Quantitative Cyber-Security

Colorado State University
Yashwant K Malaiya

CS559

Midterm Review

CSU Cybersecurity Center
Computer Science Dept

Midterm coming Tuesday

Will use canvas. Will need proper laptop/pc with camera.

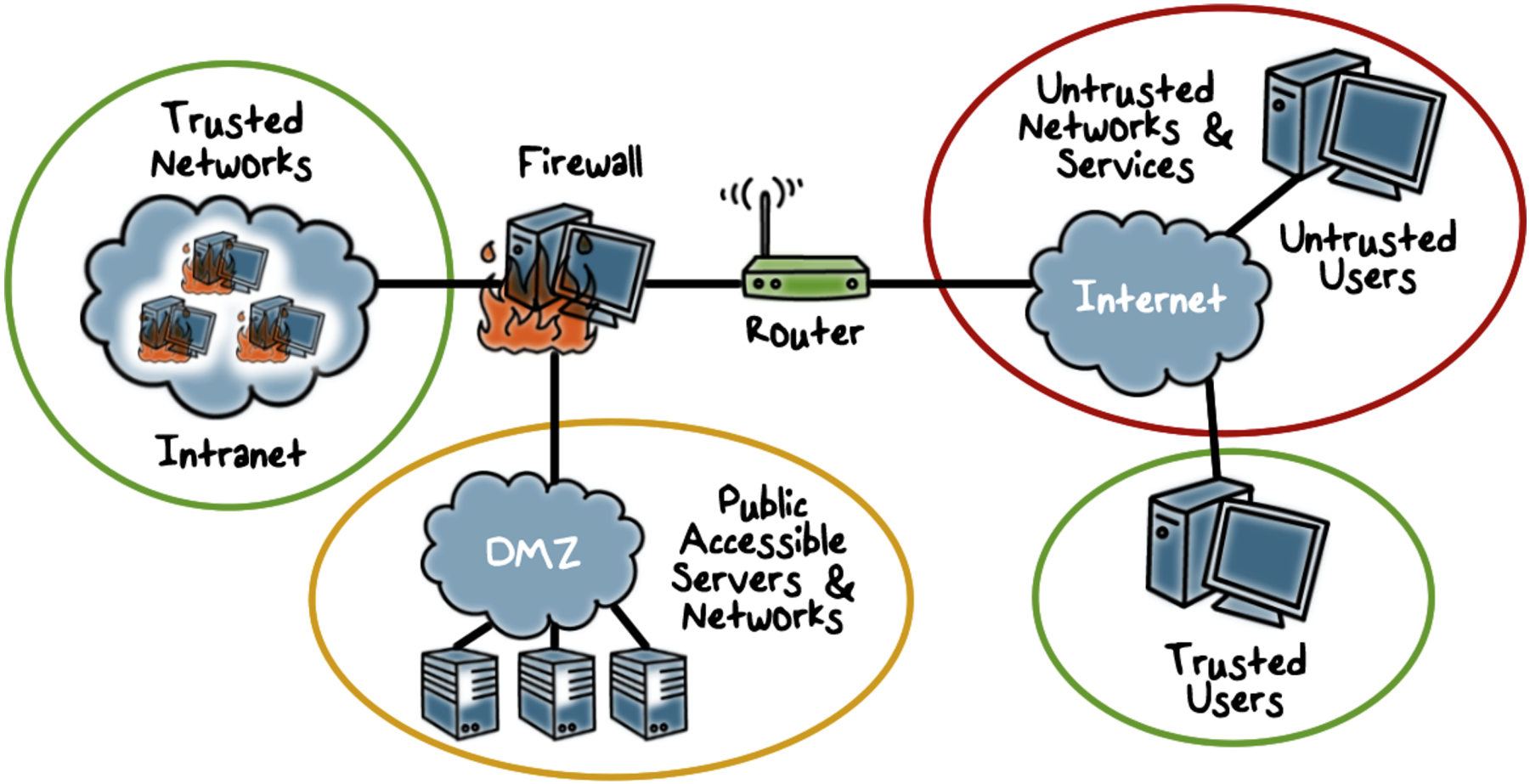
Update: Both sections will use Respondus proctoring.

- Sec 001: 3:30-4:45 PM. Tu.
- Sec 801:
  - 801 students local in Fort Collins need to take it during 3:30-4:45 PM. Tu.
  - Non-local 801 students: During 3:30-4:45 PM. Tu. – 3:30 PM Wed.
- Lockdown browser calculator permitted.
- Closed book, closed notes.

Main topics L1, L2

- Some numbers
- Security system architecture
  - Internet, trusted systems, firewalls, OSs, virtualization
- Assets, Threats, Vulnerabilities
- Cyber attack types, attack surfaces
- Malware: Viruses, worms etc
- Access Control:
  - Subjects, Objects, and Access Rights
  - Access Control Schemes
- Authentication

Firewalls



DMZ: “Demilitarized zone”, distributed firewalls, From Georgia Tech
Note multiple levels of trust.

Example: Access Control Matrix

		OBJECTS			
		File 1	File 2	File 3	File 4
SUBJECTS	User A	Own Read Write		Own Read Write	
	User B	Read	Own Read Write	Write	Read
	User C	Read Write	Read		Own Read Write

(a) Access matrix

Access Control List (ACL): Every object has an ACL that identifies what operations subjects can perform. Each access to object is checked against object's ACL.

May be kept in a relational database. Access recorded in file metadata (inode).

Main topics L3

- How to do research
  - Literature search, sources, reading papers
  - Original research
  - Publication, significance, citations
- Security frameworks
- NIST Cybersecurity Framework
  - Functions and categories
  - Implementations and priorities
- CIS Critical Security Controls
  - Basic, Foundational, Organizational

Main topics L3

- Riski = Likelihoodi x Impacti
- Risk: Possible Actions
  - Acceptance, mitigation, avoidance, transfer

$Likelihood_i = P\{A\ security\ hole_i\ is\ exploited\}.$

$= P\{hole_i\ present\}.\ P\{exploitation \mid hole_i\ present\}$

- Annual loss expectancy (ALE)

  ALE = SLE x ARO

  - Single loss expectancy SLE = AV x EF
    - AV value of the asset. EF exposure factor
  - ARO is Annualized rate of occurrence

Main topics L3

- COUNTERMEASURE_VALUE

  = (ALE_PREVIOUS – ALE_NOW) –COUNTERMEASURE_COST

- Return on Investment

  = COUNTERMEASURE_VALUE/COUNTERMEASURE_COST

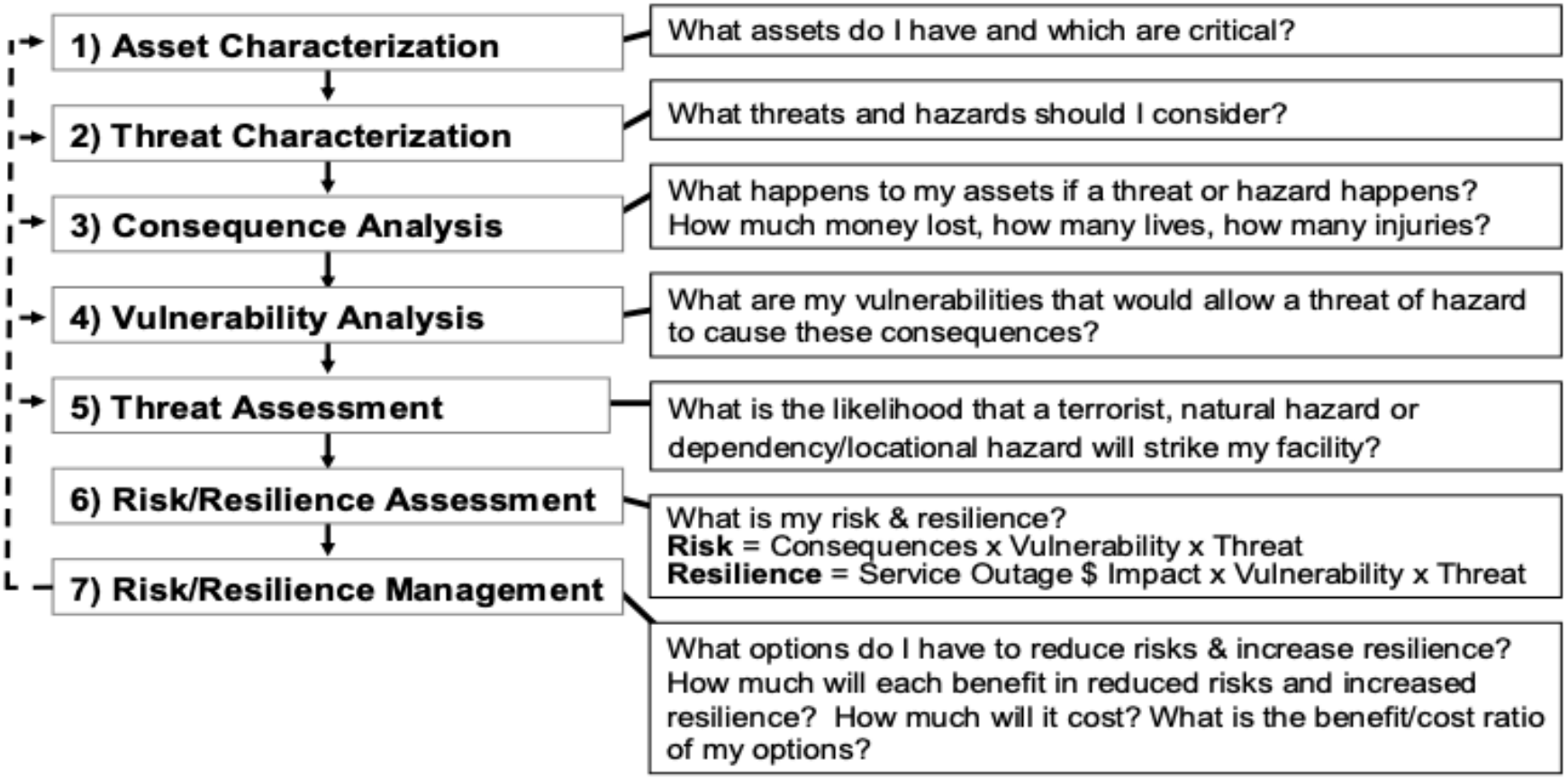
L3

- Log(Risk) = Log(Likelihood) + Log(Impact)
  - Risk score = Likelihood score + Impact score

Likelihood	Consequences				
	Insignificant	Minor	Moderate	Major	Severe
Almost certain	M	H	H	E	E
Likely	M	M	H	H	E
Possible	L	M	M	H	E
Unlikely	L	M	M	M	H
Rare	L	L	M	M	H

L4: RAMCAP

- RAMCAP Framework
  - Risk = Threat x Vulnerability x Consequence



L4: FAIR Framework

- **Factor Analysis for Information Risk**
- Risk = Probably Loss Magnitude x estimated Loss Event Frequency
  - Loss Event Frequency (LEF) = Threat Event Frequency x Vulnerability
    - Threat Event Frequency: table
  - Vulnerability (Vuln) = Threat Capability x lack of Control Strength
    - Threat Capability: table
    - Control Strength: table
- "Multiplication" achieved by using Matrices.

L4/5: **Risk management strategies**

- **Insurance: need**
- Law of large numbers
- Actuarially fair Premium: equal to expected claims

  = probability of illness in a year x average no. of utilization of services per year x unit cost of each utilization
- The loss **ratio is** the ratio of incurred losses and loss adjustment expenses to premiums earned.
- Asymmetric information
- Cyber Insurance: coverage, market, costs

Random Variables

- A random variable (r.v.) may take a specific random value at a time. For example
  - X is a random variable that is the height of a randomly chosen student
  - x is one specific value (say 5'9")
- A random variable is defined by its density function.
- A r.v. can be continuous or discrete

		continuous	*discrete*
Density function	$f(x)dx$	$P\{x \le X \le x+dx\}$	$p(x_i)$
"Cumulative distribution function" (cdf)	$F(x)$	$\int_{x\min}^{x} f(x)dx$	$\sum_{i=i\min}^{i\max} p(x_i)$
Expected value (mean)	$E(X)$	$\int_{x\min}^{x\max} x f(x)dx$	$\sum_{i=i\min}^{i\max} x_i p(x_i)$

L5: Probability

- Disjoint, independent, conditional prob.
- Bayes' rule
- Confusion matrix
  - Sensitivity = TP/(TP+FN)
  - Specificity = TN/(FP+TN)
  - Precision = TP/(TP+FP)
  - Area under the ROC curve

		Actual	
		Disease +	Disease -
Predicted	Test +ve	TP	FP
	Test –ve	FN	TN

Bayes' Rule

- Conditional probability

$$P\{A \mid B\} = \frac{P\{A \bigcap B\}}{P\{B\}} for P\{B\} > 0$$

P{A|B} is the probability of A, given we know B has happened.

- Bayes' Rule

$$P\{A \mid B\} = \frac{P\{B \mid A\}P\{A\}}{P\{B\}} \; for \; P\{B\} > 0$$

- Example: A drug test produces 99% true positive and 99% true negative results. 0.5% are drug users. If a person tests positive, what is the probability he is a drug user?

$$P\{DU \mid P\} = \frac{P\{P \mid DU\}P\{DU\}}{P\{P \mid DU\}P\{DU\} + P\{P \mid nDU)P\{nDU\}}$$

$$= \quad 33.3\%$$

L5: Distributions

- **Density and distribution functions**
  - Binomial, Poisson
  - Uniform
  - Normal, Lognormal
  - In Excel
  - Exponential, Weibull
- Variance & Covariance
- **Stochastic processes**
  - Markov process
  - Poisson process
  - Time between Two Events

L6: Intrusion detection Systems

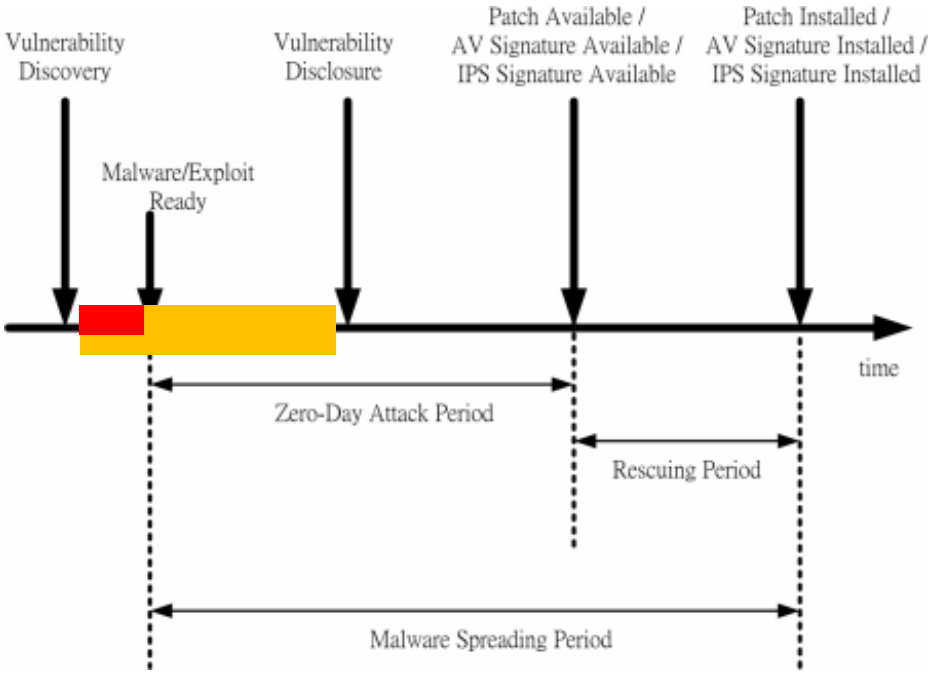
- IDS approaches
- Anomaly detection: Is this the normal behavior?
- Anomaly detection: Is this the normal behavior?
  - No clear diving line between intruder vs authorized user activity
- Rule-based heuristic
- Detections vs prevention (IPS in the path of information flow)
- Host-Based Intrusion Detection (HIDS) vs Network based

L7: Presentations

- Patch management
  - Optimal timing, tools
- Security Economics
  - Gordon-Loeb model
- Mitre ATT&CK Framework
  - Tactics (initial access to Impact for enterprises) divided into many 9-34 Techniques
  - Can be used to launch or foil attacks
  - Tools based on ATT&CK
- Ransomware
  - Attack types
  - Demand vs recovery costs

Discovery/Zero Day Timeline

- Life cycle of a zero-day vulnerability
- Time for exploitation
- Time window for developers to discover bug
  - Incredibly valuable for both attackers and defenders [1]



Colorado State University

L7-L8: Presentations

- Phishing
  - Websites
  - Trends: significant increase
  - Defenses
- Vulnerability Discovery/Zero Day Timeline
  - Time to discovery
- Vulnerability markets
  - Testing and product development cycle
  - Reward programs
  - Black markets
  - Other markets

L8

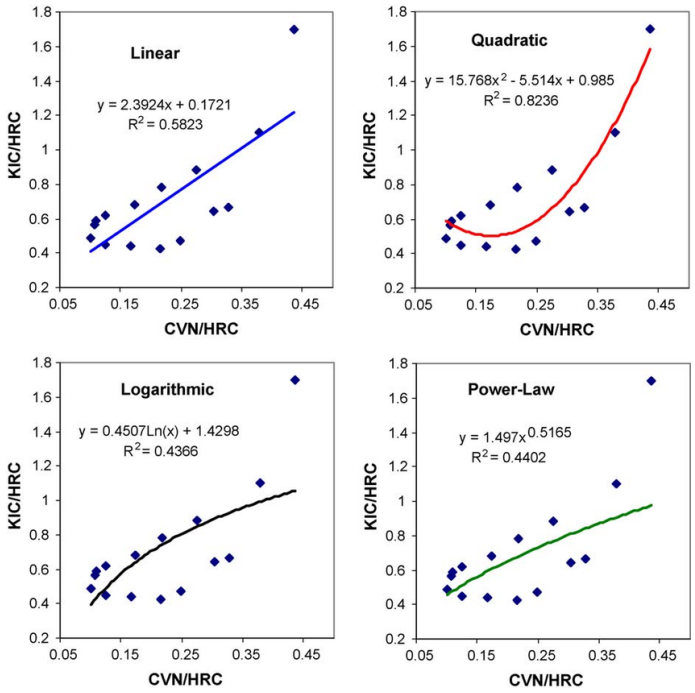
- Security Breach Costs
  - Breach timeline and costs
  - Industry dependence
  - Security Automation?
  - Costs to governments
  - Calculators and indices
- Schemes for discovering previously unknown vulnerabilities
  - Fuzzing: Black-box, white-box, gray-box
  - Fuzzer efficiency

L9: Modeling and regression

- Models: what (derived/empirical) and why
- Curve fitting, tools
- Visualization
- Linear and non-linear: polynomial, exponential, power
- Log for linearization

Empirical models

- Look at data
- See if it resembles a function
  - Linear, quadratic, logarithmic, exponential..
  - Involving 1, 2 or more parameters
- See if it fits
  - If not try something more complex
- If it fits, see if an interpretation of the parameters is possible
  - Not necessary but will be good.



23

- Defects vs vulnerabilities
- Types: software, system/physical, Personnel/procedures
- Components of Likelihood of Exploitation
  - Internal, external, interface
- Annual trends
- Vulnerability Lifecycle
- Vulnerability density and defect density
- Who discovers vulnerabilities?
- Classification of vulnerabilities

L10

- CVE numbering system
- Is it a vulnerability?
- Responsible Disclosure
  - Reward programs
  - Vulnerabilities for sale
- Data bases
- Vulnerability Lifecycle
  - Stochastic modeling
  - Zero-day attacks

L11/12

- Qualys “Laws of Vulnerabilities
  - Half-life, persistence, exploitation
- Modeling Vulnerability Discovery
- Using calendar time
  - AML model: derivation
  - Windows 98, NT
- Using equivalent effort
  - Market share
- Vulnerability density vs defect density

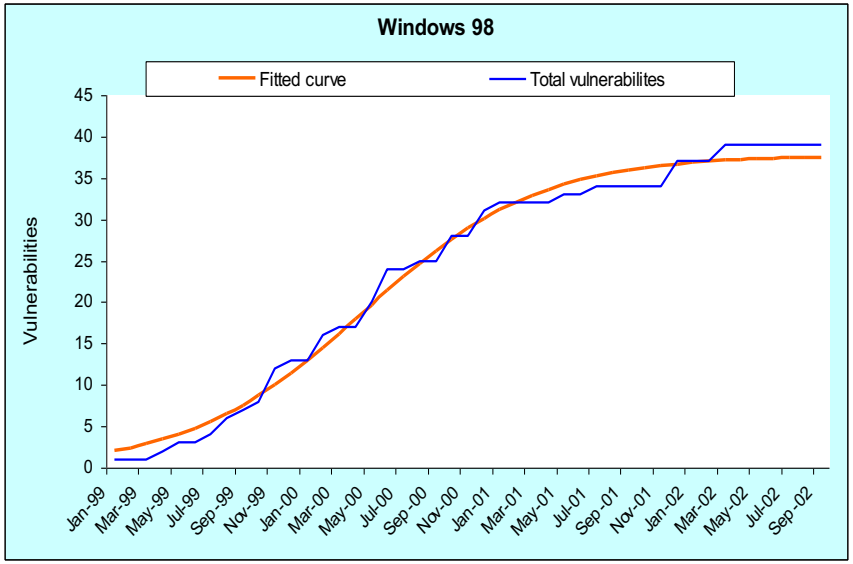
Time–vulnerability Discovery model

3 phase model S-shaped model.

- Phase 1:
  - Installed base –low.
- Phase 2:
  - Installed base–higher and growing/stable.
- Phase 3:
  - Installed base–dropping.

$$\frac{dy}{dt} = Ay(B - y)$$

$$y = \frac{B}{BCe^{-ABt} + 1}$$



Colorado State University

L12: Software Reliability Modeling

- Static metrics
- Exponential SRGM
- Usage –based vulnerability Discovery model
- Nonlinear regression using solver
- Factors Impacting Vulnerabilities
- Seasonality: testing for seasonality
  - Seasonal index analysis with test
  - Autocorrelation Function analysis

L12/13

- Is hacking legal?
- Dimensions and Approximations
- What you should question
- Software Reuse
  - Software Evolution
- Vulnerability Discovery & Evolution
  - Code Sharing & Vulnerabilities
- Multi-version Vulnerability Discovery
  - Humps vs extended linear
- Linear model
- Long Term Trends
  - Size evolution: Linus kernel

L14 Metrics

- Scales: Nominal, ordinal, Interval, Ratio
- Pendleton et al's Survey on Security metrics
- Vectors: entities, vulnerabilities, security state
- Attack-defense interactions in a computer
- Metrics Classification
  - *1. system vulnerabilities*,
  - *2. defense strength and*
  - *3. attack severity*,
  - *4. situation*

Metrics for

- Measuring User (people) Vulnerabilities
- Measuring Interface-Induced Vulnerabilities
- Measuring Software Vulnerabilities
  - Evolution, lifetime, CVSS
- Measuring the Strength of Defenses
- Attack metrics
- Situation metrics
  - Incidents
  - Damage
  - investment

L14: CVSS

Objective: prioritize effort to address vulnerabilities

- Metrics: Components by levels translated into a numerical metric
- Scores: Computed score using a set of metrics as given by formulas
- Three metric groups and associated scores;
  - Base (mandatory): intrinsic to the vulnerability
  - Temporal: time-dependent variation in risk
  - Environmental: risk component dependent on the organization's environment

CVSS Base Scores: Ratings

V 3.0 Severity Rating	Base Score Range
None	0
Low	0.1-3.9
Medium	4.0-6.9
High	7.0-8.9
Critial	9.0 - 10.0

- CVSS Metrics: Data bases
- CVE Statuses in NVD

L14

- Exploitability components:
  - Attack Vector (AV)
  - Attack Complexity (AC)
  - Privileges Required (PR)
  - User Interaction (UI)
  - Scope change
- Impact
  - Confidentiality Impact (C)
  - Integrity Impact (I)
  - Availability Impact (A)

CVSS system: How useful it is?

- What if they had multiplied exploitability and impact sub-scores instead of adding?
- Correlation among
  - CVSS Exploitability, Microsoft Exploitability metric, Presence of actual exploits: small or negative correlations
- Time to discovery? Some possible correlation
- Reward program? Significant correlation
- Time to patch: correlation
- Can metric/score determination be automated? Perhaps.
- VRP Cost effectiveness?

L15 Software testing

- Vulnerabilities are a subset of the defects (1-5%)
- Functional partitioning refers to partitioning the input space of a program.
- Structural partitioning requires the knowledge of the structure at the code level.
- A partition of either type can be subdivided into lower level partitions
- Testing: Functional (or Black-box), Structural, combined
- Random testing/fuzzing
- Coverage
- Input mix: Test Profile

How to prepare

- You have already been preparing
- Review lectures, slides, quizzes, assignments
- Focus on
  - Terms
  - Ideas and approaches
  - Solving problems
- If interested, locate references cited and read in more detail. This is a research-oriented class.
- Please review Respondus information, video. Download and install
- Note: weekend quiz likely