# Security Metrics

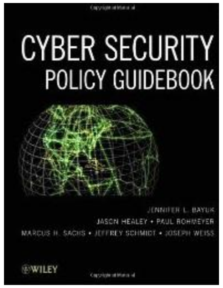
**University IT Audit Conference**

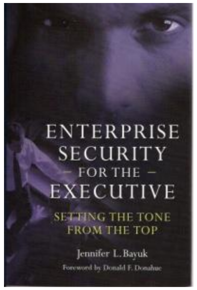
**June 10, 2013**

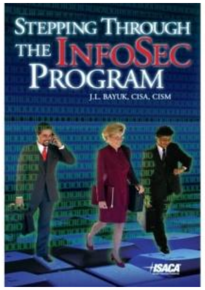
**Jennifer Bayuk**

**www.bayuk.com**

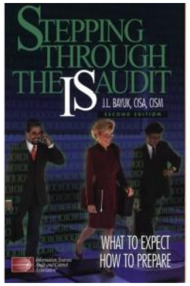
# Jennifer Bayuk

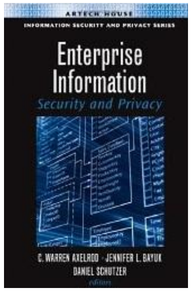







- Experienced in a wide variety of private security positions including Chief Information Security Officer.
- Author of multiple textbooks on security management topics
- Chair and contributor to multiple public and private InfoSec Boards and Committees
- CISA, CISM, CGEIT, CISSP, NJ Licensed PI, Systems Engineering Systems Engineering PhD, Thesis in Security Metrics

# Today’s Security Metrics

Today’s security metrics are typically based on two assumptions:

(i) there is a secure way to configure any system, and

(ii) the task of security management is to maintain that configuration.

However, today’s cyber attacks:

- typically do not exploit holes in configuration;
- exploit application or system functionality.

The most skilled engineers with

- have the most sincere of intentions
- use state-of-the-art techniques:
- create designs intended to accomplish security objectives, yet it is nevertheless
- build vulnerable systems.

That is, attacks are successful even though security is configured as designed.

Therefore, we are measuring the wrong thing.

*..or at least not the whole thing…*

- **Verification**
  - Target

*Did we build the system right?*

- **Validation**
  - Vulnerability
  - Usability

*Did we build the right system?*

***Correctness != Effectiveness***

Measurement is the process of ***mapping*** from the empirical world to the formal, relational world. The measure that results characterizes an ***attribute*** of some object under scrutiny. Information Security is not the object, nor a well-understood attribute.

This means you are not directly measuring security, you are measuring other things and using them to draw conclusions about security.

# Potential Evaluation Criteria for Security Metrics

| | |
|---|---|
| Valid: | data can be used to disprove a hypothesis that system is secure |
| Accurate: | data reflects the content of measurement as it was envisioned |
| Numeric: | data can be precisely quantified |
| Correct: | data is collected according to specifications |
| Consistent: | measure is independent of measurer |
| Time-based: | there is a fixed reference point of data collection |
| Replicable: | measurement repeated in same manner in same environment will yield same result |
| Unit-based: | data may be expressed in terms of a unit |
| Informative: | data provides information without additional context |

*These are actually generic criteria for metrics in any domain.*

# Example Rules for Evaluation

1. Any metric that is not accurate or not valid is weak
2. Any metric that is accurate and valid is at least neutral
3. Any metric that is valid, accurate, time-based and informative is strong

# Taxonomy of Security Metrics

| | |
|---|---|
| Target: | Metrics that have a measurable 100% target. |
| Monitor: | Metrics that monitor security processes. |
| Remediation: | Metrics that show progress toward a security objective. |
| Performance: | Metrics that demonstrate capability to accomplish system functionality. |
| Vultest: | Metrics that show susceptibility to known threats. |
| Resilience: | Metrics that demonstrate system ability to recover from harmful impact. |
| Adversary Skills: | Metrics that estimate adversary skills levels. |
| Adversary Goals: | Metrics gleaned from intelligence on adversary motivation and justification. |
| Stochastic Models: | Metrics that combine measures with probability estimates. |
| Deterministic Models: | Metrics that combine measures with inference rules to form conclusions about security. |
| Internal activity: | Metrics that chart work activity (“busyness”). |
| External activity: | Metrics that track threats (“weather”). |

# Target Example A

Measure X:

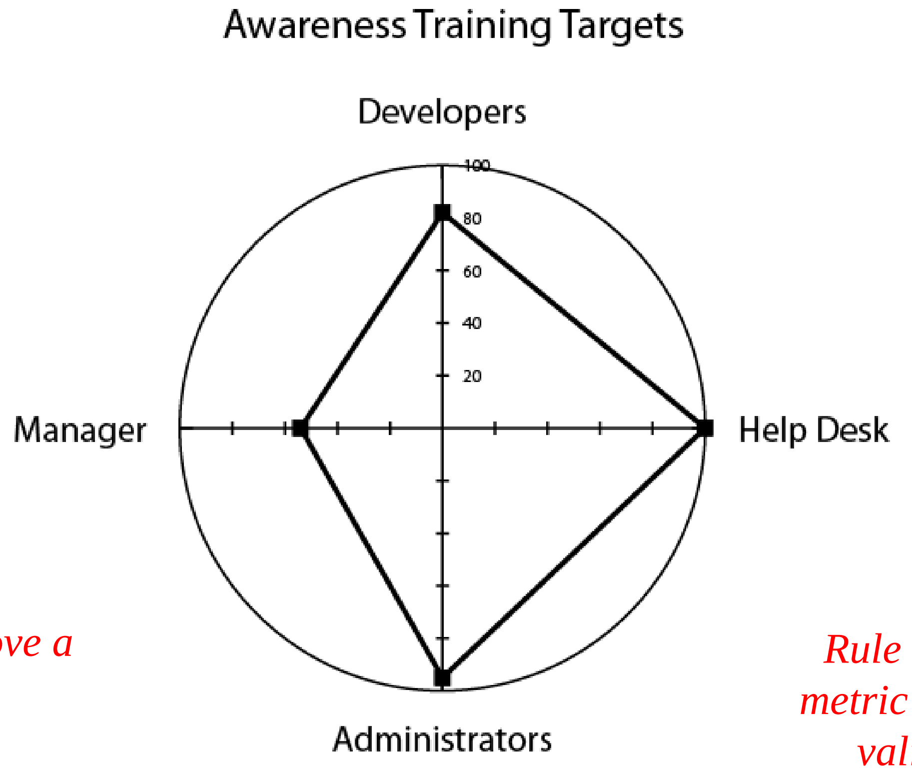
The current number of personnel in each department (the target).

Measure Y:

The number of personnel in each department who have been through security training.

Department Security Awareness Metric: Y/X

*Targets are percentages based on inventory. They derive integrity from the accuracy of both the inventory and the measurement process.*




*This measure cannot be used to disprove a hypothesis that a system is secure.*

*Rule 1 says any metric that is not valid is weak.*

| Valid | Accurate | Numeric | Correct | Consistent | Time-based | Replicable | Unit-based | Informative | Overall |
|---|---|---|---|---|---|---|---|---|---|
| No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Weak |

# Target Example B

Measure X:

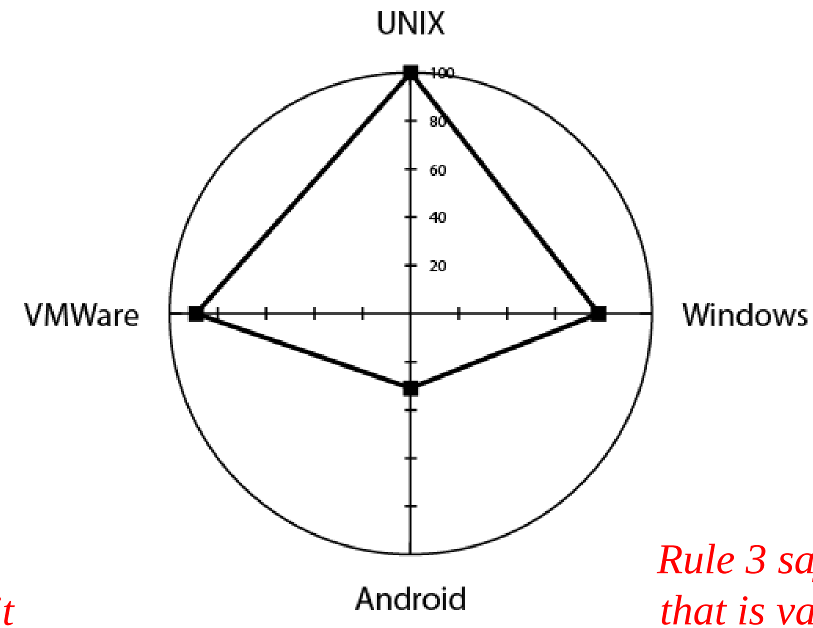
The number of computers in operation running a given operating system (OS).

Measure Y:

The number of computers in operation running a given OS that are configured as per security standards daily configuration checks.

OS Security Metric: Y/X

Operating System Security Parameter Targets




*If my hypothesis is that a system is secure if it is securely configured, then this metric can disprove it.*

*Rule 3 says any metric that is valid, accurate, informative, and time-based is strong.*

| Valid | Accurate | Numeric | Correct | Consistent | Time-based | Replicable | Unit-based | Informative | Overall |
|---|---|---|---|---|---|---|---|---|---|
| Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Strong |

# Target Example C

Daily Measure W:

The number of firewall devices in operation.

Daily Measure X:

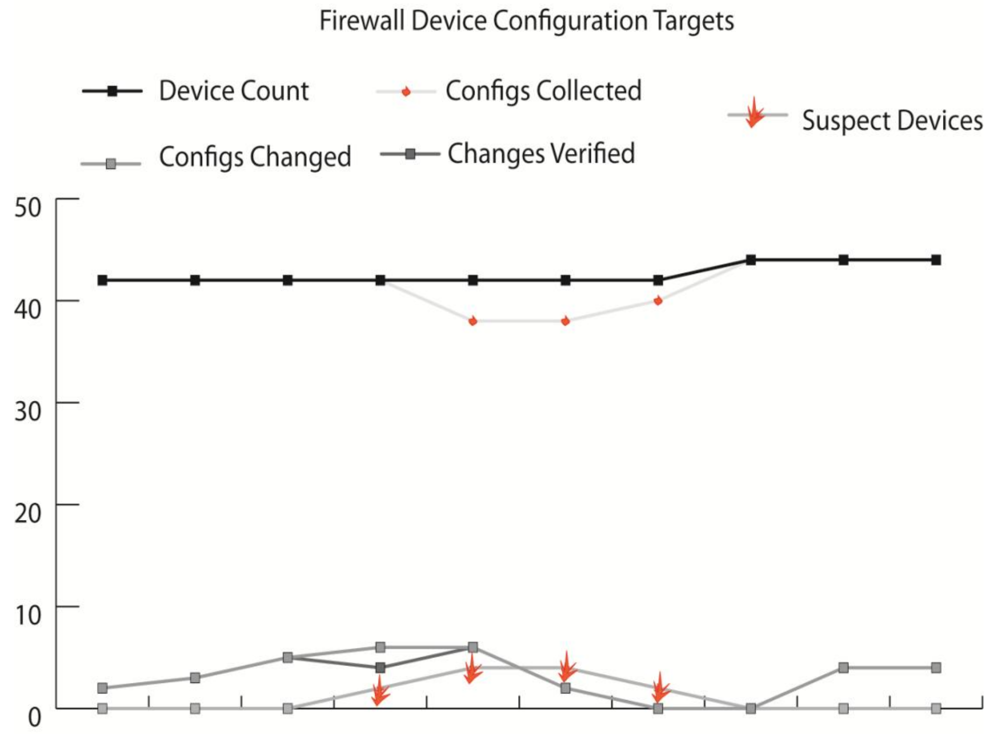
The number of firewall devices whose configuration was retrieved in past 24 hours by network management system.

Daily Measure Y:

The number of firewall devices configurations that deviate from yesterday's configuration.

Daily Measure Z:

The number of deviant device configurations where deviations directly compare to authorized planned changes.




Daily Firewall Device Metric, Suspect Devices as % of Total: ((W-X) + (Y-Z)) / W

# Monitor Example A

Measure S:

The set of work orders opened by each internal help desk person P in category “security” and subcategory “password reset” with resolution “reset” in 24 hour period.

For each W in set S,

Measure T: Elapsed time of W, between work order open to close.

Measure U: Audit log in identity management system of successful queries within elapsed time T for user U, as identified in W.

Measure R: Recordings of P asking user U for security identification code within time T, and U’s correct response.

Measure L: All P’s password resets in same 24-hour period as S.

Daily Help Desk Person Monitor Metric:

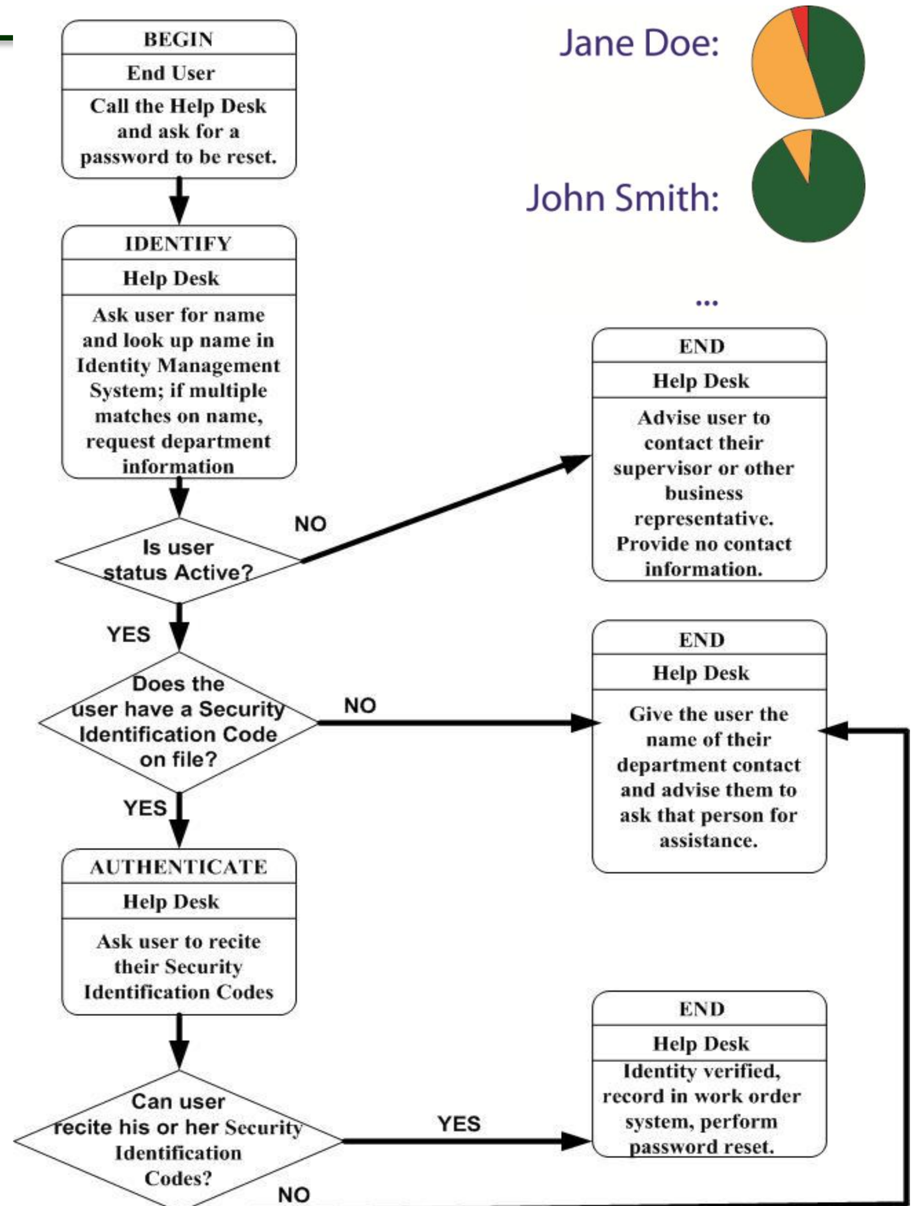
If (Count of L > Count of S), Then P = Bad

Else For each W in set S,

If (U and R exist) Then P=Good

Else If (R exists) then P=Shortcuts

Else P= Bad



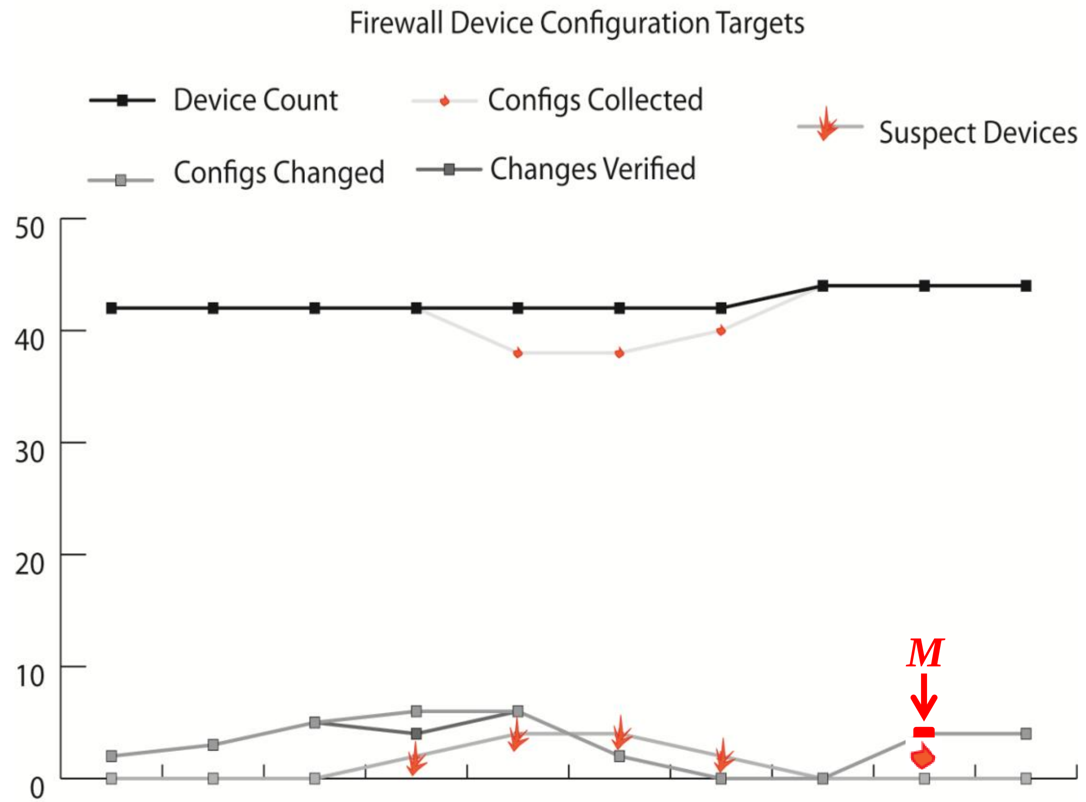
# Target Example C Monitor Overlay

Daily Measure W:
The number of firewall devices in operation.

Daily Measure X:
The number of firewall devices whose configuration was retrieved in past 24 hours by network management system.

Daily Measure Y:
The number of firewall devices configurations that deviate from yesterday’s configuration.

Daily Measure Z:
The number of deviant device configurations where deviations directly compare to authorized planned changes.



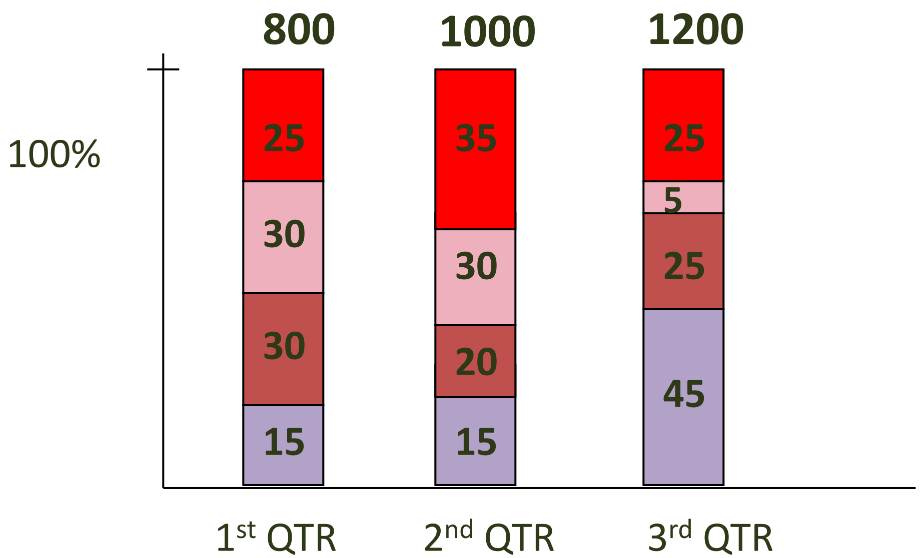

***Measure M: The number of false negative comparisons by network operations staff.***

Daily Firewall Suspect Device Metric: ((W-X) + (Y-Z)) / W

Adjusted Metric for % Expected Error rate gleaned from monitoring : ((W-X) + ((Y-Z) * 1.%M)) / W

# Remediation Example

Identity Management Deployment Progress




- estimated percent of users not yet identified
- % of users that are not mapped to an existing valid identity
- % users manually identified, but not yet configured to automatically correlate
- % users that automatically correlate to an identity management system index

# Performance Examples

Six Sigma: Target of less than 3.4 defects per million activities. Example Security Defects:

- *Fraudulent transactions processed.*
- *Unauthorized use of customer data.*
- *Outages due to hacking activities.*

Other Performance Standards:

ITIL: Service level management targets

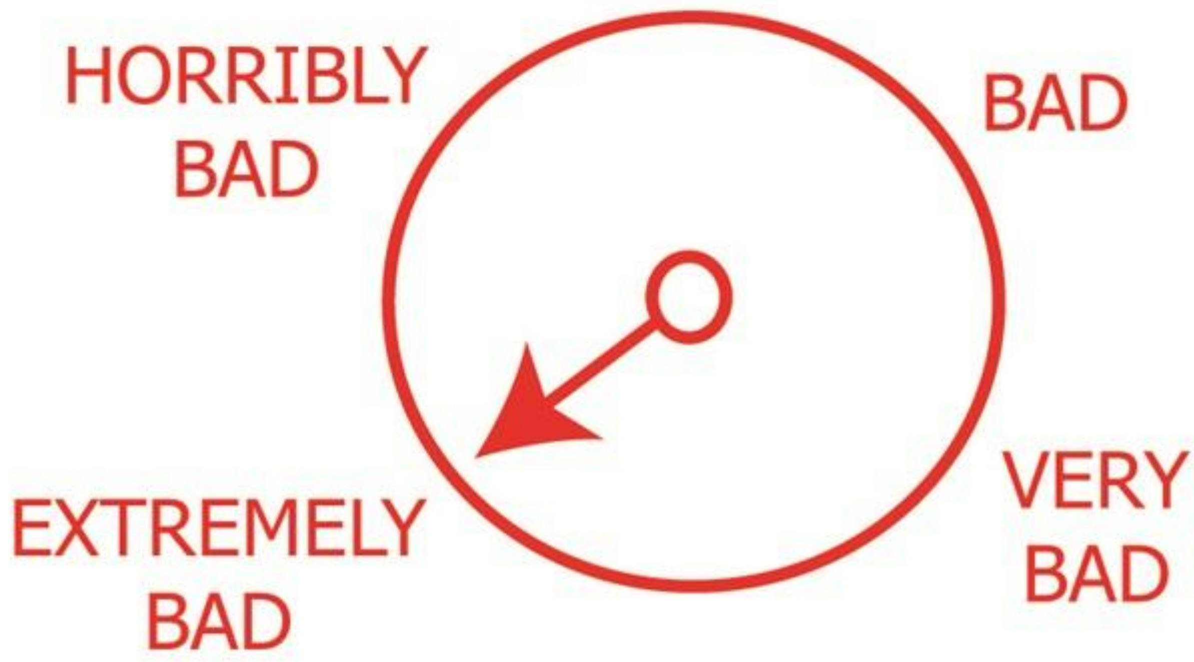
QFD: Customer satisfaction measures

*To make sense, these must be business-driven, not security-technology-driven.*
*A misconfigured device is not a performance metric unless your business is device configuration.*

# Vulntest Example

- Read team read-outs.
- Vulnerability lists

Typically
not reliable
or repeatable




*"Badness-ometers" – Gary McGraw in Software Security*

# Skills and Goals Examples

Skills and Goals metrics do not measure an implemented system, but some aspect of the system's expected interaction with an environment that includes hostile adversaries.

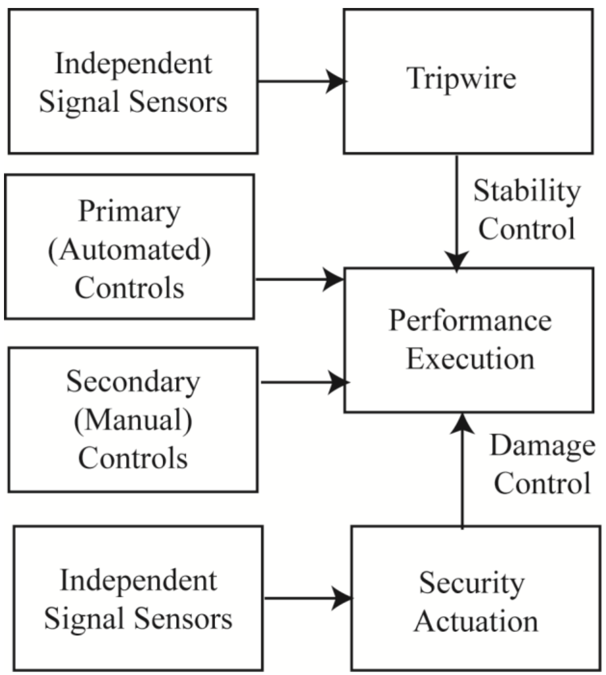
Low Risk ⟷ High RIsk; ↑ High Expertise

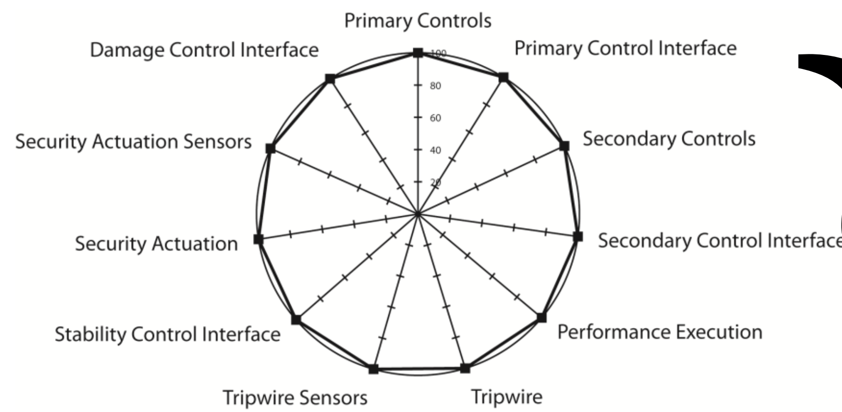
| | | | | |
|---|---|---|---|---|
| Disgruntled Insiders | voluntary workforce participation in company events | workforce employer lawsuits are above industry average | unexplained technology problems within firm are a frequent event | insider cyber attacks are a frequent event |
| Organized Criminals | software that controls financial assets is only internally accessible | publicly accessible software allows customers to control assets | publicly accessible software allows customers to transfer control of financial assets | publicly accessible software allows firm insiders to transfer control of firm and/or customer financial |
| Terrorists | domestic-only cyberspace presence | international cyberspace presence | repeated attempts by foreign nationals to cause cyber-damage to firm | declarations by terrorist(s) of intent to cause cyber-damage to firm. |
| Hactivists | consistently positive press coverage | negative press coverage related to special interests | active lobbying to government(s) against firm activities by special interests known to resort to cyber attacks. | declarations by hackivists of intent to cause cyber-damage to firm. |

***Note – such subjective measures are typically ordinal rather than numeric, but nevertheless, inform decisions***

# Resilience Examples

Resilience – End-to-end processing of failover, redundancy plus diversity.




**+**

**Failover Test Results**

HYPOTHESIS

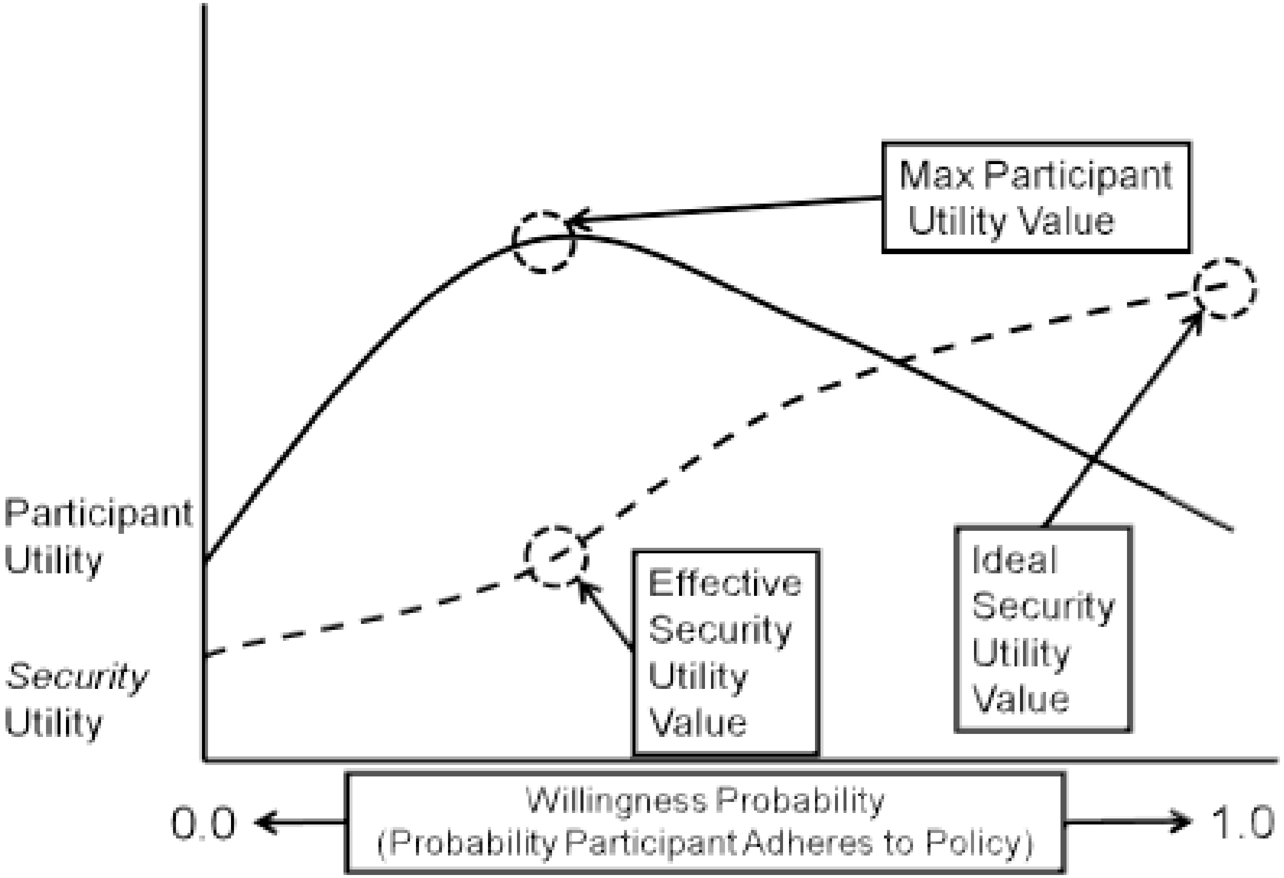
**+**

**Performance Metrics**

TEST

# Stochastic Model Example

Measures are associated with alternative probabilities of occurrence, and compared to an ideal outcome in order to determine “best” course of action.

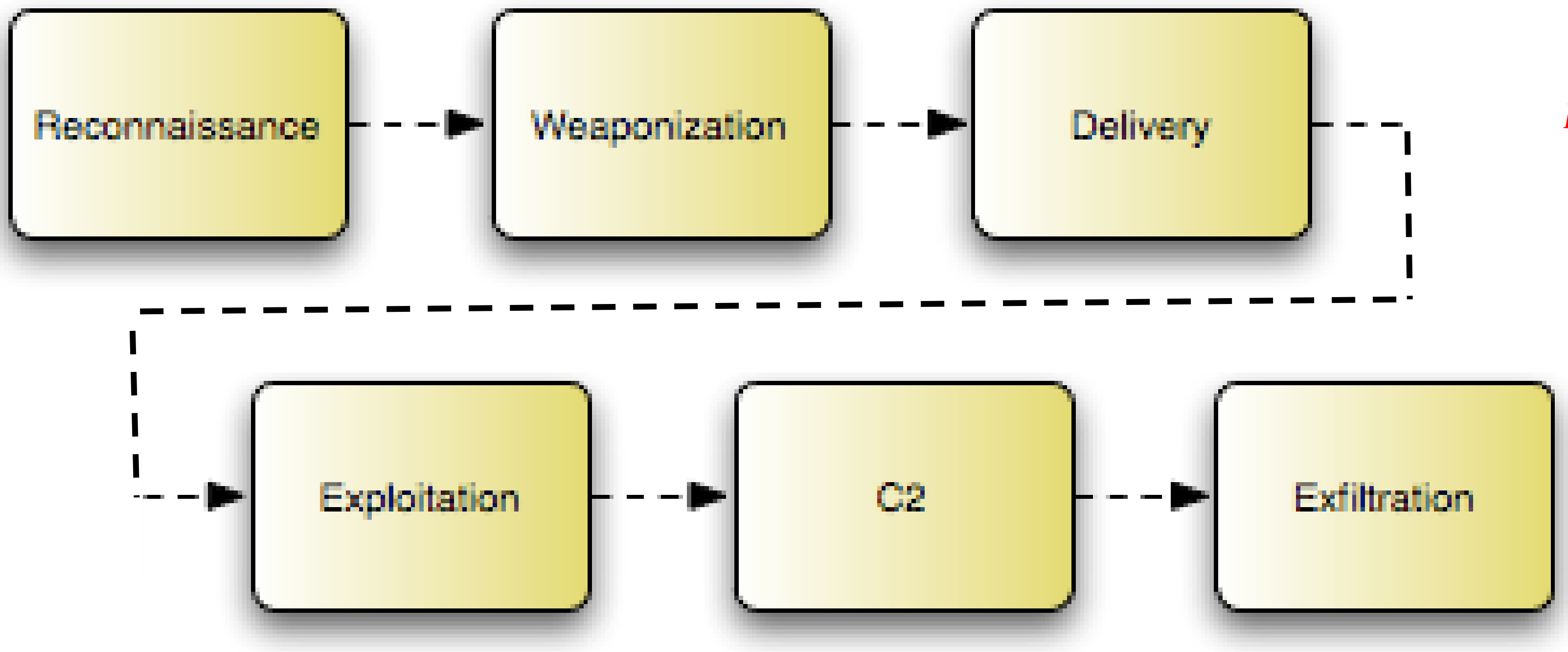



*It is improbable that all participants find utility in following policy, so ideal will never be achieved.*

*This example would support a hypothesis that security models should demonstrate high probability of conformance to procedures that are required to achieve security objectives.*

*Source: D. Eskins and W. H. Sanders, "The Multiple-Asymmetric-Utility System Model: A Framework for Modeling Cyber-Human Systems," presented at the Proceedings of the 8th International Conference on Quantitative Evaluation of SysTems, 2011.*

# Deterministic Model Example




*Deterministic model of expected attack determines placement of monitoring devices and associated correlation utilities that support enterprise security architecture detection controls.*

Measures are identified for each step using forensic techniques designed to identify attacks in progress.

*Source: M. Cloppert, "Evolution of APT State of the ART and Intelligence-Driven Response," presented at the US Digital Forensic and Incident Response Summit http://computer-forensics.sans.org, 2010.*

# Internal Activity Example

Measure W:

The number of calls to internal help desk in category "security" and subcategory "request for admin rights."

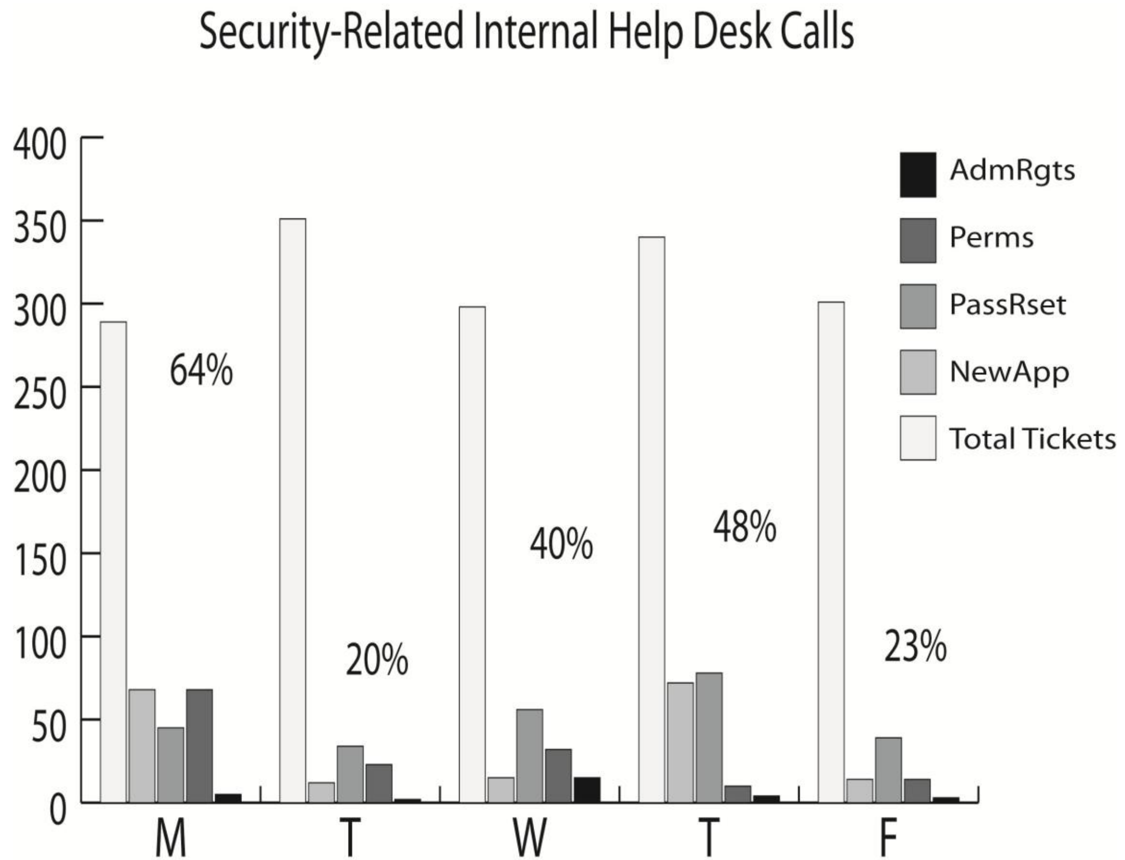
Measure X: subcategory "escalate privilege."

Measure Y: subcategory "reset password."

Measure Z: subcategory "provision application."

Measure T:

The total number of calls to internal help desk.




Security-Related Internal Help Desk Metric: (W+X+Y+Z)/T

# External Activity Example

Measure X:

The number of dropped firewall connections for a 24 hour period.

Measure Y:

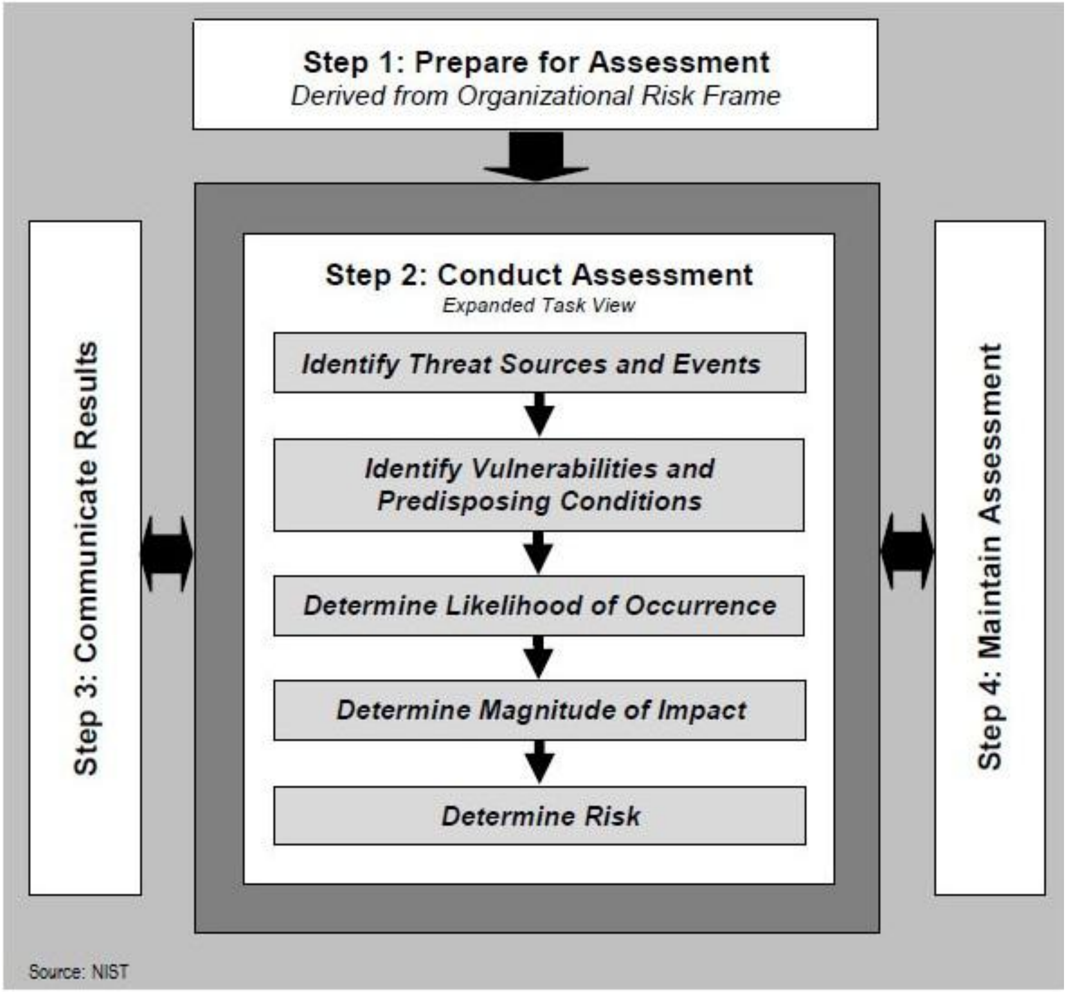
The number of dropped firewall packets for a 24 hour period coming from the same source address, or attacking the same port for that period.

Network Periphery Metric: Y/X

| Failed Source Addresses | | | |
|---|---|---|---|
| IP Address | Country | Times Appearing | Percentage |
| 202.180.216.211 | Mongolia | 765 | 11.81% |
| 81.88.194.131 | Kyrgyzstan | 532 | 8.21% |
| 95.57.171.124 | Kazakhstan | 432 | 6.67% |
| 189.194.171.109 | Mexico | 189 | 2.92% |
| 84.38.68.107 | Germany | 108 | 1.67% |
| 59.37.168.16 | China | 97 | 1.50% |
| 124.158.92.2 | Mongolia | 97 | 1.50% |
| 221.151.17.218 | South Korea | 95 | 1.47% |
| 190.22.130.38 | Chile | 87 | 1.34% |
| 211.240.39.196 | South Korea | 53 | 0.82% |
| ..... | ..... | .... | .... |

| Failed Ports Attempted | | | |
|---|---|---|---|
| Port Number | Port Name | Times Appearing | Percentage |
| 1434 | MS SQL Monitor | 1528 | 23.59% |
| 135 | Several Trojans | 963 | 14.87% |
| 1026 | Calendar Access Protocol | 904 | 13.95% |
| 1027 | ABCHlp | 726 | 11.21% |
| 1433 | MSSQL Server | 361 | 5.57% |
| 22 | SSH | 263 | 4.06% |
| 4899 | W32.RAHack | 216 | 3.33% |
| 5999 | Custom BU App | 188 | 2.90% |
| 139 | Several Trojans | 164 | 2.53% |
| 25 | SMTP | 162 | 2.50% |
| ..... | ...... | .... | .... |

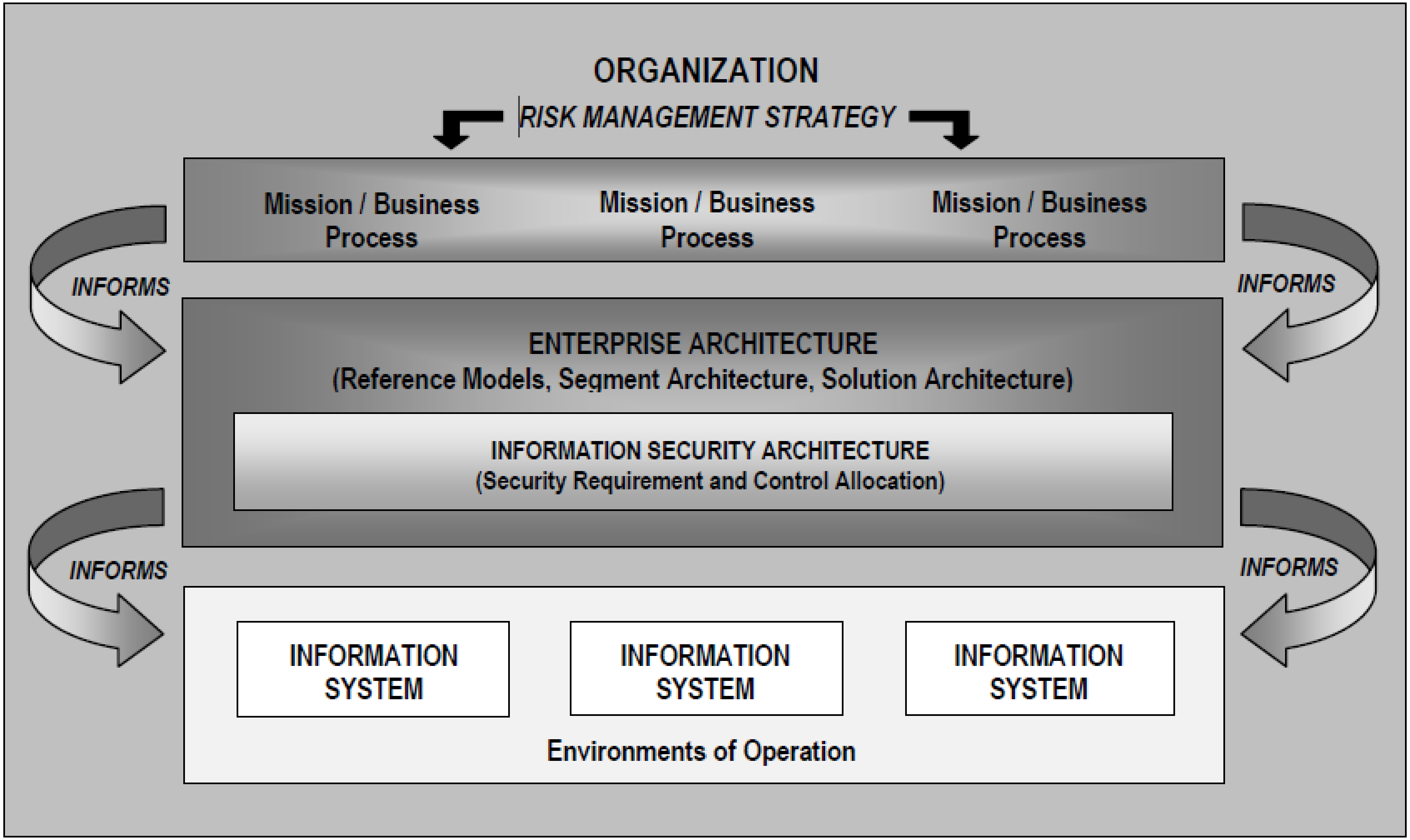
# Security Risk Analysis




The basic approach has been consistent throughout decades of variation.

Debates are not about structure of assessment, but about scope of assessments, probability measures, and appropriate communication techniques.

# Security Risk Management




NIST, "Managing Information Security Risk,“ Joint Task Force Transformation Initiative Interagency Working Group, 2011.

"The specific beliefs and approaches that organizations embrace with respect to these risk-related concepts affect the course of action selected by decision-makers."

***Security Metrics → Risk Analysis → Security Architecture***

NIST-SP800-39, *Managing Information Security Risk, Organization, Mission, and Information System View* (2011)

# Security Metrics Taxonomy

| SECURITY METRICS | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ASSESSMENT | | | | | | CONSTRUCT | | | | | |
| CONTENT | | | BEHAVIOR | | | THREAT | | MODELS | | ACTIVITY | |
| TARGET | MONITOR | REMEDIATION | PERF | VULNTEST | RESILIENCE | SKILLS | GOALS | STOCHASTIC | DETERMIN | INTERNAL | EXTERNAL |

***Construction yields a set of Measurable Security Attributes***

| Security Theory Attribute Construct (STAC) | | | | | |
|---|---|---|---|---|---|
| DESIGN VERIFICATION | | | OPERATION VALIDATION | | |
| TARGET | MONITOR | REMEDIATION | PERFORMANCE | VULNTEST | RESILIENCE |

***These become rules for evaluation, as well as potential hypothesis busters!***

Bayuk, Jennifer. "Security as a Theoretical Attribute Construct." *Computers and Security, 2013.*

# Security SME Survey Results

The most important attributes to measure included:

- Ability to articulate, maintain, and monitor system mission.
- System interfaces accept only valid input.
- Capability for incident detection and response.
- Ability to withstand targeted penetration attacks by skilled attack teams.

The least important attributes to measure included:

- Percentage of systems or components that have passed security configuration tests.
- Security standards used to set requirements.

Yet – *All measures are important!*

J. L. Bayuk. (2011). *Security Subject Matter Expert Survey on Security Metrics. Available at: http://www.bayuk.com/thesis*

To construct a theory that any given system is secure must emphasize validation, and so requires a construction of at least four *dimensions* of attributes:

1. Correct configuration, to allow for design verification.
2. Effective performance, to allow for operation validation.
3. Ability to deflect known threats, or vulntest validation.
4. Ability to adapt to unexpected harmful impact, or resiliency validation.

*Building on target example C, a simple security theory constructed from measurable system attributes is:*

"Security" **=def**

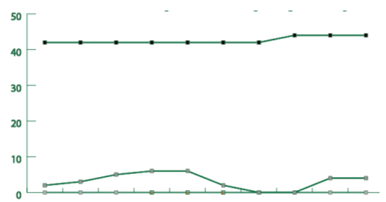
"all devices are configured as designed

**AND**

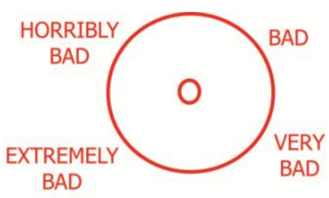
monitoring reveals no errors in execution of the process that maintains configuration

**AND** 0 vulns are found in testing for known vulns

**AND** proper failover occurs upon damaging impact"




**+**



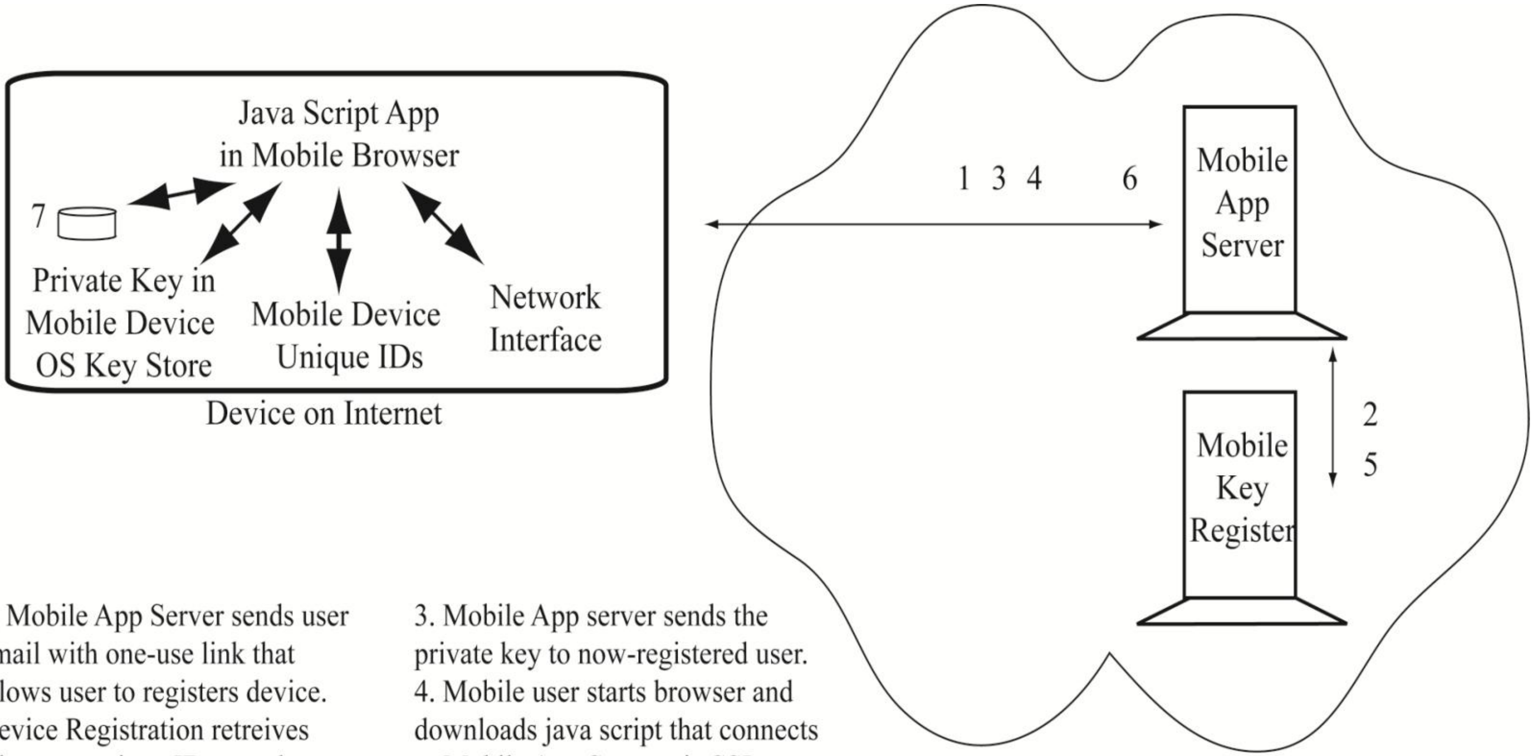

**+** 99.9999999% uptime

---

**Secure System**

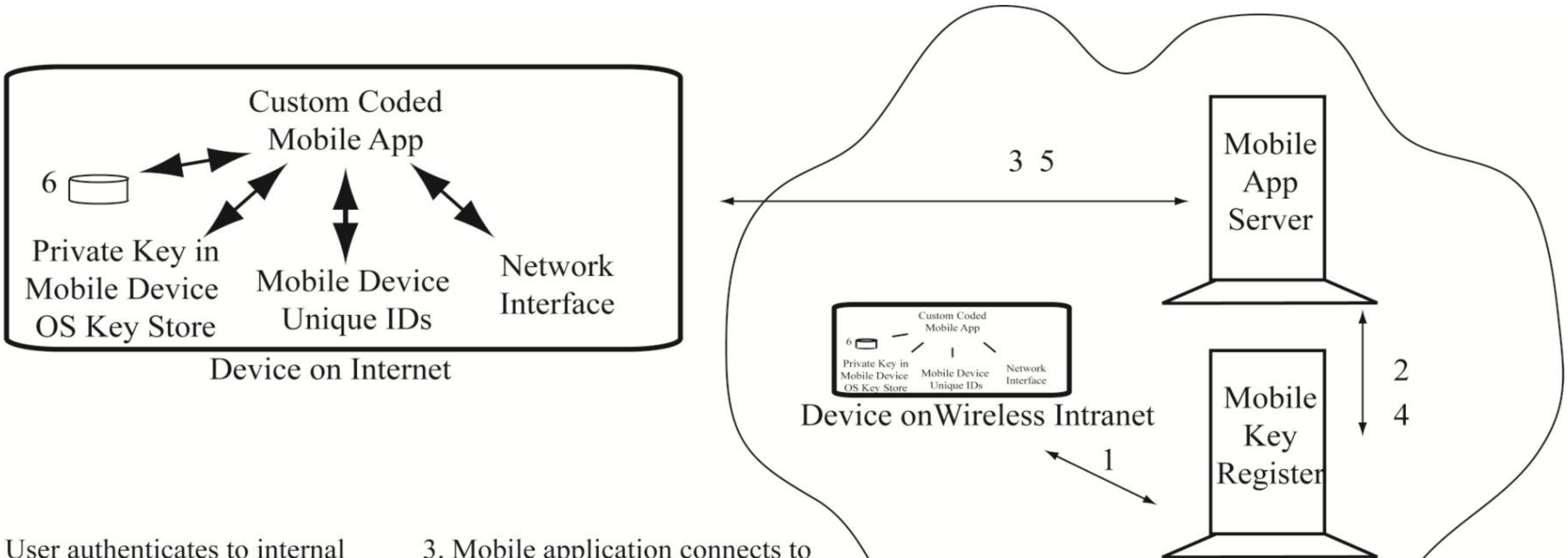
*configuration is maintained while under attack*

# Mobile Architecture Example A




1. Mobile App Server sends user email with one-use link that allows user to registers device. Device Registration retreives whatever unique IDs may be available on the device (e.g., UDID, IMEI, MSISDN) and allows user to select a user ID.
2. Mobile App Server sends user ID to Mobile Key Register and receives a private key, Register retains the public half of the key indexed by the user ID.
3. Mobile App server sends the private key to now-registered user.
4. Mobile user starts browser and downloads java script that connects to Mobile App Server via SSL on Internet and presents user ID and device IDs encrypted with random elements using private key.
5. Mobile App Server retreives public key from Register, decrypts mobile device data, compares it to that registered by user before granting requests for application data.
6. Mobile App Server encrypts data with user public key before sending, also logs transaction.
7. Local device java script app minimizes and encrypts data foot print on device.

# Mobile Architecture Example B




1. User authenticates to internal wireless network and registers device. User downloads and installs a custom mobile application, which retreives unique device identifiers (e.g., UDID, IMEI, MSISDN). Register sends a private key to the user via the application and retains the public half of the key indexed by the user ID.
2. Mobile Key Register sends the user ID and device unique IDs to Mobile App Server.
3. Mobile application connects to Mobile App Server via public Internet and presents data requests together with user ID and device IDs, all encrypted with random elements using private key.
4. Mobile App Server retreives public key from Mobile Key Register, decrypts mobile device data, compares it to that registered by user before granting request for application data.
5. Mobile App Server encrypts data with user public key before sending, also logs transaction.
6. Custom coded mobile app minimizes and encrypts data foot print on device.

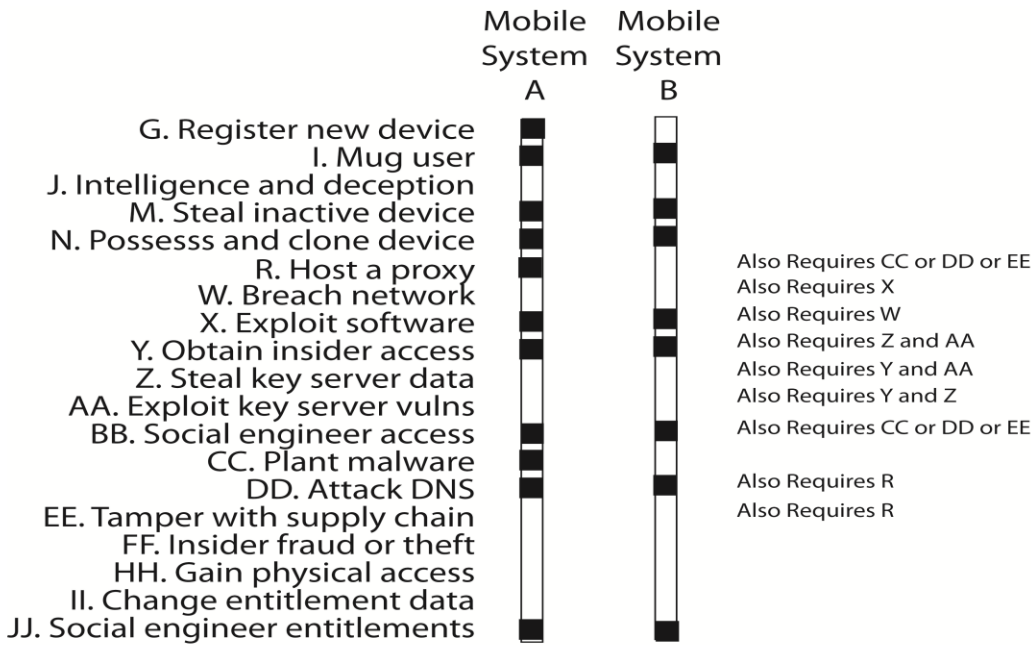
# Mobile System A *versus B*
# Security Theory Attribute Construction

***Candidate metrics for the four dimensions of the construct:***

1. Verified ability for the application server to automatically recognize only registered mobile device users minimizes risk that application data will be exposed to unauthorized individuals. *B is same as A, though different components selected, based on difference in performance requirement of #2.*
2. Users shall have access to application anywhere any time*; in B, from external networks only from preregistered devices.*
3. Vulntest shall reveal, in worst case, data exposure on lost or stolen devices would be limited to small quantities of data of relatively low sensitivity. *B is same as A.*
4. Diverse Internet architecture and agile software support structure render system flexible enough to adapt to unexpected attack. *B is same as A.*

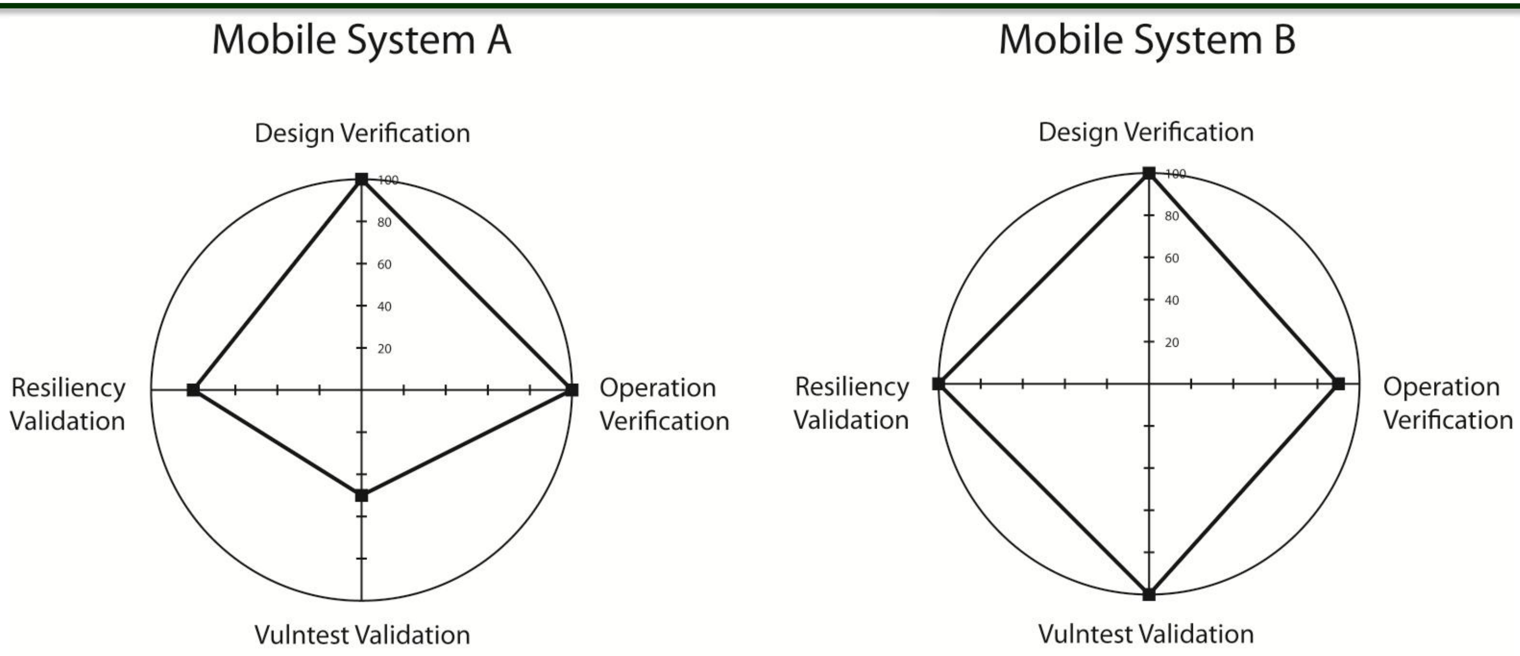
# Case Study Metrics

1. Assume design metrics as in targets and monitor examples.
2. Assume six sigma performance metrics except in cases where users with new devices are not on internal network.
3. Note different architecture would likely produce different vulntest metrics:




4. Mobile System A would be constrained in changing off-the-shelf mobile device software. This would likely affect resiliency metrics.

# Security Trade Space




- For two systems with the same mission and purpose, the performance, the vulntest and the resilience requirements may be expected to be similar enough such that the best metric score in each of these three areas would become the 100% mark for the purposed of STAC.
- Where a system is measured in isolation, the performance, the vulntest and the resilience requirements may instead be set by stakeholder expectations.

# Take-Aways

1. You cannot create a theory of what it means for a system to be secure unless you understand the mission or purpose of the system.
2. You get out of security metrics what you put into them, there is no industry standard approach that will help with validation.
3. Industry standards are focused on verification, and are useful in that capacity. But validation requires sharper focus on system purpose.

# Questions, Discussion?

jennifer@bayuk.com

www.bayuk.com

Reference: Bayuk JL, Security as a theoretical attribute construct, *Computers & Security* (2013), http://dx.doi.org/10.1016/j.cose.2013.03.006