic medical records. Srivastava et al. [167] introduced a reliable data communication between the network and storage with more advanced and lightweight cryptographic techniques like the ARX encryption scheme. They introduced the concept of ring signatures in the communication, which provides important privacy properties like Signers Anonymity and Signature Correctness. The same group of researchers introduced GHOSTDAG, a novel and unique blockchain protocol for remote patient monitoring, which uses a directed acyclic graph instead of classic long singular blockchains [168]. Researchers utilized the idea of smart contract programs from blockchain to monitor the health data of patients. However, the ownership of the current medical data is still an issue, and currently there are no rules for using blockchain in the Health Insurance Portability and Accountability Act.

Access control mechanisms. Access control mechanisms prevent unauthorized access to healthcare devices. Researchers have proposed several types of access control for healthcare systems, including proximity-based, identity-based, role-based, attribute-based, and risk-based access control [157]. Lin et al. [158] proposed an identity-based access control to protect private health data in a cloud-assisted health monitoring system. In the role-based access control scheme, the service requestor’s role determines whether the access will be granted or denied. Li et al. [159] proposed to give access rights to healthcare providers based on their roles in the wireless BAN. Attribute-based access control is an extension of identity-based access control where decisions are made based on a set of attributes (e.g., specialty, license validity). Attribute-based access policy has been employed in many research works to control access to medical data, BAN, and the cloud system [160, 161]. Risk-based access control brings real-time, risk-aware decision-making capability in the access control mechanism. The anomaly detection-based access control schemes fall into this category that constantly monitors any abnormal behavior in accessing a healthcare device [162, 163]. Proximity-based access control is based on the distance between the programmer and the device [12, 14]. Here, the programmer can generate the same key to decrypt the communication only if it is near the patient. Fu et al. [164] presented a physical obfuscated key (POK)-based IMD access control mechanism where researchers leveraged IC cards of POKs for secure credential storage. They designed a lightweight access control protocol with minimal computation and communication overhead on IMDs.

6.4 Solutions Based on Side-channel Analysis

Side-channel analysis relies on analyzing the information gained from the physical properties of the healthcare devices (e.g., energy consumption, timing analysis, or EM emanations) and compares the information with the data generated by the normal behavior of a device to detect anomalies. Most of these solutions can be broadly divided into the following sub-categories: electric current analysis, delay variation and characterization, power consumption analysis, shielding and filtering, masking, and battery-constraint mitigation.

Electric current analysis. Electric current analysis measures the current consumption in a device, allowing experts to detect anomalies in devices at the hardware level. Bhunia et al. [169] monitored current leakage from static CMOS gates to identify Trojan circuits in a healthcare device. As current leakage always remains the same for the CMOS gate, the difference in current consumption can distinguish the HT from the base circuits. However, in the case of a large circuit with a high number of gates and fewer Trojans, the current analysis fails due to an insignificant change in current readings and difficulty of performing extensive testing. Aarestad et al. [170] proposed current analysis in multiple pins instead of a single point to increase sensitivity and reduce the problem of detecting a few gates in a fraction of the total gates in the IC. However, this technique cannot detect HTs that are small in size due to the power and timing variations that HT can cause [171].

Delay variation and characterization. This HT detection method works by measuring and detecting small systematic changes in path delays introduced by capacitive loading effects or series inserted gates of HTs.
high-precision, low-overhead embedded test structure called REBEL was proposed to detect delay anomalies in HTs [172]. REBEL was capable of delivering high-resolution measurements of path delays and able to identify a wide range of delay anomalies introduced by HTs. It provides significant benefits over other traditional delay testing methods, as the digital snap-shot captured by REBEL allows glitches to be detected and can potentially speed up the path delay measurements using a small number of repeated applications of the test pattern. The detection sensitivity of REBEL was checked by varying the analog control voltage on each Trojan emulation circuit one at a time and classifying the result using regression analysis. A backtracking-based algorithm was proposed by Wei et al. [173] to identify the reconvergent locations in the circuits where the delay variations caused by HT are not observable.

A new method for IC authentication and hardware Trojan horse detection is delay characterization technique introduced by Li and Lach [174]. Such a technique measures the combination of an arbitrarily large number of register-to-register paths delays internal to the functional portion of the IC. This technique was originally developed to apply on a large number of unobservable internal combinatorial register-to-register paths to get accurate, precise data about path delays. Moreover, this technique can also be used for hardware Trojan horse detection by extracting the non-functional path delay characteristics to detect malicious circuit alterations. The delay measurement technique does not affect circuit functionality and allows monitoring of delay characteristics at runtime. However, this technique needs to assess the authentication approach across a large number of physical ICs, which introduces time latency and resource overhead. As healthcare devices are power-constraining devices, additional power and area overheads may not be accepted.

Power consumption analysis. Malware detection in healthcare devices can be performed by power consumption analysis, which compares the power consumption of traditional programs and known malware to find anomaly in the system. Clark et al. [180] proposed a behavior monitoring system, WattsUpDoc, which uses supervised learning to classify normal and abnormal power consumption in the replacement of a pre-constructed power consumption model and detect anomaly behavior in healthcare devices [180]. The key idea is a high sampling rate in power consumption tracing and high-accuracy power measurement to detect malware in healthcare devices. However, WattsUpDoc raises false alarms in case of an inconsistent base set of known behavior. In addition, achieving high-accuracy power measurement in resource-constrained medical devices is difficult, which results in a low accuracy rate in detecting malware.

Battery-constraint mitigation. As healthcare devices are resource-constraint devices, battery depletion/drainage attacks can cause severe obstruction in normal operation of the devices. However, a defense mechanism against the battery drainage attack needs to be power efficient, as otherwise the defense mechanism itself may consume more power than the attack. Researchers proposed zero-power defenses for ICDs where RF energy harvested from external sources is used for notification, authentication, and key exchange [12]. In addition, BAN protocols can be used to mitigate this problem. For instance, the IEEE 802.15.6 BAN standard allows a node and hub to negotiate their communication intervals by encoding them in authenticated messages. Accordingly, the node will not wake up without any messages that are outside of negotiated time intervals to save power in the devices [175]. Tiri et al. [176, 177] proposed novel logic styles with data-independent power consumption as circuit-level solutions against battery drainage attacks to reduce the dependence of power dissipation on input patterns. In recent work, Siddiqi and Strydis [179] proposed an adaptive zero-power defense solution using an RF power transfer mechanism based on energy harvesting against battery-depletion attacks.

Shielding and filtering. Shielding and filtering are commonly used to defend against EMI attacks. Kune et al. [88] showed that covering the exterior of a healthcare device with a conducting surface can force the attacker to transmit a $10^4$ times more powerful signal to have the same effect in an EMI attack, which can be differentiated from legitimate signals easily. Faraday cage is another countermeasure to block EM radiation and to reduce the EM
radiation signature [181]. However, it is hard to defeat EM analysis, except if the circuit and its countermeasures are overlapped.

**Masking.** Masking is a technique of hiding original data with modified content. Intermediate values of a cryptographic computation are randomized by masking, which avoids dependencies between these values and the power consumption applied in the algorithmic level. Moreover, it does not rely on the power consumption characteristics of the medical device. Researchers proposed a key masking method as a software solution against DPA attacks [183]. Although this method attempts to randomize the secret key before each execution of the scalar multiplication, power overhead is a concern here for healthcare devices. A band-pass filter [182] or a current-flattening circuit [184] can be added to the cryptosystem to suppress information leakage through the current supply pin. An internally generated random mask based on ring oscillators was proposed by Liu et al. [185] to change the power consumption dynamically.

### 6.5 Trust Management Framework for Communication Channel Attacks

A trust management framework focuses on securing information flow and communication among healthcare devices by certified software and application in the system. This framework can be categorized into biometrics, out-of-band (OOB) authentication, close-range communication, and external devices.

**Biometrics.** Biometric properties such as fingerprint, EEG, heart rate, and blood glucose can be used to authenticate a healthcare device and establish trust between two communicating devices worn on the same body [188–195]. Poon et al. [186] presented an ECG signal-based trust management system, which asserts that the time between heartbeats or interpulse interval can create a high level of randomness and can be used to generate a secret key for communication. Cai et al. [196, 197] proposed a user-specific supervised learning model to detect temporal and morphological alterations of ECG measurements using an **arterial blood pressure (ABP)** signal. As ECG and ABP signals both measure the cardiac process, different physiological signals generated by the same underlying physiological process are inherently correlated. Any unilateral change in the ECG signal without a corresponding change in the ABP signal can be detected by the proposed model. A wearable sensor is used in the work of Cornelius et al. [187] to authenticate users passively with high accuracy (>90%) by measuring their bio-impedance to alternating current of different frequencies. However, specific physiological signals within the human body can be incorrect, which may affect both the security and privacy of healthcare devices.

**OOB authentication.** An auxiliary channel or OOB communication uses audio, visual, or haptic channels for authentication that are outside the established data communication channel [136, 137, 198, 199]. Halperin et al. [12] proposed a low-frequency audio channel that enables medical devices (e.g., IMDS) to use a zero-power RF identification for generating and transmitting a key over the audio channel. Denning et al. [136] proposed visual OOB authentication (e.g., ultraviolet or visible tattoos) to record permanent IMD keys, where the keys are visible only under ultraviolet (black) lights. Although this mechanism is suitable for the emergency situation, it will be a problem for key revocation and usability. In addition, Li et al. [198] proposed a mechanism where the users are required to inspect simultaneous LED blinking visually to achieve authentication in BANs. However, this is not appropriate for emergency situations if the patient is unconscious, which makes its application limited.

**Close-range communication.** To prevent unauthorized access, restricting the communication range is an intuitive way to avoid radio attacks. If a healthcare system uses close-range communication, the attacker has to come within the range to perform radio attacks, which increases the chance of detecting the attack/attacker. There are several close-range communications (e.g., near-field communication, RF identification-based channel, near-field identification) proposed in prior works to secure the communication in a healthcare system [200, 201]. One recent close-range communication for securing healthcare devices is body-coupled communication (BCC) proposed...
by Baldus et al. [202]. BCC uses the human body as a signal propagation medium, which utilizes two different mechanisms: the transmission line approach and the capacitive approach. The transmission line approach uses the human body as a transmission line where electrodes are directly attached to the human body for directly transmitting the electrical signals. The capacitive approach uses the human body as a floating conductor, whose electric potential is changed with the electric field generated by the transmitter. However, the idea of using BCC is not ideal, as physiological signals can be read during physical contact like a handshake [203]. An alternative solution of short-range communication can be distance bounding communication between the medical device and external devices. Here, access to external devices is granted if the devices are within a safe range. The distance can be measured in various ways, including limiting the response time to a verification request, ultrasonic waves, and received signal strength indicator, among others [114, 204, 205].

External devices. For enhancing the security of existing medical devices, significant modification is required in hardware and software, which may lead to unintended changes in their behavior. To address this problem, recent studies suggest using external devices to ensure communication medium security without any modification to the existing healthcare devices. An external device can be used as an authentication module to verify service requests from the external user, which can save the battery life of main healthcare devices. Denning et al. [206] proposed Cloaker, a wearable device to block access requests from all external programmers at runtime. Cloaker allows access to only pre-authorized programmers in the normal mode, whereas any programmer, even an unauthorized one, can access the device in the emergency mode.

Gollakota et al. [207] proposed Shield, which is a personal base station placed in between the IWMD and the external programmer. Shield works as a relay that only allows communication from legitimate programmer while jamming all other direct communication to IWMDs. Furthermore, Shield provides an encryption scheme to encrypt and decrypt sensitive information shared between the programmer and the IWMD. As the communication is encrypted and unauthorized communication is jammed, the confidentiality of medical device messages is ensured by Shield.

Xu et al. [208] proposed IMDGuard, an external wearable device designed for ICDs to coordinate interactions between the ICD and other external programmers. IMDGuard uses patients' ECG signals to generate keys to share between the ICD and programmer upon their first connection. Any other external programmers need to be verified by IMDGuard before they can communicate with the ICD. Zhang et al. [163] introduced Medmon, an external device that detects abnormal communication to/from the IWMDs [165]. Medmon uses different physical characteristics (i.e., received signal strength indicator, time of arrival, differential time of arrival) to detect signal anomalies in transmission and alerts users regarding the attack. Medmon also captures behavioral abnormalities such as vital signs or commands that lie outside the historical records of the patient. However, jamming-based protection is provided by many external security devices, but it is not always effective, as a MIMO-based attack can recover jammed signal fully or partially [113].

6.6 Limitations of Current Security Solutions

In this section, we discuss the shortcoming of current security solutions, as well as highlight the limitations of healthcare devices that force the solutions to be too specific to be broadly useful:

(1) As third-party vendors manufacture most of the healthcare devices' hardware ICs, there is no specific standard to follow for IC manufacturing. As a consequence, an attacker can include an unproven intellectual property core into the IC that acts as an HT to perform malicious activities. HTs introduce power and timing variation in healthcare devices. Existing hardware-centric solutions such as current analysis and delay variation are not suitable for HT detection.

(2) Currently, there is no base set of standards for the known behavior of healthcare device power consumption. Hence, the current malware detection techniques using power consumption analysis (e.g., WattsUp-Doc [180]) are not an appropriate choice, as they can raise false alarms and have a low accuracy rate in
detecting malware. In addition, most of the healthcare applications’ source code is not open source, which makes it challenging to use static analysis for finding software-level bugs (e.g., buffer overrun, initialized variables, null pointer dereference) and defend against corresponding attacks.

(3) Healthcare devices and programmers mostly use a fixed secret key loaded during manufacturing time. Using the same pre-shared key for an extended period increases the possibility of a successful cryptanalysis attack. To solve this problem, researchers have proposed simultaneous LED blinking to be used as an authentication mechanism [198]. However, this technique is not suitable for emergency situations.

(4) As most medical devices (e.g., implantable devices, wearable devices) are resource limited, it is not easy to implement any cryptographic algorithm on the devices. One solution can be the use of cryptographic key computation from the patient’s vital state [208]. However, the computation of the key also consumes high power and reduces battery life.

(5) Medical devices collect information from multiple sensors (e.g., blood pressure, glucose, motion) to observe different vital signs of a patient. In real life, patients could become unconscious without showing any alarming symptoms, and emergency personnel would require access to healthcare devices like IMDs to collect information related to previous health-related data. Researchers have proposed using a specific key that is shared between a common group of people like doctors and emergency personnel as a backdoor solution [162]. However, a privilege escalation attack can easily reveal the secret key, and attackers can obstruct the treatment plan, causing a life-threatening situation.

(6) Current software verification tools are written in a high-level programming language that are not suitable for platform-specific and low-level applications of medical devices. These programs have to interact with medical sensors, actuators, and other hardware peripherals. In addition, system-level properties need to be verified with real-world healthcare device interfaces [133].

(7) Different healthcare devices use different communications protocols, software, and application platforms for performing their healthcare-related tasks. Hence, it is hard for security researchers to provide common security solutions for healthcare systems. Although a recent study proposed a standard solution (Health-Guard [155]) to find anomalous activities in a healthcare system, the device dependency in this solution is too complex to consider in the current healthcare domain.

7 DISCUSSION, FUTURE RESEARCH DIRECTIONS, AND RECOMMENDATIONS

In this survey, we focus on the primary security and privacy goals for the next generation of healthcare devices and analyze the most common and related security mechanisms proposed so far. To secure a healthcare system, security proposals must consider the energy, storage, and computing power constraints of the healthcare devices. Furthermore, the security solution must maintain the balance between patient safety and the security level offered. In this section, we discuss security recommendations and practices, and future research directions, that are required to be addressed for ensuring security and privacy in healthcare systems.

7.1 Recommendations

To outline future research directions and recommendations, we consider the following use case scenario of a modern healthcare system. Assume a patient, Alice, buys a smart medical device (e.g., pulse oximeter) to measure blood oxygen level and pulse rate. Alice installs the corresponding pulse oximeter app on the smartphone and connects the smartphone via wireless protocols (e.g., Wi-Fi, BLE, ZigBee) with the medical device using the device’s MAC address written on the body. The pulse oximeter sends the blood oxygen and pulse rate to the app and cloud, which can be accessible by healthcare providers. Here, the following security concerns arise: (1) system failures due to counterfeit/faulty hardware; (2) information leakage from the app and cloud server; (3) eavesdropping while transmitting data from the device to the app; (4) undesired access to the sensitive data via a backdoor in the installed app; and (5) DoS due to open, accessible device information and subpar security.
standards. Based on this use case, we have the following recommendations for device manufacturers, patients and healthcare providers, developers, and the research community.

Device manufacturers. To identify counterfeit or faulty third-party devices, medical device manufacturers can perform current analysis, delay variation, and characterization techniques to detect anomalies at the hardware level. In addition, medical device manufacturers should consider a safer way to distribute any sensitive device-related information to the consumers for avoiding unauthorized access to the smart medical devices. Moreover, smart medical devices come with high-precision sensors (e.g., pulse oximetry sensor in this use case), which can lead to different side-channel or communication-channel-based attacks such as EMI and sensor spoofing attacks. Medical device manufacturers should find a safer way to integrate sensors to evade these attacks. A possible solution in this domain would be to model the sensor behavior as presented in other works [209, 210].

Patients and healthcare providers. Patients (Alice) and healthcare providers are the main victims of different malicious attacks, as they usually have less technical knowledge on different attacks (e.g., software, hardware, communication channel). Hence, both patients and healthcare providers should be aware of the consequences of these threats and attacks and be cautious before using any risky healthcare device and app. Additionally, they should follow good security practices such as rejecting any suspicious device access or disabling automatic data sharing between apps to secure the devices and their information.

Developers. Developers can play a vital role in securing the communication between the pulse oximeter and the smartphone. They should consider a number of security factors (e.g., secure communication, authorized access) when designing a healthcare device or its app (pulse oximeter app in this use case) so that it can be both reliable and secure, as well as be easily adaptable by users. Developers should follow common healthcare data sharing policies to secure healthcare data from unauthorized access [211]. Developers should also follow good app developing practices in the healthcare industry, such as ensuring security through encryption, user authentication, and penetration testing.

Research community. Considering the use case here, the research community should follow specific communication standards based on the types of devices the user (Alice) is using. Instead of depending on the transport layer security in this medical device communication standard, researchers can enforce an authentication mechanism so that attackers cannot easily break into these devices. The research community should also help the industry to address publicly known threats and attacks against the healthcare system efficiently and propose various solutions. Researchers, along with the industry experts, should jointly propose a standard practice in app development, device manufacturing, and communication standards to minimize data abuses in smart medical devices. Furthermore, researchers should report newly found healthcare-related threats to the device manufacturer immediately to reduce the consequences.

7.2 Future Research Directions

Securing healthcare data. The healthcare industry is generating data rapidly, and this healthcare data has clinical, financial, and operational value in the market. To protect this healthcare data from breaches, effective security measures like cryptographic solutions are needed for securing wireless communication, as well as the information stored in the device or the server. In this regard, cryptographic protocols that are symmetric [212] and lightweight [148] can provide a means to control access to healthcare devices and protect against spoofing and elevation of privilege attacks. However, incorporating cryptographic mechanisms in existing healthcare devices implies that current devices like IMDs must be replaced or redesigned. In case of an emergency, communication with unauthorized personnel may be needed, which can be interrupted due to implemented cryptographic schemes. Hence, researchers should focus on developing medical-centric cryptographic solutions to meet the unique security goals of the healthcare system.

Lack of standard communication protocols. Having a communication standard for healthcare devices is a good practice, and many international standards are considered as prerequisites for the certification of healthcare
devices. These standards are limited to the development and design risk assessment process, but they are not focused on the specific security requirements within the sophisticated deployment setting. Many security flaws and corresponding vulnerabilities like SQL injection and buffer overflow are a consequence of poor software design, which may be related to communication standards used in those devices [213]. The design aspects of different medical standards, such as 62304/82304/80002, are crucial for cybersecurity and are briefly described next:

- **IEC 62304:2006—Medical device software**: Software life cycle processes define the life cycle requirements for medical device software and software used within the medical devices. It establishes a common framework for the medical device software life cycle process.
- **ISO/IEC 27032:2012—Information technology–Security techniques**: This provides cybersecurity guidelines to improve the state of security. It brings out the other aspects of security like information security, network security, Internet security, and critical information infrastructure protection, among others, to highlight the essential practices in cybersecurity.
- **IEC 82304-1:2016—Health software—Part 1**: General requirements for product safety is a standard for the security of health software products designed to operate on general computing platforms (an evolution of IEC 62304). This standard works when health software is part of—or embedded in—a physical device [214]. Standards 82304 and 62304 both focus on the process of product design, software validation, maintenance, and testing.
- **ISO/IEC 8001—Risk management of medical devices on a network**: This defines the roles, responsibilities, and activities for information technology networks incorporating medical devices.
- **IEC/TR 80002-1:2009—Medical device software—Part 1**: This provides the guidance for the application to comply with the requirements contained in ISO 14971:2007 and also provides direction for implementing a risk management process for medical device software as part of the overall risk management process. It is the principal standard for risk management regulation.
- **ISO/TR 80002-2:2017—Medical device software—Part 2**: Validation of software for medical device quality systems is a technical report under development, which considers embedded and associated software with all medical devices. It includes many types of software used in device design, testing, and component acceptance, among others.
- **IEC/TR 80002-3:2014—Medical device software—Part 3**: The reference model of the medical device software life cycle (IEC 62304) defines the software life cycle processes and associated safety class definitions that derives from IEC 62304.

Although following the standards mentioned previously is a good practice in development life cycle processes, they do not deal with the fundamental cybersecurity protection required for the medical devices. Hence, the future research direction should enforce selecting a common communication protocol standard so that researchers can provide a universal solution for any threats to healthcare communication mediums and devices.

**Fault-tolerant design.** Reliability is the top priority in life-critical healthcare systems. During the time of manufacturing, typically it is possible to identify a large number of hardware or software defects, but exhaustive testing and complete fault coverage may not be possible. Through the simultaneous detection, diagnosis, and correction of fault effects, fault-tolerant designs enable a system to continue operating in the event of faults in its components. Additionally, it can be extended to cope with software errors caused by design inadequacies [215]. Although some types of redundancy like time, hardware, or information are required for fault tolerance, this redundancy often costs performance degradation or other overhead. Trip modular redundancy can be a good example, which employs three copies of a module and uses a majority vote to determine the final output [216]. Although it costs three times more than the original circuit, fault-tolerant design techniques may be warranted in safety-critical medical devices.
Intrusion detection mechanism to detect attacks. Ensuring the safety of patients is the key reason for the need to secure healthcare devices. An intrusion detection system is usually needed to detect any types of attack. An alert can be generated to notify patients or medical staff if an adversary is threatening the healthcare system or device. An intrusion detection system would monitor incoming traffic coming from the external device to the healthcare devices based on some pre-defined rules (e.g., the delay between two successive requests coming from an external device, the length of the message payload) [217, 218]. The research community should give more focus on developing a standard IDS where the pre-defined rules can be referred to different communication protocol standards in healthcare devices.

Fine-grained access control. Current access control mechanisms, such as attribute-based policy [160] and risk-based access control [157, 219], mostly focus on preventing unauthorized access to healthcare devices. However, these healthcare data transmit through different parts of a healthcare system, where the integrity and confidentiality of the data are not ensured. As a solution to this problem, a standard access control policy should be imposed on different components of a healthcare system to identify any security violation in data access.

Lack of a general platform. Different healthcare devices are using numerous hardware specifications, application software, communication protocols, and OSs. This heterogeneity makes it hard for security researchers to study existing threats and provide a common security solution for these healthcare devices. As a result, most of the security solutions are platform specific and unable to solve the wide range of security problems. Researchers, developers, and industry should work on developing a common standard for healthcare systems that will aid researchers in developing generalized security solutions.

Privacy-preserving healthcare system. Existing privacy solutions such as user and communication anonymity [166–168] mostly focus on preventing any unauthorized access to the healthcare data and communication channel. However, there remain several issues in healthcare systems that could affect the privacy of users and healthcare data if not addressed properly. For instance, several healthcare devices have physical addresses written in their body that violates the device anonymity requirement for the healthcare system. To ensure privacy, manufacturers should find a safer way to share device information with users, and researchers should put more focus on imposing a standard privacy policy on different components of the healthcare systems.

ML and big data. Modern healthcare systems generate a large amount of heterogeneous data and information daily, making it difficult for traditional methods to analyze and process it. One way to utilize these data in different healthcare applications (e.g., prognosis, diagnosis, treatment, clinical workflow) effectively is to implement big data analytics and ML models. With the introduction of big data analytics, researchers have been more focused on disease prediction based on big data analysis [220]. However, different big data processing phases, such as data collection, transformation, and modeling, are vulnerable to different communication channels and system-level attacks. Additionally, ML models in big data analysis exhibit unpredictable and overly confident behavior outside of the training distribution. Several recent works presented several adversarial ML-based attacks to identify the pitfalls in the underlying ML and big data models of modern healthcare systems [221, 222]. Hence, the research community should ensure secure and robust ML-based big data analytics in clinical settings, prognosis, diagnosis, and treatment.

Study of smart medical devices and existing threats. In recent years, the concept of IoT has been integrated into the medical domain, making modern medical devices context aware and smart. However, there are several attacks on IoT devices that can be adapted in a smart medical platform with minimal modifications. As the security requirements of IoT devices and healthcare systems vary a lot, existing security solutions of IoT devices often cannot ensure the end-to-end security need of medical IoT devices. Hence, it is necessary to study smart healthcare platforms, as well as the medical IoT domain, to identify the underlying threats and develop security solutions specific to the healthcare domain. The research community should give more focus on smart healthcare devices to address these threats, and industry experts should propose a standard practice for device manufacturing and application development to minimize these threats.
8 CONCLUSION

In this article, we presented an overview of the existing security and privacy research in healthcare systems. Increasing functional complexity, more software programmability, and growing wireless network connectivity are general trends observed in healthcare device applications. However, there are side effects to these trends that make the healthcare devices and applications increasingly vulnerable to security and privacy issues. We analyzed various aspects of the threats and how the current solutions overcome these threats. Given the critical tasks performed by healthcare devices, these issues should be addressed aggressively and proactively by the community. We believe that this survey will have a positive impact on the medical community by documenting recent attacks and defenses and facilitating a more aware ecosystem for security and privacy in healthcare systems.

REFERENCES

A Survey on Security and Privacy Issues in Modern Healthcare Systems: Attacks and Defenses


A Survey on Security and Privacy Issues in Modern Healthcare Systems: Attacks and Defenses


Received April 2020; revised October 2020; accepted February 2021