Considering Metrics for Information Assurance

Michael Skroch
Information Systems Office
mskroch@darpa.mil
(703) 696-2375
Discussion outline

- Presentation
  - IASET interest in metrics
  - Who and what uses them?
  - Types of metrics
  - Methodology for metrics
Before we begin...

● Information presented here is intended to motivate thought in the area of metrics for IA
  ◆ we have a long way to go to make this a useful discipline and even further before it is mature

● Concepts are my opinion, based upon experience in this area, my thoughts on the subject, and ideas gathered from others in this area
  ◆ so...please provide input and constructive criticism
What: Our military and civilian information systems are at risk
  ◆ increased use, reliance, complexity, visibility

One reason why: IA is vital for reliable, secure functioning of information systems - yet IA design, assessment, and operational understanding are:
  ◆ currently unreliable, not understood, not scientific, not systematic, often non-existent

Why is IA in this condition?
  ◆ even though some people know there is a problem, we have yet to experience the wreckage of a true information system disaster

...but we cannot wait for one!
We don’t understand the science of IA in systems

- An understanding of the basic laws governing IA does not exist
- There are none to few useful ways to measure IA or its components to compare, define requirements, measure changes
- We don’t know how to compute, make decisions or otherwise utilize IA measures

We don’t know how to design and assess IA in systems

- A system-level, methodical process is rarely taken, which leads to numerous uncovered vulnerabilities
- Sufficient types and quantities of tools do not exist to allow for effective design, assessment, operation
- Designers, assessors, and operators cannot access common information about a system: no common language, tools do not work together, no common environment to express, define, communicate the attributes of a system
- Knowledge is rarely passed forward therefore we’ve often been doomed to repeat history
"...so many centuries after the Creation it is unlikely that anyone could find hitherto unknown lands of any value." - committee advising Ferdinand and Isabella regarding Columbus' proposal, 1486
- Provide a science-based environment for design and assessment that will:
  - yield improved system IA
  - allow for faster design and assessment at less cost
  - assist the designer and assessor at developing the system
  - allow the user to understand the system IA, and risks

- This environment will consist of:
  - methodologies, metrics, common languages
  - IA models with objects that carry along all information about their being
  - suite of automated tools that can operate together seamlessly within the environment
IASET addresses systems level problems, not discrete technology problems.

IASET is primarily focused on design-time; however, fundamentals will apply to operation-time work including that in the IA, AIA, and CC2 programs.

SCIENCE

1. Cyberscience
2. IA Metrics
3. Mathematics and models
4. Science-based methods for IA design and assessment

ENGINEERING TOOLS

5. Integrated environment for IA design and assessment
6. IA design and assessment tools
7. Malicious code mitigation

a view of the seven areas
Research will be started independently in each area, but results will be brought together throughout the program.

All results will be provided to other IA&S programs and the IA community.

All areas should produce transitionable technologies for DoD & industry.
Metric (’me-trik, noun)

What do I mean by metric?
- simply, a standard of measurement
  - easy definition, hard to produce
- we wish to focus on metrics which have relevance to information assurance
We don’t understand the science of IA in systems

IA metrics

- Utility of metrics
  - identify those that are important for IA
  - must be useful to the end user or some intermediate process in understanding IA
  - understand how they relate to each other, are used in calculations, can be used to make decisions
Consider the needs for metrics before deciding which to use

- Consider why metrics work
  - underlying cyberscience about how IA metrics work
  - mathematics and models which are used to manipulate IA metrics

- Consider the lifecycle of the system that will generate and utilize your metrics
Who needs metrics?

- **R&D community** needs concrete goals
  - to compare competing approaches
  - to mark progress as a function of time
- **Vendors** want products certified and way to specify performance
- **Planners** need a way to specify requirements for design or procurement
- **Designers** need them to create better systems, and systems that meet requirements (in a systems and in-between systems)
- **Assessors** need ways to measure red team evaluations, compare to requirements, measure improvement
- **Testers / Accreditors** need specifications, benchmarks and reliable data from assessors
- **Commanders / Operators / Users** need to know how well they are protected
  - in the unique environment in which they are using a system
- **Regulators**
- **Intel Community**
What needs metrics?

- **Design / production processes, tools**
  - design and assessment (IASET common environment)

- **Operational processes**
  - lifecycle: deployment, setup, operation, hardness surveillance and maintenance (HMHS), improvement

- **Decision processes, systems, tools**
  - Autonomic Information Assurance (AIA) program - reflexive defense against attacks
  - Cyber Command and Control (CC2) program - information for human-based decision

- **Analysis processes, tools**
  - strategic planning, future research, forensic, Intel
Select metrics for their utility

- Useful to meet goals of your system
  - design, assess, operation, improvement, ...
- Ease of measurement
- Ease of use (calculation, understanding, extensibility)
- Cost to obtain
- Quality:
  - precision (significant digits), uncertainty (in source)
  - consistency (between people), repeatability (over time)
- Are they relevant? (measure something we care about)
- Are they comprehensive? (measure all we care about)
We don’t understand the science of IA in systems

**IA metrics**

- **Qualitative**
  - not all measures can be reduced to numbers
  - need common frame of reference and language
  - need methods for correlation and extraction of information from qualitative metrics
  - benchmarks systems are needed

- **Quantitative**
  - measures should be science-based
  - need mathematical relationships to other metrics and the physical world
We will often need to measure more detail than the user will require to make a decision.

**Direct:** speedometer
- metric: \( v = 80 \text{ mph} \)

**Indirect:** calculation
- metric: distance \(d\)
- metric: time \(t\)
- calculate:
  - metric: \( v = \frac{d}{t} = 80 \text{ mph} \)

Hidden metrics:
- rotation rate
- wheel size
Metrics: what are we measuring?

Consider that with a single measurement you may be obtaining more than a single quantity or value...

- Value of metric (scalar, vector, concept)
- Units of metric (meaning)
- Related measures for metric (time, place, way it was measured)
- Uncertainty related to measurement
Assets Required

● To develop new metrics:
  ◆ Good imagination, patience
  ◆ Good understanding of information systems, IA requirements
  ◆ Awareness of previous efforts in the field of metrics, and of the things people want to measure
    ● Survey of metrics from other industries also relevant

● To estimate / measure metrics for a system:
  ◆ Be able to do the analyses that produce a particular metric
    ● May require a number of different skill sets
  ◆ System design and functional requirements must be captured in an accessible format
  ◆ Sometimes: need facility to actually simulate system or run a test.
    ● Required to compute some metrics (e.g., Red Team)
    ● Required to validate other metrics
Methodology for metrics

1. Understand metric basics
   - How used, how relates to other things
2. Understand application, user, system
   - Expertise, players, users
3. What are objectives and goals that must be met?
4. What are observables of above?
5. How to quantify observables?
   - Quantitative/quantitative
   - Units, bounds, relationships
   - Nature - binary, analog, range, choices...
6. How to measure metrics?
   - Methods, certainty, error bounds
   - Repeatability
Methodology for metrics

Here’s an example framework for how you could generate metrics, and go further by considering how they will be used.

**What do you want to measure?**

- System Impacts
  - Examples:
    - Human death & injury
    - Mission failure
    - Loss of confidence
    - Removed from theater

**How are these manifested?**

- Consequences
  - Examples:
    - Loss of capability
    - Modification of information
    - Unauthorized access
    - Hamper operations

**How do you measure them?**

- Metrics
  - Examples:
    - Types of service affected
    - Volume of service affected
    - Amount of time service affected

**How can you use them?**

- Models, Tools, Data
  - Analysis - search for:
    - Causes of consequence
    - Event frequency
    - Severity, extent of consequence

Examples:
- Human death & injury
- Mission failure
- Loss of confidence
- Removed from theater
- Loss of capability
- Modification of information
- Unauthorized access
- Hamper operations
- Types of service affected
- Volume of service affected
- Amount of time service affected
- Causes of consequence
- Event frequency
- Severity, extent of consequence
Other frameworks that may be of use

- Physical access analogy
  - protect, detect, delay, react, impact
- System objectives
  - access control, integrity, availability, utility, safety, non-repudiation
- System state
  - architecture, transactions, state changes, information flow, interfaces
- Views of the “universe”
  - spatial, logical, temporal, lifecycle, system
Metrics: example from AFIT*


<table>
<thead>
<tr>
<th>Constructed Definitions for Each Element of the Information Realm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Information</strong></td>
</tr>
<tr>
<td><strong>Information System</strong></td>
</tr>
<tr>
<td>with respect to decision making</td>
</tr>
<tr>
<td><strong>Information-based Process</strong></td>
</tr>
</tbody>
</table>
Metrics: example from AFIT

The Information Realm
(That Part That Affects Politico-Military Decision Making)

Information
(Data & Semantic Meaning)

Minimum Update Length
(Functional Dependency)

Age of Information
(Devaluation of Information by Time)

Potential Effect
(Effect on Type I & Type II Error)

Reject Truth
(P(Type I Error))

Accept False
(P(Type II Error))

Accept Truth
Reject False
(P(Correct Decision))

Information Systems
(Convey, Store, Process Without Adding Value)

Bandwidth
(How Much)

Throughput
(How Fast)

Fidelity
(How Accurate)

Error Content
(Bit Error Rate, SNR)

Complete Update
(Incomplete Update)

Information-Based Processes
(Processes That Add Value to Decision Making)

Efficiency
(Resources & Time)

Resources Consumed
(In Processing)
(Relative Analytical Cost)

Time Consumed
(Processing Time Consumed)

Effectiveness
( Value Added By Processing)

Timeliness
(Is There Time to Respond)

Accuracy
(Is the Analysis Appropriate)

Focus
(The Direction, Selection, and ID)

Resilience
( Robustness, Perturbation Response)

Hardness
(How Hard to Penetrate System)

Effort to Penetrate
(Physical Hardness)

Recovery Rate
(How Long to Recover)

Secure Capability of System
(Non-Physical Security Measures)

Secure Capability
(Classification)

Intrusion Detection
(Can Entry Be Identified)
<table>
<thead>
<tr>
<th>DEFINED OBJECTIVE</th>
<th>MEASURE UNIT</th>
<th>MEASURE TYPE</th>
<th>LOWER BOUND</th>
<th>UPPER BOUND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase Minimum Update</td>
<td>Level of support required to maintain awareness level</td>
<td>Category</td>
<td>User-level</td>
<td>Maintenance-level</td>
</tr>
<tr>
<td>Likely Accept False</td>
<td>Level of expected effect on adversary decision making</td>
<td>Probability</td>
<td>No Change</td>
<td>High probability</td>
</tr>
<tr>
<td>Increase Error Content</td>
<td>Percentage Error Content on System</td>
<td>Percentage</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Penetrate System</td>
<td>Level of defeat effected</td>
<td>Category</td>
<td>No Capability</td>
<td>Completely Defeat</td>
</tr>
<tr>
<td>Increase Recovery Time</td>
<td>Change Cycles Over Which System is Unable to Perform Mission</td>
<td>Quantity</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Defeat Security</td>
<td>Likelihood of Gaining Access to System</td>
<td>Category</td>
<td>No Change</td>
<td>High Probability</td>
</tr>
<tr>
<td>Defeat Detection</td>
<td>Our Expected Ability to Defeat the Adversary's Intrusion Detection</td>
<td>Category</td>
<td>Certain Detection</td>
<td>Low Likelihood of Detection</td>
</tr>
<tr>
<td>Consume Essential Resources</td>
<td>Percentage of Essential Resources Consumed</td>
<td>Percentage</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Reduce Timeliness</td>
<td>Number of Change Cycles that the Processed Product is Late</td>
<td>Quantity</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Reduce Resilience</td>
<td>Expected Ability to Reduce Resilience</td>
<td>Category</td>
<td>No Change</td>
<td>Catastrophic Failure</td>
</tr>
<tr>
<td>Minimize Collateral Damage</td>
<td>Expected level of Collateral Damage</td>
<td>Percentage</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Success Criteria
(When are we “done” creating metrics?)

- Evident utility (each metric “means something”)
- Wide adoption
- A good set of metrics should be complete enough to cover all important IA aspects of a system
Questions?
Back-ups
IA Science & Engineering tools

description - overview

- **Program components**
  - **Cyberscience**
    - We don’t understand the science of IA in systems
  - **IA Engineering**
    - We don’t know how to design and assess IA in systems
  - **Malicious Code Mitigation**
    - We are increasingly vulnerable to effects of malicious code

- **Hypotheses:**
  - Science and scientific methods applied in an environment for design and assessment will yield stronger system Information Assurance, faster design for less cost
  - Successful research approaches for Malicious Code Mitigation are now feasible due to foundational DARPA IS/IA research; we must continue IA science-based research to crack the problem
IA Science & Engineering tools  
**rational for program**

- **Primary concerns**
  - Information system users, designers, and assessors have no:
    - meaningful measure of system vulnerability, risk, assurance level
    - formal, repeatable methods for design, assessment, and specification of IA
    - ability to understand, identify, and mitigate effects of malicious code

- **Weak areas**
  - no understanding of the cyberscience which underlies the systems that are currently being created
  - few and inconsistent measures of IA
  - same mistakes made decades ago continue to show up in current systems
  - malicious code problem is increasing due to threat and rapid adoption of insecure technologies

---

**National security impact**

- Information systems present an asymmetric risk
- Complexity and connectivity (Internet, mobile code, etc.) is increasing
- Military and civilian information systems rely on COTS
- COTS currently have little incentive to include IA
- Design, assessment and user understanding of IA is not keeping pace with information technologies
Metrics: qualitative

- Qualitative metrics need frame of reference / definitions
  - promotes consistency between generators of metrics
  - promotes repeatability of metrics over time
  - provides understanding to others about what metrics mean

**Percent attack complete (red team attack scenario)**

<table>
<thead>
<tr>
<th>value</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>conceptual</td>
</tr>
<tr>
<td>20</td>
<td>preliminary</td>
</tr>
<tr>
<td>30</td>
<td>initial</td>
</tr>
<tr>
<td>40</td>
<td>developed</td>
</tr>
<tr>
<td>50</td>
<td>well developed</td>
</tr>
<tr>
<td>60</td>
<td>detailed</td>
</tr>
<tr>
<td>70</td>
<td>finely detailed</td>
</tr>
<tr>
<td>80</td>
<td>validated</td>
</tr>
<tr>
<td>90</td>
<td>simulated</td>
</tr>
<tr>
<td>100</td>
<td>tested</td>
</tr>
</tbody>
</table>

Example:

attack, no analysis performed
attack, minimal analysis performed
attack, moderate analysis, possibly some data not reviewed
attack, moderate analysis, most to all available data reviewed
attack, moderate analysis, consider high-level contingencies
attack, relatively complete analysis, many details considered
attack, relatively complete analysis, fine details considered
attack, well developed, peer reviewed, gaming or role playing used
attack, well developed, simulated through computer code or mockup
attack, physically tried the attack through field test or real thing
Level of Abstraction -- Bounds

- Includes:
  - Confidentiality, integrity, availability, authentication, non-repudiation, etc.
  - Aggregation of “small” measures
    - Component metrics ==> system metrics
    - Dissimilar metrics ==> overall “utility” score
  - Support of trending (consistency over time)
- Different metrics for different parts of a product lifecycle (design, operations, analysis, etc.)
- Excludes:
  - Considerations that are not IA-related (e.g., environmental conditions, cost, “functional” metrics) as they are the responsibility of others.
Methodological Hierarchy

- **What’s “Above” Metrics -- Who Uses Metrics**
  - Risk analysis, Decision Theory
  - Design process; composition
  - Specifications/Requirements (to the extent that they do not specifically call out metrics)

- **What’s “Below” Metrics -- Who “Measures” or “Estimates” Metrics**
  - Red teaming (“measures” the metrics)
  - Test & Evaluation; Simulation; Vulnerability analysis
Preconditions & Assumptions to Identification & Use of Metrics

- For developing new metrics: ???
- For the use of metrics: *Given* a “toolbox” of relevant metrics, I need to know:
  - What do I care about in the system? (so I know which metrics to select)
  - Where are the “bounds” on the system? (so I know where to compute those metrics)
Metrics to be Considered / Used (What Makes a Good Metric?)

- Quality of Metrics includes:
  - Metrics are computable within a time frame that is useful to decision makers
  - Makes intuitive sense (don’t fail the giggle test)
  - Repeatability / Consistency
  - Really measures what you think it measures
Metrics

Aggregation

- Composite metrics aggregate simple (basic) metrics.
- Values for weights in aggregation vary by customer -- elicitation of these is a social science task that needs research
- Mathematics of aggregation may need some research