CyberV@R: A Model to Compute Dollar Value at Risk of Loss to Cyber Attack
FloCon 2013

James Ulrich [1]
CyberPoint Labs
julrich@cyberpointllc.com
CyberPoint International LLC

January 9, 2013

[1] with contributions from Charles Cabot, Roberta Faux, Scott Finkelstein, and Mark Raugas

Goals and Motivations

- The ever-expanding threat of cyberattack presents IT administrators and CIOs with the daunting challenge of safeguarding their institutions' cyber infrastructure from breaches that could lead to catastrophic economic loss [Brenner2011], [Clarke2010], [EOPOTUS].
- Security resources remain finite, and deliberations on their wise allocation are aided by expressing risks and risk-reductions in dollar-denominated units.
- Even if we can't accurately predict overall economic loss, perhaps we can compare the relative economic benefit of alternative scenarios for resource allocation.
- So, we'd like a methodology for constructing risk models, at the organizational level, that give insight into relative, if not absolute, economic costs of cyber attack.

Proof of concept: Risk models in finance

- In finance, trading desks maintain Value at Risk (VaR) models for measuring portfolio loss exposure.
- A VaR model answers the question "what is the amount of money \$$X$, such that the odds of losing more than \$$X$, over time window T, fall below some threshold of probability P?" We call this the "P-percent VaR."
- The most vanilla case (c.f. [Hull2000]) involves a portfolio of two stocks A and B. If we know (in \$) the daily volatility σ_A and σ_B of the stock prices, and the correlation coefficient ρ describing how they move relative to each other, (typically derived from historical data), then the P-percent VaR[2] is the value of X such that:

$$\frac{P}{100} = \frac{1}{\sigma_{AB}\sqrt{2\pi}} \int_{x=-\infty}^{x=X} e^{-x^2/2\sigma_{AB}}\,dx.$$

[2] here computed from a normal distribution with mean 0 and variance $\sigma_{AB} = \sigma_A^2 + \sigma_B^2 + 2\rho\sigma_A\sigma_B$

Can we do something similar for cyber?

Goal: perform similar calculations to obtain a distribution of possible $ losses over time, but now due to cyberattack:

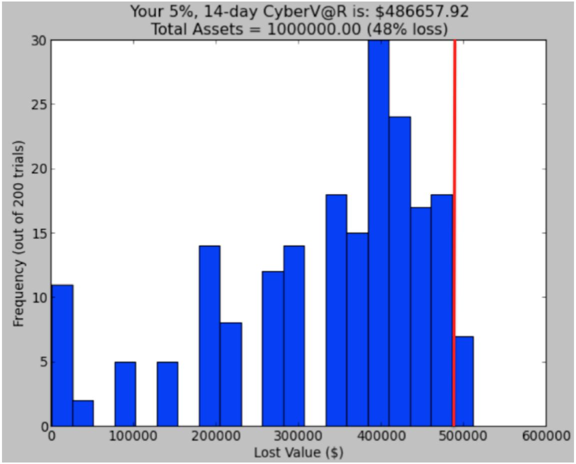


Figure 1: Loss distribution as computed by CyberV@R: red line ≈ $$X$ for P=5%. Note unlike finance example, distribution is not normal.

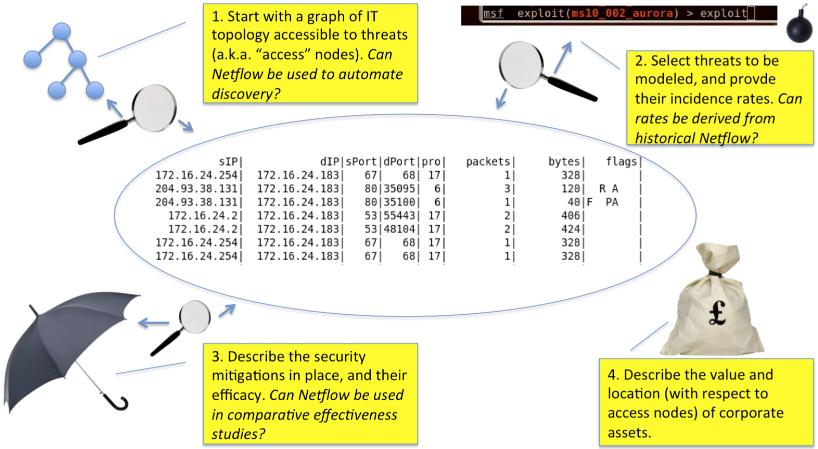
Yes: if we map from finance to cyber

In our cyber application of the finance approach, we will make the following translations:

- Financial portfolio → networked computing infrastructure (Netflow may be a data source for this) and the assets housed there.
- Market fluctuations → threats to which the network is exposed (historical Netflow may provide this).
- Trading strategies → alternative security mitigations we may enable to reduce threats (Netflow may establish historical efficacy).
- Integration over normal distribution $\mathcal{N}(\mu, \sigma) \rightarrow$ Monte Carlo sampling over a two-slice dynamic Bayesian network [3] of attack trees (c.f. [Kol2009], [Pol2012]) representing interaction of threats, network nodes, and mitigations.

[3] a DAG B_i encoding a joint probability distribution, with a rule for transforming $B_i \rightarrow B_{i+1}$

Constructing the model (in pictures):



1. Start with a graph of IT topology accessible to threats (a.k.a. "access" nodes). *Can Netflow be used to automate discovery?*

```
msf  exploit(ms10_002_aurora) > exploit
```

2. Select threats to be modeled, and provde their incidence rates. *Can rates be derived from historical Netflow?*

```
          sIP|            dIP|sPort|dPort|pro|   packets|     bytes|   flags|
172.16.24.254|  172.16.24.183|   67|   68| 17|         1|       328|        |
204.93.38.131|  172.16.24.183|   80|35095|  6|         3|       120|  R A   |
204.93.38.131|  172.16.24.183|   80|35100|  6|         1|        40|F  PA   |
  172.16.24.2|  172.16.24.183|   53|55443| 17|         2|       406|        |
  172.16.24.2|  172.16.24.183|   53|48104| 17|         2|       424|        |
172.16.24.254|  172.16.24.183|   67|   68| 17|         1|       328|        |
172.16.24.254|  172.16.24.183|   67|   68| 17|         1|       328|        |
```

3. Describe the security mitigations in place, and their efficacy. *Can Netflow be used in comparative effectiveness studies?*

4. Describe the value and location (with respect to access nodes) of corporate assets.

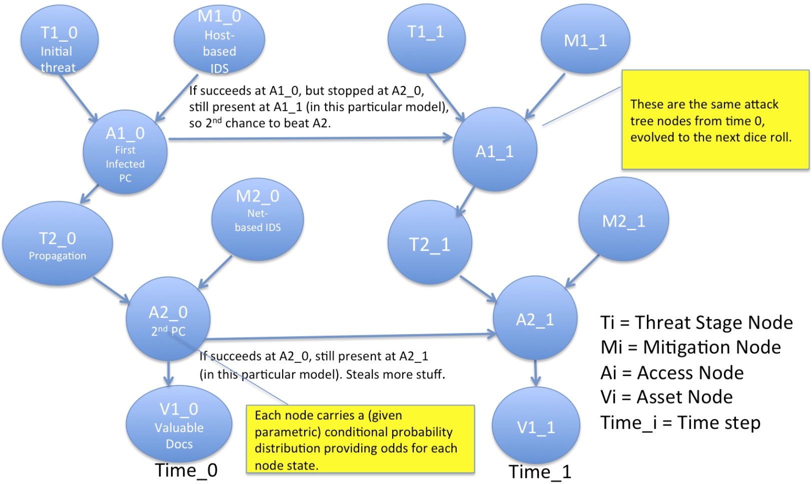
Figure 2: Model is a union of attack trees - nodes correspond to threats, security mitigations, IT infrastructure, assets of value (e.g. product designs). Each node carries a probability distribution describing its odds of being in a given state.

Constructing the model (in words):

- CyberV@R's dynamic Bayesian networks are constructed as a union of attack trees.
- Each node of each tree corresponds to a *threat stage*, a *security mitigation*, an IT element (dubbed an *access node*), or an *asset* (target of threat).
- Each node is assigned a probability distribution, conditioned on the states of its parent nodes, describing odds of the node being in a given state.[4]
- In a *trial*, the attack trees are evolved through time (via Monte Carlo sampling) to get an overall loss (value of assets reached).
- Multiple trials are conducted to produce a distribution on losses.
- The distributions are parameterized, with parameters derived empirically. Hence there is no direct training cost associated to Bayesian network construction.

[4] Threat nodes have Poisson distribution giving odds of *n* occurrences at any time step; mitigation nodes are Bernoulli, giving odds of thwarting any given threat stage occurrence. Access and asset nodes are two-state at each time step (reached/not reached; devalued/not devalued, respectively).

Simplest CyberV@R model (2 PCs; 1 threat)



If succeeds at A1_0, but stopped at A2_0, still present at A1_1 (in this particular model), so 2nd chance to beat A2.

These are the same attack tree nodes from time 0, evolved to the next dice roll.

If succeeds at A2_0, still present at A2_1 (in this particular model). Steals more stuff.

Each node carries a (given parametric) conditional probability distribution providing odds for each node state.

Ti = Threat Stage Node
Mi = Mitigation Node
Ai = Access Node
Vi = Asset Node
Time_i = Time step

Figure 3: Time evolution of a simple CyberV@R Bayesian Network

CyberV@R in the Labs

- We've constructed a CyberV@R model representing CyberPoint's internal network infrastructure at the level of routers, servers, and workstation groups ($\approx$ a dozen access nodes).
- We modeled a single threat based on Symantec's description of the Trojan.Taidoor virus (c.f. [Sym2012]).
- The model computation is implemented using CyberPoint's libPGM (see http://packages.python.org/libpgm).
- We ran the model over 100 trials, each covering a 24-month time step, in the presence and absence of hypothetical workstation software that would remove the virus if found.
- Presence of the AV software led typically to $\approx$ 35% reduction in 5% VaR.
- Computation time less than a minute.

Attack Flow for Single Threat

Trojan.taidoor exploits WinOS vulnerabilities in PDFs, such as BID47314. Drops a backdoor that attaches to services.exe, svchost.exe, communicates with a C&C server. Used to target think tanks involved in U.S./Tawain relations & policy formulation. -Symantec, 2012.

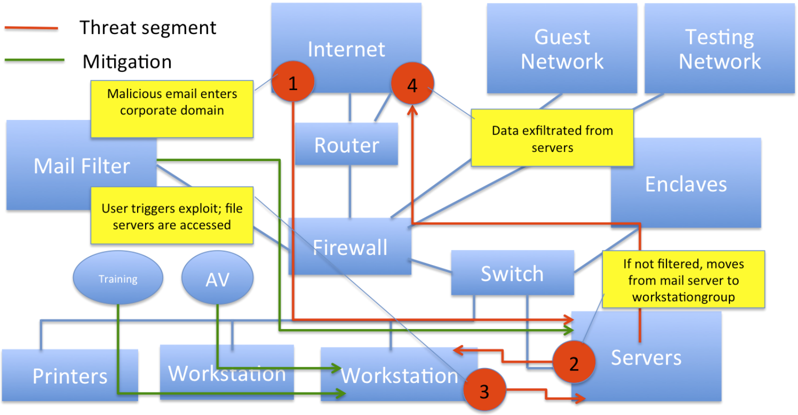

Threat segment
Mitigation
Internet
Guest Network
Testing Network
1
4
Malicious email enters corporate domain
Data exfiltrated from servers
Router
Mail Filter
Enclaves
User triggers exploit; file servers are accessed
Firewall
Switch
If not filtered, moves from mail server to workstationgroup
Training
AV
Printers
Workstation
Workstation
Servers
2
3

Figure 4: Attack flow of Trojan.Taidoor

Corresponding Attack Tree

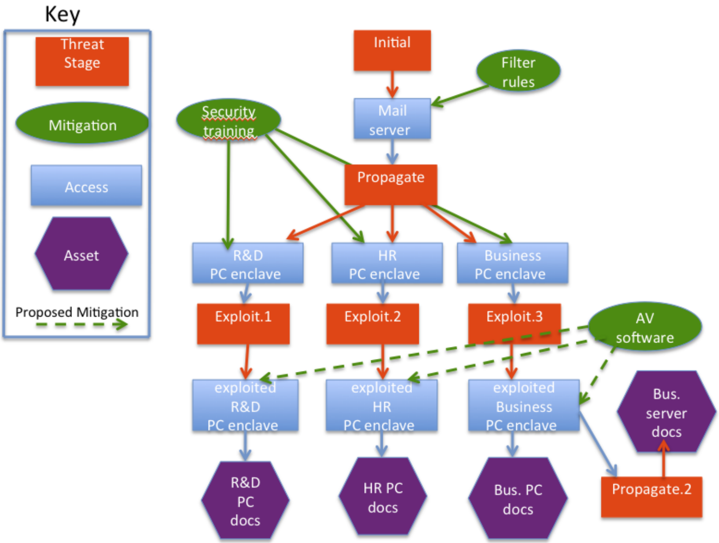


Figure 5: Partial attack tree for one time-step of evolution

Reduction in CyberV@R

We see from the graphs that the $ amount of the 5% VaR, expressed as a percentage of total projected value of intellectual property, is reduced by ≈ 37 percentage points, when virus-removing software is introduced on each workstation node (giving the virus less opportunity to spread).

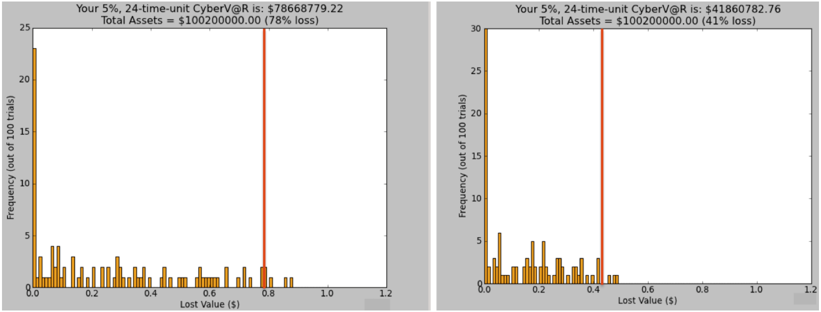


Figure 6: Computed reduction in VaR when AV added to workstations

Scaling CyberV@R

- We're exploring use of Netflow and related tools to automate construction of the IT infrastructure input to the dynamic Bayesian networks.
- Historical Netflow data might be sampled and categorized with aid of visualization tools, to uncover empirical incident rates for threat types. See for example [Yin2005]. This could be automated as well.
- For organizations with 100,000s of nodes, CyberV@R computation can be deconstructed as a series of iterated MapReduce jobs. Each iteration covers one time step. The map jobs each work independently on one subnet's worth of information. A single reduce instance combines the jobs into a new Bayesian network.
- Reducer can replace sufficiently infected subnets from the computation chain with a single threat node added to each remaining peer subnet. A large network reduces to a few "last standing" subnets after several iterations.

Thanks and Questions

- *I thank you for your time and attention.*
- *I also thank the FloCon 2013 organizers for the opportunity to present.*
- *Your questions and comments will be appreciated!*
- *Follow the links at www.cyberpointllc.com for the full CyberV@R technical report.*

More Details

ADDITIONAL DETAIL SLIDES FOLLOW.

Proof of concept: Risk models in finance

- The canonical value at risk model (c.f. [Hull2000]) involves a portfolio of stocks; say for exampe U.S. \$10,000 in shares of company A and U.S. \$20,000 in shares of company B.
- Say, based on historical data, the daily volatility σ_A of A's stock price is 5%, and the daily volatility σ_B of B's price is 10%. Assume also that fluctuations in stock price over a time horizon of T days are modeled as $\mathcal{N}(0, \sigma^2 T)$[5]. So the T-day standard deviation for the A holding is given by:

$$\sigma_A = 10,000 \times 0.05 \times \sqrt{T}$$

and similarly the standard deviation for B is given by:

$$\sigma_B = 20,000 \times 0.10 \times \sqrt{T}.$$

[5] a normal distribution with mean 0 and variance $\sigma^2 T$

Risk models in finance (continued)

- Say ρ gives the correlation of stock price movements in A and B. Then the T-day distribution for the change in value Δp of our portfolio is given by $\mathcal{N}(0, \sigma_{AB} = \sigma_A^2 + \sigma_B^2 + 2\rho\sigma_A\sigma_B)$.
- Using this information, one can find X s.t. $P(\Delta p < X) = 0.02$, that is:

$$X \text{ s.t. } 1 - \frac{1}{\sigma_{AB}\sqrt{2\pi}} \int_{x=X}^{x=\infty} e^{-x^2/2\sigma_{AB}}\, dx = 0.02.$$

- We say that X is our 2% VaR (that is, any losses greater in magnitude than $|X|$ fall in the 2% tail of likelihood) . For $T = 10$ and $\rho = 0.75$, $X \approx -\$6382.00$.
- In our CyberV@R model, we will want to perform similar calculations over distributions of possible losses of intellectual property (or incurring of liabilities) over time, due to various forms of cyberattack on our organization's computing infrastructure.

CyberV@R: specification of the model

A CyberV@R model is:

- A particular JSON encoding of a two time-slice dynamic Bayesian network in which each node is one of four types (threat stage, mitigation, access, and asset).
- The Bayesian network describes a union of time-evolving attack trees, one per threat type of interest.
- The edges of the network observe a set of constraints designed to model the flows of multi-stage attacks throughout the IT infrastructure.
- Each node is labelled with a conditional probability distribution; VaR is computed by Monte Carlo sampling over the joint distribution.
- All conditional probability distributions are parameterized, with parameters derived from empirical estimates passed as input to the model. Within the model itself, there is no learning cost associated to discovering / fitting the prior distributions.

CyberV@R: threat stage nodes

- A *threat stage node* represents a particular stage of a particular threat, and is identified by a node id and a time index.
- The associated conditional probability distribution is Poisson: P(n attempts at executing stage at t) $= \frac{\lambda_t^n}{n!} e^{-\lambda_t}$ (this represents the odds of there being n attempts to execute the stage, between time t and $t+1$).
- A threat stage node optionally connects (upstream) to an access node (defined later), and connects downstream to an access node, having the same time index.
- In practice, mitigation nodes might be active threat types as listed by an AV provider, known to exploit certain CVEs (as listed in the National Vulnerabilities Database).
- If an organization has access to historical Netflow data, these might be mined and categorized with aid of visualization tools, to uncover empirical incident rates for threat types. See for example [Yin2005].

CyberV@R: mitigation nodes

- A *mitigation node* represents a security mitigation (IPS, AV software, patch set, etc.). It is identified by a node id and a time index.
- The corresponding probability distribution will be a Bernoulli variable (independent of time) giving the odds of the mitigation thwarting any given attempt by a threat stage of type τ; e.g. P(attempt blocked) $= M$ where $0 \leq M \leq 1$.
- Mitigation nodes have outgoing edges to access nodes only (see below).
- As above, statistical analysis of Netflow data might be used to gauge effectiveness empirically by examining historical data in the presence and absence of comparable mitigations.

CyberV@R: access nodes

- An *access node* represents an element of the IT infrastructure (a router, hub, server, or workstation, or cluster thereof). It is identified by a node id and a time index.
- At time t, an access node is reached by a threat stage with odds given by:

  $$\mathrm{P}(\text{access}) = \sum_{n=1}^{n=\max} \frac{\lambda_t^n}{n!} e^{-\lambda_t} [1 - (1 - (1 - M_{j_1}) \cdots (1 - M_{j_N}))^n] \; ,$$

  i.e. at time t there are N mitigations in place, up to "max" threat stage execution attempts occur, and at least one gets by all the mitigations.
- An access node has as parents a single threat stage node, and zero or more mitigation nodes. It connects to a follow-on threat stage node, or an asset node (the object of the attack). Netflow data can be mined to discover these nodes.

CyberV@R: asset nodes

- An *asset node* represents an aspect of the organization (intellectual property, operational continuity, absence of legal liability) that is at risk due to cyberattack.
- At time t it carries a dollar-denominated value $V_I(t)$, where I is the node id. It has access nodes for parents, and no children.
- The conditional distribution is simple: if a parent access node is reached at time t, then a fixed amount δV_I is taken from the asset node value. Otherwise the asset node value remains as it was.
- The arrangement of threat stage,mitigation, access, and asset nodes over all threat types, at an initial time point, constitutes the starting state of the Bayesian network. One evolves the network through time by sampling each node according to its distribution (always sampling parents before children).

Computing value at risk via Monte Carlo

In outline form, the VaR computation then reduces to Monte Carlo sampling over the network:

Procedure:estimate P-% CyberV@R

Input: JSON-encoded Bayesian Network, # of trials N, # of time steps T, percentage P

Method:

```
LossArray = []
Sort Bayesian Network in topological order
FOR $n = 0 \cdots , N - 1$
    trialLosses = 0
    FOR $t = 0, \cdots , T - 1$
        FOR each threat type:
            Sample each node in order, according to node's CPD
                IF asset node $l$ is reached, trialLosses $+= \delta V_l$.
    LossArray.insert[trialLosses]
sort LossArray(ascending)
return LossArray[floor(P*N)]
```

Bibiolography

- J. Brenner, *America the Vulnerable*, Penguin Press, New York: 2011.
- R. Clarke, *Cyber War: The Next Threat to National Security and What to Do About It*, HarperCollins e-books: 2010.
- Executive Office of the President of the United States, *The Comprehensive National Cybersecurity Initiative*, available at http://www.whitehouse.gov/sites/default/files/cybersecurity.pdf, accessed February 9, 2012.
- J. Hull, *Options, Futures, and Other Derivatives*, 4th ed, Prentice Hall, Upper Saddle River, NJ: 2000.
- D. Koller and F. Friedman, *Probabilistic Graphical Models: Principles and Techniques*, MIT Press, Cambridge MA: 2009.
- National Vulnerabilities Database, version 2.2, available at: http://nvd.nist.gov/home.cfm/
- N. Poolsappasit, et. al., Dynamic Security Risk Management Using Bayesian Attack Graphs, *IEEE Transactions on Dependable and Secure Computing*, Vol. 9, No. 1, January/Februrary 2012.
- S. Doherty, P. Krysiuk, *Trojan.Taidoor: Targeting Think Tanks*, Symantec Security Response, 2011, available at: http://www.symantec.com/security_response/whitepapers.jsp, accessed June 25, 2012.
- X. Yin, et. al., "VisFlowConnect-IP: An Animated Link Analysis Tool For Visualizing Netflows," FloCon 2005.