Security Metrics for Cyber Insurance

Sachin Shetty, Ph.D.
Associate Professor
Virginia Modeling, Analysis and Simulation Center
Department of Modeling, Simulation and Visualization Engineering
Old Dominion University
Norfolk, VA USA

Collaborators- Jay Kesan and Linfeng Zhang (UIUC)
Cybersecurity is a National Concern

• Cybersecurity is tied to the health of the U.S. economy. Malicious cyberattacks could throw the financial industry into chaos.
  – The World Economic Forum estimates that ineffective cybersecurity may cost the world’s economy as much as $3 trillion by 2020.

• Cybersecurity is also national security. Critical infrastructure systems, from transportation to nuclear power, are vulnerable to cyberattacks.
  – Hospitals and police departments have been targeted with ransomware that severs access to vital information.
Backward-Looking Cybersecurity

• Current approaches to cybersecurity are often backward-looking.
  – Cyberattack happens → Analyze the hacked system to identify what vulnerability was exploited.
  – The presence of a cyberattack is essentially treated as evidence that cybersecurity procedures were inadequate.
• There is a need for a holistic, forward-looking approach – one that is premised on risk assessment (which leads to mitigation and resilience).
Solving the Cybersecurity Crisis (the “3 Ds”)

• **Defend**
  – Strengthening system security
  – Supporting cybersecurity research
  – Addressing software vulnerabilities as they are discovered
  – Adoption of cybersecurity standards like NIST’s Cybersecurity Framework

• **Deter**
  – Deterrence by denial
  – Criminal laws and punishment

• **De-escalate**
  – Resilience, which is also addressed in the NIST Cybersecurity Framework
  – A cyber insurance regime can improve resilience through financial resources and connecting policy holders to technical resources
“Substitutes” (High demand for one lowers demand for the other)

“Complements” if premiums are tied to self-protection level (Cyberinsurance increases self-protection, i.e. no moral hazard)

“Substitutes” (Availability of one would discourage the other. Self-insurance likely to create a moral hazard)

Source: Kesan, Majuca, and Yurcik, “The Economic Case for Cyberinsurance,” 2004
Cyberinsurance as a Risk Transference Vehicle

• The only existing financial instrument designed for transferring cyber-risk.
• Cyberinsurance preserves market autonomy.
• Cyberinsurance provides carrots and sticks for policy holders to maintain their systems, potentially benefiting society as a whole.
Current Issues with the Cyberinsurance Market

- Information Asymmetry
  - Adverse selection
  - Moral hazard

- Data Scarcity
  - Nature of cyber-risk
  - Correlation
Issues with Cyberinsurance Market ("PEC" Strategy)

- Information Asymmetry (Moral hazard, Adverse Selection)
  - High Premium
    - Few buyers
      - Poor data
      - Bad risk pooling
        - Complexity in products (exclusions, limits)
Underwriting – Process of Distinguishing Risk Levels

• Underwriter looks at several attributes of an applicant and then decides:
  – Whether to take the risk
  – How much coverage to offer
  – How much premium to charge

• Underwriting attributes vary among insurance companies

• Key attributes are not identified

• Underwriting processes in cyberinsurance are mostly questionnaire and/or interview-based
Risk Assessment is the Key

• The essential solution to the informational asymmetry problem is **better risk assessment**.
  – Narrow down the information gap between the insureds and the insurers.
  – Separate insureds with different risk levels and price policies more accurately.
  – Periodical risk assessment helps insurers monitor their risk exposures.
  – Better estimation of cost after losses.

• **We can address this problem from several angles:**
  – Technological Risk
  – Portfolio Risk
  – Legal Risk
TECHNOLOGICAL RISK
Research Goals

Economic Modeling
- Empirical analysis of insurance sector and clients
- Understand market-based and regulatory incentives

Cyber Risk Management
- Modeling technique to quantify cyber risk
- Real time analytics, analyze historical datasets and predict future risks

Measure and Price Risk in Cyber Insurance
Limitations of IDSs

- Cognitive overload to the decision maker
  - Overwhelming number of alerts
- False Positives
  - Concept Drift Issues
- Isolated Alerts
  - Lack of strategies to integrate and correlate alerts
- Missing Information
  - Requires reach-back and updates from higher command levels
- Intent of the attacker
  - Stepping Stones/Pivot Points
Enterprise Security Metrics Challenges

• Need for an **objective way of measuring the security risk** in a network
  – If I choose network configuration A versus B or security control X versus Y, can I measure the risk reduction?
  – How secure is the database server in a given network configuration?
  – How much security does a new configuration provide?
  – How can I plan on security investments so it provides a certain amount of security?
• Need for automated security risk modeling and analysis tool
Enterprise Security Metrics Challenges

• Metric for individual vulnerability exists
  – Impact, exploitability, temporal, environmental, etc.
  – E.g., the Common Vulnerability Scoring System (CVSS) v2 released on June 20, 2007\(^1\)
  – Vulnerabilities have different importance
  – Scoring of a vulnerability depends on context and configuration

• How to \textit{compose vulnerabilities} for the overall security of a network system

• Need to integrate individual vulnerability metrics to compute overall \textit{enterprise security metric}

Attack Graphs

• Adversaries penetrate network through a chain of exploits
  – Each exploit lays foundation for subsequent exploits
• Chain is called an attack path
• All possible attack paths form an attack graph
• Generate attack graphs to mission critical resources
• Report only those vulnerabilities associated with the attack graphs
Common Vulnerability Scoring System

- Provides an open framework which assess the severity level of IT vulnerabilities
- Commissioned by NIAC / Maintained by FIRST
- Quickly becoming the *standard* for application vulnerability risk modeling
- Associates a severity score (CVSS score) to each IT vulnerability, which ranges from 0.0 to 10.0
- Computed by the use of metrics, which result from three vectors
  - Base, Temporal, and Environmental
CVSS Metrics

- CVSS is composed of three metric groups: Base, Temporal, and Environmental, each consisting of a set of

**Base**: Captures intrinsic characteristics of a vulnerability

**Temporal**: Represents characteristics of a vulnerability that change over time

**Environmental**: Represents the characteristics of a vulnerability that are unique to user's environment.

Limitations of CVSS

- Limited to comparing relative severity of different vulnerabilities
  - Does not factor full likelihood of attack (value of asset, combination of vulnerabilities)
  - Assesses atomic attacks.
  - Assumes user environment is safe.
  - Does not take into account security controls
  - Potential for errors or inaccuracies in scoring

- National Vulnerability Database assigns rankings according to CVSS base scores
  - Low: 0.0 to 3.9
  - Medium: 4.0 to 6.9
  - High: 7.0 to 10.0
Cyber Risk Scoring and Mitigation (CRISM)

- **Problem**
  - Need to understand and quantify the impacts of threats to cyber systems and networks in order to develop prioritized mitigation plan

- **Approach**
  - Built over a platform optimized for vulnerability detection, attack graph analysis, and risk assessment.
  - Provides quantitative risk assessment and categorizes attack paths based on the impact of vulnerabilities.

- **Technology Transition**
  - Cyber Risk Scoring and Mitigation (CRISM) tool developed
  - CRISM Commercialization for IT and OT sectors

- **Funding**
  - Critical Infrastructure Resilience Institute, A Department of Homeland Security Center of Excellence
Measuring Security Risk

Cyber Risk Modeling and Analysis
Measuring Enterprise Risk

Bayesian Attack Graph

• Extract topology and vulnerability information for enterprise systems and network from scanning, vulnerability tests and vulnerability databases

• Generation of attack graph to model security state of enterprise system and network

• Developed probabilistic security metric using Bayesian Networks by leveraging attack graph

• Encoding contribution of different security conditions during system compromise.
Measuring Enterprise Risk

Bayesian Attack Graph Analysis

- Examples of vulnerabilities-
  - Unsafe security policy, corrupted file, memory access permission, unsafe firewall properties, unauthorized access

- System and network states represented as attributes and modeled using Bernoulli distribution

- Attacker success to reach goal depends on state of attributes

- Bayesian Attack Graph captures cause-consequence relationships between attributes
Future Technological Risk Work

• Model criticality of assets in the risk scoring

• Provide prioritized mitigation plan

• Simulating security remediation scenarios based on patching vulnerabilities vs enforcing security controls
PORTFOLIO RISK
Data Description

• Analysis are done using the cyber-cases datasets provided by:
  – Veris Community Database, the largest non-commercial cyber-incident database
    • Over 8000 cyber-incidents recorded
    • For each incident, it has information about:
      – Actors, Actions, Affected Assets, Outcome
    • Good for profiling attackers and actions they took
  – Advisen Ltd., a leading provider of data for the commercial property and casualty insurance market
    • Over 40,000 cyber-cases are recorded in the dataset, and over 35,000 cases took place in US
    • The record for each case contains:
      – Timeline (first notice date, report data, etc.)
      – Case characteristics (case type, causes, etc.)
      – Legal information (juris trigger, court, etc.)
      – Outcome (Loss amounts, injuries, etc.)
      – Victim company information (name, sector, size, etc.)
      – Detailed description
      – Possibly applicable line of insurance business
    • Good for studying victim companies and the legal aspect of each incident
Risk Profile Across Different Sectors

Data Source: Advisen Ltd.
Risk Profile

Data Source: Advisen Ltd.
Risk Profile

- Around 14.2% of the cases in US happened to companies located in California

Data provided by Advisen Ltd.
Risk Profile

(Chart in log scale)

- We use number of employees as a proxy for company size
- Overall, loss amount goes up for larger companies, but variance is large
- Small companies can often experience the same or even larger amount of loss than large companies

Data provided by Advisen Ltd.
Risk Profile

Monthly incident count since 2010

Monthly cyber-incident number since 2010
Data source: Veris Community Database (VCDB)
Risk Profile

• Holt’s linear method with multiplicative errors - ETS(M, A, N) best describes the properties of the incident count over time
  – Estimated parameters:
    • Alpha = 0.42 (larger the value, less influence from the past)
    • Beta = 0.07 (trend component)

• Relatively Short memory
  – Incidents from distant past has little influence on the future incidents

• Positive trend component
  – The overall incident number grows at a slow pace
  – The incident count is getting more volatile
    • Possibly because Conflict between lower entry barrier for being an attacker and higher cyber-security awareness

• No seasonal component
  – Making cyber-risk unlike natural catastrophes such as floods and earthquakes that appear to be seasonal
  – No easy way to predict the incident number
Risk Profile

- Attack clustering
  - Discussed in *A Closer Look at Attack Clustering* (Bohme et al., 2006)

  - Clustered attacks arise after a vulnerability is exploited due to the interaction between attacker and defender strategies
Risk Profile

- Attack clustering
  - Similar phenomenon appears on a macro level

Outburst of attacks in a short period
- Vulnerabilities discovered
- New attacking technique
- Etc.

Slowly dies out
- Slow adoption of patches
- Workaround attacks are used until completely blocked
- Etc.
Risk Profile

• Clustering indicates interdependency in attacks
  – Different attacks may have the same cause:
    • Vulnerability
    • Tool/method
    • Actor
    • Target

• From a cyber-insurance aspect, correlation among policies may amplify the losses
  – Several policyholders can be affected in a single incident. e.g.
    • XSS attack affects site visitors on a large scale
    • Vulnerabilities in popular backend platforms can expose the sites using these platforms in danger
Methodology for Analyzing Portfolio Risk

• Assumptions:
  – Every incident results in a claim
  – Claims are all covered.
  – Portfolios are static (not changing in different years)
  – It’s far from being realistic, but simple and good enough for a start

• Generate a list of publicly traded companies in US
  – Exchanges we considered include Nasdaq, NYSE and AMEX
  – 5700+ companies in total after removing duplications
    • different classes of stock or different divisions from the same company are seen as duplications.
    • 6600+ companies before removal
  – Data source: NASDAQ( http://www.nasdaq.com/screening/companies-by-name.aspx)
Methodology for Analyzing Portfolio Risk

• We randomly sample 100 companies from the list to form a portfolio of policyholders
  – Assuming each company has the same likelihood of purchasing cyber-insurance
  – Repeat the process for 1000 time to create 1000 portfolios.

• Using name matching algorithm, we look up the companies from each portfolios in the incident database (VCDB) to see how many of them have incidents in a given year
  – we use data points from 2011 to 2014, since they have the best quality

• Log the incident number of each portfolio. Since we have 1000 portfolios in each year, we get a pretty good distribution of portfolio risk in terms of claim counts.
• The claim count from a portfolio generally follows Poisson distribution (mean ≈ 3.5 out of 100 in 2011)

• Empirical distribution has a larger tail than the fitted one, indicating a higher probability of large claim numbers
Incident Number Distribution of Portfolios of Size 100 in Different Years

2011 (mean=3.497)

2012 (mean=8.965)

2013 (mean=18.337)

2014 (mean=10.936)
Portfolio Risk in Different Years

• Distribution parameter does not hold constant

• The expected value of claim counts can be very different in different years
  – i.e. in the first three years, the mean number keeps growing, but in the fourth year it declines

• Tail is getting bigger over time, which means large claim count is becoming more likely to occur
Portfolio Risk

- Typical methods to model the correlation in portfolios
  - Common shock model
  - Copulas
  - Bayesian Belief Network (BBN)

- Common shock model gets messy when the number of dimensions increases (e.g. a large group of policies that are interdependent)

- Copulas are simple and popular in credit risk analysis
  - Credit risk and cyber-risk are similar in many ways
    - Both bond portfolio and cyber-insurance portfolio have some diversity but correlated risk arises from changes in the big environment
    - We can see a cyber-incident as the consequence of defense level going below some threshold, which is similar to a ‘default’
  - For this reason, many researchers have proposed using copulas to measure the risk in cyber-insurance portfolios
  - Downside: many implementations of these methods need large amount of historical data
Portfolio Risk

• Bayesian belief network
  – It updates our belief based on evidence, so we can easily incorporate new data in analysis
  – Fairly new in portfolio management
  – Few literatures discuss the application of such method in details
  – A possible structure of the network
Portfolio Risk

• Probabilities in the network are updated according to the new information, e.g.
  – Assuming we know A is attacked
  – From this piece of information we induct that possibly there is a vulnerability in the system that A uses being exploited
  – Because B used the same system, the possibility of B being attacked goes up

• Application
  – Update portfolio risk profile in real-time
  – Insurers can change premium rates and reserves accordingly
Thank you