Data Security and Privacy

DSE 3258

Security metrics: Design, Data sources, Analysis of security metrics data, Measuring security cost and value, Different context for security process management

L12 –Security Metrics

TB 7 - Lance Hayden, *IT Security Metrics: A Practical Framework for Measuring Security & Protecting Data*, Tata McGraw Hill, 2016

What are Security metrices?

- As defined by the **National Institute of Standards and Technology (NIST**), metrics are tools that are designed to facilitate decision-making and improve performance and accountability through collection, analysis, and reporting of relevant performance-related data.
- Security metrics can be naturally interpreted as a standard (or system) used for quantitatively measuring an organization's security posture.
- Security metrics are quantifiable measurements used to understand the status of systems and services through the collection, analysis and reporting of relevant data.
- They are based on security objectives that help inform decisions on how to improve the security of all components involved in delivering services and processing data.
- Without good metrics, analysts cannot answer many security related questions.
  - Some examples of such questions include "Is our network more secure today than it was before?"
- *IT Security Metrics* provides a comprehensive approach to measuring risks, threats, operational activities, and the effectiveness of data protection in the organization.

Why are Security metrics needed?

Security metrics are needed to:

- Provide a quantitative and objective basis for security operations,
- Support decision making, e.g. is investment in more security controls needed?
- Support software quality since software security is part of software quality,
- Support the reliable maintenance of security operations, e.g. how often do users need to change their passwords?
- Support the incremental improvement of software's resistance to attacks

Most security metrics follow the SMART structure, which stands for:

- **Specific**

The data must be targeted to the exact area being measured.

- **Measurable**

To be used as a security metric, the data needs to be accurate as well as complete.

- **Actionable**

Data should be easy to understand so action can be taken as soon as possible.

- **Relevant**

All metrics being measured should be important to the data being protected.

- **Timely**

The data should be available when you need it during an analysis.

Good vs bad metric

Collecting valuable data is important, however, if the generation and selection of metrics is done without care, all the data collected will produce useless and meaningless SM.

list of criteria for good metrics

1) Consistently measured, without subjective criteria;

2) Cheap to gather, preferably in an automated way;

3) Expressed as a cardinal number or percentage, not in a qualitative label like “high”, “medium” and “low”;

4) Expressed using at least one unit of measure, such as “defects”, “hours”, or “dollars”;

5) Contextually specific, and relevant enough to decision-makers that they can act.

- Designing Security Metrics:
  - Designing effective security metrics requires an understanding of the business objectives, security goals, and the types of data that are available.
  - Metrics should be designed to measure the effectiveness of security controls and policies and to provide insights into the security posture of the organization.
- Data Sources:
  - The data sources for security metrics can come from a variety of sources, including security tools, logs, and other data sources.
  - The data should be reliable, accurate, and relevant to the security metrics being measured.
- Analysis of Security Metrics Data:
  - Analysis of security metrics data is essential to identify trends, patterns, and anomalies. It also helps to identify areas where security controls need to be improved.
  - Data analysis should be done using appropriate tools and techniques to ensure the accuracy and reliability of the results.

- Measuring Security Cost and Value:
  - Measuring the cost and value of security controls and policies is an important aspect of security metrics. This involves measuring the cost of implementing security controls and policies, as well as the value that they provide to the organization.
- Different Context for Security Process Management:
  - Security metrics can be used in different contexts, such as risk management, compliance management, incident management, and vulnerability management. In each of these contexts, different metrics may be relevant, and different data sources may be used.

Designing Security Metrics

Choosing Good Metrics:

- Nothing Either Good or Bad, but Thinking Makes It So
  - As you develop your security metrics, you should be less concerned with what makes a metric intrinsically good or bad and much more concerned with how you develop measurement projects that provide value and organizational benefits to your security program.
  - This means taking the time to develop metrics that are based on your unique requirements and not relying on “out-of-the-box” metrics that you apply without thinking about what the measurement is supposed to achieve.

Purpose

- Designing security metrics serves several purposes, including:
  - **Improving Security Posture:** Security metrics can be used to identify areas of weakness in an organization's security posture and to track progress in implementing security controls and addressing vulnerabilities.
  - **Demonstrating Compliance:** Security metrics can be used to demonstrate compliance with regulatory requirements and industry standards, such as ISO 27001 or NIST Cybersecurity Framework.
  - **Supporting Risk Management:** Security metrics can be used to support risk management efforts by identifying and monitoring key risk indicators and measuring the effectiveness of risk mitigation strategies.
  - **Providing Management Insight:** Security metrics can be used to provide management with insight into the effectiveness of security controls and to identify areas where additional investments in security may be necessary.
  - **Improving Incident Response:** Security metrics can be used to measure the effectiveness of incident response processes and to identify areas where improvements can be made.

Methods for Deriving Security Metrics from Security Goals

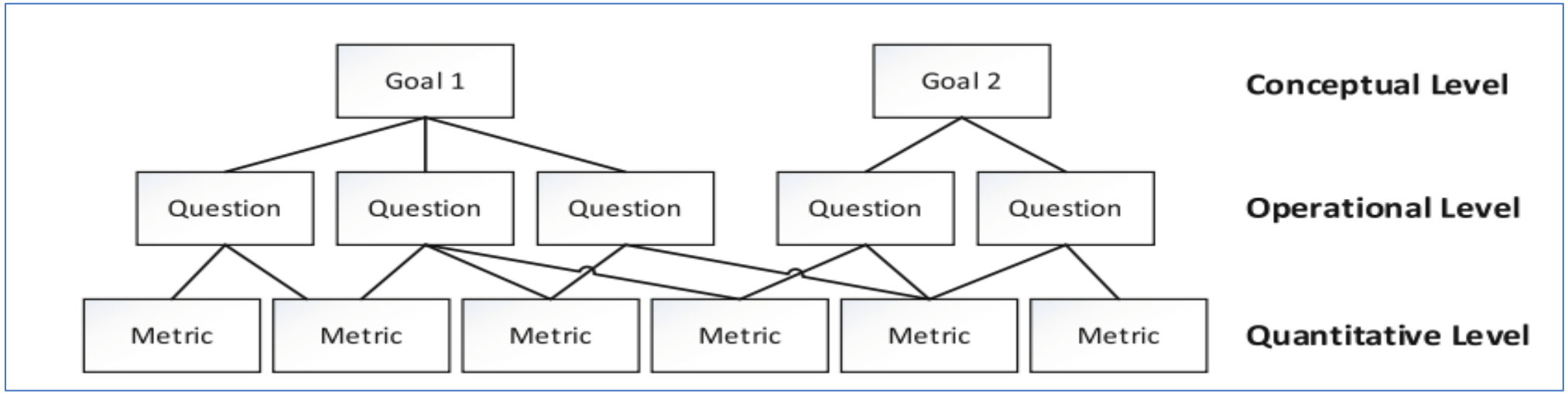
- Three approaches that support metrics derivation from goals:

Method	Proposed Year	Description
GQM (Goal-Question-Metric)	1994	Provides an outline of process that defines goals, refining them into questions and then specifying measurements and finally data to be collected.
BSC (Balanced Scorecard Framework)	2000	Goal-oriented methodology for defining measurement plans.
GAM (Goal-Argument-Metric)	2008	Framework that look into several dimensions for describing, implementing and managing strategy at different levels of an organization by linking objectives, initiatives and measures to an organization's strategy

Goal-Question-Metric (GQM) Approach

- GQM is a simple, three-step process for developing security metrics.
- The first step in the process involves defining specific goals that the organization hopes to achieve. These goals are not measurement goals, but objectives that measurement is supposed to help achieve.
- The goals are then translated into even more specific questions that must be answered before assessing whether the organization has achieved or is achieving the goals.
- Finally, these questions are answered by identifying and developing appropriate metrics and collecting empirical data associated with the measurements.
- The method ensures that the resulting metrics data remains explicitly aligned with the higher level goals and objectives of the measurement sponsors.

GQM Model is a hierarchical structure



- **Conceptual level (Goal)**
  - A goal is defined for an object, for a range of reasons, with respect to different models of value, from different perspectives and relative to a specific domain.
- **Operational level (Question)**
  - A set of questions is utilized to define models of the object of study and after that emphases on that object to describe the evaluation or accomplishment of a particular goal.
- **Quantitative level (Metric)**
  - A set of measurements, taking into account the models, associated with every question in order to answer it quantifiably.

Setting Goals

- Goals give GQM measurements their power, so setting appropriate goals becomes the most important part of the metrics process.
- The GQM method includes a basic template concept for articulating the goals of a security measurement or improvement project quickly and succinctly.

Goal Component	Description	Example
Outcome	The purpose of the project, what will be achieved	Improvement, assessment, understanding
Elements	The boundaries and objects (systems, processes, characteristics) involved in or impacted by the goal	Vulnerabilities, network components, regulatory compliance, system users
Perspective	The point of view taken to understand the goal	External attackers, compliance auditors

Asking Questions

- Goal statements are conceptual in nature. They do not define how the attributes and targets of the goal will be operationally addressed.
- Individual goals are translated into a series of questions that enable the components of the goal to be achieved or evaluated for success.
- These questions articulate the goal and the measurement project in terms of what objects or activities must be observed and what data must be collected to address the individual components of the goal statement.

Assigning Metrics

- After questions have been developed to define the goal operationally, the goal can begin to be characterized at a data level, and metrics can be assigned that will provide answers.
- Designing metrics becomes much more intuitive, because only certain measurements will produce the data necessary to answer the very specific questions that the goal has produced.

Example: Security-Related Downtime

Understanding how long your systems are up and available to users is a common IT metric. Understanding how security impacts availability is also important, particularly when you need to compare security to other IT challenges.

Goal Statement	*The goal of this project is to understand security impacts on system availability by comparing security-related downtime to general availability from the perspective of the security team.*
Question	How often is the system down due to failure?
Metrics	Time between failures Failure duration Mean system availability
Question	How often is the system down due to maintenance?
Metrics	Time between maintenance Maintenance duration Mean system availability
Metrics	How often is downtime the result of a security event?
Question	Number of security events in time period Duration of event remediation

Goal Statement	*The goal of this project is to evaluate the company's compliance with the HIPAA security regulations by comparing company knowledge and activities to the HIPAA compliance guidance for IT systems provided in NIST SP 800-66 from the perspective of regulatory auditors.*
Question	Does the company have a security management process?
Metrics	Number of assets and information systems that create, receive, transmit, or maintain electronic personal health information (EHPI) Number (percentage) of assets and information systems that have not been assessed for EHPI
Question	What are the risks to EHPI under the company's custodianship?
Metrics	Number of risk assessments performed by the company in previous 12 months Mean time between risk assessments
Question	How does the company manage risks to EHPI?
Metrics	Number of approved controls in the company's security controls baseline Ratio of addressable or supplementary to required security controls and implementation specifications

Table 2-5. GQM Project for HIPAA Compliance Using NIST SP 800-66

Understanding Data

Data Types

- Quantitative Data
  - Expressed with numbers and analysed statistically
- Qualitative Data

Data Types

Quantitative Data

Quantitative data is expressed with numbers and analyzed statistically.

Nominal, Ordinal, interval, Ratio

Qualitative Data

Qualitative data **describes qualities or characteristics**. It is collected using questionnaires, interviews, or observation, and frequently appears in narrative form.

If you recorded your interview, the video, audio, and transcripts would also be qualitative data. Analysis of qualitative data is very different than quantitative analysis, as the data is messier, more complex, and requires more interpretation.

Data Sources for Security Metrics

- System Data
  - System and event logs
  - System configurations
  - Source code
  - Test results such as vulnerability assessments or patch testing
- Process Data
  - Activity reporting (budgets, time tracking, training records, meeting minutes)
  - Process tracking (trouble tickets, support call records, compliance monitoring)
  - Workflow breakdowns
  - Business process diagrams (a visual representation of a process that company carries out to achieve a goal.)

Data Sources for Security Metrics (Cont..)

- Documentary Data
  - Security policies and procedures
  - Other policies (which might have an impact on security operations)
  - Audit and review reports
  - Project plans and stakeholder documents
  - Corporate records (financial statements, customer lists, contracts, e-mail)
  - Corporate documents (annual reports, shareholders briefings, SEC filings)
  - Industry reports (analyst research, government reports, market research)
- People Data
  - Surveys and questionnaires (internal and external)
  - Interviews and focus groups
  - Case studies
  - Direct observations

Analysing Security Metrics Data

Reasons for Analysis

- Security metrics data analysis is designed to answer a known, specific question about an aspect of the security program

Date	Time	Action	IN/OUT	Source IP	Destination IP	Service
Oct 28	09:34:20	Accept	OUT	xxx.xxx.110.25	xxx.xxx.200.33	HTTP
Oct 28	09:34:50	Deny	IN	xxx.xxx.66.78	xxx.xxx.110.119	ICMP
Oct 28	09:35:01	Accept	OUT	xxx.xxx.110.25	xxx.xxx.200.33	HTTP
Oct 28	09:35:15	Drop	OUT	xxx.xxx.66.92	xxx.xxx.125.10	FTP
…						

Consider a situation in which a firewall administrator must report monthly on the number of accepted and rejected connections through the corporate perimeter.

Analysing Security Metrics Data

Types:

- Applied Analysis
  - When your security metrics data analysis is designed to answer a known, specific question about an aspect of the security program, is applied analysis.
  - Examples include analyses such as those mentioned in the preceding section, in which statistics on events or security operations are needed for reporting or compliance purposes.
- Exploratory Analysis
  - When you analyze data for the purposes of answering new questions, or even for developing those new questions on the basis of existing information or knowledge, you begin to move from applied analysis to exploratory analysis

Preparing for Data Analysis

- Source of the Data
  - System logs
  - Security event and incident management (SEIM) systems
  - Scanners and analysis tools
  - Audit reports
  - User surveys
  - Company databases (operational and historical)
  - Policies and other records and documents
- In some cases, you may be pulling data from one source that has been collected or aggregated from another source.

Scale of the Data

- **Nominal** Names or labels only, with no quantitative meaning involved even if numbers are used;

Category Value	Operating System
1	Windows XP
2	Windows Vista
3	HP-UX
4	Solaris
5	Linux
6	Mac OS X

Table 3-1. Nominal Categories for OS Type

- For analytical purposes, can use nominal data to build frequency distributions and perform cross-tabulation if you have more than one set of nominal data.
- It is not appropriate to use statistical techniques such as the *mean* (commonly called the average, although the two are different), or the *median* (the middle value) on nominal data

Scale of the Data

- **Ordinal** Indicates ranking order, but with no insight into the differences between rankings; first, second, and third place race results.
- That is Ordinal data doesnot provide any information regarding the amount of difference between the rankings, such as how much faster the winner of the race was compared to the runner up.
- Ordinal data uses numbers to describe a more complex relationship between the targets of observation that is found in nominal data
- Analysis techniques for ordinal data are much like those of nominal data, involving counts of which observations fall into which ranks and the distribution of the data.
- Although people often do it, it is still inappropriate to apply means or averages to ordinal data, because the ordinal scale does not give any insight into the differences between ordinal rankings.

Scale of the Data

- **Interval**
- Where ordinal data describes a ranking relationship, but with no real measure of the distance between individual rankings, interval data involves increases in rank in which the distance between the ranks is measured in some sort of standard unit. Thus the amount of difference between ranks means something.
- Measures of temperature on the Celsius and Fahrenheit scales are good examples of interval data, because the difference between 10 degrees and 20 degrees is the same as the distance between 0 degrees and 10 degrees on each scale .
- It is possible to do more analytically with interval data than with nominal or ordinal data because we are now playing with real numbers.
- We can add, subtract, and multiply measurements.
- We cannot divide or develop ratios between data, however, since the zero point on an interval scale is arbitrary and it is possible to use negative numbers (as with temperature), although this is not always part of the scale (as with academic grades).
- But most common statistical techniques become available with interval data, including the mean, the median, the mode, and the standard deviation.
- Interval data allows us to analyze dispersion, or how "spread out" our data is, and this in turn opens up some interesting probabilistic analysis techniques and the possibility of inferential statistics (those that generalize and predict) rather than more simple descriptive).

Scale of the Data

- **Ratio** Ratio data is pretty much the same as interval data, with the addition of an absolute zero point where nothing exists to measure.
- On a ratio scale, not only is the difference between 0 and 1 the same as the difference between 1 and 2 (as with interval data), but the difference between 0 and 1 is also half the difference between 0 and 2.
- Measurements such as weight and length are measured on ratio scales.
- Analytically, ratio and interval data are very similar, because the data is truly quantitative and allows for a variety of statistical techniques to be performed.

Qualitative Data

- **Data from observations**: Empirical data is based on direct observation. Qualitative data can be highly empirical. Qualitative data of this kind may include written research notes, photographs and drawings, video or audio recordings, and transcriptions of such data.
- **Data from responses**: Response data comes from interviews and interactions with people as individuals and as groups. This type of qualitative data is in the form of records of these interactions, with one person asking questions that are answered by others. The data is still empirical, based on direct observation of the interviewees' responses, but response data tends to be more structured and specific than ethnographic observations,
- **Data from Records and Artifacts:** The third type of qualitative data comprises information produced by our activities. Written documents and texts are common examples of qualitative data, from books and periodicals, to policy documents and corporate reports, to HTML pages and source code. This type of data reflects what you are measuring or observing.

Cleaning or Normalizing

- Does the Data Require Cleaning or Normalizing?
  - For Consistency and Accuracy
  - Missing Data and Outliers
  - Transforming Data

Analysis Tools and Techniques

- Techniques
  - Descriptive Statistics
    - Distribution
    - Central Tendency
      - Mode
      - Median
      - Mean
    - Dispersion
      - Range
      - Variance
      - Standard Deviation
- Tools
  - Spreadsheets
  - **Statistical Software**