Impact of Vulnerability Disclosure and Patch Availability - An Empirical Analysis

Ashish Arora, Ramayya Krishnan, Anand Nandkumar , Rahul Telang and Yubao Yang

WEIS 2004

How to handle vulnerability information and its disclosure

Full disclosure Vs. *partial disclosure*

- Pros of full disclosure
  - Information on vulnerabilities enable users to take precaution to reduce losses from breaches
  - Presses vendors to patch earlier
- Cons
  - Leaves users defenseless against attackers who exploit the vulnerability
- Lack of empirical data to validate the pros and cons of *full* and *partial disclosure* – (Gordon et al. [1999])

This research: Empirically tests the impact of vulnerability information disclosure and availability of patches on

- Number of attacks seeking to exploit the vulnerability
- How promptly vendors release patches

Motivation

- Understanding optimal disclosure policy requires measuring
  - Likelihood / frequency of exploit attempts, as function of status of vulnerability
  - Loss from exploits θ
    - Also, share internalized by vendor (λ)
  - Rate of diffusion of patches
  - Cost of patching as function of time and quality
- This research has begun to explore some of these elements.
  - Specifically, estimate how exploit attempt vary with status
  - Estimate a reduced form of vendor decision, reflecting both patching costs and customer losses internalized by the vendor.

Part-1: **Impact of vulnerability disclosure and Patch availability on attack frequency**

Data sources — Attack data

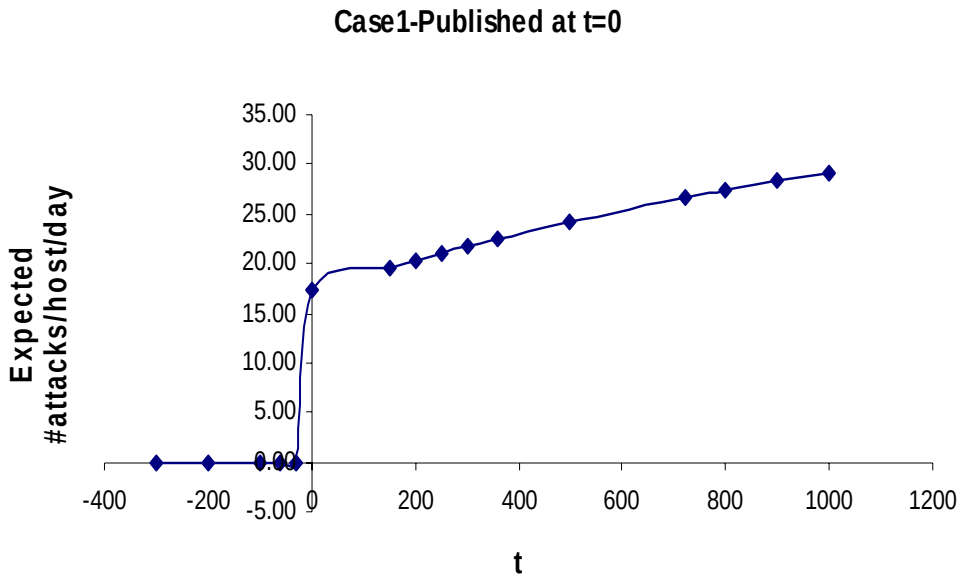
- Data from Honeypots - systems that emulates a computer that is connected to the Internet
- TCP/IP traces from 14 honeypots operating on different operating environments – Linux, Solaris, OpenBSD and Windows – collected for several weeks over the course of a year (2003)
- From the traces, “attacks” created by matching attacking signatures acquired from public data sources
- This data provided us with count of the number of attempts to exploit a specific vulnerability - *the number of attacks per day*
- Randomly selected 308 vulnerabilities from (CVE) database
- Dataset characteristics
  - Unit of observation is vulnerability *i* at time *t*
  - Unit of measure is attacks per day
    - We average within week so variation over time is from week to week
  - 2772 observations over 9 weeks from Nov. 2002 to Dec 2003
- Classify vulnerabilities as *secret, published* or *patched vulnerabilities*
  - A given vulnerability can change status over time

Average effects of patching and publishing

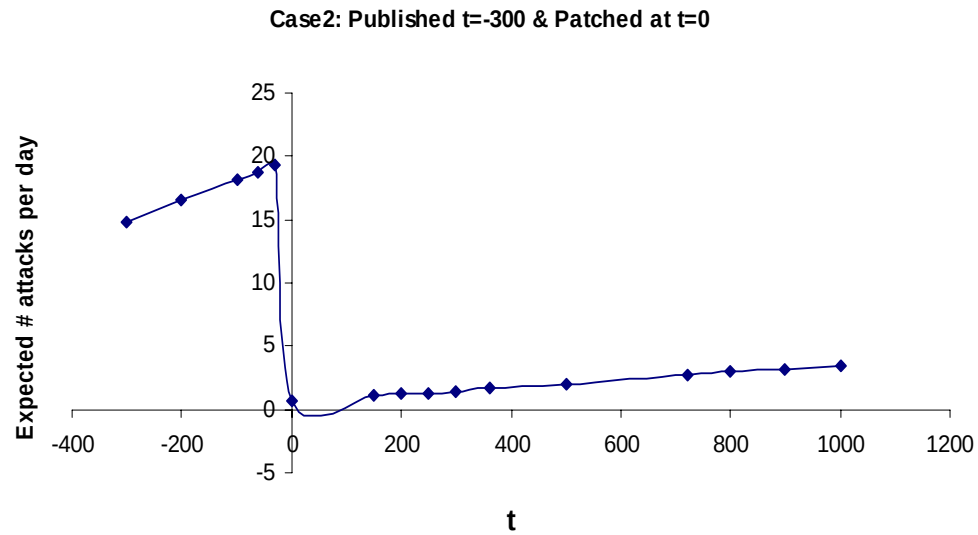
- Non – parametric (*Diff-in-Diff*)
  - Both publishing and patching increase attacks by 0.02 per host per day
    - Identified off only those vulnerabilities that change status
    - No time varying effects controlled for.
- OLS results
  - With vulnerability characteristics:
    - Publishing increases the number of attacks per day per host by 0.26
    - Patching decreases number of attacks by 0.51 attacks per host per day
  - With vulnerability dummy variables
    - Publishing increases the number of attacks per day per host by 0.03
    - Patching decreases the number of attacks per day per host by 0.17

Time effects (From a Tobit Specification)

- Publishing a vulnerability sharply increases the expected number of attacks when published and then increases more gradually with time. (Case 1)

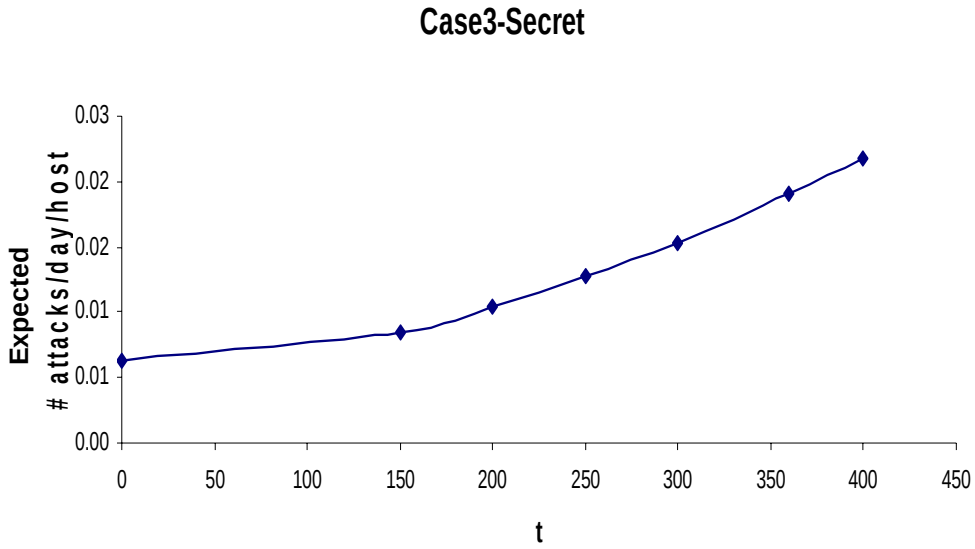


- Patching deters attackers though after patching, with elapsed patch days the expected number of attacks per day per host increases with time, which may be an artifact of our specification (Case2)

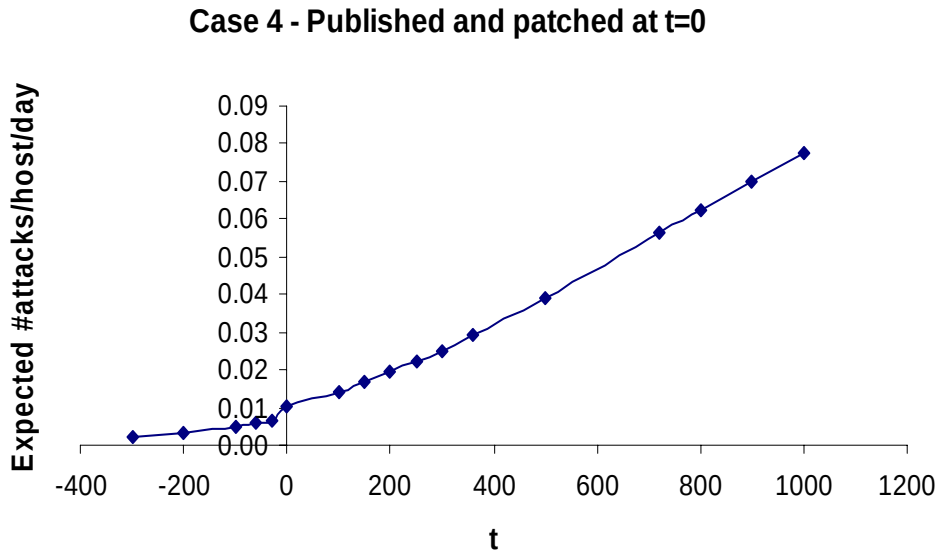


…Time effects

- *Secret* vulnerabilities attract the fewest attacks, though the closer the vulnerability gets to being published the expected number of attacks per day per host increases (Case3)

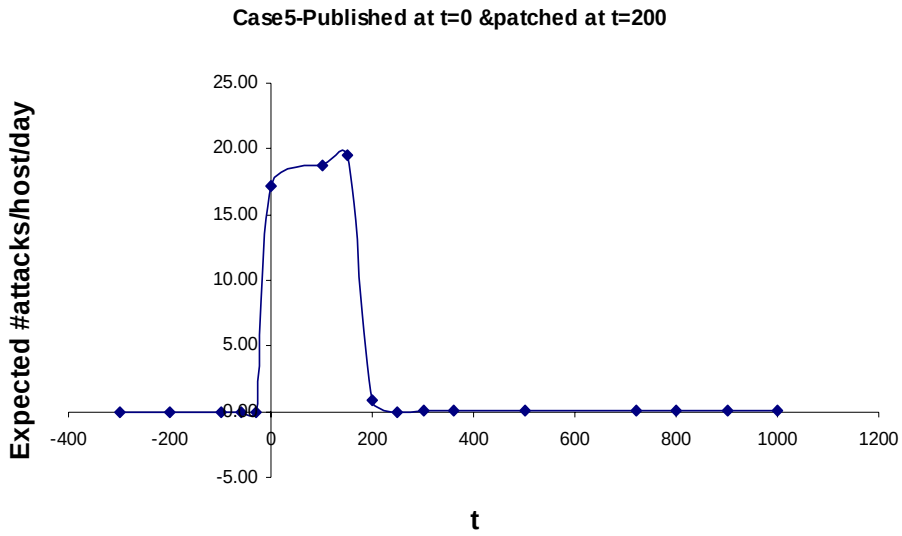


- Vulnerabilities that were patched and published on the same day still get exploited more by attackers than *secret* vulnerabilities (case4)



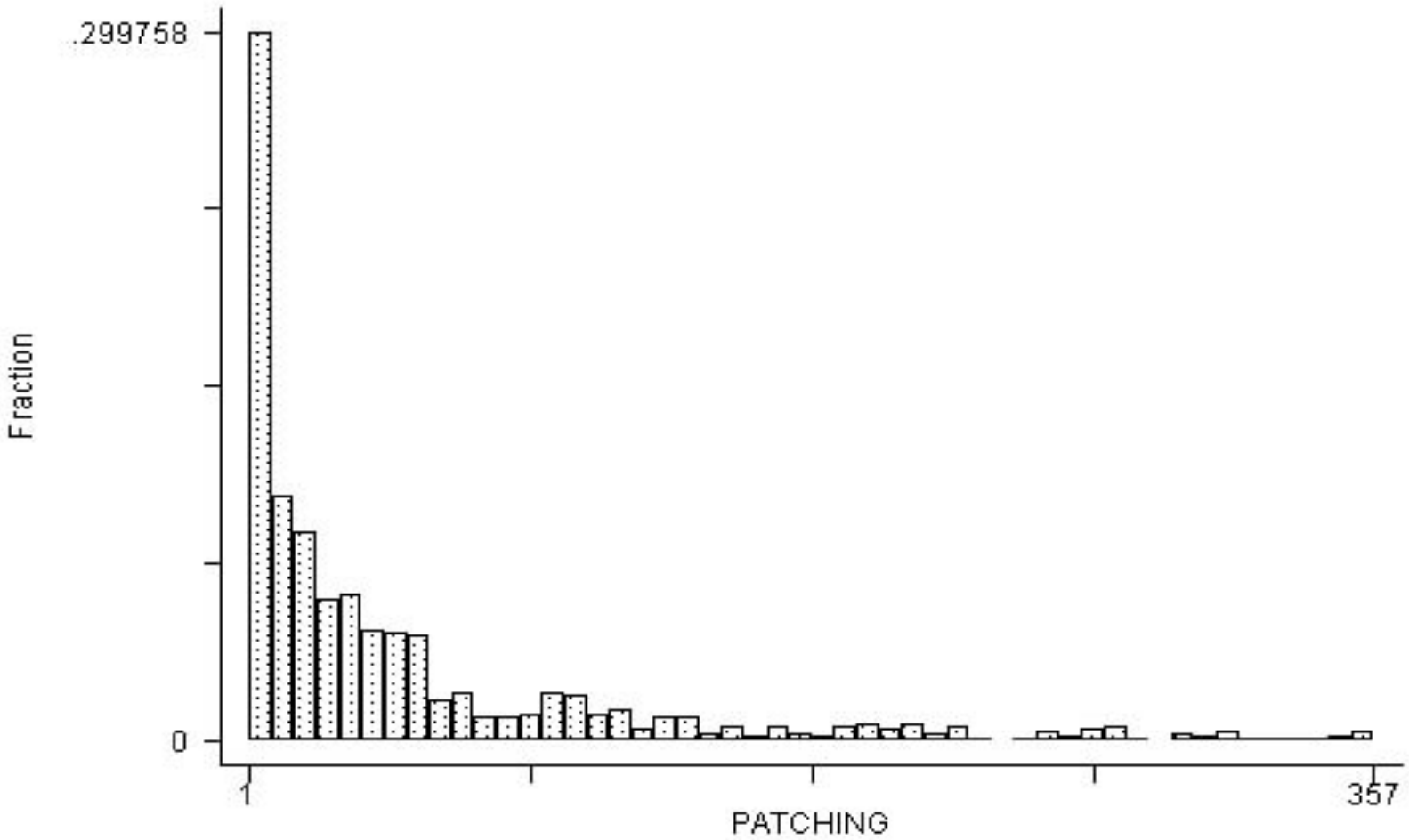
Summary

- Both *patched* and *published* vulnerabilities attract more attacks than the *secret* vulnerabilities
- *Published* vulnerabilities that have no patches tend to get the most attacks.
- With time, the number of attacks per day per host increases in the case of *patched* vulnerabilities but the number of attacks per day per host decreases with time in the case of *published* vulnerabilities



Part-2: **Impact of vulnerability disclosure source on vendor patching behavior**

Data description The distribution of patching time, (only vuls patched within one year)

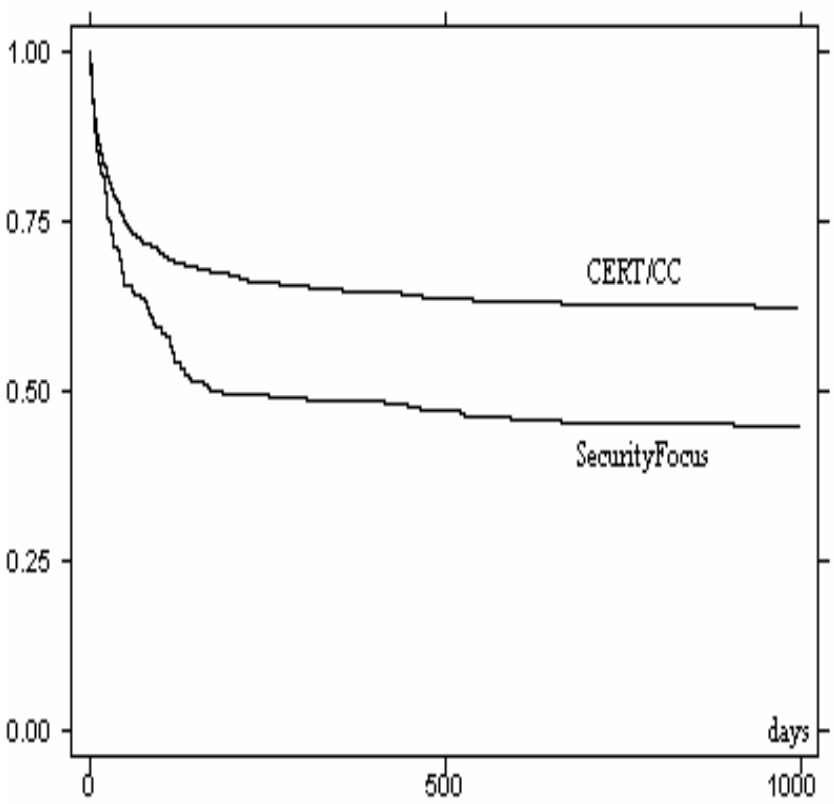


Data description

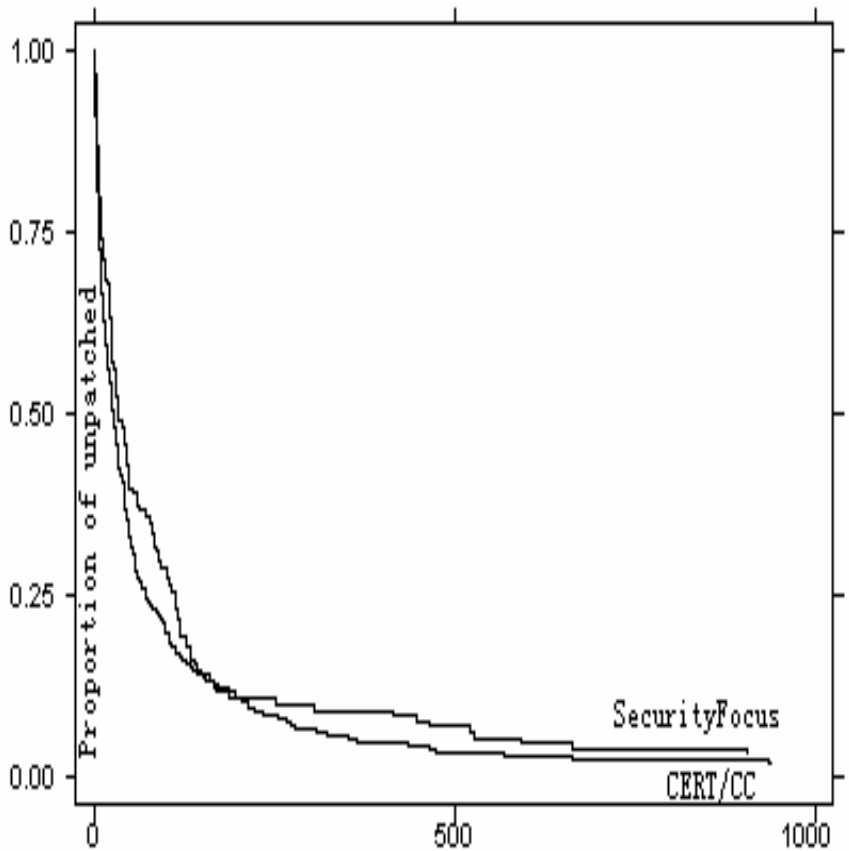
A comparison between policies and open/close sources

	(CERT/CC) Partial disclosure policy	(Bugtraq) Instant disclosure policy	Total
Avg. days to patch	101.21 (373.36)	126.38 (271.91)	103.3€ (365.77
Patched/Total	40%	57%	41%
Employee Size (k)	43.07 (91.14)	40.46 (67.40)	42.92 (89.98
Open source/Total	12%	11%	12%
Severity	25.46 (20.31)	20.35 (12.95)	25.16 (19.98
	UN-PATCHED	**PATCHED**	**Total**
Employee Size (k)	43.77 (95.66)	41.73 (81.33)	42.92 (89.98
Open source/Total	11%	14%	12%
Severity	24.40 (18.67)	26.25 (21.68)	25.16 (19.98
	Close source	**Open source**	**Total**
Avg. days to patch	111.50 (387.44)	51.25 (162.31)	103.3€ (365.77
Patched/Total	41%	47%	41%
Severity	25.75 (20.34)	20.88 (16.57)	25.16 (19.98
Employee size (k)	46.48 (92.60)	3.89 (34.52)	42.92 (89.98

SecurityFocus Vuls: More likely to be patched than CERT/CC Vuls, but conditional on patching, takes about the same time.

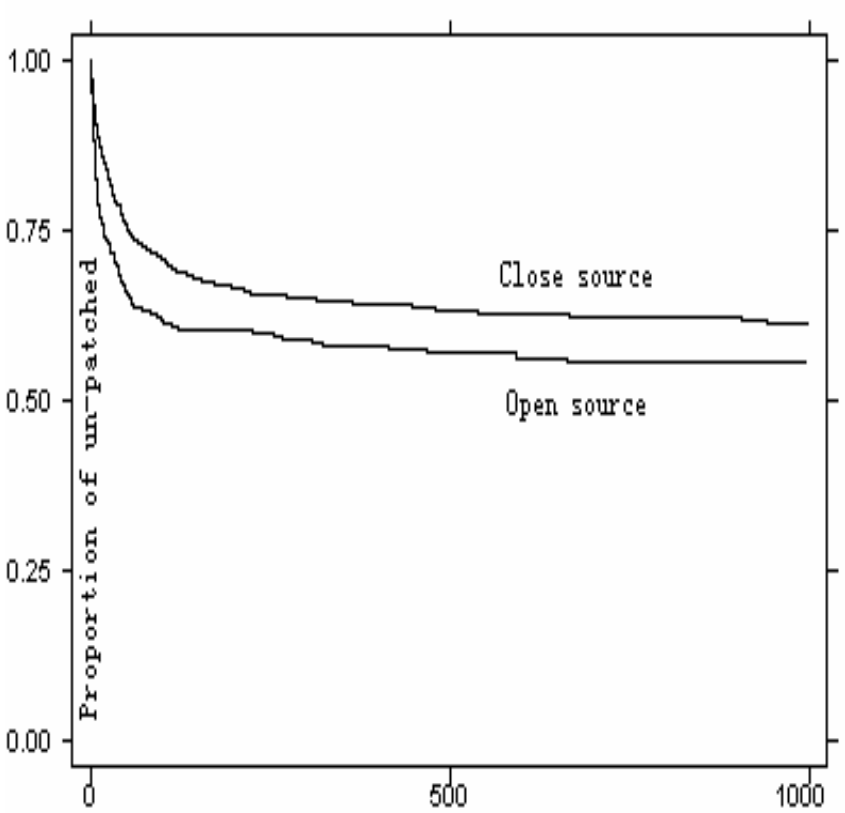


All observations

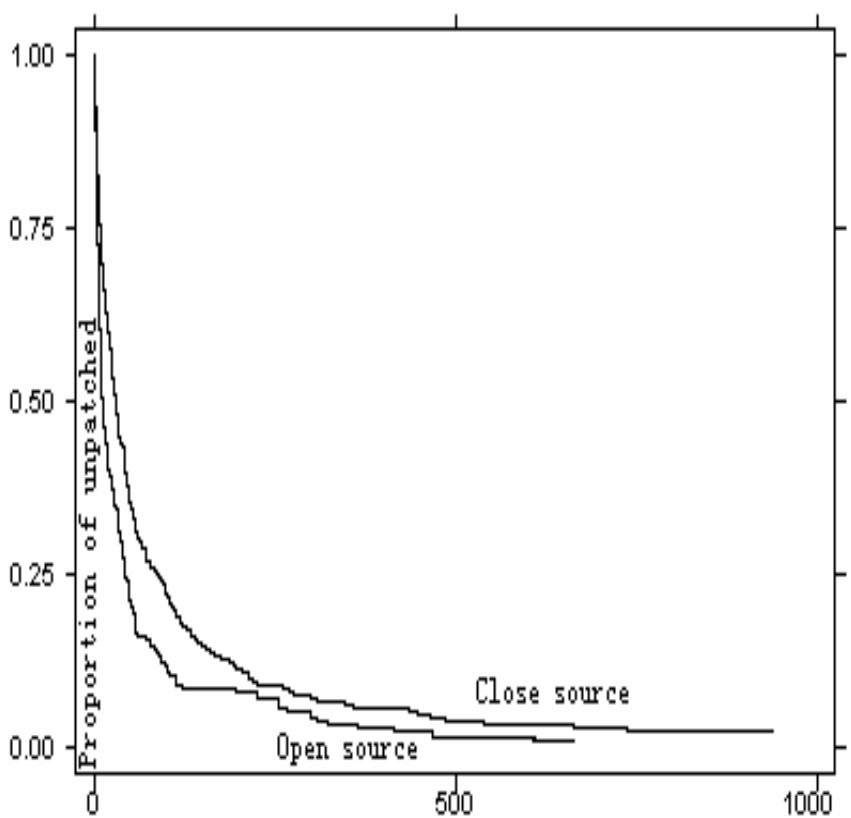


Patched only

Open source vendor are more likely to patch, and patch more quickly.



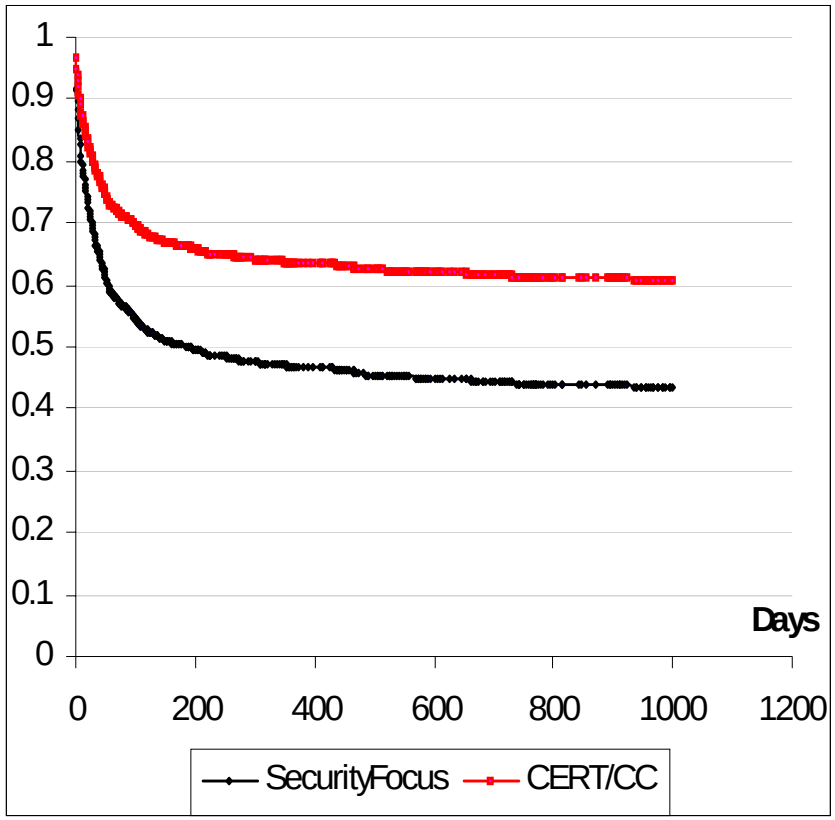
All observations



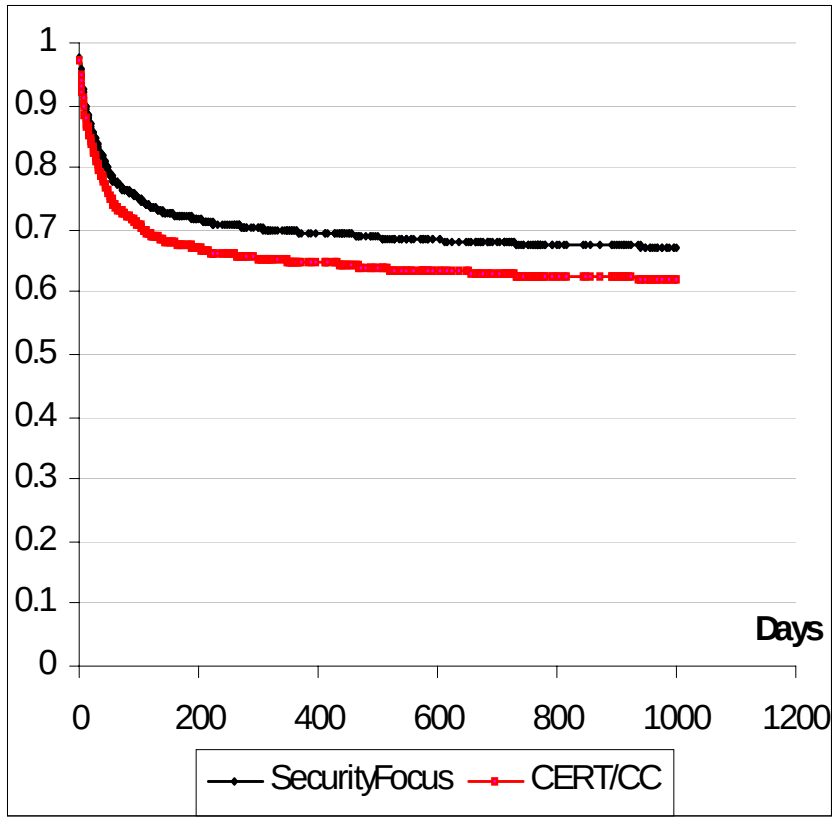
Patched only

Effect of disclosure policy: Duration model results:

Vulnerability type : Unix system server; close source, severity score =25, remote launch; possible availability loss



All observations



Patched Only

Results of parametric model estimation

Do they patch?

- Instantaneous disclosure policy increase the probability of patching by about 28.7% for average vulnerability in this sample
- Larger vendors more likely to patch.
- Severe vulnerabilities are more likely to be patched by the vendors.
- Open source vendors are more likely to patch

When do they patch?

- Instant disclosure policy is 1.33 days slightly slower than CERT/CC on average
- open source vendors are quicker to patch
- large vendors quicker to patch
- Sever vulnerabilities are patched quicker

Summary

- Disclosure increases attacks
- Patching decreases attacks
  - Patch for a known vulnerability decreases the number of attacks .
- Patching and disclosure increases attacks
  - even when a patch is available, disclosing a vulnerability increases the frequency of attacks.

- Instantaneous disclosure increases the probability of patching.
- Conditional on patching, time to patch is roughly same for both types of disclosure policy.
- Open source vendors are quicker to patch
- Large vendors are more responsive than small.
- Severe vuls more likley to be patched, and more likely to be patched quicker.

Backup: Time effect of patching and publishing

- We include no. of elapsed days from publishing a vulnerability and no. of elapsed days from release of patch, and quadratic terms for each.
- Estimate as Tobit

$$A_{it}^{*} = \alpha_i + \tau_t + \delta_1 nopatch_{it} + \delta_2 published_{it} + \delta_3 patched_{it} + \delta_4 (1 - nopatch_{it}) t_{patch}$$

$$+ \delta_5 [(1 - nopatch) t_{patch}]^2 + \delta_6 t_{pub} + \delta_7 t_{pub}^{\ 2} + u_{it} \equiv (\mathbf{X\delta} + u_{it})$$

$$A_{it} = \max(0, A_{it}^{*})$$

Backup -1

Difference in means of average number of attacks per host per day								
	Patched (1)	**Difference in (1) between periods (2)**	**Published (3)**	**Difference in (3) between periods (4)**	**Effect of patching = (2) – (4) (5)**	**Secret (6)**	**Difference in (6) between periods (7)**	**Effect of publishing Col.(7) – col.(2) (8)**
Period 1 (Nov 2002)	0.05 (0.002)	-	0.22 (0.11)	-	-	0.004 (0.001)	-	-
Period 2 (Jan 2003)	0.06 (0.001)	0.01 (0.003)	0.53 (0.03)	0.31 (0.14)	-0.30 (0.14)	0.06 (0.005)	0.056 (0.01)	0.05 (0.01)
Period 3 (Jan 2003)	0.06 (0.002)	0	2.29 (0.23)	1.76 (0.26)	-1.76 (0.26)	0.04 (0.004)	-0.02 (0.01)	-0.02 (0.01)
Period 4 (Jan 2003)	0.07 (0.004)	0.01 (0.01)	0.18 (0.17)	-2.11 (0.40)	2.10 (0.41)	0.10 (0.009)	0.06 (0.01)	0.05 (0.02)
Period 5 (Mar 2003)	0.11 (0.000)	0.04 (0.004)	0.32 (0.010)	0.14 (0.18)	-0.10 (0.18)	0.06 (0.004)	-0.04 (0.01)	0
Period 6 (May 2003)	0.13 (0.002)	0.02 (0.002)	0.37 (0.014)	0.05 (0.02)	-0.03 (0.02)	0.11 (0.017)	0.05 (0.02)	0.03 (0.02)
Period 7 (Sep 2003)	0.01 (0.000)	-0.12 (0.002)	0.01 (0.007)	-0.36 (0.02)	0.24 (0.02)	0	-0.11 (0.02)	0.01 (0.02)
Period 8 (Nov 2003)	0.01 (0.005)	0 (0.005)	0.02 (0.002)	0.01 (0.01)	-0.01 (0.01)	0	0	0
Period 9 (Dec 2003)	0.01 (0.004)	0	0.001 (0.001)	-0.02 (0.003)	0.02 (0.003)	0	0	0
Average effect					**0.02 (0.13)**			**0.02 (0.003)**

Backup-2

Impact of Patching and Publishing (OLS estimates)

Variables	Specification 1: vulnerability characteristics+	Specification 2: vulnerability fixed effects
Windows	0.15 (0.96)	-
UNIX	-0.04 (-0.26)	-
All	0.02 (0.13)	-
Linux	0.01 (0.05)	-
Secret	-0.25*** (-2.92)	-0.20* (1.67)
Published	0.51* (1.72)	-0.17** (1.90)
Location	-0.43 (-0.72)	-0.06 (0.67)
Time dummies (8)	Yes	Yes
N	2772	2772

*Notes: *** p < 0.01 * p < 0.10. +Estimates include security protection, confidentiality, integrity, availability, input_validation, boundary_condition, buffer_overflow, access_validation, exceptional_condition, config error, other_vuln. besides Windows, UNIX, all, Linux and others, which denote vulnerability specific effects.*

Backup-3

Table 6 Tobit regression – Effect of elapsed patch days and elapsed publish months Dependent variable – Average number of attacks per day per host (z-statistics in parentheses)	
Variable	**Coefficient Estimate**
Not Patched dummy	-35.81*** (-4.72)
Patched & Published dummy	-34.32*** (-5.06)
Published & Not Patched dummy	19.39***(2.56)
Elapsed patch months**(1-Not Patched)*	0.39 (0.80)
Elapsed publish months	0.10 (0.25)
Elapsed publish months squared	-0.002 (-0.41)
Elapsed patch months squared**(1-Not Patched)*	-0.002 (-0.32)
Time dummy variable included (8)	Yes
Vulnerability dummy variable (59)	Yes
Vulnerability technical characteristics included	No
No. of observations	2772
Log likelihood	-632.20
No of vulnerabilities	308
σ (Std deviation)	12.69