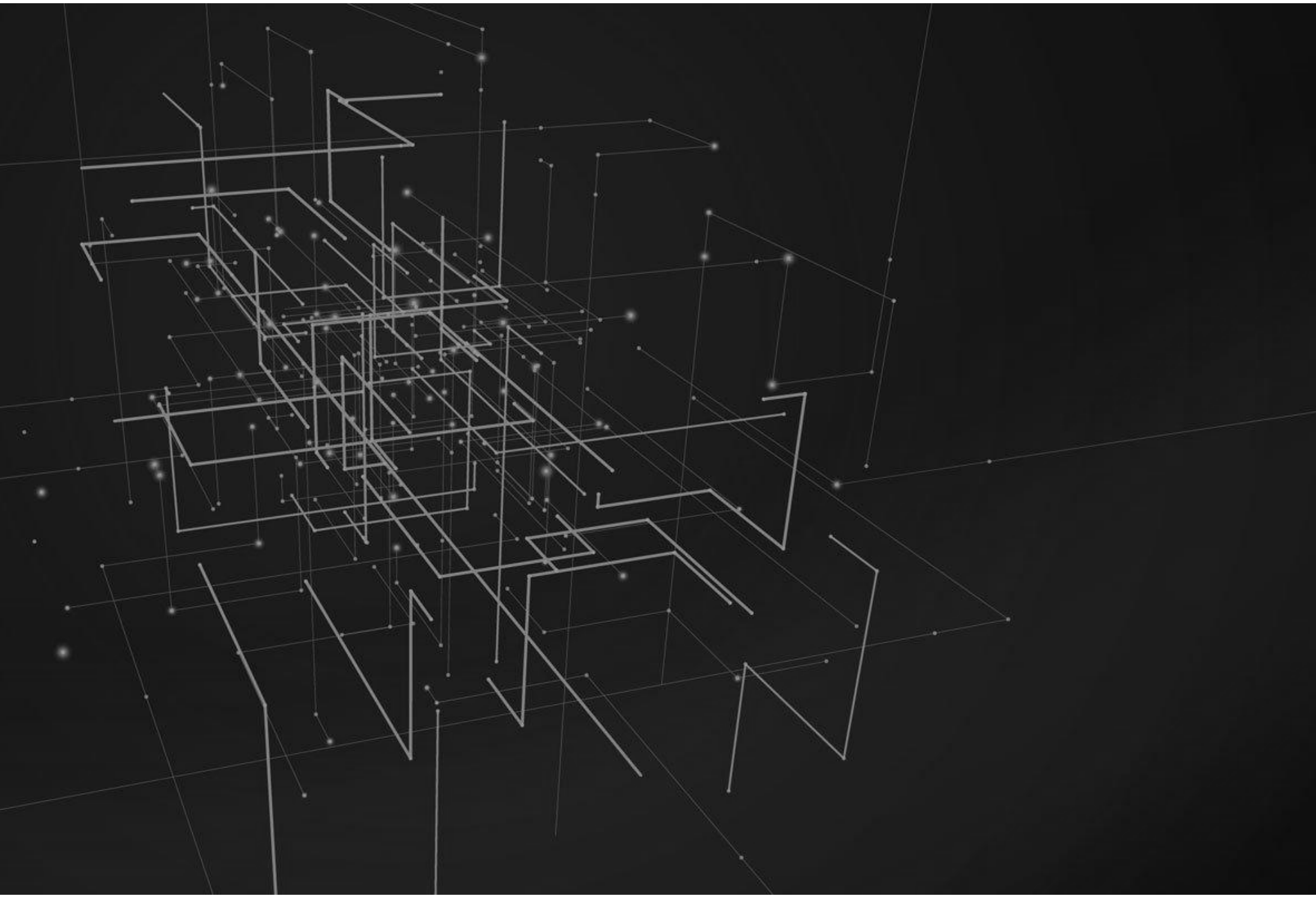
REPORT

2018 State of Vulnerability Risk Management



NOPSEC

Recorded Future

nopsec.com | hello@nopsec.com

Table of Contents

Executive Summary

This report offers an analysis into current trends in vulnerability risk management. It examines the attributes of security vulnerabilities viewed through a variety of lenses:

- Attributes of vulnerabilities published since 2002 versus those only recently published
- Attributes of all vulnerabilities published in the National Vulnerability Database (NVD) in contrast with only those uploaded into our platform by our clients
- Vulnerabilities broken down by industry vertical, CVSS score, product vendor and active exploitation in the wild

In building this report, Nopsec examined anonymized data collected from clients using Unified VRM, the company's flagship vulnerability risk management (VRM) product. To get a broader picture of the landscape, the report looks at data from a variety of sources, including commercial threat intelligence and vulnerability management platforms.

The report highlights several ongoing trends that are worthy of note, particularly to those working in remediation. In particular, it shows that prioritizing vulnerabilities solely by CVSS score severity isn't always appropriate.

The evidence shows that a surprisingly high portion of vulnerabilities incorporated into malware or exploit kits are ranked low or medium severity. Counter to commonly-accepted practices, focusing only on high-severity vulnerabilities and setting a ‘cut-off’ point for lower scored issues, is not a safe or effective strategy.

In addressing the unreliability of CVSS scoring, the report explores the application of machine learning, natural language processing and other techniques in search of better indicators of risk. By analyzing previous trends in vulnerabilities, this novel approach is able to better predict the threat posed by a freshly-discovered vulnerability.

Key Findings

1. We found that approximately 21% of CVEs published have associated exploit code in the Exploit Database alone. However, only 1.6% have associated Metasploit modules. Less than 2% (1.92%) have been linked to malware. Roughly 95% of vulnerabilities ranked as high have never been linked to malware seen in the wild.
2. 44% of CVEs associated with malware were scored as medium or low on the CVSS scale, suggesting that focusing solely on CVEs with high scores (7+) would be a mistake.
3. NopSec has found that the language used in CVE descriptions lends clues to the fate of vulnerabilities. For example, approximately half of all descriptions of vulnerabilities linked to malware include words “allows remote”.

4. Vendors most likely to be associated with malware vary significantly, depending on whether all CVE data is taken into consideration, or just the last 18 months’ worth. For example, OpenSSL is most commonly associated with malware when considering all CVEs, whereas Canonical (Ubuntu) takes the top spot when considering only recently published CVEs.
5. Only half of the Top 20 vulnerabilities derived from NopSec client data can be fixed with a patch. The remainder represent configuration issues to be fixed or insecure cryptographic algorithms or protocols to be disabled.
6. Microsoft is the biggest source of vulnerabilities for Financial Services organizations. Healthcare, however, has more to worry about from BSD and Linux. All industries have a significant number of Oracle vulnerabilities.

Introduction

Each new year challenges those working on the front lines of the security field, but for some reason, 2018 feels especially arduous – and it isn’t even over yet. In many respects, a large number of the challenges faced are hangovers from 2017, only scarier.

Cryptojacking, for example, is no longer confined to the shadier parts of the internet, like torrent sites and adult video properties. It has matured, and is now finding its way into previously unanticipated nooks-and-crannies, like the Internet of Things, Mobile Devices, and serverless applications.

Another worrying trend is the use of leaked, military-grade exploits in – for lack of a better phrase – mass-market, consumer-grade attacks. There's still life left in the NSA's EternalBlue exploit, long after it wrecked havoc during the WannaCry debacle, and it's no longer solely used by nation state actors.

That's on top of the other attacks - the usual malware, ransomware, bugs, and assorted nasty detritus that practitioners are laden with.

Suffice to say, 2018 has been a tough year for our industry. This report will help you make sense of the story so far. It will guide you through the contemporary vulnerability landscape, where we see the biggest threats coming from, and the challenges faced in comprehending risk. We hope you'll find it useful

NopSec Note: As with previous years, one of our partners has been gracious enough to provide us with some insight for our report. This year, we've moved it to the front, as Recorded Future's sobering view into how geopolitics affect government-run vulnerability databases is an apt primer for the remainder of the report.

Vulnerability Views Vary By Province

Information about new vulnerabilities should be a relatively mundane topic that is bereft of the geopolitics that consumes so much of the world, but it turns out that is not always the case. In fact, exposure to new vulnerabilities can vary greatly from country to country and the risk of exploitation by a new vulnerability can often be highly dependent on where the victim lives.

The World of National Vulnerability Databases

The authoritative record for new vulnerabilities generally rests in the National Vulnerability Database (NVD). Most large countries maintain their own version of an NVD and as new software vulnerabilities are discovered and reported the NVD catalog those vulnerabilities, along with relevant information, such as impacted Common Platform Enumeration (CPE) and the Common Vulnerability Scoring System (CVSS) score. The United States' NVD is hosted by NIST and is available on the NIST website (https://nvd.nist.gov/).

Vulnerability scanners and vulnerability management teams (VMT) rely heavily on NIST's NVD to build a base of information that allows the VMT to scan their networks for new vulnerabilities and to prioritize necessary patching.

In addition to the United States, other countries maintain their own NVDs, but information is not equally populated across all NVDs. For example, China's NVD (CNNVD) typically adds new vulnerabilities in 12 days from initial reporting, whereas NVD takes on average 27 days from initial reporting to appearing in the the NVD, though based on Recorded Future's research both

Recorded Future

About Recorded Future

Recorded Future delivers the only complete threat intelligence solution powered by patented machine learning to lower risk. They empower organizations to reveal unknown threats before they impact business and enable teams to respond to alerts 10 times faster.

To supercharge the efforts of security teams, Recorded Future's technology automatically collects and analyzes intelligence from technical, open web, and dark web sources, and aggregates customer-proprietary data.

Recorded Future delivers more context than threat feeds, updates in real time so intelligence stays relevant, and centralizes information ready for human analysis, collaboration, and integration with security technologies. 91 percent of the Fortune 100 use Recorded Future.

Learn more at
www.recordedfuture.com
and follow them on Twitter at
@RecordedFuture.

NVDs are getting faster at reporting new vulnerabilities. Part of this is due to the fact that NVD relies on voluntary vendor reporting, whereas the CNNVD uses web reporting, so when a vulnerability is first announced it is added to the CNNVD.

However, there are some interesting quirks with the CNNVD. For example, the CNNVD altered reporting for some of their vulnerabilities that were suspected of being in use by the Ministry of State Security (MSS). It appears that the MSS held reporting for these vulnerabilities as they evaluated their value for exploitation. When MSS finished their evaluation process they added the vulnerabilities to CNNVD and backdated their release so it appeared the vulnerabilities had been in the CNNVD the entire time.

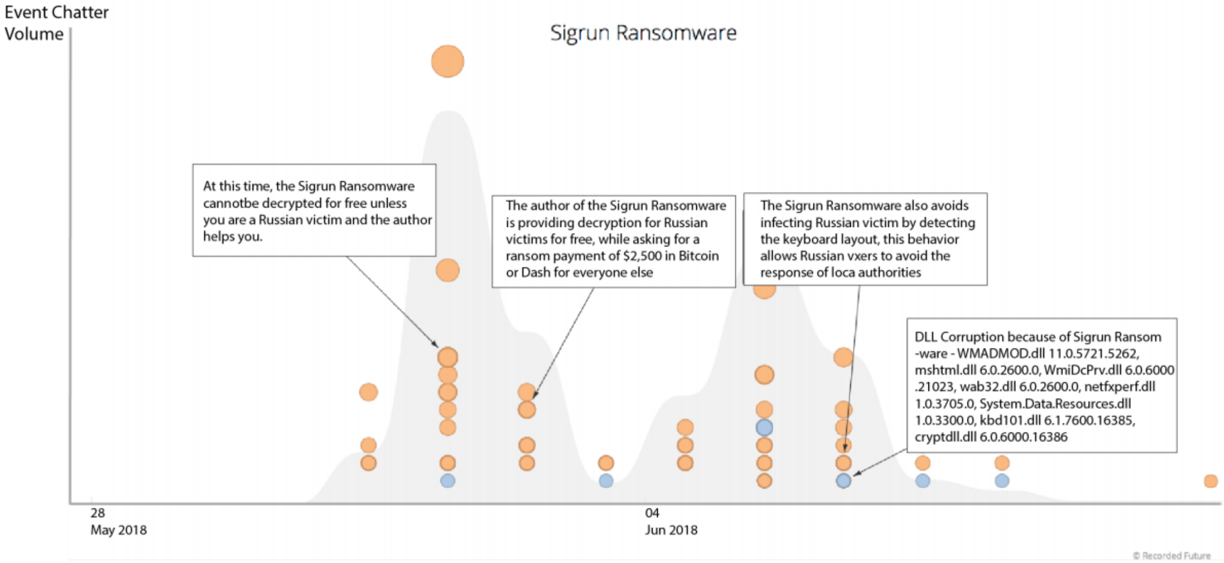
Russia is another example of oddities in the NVD process. The Russian NVD is known as the BDU and it lags significantly behind both the United States and China. Where the United States has almost 108,000 vulnerabilities in its NVD, the BDU only have a little more than 11,000, roughly 10% of the vulnerabilities reported in the United States. The BDU is also significantly slower to report new vulnerabilities, where China takes, on average, 13 days and the United States takes 33 those vulnerabilities that do get added to the BDU take 83 days on average.

BDU appears to focus vulnerability reporting on those systems that are deemed part of Russia's critical infrastructure. Rather than being run as a public service, the way NVD and CNNVD are, the primary purpose of the BDU is to keep the Russian government safe.

It appears that no matter where a VMT resides, relying solely on a government NVD for information about newly discovered vulnerabilities could leave an organization exposed to new exploits.

Best Protection Against Malware: Cyrillic Keyboard

The difference in vulnerability exposure is not limited to just the NVDs. Many types of malware have mechanisms that prevent them from exploiting or installing themselves on victim machines in certain countries. This activity is most often associated with Russia.



At this time, the Sigrun Ransomware cannotbe decrypted for free unless you are a Russian victim and the author helps you.

The author of the Sigrun Ransomware is providing decryption for Russian victims for free, while asking for a ransom payment of $2,500 in Bitcoin or Dash for everyone else

The Sigrun Ransomware also avoids infecting Russian victim by detecting the keyboard layout, this behavior allows Russian vxers to avoid the response of loca authorities

DLL Corruption because of Sigrun Ransomware - WMADMOD.dll 11.0.5721.5262, mshtml.dll 6.0.2600.0, WmiDcPrv.dll 6.0.6000.21023, wab32.dll 6.0.2600.0, netfxperf.dll 1.0.3705.0, System.Data.Resources.dll 1.0.3300.0, kbd101.dll 6.1.7600.16385, cryptdll.dll 6.0.6000.16386

For example, the Sigrun ransomware, which first appeared in late May or early June of 2018. The ransomware primarily targeted users in the Ukraine and Western Europe, but if it accidentally landed on a Russian computer it would not install itself, the ransomware looked for a cyrillic keyboard layout and if it detected it would not install.

The Sigrun author even went a step further, if a victim wrote to him claiming to be Russian he would provide the victim with a key to unlock encrypted files.

This is a surprisingly common tactic for malware, especially ransomware. The Cerber ransomware, one of the most successful ransomware campaigns of 2016 and 2017 looked for 15 different keyboard layouts to avoid, including Russian, Uzbek and Georgian.

There are two methods that attackers use to determine location information of their victim machines and some malware families switch from one to the other depending on its effectiveness. The Rapid 2.0 Ransomware used the GetLocaleInfo Microsoft API call when it was first uncovered in March of 2018. Later, in May of 2018, Rapid 3.0 Ransomware switched to using the Microsoft GetKeyboardLayoutList API call.

GetLocaleInfo/GetLocaleInfoEx is used to determine the location of the system, based on user entered information. GetKeyboardLayoutList is used to get information about the installed keyboard layout. One challenge in trying to determine how widespread this practice is is that these API calls are common and so don't show up in typical malware analysis reports.

These precautions in malware are not limited to ransomware. The Zeus Panda banking Trojan, first discovered in November 2017, would not install on computers with Russian, Belarusian, Kazak, or Ukrainian keyboard layouts.

This is not a new trend either. Going back to 2009, W2/Conficker.A would not install on computers with a Ukrainian layout. It is also not a well-kept secret many advertisements selling malware in underground forum explicitly state that the malware will not install on machines in Russia or using Russian language keyboards.

This type of behavior is normally expected from nation state actors, but as Russia has imposed harsh sentences on cyber criminals who operate in Russian IP space, many criminals feel the risk is not worth whatever monetary gains they may make.

Overall Vulnerability Landscape (NVD)

Vulnerability Counts

Any vulnerability worth its salt ends up ascribed a CVE number. These are a bit like social security numbers, but for vulnerabilities, allowing each to be uniquely identified. Keeping with the trend of

2018 being a landmark year, in April, the MITRE corporation published its hundred-thousandth CVE identifier. Although CVE-2017-2906, discovered by Cisco Talos, was a landmark numerically, it wasn’t anything particularly dangerous. Talos found an integer overflow error in Blender, the popular open-source 3D modelling application, which could result in arbitrary code execution by an attacker.

That said, it is impressive to ponder how we managed to hit this landmark CVE. Every vulnerability reported to MITRE is the culmination of hard work, totalling countless man-hours, and an unknowable amount of money spent.

There’s an observed trend showing vastly more vulnerabilities being discovered in recent years (Figure 1). 2017, for instance, saw double the number of CVEs published relative to 2016. In total, 14,643 CVEs were published (and assigned a CVSS score) in 2017, which works out to 40 new CVEs each day.

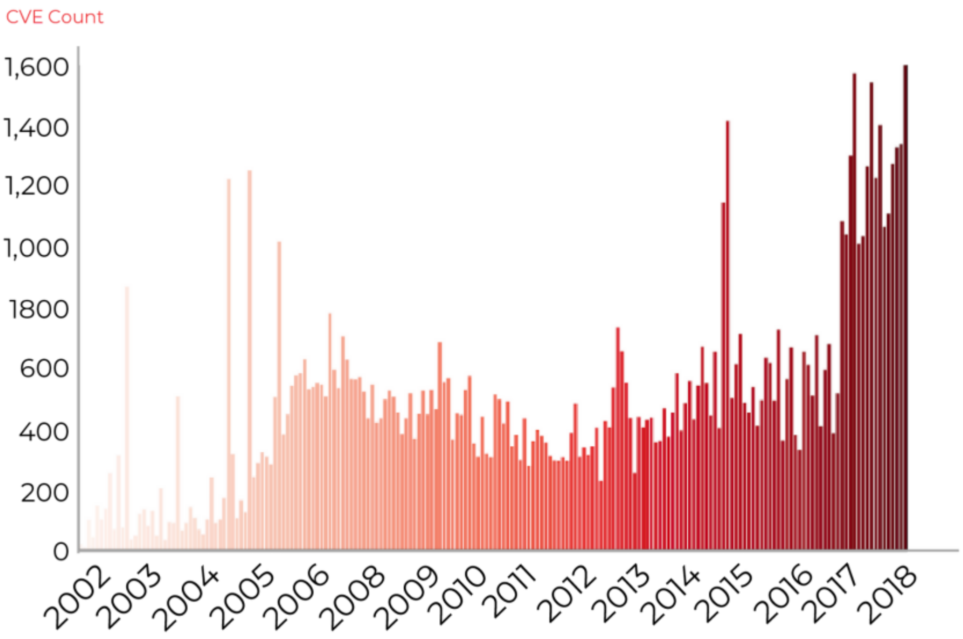


Figure 1:
Recent months have seen a spike in the number of CVEs published in NVD. Represented are all CVEs published between January 2002 and June 2018 which have had CVSS V2 scores assigned to them. Since the beginning of 2017 there have been a 1000+ vulnerabilities published each month.

This trend is partially attributed to a backlog of vulnerabilities reserved, but not published. In 2017, 26 percent of CVEs were dated to previous years. That's not the whole story, however. The vulnerability spike could also be blamed on a plethora of hopelessly vulnerable IoT devices, as well as companies being better at engaging with the security community, and the welcome proliferation of many private and public bug bounty programs.

Most Frequent Vulnerabilities

Almost three quarters (71%) of all vulnerabilities published in 2017 and 2018 so far (as of June 19, 2018) fall into just 10 Common Weakness Enumeration (CWE) categories (Figure 2). Few surprises here, this list includes buffer overflows, XSS, SQLi and a general lack of input validation. Information exposure has also always been among the most common security weaknesses.

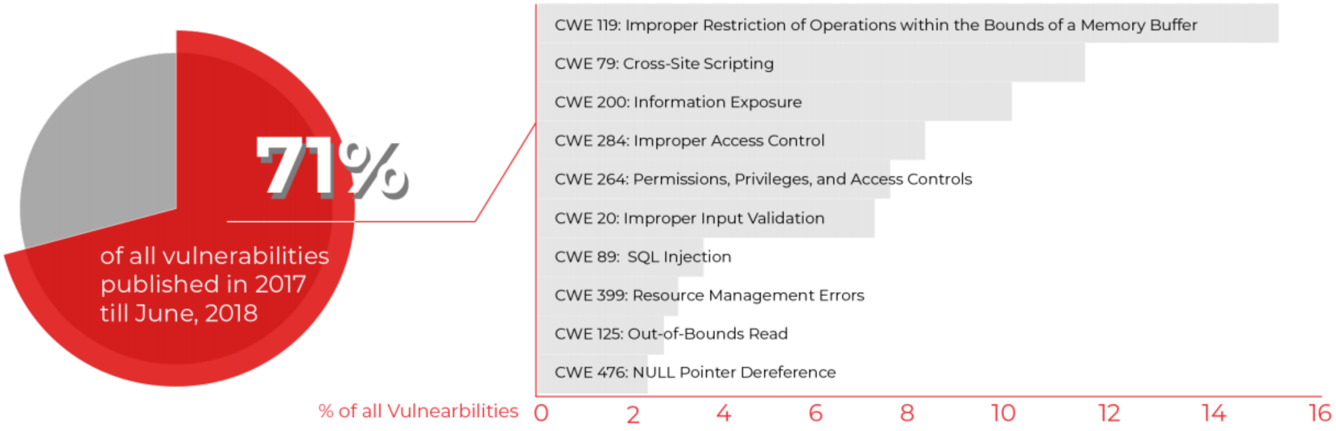


Figure 2:

Most common types of vulnerabilities based on CWE. Analyzed are vulnerabilities disclosed since the beginning of 2017. These ten categories describe > 70% of all vulnerabilities published in NVD in this period.

CVSS Score Analysis

The Common Vulnerability Scoring System (CVSS) offers a straightforward way to communicate the characteristics, risks, and potential impact of software vulnerabilities. It was initially introduced to help managers convert information into a simple, comprehensible score, which can easily be communicated to non-technical colleagues.

Note: This report assumes a basic understanding of CVSS. For an introduction or refresher, the NVD website has a concise explanation with examples (https://nvd.nist.gov/vuln-metrics/cvss). Throughout this report, we will be using and referring to CVSS version 2.

The below chart shows the distribution of all CVSS v2 scores for all publicly disclosed vulnerabilities. Most tend to skew towards medium and high severity. Approximately 38 percent of all historical vulnerabilities are rated high severity, while 54 percent are classified medium severity. Only 8 percent fall under low severity.

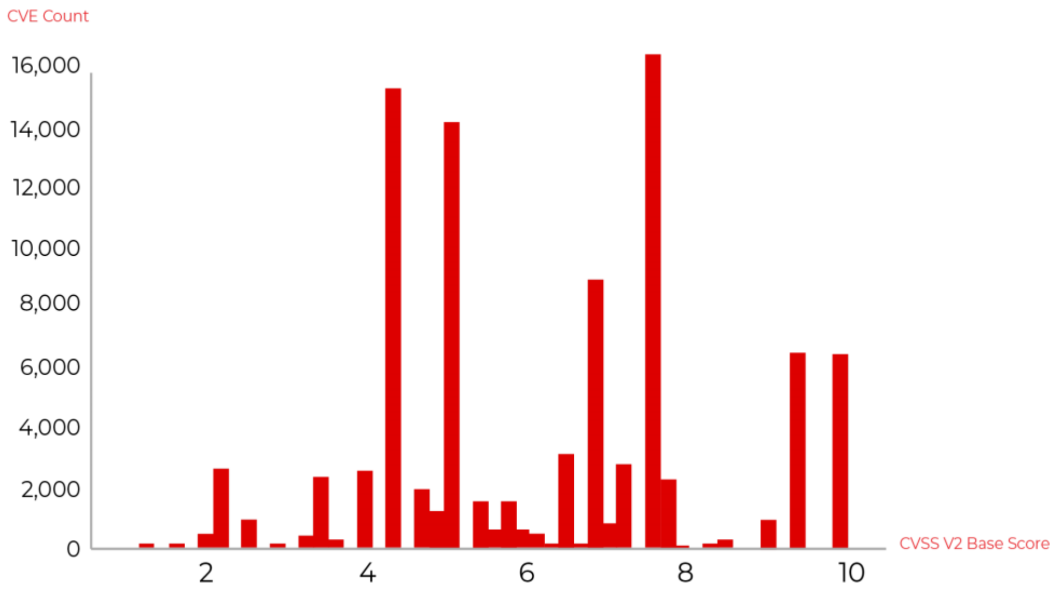


Figure 3:
Distribution of CVSS V2 Base Score for all publicly disclosed vulnerabilities since 2002.

This chart of all CVE ratings will serve as a baseline throughout the remainder of this report as we explore the true risk of these vulnerabilities and how they differ from the assigned scores.

Recently, there has been an increase in the portion of vulnerabilities rated medium, and a matching decrease in the portion assigned to high severity category (Figure 4). While the reason for this isn't clear, our data shows a significant portion of CVEs correlating with malware are mediums. An increasing population of mediums suggests that an increasing number of dangerous vulnerabilities may be getting ignored.

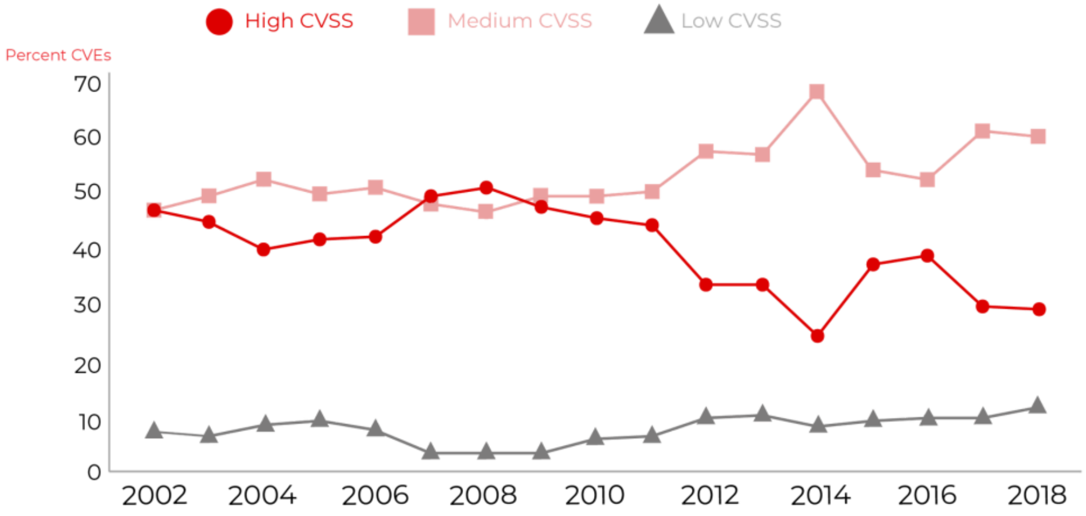


Figure 4:
Percent of vulnerabilities labeled as High based on CVSS V2 Base Score has slightly declined in recent years, while an increasing portion of vulnerabilities are ranked as Medium.

Overall Vulnerability Landscape (NVD)

Exploit and Malware Analysis

NopSec collects and aggregates data regarding existence of exploit code utilizing specific CVEs from many publicly available open source and commercial exploit code repositories and penetration testing tools. These include (but are not limited to):

- Exploit Database (EDB)
- Metasploit
- PacketStorm
- ZeroDayToday
- Immunity (Canvas, D2Square)
- CoreSecurity
- Saint

NopSec also uses threat intelligence provided by Recorded Future, Symantec, and AlienVault, among others. CVEs are correlated with malware, ransomware, remote access trojans, targeted attacks, exploits and exploit kits.

For simplicity here, we will refer to malware, ransomware, remote access trojans and exploit kits collectively as 'malware and exploit kits'.

We find that approximately 21% of CVEs published up to date have some exploit code published in Exploit Database, but only about 1.6% have Metasploit modules available. The portion of CVEs with exploit code when combining our sources listed above is approximately 24%. However, less than 2% (1.92%) have been historically or recently incorporated into malware and exploit kits, having therefore reached its riskiest state where they could easily be used in targeted attacks (Figure 5).

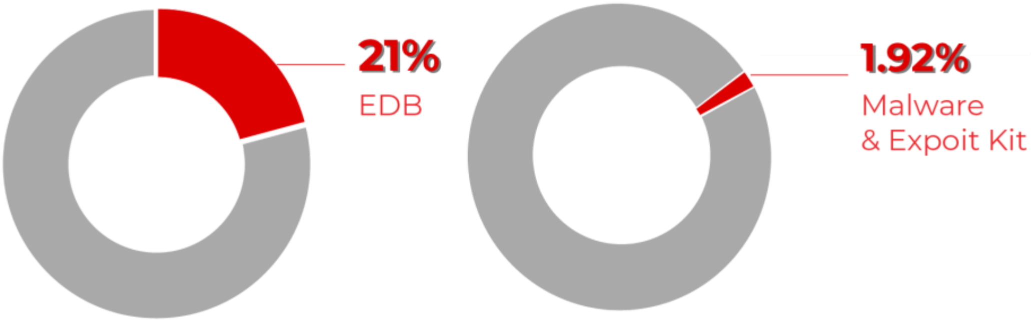


Figure 5:
Portions of all CVEs that have exploit code in EDB, and that have been incorporated into malware and exploit kits.

We find that when looking at the year a CVE is published in, malware and exploit kit incorporation rates have risen from approximately 0.2% all CVEs published in 2005 to 3.4% (so far) of CVEs published in 2017 (Figure 6).

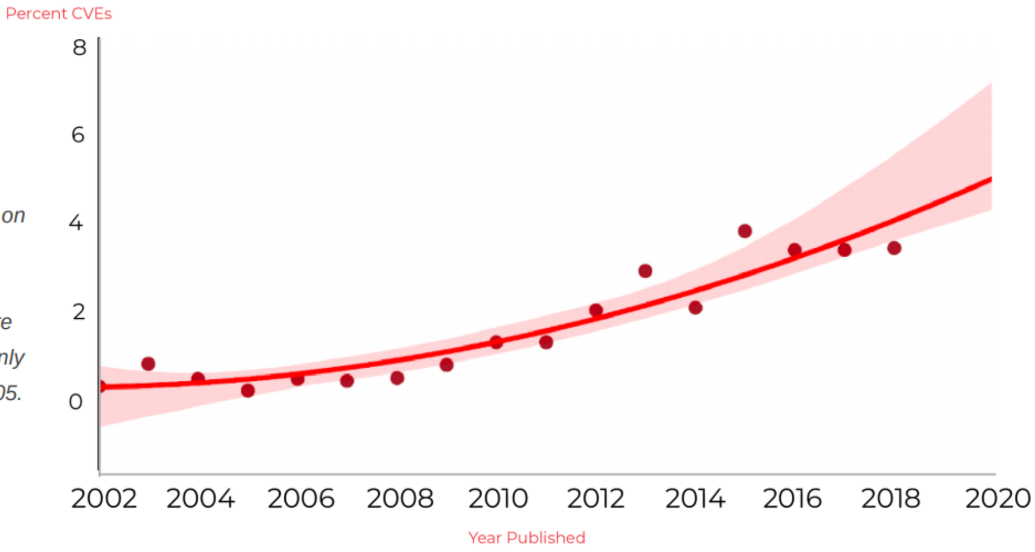


Figure 6:
Malware and exploit kits incorporation rates have been on a rise. Around 3.4% of CVEs published in 2017 have so far been incorporated into malware and exploit kits compared to only 0.2% of those published in 2005.

It is important to note that while 38% of all CVEs are marked as high severity based on CVSS, only under 2% have so far reached the most dangerous state of being used in malicious code and commoditized in exploit kits. This means that when prioritizing vulnerabilities associated with malware and exploit kits to remediate, choosing solely based on a high CVSS score would result in many false positives. In other words, many vulnerabilities that will never reach that level of danger would be incorrectly labeled as such.

Furthermore, the CVSS base score distribution for those CVEs that actually have malware and exploit kit association very much resembles the CVSS base score distribution for the entire CVE population (Figure 3), with 44% of CVEs with Malware & Exploit Kit association having a severity of medium or low (Figure 7). This means that choosing exclusively based on CVSS scores would lead to many false negatives, as well - dangerous vulnerabilities would be omitted by focusing only on items with high CVSS ratings.

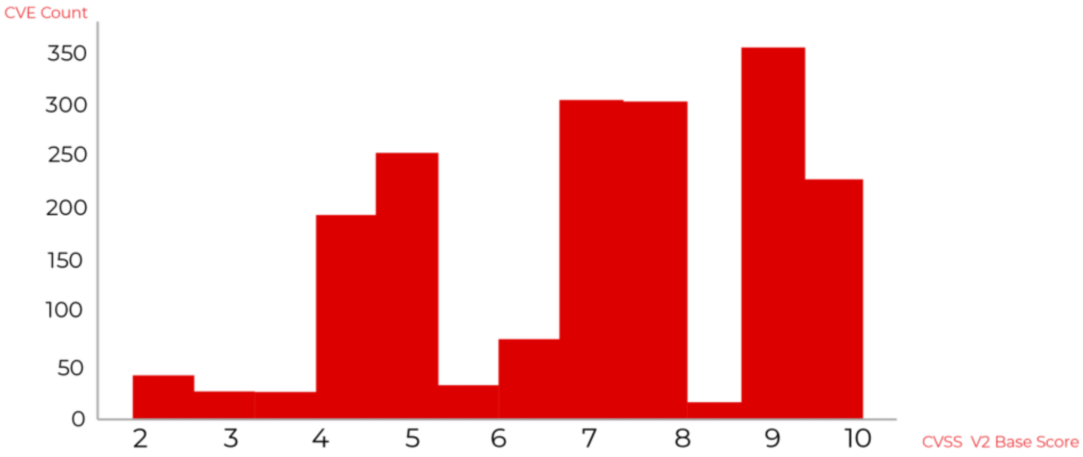


Figure 7:
Distribution of CVSS V2 Base Score for CVEs with Malware and Exploit Kits association - 44% have a CVSS Base Score less than 7 and are ranked as medium or low.

Analysis of Vulnerabilities by Vendor

Large vendors such as Microsoft, Google, Apple, Oracle, and IBM have consistently been associated with discovered vulnerabilities (Figure 8). However, while vulnerability counts for these vendors and their products are especially high relative to other vendors, this may be a reflection of their widespread use and size of code base rather than any inherent weakness or danger. In fact, some of these may have below average rates of exploits and malware presence.

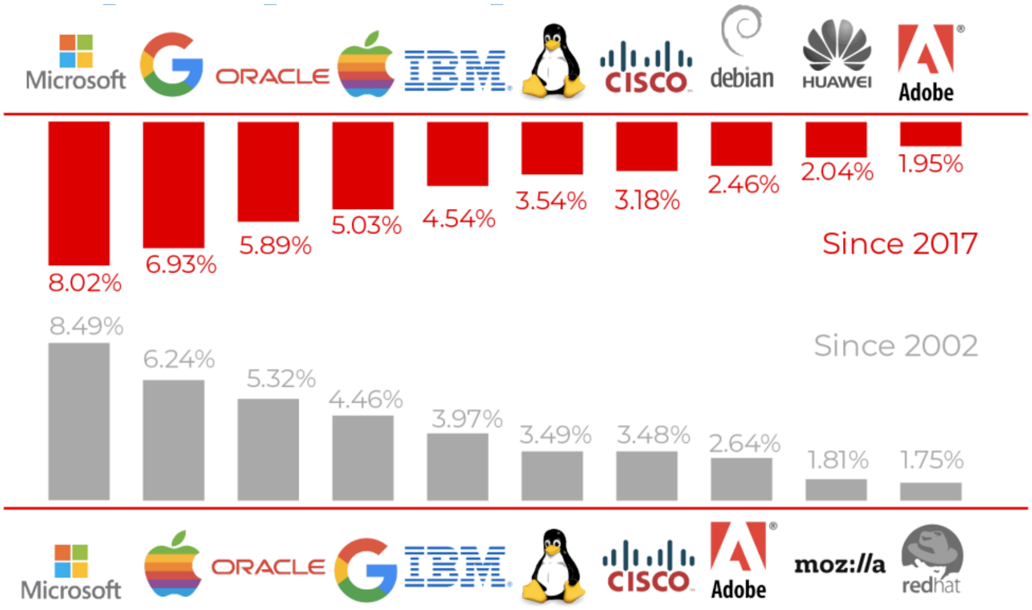


Figure 8: *Vendors with most vulnerabilities (CVEs). Bottom: since 2002, top: vulnerabilities disclosed since the beginning of 2017.*

Perhaps a more important question here is whether there are any vendors that have higher than average malware and exploit rates for vulnerabilities in their products (Figure 9).

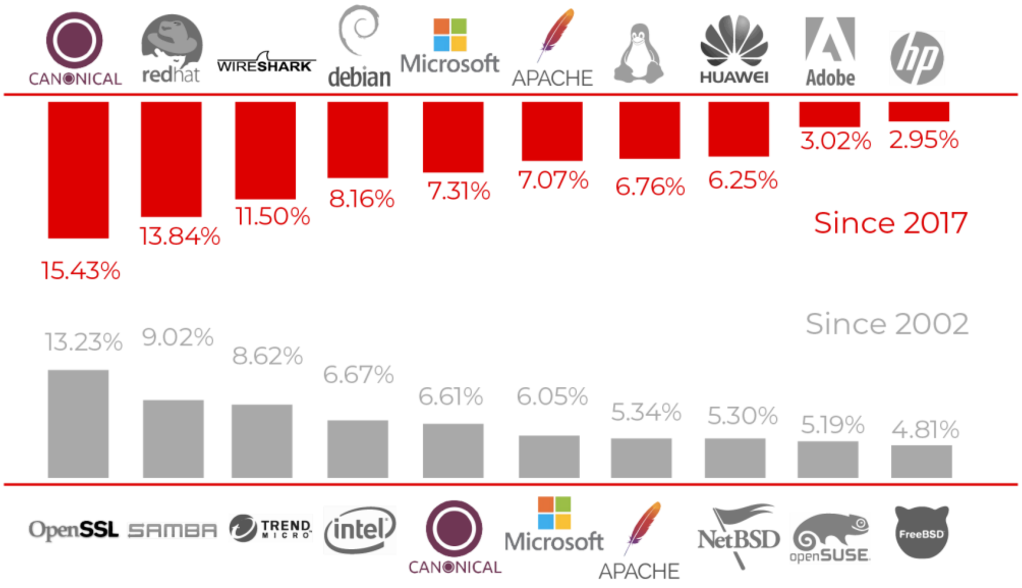


Figure 9: *Vendors with highest rates of malware and exploit kit incorporation (percent of CVEs with malware or exploit kit association). We analyzed only those vendors with more than 100 vulnerabilities overall (bottom) or more than 100 vulnerabilities since the beginning of 2017 (top). For reference, 1.92% of all CVEs since 2002 have malware association and 3.38% of all CVEs since the beginning of 2017 have malware association.*

In the above chart, we looked solely at vendors with more than 100 disclosed vulnerabilities overall, or those that have seen 100 disclosed since the start of 2017. It is worth noting that some vendors who weren't included had higher than average malware and exploit incorporation rates for their vulnerabilities.

While it might seem odd that Microsoft isn't at the forefront of each of these lists, consider the approach. Since we're counting CVEs in these totals, not malware, 10,000 malware variants leveraging a single vulnerability wouldn't affect Microsoft's ranking in this chart. That's an insight worth pondering - though Microsoft's Windows operating system has always been a malware magnet, malware tends to leverage a relatively small number of vulnerabilities (CVEs).

Natural Language Processing (NLP) Analysis

NopSec believes that language used in vulnerability descriptions carries important information that goes beyond the CVSS score in indicating how likely a vulnerability is to be exploited.

We will show words and combinations of words used in vulnerability descriptions (in NVD) that could be useful to differentiate CVEs associated with malware & exploit kits from those that are not.

We examined words, bigrams and n-grams (defined as "a contiguous sequence of n items from a given sample of text or speech") after removing stop words (commonly used English words such as ‘a’, ‘the’, ‘and’, etc.) and after stemming in order to consolidate expressions with similar meaning (keeping only stems of words so that, for example, “attacks” equalsbecomes “attack” and expressions such as “causing denial of service” and “cause denial of service” are counted together rather than as two different n-grams).

When considering the entire corpus of vulnerabilities’ descriptions (roughly 96,000 CVEs) we find that the most common words are “allow”, “attack”, “via”, “remote”, “vulnerability”, “arbitrary”, and “execute.” These terms appear in a high number of vulnerability descriptions. More importantly, we find that certain words, bigrams, etc. have much higher rates of malware and exploit kits association compared to the the numbers presented above (1.92% of all vulnerabilities have malware association).

Figure 10 shows bigrams appearing in more than 1000 vulnerability descriptions with highest percentage of malware and exploit kit correlation:

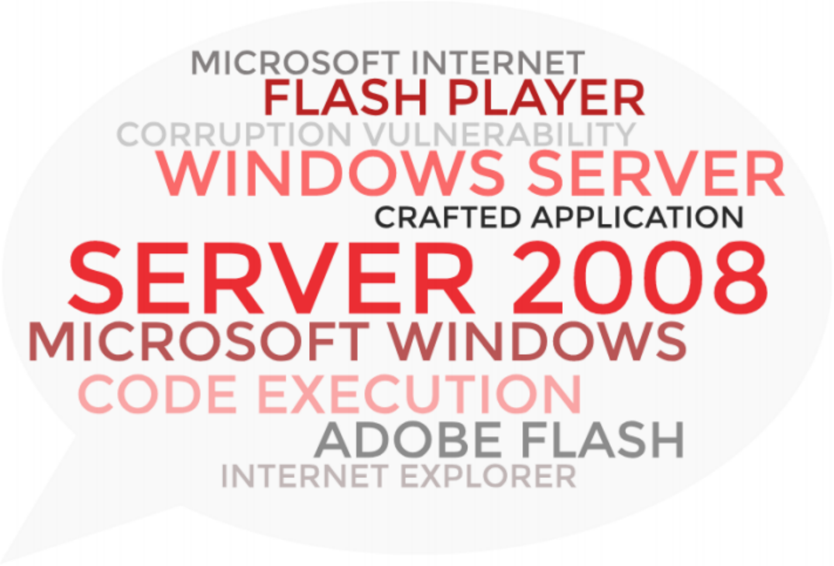


***Figure 10**: Bigrams appearing in >1,000 descriptions with a high percentage of malware association.*

For example, vulnerabilities with descriptions containing ‘server 2008’ or ‘Windows server’ have a malware incorporation rates of 10-11%, descriptions containing ‘Microsoft Windows’ are at 8.6% and descriptions containing ‘code execution’ have a malware incorporation rate of 8.4% - all of which are much higher than the overall malware incorporation rate.

A similar analysis on longer strings of contiguous words in vulnerability descriptions showed that descriptions containing talk about allowing remote attackers to execute arbitrary code and/or cause denial of service (memory corruption) had higher than expected rates of malware (Figure 11).

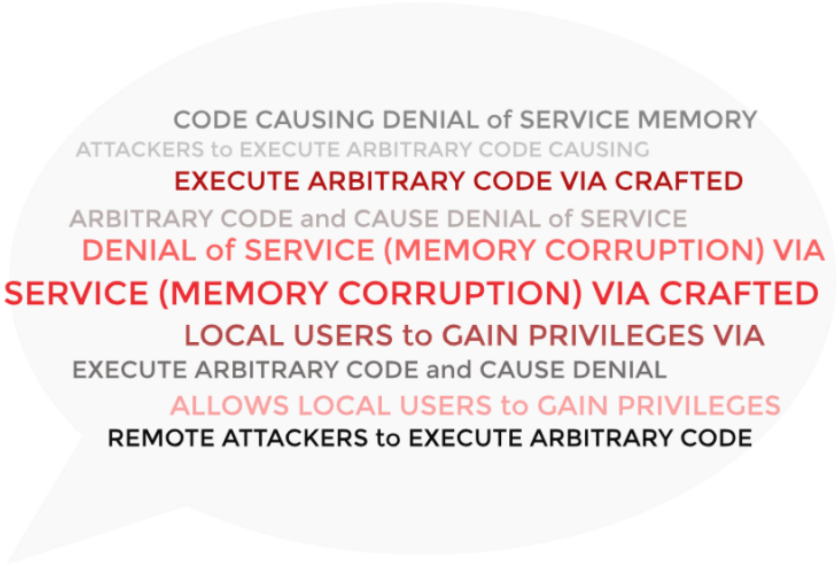


***Figure 11**: N-grams (n = 5) appearing in > 1000 descriptions with a high percentage of malware association.*

In fact, if we only look at past vulnerabilities for which we know that there is a recent or historical association to malware and exploit kits, the most commonly appearing bigrams in their descriptions are “allows remote” (found in 51% of all such CVEs), “remote attackers” (48%), “execute arbitrary” (35%), “arbitrary code” (31%), and so on (more complete list in Figure 12). We see from here that at least 50% of all such vulnerabilities involve remote code execution.

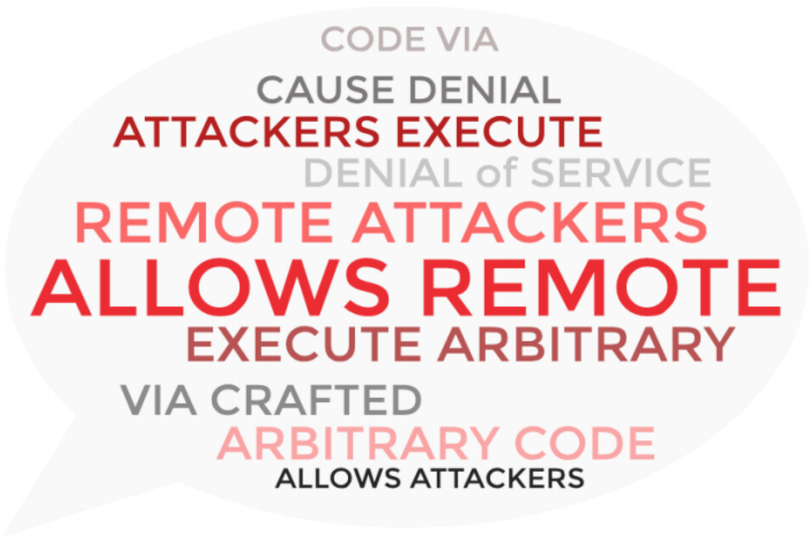


***Figure 12**:*
Most commonly seen bigrams in descriptions of vulnerabilities that have been associated with malware and exploit kits.

Furthermore, by analyzing more than two consecutive words, we can clearly see that the most common expressions indicate that these vulnerabilities are allowing remote attackers to execute arbitrary code and cause denial of service (Figure 13).

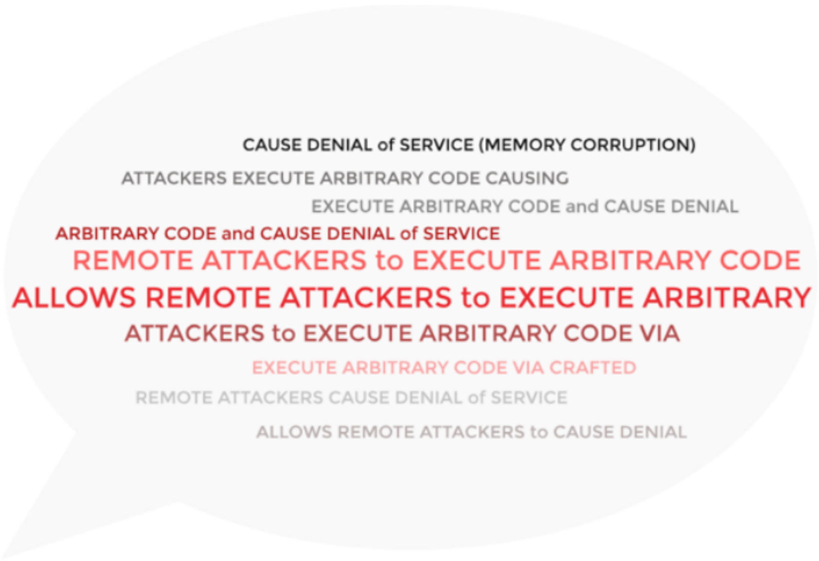


***Figure 13**:*
Most commonly seen n-grams (n = 5) in descriptions of vulnerabilities that have been associated with malware and exploit kits.

Vulnerability Landscape Based on NopSec Client Data

NopSec Client Data: Vulnerability Counts

Numbers and figures in the NVD portions of this report refer to vulnerabilities as identified by their unique CVE IDs. This is not a perfect approach, as the same vulnerability may carry different risks depending on the context. This context – both environmental and temporal – may carry just as much relevance as the intrinsic attributes.

As in our previous reports, when referring to our clients' data, we define a unique vulnerability as a unique combination of client, asset, vulnerability ID (scanner plugin identifier), and port affected. This part of the report is based on the analysis of aggregated anonymized NopSec Unified VRM client data.

Given this sea of vulnerabilities documented in NVD, a full scan of all of a company's assets may reveal thousands of vulnerabilities to address. In fact, our current (active) clients have faced over 1.5 million unique vulnerabilities in the last year alone!

For the sake of clarity, it's worth noting that this means there were 1.5 million unique vulnerabilities in scans from the last year. It doesn't necessarily mean they were first found last year. Some are leftover vulnerabilities previously discovered, but not remediated.

Our clients span a wide range of industries, but for the purposes of this report, as in our 2017 report, we have grouped them into one of four broad industry categories: Financial, Technology, Healthcare, and Other.

It is important to note that our analysis comes from a convenience sample of our clients – as such, we do not claim that this is a definitive analysis of all possible threats. The possibility of sample bias exists, and this should be kept in mind throughout the report.

However, we believe that our research offers important insight into how companies in various industries address vulnerabilities, universal weaknesses companies across industries share, and factors that should be incorporated into a comprehensive threat detection and remediation program.

In one year's period of time, our typical client in the Technology category has seen 9,155 unique vulnerabilities (median number of unique vulnerabilities per client by industry), followed by 3,939 vulnerabilities for clients in Healthcare and 834 for those in Financial industry (Figure 14).

Since these numbers are highly dependent on the number of assets a company has (or chooses to have scanned), and since our clients vary from very small to very large companies, and as our groups are unevenly represented (with most of our clients being in the Financial industry), perhaps a better representation of what a company in a certain industry could expect in terms of vulnerability counts is the average number of unique vulnerabilities per asset (by industry).

The Financial services industry had the highest number of unique vulnerabilities discovered per asset in the last year with 93 vulnerabilities on average. They were followed by Healthcare with 13, Technology with 7, and Other with 5 vulnerabilities per asset on average (Figure 15).

Given that large companies may have thousands of assets, these numbers can easily translate to hundreds of thousands of vulnerabilities per scan, even after discarding duplicates. In the past, we have seen individual scans with close to 300,000 unique vulnerabilities.

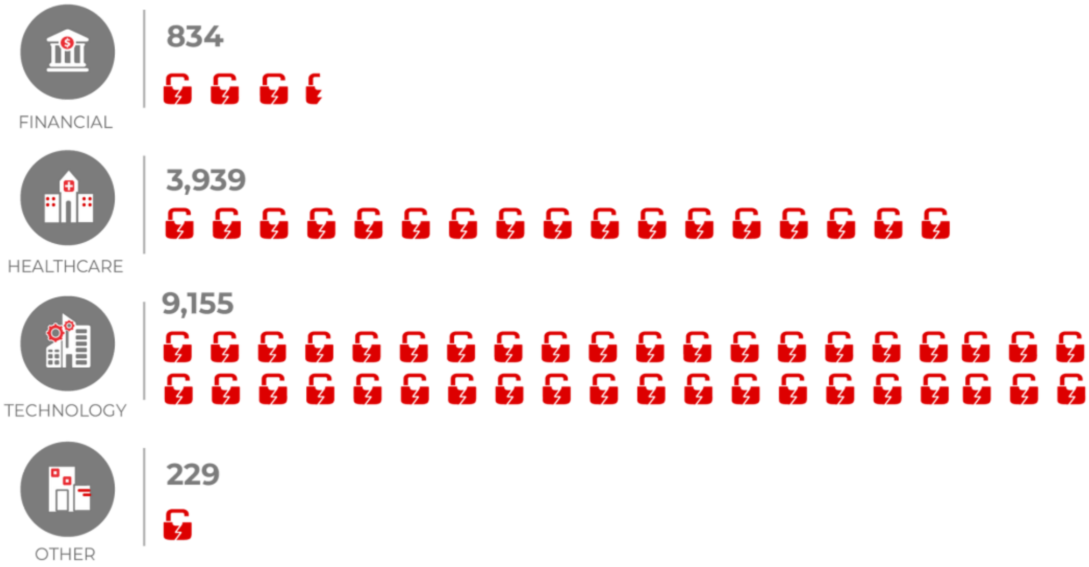


Figure 14:
Median number of unique vulnerabilities discovered per client by industry in the last year.

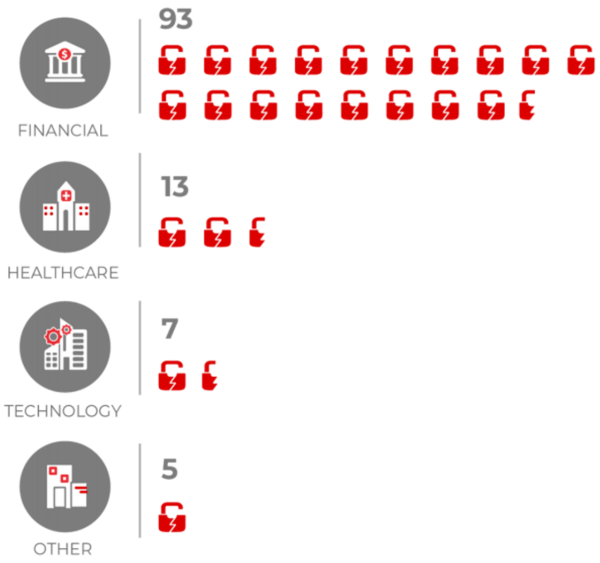


Figure 15:
Average number of unique vulnerabilities discovered per asset by industry in the last year.

NopSec Client Data: Most Frequently Found Vulnerabilities

Our data sample contains 10,735 unique CVE IDs. These were collected over the duration of one year. Distribution of CVE-year for these CVE IDs is shown below (Figure 16).

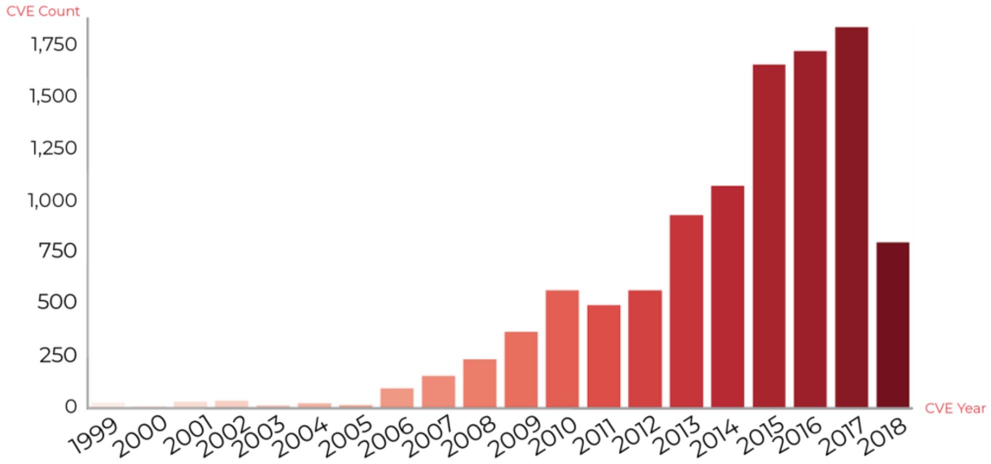


Figure 16:
CVE-year distribution for unique CVEs found in our clients' data in the last year (June 15, 2017 - June 15, 2018). Most are recent vulnerabilities (published in the last three years), but some go as far back as 1999.

These are the top 20 most frequently found vulnerabilities across all NopSec clients. A number of interesting insights can be gleaned from this top 20 and are mentioned following the list.

Top 20 Most Frequently Found Vulnerabilities:

1. **CVE-2000-1200 -** A true classic, this is more of a ‘feature’ or misconfiguration than a vulnerability. Allows anonymous unauthenticated attackers to list domain users or local users.
2. **CVE-2015-2808, CVE-2013-2566 -** Issues with the RC4 algorithm as used in TLS and SSL protocols.
3. **CVE-1999-0520, CVE-1999-0519 -** Inappropriate access controls or insecure passwords on NETBIOS/SMB shares. Like the first most common ‘vuln’, this is more of a configuration issue than a vulnerability. By current standards, it would not receive a CVE number and would not be considered a vulnerability.
4. **CVE-2004-2761 -** MD5 collisions. Another protocol/algorithm-level issue - not really a software bug that can be patched.
5. **CVE-2018-4877, CVE-2018-4878 -** Two Adobe Flash Player use-after-free vulnerabilities in the Primetime SDK, which can lead to arbitrary code execution.
6. **CVE-2018-2677 and 19 other CVEs -** A large bundle of Oracle Java vulnerabilities in a variety of Java components with CVSS scores ranging from 2.x to 6.x. Oracle tends to bundle large groups of fixes together, resulting in many scanners detecting 20 or more CVEs as a single, patchable issue.

7. **CVE-2018-4919, CVE-2018-4920 -** Two more Adobe Flash Player vulnerabilities resulting in arbitrary code execution. Both are scored a 10.0 on the CVSS V2 scale.
8. **CVE-2014-3566 -** SSLv3 uses nondeterministic CBC padding, making it unsuitable for protection against MITM attacks. Better known by the "POODLE" moniker.
9. **CVE-2017-11506 -** Some versions of Nessus scanners and agents don't verify TLS certificates when linking to Tenable.io, which creates an opportunity for MITM attacks.
10. **CVE-2015-4000 -** Issue with TLS 1.2 protocols and earlier that potentially allows to downgrade cipher selections. Also known as the "Logjam" issue.
11. **CVE-2004-0230 - I**ssue with TCP when using a large Window Size, allowing attackers to guess sequence numbers. This creates an opportunity to inject TCP RST packets, effectively causing a denial-of-service.
12. **CVE-2013-1609, CVE-2014-0759, CVE-2014-5455 -** Services on a remote Windows host were found to use an unquoted service path containing at least one whitespace. This creates an opportunity for local privilege escalation by inserting an executable file in the path. This is a generic test for most scanners that will return any services suffering from this issue.
13. **CVE-2018-2826 and 14 other CVEs -** Another Oracle Java vulnerability rollup, affecting a variety of components. Issues range from 2.x to 5.x on the CVSSv2 scoring scale.
14. **CVE-2008-5161 -** Issue in many products' implementation of SSH protocol error handling, making it possible for remote

attackers to recover plaintext from ciphertext in some scenarios.

15. **CVE-2016-6329, CVE-2016-2183 -** Various cryptographic attacks resulting in plaintext data recovery. Affects OpenVPN in some cases and DES/3DES in various protocols and products.
16. **CVE-2018-2932 and 5 other CVEs -** Various vulnerabilities in Adobe Flash Player on Windows.
17. **CVE-2018-0851 -** Memory corruption vulnerability in how the Windows scripting engine handles objects in memory, allowing arbitrary code execution in Internet Explorer and Microsoft Edge browsers.
18. **CVE-2018-0772, CVE-2018-0762 -** Remote code execution vulnerabilities in Internet Explorer.
19. **CVE-2018-0852, CVE-2018-0850 -** Arbitrary code execution and privilege escalation vulnerabilities in Microsoft Outlook.
20. **CVE-2018-0764, CVE-2018-0786 -** Two mid-severity (5.0) vulnerabilities in the .NET Framework

It might be surprising to hear that only just over half of these vulnerabilities are software bugs. Six are broken cryptographic algorithms or protocols implementing encryption in-transit. The remainder are configuration issues. This is significant, because only half of the top 20 vulnerabilities are issues that can be fixed with a patch. The rest require disabling, decommissioning or reconfiguring systems or software.

Of the software bugs on the top 20 list, there are no surprises. Java, Flash and Microsoft Office take their familiar place as some of the most persistent vulnerabilities present in enterprise environments. Internet Explorer and Microsoft's .NET framework also make appearances.

Regarding the types of vulnerabilities in the top 20, the presence of insecure cryptography makes man-in-the-middle attacks the top attack vector on this list. Remote and arbitrary code execution represent most of the software bugs. Information disclosure, privilege escalation and denial-of-service round out the vulnerability attack vectors.

NopSec Client Data: CVSS Score Analysis

The distribution of CVSS V2 Base Scores (Figure 17) skews to the side of High Severity, with a median CVSS Base Score of 7.2. This means that more than half of all detected vulnerabilities are rated High Severity per CVSS V2 guidelines).

In fact, 53% of all vulnerabilities on our clients' systems were scored as 7 or higher (high severity) and 26% as 9 or higher (critical). This is an even higher portion of high and critical severity vulnerabilities than what would be expected based on the analysis of the entire NVD dataset above (Figure 18). These results can be attributed to the fact that most scanner plugins are optimized to detect medium and high severity vulnerabilities (and not necessarily low severity ones.

The verticals particularly vulnerable appear to be the Financial industry with median base score of 7.6. The median scores for

Health, Technology, and Other groups are 5.1, 5.0, and 5.0, respectively.

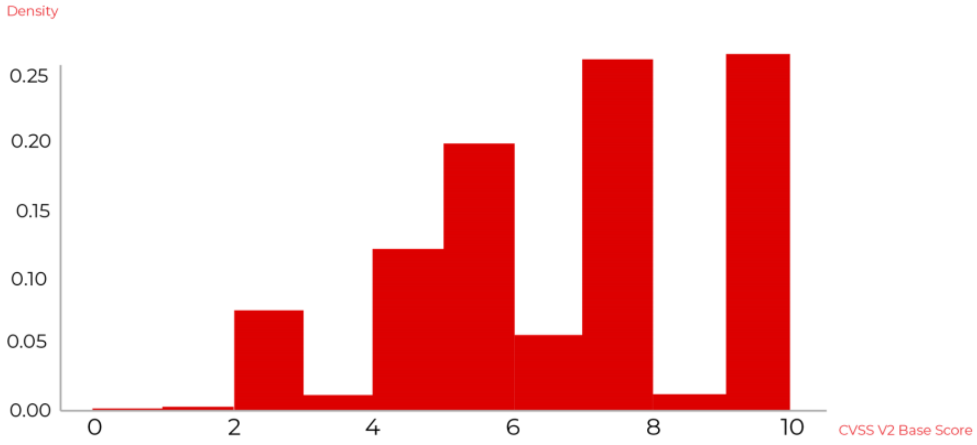


Figure 17: *Distribution of NopSec client vulnerability CVSS Base Scores.*

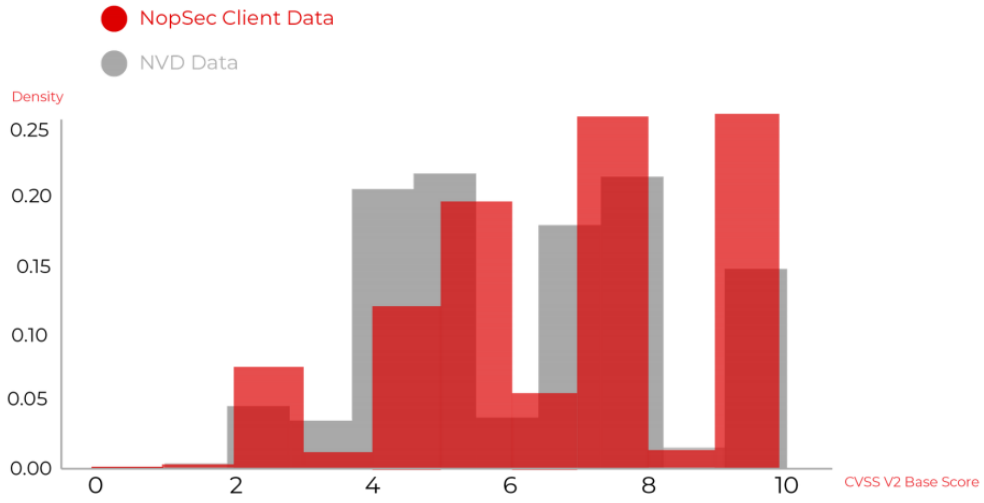


Figure 18: *Distribution of NopSec client vulnerability CVSS Base Scores (red) overlaid on distribution of CVSS Base Scores for all documented CVE IDs in NVD (grey).*

NopSec Client Data: CVE-Correlated Exploit and Malware Analysis

As with the NVD data, for simplicity, we focus on two well-known open sources of exploit code - Exploit Database and Metasploit.

In addition, we aggregate information on malware, ransomware, remote access trojans and exploit kits into the 'malware and exploit kits' category.

Nopsec observed a higher percentage of CVEs detected in our client data as having functional exploits (Metasploit) and malware linked to them compared to the entire NVD database CVE population.

This isn't especially surprising. Most of these CVEs are recent (and we showed that malware incorporation rates are on the rise) and have, on average, higher CVSS scores.

Upon analysis, we found that 12.5 percent have code in Exploit Database, 2.6 percent have Metasploit modules, and 6.4 percent have malware and exploit kits linked to them (Figure 19). These are still all low percentages compared to over 50 percent identified as high severity based on CVSS score.

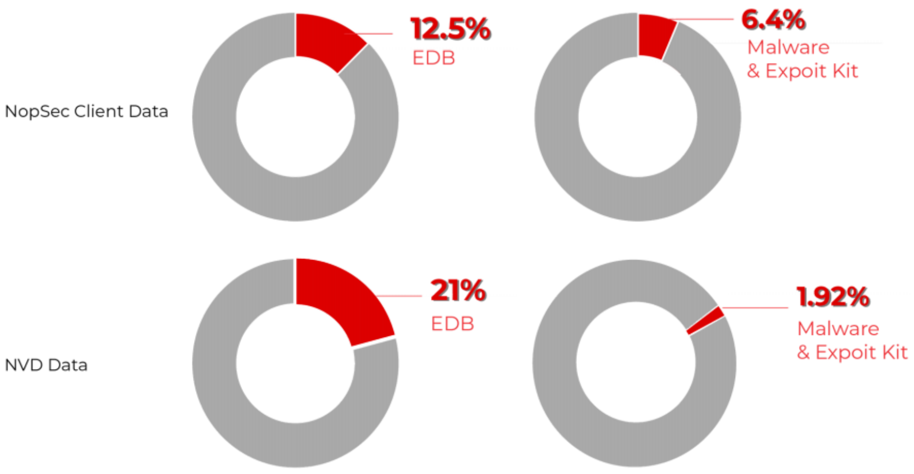


Figure 19:
Portion of CVEs found in our client data that are in EDB or have malware and exploit kits association (top), and portion of of all CVEs that are in EDB or have associated malware (bottom).

Moreover, when looking at CVSS Base Score distribution for those CVEs that have malware and exploit kits associated with them, we see many vulnerabilities that are not ranked high severity based on CVSS V2 (Figure 20). About 35% of CVEs identified in our clients' data with associated malware and exploit kits have a CVSS V2 score under 7 – ranking them as medium or low.

Again, choosing solely based on high CVSS score to remediate vulnerabilities would lead to many false negatives.

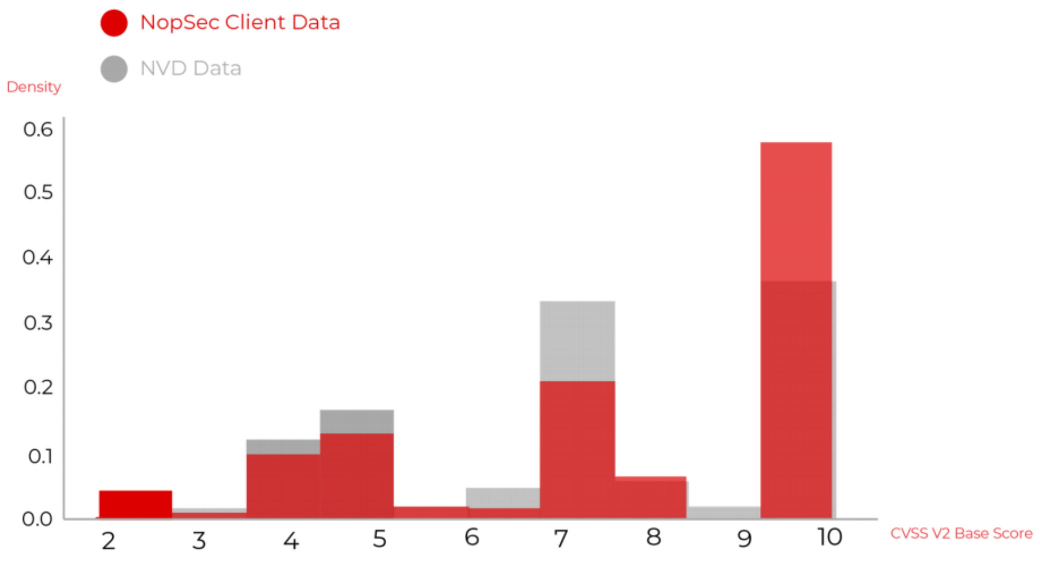
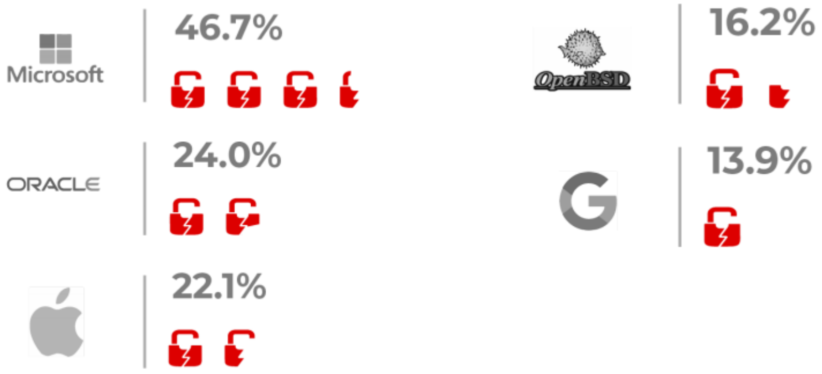


Figure 20:
NopSec client data: CVSS Base Score distribution for CVEs with malware and exploit kits associated to them resembles the Base Score distribution for the entire CVE population - many dangerous CVEs have medium and low CVSS Base Score.

NopSec Client Data: Vendor Analysis

Across all industries and clients, the most commonly affected vendor was Microsoft (~47% of all vulnerabilities), followed by Oracle (24%), and Apple (22%) (Figure 21). While these counts, as one would expect, reflect most widely used vendors (vendors with most products and presence across industries), there are notable differences between industries.

__Figure 21:__
Vendors with highest percentage of detected vulnerabilities across all clients and industries.



For example, the vast majority of vulnerabilities found in Financial companies (almost 70%) are Microsoft vulnerabilities, while as in Healthcare and Other groups Microsoft does not even make it to top 10 vendors by percent of all detected vulnerabilities (Figure 22).

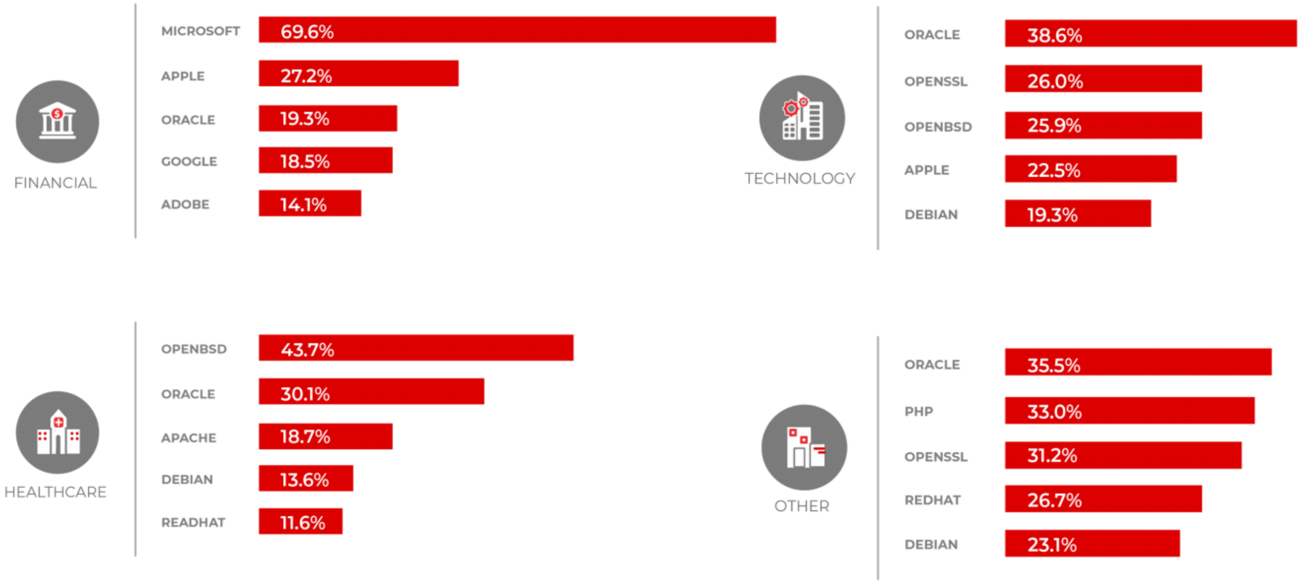


***Figure 22:** Vendors accounting for highest portion of detected vulnerabilities for clients in the Financial industry, Healthcare group, Technology group, and Other group.*

NopSec Risk Prediction

Overview

Unpatched known vulnerabilities may represent a significant security risk. Software vendors are under pressure to quickly provide patches and users are urged to install them as soon as they are available. Since remediation requires considerable efforts, and as scans may reveal hundreds of thousands of vulnerabilities, it is desirable to prioritize the remediation of vulnerabilities that are likely to be targeted and exploited in the wild.

In fact, based on our data, only a small fraction of vulnerabilities documented in the NVD have functional exploits published, and even fewer are exploited in real-world attacks. These account for fewer than than 5% of all vulnerabilities, in agreement with multiple published data on the topic.

How can we then assign a risk to a newly released or existing vulnerability to account for this? Can we use historic data to predict if a newly published vulnerability will be exploited in the wild?

Given the serious shortcomings of the CVSS (base) score described above, NopSec aims to provide an improved prioritization of vulnerabilities through a more accurate assessment of real-world risk each vulnerability carries.

NopSec achieves this by incorporating additional vulnerability information from many threat feeds and social media into a machine learning model to obtain the probability that a vulnerability will be used in real-world attacks. We use this as a proxy of risk rather than a simple existence of publicly available exploit code (for example, in the Exploit Database) as that may not be sufficient for real-world cyberattacks.

Despite it potentially being a facilitating factor, a proof-of-concept (PoC) exploit code may require significant additional research and validation by a highly skilled and motivated actor in order to be weaponized and used successfully for malicious purposes.

This assumption is corroborated by the fact that, based on our aggregated data sources, about 24 percent of CVEs published since 2002 have had some sort of exploit code published till now, but only around 2 percent have been highly weaponized in malicious code.

The basis for the risk factor which NopSec's Unified VRM assigns to an asset, a group of assets, a business module, or the entire business itself is what we refer to as a technical risk score. Technical risk score is a value between zero and one quantifying the likelihood that a vulnerability will be used in malware/exploit kits/targeted attacks.

The starting point for technical risk score calculation is a publicly disclosed vulnerability uniquely identified by a CVE ID. Asset and/or business - level risk score is then derived from the technical risk score by taking into account not only all the vulnerabilities found on the asset(s), but also client's unique environment and asset importance (confidentiality, integrity and availability factors as provided by the client).

Risk Scoring Algorithm (Machine Learning)

This task of assigning the probability of attack incorporation to a CVE may be framed as a supervised machine learning (classification) problem. First, we provide a set of training examples (CVE-IDs/vulnerabilities) with some attributes/features:

- X (e.g., CVSS score, vulnerability age, vulnerability type, vulnerability description, existence of exploit code, social media presence, etc.)
- Y - a set of corresponding targets/labels/the ground truth. (e.g., "no malware/exploit kits/targeted attacks" versus "linked to malware/exploit kits/targeted attacks").

Then, we find the best possible function/model that maps X to Y. From here, we can apply this model to a separate testing set.

The core idea of the model in this context is to use historic data about vulnerabilities with known labels to learn the best (potentially highly nonlinear) combination of their features that can be used to distinguish those that have had malware associated to them from those that have not.

By then applying this model to a previously unseen (test) dataset, we can estimate its predictive power for future vulnerabilities. Our data are continuously updated to reflect all available information since the vulnerability landscape is ever-changing.

Model parameters are tuned on an additional cross-validation set (taken from the training data), and model performance is evaluated on the test set - making sure that there is no drastic decrease in performance going from training to test. This ensures that the model generalizes well to new examples (new vulnerabilities) and does not ‘overfit’. A model that contains too many parameters could fit the training data very well by modeling the noise, but fail to perform well on any new data.

Since the vast majority of vulnerabilities will never be associated with malware, exploit kits, or targeted attacks, looking at accuracy as a measure of performance may be problematic. Simply predicting all new vulnerabilities to have no malware association would lead to very high accuracy. Hence, rather than measuring accuracy, we look at the types of mistakes the model is making - the false positives and the false negatives (and try to maximize the related measures of precision and recall). As a reminder, by false positives here we mean all those vulnerabilities that will not have malware association but that our algorithms mistakenly labels as such. False negatives would be those vulnerabilities that will have malware association, but mistakenly get labeled as the opposite. We choose the model parameters based on the performance on the validation set and the final model performance is measured on the test set. This way, when a new vulnerability comes in, we can assign risk to it and predict whether if it will be exploited in the wild with confidence based on the performance of the optimized model on the test data.

Model Features and Natural Language Processing

The basic unit for analysis is a vulnerability as defined by a CVE-ID. As mentioned before, the target of our classification problem is existence of CVE-associated malware/exploit kits/targeted attacks. For each vulnerability, we extract a list of attributes/features that we suspect to be informative in determining whether this vulnerability is risky or not.

For each vulnerability we utilize hundreds of features: we combine the intrinsic information about a vulnerability coming from NVD (attributes such as vulnerability age, time since it was last modified, six CVSS V2 vectors, impact, exploitability and base scores, vendor and product information, references, and description) with data from multiple threat feeds/exploit databases/penetration testing frameworks and social media.

For each vulnerability, some of the numerical attributes we use include vulnerability age (relative to disclosure date), time since last modification, CVSS base score and impact and exploitability subscores, number of different vendors and products it affects, vendor and product prevalence/popularity - how many vulnerabilities these vendors/products have had in the past, number of mentions in various exploit sources and social media.

In addition to these numerical features, motivated by our analyses that show that certain vendors and products have higher than average portions of vulnerabilities linked to malware, we also include specific vendor and product names as features (for example, a feature from this category may be able to answer a question like this: “is this vulnerability a Microsoft vulnerability?").

Finally, as explained in the NLP section above and motivated by our finding that some words or combinations of words appear to be indicative of increased risk, we use training data to extract words and n-grams from vulnerability descriptions and use them as features in our model (for example, a feature from this category may answer a question such as “does this vulnerability description include remote code execution?").

Conclusion

Risk isn’t straightforward. This is an issue that security professionals grapple with on a daily basis. As we discussed, relying on mere CVSS scores to understand the threat posed by a vulnerability isn’t enough. Relying solely on the existence of exploit code is also not enough, as data imply that most of exploit code never gets weaponized and used in real world attacks for malicious purposes. There are other factors at play, all arguably equally or more significant.

There's the context in which a vulnerability exists. A machine running Windows XP SP2 might be filled with easily-exploitable holes, but if it's sat in a corner, disconnected from the Internet, and operating a SCADA controller, it is probably best to turn your attention elsewhere.

Similarly, focusing on the most severe of vulnerabilities is folly, especially if they're not presently exploitable.

A way forward is to combine our current knowledge about existing vulnerabilities – from threat-intelligence sources and vulnerability databases – and augment it with cutting-edge machine learning technology, which takes advantage of historical observable trends in vulnerability disclosure and research to find patterns and make predictions about the future.

Only then will risk start to make sense.

Find out how NopSec's Unified VRM can help you think like a hacker and stay ahead of the trends. Visit www.nopsec.com or email hello@nopsec.com for additional information or to request a demo.

About NopSec

NopSec operates with one mission: to help people make better decisions to reduce security risks. Our team is passionate about building technology to help customers simplify their work, manage security vulnerability risks effectively, and empower them to make more informed decisions. Our software-as-a-service approach to vulnerability risk management offers an intelligent solution to dramatically reduce the turnaround time between identification of critical vulnerabilities and remediation.

NopSec helps security professionals simplify their work, effectively manage and prioritize vulnerabilities, and make better informed decisions.

NopSec's Unified VRM is an innovative threat and vulnerability management solution that addresses the need for better prioritization and remediation of security vulnerabilities in a single platform.

NopSec Inc. • www.nopsec.com • info@nopsec.com