Software Diversity: Security, Entropy and Game theory

Saran Neti, Anil Somayaji, Michael Locasto

CCSL, Carleton University (Neti, Somayaji)
University of Calgary (Locasto)

August 7, 2012

Motivation to study diversity

If you **don't** want to keep all your eggs in one basket, then

- How many baskets do you need?
- How should you distribute eggs among baskets?

Is it possible to quantify this popular notion?

Outline

- **Model of a software ecosystem**
- **Diversity measures**
- **Anti coordination games**
- **Capture the diversity**

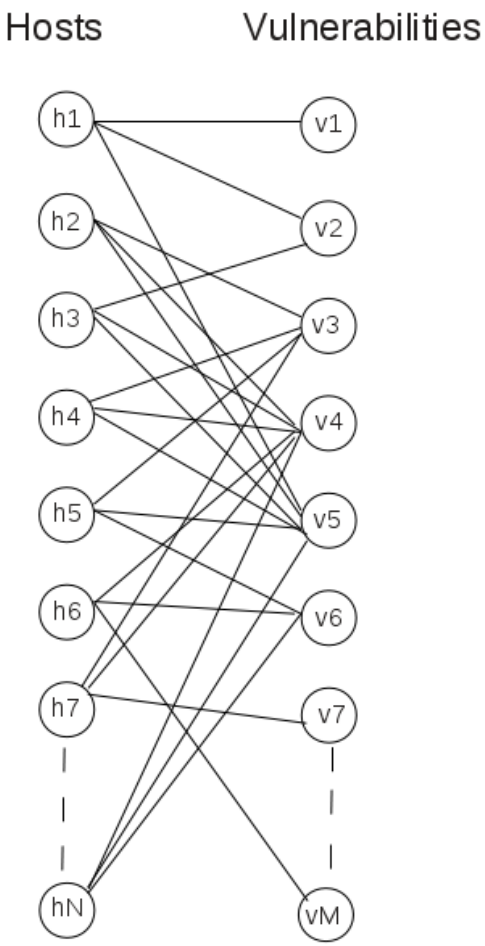
Model - Sets

Software is a bag of vulnerabilities as far as security is concerned

- **Graph** with n hosts, m vulnerabilities, and k vulnerabilities/host.
  - Vertices - $H = h_1, h_2, ..., h_n, V = v_1, v_2, .., v_m$
  - If a host has a vulnerability, there is an edge connecting their vertices.
  - A bipartite graph with kN edges distributed over M vulnerabilities.
- **Software** is a set of vulnerabilities.
  - If a vulnerability v_i is exploited, then $deg(v_i)$ hosts are affected.
  - A host's **strategy** is a subset of vulnerabilities $S = w_1, w_2, ..w_l \subset V$

Model - Graph

If vulnerabilities are what matter for security, lets focus on that



Assumptions

Make all the assumptions you need

1. **Asset value** is uniform. *(Low value targets)*
   - All hosts are equally valuable.
   - All hosts have same number, k, of vulnerabilities.
2. **Residual vulnerabilities** don't change with time. *(Steady state)*
   - If a vulnerability is discovered and patched, nothing changes.
   - Software and vulnerabilities are synonymous.
3. **Vulnerability criticality**
   - An exploit results in complete host compromise.
   - Targeted attacks are not considered.

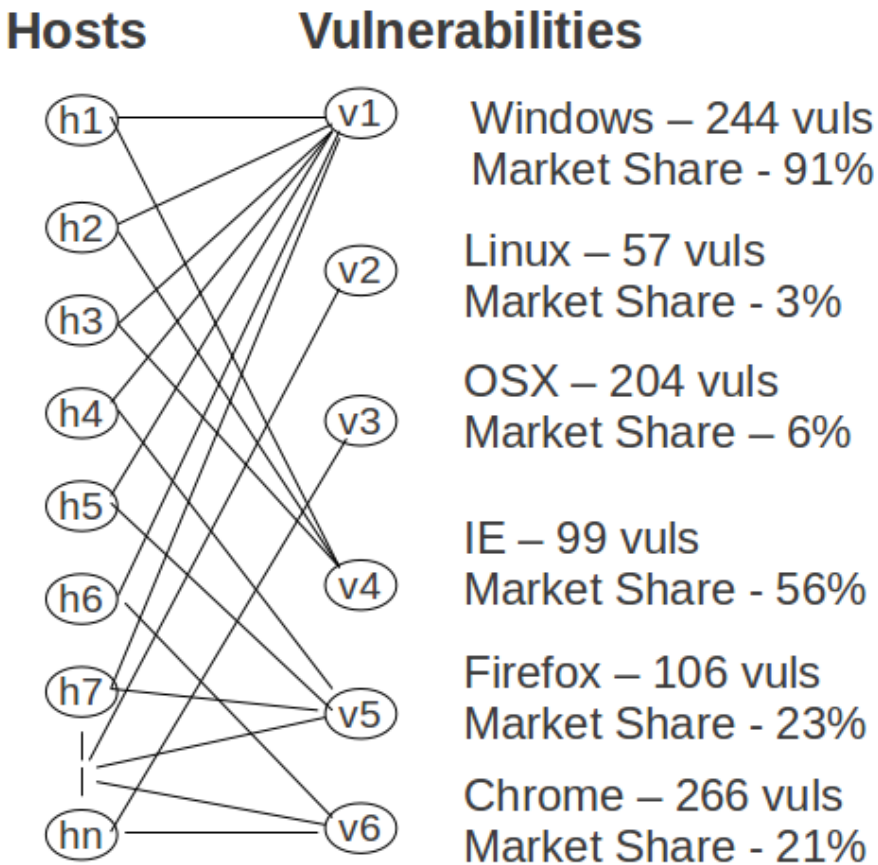
Diversity measures

(1) How many varieties are needed? (2) How are they distributed?

- Set of vulnerabilities $V = \{v_1, v_2, ...v_m\}$
- Define $p_i = \frac{deg(v_i)}{nk}$
- Diversity number $= N_a = (\sum_{i=1}^{m} p_i^a)^{1/1-a}$
- Renyi Entropy $= log(N_a)$
- Shannon Entropy is a special case when $a = 1$

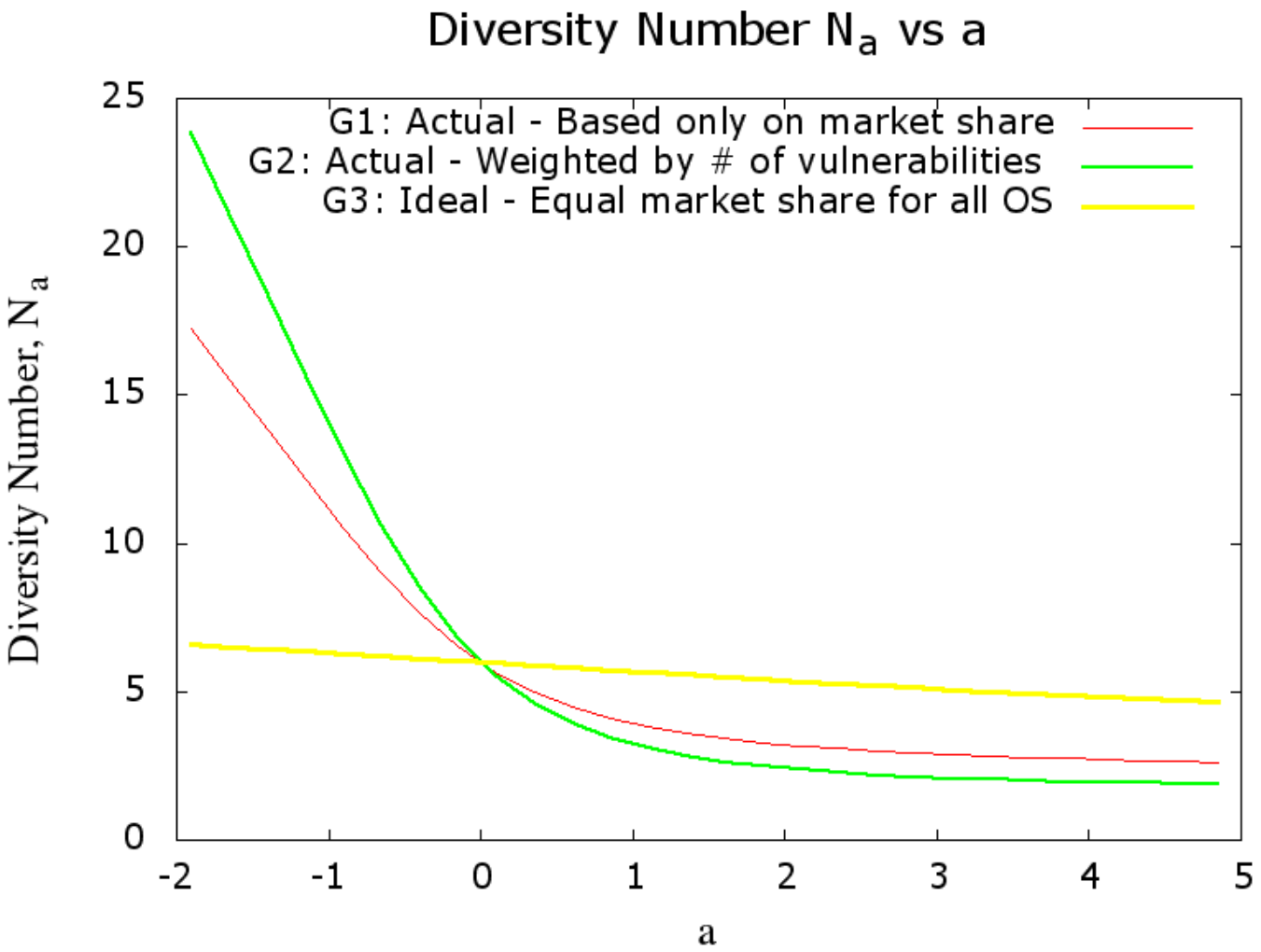
Can we actually calculate diversity?

Market share and Vulnerability data taken from Netmarketshare and NVD



Diversity number $N_a = (\sum_{i=1}^{m} p_i^a)^{1/1-a}$ versus a

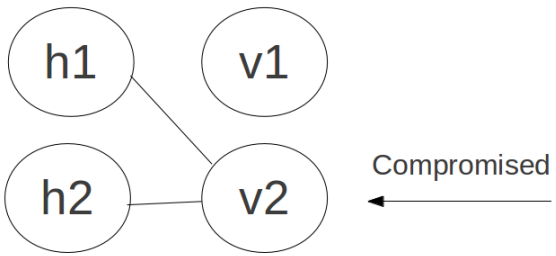
The parameter 'a' changes the relative importance of Factors 1 and 2.



Anti Coordination Game

Why should we care about diversity measures or entropy?

- What kinds of games do we play in software security?
- Do those games have high entropy outcomes?



	Stay	Switch
Stay	$-c_2, -c_2$	$0, -c_w$
Switch	$-c_w, 0$	$-(c_2 + c_w), -(c_2 + c_w)$

- When $c_w < c_2$, there are two pure strategy Nash equilibria
- But there is also a mixed strategy nash equilibrium.

Dispersion Games

How does vulnerability risk grow as market share increases?

- Same game, but with n players, m vulnerabilities.
- Let $\pi(n)$ be the $risk - profit$ payoff multiplier.
- Let c_0 = unit cost of a vulnerability and c_w = cost of switching out a vulnerability.
- Then at equilibrium, $\pi(n_s) = \pi(n - (m-1)n_s) - c_w/c_0$.
- where $(m-1)n_s$ = number of players who switch.

Capture the diversity

What kind of π do people want?

Capture the flag

- Several teams have identical VMs with buggy services.
- Ex. mail server, web server etc.
- Points for attacking, defending and to keep **all** services running.

Capture the diversity

- Simulate options.
- Ex. 2 mail servers : exim or postfix. 2 web servers.
- Keep 1 **out ot** 2 services running.
- Which software will teams select? Most popular? Least popular?
- What will the Host - Vulnerability graph look like?
- How does entropy change with time?

Conclusion

- Diversity := Number of varieties and Distribution of varieties.
- Entropy measures the tradeoff in uncertainity.
- Game theory analyses the tradeoff in stratgies.
- Capture the diversity emperically determines the tradeoff in user choice.

Thank You