Can diversity modeling help security studies?

Bev Littlewood

with Robin Bloomfield, Lorenzo Strigini

(who will answer any difficult questions…)

Centre for Software Reliability, City University

Presentation to IFIP 10.4, Monterey, June 2003

City University London

Introduction and background

- Note question mark in title
  - This is work not-yet-in-progress, but proposed
  - Some ideas about stochastic modeling work we'd like to do
- We are not security experts
  - But we know people who are!
  - And we've spoken to some of you about collaboration
- The idea: much of the work on formal probability modeling of reliability and safety may be applicable to security
  - And there are some interesting needs for extending these models to cope with the particular problems of security
  - Indeed much of our work on dependability & safety cases, modeling human-machine systems (DIRC project), may be applicable

Examples of applicable models

- Reliability growth models
  - Seem directly applicable to security
  - Main problem is identification of a suitable ‘exposure’ variable
    - But ‘time’ may sometimes be OK, at least over a large population of users
  - Will not address these models today
- Models of diversity
  - Already applied to: system design diversity; process diversity; and now argument diversity (my talk at DSN)
  - We think they will apply to diverse intrusion detection

What do diversity models do?

- Informally, diversity is clearly ‘a good thing’
  - But ‘independence’ is not believable, so simple mathematical calculations of efficacy are not available
  - Therefore issue is ‘*how good* is it’
    - How reliable will a diverse system be?
- Why can’t we claim independence? What is the nature of the dependence?
  - Models for (software) diversity by Eckhardt and Lee, Littlewood and Miller, gave insight into this via *difficulty function*
  - See Littlewood, Popov, Strigini, *ACM Computing Surveys*, 2002 for an up-to-date account of all this stuff. But briefly…

Difficulty function - quick intro

- Idea here (using *reliability* terminology) is that inputs to a programme vary in 'difficulty'
  - Here 'difficulty' can be thought of as 'propensity to failure'
  - More precisely, $\theta(x)$, the chance that a program fails on x, is a function of x (the input being executed)
  - Some inputs are intrinsically harder to execute correctly than others
- Eckhardt and Lee model says 'independent programs fail *dependently*'
  - Program *A* fails on a randomly selected input - this means it's probably a 'hard input' - therefore increased chance that *B* will also fail

Difficulty function (2)

- Littlewood and Miller model generalises this to 'forced diversity'
  - Design method *A* tries to overcome (some of) weaknesses of method *B*
  - In the ideal case, inputs that are difficult for program *A* will be easy for *B* and vice versa
- Detailed mathematics depends upon *first and second moments* of the random variables $\theta_A(X)$ and $\theta_B(X)$
  - Probability of failure of a 1-out-of-2 *(A, B)* system is

$$E(\theta_A(X)).E(\theta_B(X)) + Cov[\theta_A(X), \theta_B(X)]$$

  - First term here is naïve 'independence' result

What can security learn from these?

- Dependence of failures between versions comes from subtle interplay between *A* and *B* difficulty variation
  - Apply this to diverse intrusion sensors?
    - Difficulty function over intrusions?
  - Qualitative results presumably carry over
    - E.g. ‘independent’ intrusion sensors do not show independent failures
    - Is anyone in security community assuming they do…?
  - Can we get a handle on *quantitative* efficacy of single detectors, and of dependence between them?

Example

- Suppose you have *n* possible intrusion sensors
- You want to use the most effective *m*-fold diverse detector (*m*<*n*)
  - For example, because of the likely excessive number of false alarms if you used all *n* sensors
- How do you select the *m* sensors to use?
- This decision requires knowledge of the individual efficacies of the sensors *and of dependencies between them*
- Diversity models would help here *if we knew the model parameters*

Some novel modeling issues

- Do models apply also to diversity of intruders?
  - How do you pick the best red team of size *m* from *n*?
  - Ditto selection of intrusion procedures
- Diverse intruders with diverse sensors
  - Can we model this? Not generally an issue in reliability: nature does not mount diverse threats….does she?
  - How do the two types of diversity interact?
- Issue of false alarms - can models be extended to deal with trade-off here (sensitivity versus specificity)?
  - This has not been done in reliability versions of the models, surprisingly (it's needed there too)

Availability of data is crucial

- Real data needed
  - Historical? Honey-traps?
- How ‘strong’, how ‘diverse’, are real systems?
- Can we estimate from data the parameters of our models?
- Validation issues
  - Can we check our models against reality?
  - Can we check model predictions against reality?
  - Can we learn, and improve, from feedback?

Draft paper available

- "Redundancy and diversity in security"
- comments welcome