Design Challenges for Secure Implantable Medical Devices

Wayne Burleson
Department of Electrical and Computer Engineering
University of Massachusetts Amherst
Amherst, MA 01003
burleson@ecs.umass.edu

Shane S. Clark, Benjamin Ransford, Kevin Fu
Department of Computer Science
University of Massachusetts Amherst
Amherst, MA 01003
{ssclark,ransford,kevinfu}@cs.umass.edu

ABSTRACT

Implantable medical devices, or IMDs, are increasingly being used to improve patients’ medical outcomes. Designers of IMDs already balance safety, reliability, complexity, power consumption, and cost. However, recent research has demonstrated that designers should also consider security and data privacy to protect patients from acts of theft or malice, especially as medical technology becomes increasingly connected to other systems via wireless communications or the Internet. This survey paper summarizes recent work on IMD security. It discusses sound security principles to follow and common security pitfalls to avoid. As trends in power efficiency, sensing, wireless systems and bio-interfaces make possible new and improved IMDs, they also underscore the importance of understanding and addressing security and privacy concerns in an increasingly connected world.

Categories and Subject Descriptors
J.3 [Computer Applications]: Life and Medical Sciences—Medical information systems; C.3 [Computer Systems Organization]: Special-Purpose and Application-Based Systems—Real-time and embedded systems

General Terms
Security, Design

Keywords
Implantable Medical Devices, IMD Security

1. INTRODUCTION

Implantable medical devices (IMDs) perform a variety of therapeutic or life-saving functions ranging from drug infusion and cardiac pacing to direct neurostimulation. Modern IMDs often contain electronic components that perform increasingly sophisticated sensing, computation, and actuation, in many cases without any patient interaction. IMDs have already improved medical outcomes for millions of patients; many more will benefit from future IMD technology treating a growing number of ailments.

Because of their crucial roles in patient health, IMDs undergo rigorous evaluation to verify that they meet specific minimum safety and effectiveness requirements. However, security is a relatively new concern for regulatory bodies; bug-averse manufacturers have traditionally had little incentive to add security mechanisms that might cause problems or slow down regulatory approval. Perhaps not surprisingly in light of this situation, recent security research has demonstrated that some IMDs fail to meet appropriate expectations of security for critically important systems.

The key classes of IMD vulnerabilities researchers have identified are control vulnerabilities, in which an unauthorized person can gain control of an IMD’s operation or even disable its therapeutic services, and privacy vulnerabilities, in which an IMD exposes patient data to an unauthorized party. Both kinds of vulnerabilities may be harmful to patients’ health outcomes, and both kinds are avoidable.

As designers realign themselves with incentives for better security, there are ample opportunities to adapt well-tested security principles to IMD design. This survey paper’s goals are to (1) outline design principles for IMD security; (2) highlight the security challenges in designing implantable medical devices, some of which remain open problems; and (3) sketch the defensive measures that researchers have proposed and implemented.

1.1 Security Goals for IMD Design

The term security refers to the goal of well-defined, correct system behavior in the presence of adversaries.1 Security and reliability, both of which define policies and actions under a variety of conditions, form the basis of trustworthiness [9]. IMD designers can follow well-founded security practices to avoid pitfalls (§3) and build trustworthy systems. In short, designers should:

- Consider security in early design phases.
- Encrypt sensitive traffic where possible.
- Authenticate third-party devices where possible.
- Use well-studied cryptographic building blocks instead of ad-hoc designs.
- Assume an adversary can discover your source code and designs; do not rely on security through obscurity.
- Use industry-standard source-code analysis techniques at design time.
- Develop a realistic threat model (§1.2); defend the most attractive targets first.

1See Bishop’s textbook [4] for an introduction to security.
These design principles are not specific to IMDs; they are fundamental security ideas. Applying them to the IMD domain requires special consideration of IMDs' use cases and limitations. For example, the choice of cryptographic system to implement on a tiny biosensor or nonrechargeable heart device can have major implications for device longevity.

Halperin et al. detail some holistic design considerations related to medical-device security [14]. In contrast, this paper focuses on device-level concepts, relating the above principles to three specific classes of IMD (§2).

1.2 Threat Modeling

Threat modeling, which entails anticipating and characterizing potential threats, is a vital aspect of security design. With realistic models of adversaries, designers can assign appropriate priorities to addressing different threats.

The severity of vulnerabilities varies along with the sensitivity of the data or the consequences of actuation; there is no "one size fits all" threat model for IMDs. A non-actuating glucose sensor incurs different risks than a defibrillator that can deliver disruptive electrical shocks to a heart.

Adversaries are typically characterized according to their goals, their capabilities and the resources they possess. Security designers evaluate each threat by considering the value of the target and the amount of effort necessary to access it. Recent work analyzing IMD security and privacy has posited several classes of adversaries, described below.

An eavesdropper who listens to an IMD's radio transmissions, but does not interfere with them, can often learn private information with minimal effort. Such a passive adversary may have access to an oscilloscope, software radio, directional antennas, and other listening equipment. Several studies have considered this type of adversary and demonstrated that eavesdropping on unencrypted communications could compromise patients' data privacy [15, 20, 23, 26, 28].

An active adversary extends the passive adversary's capabilities with the ability to generate radio transmissions addressed to the IMD, or to replay recorded control commands. Halperin et al. demonstrated that an active adversary with a programmable radio could control one model of implantable defibrillator by replaying messages—disabling programmed therapies or even delivering a shock intended to induce a fatal heart rhythm [15].

Adversaries are typically characterized according to their goals, their capabilities and the resources they possess. Security designers evaluate each threat by considering the value of the target and the amount of effort necessary to access it. Recent work analyzing IMD security and privacy has posited several classes of adversaries, described below.

An eavesdropper who listens to an IMD’s radio transmissions, but does not interfere with them, can often learn private information with minimal effort. Such a passive adversary may have access to an oscilloscope, software radio, directional antennas, and other listening equipment. Several studies have considered this type of adversary and demonstrated that eavesdropping on unencrypted communications could compromise patients’ data privacy [15, 20, 23, 26, 28].

An active adversary extends the passive adversary’s capabilities with the ability to generate radio transmissions addressed to the IMD, or to replay recorded control commands. Halperin et al. demonstrated that an active adversary with a programmable radio could control one model of implantable defibrillator by replaying messages—disabling programmed therapies or even delivering a shock intended to induce a fatal heart rhythm [15].

Another adversarial capability is binary analysis, the ability to disassemble a system’s software and in some cases completely understand its operation. By inspecting the Java-based configuration program supplied with his own insulin pump, researcher Jerome Radcliffe reverse-engineered the pump’s packet structure, revealing that the pump failed to encrypt the medical data it transmitted or to adequately authenticate the components to one another [26].

In contrast to design-time static analysis of source code, a crucial practice that may expose flaws before devices are shipped [18], binary analysis involves inspecting compiled code; it can expose flaws in systems that erroneously depend on the supposed difficulty of reverse engineering to conceal private information.

In the context of medical conditions, it may be difficult to comprehend why a malicious person would seek to cause harm to patients receiving therapy, but unfortunately, it has happened in the past. For example, in 2008, malicious hackers defaced a webpage run by the nonprofit Epilepsy Foundation, replacing the page’s content with flashing animations that induced migraines or seizures for some unsuspecting visitors [24]. Although we know of no reports of malicious attacks against IMDs “in the wild,” it is important to address vulnerabilities before they become serious threats.

2. DEVICES IN DEPTH

To illustrate the complexity of the design space for IMD security, we offer three examples of IMD systems that pose different security challenges because of their different design and usage. The common thread among all three devices—insulin pump systems, implantable cardioverter defibrillators, and subcutaneous biosensors—is that security is a crucial design concern. Section 2.4 explores commonalities and defensive concepts.

2.1 Insulin Pump: Open-Loop System

Insulin pump systems straddle the boundary between implanted and external systems, including some components that are physically attached to a patient and others that are external. A typical modern insulin pump system (IPS) may include: an insulin infusion pump with wireless interface that subcutaneously delivers insulin, a continuous glucose monitor (CGM) with wireless transmitter and subcutaneous sensor for glucose measurement, and a wireless remote control that the patient can use to alter infusion pump settings or manually trigger insulin injections. The CGM automatically takes frequent glucose readings, presenting the data to the user via a screen or PC, or sending data directly to the pump. The pump automatically provides basal doses for insulin maintenance and can also administer larger bolus doses to compensate for large insulin spikes that may result from, e.g., a meal. Finally, the remote control provides a convenient interface for the user to adjust pump settings without using the pump controls and screen typically attached at the abdomen. Figure 1 shows a block diagram of an IPS.

Insulin pump systems exemplify open-loop IMDs: they require patient interaction to change pump settings. Specifically, the patient’s remote control—but not the CGM—directly controls pump actuation. Because the remote-control interface carries crucial information and control signals, initial security studies have focused on finding vulnerabilities at this interface. Li et al. discovered that one IPS’s communications were unencrypted, leading to potential disclosure of private patient information (e.g., glucose levels) [20]. They also found that the components failed to check their inputs appropriately, allowing the researchers to inject forged packets reporting incorrect glucose levels to the patient and pump—and more alarmingly, to issue unauthorized pump-control commands. Soon thereafter, two security researchers independently demonstrated full control of IPSes via circumventing authentication mechanisms: Radcliffe compromised the wireless channel of his own (unspecified)
power consumption. Instead of using the ICD’s radio in a low-power sleep mode, with a small chance of overheating of tissue that might occur during battery recharging. In contrast, the ICD’s radio in a high-power active mode, indefinitely transmitting packets at a regular rate and dramatically increasing the ICD’s power consumption.

Concerning the battery, the study found that a sequence of transmissions from the software radio could keep the ICD’s radio in a high-power active mode, indefinitely transmitting packets at a regular rate and dramatically increasing the ICD’s power consumption. Although the ICDs spend most of their time in low-power sensing states, they also include radios for clinical programming and at-home status reporting.

### 2.2 Defibrillator: Closed-Loop System

Like an artificial pacemaker, which continually issues small electrical pulses to heart muscle to maintain a healthy rhythm, an **implantable cardiac defibrillator (ICD)** is implanted under the skin near the clavicle. ICDs extend the capabilities of artificial pacemakers with the ability to issue large (tens of joules) shocks to “reset” an unsustainable heart rhythm (arrhythmia). Figure 2 shows a block diagram of an ICD.

Unlike an insulin pump that accepts patient input via a user interface, a fully implanted device such as an ICD is a **closed-loop** system: under normal circumstances, its sensing function alone dictates its actuation activities. **Closed-loop IMDs typically also have special modes for in-clinic configuration and operation.** Halperin et al. enumerated the security and privacy challenges of closed-loop implanted systems in a 2008 article [14], focusing primarily on the tensions between security and utility.

ICD implantation currently requires invasive surgery with a risk of complications (infection or death) [11], so ICDs are designed to last for at least five years once implanted—resulting in long design and deployment cycles for manufacturers. ICDs draw power from single-use batteries, sealed inside the case, to provide uninterrupted monitoring throughout the device’s lifetime and to avoid the heating of tissue that might occur during battery recharging. In conformity with these design choices, ICDs spend most of their time in low-power sensing states. They also include radios for clinical adjustments and at-home status reporting.

A 2008 security analysis of a commercial ICD found vulnerabilities in multiple subsystems [15], including those listed above. Focusing on the ICD’s radio link, researchers used open-source software-radio tools to record transmissions between the ICD and a clinical programming console. Offline analysis of these traces revealed patient information in clear text without evidence of encryption. They replayed recorded traces of clinical therapy commands and found that they could control or disable the ICD’s therapies with their software radio. Concerning the battery, the study found that a sequence of transmissions from the software radio could keep the ICD’s radio in a high-power active mode, indefinitely transmitting packets at a regular rate and dramatically increasing the ICD’s power consumption.

### 2.3 Biosensors for Data Acquisition

Implantable biosensors (Figure 3) are **IMDs** that measure biological phenomena and send data to a more powerful device for storage or analysis. Biosensors are a broader device category than insulin pump systems or defibrillators, representing a wide range of both signals and signal processing techniques. They are subject to a third set of security and privacy challenges that does not completely overlap with those mentioned above.

Biosensors range from high-data-rate imaging devices for the eye [3] or brain [25] to extremely low-data-rate sensors for glucose [12] or other metabolites in the blood [5]. Actuators that consume biosensor data can control potentially lethal drug-delivery systems [23] or electrical therapies [15]. Keeping biosensor data confidential is important because it can be used in illegal or unethical ways including insurance fraud or discrimination. The provenance (origin) and timestamp information that accompany biosensor readings are also critically important for medical care and must be protected from tampering.

Subcutaneous biosensors present a special set of security and privacy requirements. A subcutaneous sensor [5] involves an implanted biosensor that acts as a **lab on chip**, conducting a small experiment at the molecular or electro-chemical level on the sensor. Current subcutaneous sensors can detect drugs, bio-markers, and antibodies and may eventually examine DNA, and simultaneously log temperature, pH, and other phenomena.

Recent examples of subcutaneous biosensors include **injectable** subcutaneous devices that are remotely powered by a bandage-like patch that also provides a data link to a higher-level wearable device, possibly a body-area-network (BAN) or eventually a higher-level health information system. A related class of devices are **low-cost disposable biosensors** for detecting infectious disease or critical levels of glucose and lactate in a battlefield or other trauma situation [13]; such devices penetrate the skin for communication and power. These two classes of devices support different threat models because of their different usage parameters.

Biosensors that are fully implanted must communicate wirelessly to transmit through tissue. (Some receive power through tissue as well; recent work has shown that remotely powering biosensors is feasible at gigahertz frequencies that enable millimeter-sized antennas [22].) A key problem with fully implanted sensors is that small, infrequent wireless transmissions may pose a greater privacy risk than large or continuous transmissions. For example, a sensor may take several minutes to complete its task, then deliver only a few bytes of data—giving this information a high value per bit that may make it an attractive target. Short data transmissions necessitate careful use of a cipher, especially if the plaintext sensor data may take only a few different values. The small amount of data also...
has little inherent redundancy, making error-correction necessary. When a biosensor includes a patch that is meant to pair with the sensor, additional risks arise. Although eavesdropping on a properly operating tag may be unlikely because of the short (several millimeters for a subcutaneous sensor) nominal transmission range, impersonation of both the clinical reader and the patch are plausible concerns. For example, the patch of an unconscious patient can be removed and replaced by another patch. Similarly, a rogue sensor can upload fraudulent data to a trusted patch. All components involved should authenticate one another using well-studied cryptographic mechanisms, especially during the critically important period when a sensor is first being tested or calibrated.

Biosensors present a diverse set of challenges for security and privacy and a unique combination of constraints. Open problems include: 1) developing more detailed threat models; 2) exploring design alternatives that effectively trade off safety, security, and utility; 3) understanding energy issues, including power depletion and side-channel attacks that exploit the lightweight nature of the biosensor; 4) implementing multiple layers of security to accommodate the multiple stages required to access data from a lightweight sensor (implant to patch to wearable to internet); and 5) understanding the security and privacy implications of future biosensing devices that provide an unprecedented view into the (presumably private) inner workings of the human body. Future devices are likely to include more storage, more complex signal processing, integrated software control, and use of multiple intercommunicating sensors, all of which will complicate security and privacy issues.

2.4 Common Threads

Different classes of IMDs have distinct hardware and usage constraints, but there are important security considerations that apply to many IMDs. Researchers investigating the security and privacy of IMDs have also proposed several domain-specific mechanisms that apply broadly. This section discusses some of these common threads in the context of our example IMD systems.

All of the IMD vulnerabilities disclosed thus far could be mitigated by the use of encryption on radio links. Hosseini-Khayat presents a lightweight wireless protocol for IMDs [17] that leverages well-studied wireless and cryptography technologies and emphasizes low-energy computation. The choice of encryption scheme should consider the nature of the data as well as the device constraints. Fan et al. contribute hardware implementations of the stream cipher Hummingbird [7, 8]. Beck explores the use of block ciphers in IMD security [3].

Unfortunately, encryption is not a panacea for IMD security and privacy vulnerabilities; many questions remain. If the radio link were to use encryption, how would the necessary secret key material be distributed, and by whom? How should an IMD authenticate external entities, and how should it determine whether a particular entity is allowed to communicate with it? Even assuming that each of these questions can be answered, successful implementation of encryption would not completely address known risks. Encryption alone fails to address replay attacks, and previous work has demonstrated that encryption may not sufficiently conceal characteristic traffic patterns [16]. Furthermore, since some IMDs must “fail open” to allow emergency access (e.g., to disable the IMD during emergency procedures), how can it also provide security in non-emergency situations? Should an IMD raise an alarm (perhaps tactile or audible) when a security-sensitive event occurs? These questions are largely open.

Recent research toward addressing these design tensions has proposed new techniques and auxiliary devices to provide fail-open security for IMDs. Rasmussen et al. proposed the use of ultrasonic distance bounding to enforce programmer proximity [27]. Li et al. proposed body-coupled communications for the same purpose [20], hoping to prevent an adversary from launching a long-range radio-based attack. Both of these distance-bounding techniques require new hardware, but this constraint may not represent a major stumbling block for IPSes or biosensor systems, which are short-lived and non-invasive compared to ICDs.

Researchers have also proposed defenses specifically targeted toward existing ICDs, but which may be useful for other IMDs. Denning et al. proposed that an IMD be paired with a cloaker that would provide authentication services whenever it was present, and allow open communication otherwise [6]. Xu et al. proposed the Guardian, a device that would pair with an IMD and use radio jamming to defend against eavesdropping and unauthorized commands [30]. Gollakota et al. independently proposed an auxiliary device called the shield that would use “friendly” radio jamming to proxy an ICD’s communications to an authorized reader [10]. The shield is designed for compatibility with devices that are already implanted, reducing the burden on device designers to address the security vulnerabilities in devices that have not completed their deployment lifecycles.

3. SECURITY PITFALLS

Designing for security has many subtleties. In the context of IMDs, where devices may be physically inaccessible for years, it is particularly important to avoid design errors that lead to failures or recalls later. One common error is believing in security through obscurity—relying entirely on proprietary ciphers or protocols for secrecy.

Security through obscurity—relying entirely on the secrecy of proprietary ciphers or protocols—is a common fallacy. Sound security principles dictate that a system’s security must not depend on the secrecy of the algorithm or hardware; it is better to use well-studied standard ciphers and spend more design effort protecting cryptographic keys. This principle, commonly known as Kerckhoff’s principle,2 is a fundamental guideline for security design. Following it is essential for resistance against reverse-engineering adversaries.

A recent example that illustrates the hazards of security through obscurity is that of the NXP Mifare Classic smart-card chipset, which is widely used for transit ticketing systems. Nohl et al. reverse-engineered the Mifare Classic hardware and analyzed the underlying cipher and protocol, discovering that it used a flawed implementation of a cipher called Crypto-1 [21]. Crypto-1 supports only a limited key size; the Mifare Classic hardware also implements a predictable random-number generator. These factors combine to allow an adversary to clone a tag in a matter of seconds. The Mifare Classic tag could have addressed these flaws by using established, publicly studied cryptographic primitives rather than ad-hoc proprietary systems.

4. OPEN PROBLEMS IN IMD DESIGN

IMDś are first and foremost intended to improve patients’ quality of life. To this end, the primary focus for designers must be device safety and utility. We argue that security and privacy are also important properties that must be part of the design process, but there is the potential for direct conflict between these two sets of properties.

The issue of emergency access highlights some of the tensions
that exist among these properties. Requiring users to authenticate to a device before altering its functionality is a boon for security, but it introduces risks in the case of an emergency. A medical professional may need to reprogram or disable a device to effectively treat a patient. As discussed in Section 2.4, encryption or other strong authentication mechanisms could make such emergency measures impossible if the patient is unconscious or the facility does not possess a programming device with a required shared secret.

For some IMDs, including both IPSes and ICDs, designers must carefully weigh the energy costs of encryption against safety and utility. A heavyweight encryption scheme could potentially drain enough energy to require more frequent device replacement—a surgical procedure for ICD patients and a persistent burden for IPS users. Costly encryption could even make the construction and deployment of some subcutaneous biosensors infeasible. It remains to be seen whether ASIC implementations of lightweight algorithms can effectively mitigate this issue because of the lack of public deployments to date [17, 3].

There are no clear-cut methods for resolving these tensions, and there is little publicly available information about whatever steps manufacturers have already taken. While cryptographers and security researchers have long embraced Kerckhoff’s principle, device manufacturers employ proprietary systems and generally do not comment (for business reasons) on security measures that they may employ. These closed ecosystems hamper industry-wide progress on shared issues such as security and privacy. Research into whole-system modeling and formal analysis of medical devices [19, 2, 1] offers hope that future IMDs will integrate sound security principles at design time, but the time horizon for industrial adoption may be long.

Recent analyses of implantable medical devices have revealed a number of security and privacy failings, but researchers are developing novel solutions to the problems IMD designers face. By incorporating security and privacy design principles into the development process, IMD designers have the opportunity to address these issues before they become larger threats.

5. ACKNOWLEDGMENTS

This material is based upon work supported by: the Armstrong Fund for Science; the National Science Foundation under Grants No. 831244, 0923313 and 0964641; Cooperative Agreement No. 90TR0003/01 from the Department of Health and Human Services; two NSF Graduate Research Fellowships; and a Sloan Research Fellowship. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of DHHS or NSF.

6. REFERENCES


