Assuring Critical Systems

Bob Blakley
Tivoli Systems, Inc.
blakley@us.ibm.com

1. PANEL DESCRIPTION

This position paper was prepared in response to the call for positions for the "Assurance in Life- or Nation-Critical Endeavors " panel at the 2002 ACM New Security Paradigms Workshop.

The call for positions read as follows:

"GIVEN that biometrics and other intertwined technologies will be used to supplement the work of people, and GIVEN that misinterpretation and error are highly likely to result in the deprivation of life, liberty, and happiness (AKA these systems are "criticality one," or "man rated"), THEN we need to identify the technical measures that are needed in the fundamental systems."

2. WHAT SHOULD BE ASSURED?

In order to identify what technical measures might protect humans against deprivation of life, liberty, and happiness in the face of failure (or success!) of security-related technologies, including biometrics, we must first identify what these technical measures should do.

Presumably a general statement of what technical measures should do is "support correct operation of the protective functions of the critical systems in which they are imbedded, while preventing harm to people arising from malfunctions of or unanticipated side-effects of the operation of the system ".

In order to support correct operation of the system's protective functions, technical and other measures should assure:

2.1 THAT THE SYSTEM ADDRESSES THE CORRECT PROBLEM.

For example, imagine a system whose goal is to prevent people who intend to perform terrorist acts from boarding an aircraft. One could design a system which uses biometric technology to identify passengers before they board the aircraft. This system might correctly identify previously known or suspected terrorists and keep them off the plane. It would certainly not keep a previously unknown person, with no known or suspected terrorist links (but with the intention to perform a terrorist act) from boarding the plane. The issue here is that the problem which really needs to be solved is not determining identity, but determining intent. A system designed to determine identity will solve part of the problem, but not all of it.

When investigating the problem statement, the following questions should be asked:

What security property or properties does the system need to ensure?

- Availability of the system and its services to authorized users
- Integrity of the system's resources
- Integrity of the system itself (i.e. correct functioning)
- Identification of users
- Identification of system components to users
- Prevention of unauthorized use
- Accountability of users for their actions in the system

How does the system approach ensuring security properties?

- Transfer Liability
- Indemnify Against Loss
- Mitigate Risk
- Prevent damage
- Detect threats
- Detect incidents
- Respond to incidents
- Recover from incidents

What should the system do if it cannot ensure the required security properties?

- Halt
- Pause and generate an alert to a human
- Continue and make a record of the event

Under what circumstances are other quality attributes more important than security?

- Threat to human life or safety
- Threat to individual privacy
- Threat to property
- Threat to national security
- Threat to organizational liability

2.2 THAT THE SYSTEM IS EFFECTIVE IN REDUCING THE SEVERITY OF THE PROBLEM.

Return to the example in the previous paragraph. Many biometric technologies are easily spoofed (fingerprint scanners can be deceived by using gelatin fingerprint molds;

New Security Paradigms Workshop '02, September 23-26, 2002, Virginia Beach, Virginia.

face recognition scanners can be fooled using pictures of faces; iris scanners can be fooled using pictures of irises with the pupils cut out, etc....) Here technical means can be of use; presumably improved technology can increase the difficulty of spoofing the system. Non-technical means may be easier and cheaper to deploy, however. It's fairly easy for a human guard to tell if you're holding your face (as opposed to a picture of someone else's face) up to a face recognition scanner.

2.3 THAT ALL RELEVANT LEGAL SYSTEMS PERMIT OPERATION OF THE SYSTEM IN A MODE WHICH MEETS THE SYSTEM'S GOALS.

If the system's goals include supporting civil or criminal prosecutions, some mechanism for establishing intent must be designed into the system. Here, for example, the difference between identification (the establishment of an individual's identity without requiring the voluntary participation of the individual) and authentication (the establishment of an individual's identity based on a voluntary affirmative act performed by the individual) is critically important. Technical mechanisms may be required to insure that individuals' intent can be established with confidence and demonstrated convincingly to a third party after the fact.

Technical mechanisms must be designed to support legal deployment and use of the system where it needs to be used. So, for example, systems which rely upon cryptography must be designed with laws regulating export, import, deployment, and use of cryptography in mind.

If the system's goals include supporting detention of individuals or generation of evidence to be used in prosecutions, technical mechanisms may need to be used to assure admissibility of the evidence the system generates or collects. Technical mechanisms may also be required to insure integrity of the evidence, and to maintain a trustworthy chain of custody of the evidence.

2.4 THAT THE SYSTEM'S DETECTION-RESPONSE CYCLE IS SUFFICIENTLY SHORT TO DEFEAT ANTICIPATED THREATS.

If the system's goals include prevention of losses due to changing risks or threats, technical mechanisms may be required to insure that the emergence of new risks or threats is detected in a timely fashion, that the emergence of new risks or threats is communicated to responders in a timely fashion, and that countermeasures designed to prevent losses can be implemented in a timely fashion.

In order to prevent harm to people arising from malfunctions of or unanticipated side-effects of the operation of the system, technical and other measures should assure:

2.5 THAT FAILURES OF THE SYSTEM ARE RECOVERABLE AND THEIR CONSEQUENCES ARE CONTAINABLE.

Failures of the system should not have irreversible negative consequences for individuals' life, liberty, or happiness. So, for example, systems which need to use sensitive private information about individuals should use technical and other mechanisms to insure the continued protection of any retained sensitive private information even in the event of a system failure - because once private information becomes public, the damage cannot be undone.

Failures of the system should not propagate into other systems, and failure should be both graceful and gradual so that measures can be taken (within the capability of humans to respond in a timely way) to limit the damage a failure causes. So for example, systems should not be designed under the assumption that other related systems will never fail.

2.6 THAT THE SYSTEM CAN BE TURNED OFF IF IT DOESN'T WORK, OR STOPS WORKING.

Failures of the system should not create a "Hobson's choice " between turning off the system (and risking serious damage because of the loss of the protection the system provides) and continuing to operate the system (and suffering ongoing damage from the system's failure). Technical and other mechanisms should insure that an independent backup system or backup mode of operation is available in case of a serious failure which causes the system itself to cause serious harm to individuals on an ongoing basis.

2.7 THAT NO UNANTICIPATED ADVERSE CONSEQUENCES ARE LIKELY OR POSSIBLE.

Unanticipated consequences come in many forms. A technical privacy mechanism might prevent a physician from accessing a patient's record during a medical emergency, waiting for the (unconscious) patient to give consent. A law intended to prevent software piracy might be used by software vendors to suppress discussion of flaws in their offerings.

A specific class of unanticipated consequence should be considered in the design of security systems - the problem of moral hazard. Moral hazard is the property of a system of protection that its use encourages risky behavior which would be unlikely in the absence of the system of protection. Security systems should be designed to minimize the possibility of users' developing a false sense of security, and strong technical accountability mechanisms should be used to discourage risky behavior which might otherwise be thought of as "safe."

2.8 THAT ALL THE SYSTEM'S DEPENDENCIES ON OTHER SYSTEMS ARE EXPLICIT AND HAVE BEEN ANALYZED.

System failures often result from an incorrect and implicit assumption about the system's environment. When the environment changes in a way which isn't consistent with the assumption, the system fails. System dependencies on environmental conditions or other (external) systems need to be carefully analyzed and made explicit in the system's documentation. Technical mechanisms (developed for formal verification of system correctness) can help with both the analysis and the documentation.

3. HOW CAN WE ACHIEVE ASSURANCE?

Here are a few rules which will make systems easier to assure:

1. Decide what problem the system needs to solve. Then design the system so that it doesn't do anything else. Secondary uses (uses of the system for other than its primary purpose) are security failures waiting to happen.
2. Decide what you need to assure. Is it security? Safety? Availability? Think about this property while you're designing and building the system.
3. If something's impossible, don't do it. Don't deceive yourself about what's possible. For example, if your system is supposed to help prevent terrorism, don't delude yourself that a computer can determine the intent of a human.
4. Figure out how to identify at least one correct state of the system. (You might need to get back to it sometime, after something bad has happened.)
5. Don't use general-purpose computers. General-purpose computers, by definition, are capable of infinitely much undesirable, unsafe, or insecure behavior. Use special-purpose devices instead. If possible, use obsolete special-purpose devices. Obsolete stuff is cheap, and it probably works. Nobody is likely to lie about it for commercial gain. People probably understand it, and it's not likely to change anymore.
6. Make sure the system operates at a human pace. Make sure it fails slowly, visibly, and publicly. That way, somebody might notice and have time to fix the problem.
7. Drive attackers into the open. If people are going to do dangerous things, make sure they have to do them in public; that way somebody might notice. The dumbest security guard can tell the difference between you holding your face up to the camera and you holding a picture of somebody else's face up to the camera. While you're at it, make sure that people who want to attack the system have to put themselves in danger to do it. The self-destruct switch should be inside the bomb, and it should not have a countdown timer.
8. It'll work better if people are there.