CS378: Information Assurance and Security

Metrics for IA

Dr. Bill Young
Department of Computer Sciences
University of Texas at Austin

Last updated: March 23, 2015 at 08:42

Metrics for IA

Things have their due measure; there are ultimately fixed limits, beyond which, or short of which, something must be wrong. –Horace

It is a well-known fact that it is nearly impossible to manage, monitor, control or improve anything without measuring it. –Debra Herrmann, Complete Guide to Security and Privacy Metrics, p. 4

Is this True?

The following is quoted from *How to Measure Anything* by Douglas W. Hubbard:

1. Your problem is not as unique as you think.
2. You have more data than you think.
3. You need less data than you think.
4. There is a useful measurement that is much simpler than you think.

What do you think about this quote? To which domains does it apply?

Metrics for IA

The goal of security metrics is to provide the information an organization needs to prevent attack by establishing a quantitative basis for measuring security.

Name some useful security metrics.

What do you think are the fundamental difficulties in defining useful security metrics?

Metrics for IA

According to Herrmann, there are three major types of security/privacy metrics:

Compliance metrics: measure compliance with current security and privacy regulations and standards, such as HIPAA, Sarbanes-Oxley, GLBA, etc.

Resilience metrics: measure the resilience of controls relating to physical security, personnel security, IT security, and operational security both before and after a product, system or network is deployed.

Return on investment (ROI) metrics: measure the ROI in physical, personnel, IT, and operational security controls to guide capital investment.

Metrics

The collection and validation of security and privacy metrics should be embedded in the security engineering life cycle and tied to measurement goals and decision making. –Herrmann, p. 36

Patient, accurate data collection and recording is the key to producing accurate measurements. If not well organized, data collection can become very expensive. To contain these costs, the information gathering should be integrated with verification and validation activities. –IEEE Standards 982.1 and 982.2

If there are no prestated objectives on how the results will be used, there should be no measurements. –IEEE Standard 982.2

Discuss these three assertions. Are they reasonable? Do you agree or disagree? Why?

Terms

Metric: a proposed measure or unit of measure that is designed to facilitate decision making and improve the performance and accountability through collection, analysis, and reporting of relevant data.

Measurement: process by which numbers or symbols are assigned to entities in the real world in such a way as to describe them according to clearly defined rules. The comparison of a property of an object to a similar property of a standard reference.

Primitive: data relating to the development or use of software that is used in developing measures of quantitative descriptions of software. Primitives are directly measurable or countable. Examples include error, fault, failure, time interval, date, and number of an item.

Does this relate to the distinction between data and information? Be able to explain the relationships among these three notions.

Goals for Metrics

Metrics should provide a means for comparison:

- Between alternatives
- Change over time
- Relative to others: "benchmarking" an organization

Explain these goals. Give examples of each.

Goals for Metrics

Metrics and process-based measures together allow organizations to compare themselves to others. Difference between an organization and others is a "performance gap".

- *Best practices* are often cited as general recommendations.
  - *Gold standard* are the best possible practices.
  - *Best current practice* (BCP) are often recommended based on current technology or environments.
- *Standard of due care:* what any organization would do in similar circumstances.
- *Due diligence* is the process that an organization ensures standards provide adequate protections.

Why do organizations care about these? Which of these are necessary? Which are most desirable?

Regulations and Best Practice

Recall from our slideset on IA in Business:

Regulations often define "best practice" within a *particular industry.* Different industries have different standards. Why do you suppose that is?

> *"There is no single definition of the 'best practices' for an information security program. ... In fact, the term 'best practices' for information security is really a misnomer or even could be considered a myth."* (Landoll)

Due Diligence

Due diligence is a term used for a number of concepts involving either the performance of an investigation of a business or person, or the performance of an act with a certain standard of care.

First came into common use as a result of the U.S. Securities Act of 1933. So long as broker/dealers conducted a "Due Diligence" investigation into the company whose equity they were selling, and disclosed to the investor what they found, they would not be held liable for nondisclosure of information they failed to uncover in the process of that investigation.

What is the driver for this? I.e., what does it require and why do organizations strive for due diligence?

Malfeasance

Nonfeasance is failure to perform an official duty or legal requirement. *Malfeasance* is misconduct or wrongdoing esp. by a public official.

Information security due diligence is often undertaken during the information technology procurement process to ensure risks are known and managed.

Using Metrics

There are two important aspects of metrics:

- How you compute them
- How you use them

Think about the metrics used to evaluate the competence of K–12 teachers. Are they susceptible to misuse?

Classifying and Prioritizing Threats

An important fact to measure is *how secure your system is*. Various frameworks have been devised for this, some more successful than others.

The driving reality is that *not all vulnerabilities are equally serious.*

> *In a perfect world [...] we would tell you that if there's even the slightest chance of a single attacker being able to compromise a single user for even the smallest nuisance attack, that you should hold off the product release until every single possible vulnerability has been eliminated from the code. ... You'd never actually ship any code. –Sullivan and Liu, A Beginner's Guide to Web Application Security*

STRIDE

STRIDE is a threat classification system designed by Microsoft. It does not attempt to rank or prioritize vulnerabilities, only to classify them:

- Spoofing
- Tampering
- Repudiation
- Information Disclosure
- Denial of Service
- Elevation of Privilege

Many vulnerabilities may cross boundaries and some are "threshold attacks" that lead to others.

IIMF

The IIMF model classifies potential vulnerabilities according to four categories:

- Interception
- Interruption
- Modification
- Fabrication

Again, there may be overlap. Compare and contrast IIMF and STRIDE.

CIA

CIA classifies potential vulnerabilities according to violations on the traits we want a system to have:

- Confidentiality
- Integrity
- Availability

Some groups (notably NSA) add two more high level security goals to derive CIA-AN:

- Authenticity
- Nonrepudiation

Common Weakness Enumeration (CWE)

The Common Weakness Enumeration (maintained by the MITRE Corporation) is a general list of types of software vulnerabilities, such as:

- SQL injection (CWE-89)
- Buffer overflow (CWE-120)
- Missing encryption of sensitive data (CWE-311)
- Cross-site request forgery (CWE-352)
- Use of a broken or weak crypto algorithm (CWE-327)
- Integer overflow (CWE-190)

MITRE (in conjunction with the SANS Institute) also publishes an annual list of the top 25 most dangerous CWE issues. Find and investigate the CWE list.

DREAD

DREAD is another system from Microsoft. Unlike STRIDE, it not only classifies potential threats, but also ranks them. Compute 5 potential subscores for each vulnerability (scale 1..10) and for each of:

- Damage potential
- Reproducibility (or Reliability)
- Exploitability
- Affected users
- Discoverability

Add scores together and divide by 5 to get an overall DREAD score. Critique this approach. Having Discoverability as a component of the score is often criticized. Can you see why?

Critique of DREAD

- Not all of the DREAD criteria are necessarily equal in ranking any vulnerability, even though they are treated that way in the method.
- The ratings are highly subjective.
- In security, it might be best to take a pessimistic view and assume that if the vulnerability is present, it will be discovered and exploited.

DREAD has fallen out of favor, even within Microsoft.

Common Vulnerability Scoring System (CVSS)

A more common metric for rating vulnerabilities is the CVSS, created by a consortium of software vendors and security organizations, including: CMU CERT, Cisco, DHS, MITRE, eBay, IBM, Microsoft, others.

CVSS is currently maintained by the Forum of Incident Response and Security Teams (FIRST).

Rate each vulnerability in three dimensions on a scale 0..10:

- Base equations: objective characteristics of the vulnerability
- Temporal score: how may the risk change over time
- Environmental score: how is the vulnerability specific to your organization

CVSS

Unlike DREAD, the components are weighted according to a complicated formula.

Rather than compute by hand, use one of the CVSS online calculators. For example, one is hosted at the National Vulnerability Database (NVD) site:

`nvd.nist.gov/cvss.cfm?calculator`

Cost Metrics for Security and IA

Multiple costs combine into *Total Cost of Ownership* (TCO):

- direct costs: labor, equipment
- indirect costs: contractor overhead
- hidden costs: inefficiency and opportunity costs

What are opportunity costs?

Cost metrics are often used to compare similar companies:

- Number of staff, or number of staff-hours
- Amount spent per project, year or total on IA

Costs of Security and IA

Within a typical large company or government organization, security is a small part of the overall budget.

- Joint Security Commission of 1994 found security, in general, to be less then one percent of total operating costs. Do you think that's still true?
- However, this rule does not hold for SECRET/TOP SECRET projects: in unacknowledged programs 40% of the costs may be security.

According to Security500,2/14/11:

> *IT security staff at small and midsize businesses spend 127 hours every month managing their on-premises security infrastructure, according to a new survey by Webroot.*

Asset Value and Cost-Benefit Analysis

Benefit is the value an organization receives by using controls and countermeasures associated with a specific vulnerability.

- Typically by valuing the exposed asset and finding the percentage exposed.
- Asset valuation is often difficult or inconsistent across organizations.
  - Assets may have intrinsic value as well as acquired value (a storage array may have a higher value than the value of the underlying disks and controllers).
  - Intellectual property valuation may depend on the cost to acquire or produce the intellectual property, its worth (e.g. how much it can make over its lifetime), and how much it costs to protect it.

Some Metrics

Single Loss Expectancy: depends on percentage loss if a vulnerability is exploited; (SLE = asset value * exposure factor).

Exposure factor is the impact of the risk over the asset, or percentage of asset lost.

Annualized Rate of Occurence (ARO): how many times the attack will occur.

Annualized Loss Expectancy (ALE): (ALE = SLE * ARO)

Cost Benefit = (ALE_{before} - ALE_{after}) - Cost of safeguard

Identifying the Risk

Security planning is really about *risk management*—how much to spend on protection against what. This should be done within the broader framework of managing non-IT risks.

One common tool is the *annualized loss expectancy* (ALE), which is a table of possible losses, their likelihood, and potential cost. Example, for a bank with the following ALE, where should they put their security dollars?

Loss type	**Amount**	**Incidence**	**ALE**
SWIFT fraud	$50,000,000	.005	$250,000
ATM fraud (large)	$250,000	.2	$100,000
ATM fraud (small)	$20,000	.5	$10,000
Teller theft	$3,240	200	$648,000

Financial Metrics

38% of organizations in CSI/FBI survey use *Return On Investment* (ROI) to compare security investments. ROI is a percentage (benefit/cost) over time, e.g. 3 years:

$$\frac{[(b/(1+d)) + (b/(1+d)^2) + (b/(1+d)^3)]}{c}$$

where b is the benefit, d is the discount rate, and c is the initial cost.

18% use *Net Present Value (NPV)*. NPV is a currency value of expected return (benefits-cost) over time.

19% use *Internal Rate of Return (IRR)*. IRR is a discount rate chosen so that NPV would be 0. The investment with the highest IRR is preferred.

MTBF

Failure rate is the number of failures of a part in a population, divided by the sum total of time expended by all the parts in the population.

Mean-time-between-failure (MTBF) is the reciprocal of the failure rate. MTBF of a system depends on the interrelationship of the parts.

MTBF (continued)

In an architecture with the possibility of serial failures, failure of any one component part will fail whole system.

- Serial failures combine the failure rates of each of the parts.
- Redundancy allows for continued operation even if one part fails, so long as
  - there is a redundant path (e.g. two power supplies plugged in and both connected to the bus),
  - there are adequate standby parts which have not failed to take over, and
  - the failure detection circuit is working correctly, and is able to lock out the failed part. Typically redundancy is used to protect a few subassembles which have a high failure rate but it is possible to detect and recover, such as a bad power supply, bad network cable, or disk in a RAID array.

MTBF (continued)

MTBF only provides an estimate, useful during the design phases, to highlight areas where high failure rates might be predicted. Actual experiences may differ from MTBF as:

- Parts quality may not be the same as the part for which MTBF was originally estimated.
- Parts may be subjected to conditions outside of the bounds by which MTBF was estimated.

Recovery, MTTR and Availability

Mean time to repair (MTTR) after a failure (or attack). Goal is to reduce downtime costs

- In 2000, Amazon estimated losses due to downtime at $180,000 per hour.
- Ebay estimates $225,000 per hour (and had a 22 hour outage).

Availability is the percentange of uptime. Inherent availability is

$$(A = MTBF / (MTBF + MTTR))$$

Marketing literature quotes "5 nines" availability: 99.999%

- translates to 5 minutes downtime/year;
- however, requires "pre-defined, pre-tested environments"

If availability is particularly important to your organization, how might you enhance it?

Recovery, MTTR and Availability (continued)

Operational availability must take into account the time between maintenance (not just failure), and the downtime due to maintenance.

Availability metrics may no longer be valid if environment changes.

- In particular the introduction of new threats!
- Or when running in a "degraded capability."

Quality

ISO 9000/9001 quality typically implies that organizations have and follow documented, measurable procedures.

Data quality metrics are often used when reviewing the state of a database, such as when reviewing the database for use in a new project.

- Data decay: data in a database may lose currency or value over time as it becomes out of date. A significant issue affecting Customer Relationship Management (CRM) databases.
- Attacks may corrupt database. Yet few metrics have been defined for the quality of security-sensitive databases (password files, etc.)

ISO 9000

ISO 9000 is a family of standards for quality management systems, maintained by the International Organization for Standardization and is administered by accreditation and certification bodies.

- a set of procedures covering all key processes in the business;
- monitoring processes ensuring effectiveness;
- keeping adequate records;
- checking output for defects, with appropriate corrective action;
- regularly reviewing individual processes and the quality system itself for effectiveness; and
- facilitating continual improvement.

A "product," in ISO vocabulary, can mean a physical object, or services, or software, but "service sectors now account by far for the highest number of ISO 9001:2000 certificates - about 31% of the total."

ISO 9001: Key Features

- Quality manual is required
- Decisions made on recorded data and regularly audited
- Records show how and where materials acquired and processed
- Documented procedure to control quality documents
- Provide suitable infrastructure, resources, information, equipment, conditions
- Clear quality objectives for each product
- Determine key monitoring points in the process
- Set clear requirements for purchased product
- Establish procedures for customer information
- Plan stages of development for new products
- Review performance through internal audits
- Document procedures for nonconformance

Quality (continued)

Quality of service (QoS) metrics are often incorporated by contracts with service providers.

- Service providers face financial and indirect penalties if QoS cannot be met.
- May not be able to meet levels while under attack, after a natural threat has occurred, and during recovery.

Effectiveness of Protection

Monitoring is essential to determine if countermeasures are having the desired effect. However, determining the appropriate metrics may be difficult as the underlying landscape is changing.

- New threats not anticipated by current countermeasures may cause a change of metric to be required, which hinders long-timescale or cross-organization comparison.

Often, organizations may wish to compare the number of successful attacks to the number of total attempts.

- However, this metric needs to take into account the impact of the attack. For example, many organizations do not consider attacks against their external web site to be significant unless they succeed.
- Furthermore, finding the total number of attack attempts may be difficult. Is every packet that a firewall detects coming from a hacker scanning the Internet for vulnerabilities an 'attack'?

Effectiveness of Protection (continued)

Metrics are not yet consistent for other events, such as human error, and in many cases the underlying protection mechanisms do not yet give adequate reports.