Editors: Mohamed Kaâniche, mohamed.kaaniche@laas.fr | Aad van Moorsel, aad.vanmoorsel@ncl.ac.uk

Security and Interoperable-Medical-Device Systems, Part 2: Failures, Consequences, and Classification

Eugene Y. Vasserman | Kansas State University
Krishna K. Venkatasubramanian | Worcester Polytechnic Institute
Oleg Sokolsky and Insup Lee | University of Pennsylvania

Medical devices are gaining considerable communication capabilities, allowing them to interact with the devices around them. This interoperability presents many benefits for clinical workflows and patient care outcomes. Examples include increased safety, usability, and decision support, as well as decreased false alarms and clinician cognitive workload.[1]

In such open interoperable medical device (IMD) environments, security becomes crucial for safe operation because of an increased adversarial presence in the network. Efforts are underway to proactively develop interoperability standards, involving all key stakeholders (manufacturers, clinical facilities, regulating agencies, and so on).[2–5] We must work to ensure that the resulting systems and protocols prevent unauthorized, unsafe interaction.

In the first part of this two-part article,[6] we defined an abstract model for IMD environments and an associated attack model in the context of the integrated-clinical-environment architecture.[2] The IMD environment consists of a *coordinator* (which facilitates interoperability), a network connecting the coordinator and medical devices, and an alarm system. The alarm alerts clinicians to issues both functional (for example, loss of a device or the coordinator) and medical (for example, an abnormally low patient heart rate). Individual medical devices might have their own alarms for functional and medical issues, which would complement the IMD alarm system's capabilities. Compare that with today, when devices have their own dedicated alarms, sometimes complemented by alerts sent to a central location such as the nurses' station.

In this part, we define a failure model, or the specific ways in which IMD environments might fail when attacked. In addition, an *attack-consequences model* expresses the combination of failures experienced by IMD environments for each attack vector. This analysis leads to interesting conclusions about regulatory classes of medical devices in IMD environments subject to attacks.

A Failure Model

Attacks might cause the IMD environment to fail in arbitrary ways. For our purposes, a failure is an adverse effect on a patient due to an adversary's actions. This includes leakage of sensitive patient information; untimely, incorrect, or no treatment (actuation); untimely or no monitoring; or alarm deactivation. In part one of this article, we defined five kinds of attacks:

- *Destroy* attacks try to physically destroy a device or its components.
- *Disturb* attacks try to disturb and alter the functionality of a device or the IMD environment.
- *Reprogram* attacks try to reprogram a device (this is a subset of the disturb attack).
- *Denial of service* (DoS) attacks try to deny service to devices or the entire IMD environment.
- *Eavesdrop* attacks try to eavesdrop on communication.[6]

Researchers have proposed numerous failure models in computing[7,8] and noncomputing settings,[9] but these models didn't take into account device interoperability and

 Copublished by the IEEE Computer and Reliability Societies

cyberphysical systems. We propose a (limited but extensible) failure model for IMD environments under attack, expressed in terms of the system's failure. That is, an IMD environment as a whole fails in a specific way if an individual component fails.

Failure Modes

Our failure model involves the following four failure modes.

In *fail-stop*, one or more IMD environment components abruptly stop operating and can't be restarted easily. For example, someone burns out an infusion pump's motor by forcing excessive use of the pump.

In *fail-safe*, one or more IMD environment components stop or alter their operation, enter a safe state, and can be restarted easily. That is, a device

- goes offline and stops operating if stopping won't harm the patient, or
- continues operating if it's administering treatment crucial to the patient's health.

For example, this might occur if an attack destroys the coordinator, causing an x-ray scanner to deactivate.

In *fail-loud*, an alarm sounds in response to the stoppage, alteration, or degradation of functionality of one or more components, including the network. Generally, the coordinator controls these alarms. For example, if a device suddenly stops or misbehaves, the coordinator, observing this, instructs the alarm system to sound. An individual device might likewise generate an alarm if it detects a problem, such as unexpected disconnection from the network.

In *fail-quiet*, an IMD environment component stops, alters, or degrades its operation quietly, without raising an alarm. For example, a wireless blood pressure monitor starts broadcasting unencrypted copies of all its data, allowing an adversary-controlled (unauthorized) device to eavesdrop on it.

Intersections

The two basic types of failures are fail-loud and fail-quiet. Within these, there are two overlapping subsets: fail-stop and fail-safe. Five intersections are possible:

- Fail-stop ∩ fail-safe. Some devices, such as x-ray scanners and patient-controlled-analgesia pumps, are designed to be fail-safe when stopped abruptly. In an IMD environment, this can lead to fail-loud or fail-quiet, depending on whether the alarm system continues to function or fails.
- Fail-quiet ∩ fail-stop. Devices might fail both abruptly and quietly. This is especially true when failure of both functionality and the alarm system occurs simultaneously.
- Fail-quiet ∩ fail-safe. This case occurs when the alarm fails but the IMD environment enters a safe state.
- Fail-loud ∩ fail-stop. This case occurs when a device fails but the alarm alerts clinicians to the failure.
- Fail-loud ∩ fail-safe. Some IMD environments will sound an alarm when a device enters a safe state.

Fail-loud ∩ fail-quiet is, of course, impossible because the alarm system either generates an alarm or stays quiet.

An IMD environment can potentially experience multiple combinations of failures due to an attack. In such cases, we address the combination that's most dangerous to the patient or caregiver.

The Attack-Consequences Model

Previous research in attack-centric modeling focused on attack trees,[10] attack-intention models,[11] and capability-vulnerability models.[12] These models weren't tailored to handle cyberphysical systems such as medical devices, which can have attacks and consequences in both the cyber and physical realms.

Here, we discuss combinations of the attack vectors we described in part one and their consequences in terms of failures in the IMD. Owing to space limitations, we can't enumerate all such attacks, so our list comprises 10 representative scenarios:

- Scenario 1a involves a destroy or DoS attack on the coordinator. The coordinator's inability to respond causes the alarm system to sound. Individual devices, unable to reach the coordinator, go offline and might sound their internal alarms.
- Scenario 1b involves a disturb or reprogram attack on the coordinator. The alarm system might eventually sound an alarm if it detects abnormal patient health indicators.
- Scenario 1c is an extension of 1b in which both the coordinator and alarm system are compromised. The alarm system should be designed so that attackers cannot silence it without attacking it directly.
- Scenario 2a involves a destroy or DoS attack that causes one or more devices to stop abruptly.
- Scenario 2b is an extension of 2a in which one or more devices and the alarm system are compromised, leading to fail-quiet for the IMD.
- Scenario 2c involves a disturb or reprogram attack causing one or more devices to misbehave. The alarm system isn't attacked.
- Scenario 2d is an extension of 2c in which the alarm system is also compromised, leading to fail-quiet for the IMD.
- Scenario 3a involves a disturb attack on the network—for example, modifying the packets being sent or selectively dropping them.

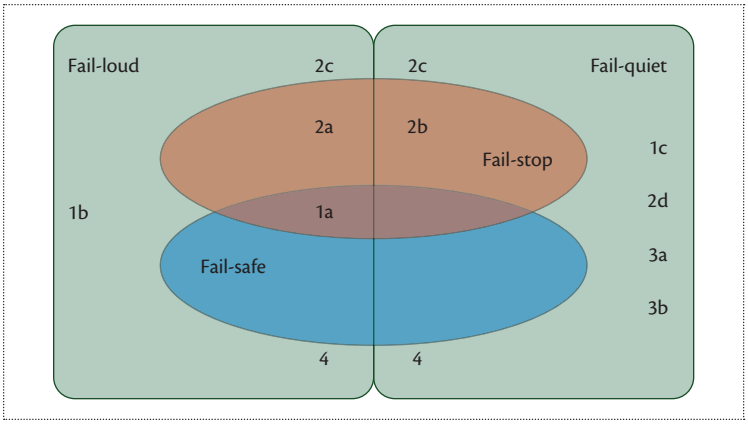

Figure 1. The attack-consequences model comprises four failure modes. Each number or combination of a number and letter indicates a scenario described in the main article.

- Scenario 3b involves an eavesdrop attack on the network—for example, listening in on communication between entities.
- Scenario 4 is when the alarm system is compromised and fails completely or partially.

Figure 1 shows where these scenarios fit in the attack-consequences model.

We don't consider cases in which the coordinator, medical device, and network fail in various combinations in conjunction with the alarm system. That's a subject for future research.

Device Classification Consequences

The attack-consequences model only lists the possible failures resulting from various attacks; it provides no information on the failures' extent. The extent depends on, among other things, the device's regulatory class—its capability to do harm.

The US Food and Drug Administration (FDA) ranks devices into three regulatory classes:

- Class I devices can't cause temporary discomfort or permanent harm if they malfunction or are used incorrectly.
- Class II devices might cause temporary discomfort.
- Class III devices might cause permanent damage.[13]

The FDA currently categorizes devices as either stand-alone units or fixed interoperability configurations. An IMD environment can't add new device types without reevaluating and potentially reclassifying the entire system.[14]

When devices are placed in an IMD environment subject to attacks, classification issues arise. For instance, the regulatory class of what was a stand-alone device might spill over to other IMD environment components. A simple example involves a hospital's internal network. If a class III device requires the network for correct functionality, the entire network might have to be deemed a class III device.

Furthermore, when considering potential threats in an IMD environment, a device's classification might change. For instance, in scenario 1c, the alarm system's failure will lead to fail-quiet, with potentially the most severe consequences of any failure. So, the alarm system is a class III device because many other devices' safety depends on it functioning correctly, irrespective of its regulatory status as a stand-alone device.

Toward a Revised Classification

Owing to the spillover of device classifications and the changing nature of device classes in IMD environments, perhaps a more fine-grained medical-device classification is warranted. Such a refinement would seem especially useful when considering attacks that would induce deviations from the IMD environment's expected behavior. Attacks' consequences might differ depending on the affected medical device's importance to patients' health and the time between the attack and damage to health.

One approach could be to amend the FDA classification scheme to define at least one new class: IIIa. Devices in this class might cause permanent harm if they malfunction without notice for longer than, for example, 15 minutes. Current class II devices would be mostly unaffected by this change. However, some that are fail-quiet could move to IIIa because the period of time over which they malfunction is unbounded. One possible example is a stationary x-ray scanner, currently classified as class II, possibly owing to the large time scale required to deliver dangerous doses of radiation. Over time, however, patients or clinicians might receive excessive exposure if the radiation source doesn't disengage. This device could, with adversaries present, become IIIa.

In an IMD environment, a successful attack might change a device's behavior. The coordinator will raise an alarm if other devices connected to the patient observe degradation of vital signs. If the network or coordinator is likewise attacked, alerts might never propagate to the alarm system.

To keep the benefits of fail-loud, we must ensure that any attack affecting a device or the coordinator would affect the alarm system in a way that activates it. This implies two things. First, the IMD environment

won't survive a destroy attack on the alarm. Second, a nonreprogrammable alarm subsystem will be impervious to disturb or reprogram attacks and will handle DoS attacks on the network or coordinator by activating an alarm. Disturb or reprogram attacks on the coordinator are more challenging. However, they're solvable as long as the alarm system listens to device messages on the network independently of the coordinator. If we can be certain that, when an IMD component fails, an alarm will sound in some bounded time interval, the consequences of the IMD environment deviating from the expected behavior are mitigated. Patient health will be preserved as long as someone can hear the alarm and fix the problem.

So, medical-device classification should take into account time and the potential for human action. The potential for immediate patient harm makes a device high-hazard, but delayed harm with guaranteed fail-loud is a lower-hazard classification. However, this system architecture can't guarantee fail-loud behavior if all component devices as well as the coordinator are reprogrammed to give false readings. Because devices are falsifying reported data, the alarm will never sound if it neither gets a command to do so nor detects problems with patients' health.

Therefore, we plan to develop a more fine-grained classification scheme for IMDs. Once the classification system is sufficiently expressive, we'll be able to abstract out the properties of individual IMD components and reason about them in terms of their classes. We can then expand the attack-consequences model to include the information on the consequences' extent. ■

Acknowledgments

US National Institutes of Health grant 1U01EB012470-01 and US National Science Foundation award CNS-1035715 partly supported this research.

References

1. D. Arney et al., "Biomedical Devices and Systems Security," *Proc. 33rd Ann. Int'l Conf. IEEE Eng. in Medicine and Biology Soc.* (EMBC 11), IEEE, 2011, pp. 2376–2379.
2. *ASTM F2761 - 09 Medical Devices and Medical Systems—Essential Safety Requirements for Equipment Comprising the Patient-Centric Integrated Clinical Environment (ICE)—Part 1: General Requirements and Conceptual Model*, ASTM F29.21, ASTM Int'l, 2009.
3. M. Clarke et al., "Developing a Standard for Personal Health Devices Based on 11073," *Proc. 29th Ann. Int'l Conf. IEEE Eng. in Medicine and Biology Soc.* (EMBC 07), IEEE, 2007, pp. 6174–6176.
4. "Introduction to HL7 Standards," Health Level Seven Int'l, 2012; www.hl7.org/implement/standards.
5. *Integrating the Healthcare Enterprise*, IHE Int'l, 2012; www.ihe.net.
6. K.K. Venkatasubramanian et al., "Security and Interoperable-Medical-Device Systems, Part 1," *IEEE Security & Privacy*, vol. 10, no. 5, 2012, pp. 61–63.
7. B.M. O'Halloran, R.B. Stone, and I.Y. Tumer, "A Failure Modes and Mechanisms Naming Taxonomy," *Proc. 2012 Ann. Reliability and Maintainability Symp.* (RAMS 12), IEEE, 2012.
8. S.J. Uder, R.B. Stone, and I.Y. Tumer, "Failure Analysis in Subsystem Design for Space Missions," *Proc. 16th Int'l Conf. Design Theory and Methodology*, ASME, 2004, pp. 201–217.
9. J.A. Collins, *Failure of Materials in Mechanical Design: Analysis, Prediction, Prevention*, 2nd ed., Wiley Interscience, 1981.
10. S. Mauw and M. Oostdijk, *Foundations of Attack Trees*, LNCS 3935, Elsevier, 2005.
11. P. Wu, W. Zhigang, and C. Junhua, "Research on Attack Intention Recognition Based on Graphical Model," *Proc. 2009 5th Int'l Conf. Information Assurance and Security* (IAS 09), vol. 1, IEEE CS, 2009, pp. 360–363.
12. S. Song et al., "Capability-Centric Attack Model for Network Security Analysis," *Proc. 2nd Int'l Conf. Signal Processing Systems* (ICSPS 10), vol. 2, IEEE, 2010, pp. 372–376.
13. "Device Classification Panels," US Food and Drug Administration, 2012; www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/ClassifyYourDevice/ucm051530.htm.
14. J. Hatcliff et al., "An Overview of Regulatory and Trust Issues for the Integrated Clinical Environment," *Proc. 2011 High-Confidence Medical Device Systems and Software Workshop*, 2011.

Eugene Y. Vasserman is an assistant professor in Kansas State University's Department of Computing and Information Sciences. Contact him at eyv@ksu.edu.

Krishna K. Venkatasubramanian is an assistant professor in the Worcester Polytechnic Institute's Department of Computer Science. Contact him at kven@wpi.edu.

Oleg Sokolsky is a research associate professor in the University of Pennsylvania's Department of Computer and Information Science. Contact him at sokolsky@cis.upenn.edu.

Insup Lee is the Cecilia Fitler Moore Professor of Computer and Information Science at the University of Pennsylvania. Contact him at lee@cis.upenn.edu.