Underground Economics for Vulnerability Risk

Luca Allodi
Eindhoven University of Technology
Eindhoven, the Netherlands

Abstract

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, we infiltrated a prominent Russian cybercrime market where the most prominent attack technology is traded. Supported by direct observations of market activity, in this work we investigate the economic factors that drive the adoption of new attacks *at scale*, and their effect on risk of attack in the wild. As market participants, we have access to the full spectrum of attack services offered to all members. In particular, in this work we look at the market economics of vulnerability exploitation.

Introduction

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.

Are all vulnerabilities (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 [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 works showed that this huge skew in attack distribution is present also for 0-day vulnerabilities [6]: in this analysis, across twenty 0-day vulnerabilities two were reportedly responsible for millions attacks worldwide, one for twenty thousands, and the remaining seventeen for a few dozens only. These results are confirmed in follow-up empirical studies that estimate 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].

Whereas this is in sharp contrast with the current narrative in the information security community (according to which every new severe vulnerability loosely resembles *Doomsday*), industry studies recently started to acknowledge this effect as well (e.g. in the last few editions of the 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 at this point important to clarify the nature of the data leading to the observations above, and its relation with different attacker types. In general, field data concerns attacks of an 'untargeted' nature, where attackers in possess 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 this case attackers adapt their exploit portfolio to the desired target system (as opposed to relying on a fixed set of exploits). These however concern a very limited set of Internet system, 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). For this reason, in this article we specifically refer to *risk of untargeted attacks at scale*.

A dive into exploit economics

This distinction between 'untargeted' and 'targeted' attacks has became 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 Prof. Fabio Massacci at the University of Trento (Italy) and Prof. Julian Williams at the Durham Business School (UK), in 2011 I started evaluating different underground markets in the English and Russian hacking communities. 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 gained first access to `RuMarket` in 2011, and carried '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 supervisioned by the market administration, whereby access was granted only if either:

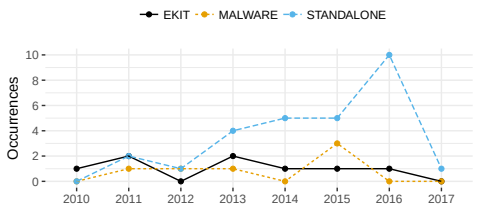

Figure 1: Release of exploit packages by type per year

1. A trusted member of the market vouches the entry request, effectively implementing a *pull-in* mechanism;

2. The request for market entry is backed up by evidence that the requestor is a reputable member of the Russian hacking community.

As we had no contacts inside the market, to re-gain 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 required some proficiency in Russian in the discussion boards, and 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 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 in three categories: `EKIT` (Exploit Kits), `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 model 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 ten 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 counter-intuitive effect, whereby only 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, 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 Fig. 1.

Table 1 reports descriptive statistics of package prices. Prices for rented EKITs are averaged over a period of 3 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 the 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 in time. This reflects the 'consumer' nature of EKIT products, that 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. Unsurprisingly, in RuMarket we find exploits for Microsoft, Oracle, and Adobe software, that 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 Hyper Cycle* describing the introduction of new technologies in a market: a first inflation in the expectations associated with that technology causes a burst in 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.

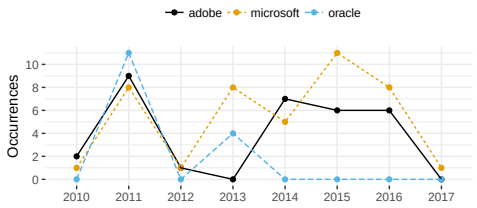

Figure 2: Occurrences of exploit publication by year

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 [1, Sec. 3.2] for a discussion on why is this the case). All MALWARE samples included an exploit for the same vulnerability, that 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

Table 1: Package prices.

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

Type	n	Min	Mean	Med.	Max
EKIT	25	1	372.48	294	1745
MAL	1	185	185	185	185
STDL	29	1	147.34	75	934
All	55	1	250.36	93	1745

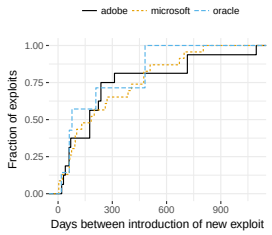

Figure 3: Distribution of days between exploit introduction

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, whereby STANDALONE exploits are the most expensive in the market, and EKITs the cheapest. A more formal analysis reveals indeed 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.

Another important aspect in the overall threat scenario is *how often* are exploits for a software platform updated in the market. Figure 3 reports the cumulative distribution function of the time that passes in between the introduction of a new exploit 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 are 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 the estimation of 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 in 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 above figures 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 age of the vulnerability at time of the estimation and the software vendor; 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.

Whereas 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's the attackers' decision process 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 provide an important building block for the devision of workable and effective security practices.

References

[1] Luca Allodi. "Economic Factors of Vulnerability Trade and Exploitation". In: *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*. CCS '17. Dallas, Texas, USA: ACM, 2017, pp. 1483–1499. ISBN: 978-1-4503-4946-8. DOI: 10.1145/3133956.3133960. URL: http://doi.acm.org/10.1145/3133956.3133960.

[2] Luca Allodi. "The Heavy Tails of Vulnerability Exploitation". In: *Proc. of ESSoS'15*. 2015.

[3] Luca Allodi, Marco Corradin, and Fabio Massacci. "Then and Now: On the Maturity of the Cybercrime Markets." In: *IEEE Transactions on Emerging Topics in Computing* (2015).

[4] Luca Allodi and Fabio Massacci. "Comparing vulnerability severity and exploits using case-control studies." In: *ACM Transaction on Information and System Security (TISSEC)* 17.1 (Aug. 2014).

[5] Luca Allodi, Fabio Massacci, and Julian Williams. "The Work-Averse Cyber Attacker Model. Evidence from two million attack signatures". In: *WEIS 2017. Available at https://ssrn.com/abstract=2862299*. 2017.

[6] Leyla Bilge and Tudor Dumitras. "Before we knew it: an empirical study of zero-day attacks in the real world". In: *Proc. of CCS'12*. Raleigh, North Carolina, USA: ACM, 2012, pp. 833–844.

[7] Barry Charles Ezell et al. "Probabilistic Risk Analysis and Terrorism Risk". In: *Risk Analysis* 30.4 (2010), pp. 575–589. ISSN: 1539-6924. DOI: 10.1111/j.1539-6924.2010.01401.x. URL: http://dx.doi.org/10.1111/j.1539-6924.2010.01401.x.

[8] Chris Grier et al. "Manufacturing compromise: the emergence of exploit-as-a-service". In: *Proc. of CCS'12*. ACM, 2012, pp. 821–832.

[9] C. Herley and D. Florencio. "Nobody sells gold for the price of silver: Dishonesty, uncertainty and the underground economy". In: *Springer Econ. of Inf. Sec. and Priv.* (2010).

[10] Kartik Nayak et al. "Some Vulnerabilities Are Different Than Others". In: *Proc. of RAID'14*. Springer, 2014, pp. 426–446.