Underground Economics for Vulnerability Risk

LUCA ALLODI

Luca Allodi is an Assistant Professor at the Technical University of Eindhoven, the Netherlands. His research interests lie around the quantitative characterization of IT risk and attacker decisions. l.allodi@tue.nl

The estimation of vulnerability risk is at the core of any IT security management strategy. Among technical and infrastructural metrics of risk, attacker economics represent an emerging new aspect that several risk assessment methodologies propose to consider (e.g., based on game theory). Yet the factors over which attackers make their (economic) decisions remain unclear and, importantly, unquantified. To address this, I infiltrated a prominent Russian cybercrime market where the most prominent attack technology is traded. Supported by direct observations of market activity, I investigate in this work the economic factors that drive the adoption of new attacks *at scale* and their effect on risk of attack in the wild. As a market participant, I have access to the full spectrum of attack services offered to all members and, in particular, look at the market economics of vulnerability exploitation [1].

Software vulnerabilities are one of the main vectors of attack used to infect systems worldwide. As such, an effective management of vulnerability fixes is desirable on any system. Unfortunately, due to technical and budgeting restrictions, applying *all* fixes as soon as they are available is oftentimes not possible. For this reason, prioritizing patching work is a key aspect of any vulnerability management policy. The goal is clear: identify which vulnerabilities carry the highest risk and need immediate treatment.

Several methodologies to estimate this “potential risk” of vulnerability exploit exist, including technical measures of vulnerability severity (e.g., the Common Vulnerability Scoring System, CVSS), attack graphs, attack surfaces, and game-theoretic approaches that, for example, assign probabilities to specific attacker strategies in response to a certain set of defender decisions.

Importantly, and across all current approaches, the probability assigned to the materialization of an exploit mainly depends on vulnerability characteristics or specific “contextual” aspects such as network topology, deployed security controls, and vulnerability chaining. This, in turn, implicitly assumes that, all other factors being the same, attackers will be indifferent to which vulnerability to exploit.

An implication of this model is that all “high severity” vulnerabilities on a certain system or software will be equally likely to be exploited. Oftentimes, due to the high prevalence of severe vulnerabilities, exploit estimations will not be dramatically different across systems and vulnerabilities. This ultimately leads to inefficient vulnerability patching strategies [4], as most vulnerabilities are “indistinguishable” in terms of posed risk, and therefore all need immediate treatment.

All Vulnerabilities Are Not Equally Important

On the other hand, recent research developments reveal that the vast majority of attacks seem to be driven by a handful of vulnerabilities only. In [2], across most software types, the top 10% of vulnerabilities are reported to carry 90% of attacks across 1M Internet users worldwide, approximating a power law distribution. Other research has shown that this huge skew in attack distribution is present also for zero-day vulnerabilities. In this analysis [6], of 20 zero-day vulnerabilities, two were reportedly responsible for millions of attacks worldwide, one for twenty thousand, and the remaining 18 for a few dozen only. These results are confirmed in follow-up empirical studies estimating that approximately 15% of disclosed vulnerabilities are exploited in the wild, and that this fraction is decreasing for recent vulnerabilities [10]. Similarly, recent work showed that the *refresh time* of exploits is very slow, with exploits being actively deployed in the wild up to two or three years before being substituted at scale by a different exploit [5].

These observations are in sharp contrast with the current narrative in the information security community, where every new severe vulnerability loosely resembles *Doomsday*. Industry studies recently started to acknowledge this effect as well: for example, in the last few editions of Verizon's Data Breach Investigations Report. Overall, empirical data clearly shows that a handful of vulnerabilities carry disproportionately more risk (by several orders of magnitude) than most vulnerabilities. It seems therefore that factors other than the characteristics of the vulnerability should be considered to explain this phenomenon.

Vulnerability Risk and Attacker Types

It is important to clarify the nature of the data leading to the observations above and its relation to different attacker types. In general, field data concerns attacks of an "untargeted" nature, where attackers in possession of a "fixed" set of exploits deliver attacks in the wild against the population of Internet users as a whole. These attacks are the most common and involve high attack automation, exploitation as-a-service [8], and delivery infrastructures based on spam or redirection of Internet traffic.

Attacks of a more "targeted" nature are radically different from the previous scenario: in such cases attackers adapt their exploit portfolio to the desired target system (as opposed to relying on a fixed set of exploits). Targeted attacks affect a very limited set of Internet systems and entail high levels of variability as attackers are (un)bounded by resource constraints, technical capabilities, and access rights to the network. Hence, in the case of targeted attacks, assigning probabilities to compute risk levels may not be a meaningful approach [7] as the notion itself of *probabilistic* risk does not apply anymore. In this article, I specifically refer to *risk of untargeted attacks at scale*.

A Dive into Exploit Economics

This distinction between "untargeted" and "targeted" attacks has become more and more relevant with the establishment of an underground economy driving the commodification of attacks at scale [8]. By outsourcing the complexity of attack engineering to the technically proficient sections of the underground, the technical difficulty of engineering and deploying an attack significantly decreased for those who participate in this economy. The acquisition of "off the shelf" attack tools represents a "multiplier factor" whereby a single attack technology (e.g., malware or vulnerability exploit) is shared among a multitude of attackers.

For example, exploit kits are known to be responsible for a significant share of the overall attack scenario by providing a ready-to-use, easy-to-configure attack framework that covers all steps of the attack process, from selection and redirection of vulnerable traffic, to vulnerability exploitation and malware delivery. Hence, buyers of these attack technologies may, potentially, jointly deliver a large fraction of attacks in the wild by sharing the same attack vectors and infrastructure.

I propose that the adoption of attack techniques traded in the cybercrime markets may explain the disproportionate concentration of attacks over a small set of vulnerabilities discussed above. Hence, under this hypothesis, it becomes central to understand the relation between deployment of an attack at scale and attackers' economic activities [1]. For example, pricier exploits may be adopted less widely by attackers, and vulnerabilities that are seldom substituted in the markets may remain exploited at scale for longer periods of time.

Market Identification and Infiltration

One of the difficulties associated with studying the underground economy is to identify active, well-functioning underground markets where prominent attack tools are traded. The underground economy is indeed fragmented in a multitude of markets, both in the so-called "deep web" as "onion services" and in the "open Internet." Whereas finding these markets is not a challenge per se, finding *credible* markets is: one should expect most markets to be places where gullible "wanna-be" criminals get scammed and no real technological innovation happens; Herley and Florencio provide an excellent coverage of the foundational economic reasons why this is the case [9].

Following Herley and Florencio's guidelines, and jointly with Professor Fabio Massacci at the University of Trento (Italy) and Professor Julian Williams at the Durham Business School (UK), I started evaluating different underground markets in the English and Russian hacking communities in 2011. One (Russian) community, above all, emerged as a prominent market where we find convincing evidence of severe trade regulation enforcement, credible trade activities, and the most prominent attack tools reported by the security industry, including exploit kits such as *RIG* and *Blackhole*, malware platforms, malware packers, and so on. We refer to this market under the fictitious name of RuMarket. All other markets in our analysis have been discarded for not meeting at least one of these criteria; [3] reports an example comparison.

We first gained access to RuMarket in 2011 and carried out "under-the-radar" observations of the activity therein, without performing any interaction with the market members. At the time, access to the market was only as difficult as registering to the corresponding forum platform under a fictitious identity.

This changed rather abruptly in 2013 when a prominent member of the market was arrested by the Russian authorities. The market reacted by ejecting all non-active participants and by significantly increasing the entry barrier to the market. Uncontrolled access to the market was replaced by a more strict process supervised by the market administration whereby access was granted only if either:

1. A trusted member of the market vouched for the entry request, effectively implementing a *pull-in* mechanism.
2. The request for market entry was backed up by evidence that the requestor was a reputable member of the Russian hacking community.

As we had no contacts inside the market to regain access, we chose to follow (2). This required extensive research to identify communities affiliated with RuMarket with more loose access barriers and build our identity from there. This, in turn, called for some proficiency in Russian in the discussion boards but did not involve the execution or support of criminal activities.

We gained new access to RuMarket in 2014 after more than six months of activity in the affiliated communities. We have been observing the market ever since. In this article, we look at the economics of vulnerability exploit trading [1].

Market Activity and Exploit Packages

In RuMarket, vulnerability exploits are traded in *packages*, or bundles. These can be classified using three categories: EKIT (exploit kit), Malware, and Standalone exploits. Figure 1 reports on the introduction of new exploit packages per year. Standalone packages are clearly on the rise, whereas Malware and EKIT packages are introduced or updated at a steady rate each year. This difference can be explained by looking at the different business models behind the bundles: Malware and EKIT are typically service-oriented products that require a prolonged contractual agreement between the buyer and the seller and are very popular in the market (in particular, the average EKIT advertisement receives approximately 10 times more replies from the community than the average Standalone or Malware package). As such, vendors tend to regularly update their products (e.g., with new or more reliable exploits) as opposed to substituting the whole package with a new one. This creates a perhaps slightly counterintuitive effect in which only a few players sell EKITs (despite these being very attractive products in the market): the prolonged contractual form requires high levels of trust between market participants, a condition only well-established vendors can meet, and hence the low rate of new kits each year. As most malware in RuMarket is not advertised to exploit any specific vulnerability, Malware products have low introduction rates in Figure 1.

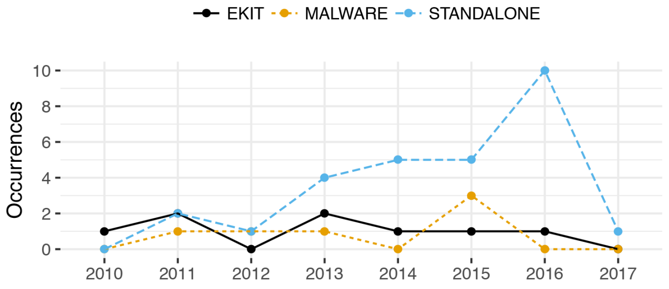

Figure 1: Release of exploit packages by type per year

Type	No.	Min	Mean	Median	Max
EKIT	6	150	693.89	400	2000
Malware	6	420	1735	1250	4000
Standalone	26	100	2972.69	3000	8000
ALL	38	100	2417.46	1500	8000

Table 1: Package prices (in USD)

Table 1 reports descriptive statistics of package prices. Prices for rented EKITs are averaged over a period of three weeks, following the duration of typical malware delivery campaigns. We can observe that EKIT products are by far the cheapest, with a mean price of 700 USD, whereas Malware and Standalone products are significantly more expensive at 2000–3000 USD on average. This difference is stressed at the right-end tail of the distributions, where Standalone packages peak at 8000 USD, Malware at 4000, whereas EKITs stop at 2000 USD. Prices do not show a significant correlation with the number of embedded exploits, suggesting that other aspects, such as the business model behind the trade, or the age of the embedded exploits, may play a factor. An evaluation of the trend in pricing for each package type reveals that prices are clearly inflating for Standalone and Malware products, whereas EKIT prices are decreasing over time. This reflects the "consumer" nature of EKIT products, which are becoming more and more available to a larger pool of buyers, whereas the prices for Standalone exploits reflect a "niche" part of the market and are inflating.

Vulnerability Exploits

With the aim of evaluating the effect of exploit economics on vulnerability risk, it is useful to look at a breakdown of exploits bundled in a package, as opposed to the bundle "as a whole." Figure 2 reports the rate of introduction of single exploits in the market aggregated by vendor of the vulnerable software. Unsur-

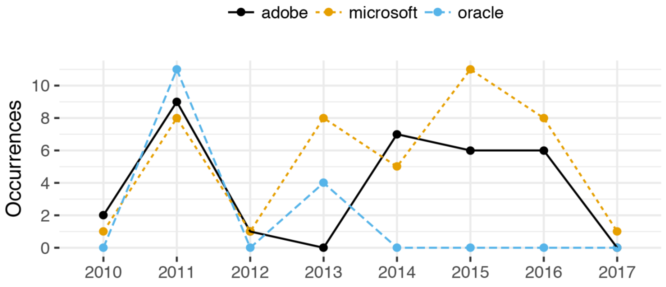

Figure 2: Occurrences of exploit publication by year

Type	No.	Min	Mean	Median	Max
EKIT	25	1	372.48	294	1745
Malware	1	185	185	185	185
Standalone	29	1	147.34	75	934
ALL	55	1	250.36	93	1745

Table 2: Exploit age (days) at time of first appearance in RuMarket

prisingly, in RuMarket we find exploits for Microsoft, Oracle, and Adobe software, which can be expected to cover the vast majority of user systems in the wild. The first observation we make is that the first "burst" of exploits appears in 2011, which corresponds to the appearance of "exploitation-as-a-service" as a new attack model [8]. After 2011 the market experienced a relative drop in number of introduced exploits to then stabilize around an average level of 6–8 new exploits per software vendor per year. This trend loosely resembles the *Gartner Hype Cycle* describing the introduction of new technologies in a market: a first inflation in the expectations associated with that technology causes a burst of interest in the market, followed by a "disillusionment" phase and, finally, by what Gartner calls the *plateau of productivity*, where the technology reaches maturity and its true value.

Table 2 reports the age, in days, of the exploits first introduced in RuMarket relative to the date of their publication in the National Vulnerability Database (NVD). As all collected exploits are associated with a Common Vulnerabilities and Exposures (CVE) identifier, no vulnerability is published in RuMarket before its publication on NVD. Interestingly, reporting the vulnerability's CVE is also the de facto standard for exploit advertisement in RuMarket (see sec. 3.2 in [1] for a discussion of why this is the case). All Malware samples included an exploit for the same vulnerability, which allows the malware to escalate to a higher privilege group on the victim system.

EKIT and Standalone exploits account for most of the variability in the market. EKIT exploits are by far the older ones at time of publication; 50% of Standalone exploits arrive two months after disclosure, whereas the faster 50% of EKIT's make it to the market after more than nine months. This has a clear correspondence with the package prices reported in Table 2, in which Standalone exploits are the most expensive in the market and EKITs the cheapest. A more formal analysis indeed reveals a strong correlation between exploit price and exploit age, with significantly different rates associated to different vulnerable software platforms: for example, exploits for Microsoft and Adobe products appear to better retain their value as they age than exploits for Oracle products.

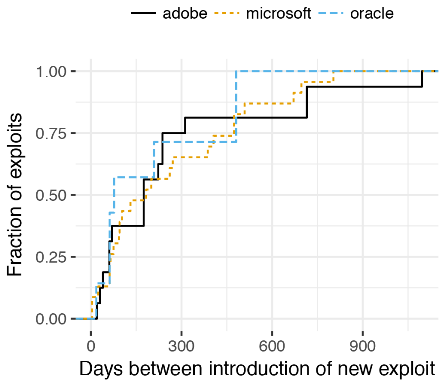

Figure 3: Distribution of days between exploit introduction

Another important aspect in the overall threat scenario is *how often* exploits for a software platform are updated in the market. Figure 3 reports the cumulative distribution function of the time that passes between new exploits for a specific software, grouped by vendor. Irrespective of software vendor, we observe that in the median case, exploits are substituted six months after first introduction. The slowest update rate of exploits is around two years. This figure is well in line with previous findings on measurements of exploit appearance in the wild [5, 10] and underlines the importance of considering attacker activity in estimating vulnerability risk.

Economic Factors of Vulnerability Exploitation

To evaluate the relation between market activity and risk of exploit, we rely on data from Symantec on the presence of an exploit at scale [4]. Note that whereas an exploit for a vulnerability might well exist even if not reported by Symantec, it is unlikely for an exploit that delivers on the order of hundreds of thousands or millions of attacks to remain unnoticed and unreported.

We consider exploit package price, market activity around an exploit (measured in terms of the number of RuMarket responses to the ad reporting the exploit), and vulnerability severity as factors that may affect the probability of finding an exploit at scale. A formal analysis reveals that all effects significantly affect the change in odds of exploitation in the wild for the respective vulnerability. Whereas a full description of the technical analysis is given in [1], as a rule-of-thumb the following emerges:

1. As market activity around an exploit doubles, so do the odds of finding an exploit at scale for the corresponding vulnerability.
2. As price of exploit acquisition doubles, the odds of exploit at scale halve.
3. Once we consider exploits traded in the markets, vulnerability severity becomes a significant predictor for exploitation in the wild.

Whereas the figures above are only indicative, a fully quantitative model can be obtained by plugging the coefficients reported in [1] in any vulnerability risk model. Importantly, a first approximation can be obtained without any direct insight from the markets. For example, exploit price can be estimated by considering the software vendor and the age of the vulnerability at the time of the estimate; this price can then be used, in conjunction with the vulnerability's severity, to estimate the change in the risk profile of the vulnerability if introduced in the market and how this evolves as time passes.

Although these conclusions are necessarily limited to RuMarket, and therefore the specific quantitative estimations may vary by considering other markets (e.g., trading vulnerabilities affecting different software vendors, or aiming at a larger English-speaking community), the qualitative conclusion remains: attacker economics are clearly correlated with risk of attack. Further research is needed in this direction: what is the attacker's process in deciding on which exploit to introduce and when? What determines whether an exploit can be expected to be traded in a market, as opposed to being used privately, or not being used at all? I believe that a characterization of these aspects can fundamentally change our perspective on cyber-risk and can provide an important building block for the division of workable and effective security practices.

References

[1] L. Allodi, "Economic Factors of Vulnerability Trade and Exploitation," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS '17),* pp. 1483–1499: https://acmccs.github.io/papers/p1483-allodiA.pdf.

[2] L. Allodi, "The Heavy Tails of Vulnerability Exploitation," in *Proceedings of 7th International Symposium on Engineering Secure Software and Systems (ESSoS '15)*, pp. 133–148.

[3] L. Allodi, M. Corradin, and F. Massacci, "Then and Now: On the Maturity of the Cybercrime Markets," *IEEE Transactions on Emerging Topics in Computing*, vol. 4, no. 1 (Jan.–March 2016): http://ieeexplore.ieee.org/document/7044581/.

[4] L. Allodi and F. Massacci, "Comparing Vulnerability Severity and Exploits Using Case-Control Studies," *ACM Transaction on Information and System Security (TISSEC*), vol. 17, no. 1 (August 2014): http://disi.unitn.it/~allodi/allodi-tissec-14.pdf.

[5] L. Allodi, F. Massacci, and J. Williams, "The Work-Averse Cyber Attacker Model: Evidence from Two Million Attack Signatures," in *Workshop on the Economics of Information Security (WEIS '17)*: https://ssrn.com/abstract=2862299.2017.

[6] L. Bilge and T. Dumitras, "Before We Knew It: An Empirical Study of Zero-Day Attacks in the Real World," in *Proceedings of the 2012 ACM Conference on Computer and Communications Security* (*CCS '12),* pp. 833–844: http://users.ece.cmu.edu/~tdumitra/public_documents/bilge12_zero_day.pdf.

[7] B. C. Ezell, S. P. Bennett, D. von Winterfeldt, J. Sokolowski, and A. J. Collins, "Probabilistic Risk Analysis and Terrorism Risk," *Risk Analysis*, vol. 30, no. 4 (2010), pp. 575–589: https://www.dhs.gov/xlibrary/assets/rma-risk-assessment-technical-publication.pdf.

[8] C. Grier, L. Ballard, J. Caballero, N. Chachra, C. J. Dietrich, K. Levchenko, P. Mavrommatis, D. McCoy, A. Nappa, A. Pitsillidis, N. Provos, M. Z. Rafique, M. A. Rajab, C. Rossow, K. Thomas, V. Paxson, S. Savage, G. M. Voelker, "Manufacturing Compromise: The Emergence of Exploit-as-a-Service," in *Proceedings of the 2012 ACM Conference on Computer and Communications Security* (*CCS '12),* pp. 821–832: http://cseweb.ucsd.edu/~voelker/pubs/eaas-ccs12.pdf.

[9] C. Herley and D. Florencio, "Nobody Sells Gold for the Price of Silver: Dishonesty, Uncertainty and the Underground Economy" in *Economics of Information Security and Privacy* (Springer, 2010).

[10] K. Nayak, D. Marino, P. Efstathopoulos, T. Dumitras, "Some Vulnerabilities Are Different Than Others," in *Proceedings of the 17th International Symposium on Research into Attacks, Intrusions, and Defenses (RAID '14)*, pp. 426–446: http://www.umiacs.umd.edu/~tdumitra/papers/RAID-2014.pdf.

27TH USENIX SECURITY SYMPOSIUM

BALTIMORE, MD, USA

Co-located Workshops

WOOT '18

12th USENIX Workshop on Offensive Technologies
August 13–14, 2018
Submissions due May 30, 2018
www.usenix.org/woot18

WOOT '18 aims to present a broad picture of offense and its contributions, bringing together researchers and practitioners in all areas of computer security. Offensive security has changed from a hobby to an industry. No longer an exercise for isolated enthusiasts, offensive security is today a large-scale operation managed by organized, capitalized actors. Meanwhile, the landscape has shifted: software used by millions is built by startups less than a year old, delivered on mobile phones and surveilled by national signals intelligence agencies. In the field's infancy, offensive security research was conducted separately by industry, independent hackers, or in academia. Collaboration between these groups could be difficult. Since 2007, the USENIX Workshop on Offensive Technologies (WOOT) has aimed to bring those communities together.

ASE '18

2018 USENIX Workshop on Advances in Security Education
August 13, 2018
Submissions due May 8, 2018
www.usenix.org/ase18

ASE '18 is intended to be a venue for cutting-edge research, best practices, and experimental curricula in computer security education. The workshop welcomes a broad range of paper and demo submissions on the subject of computer security education in any setting (K–12, undergraduate, graduate, non-traditional students, professional development, and the general public) with a diversity of goals, including developing or maturing specific knowledge, skills and abilities (KSAs), or improving awareness of issues in the cyber domain (e.g., cyber literacy, online citizenship). ASE is intended to be a venue for educators, designers, and evaluators to collaborate, share knowledge, improve existing practices, critically review state-of-the-art, and validate or refute widely held beliefs.

CSET '18

11th USENIX Workshop on Cyber Security Experimentation and Test
August 13, 2018
Submissions due May 10, 2018
www.usenix.org/cset18

CSET '18 invites submissions on cyber security evaluation, experimentation, measurement, metrics, data, simulations, and testbeds. The science of cyber security poses significant challenges. For example, experiments must recreate relevant, realistic features in order to be meaningful, yet identifying those features and modeling them is very difficult. Repeatability and measurement accuracy are essential in any scientific experiment, yet hard to achieve in practice. Few security-relevant datasets are publicly available for research use and little is understood about what "good datasets" look like. Finally, cyber security experiments and performance evaluations carry significant risks if not properly contained and controlled, yet often require some degree of interaction with the larger world in order to be useful.

FOCI '18

8th USENIX Workshop on Free and Open Communications on the Internet
August 14, 2018
Submissions due May 24, 2018
www.usenix.org/foci18

FOCI '18 will bring together researchers and practitioners from technology, law, and policy who are working on means to study, detect, or circumvent practices that inhibit free and open communications on the Internet. Internet communications drive political and social change around the world. Governments and other actors seek to control, monitor, and block Internet communications for a variety of reasons, ranging from extending copyright law to suppressing free speech and assembly. Methods for controlling what content people post and view online are also multifarious. Whether it's traffic throttling by ISPs or man-in-the-middle attacks by countries seeking to identify those who are organizing protests, threats to free and open communications on the Internet raise a wide range of research and interdisciplinary challenges.

HotSec '18

2018 USENIX Summit on Hot Topics in Security
August 14, 2018
Lightning talk submissions due June 11, 2018
www.usenix.org/hotsec18

HotSec '18 aims to bring together researchers across computer security disciplines to discuss the state of the art, with emphasis on future directions and emerging areas. HotSec is not your traditional security workshop! The day will consist of sessions of lightning talks on emerging work and positions in security, followed by discussion among attendees. Lightning talks are 5 MINUTES in duration—time limit strictly enforced with a gong! The format provides a way for lots of individuals to share ideas with others in a quick and more informal way, which will inspire breakout discussion for the remainder of the day.

Registration will open in May 2018.