Disentangling Joint Responsibility for Web Security in Shared Hosting

Samaneh Tajalizadehkhoob, Tom Van Goethem, Maciej Korczynski,
Arman Noroozian, Rainer Böhme, Tyler Moore,
Wouter Joosen, Michel van Eeten

MENOG
April 24, 2018

Outline

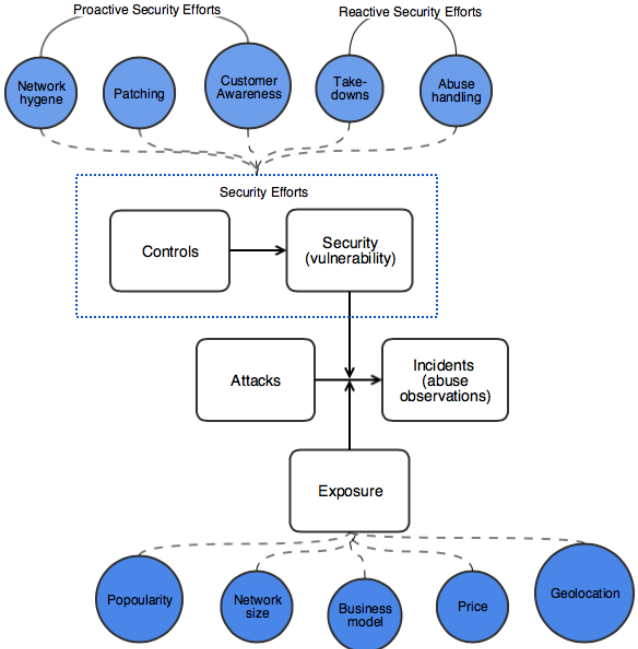
1. Motivation
2. Data collection methodology
3. Descriptive findings about the landscape
4. Impact of security efforts on abuse
5. Concluding remarks

Remedying web compromise

Providers are regularly faulted for not doing enough to combat different forms of compromise such as phishing, malware, botnet C&C.

But how much abuse can providers realistically prevent?

- Compromise rates are at least partially driven by factors outside providers' immediate control (i.e., the security practices of their clients)
- Concentration at providers may reflect attacker preferences as much as defender ineptitude
- It remains unclear how much the security efforts of hosting providers actually influence abuse prevalence

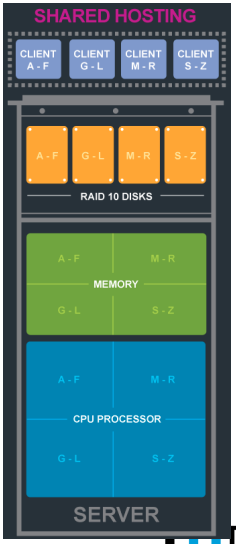

Proactive Security Efforts
Reactive Security Efforts
Network hygene
Patching
Customer Awareness
Take-downs
Abuse handling
Security Efforts
Controls
Security (vulnerability)
Attacks
Incidents (abuse observations)
Exposure
Popoularity
Network size
Business model
Price
Geolocation

Outline

1. Motivation
2. Data collection methodology
3. Descriptive findings about the landscape
4. Impact of security efforts on abuse
5. Concluding remarks

Shared hosting properties

- Physical server and server resources are shared among multiple customers
- Customers operate under restricted privileges
- Hosting providers maintain administrator privileges and can typically regulate what software is installed and whether it is updated



Collecting features from shared hosting providers

- Sampling domains from shared hosting providers
  - For each shared hosting provider, we randomly sampled 500 domain names
  - Our final set contains **442,684** domains distributed over **1,259** hosting providers
- Large-scale measurement of features
  - Distributed crawling infrastructure visited up to 20 pages per domain using headless browser PhantomJS
  - Used off-the-shelf tools to extract security features (e.g., Zonemaster, SSlyze, WhatWeb, WPscan)
  - **7,463,682** web pages were visited over between November 20-27, 2016

Feature list

Security features
`HttpOnly` cookie (+)
`X-Frame-Options` (+)
`X-Content-Type-Options` (+)
Mixed-content inclusions (−)
`Secure` cookie (+)
`Content-Security-Policy` (+)
HTTP `Strict-Transport-Security` (+)
SSL-stripping vulnerable form (−)
Weak browser `XSS` protection(−)

Software features (presence and version)
HTTP server
SSL
Admin panel
PHP
OpenSSH
CMS

TU Delft Delft University of Technology

Abuse data

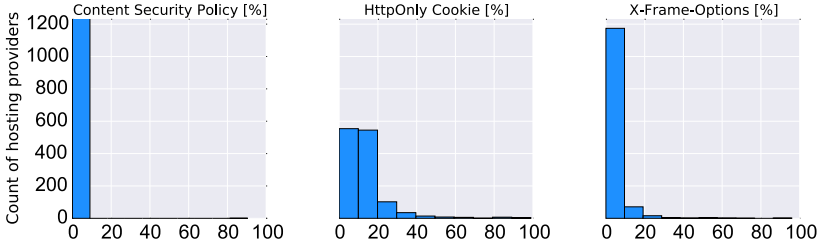
- Phishing data
  - Data collected from Anti-Phishing Working Group and PhishTank
  - 62K distinct domains for June-Dec 2016
  - 49K domains hosted by one of 968 shared providers
- Drive-by-download malware
  - Data from Google Safe Browsing as reported to StopBadware
  - 362K distinct domains for June-Dec 2016
  - 97K domains hosted by one of 1,050 shared providers

Outline

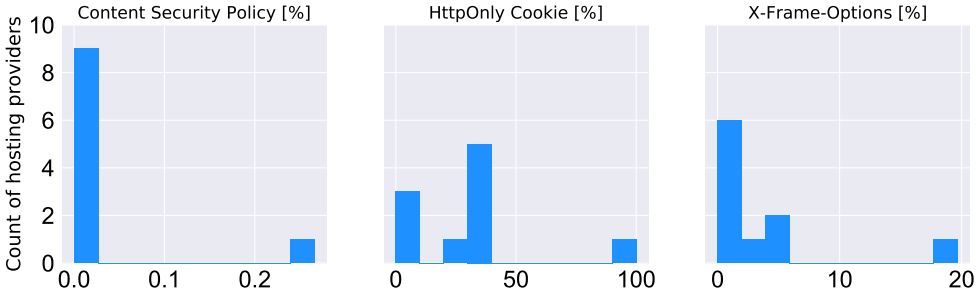
1. Motivation
2. Data collection methodology
3. Descriptive findings about the landscape
4. Impact of security efforts on abuse
5. Concluding remarks

Distribution of security features

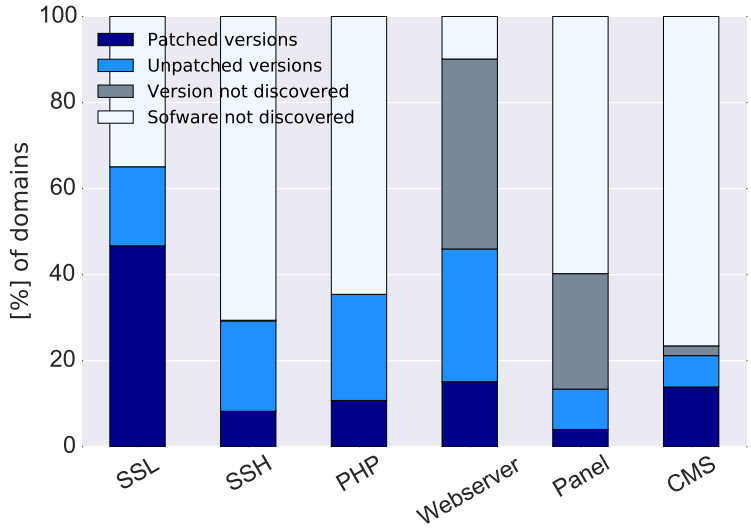
Feature	# of domains	% of domains
`HttpOnly` cookie (+)	57,696	13.04
`X-Frame-Options` (+)	22,212	5.02
`X-Content-Type-Options` (+)	8,685	1.96
Mixed-content inclusions (−)	2,107	0.47
`Secure` cookie (+)	1,378	0.31
`Content-Security-Policy` (+)	894	0.20
HTTP `Strict-Transport-Security` (+)	847	0.19
SSL-stripping vulnerable form (−)	515	0.11
Weak browser `XSS` protection (−)	376	0.08



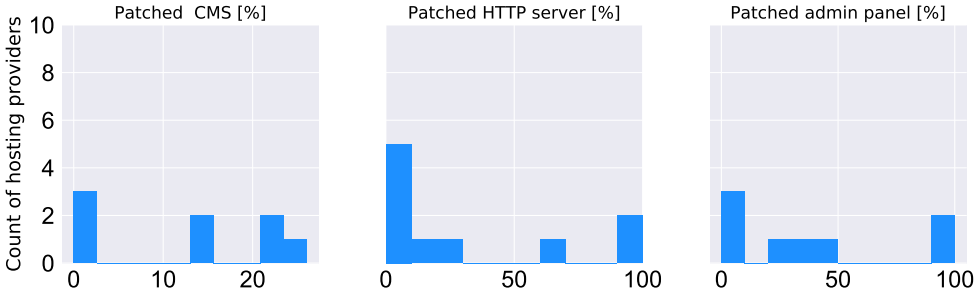
The case of IRAN



Patching practices across domains



The case of IRAN



Factor analysis: Which features correlate?

	MR1	MR2	MR3	MR4
`X-Content-Type-Options`	**0.87**	0.11	0.14	-0.01
`Content-Security-Policy`	**0.80**	0.23	-0.01	0.37
`X-Frame-Options`	**0.83**	0.09	0.10	-0.16
HTTP `Strict-Transport-Security`	**0.61**	0.50	0.04	0.03
Mixed-content inclusions	0.26	**0.76**	-0.01	-0.24
Weak browser `XSS` protection	-0.39	**0.68**	0.24	0.29
SSL-stripping vulnerable form	0.08	**0.60**	-0.05	-0.38
`HttpOnly` cookie	0.13	**0.65**	0.14	0.12
`Secure` cookie	0.36	**0.86**	0.03	0.11
Patched HTTP*	0.09	0.05	**0.74**	-0.11
Secure SSL implementation*	-0.15	-0.09	**0.74**	-0.10
Patched SSH*	-0.07	0.04	**0.42**	0.35
Patched PHP*	0.09	-0.12	0.13	**0.55**
Patched CMS*	-0.14	0.01	-0.23	**0.78**
Patched Admin panel*	0.08	0.08	0.10	**0.58**
Loadings' sum of squares	2.90	2.92	1.48	1.90
Proportion of variance explained	0.19	0.19	0.10	0.13
Cumulative variance explained	0.19	0.39	0.49	0.62

* Scale from least to most secure: 0 older versions, 1 latest or no version, 2 no software

- MR1: Content security practices
- MR2: Webmaster security practices
- MR3: Infrastructure security practices
- MR4: Web application security practices

Whose security effort: hosting providers or webmasters?

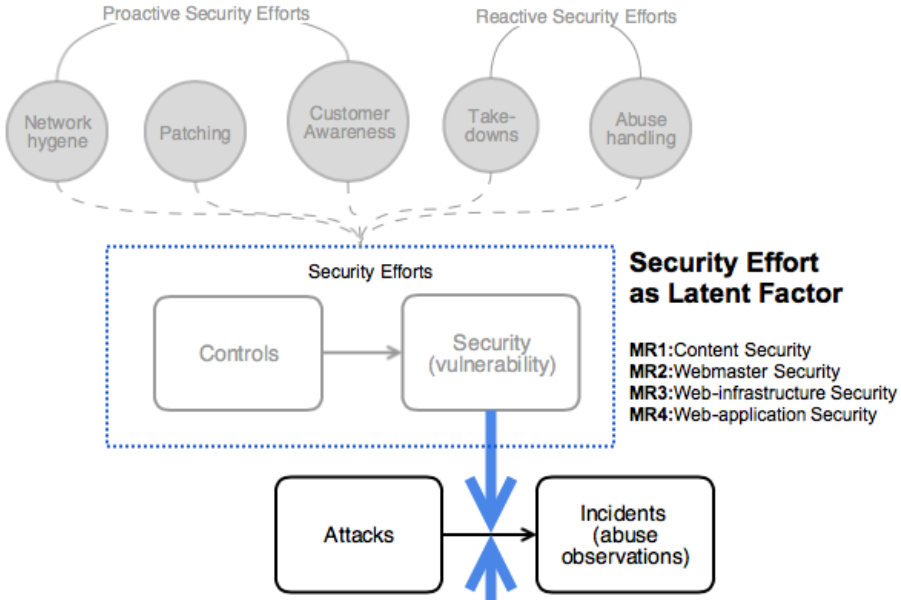
	Response Variable: Security Factor(s)			
	MR1	MR2	MR3	MR4
	(1)	(2)	(3)	(4)
Hosting provider fixed-effect	yes	yes	yes	yes
Constant	-0.250^{***}	-0.300^{***}	0.100^{*}	0.420^{***}
	(0.064)	(0.066)	(0.043)	(0.051)
Observations	442,075	442,075	442,075	442,075
R^2	0.077	0.066	0.270	0.200
Adjusted R^2	0.075	0.064	0.270	0.200
Residual Std. Error (df = 440801)	1.400	1.400	0.920	1.100

*p<0.05; **p<0.01; ***p<0.001
Standard errors in brackets

Outline

1. Motivation
2. Data collection methodology
3. Descriptive findings about the landscape
4. Impact of security efforts on abuse
5. Concluding remarks

Model fitting



Impact of security efforts on abuse: Phishing

- **Size** of shared hosting infrastructure explains majority of abuse concentrations in the network of providers;
- **Content Security**, **Webmaster security** and **Web application security** show a strong significant relation with abuse concentrations;
- This implies that after size, **strong** security regarding content, webmaster features and application **reduces** website abuse the most;

Outline

1. Motivation
2. Data collection methodology
3. Descriptive findings about the landscape
4. Impact of security efforts on abuse
5. Concluding remarks

Conclusions

- Most security features are sparsely implemented by webmasters or providers
- Higher levels of the web stack (CMS, admin panel) are more up-to-date than infrastructure software (SSH, PHP)
- We showed via Indirect measurement of security effort that shared hosting providers influence web-application security and infrastructure security practices
- Both webmasters and providers can inhibit malware and phishing abuse
- Shared hosting providers exert influence high up the web stack, where applications such as CMSes are mostly managed by clients

Practical implications

- What affects the security outcome of providers are the nature of their business (Network size, service type (e.g. shared vs dedicated hosting))
- After that, proactive security efforts can reduce abuse in providers networks (patching, secure configurations)
- Customer level efforts are as important as provider level efforts
- Therefore, providers should put more force on improving client side security, by indirect measures, using the power of defaults, notifications, etc.

Future work

- Collect additional discriminating features to explain more variance
- Model effects at individual level, rather than aggregated by provider
- Apply method to other areas of joint responsibility for security, such as between cloud hosting providers and tentants, or corporate system administrators and end users

Thanks for your attention

Questions?

Dr. Samaneh Tajalizadehkhoob
s.t.tajalizadehkhoob@tudelft.nl

Acknowledgments

We thank

- Anti Phishing Working Group (APWG)
- PhishTank
- StopBadware
- Farsight Security

This work was supported by NWO (grant nr. 12.003/628.001.003), the Dutch National Cyber Security Center (NCSC) and SIDN, the .NL Registry, and Archimedes Privatstiftung, Innsbruck.