S

Security Reputation Metrics

Maciej Korczyński[1] and Arman Noroozian[2]
[1]University Grenoble Alpes, CNRS, Grenoble INP, LIG, Grenoble, France
[2]Delft University of Technology, Delft, Netherlands

Definitions

Security reputation metrics (*aka.* security metrics) quantify the security levels of organization (e.g., hosting or Internet access providers) relative to comparable entities. They enable benchmarking and are essential tools for decision- and policy-making in security and may be used to govern and steer responsible parties toward investing in security when economic or other decision-making factors may drive them to do otherwise.

Background

We increasingly interact with online digital content, which relies on services provided by the so-called *Internet intermediaries*, among them Internet service providers (ISPs), hosting providers, domain name registrars, search engines, payment providers, certification authorities, cloud service providers, social network operators, and e-commerce suppliers. Simultaneously, miscreants have been abusing the products and infrastructure of such service providers toward cybercrime by compromising their security and using them in unintended ways.

A wealth of research into cybercrime points to how cybercriminals misuse hosting services (Nikiforakis et al. 2011; Noroozian et al. 2019), domain names (Felegyhazi et al. 2010; Liu et al. 2011; Hao et al. 2011, 2013; Szurdi et al. 2014; Le Pochat et al. 2020), DNS services (Bilge et al. 2011; Canali et al. 2011), and mail servers (Stone-Gross et al. 2011; Levchenko et al. 2011) to name a few examples.

A typical situation with intermediary services is that their consumers are at an inherently disadvantaged knowledge position in which they do not know much about how secure these services are. In contrast the service providers themselves know much more. Without additional security-related information, users typically base their decision to subscribe to a particular service on more readily available information, for instance, pricing in case of hosting or available bandwidth in case of broadband ISP. In other words, it is difficult for other businesses, consumers, and regulators to reliably assess how secure intermediary services are. This so-called *information asymmetry* about the security of intermediary services, combined with the fact that, typically, parties other than the intermediaries themselves bear the cost of the cybercrime enabled through their services (Anderson et al. 2012), leads to an erosion of their incentives to adequately invest in security.

S. Jajodia et al. (eds.), *Encyclopedia of Cryptography, Security and Privacy*,
https://doi.org/10.1007/978-3-642-27739-9_1625-1

In other words their economic incentives are misaligned with security goals.

Thus, systematically comparing the security performance of digital services, and intermediaries, may help reduce the security information asymmetries from a consumer perspective. Security reputation metrics are essential to this end and may help in reducing cybercrime, which is as much a technical issue as a problem of economic incentives (Anderson 2001). Various stakeholders such as network service providers (Asghari et al. 2015b), domain registries (Korczyński et al. 2017a), law enforcement agencies (Noroozian et al. 2015), and even policy-makers (Korczyński et al. 2017b) employ security metrics to answer questions like which are the worst service providers and what actions should be taken to steer market-driven economies toward improved security outcomes.

Existing metrics typically compare security based on either (i) how frequently abuse incidents occur (or vulnerabilities are discovered), i.e., are based on counting the number of incidents (vulnerabilities), or (ii) how timely incidents are remediated once they have occurred. The number of maliciously registered domain names, compromised end-user machines, and machines running outdated software per service provider are examples of the former case. The amount of time required to remediate, block, and remove phishing or malware spreading webpages is an example of the second metric type.

Count-based metrics are typically normalized by estimates of the size of each intermediary's potential *attack surface* to control for more exposed intermediaries that have higher probability of experiencing incidents. This enables apples-to-apples comparisons between intermediaries of various exposures. The number of advertised IP addresses by a hosting provider or the number of domains that it hosts, for instance, may be used to estimate a hosting provider's potential attack surface (Noroozian et al. 2015; Tajalizadehkhoob et al. 2016, 2018).

Application

The following subsections present examples of security reputation metrics for different types of providers and show how they are used by Internet stakeholders in reducing cybercrime and aligning economic incentives with better security.

Hosting Providers

Hosting providers are companies that provide servers via which customers can make content or services available on the Internet, e.g., websites, email, or even sharing of files. As with all services on the Internet, they are also abused for criminal purposes. Think of phishing sites, command-and-control servers for botnets, child sexual abuse material (CSAM), malware distribution, and spam servers.

In theory, hosting providers can mitigate or prevent the abuse of their infrastructure by following security best practices set forth by organizations like the Messaging, Malware and Mobile Anti-Abuse Working Group (M3AAWG), for example, by vetting customers, monitoring their infrastructure for signs of compromise, or even running antivirus software to name a few recommended practices. Yet following such advice remains voluntary, and there is considerable variation in how hosting providers choose to act when it comes to abuse and securing their infrastructure. Naturally then, some providers are abused more often than others.

Therefore, the question of which hosting providers are better at securing their infrastructure is one that may be answered through security metrics that quantify how effective each provider is at curbing abuse.

A systematic approach to metrics development within the hosting market is presented in a study by Noroozian et al. (2015), which is scoped to the Netherlands and the result of a collaboration between several local authorities. The aim of the study being to answer the following question: which are the worst hosting providers in Dutch jurisdiction? Metrics developed within this study were used to steer the Dutch hosting market

toward more effective security practices via involving several local authorities (Noroozian et al. 2015). Furthermore, Noroozian et al. (2017) have developed metrics for comparing the security of hosting providers globally (Noroozian et al. 2017), which are shown to have considerable predictive power. Other work discusses how multiple approaches to curbing abuse within the hosting market, from loose market-based approaches to stricter regulation-based approaches, may benefit from the employment of security metrics (Fryer et al. 2015; Noroozian 2020).

Top-Level Domains

There exists little empirical information about the security of entire top-level domains such as .com, .nl, or .top. Korczyński et al. (2017a) were first to present security metrics for this ecosystem and have measured their operational values. They compared entire TLDs against the rest of the market. However, they have explicitly distinguished those metrics from the objective of measuring the security performance of the registry operators. The reason is that a TLD is not a single organization but constitutes an entire "domain name ecosystem" of different types of intermediaries such as domain registries, registrars, or hosting providers.

The follow-up study requested by ICANN, spanning the period up to the end of 2016, investigated the following research question: how do abuse rates in the new gTLDs (e.g., .top, .science, .bank, or .site) compare to legacy gTLDs (e.g., .com, .net, .edu, or .org), since the introduction of the new gTLD program in 2013? To determine the distribution of abusive activities across the gTLDs, Korczyński et al. (2018) have analyzed the number of reported domains from reputed URL and domain blacklists normalized by the size of their respective TLD, calculated as the number of 2nd-level domain names present in a zone file for each gTLD.

They have compared abuse rates separately for compromised and maliciously registered domains (Maroofi et al. 2020). Reputation metrics reflecting spam activity in the new and legacy gTLDs have revealed an interesting trend: miscreants seem to be switching from abusing legacy to new gTLDs when it comes to maliciously registered spam domains. In the last quarter of 2016, new gTLDs collectively had approximately one order of magnitude higher rate of spam domains per 10,000 registrations compared to legacy gTLDs. Moreover, as many as 15 most abused new gTLDs had more than 10% of all registered domain names blacklisted by Spamhaus at the end of 2016. Finally, as many as 51.5%, 47.6%, and 33.4% of all .science, .stream, and .study new gTLDs, respectively, were maliciously registered by cybercriminals and blacklisted by Spamhaus.

ICANN has used the calculated reputation metrics to review the existing anti-abuse safeguards in new gTLDs and to introduce more effective ones before an upcoming new gTLD rollout.

Internet Service Providers

A significant amount of scientific work addresses the role of Internet service providers (ISPs)—network access providers—in mitigating cybercrime, for instance, through botnet mitigation by employing security metrics (van Eeten et al. 2010, 2016; Asghari et al. 2015a). ISPs typically provide Internet connectivity to customers and thus are in a unique position to mitigate certain forms of cybercrime at their origin.

ISPs should also follow security best practices to mitigate abuse. Examples of some security best practices for ISPs include the use of walled gardens to quarantine and isolate infected machines connected to the Internet (Çetin et al. 2018, 2019) or deploying source address validation, also known as BCP38, to prevent distributed denial-of-service (DDoS) attacks from being launched via their infrastructure (Luckie et al. 2019; Korczyński et al. 2020; Lone et al. 2017) Yet, again, the voluntary nature of implementing such best practices results in certain ISPs experiencing a higher level of abuse than others due to having laxer security practices.

Differences among ISPs in mitigating botnet infections, for instance, have been quantified in several studies. van Eeten et al. (2010), for example, found that just 50 ISPs account for over half

of all spam sources suggesting concentrations of spambots within a few ISPs worldwide.

A typical approach in such studies is to first process global or national datasets of botnet activity from available sinkholes and to extract IP addresses of infected end-user machines. The methods employed typically map each bot-infected IP address to an ISP and then counts the IP addresses seen in each ISP per day to account for IP churn (Moura et al. 2015). Security metrics to compare botnet mitigation among ISPs are then calculated by dividing this count by the numbers of subscribers of each ISP, where a larger number indicate less effective mitigation by the ISP. Such security metrics for botnet mitigation among ISPs have been employed successfully to incentivize ISPs, within the Netherlands, for example, to subscribe to threat intelligence data feeds and deal with bot infections within their networks (van Eeten et al. 2016).

Open Problems and Future Directions

Empirical measurements and analysis of security indicators leading to reliable security reputation metrics has proven to be quite challenging. The challenge partly lies in limitations of data: e.g., coverage, measurement errors, and biases that are invariably linked to metric limitations. For example, the construction of security metrics typically depends on chaotic Internet operations data, such as WHOIS information that are error prone and incomplete, or on ever-evolving dynamic BGP routing data for attributing security incidents to the responsible entities.

The security incidents themselves are observed through various opaque abuse feeds with no clear documentation of their collection methodologies, accuracy, and biases. Abuse feeds contain various degrees of false-positive incident information or carry biases that are not well documented or understood. Abuse feed coverage is also limited with an unknown number of security incidents that goes unnoticed as false negatives for each entity, thereby affecting metric outcomes.

In addition to limitations in data, methodological challenges in constructing metrics also exist. A particularly challenging methodological aspect is that of identifying service providers within certain markets. For example, in the case of hosting providers, there is no maintained authoritative list of companies, even at a country level, to identify the companies that offer hosting services. This problem is worsened by layers of smaller companies that resell the services of larger hosting providers. Similar problems exist for domain registrars. Yet, such information is vital for the construction and comparison of service providers against their competitors at a market level to make security metrics more useful.

Another methodological challenge is an incomplete causal understanding of the factors that drive abuse across online services. A better and more complete causal understanding of such drivers enriches and allows for the construction of security metrics that are better interpretable, easier to understand, and more useful.

Finally, there are practical limitations as well including the fact that metrics do not reflect the intent of bad service providers whether it be negligent or criminal behavior, for instance, in the case of bullet-proof services that cater to cybercriminals with the promise of ignoring or delaying lawful take down requests. As such, security metrics can be gamed by coordinated criminals, and thus there are limits to how they may be interpreted.

References

Anderson R (2001) Why information security is hard – an economic perspective. In: ACSAC

Anderson R, Barton C, Rainer B, Clayton R, van Eeten M, Levi M, Moore T, Savage S (2012) Measuring the cost of cybercrime. In: WEIS

Asghari H, Ciere M, Van Eeten MJG (2015a) Post-mortem of a zombie: conficker cleanup after six years. In: USENIX Security

Asghari H, van Eeten MJ, Bauer JM (2015b) Economics of fighting botnets: lessons from a decade of mitigation. IEEE S&P 13(5):16–23

Bilge L, Kirda E, Kruegel C, Balduzzi M (2011) EXPOSURE: finding malicious domains using passive DNS analysis. In: NDSS

Canali D, Cova M, Vigna G, Kruegel C (2011) Prophiler: a fast filter for the large-scale detection of malicious web pages. In: WWW

Çetin O, Gañán C, Altena L, Tajalizadehkhoob S, van Eeten M (2018) Let me out! evaluating the effectiveness of quarantining compromised users in walled gardens. In: USENIX SOUPS

Çetin O, Gañán C, Altena L, Kasama T, Inoue D, Tamiya K, Tie Y, Yoshioka K, van Eeten M (2019) Cleaning up the internet of evil things: real-world evidence on ISP and consumer efforts to remove mirai. In: NDSS

Felegyhazi M, Kreibich C, Paxson V (2010) On the potential of proactive domain blacklisting. In: USENIX LEET

Fryer H, Stalla-Bourdillon S, Chown T (2015) Malicious web pages: what if hosting providers could actually do something. Comput Law Secur Rev 31(4):490–505

Hao S, Tech G, Feamster N, Tech G (2011) Monitoring the initial DNS behavior of malicious domains. In: ACM IMC

Hao S, Thomas M, Paxson V, Feamster N, Kreibich C, Grier C, Hollenbeck S (2013) Understanding the domain registration behavior of spammers. In: ACM IMC

Korczyński M, Tajalizadehkhoob S, Noroozian A, Wullink M, Hesselman C, van Eeten M (2017a) Reputation metrics design to improve intermediary incentives for security of TLDs. In: IEEE Euro S&P

Korczyński M, Wullink M, Tajalizadehkhoob S, Moura GC, Hesselman C (2017b) Statistical analysis of DNS abuse in gTLDs final report. Technical report

Korczyński M, Wullink M, Tajalizadehkhoob S, Moura GC, Noroozian A, Bagley D, Hesselman C (2018) Cybercrime after the sunrise: a statistical analysis of DNS abuse in new gTLDs. In: ACM ASIACCS

Korczyński M, Nosyk Y, Lone Q, Skwarek M, Jonglez B, Duda A (2020) The closed resolver project: measuring the deployment of source address validation of inbound traffic. In: CoRR

Le Pochat V, Van hamme T, Maroofi S, van Goethem T, Preuveneers D, Duda A, Joosen W, Korczyński M (2020) A practical approach for taking down avalanche botnets under real-world constraints. In: NDSS

Levchenko K, Pitsillidis A, Chachra N, Enright B, Felegyhazi M, Grier C, Halvorson T, Kanich C, Kreibich C, Liu H, McCoy D, Weaver N, Paxson V, Voelker GM, Savage S (2011) Click trajectories: end-to-end analysis of the spam value chain. In: IEEE S&P

Liu H, Levchenko K, Félegyházi M, Kreibich C, Maier G, Voelker GM, Savage S (2011) On the effects of registrar level intervention. In: USENIX LEET

Lone Q, Luckie M, Korczyński M, van Eeten M (2017) Using loops observed in traceroute to infer the ability to spoof. In: PAM

Luckie MJ, Beverly R, Koga R, Keys K, Kroll JA, Claffy KC (2019) Network hygiene, incentives, and regulation: deployment of source address validation in the internet. In: ACM CCS

Maroofi S, Korczyński M, Hesselman C, Ampeau B, Duda A (2020) COMAR: classification of compromised versus maliciously registered domains. In: IEEE Euro S&P

Moura GCM, Gañán C, Lone Q, Poursaied P, Asghari H, van Eeten M (2015) How dynamic is the ISPs address space? Towards internet-wide DHCP churn estimation. In: Networking

Nikiforakis N, Joosen W, Johns M (2011) Abusing locality in shared web hosting. In: EUROSEC

Noroozian A (2020) Evaluating hosting provider security – through abuse data and the creation of metrics. Doctorol thesis, TU Delft, Faculty of Technology, Policy and Management

Noroozian A, Korczyński M, Tajalizadehkhoob S, van Eeten M (2015) Developing security reputation metrics for hosting providers. In: USENIX CSET

Noroozian A, Ciere M, Korczyński M, Tajalizadehkhoob S, Eeten MV (2017) Inferring the security performance of providers from noisy and heterogenous abuse datasets. In: WEIS

Noroozian A, Koenders J, van Veldhuizen E, Gañán C, Alrwais S, McCoy D, van Eeten M (2019) Platforms in everything: analyzing ground-truth data on the anatomy and economics of bullet proof hosting. In: USENIX Security Symposium

Stone-Gross B, Holz T, Stringhini G, Vigna G (2011) The underground economy of spam: a Botmaster's perspective of coordinating large-scale spam campaigns. In: USENIX LEET

Szurdi J, Kocso B, Cseh G, Felegyhazi M, Kanich C (2014) The long taile of typosquatting domain names. In: USENIX Security Symposium

Tajalizadehkhoob S, Korczyński M, Noroozian A, Gañán C, van Eeten M (2016) Apples, oranges and hosting providers: heterogeneity and security in the hosting market. In: IEEE/IFIP NOMS

Tajalizadehkhoob S, Böhme R, Gañán C, Korczyński M, van Eeten M (2018) Rotten apples or bad harvest? What we are measuring when we are measuring abuse. ACM TOIT 18(4):1–25

van Eeten M, Bauer JM, Asghari H, Tabatabaie S, Rand D (2010) The role of internet service providers in botnet mitigation: an empirical analysis based on spam data. In: WEIS

van Eeten M, Lone Q, Moura G, Asghari H, Korczyński M (2016) Evaluating the impact of AbuseHUB on botnet mitigation. In: CoRR