2014 DATA BREACH INVESTIGATIONS REPORT

92%

THE UNIVERSE OF THREATS MAY SEEM LIMITLESS, BUT 92% OF THE 100,000 INCIDENTS WE'VE ANALYZED FROM THE LAST 10 YEARS CAN BE DESCRIBED BY JUST NINE BASIC PATTERNS.

Conducted by Verizon with contributions from 50 organizations from around the world.

2014 DBIR Contributors

(see Appendix C for a detailed list)

CONTENTS

Questions?
Comments?
Brilliant ideas?

We want to hear them. Drop us a line at dbir@verizon.com, find us on LinkedIn, or tweet @VZdbir with the hashtag #dbir.

INTRODUCTION

50
CONTRIBUTING GLOBAL ORGANIZATIONS

1,367
CONFIRMED DATA BREACHES

63,437
SECURITY INCIDENTS

95
COUNTRIES REPRESENTED

Welcome to the 2014 Data Breach Investigations Report (DBIR).[1] Whether you're a veteran reader who's been with us since our initial publication back in 2008 or a newbie to our annual data party, we're sincerely glad you're here. We hope that this year's submission will improve awareness and practice in the field of information security and support critical decisions and operations from the trenches to the boardroom.

For DBIR veterans, a cursory look at the table of contents will reveal some significant changes to the report structure you've gotten used to in years past. Rather than our signature approach organized around actors, actions, assets, timelines, etc., we've created sections around common incident patterns derived directly from the data itself (more on that later). Within each of those patterns, we cover the actors who cause them, the actions they use, assets they target, timelines in which all this took place, and give specific recommendations to thwart them. The drive for change is three-fold: first, we realized that the vast majority of incidents could be placed into one of nine patterns; second, we can (and did) draw a correlation between these incident patterns and industries; and third, we wanted to challenge ourselves to look at the data with a fresh perspective. The ultimate goal is to provide actionable information presented in a way that enables you to hash out the findings and recommendations most relevant to your organization.

We all know that data doesn't grow on trees, and we must express our gratitude to the 50 organizations that contributed to this report, representing public and private entities from around the globe. We're proud to work with these organizations and feel that what you're now reading is proof of the benefits of coordinated incident data sharing. For the full list of 2014 DBIR contributors, check out Appendix C.

The dataset that underpins the DBIR is comprised of over 63,000 confirmed security incidents — yep, over Sixty-Three Thousand. That rather intimidating number is a by-product of another shift in philosophy with this year's report; we are no longer restricting our analysis only to confirmed data breaches. This evolution of the DBIR reflects the experience of many security practitioners and executives who know that an incident needn't result in data exfiltration for it to have a significant impact on the targeted business.

So prepare to digest what we hope will be some very delicious data prepared for you this year. The Methodology section, normally found near the beginning of the report, is now in Appendix B. We'll begin instead with a review of 2013 from the headlines, then provide a few sample demographics to get you oriented with the dataset. The following section — a summary of our 10 years' of incident data — might just be our favorite. (but please don't tell the other sections that). We'll then provide analysis of the aforementioned incident classification patterns and end with some conclusions and a pattern-based security control mapping exercise. So let's get started!

2013 YEAR IN REVIEW

The year 2013 may be tagged as the "year of the retailer breach," but a more comprehensive assessment of the InfoSec risk environment shows it was a year of transition from geopolitical attacks to large-scale attacks on payment card systems.

2013 may be remembered as the "year of the retailer breach," but a comprehensive assessment suggests it was a year of transition from geopolitical attacks to large-scale attacks on payment card systems.

JANUARY

January saw a series of reports of targeted attacks by what were probably state-sponsored actors. The Red October cyber-espionage campaign was exposed and responsible for targeting government agencies and research institutions globally, but in Russian-speaking countries in particular. Intelligence then connected it to actors using the Elderwood framework, and also a complex series of attacks beginning with a "watering hole" attack on the Council on Foreign Relations web site (cfr.org) that began on Boxing Day 2012. Meanwhile, the Izz ad-Din al-Qassam Cyber Fighters (QCF) were almost a month into Phase II of Operation Ababil Distributed Denial of Service (DDoS) attacks on U.S. financial services companies.

FEBRUARY

The segue into February was provided by The New York Times and the Wall Street Journal, with new reports of targeted cyber-espionage. And Sophos reported a new Citadel-based Trojan crafted to attack Point-of-Sale (POS) systems using a Canadian payment card processor. We would soon learn that www.iphonedevsdk.com became a watering hole, using a surprise attack on Java late in the month. Most InfoSec professionals well remember February as the month Mandiant (now FireEye) released its superb APT1 report. February was also the start of reports of data breaches from large enterprises, courtesy of the aforementioned iPhoneDevSDK: Facebook, Twitter, Apple, and Microsoft were all victims. Noteworthy retailer POS data breaches were reported by Bashas' and Sprouts, two discrete grocery chains in the U.S. Southwest. Bit9 reported a data breach that began in July 2012, attacking its code-signing infrastructure.

MARCH

Fifty million Evernote users remember that March was the month they were forced to change their passwords. On March 20, the Republic of Korea suffered a large-scale cyber-attack that included disk corruption. We remain skeptical that the Cyberbunker-CloudFlare-Spamhaus DoS attack almost broke the Internet at the end of March. Group-IB reported "Dump Memory Grabber" (a.k.a. BlackPOS), a new POS Trojan that would go on to make headlines when news broke of Target Stores' breach in December.

This section is a compilation of the weekly INTSUM lead paragraphs posted to our blog and is 100% based on open source intelligence (OSINT). We maintain a very strong policy against identifying Investigative Response clients, and mentions of organizations in this section in no way imply that we conducted an investigation involving them or that they are among the victims in our dataset.

APRIL
In April, another U.S. grocery retailer, Schnucks, reported a POS data breach. The Syrian Electronic Army (SEA) did some damage when it hijacked the Associated Press' Twitter account, sending a tweet reporting an explosion at the White House and causing a spasm on Wall Street. Operation Ababil continued, but OSINT cannot support attributing DoS attacks on several European banks to the QCF.

MAY
Cyber-espionage continued in May, with reports from QinetiQ and the U.S. Army Corps of Engineers. The SEA hijacked the Twitter accounts of both The Guardian and The Financial Times. A watering hole attack targeted nuclear weapons researchers in the U.S. for cyber-espionage, probably from China. More cyber-espionage campaigns reported in May included Operation Hangover, targeting Pakistan; Safe, targeting Mongolia; and operations by the Sunshop actors against Tibetan activists. The U.S. Department of Justice shut down Liberty Reserve, the go-to bank for cyber-criminals.

JUNE
Early in June, Raley's, yet another U.S. grocer with stores in California and Nevada, reported its payment card systems were breached. NetTraveller, a global cyber-espionage campaign targeting diplomats in countries with interests not aligned with China occurred. A day later, The Guardian published the first intelligence leaked by Edward Snowden… and then InfoSec intelligence became the "All-Snowden-All-the-Time" channel.

JULY
July's largest retailer data breach was reported by Harbor Freight, a U.S. tool vendor with 445 stores – nearly 200 million customers and we still don't know how many records were compromised. The QCF initiated Phase IV of Operation Ababil. The SEA breached Viber, Tango, and the Daily Dot. The U.S. Department of Justice indicted four Russians and one Ukrainian for high-profile data breaches, including Heartland and Global Payments.

AUGUST
In August, the SEA hijacked the Twitter accounts of CNN, The Washington Post, Time Magazine, SocialFlow, and both The New York Times and New York Post. Attendees of the G-8 Summit in St. Petersburg, Russia, were targeted for cyber-espionage by the Calc Team actors.

SEPTEMBER
In September, Vodafone notified two million customers their personal and financial information had been breached. Espionage reported in September involved the EvilGrab Trojan and separately, the Hidden Lynx actors who seem to engage in both espionage and cybercrime. New intelligence linked the Bit9 attack from February with Operation Deputy Dog, Hidden Lynx, and watering hole attacks on Japanese financial institutions. At the end of the month Brian Krebs began his reports on intelligence extracted from ssndob[dot]ms. The site was home to data stolen from some of America's largest data brokers: Lexis-Nexis, Kroll, and Dun & Bradstreet. Cryptolocker made its first appearance in September, extorting money from victims that were willing to pay to decrypt their essential files.

OCTOBER
On October 3, Adobe announced its systems had been breached; eventually 38 million accounts were identified as affected. Intelligence connected this to the ssndob[dot]ms actors. Nordstrom, the luxury U.S. department store, discovered skimmers on some of its cash registers. Two of 2013's big wins also occurred in October: Dmitry "Paunch" Fedotov, the actor responsible for the Blackhole exploit kit, was arrested in Russia, and Silk Road, an online fraud bazaar, was taken down.

NOVEMBER
The proverbial calm before the storm, November was fairly quiet. Banking malware evolved with reports of Neverquest and another version of IceIX. BIPS, a major European bitcoin payment processor, was the victim of one of the largest bitcoin heists recorded up to that point in time.

DECEMBER
The last significant entry under cyber-espionage for 2013 was the targeting of foreign ministries in European countries by Operation Ke3chang. The Washington Post reported its second breach of the year. And then InfoSec intelligence became the "All-Target-All-the-Time" channel. Although the breach of this major U.S. retailer was a little more than half the size of Heartland and three-fourths the size of TJX, it's vying to become the event for which 2013 will always be remembered.

Questions?
Comments?
Brilliant ideas?

We want to hear them. Drop us a line at dbir@verizon.com, find us on LinkedIn, or tweet @VZdbir with the hashtag #dbir.

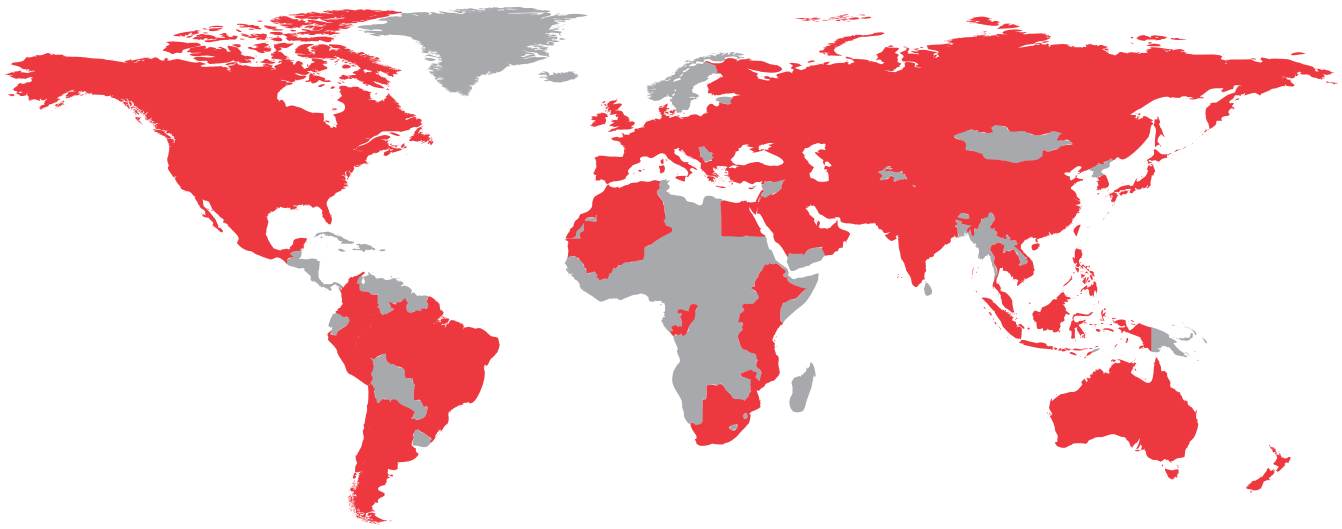
VICTIM DEMOGRAPHICS

Readers of the DBIR frequently approach us with two important questions. How generally representative are the findings of this report? Are these findings relevant to my organization? To help get you oriented with this year's report, let's see what the data has to show us.

The 2013 DBIR featured breaches affecting organizations in 27 countries. This year's report ups that tally by 350%, to 95 distinct countries (Figure 1). All major world regions are represented, and we have more national Computer Security Incident Response Teams (CSIRTs) than ever before. Our ability to compare global trends has never been higher.

But it's not quite that simple. The charter, focus, methods, and data differ so much between CSIRTs that it's difficult to attribute differences to true variations in the threat environment.[2] However, regional blind spots are getting smaller thanks to our growing list of contributors (see Appendix C), and we're very happy with that.

Figure 1.
Countries represented in combined caseload



Countries represented in combined caseload (in alphabetical order): Afghanistan, Albania, Algeria, Argentina, Armenia, Australia, Austria, Azerbaijan, Bahrain, Belarus, Belgium, Bosnia and Herzegovina, Botswana, Brazil, Brunei Darussalam, Bulgaria, Cambodia, Canada, Chile, China, Colombia, Congo, Croatia, Cyprus, Czech Republic, Denmark, Egypt, Ethiopia, Finland, France, Georgia, Germany, Greece, Hong Kong, Hungary, India, Indonesia, Iran, Islamic Republic of, Iraq, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Kenya, Korea, Republic of, Kuwait, Kyrgyzstan, Latvia, Lebanon, Lithuania, Luxembourg, Macedonia, the former Yugoslav Republic of, Malaysia, Mali, Mauritania, Mexico, Moldova, Republic of, Montenegro, Morocco, Mozambique, Nepal, Netherlands, New Zealand, Oman, Pakistan, Palestinian Territory, Occupied, Peru, Philippines, Poland, Portugal, Qatar, Romania, Russian Federation, Saudi Arabia, Singapore, Slovakia, Slovenia, South Africa, Spain, Switzerland, Taiwan, Province of China, Tanzania, United Republic of, Thailand, Turkey, Turkmenistan, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States, Uzbekistan, Vietnam, Virgin Islands.

Figure 2.
Number of security incidents by victim industry and organization size, 2013 dataset

Industry	Total	Small	Large	Unknown
Accommodation [72]	212	115	34	63
Administrative [56]	16	8	7	1
Agriculture [11]	4	0	3	1
Construction [23]	4	2	0	2
Education [61]	33	2	10	21
Entertainment [71]	20	8	1	11
Finance [52]	856	43	189	624
Healthcare [62]	26	6	1	19
Information [51]	1,132	16	27	1,089
Management [55]	10	1	3	6
Manufacturing [31,32,33]	251	7	33	211
Mining [21]	11	0	8	3
Professional [54]	360	26	10	324
Public [92]	47,479	26	47,074	379
Real Estate [53]	8	4	0	4
Retail [44,45]	467	36	11	420
Trade [42]	4	3	0	1
Transportation [48,49]	27	3	7	17
Utilities [22]	166	2	3	161
Other [81]	27	13	0	14
Unknown	12,324	5,498	4	6,822
Total	**63,437**	**5,819**	**47,425**	**10,193**

For more information on the NAICS codes [shown above] visit:
https://www.census.gov/cgi-bin/sssd/naics/naicsrch?chart=2012

Next, let's review the different industries and sizes of victim organizations in this year's dataset (Figure 2). The Public sector's astronomical count is primarily a result of U.S. agency reporting requirements, which supply a few of our contributors with a vast amount of minor incidents (more on that later), rather than a sign of higher targeting or weak defenses. Figure 3 filters out the minutiae by narrowing the dataset to only those incidents involving confirmed data compromise. Moving beyond the Public sector outlier, both Figure 2 and Figure 3 show demographics relatively similar to prior years.

Figure 3.
Number of security incidents with confirmed data loss by victim industry and organization size, 2013 dataset

Industry	Total	Small	Large	Unknown
Accommodation [72]	137	113	21	3
Administrative [56]	7	3	3	1
Construction [23]	2	1	0	1
Education [61]	15	1	9	5
Entertainment [71]	4	3	1	0
Finance [52]	465	24	36	405
Healthcare [62]	7	4	0	3
Information [51]	31	7	6	18
Management [55]	1	1	0	0
Manufacturing [31,32,33]	59	6	12	41
Mining [21]	10	0	7	3
Professional [54]	75	13	5	57
Public [92]	175	16	26	133
Real Estate [53]	4	2	0	2
Retail [44,45]	148	35	11	102
Trade [42]	3	2	0	1
Transportation [48,49]	10	2	4	4
Utilities [22]	80	2	0	78
Other [81]	8	6	0	2
Unknown	126	2	3	121
Total	**1,367**	**243**	**144**	**980**

Small = organizations with less than 1,000 employees,
Large = organization with 1,000+ employees

We saw some increases where we added new industry-specific contributors, so pieces of the puzzle are filling in. Certain sectors will always skew higher in the victim count given their attractiveness to financially motivated actors — i.e., those that store payment card or other financial data. But even discounting that, we don't see any industries flying completely under the radar. And that's the real takeaway here — everyone is vulnerable to some type of event. Even if you think your organization is at low risk for external attacks, there remains the possibility of insider misuse and errors that harm systems and expose data.

So, we can't claim to have unbiased coverage of every type and size of organization on the planet (fingers crossed for next year, though!). But we dare say that the majority of readers will be able to see themselves or something that looks enough like them in this sample.

A DECADE OF DBIR DATA

Long-time readers of this report will know that we're not very good at maintaining the status quo. The sources of data grow and diversify every year. The focus of our analysis shifts. The way we visualize data and organize results evolves over time. And with the 2014 DBIR, we're really gonna shake things up.

This section attempts to create an "as-comparable-as-possible" set of findings to previous DBIRs. It "only" includes breaches from 2004-2012, plus the 1,367 incidents for which data compromise was confirmed in 2013.

While this does make it hard to meaningfully compare trends across time, it has the positive effect of shining light into new and shadowy areas each year. The truth of the matter is that we're more interested in exploring and learning than churning out the same 'ol stuff each time just to measure deltas.

That said, measuring deltas has value and we know readers appreciate some level of continuity between reports. Thus, this section attempts to create an "as-comparable-as-possible" set of findings to previous DBIRs. It "only" includes breaches from 2004-2012, plus the 1,361 incidents for which data compromise was confirmed in 2013. It's worth noting that this represents the high mark in ten years of data breaches, and is the first time we've crossed 1,000. (Give a round of applause to all those contributors who keep adding fuel to the bonfire.)

We began writing a lot of commentary for this section, but then changed our minds. Instead, we'll churn out some eye candy for you to chew on as long as you like with only a few general observations from us.

We began writing a lot of commentary for this section, but changed our minds. Instead, we'll churn out some eye candy for you to chew on as long as you like, with only a few general observations from us.

A BRIEF PRIMER ON VERIS AND VCDB

The Vocabulary for Event Recording and Incident Sharing (VERIS) is designed to provide a common language for describing security incidents in a structured and repeatable manner. It takes the narrative of "who did what to what (or whom) with what result," and translates it into the kind of data you see in this report. Because we hope to facilitate the tracking and sharing of security incidents, we released VERIS for free public use. Get additional information on the VERIS community site ; the full schema is available on GitHub. Both are good companion references to this report for understanding terminology and context.
www.veriscommunity.com | github.com/vz-risk/veris

Launched in 2013, the VERIS Community Database (VCDB) project enlists the cooperation of volunteers in the security community in an attempt to record all publicly disclosed security incidents in a free and open dataset.

We leverage VCDB for a few sections in this report, which are clearly marked. Learn more about VCDB by visiting the website below.
vcdb.org

Figure 4.
Number of breaches per threat actor category over time

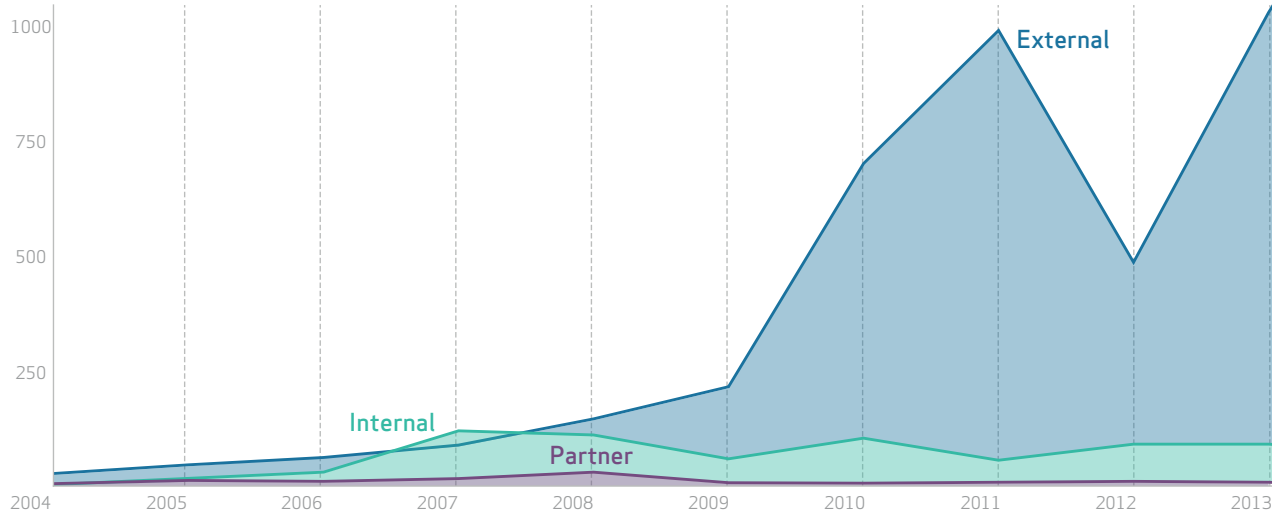


Figure 5.
Percent of breaches per threat actor category over time

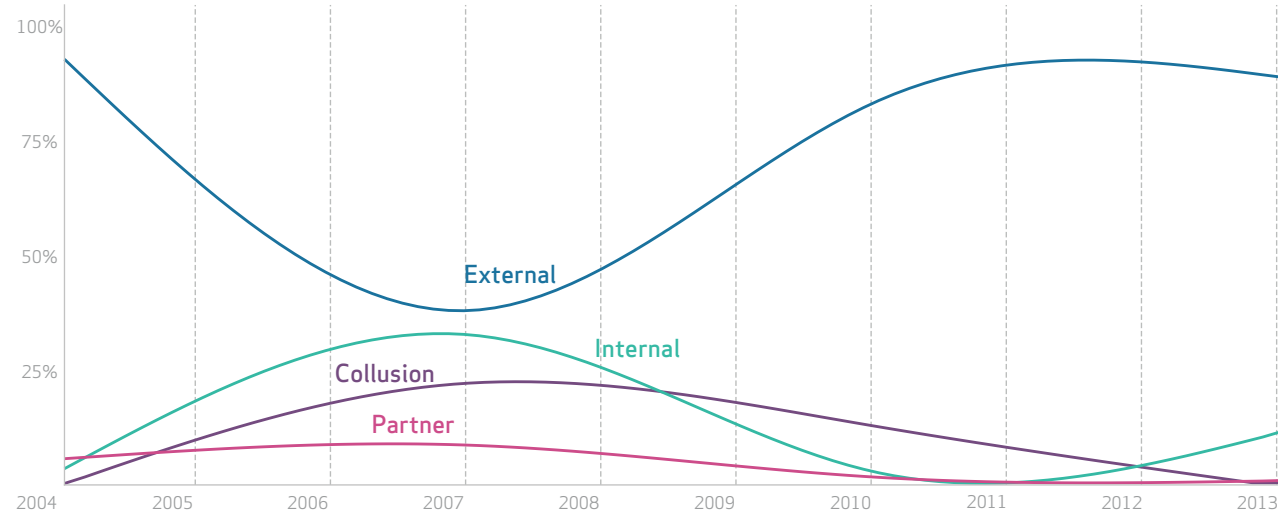


Figure 4 depicts the raw count of breaches attributed to external, internal, and partner actors over the 10-year history of our breach data. Figure 5 shows this as a proportion of all breaches and rearranges the categories to highlight exclusivity and overlap among them. It uses a third-degree polynomial trend line to make it nice and smooth, so we can see the basic behavior over time. Together they help answer our primary questions of interest — which actors perpetrate the most breaches and what's the relative change over time?

BREACHES VS INCIDENTS?

This report uses the following definitions:
Incident: A security event that compromises the integrity, confidentiality, or availability of an information asset.
Breach: An incident that results in the disclosure or potential exposure of data.
Data disclosure: A breach for which it was confirmed that data was actually disclosed (not just exposed) to an unauthorized party.

Since we're letting the visualizations do most of the talking here, we'll only make a few observations and leave the rest for homework.

- Ten years offers some nice min/max/most likely estimates for you modelers out there. Barring 2006-2008, the overall ratio is relatively stable, especially when you consider the dramatic changes in total breaches and sources in scope each year.
- 2007 is the only year showing an insider majority in Figure 4. This is primarily the result of an unusually small Verizon caseload for confirmed breaches and an influx of U.S. Secret Service data from 2006-2008. We essentially crashed two equally sized – but very different – samples together.
- That giant dip for external actors in 2012 seen in Figure 4 coincides with an overall drop in breach count that year, mainly due to fewer large, multi-victim POS intrusion sprees targeting SMBs in the dataset.
- Thanks to several new partners who focus on insider crimes, the proportional trend line for internal swings up over the last couple years while external turns downward. If we removed the polynomial curving, however, you'd see a positive regression for outsiders and a slightly negative one for insiders.

Figure 6.
Percent of breaches per threat actor motive over time

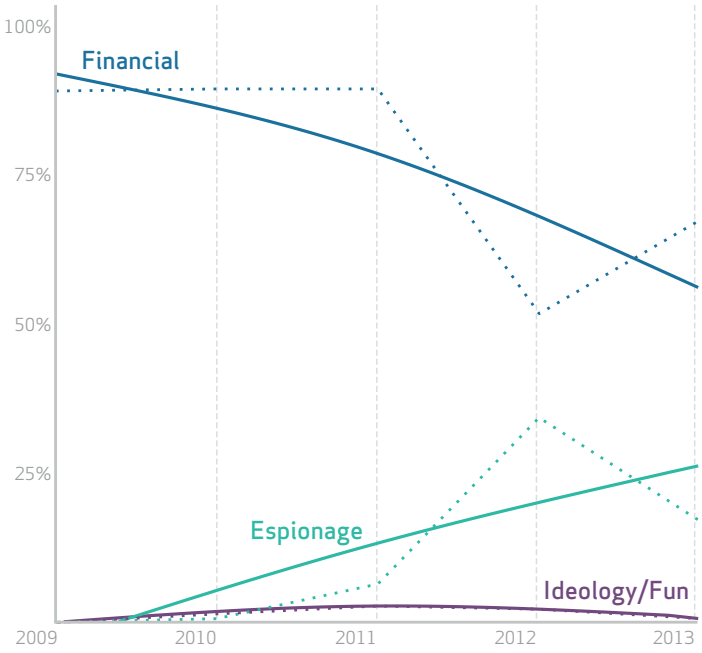
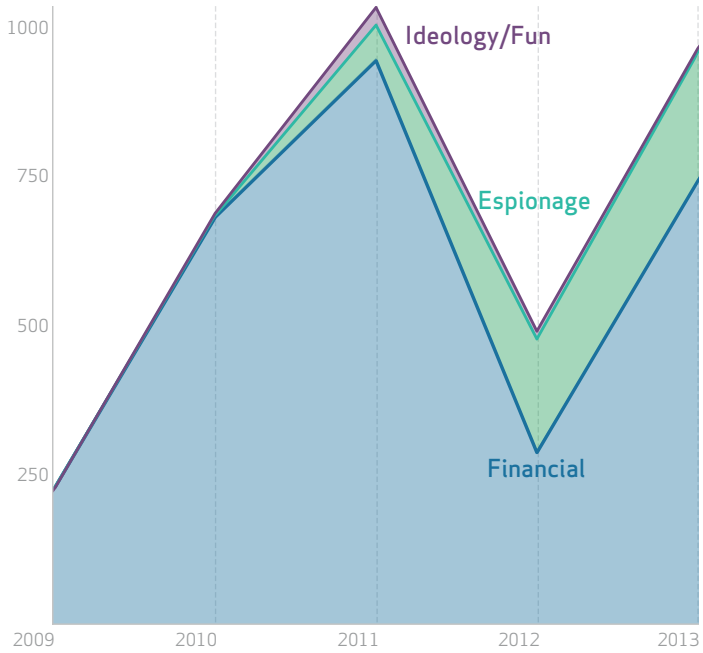


Figure 7.
Number of breaches per threat actor motive over time



Two different views indicating how the motives of threat actors have changed over the last five years appear in Figure 6 and 7. The line chart (Figure 6) gives the relative percentage of the top three motives in our dataset, while Figure 7 uses an area plot of total incident counts.

- We knew espionage had been rising over the last few years, but the trend line chart surprised us by the degree of convergence with financial motives. Will that continue?
- Is this finding merely the result of adding contributors to the DBIR who specialize in espionage, or is money truly diminishing as the prime driver of crime on the Internet? We have an easier time believing the former than the latter, but it certainly makes us want to continue widening our collection of breach data in the future.
- The area plot reminds us that money-motivated breaches still outnumber others by a good margin. To borrow from Pink Floyd, most actors still want to "grab that cash with both hands and make a stash."

Figure 8 has the challenging job of exhibiting 10 years of threat actions leading to data breaches. We experimented with alternate ways to visualize this, but thought the simplicity of this chart worked best. Keep in mind that actions aren't mutually exclusive; several can contribute to an incident (the same goes for actors and assets).

- This chart does a superb job underscoring the value of data sharing. You can see the number of breaches and diversity of threats grow as the DBIR transitions from single sample to a meta study.
- But it's not all because of changes in the sample set. Notice how the hacking and malware categories explode upward in 2009 and social tactics begin to climb in 2010. These have parallel stories in the real world (e.g., better automated attack tools and DIY malware kits), and it's fascinating to see them reflected in the data.

Figure 8.
Number of breaches per threat action category over time

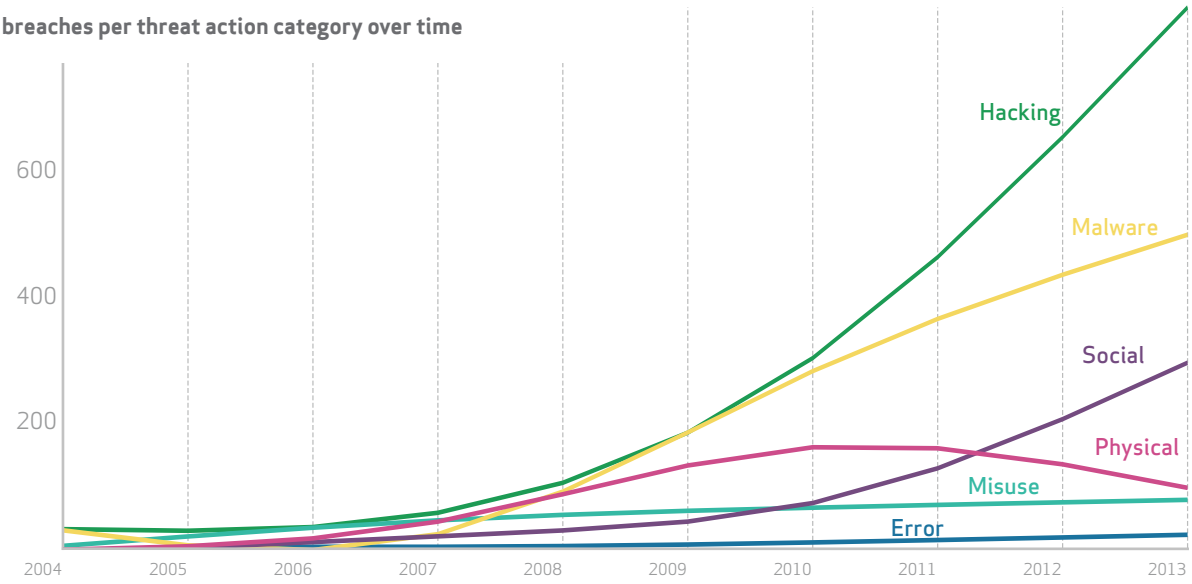


Figure 9.
Top 20 varieties of threat actions over time

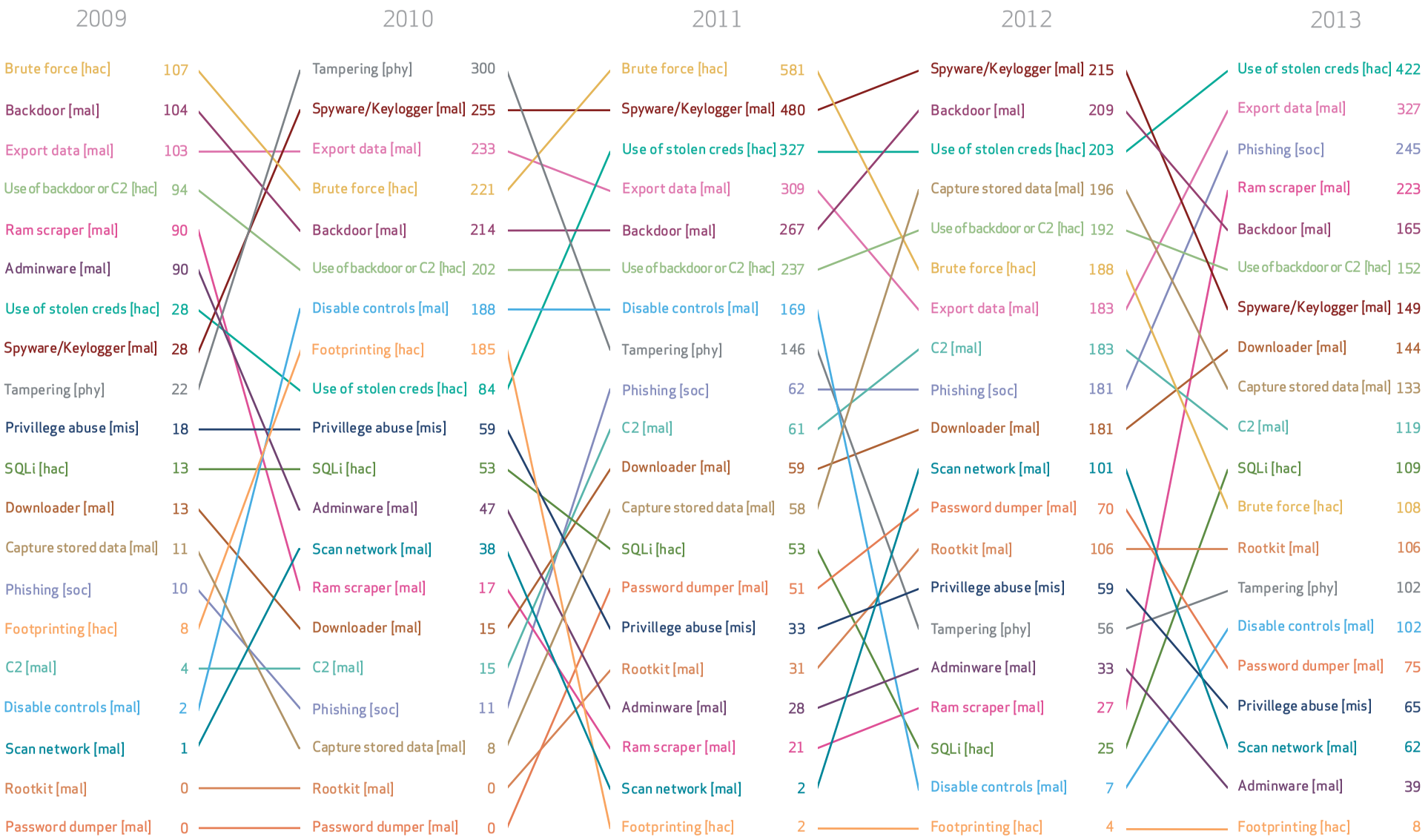


Figure 9 dives deeper into the specific varieties of threat actions observed over the last five years. The overall top twenty across the five-year span is listed in successive columns, and the lines connecting columns highlight how each action changes over time.. To be honest, concise commentary on this visualization may be impossible. Yes, it's incredibly busy, but it's also incredibly information-dense. Let your eyes adjust and then explore whatever strikes your fancy. As an example, follow RAM scrapers through the years. They start at #5 in 2009, drop way down over the next few years and then shoot up the charts to the #4 spot in 2013. We talk about that resurgence in the POS intrusions section of this report. Literally every item in Figure 9 has a story if you care to look for it. Enjoy.

Figure 10.
Percent of breaches per asset category over time

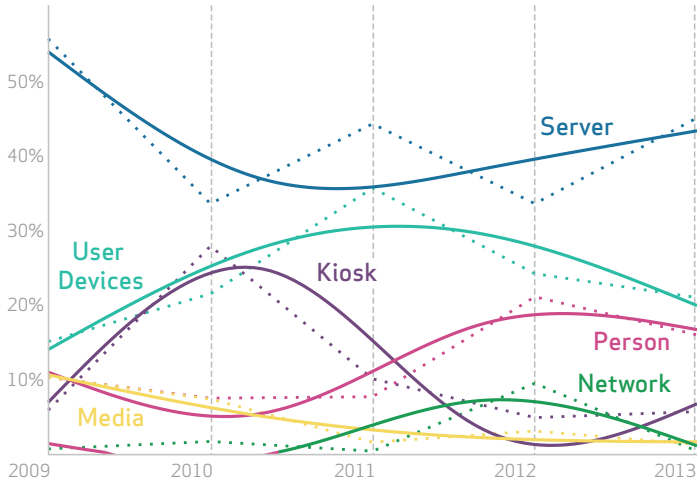
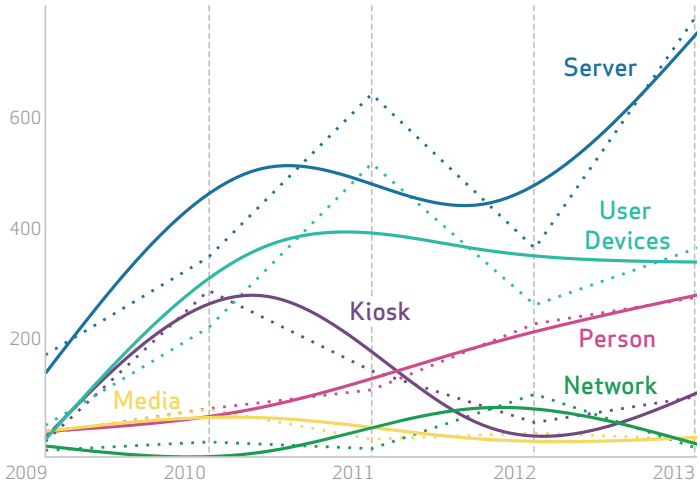


Figure 11.
Number of breaches per asset category over time



Figures 10 and 11 show how the mix of compromised assets has changed over time. It's useful because it reveals the "footprint" of attackers as they travel through the victim's environment in search of data. As defenders, it gives us a sense of what may need extra attention or protection.

- Servers have typically been on top, probably because attackers know that's where the data is stored.
- User devices have been growing over time, probably because they offer an easy foot in the door.
- Media is the only asset category trending down, probably because of an unusually high concentration of (partially-related) cases in 2009 that involved numerous thefts of documents and digital media.
- Many ask why the Network category is so low, given that most of these breaches take place over the network. In view here are specific network devices like routers, switches, etc. Malicious traffic definitely passes through those, but they're not typically compromised during a breach.

Figure 12.
Breach count by data variety over time

It would be hard to give proper treatment to a decade of data theft without covering the varieties of data stolen over that time period. Thankfully, Figure 12's got us covered in that department.

- If you compare these trends with those of actor motives from Figure 6 and 7, you'll see some parallels. Financially-motivated criminals will naturally seek out data that is easily converted to cash, such as bank information and payment cards, while espionage groups target internal corporate data and trade secrets.
- The trend for payment card theft is quite fascinating; it rises quickly to a peak in 2010, and then exhibits a negative slope. There's an uptick in 2013, but it was still the first year in the history of this report where the majority of data breaches did not involve payment cards.
- Authentication credentials are useful in both the criminal underground and the shadowy world of the clandestine, and that demand is reflected here.

Figure 13.
Percent of breaches where time to compromise (red)/time to discovery (blue) was days or less

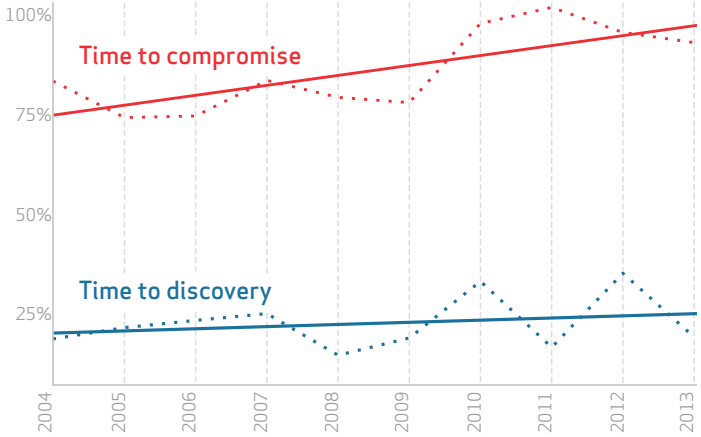
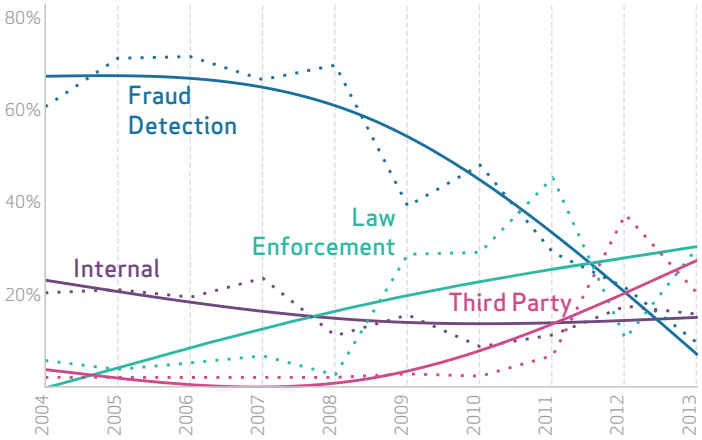


Figure 14.
Breach discovery methods over time



Take a deep, calming breath before diving into this last one; it may result in mental or even bodily harm. In Figure 13, we're contrasting how long it takes the attacker to compromise an asset with how long it takes the defender to discover this. We chose to peg this on "days" to keep things simple and stark (one might also add "sad" to that alliteration).

- Ignore the behavior of the lines for a minute and focus on the wide gap between percentages for the two phases. It smacks us with the fact that the bad guys seldom need days to get their job done, while the good guys rarely manage to get theirs done in a month of Sundays.
- The trend lines follow that initial smack with a roundhouse kick to the head. They plainly show that attackers are getting better/faster at what they do at a higher rate than defenders are improving their trade. This doesn't scale well, people.
- We thought about superimposing "total spending on network monitoring," "number of security products on the market," and "number of Certified Information Systems Security Professionals (CISSPs) in the workplace," but we were concerned it would result in much self-inflicted harm within the security community. And we'd much rather you guys and gals stick around and help us fix this.

Having dealt that last blow regarding timelines, readers familiar with the traditional flow of the DBIR may expectantly hear that Mortal Kombat imperative of "Finish Him!" in their heads as we head into discussion of breach discovery methods. But there will be no triumphant "Fatality!" announcement here; we're going to show mercy instead and end on a positive note.

- We're thrilled to see that internal discoveries outnumber external fraud detection for the first time in DBIR history!
- It's great that law enforcement is steadily getting better and better at detecting breaches and notifying victims!
- Unrelated third parties, like CSIRTs and threat researchers, are quickly rising as an important and prominent way that victims — especially espionage victims — come to learn about breaches. Keep up the good work, folks; we're making a dent!

We hope you enjoyed this little ten-year trip down memory lane as much as we did. This small band of geeks is grateful to Verizon for allowing us to spend so much time in our playground of breach information. We're also grateful to the many organizations that have participated in making it possible; without your contributions the data would have gotten stale years ago. And finally, thanks to all you readers out there who download this document and consider these trends as you fight the good fight of protecting information and customers. May the next ten years find us all on the winning side of that battle.

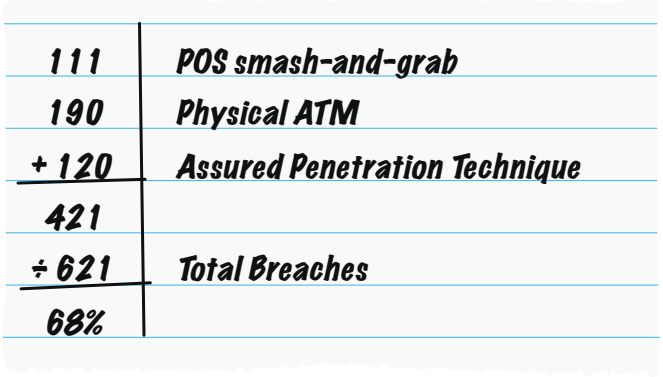
Questions? Comments? Brilliant ideas?

We want to hear them. Drop us a line at dbir@verizon.com, find us on LinkedIn, or tweet @VZdbir with the hashtag #dbir.

RESULTS AND ANALYSIS

he seeds of our approach to the 2014 DBIR began to grow during the final phase of drafting the 2013 report. When trying to present statistics around threat actions in a simple and meaningful way, we noticed certain combinations of actors, actions, and assets frequently occurring together within an incident scenario. We gave names to three of these and included some "scratch paper" calculations showing they collectively described 68% of the entire dataset (Figure 15). Production deadlines prevented further exploration into that phenomenon, so we left readers with this thought: "We may be able to reduce the majority of attacks by focusing on a handful of attack patterns." But as soon as the 2013 DBIR was put to bed, we returned to the notion of incident patterns and began studying our dataset from a very new perspective with a new set of techniques.

Figure 15.
Scratch paper calculations from the 2013 DBIR for commonly-observed incident patterns



Now, fast forward to the 2014 DBIR. We have more incidents, more sources, and more variation than ever before – and trying to approach tens of thousands of incidents using the same techniques simply won't cut it. Not only would the dominant incident characteristics drown out the subtleties of the less frequent varieties, but we cannot continue to study those characteristics as though they occur in isolation. A list of the top 20 actions is helpful, but even more helpful is an accounting of the actors that perform them, other actions used in combination with them, and the assets they tend to target. To reel in that prize, we're going to need a bigger boat. Full throttle, Mr. Hooper!

And that brings us back to recurring combinations of actors, actions, assets, and attributes, or more formally, incident classification patterns. In order to expose these latent patterns in the data, we applied a statistical clustering technique (the bigger boat) by creating a matrix aggregating incidents within each of the common VERIS enumerations and calculating the numeric "distance" between them. This enabled us to find clusters, or patterns, of strongly related VERIS enumerations within the incident dataset. "Strongly related" here essentially means they often occur together in the same incidents and are distinct in some way from other combinations.

The first time through, we tossed everything in and looked at the clustering (of the hierarchical type if you're into that) of VERIS enumerations. Some clusters were obvious, like the social action of phishing with the social vector of email. However, we were looking for clusters that describe comprehensive incident classifications rather than just frequent pairings. For example, incidents involving physical tampering of ATMs by organized criminal groups to steal payment cards stood out like a Wookie among Ewoks. So we labeled that pattern "skimmers," removed the matching incidents, and reran the cluster analysis on the remaining incidents to look for the next pattern.

In the end, we identified nine patterns that together describe 94% of the confirmed data breaches we collected in 2013.

Nine out of ten of all breaches can be described by nine basic patterns.

But (using our best infomercial voice) that's not all! When we apply the same method to the last three years of breaches, 95% can be described by those same nine patterns.

But wait — there's more! Act now, and we'll throw in all security incidents — not just breaches — from all partners and the VERIS Community Database (VCDB) over the last ten years — for free! Yes, all for the same price of nine patterns, you can describe 92% of 100K+ security incidents!

Remember that promise from last year — "We may be able to reduce the majority of attacks by focusing on a handful of attack patterns?" Consider it fulfilled. To us, this approach shows extreme promise as a way to drastically simplify the seemingly endless array of threats we must deal with to protect information assets.

We dig into each incident pattern in the following sections, but you can see from Figure 16 that POS intrusions, web app attacks, cyber-espionage, and card skimmers are among the top concerns when we focus on data disclosure. However, it's not enough to just identify and count the patterns as a whole.

Figure 16.
Frequency of incident classification patterns

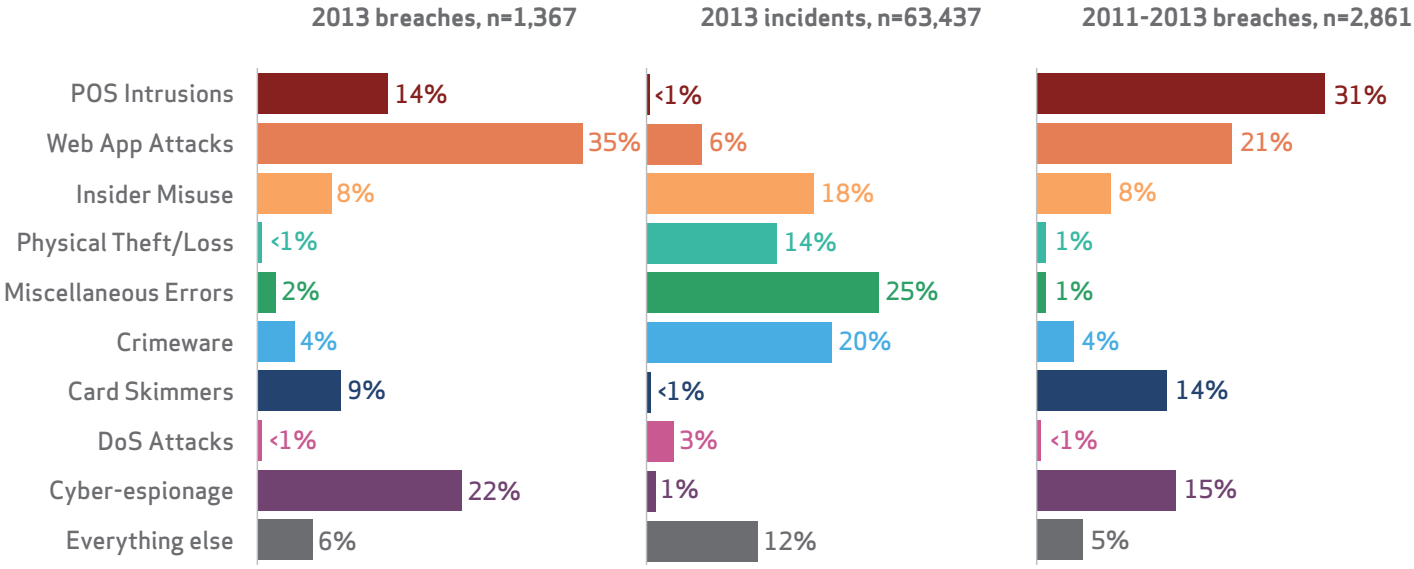


Figure 17.
Number of selected incident classification patterns over time

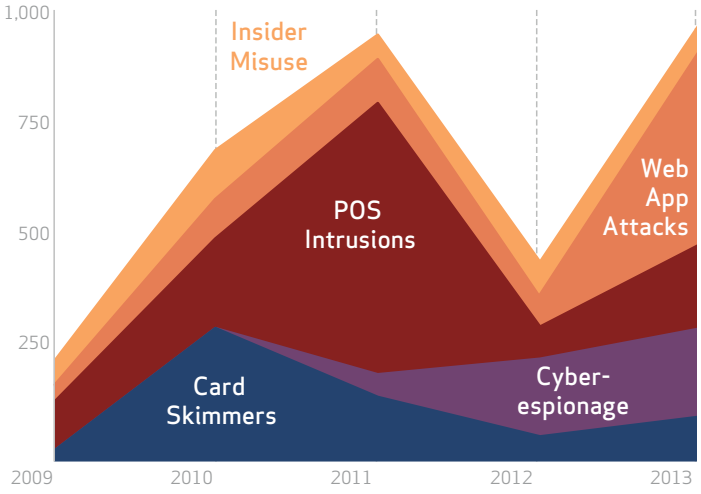
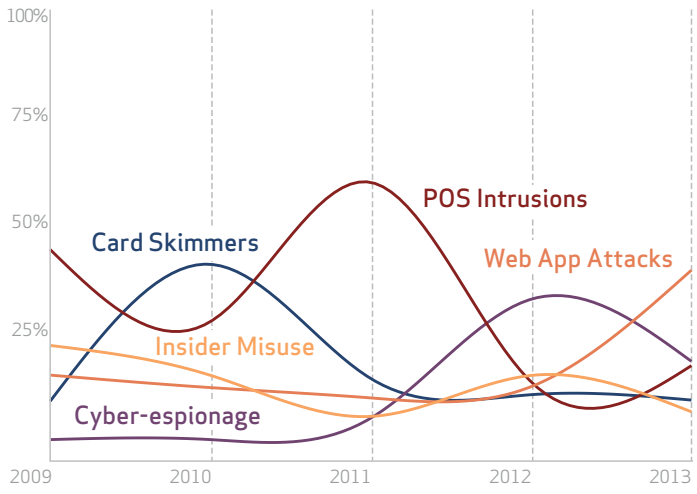


Figure 18.
Percent of selected incident classification patterns over time



Obviously not every organization needs to focus on point of sale attacks. To make the analysis actionable, we pulled all incidents within each industry and then applied the patterns to create the work of art that is Figure 19. It shows the proportion of incidents within each industry represented by the nine patterns over the last three years.

In order to use Figure 19, identify your industry in the left hand column. Refer to the NAICS website if you're unsure where your organization fits. The percentages are relative to each industry. For example, 10% of all Retail incidents fall within the "web app attack". The coloring should help you quickly identify "hot spots" for your industry and/or discern differing threat profiles across multiple industries.

Before continuing on to the detailed discussion of each pattern (which appear in order according to Figure 18), you may want to study Figure 19. Look up the industry (or industries) that matter to you, identify which patterns are most relevant, and pay special attention to those sections in the report (you'll still want to read the whole thing, of course). For those curious about how these incident patterns trend over time, we've retrofitted them to pre-2013 data to produce Figures 17 and 18.

We've heard the (constructive) criticism from some of you noting that it's difficult to pick out exactly which findings from the DBIR apply to your organization, and we spent a lot of time figuring out how to address that. We hope you'll agree this is a step in the right direction, not only for this report, but also for threat analysis and decision support in general.

Figure 19.
Frequency of incident classification patterns per victim industry

INDUSTRY	POS INTRUS-ION	WEB APP ATTACK	INSIDER MISUSE	THEFT/ LOSS	MISC. ERROR	CRIME-WARE	PAYMENT CARD SKIMMER	DENIAL OF SERVICE	CYBER ESPION-AGE	EVERY-THING ELSE
Accommodation [72]	75%	1%	8%	1%	1%	1%	<1%	10%		4%
Administrative [56]		8%	27%	12%	43%	1%		1%	1%	7%
Construction [23]	7%		13%	13%	7%	33%			13%	13%
Education [61]	<1%	19%	8%	15%	20%	6%	<1%	6%	2%	22%
Entertainment [71]	7%	22%	10%	7%	12%	2%	2%	32%		5%
Finance [52]	<1%	27%	7%	3%	5%	4%	22%	26%	<1%	6%
Healthcare [62]	9%	3%	15%	46%	12%	3%	<1%	2%	<1%	10%
Information [51]	<1%	41%	1%	1%	1%	31%	<1%	9%	1%	16%
Management [55]		11%	6%	6%	6%		11%	44%	11%	6%
Manufacturing [31,32,33]		14%	8%	4%	2%	9%		24%	30%	9%
Mining [21]			25%	10%	5%	5%	5%	5%	40%	5%
Professional [54]	<1%	9%	6%	4%	3%	3%		37%	29%	8%
Public [92]		<1%	24%	19%	34%	21%		<1%	<1%	2%
Real Estate [53]		10%	37%	13%	20%	7%			3%	10%
Retail [44,45]	31%	10%	4%	2%	2%	2%	6%	33%	<1%	10%
Trade [42]	6%	30%	6%	6%	9%	9%	3%	3%		27%
Transportation [48,49]		15%	16%	7%	6%	15%	5%	3%	24%	8%
Utilities [22]		38%	3%	1%	2%	31%		14%	7%	3%
Other [81]	1%	29%	13%	13%	10%	3%		9%	6%	17%

For more information on the NAICS codes [shown above] visit: https://www.census.gov/cgi-bin/sssd/naics/naicsrch?chart=2012

POINT-OF-SALE (POS) INTRUSIONS

AT A GLANCE

Description

Remote attacks against the environments where retail transactions are conducted, specifically where card-present purchases are made. Crimes involving tampering with or swapping out devices are covered in the Skimming pattern.

Top industries

Accommodation and Food Services, Retail

Frequency

198 total incidents

198 with confirmed data disclosure

Key findings

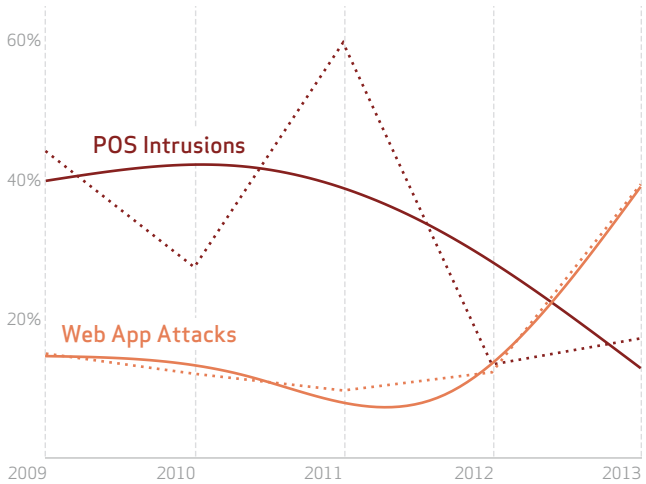
Given recent headlines, some may be surprised to find that POS intrusions are trending down over the last several years. That's mainly because we've seen comparatively fewer attack sprees involving numerous small franchises. Brute forcing remote access connections to POS still leads as the primary intrusion vector. A resurgence of RAM scraping malware is the most prominent tactical development in 2013.

We know many of you will come to this section hoping to find all the particulars and dirty laundry of a certain breach involving a major U.S. retailer in late 2013. Prepare to be disappointed; we don't name victims in this report nor do we divulge client-specific information on any breaches handled by any of the DBIR contributors. If you want up-to-the-minute news on particular breaches, you'd best look elsewhere. As a consolation prize, however, we hope you'll accept our overall analysis of two hundred POS intrusions that occurred in 2013, along with recommendations on how you can avoid increasing that number in 2014.

The industries most commonly affected by POS intrusions are of no surprise: restaurants, hotels, grocery stores, and other brick-and-mortar retailers are all potential targets. Recent highly publicized breaches of several large retailers have brought POS compromises to the forefront. But at the risk of getting all security-hipster on you — we've been talking about this for years. In fact, this is the main cause of the large dip in 2012 seen in many of the "over time" charts in this report. We were writing about RAM scrapers before anyone heard of them and we're quite frankly not all that into them anymore because they've sold out and gone mainstream.

Jokes aside, while POS hacks are getting more press recently, they really have been going on for years and we really have talked quite a bit about them in previous DBIRs. The media frenzy makes quite a splash, but from a frequency standpoint, this largely remains a small-and-medium business issue. Focusing too much on outliers and headlines can reflect cognitive bias. For instance, some may be surprised that the number of POS attacks in 2012 and 2013 is substantially lower than the number recorded in 2010 and 2011 (despite having ten times more contributors in the latter years). Figure 20 reminds us that our understanding of risk should always come back to the data, not what makes good headlines and marketing fodder.

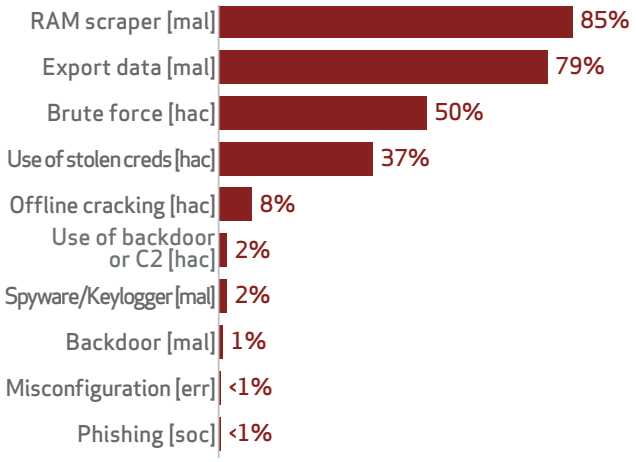
Figure 20.
Comparison of POS Intrusions and Web App Attacks patterns, 2011-2013



From an attack pattern standpoint, the most simplistic narrative is as follows: compromise the POS device, install malware to collect magnetic stripe data in process, retrieve data, and cash in. All of these attacks share financial gain as a motive, and most can be conclusively attributed (and the rest most likely as well) to organized criminal groups operating out of Eastern Europe.[3] Such groups are very efficient at what they do; they eat POSs like yours for breakfast, then wash 'em down with a shot of vodka. While the majority of these cases look very much alike, the steps taken to compromise the point-of-sale environment offer some interesting variations.

Let's start with the most frequent scenario, which affects small businesses that may or may not realize just how lucrative a target they are. This event chain begins with the compromise of the POS device with little to no legwork; the devices are open to the entire Internet and, to make matters worse, protected with weak or default passwords (and sometimes no passwords).

Figure 21.
Top 10 threat action varieties within POS Intrusions (n=196)



The top three threat actions tell the story rather well (Figure 21). The perpetrators scan the Internet for open remote-access ports and if the script identifies a device as a point of sale, it issues likely credentials (Brute force) to access the device. They then install malware (RAM scraper) to collect and exfiltrate (Export data) payment card information.

One finding that intrigued us is the renaissance of RAM scraping malware as the primary tool used to capture data. RAM scrapers allow payment card data to be grabbed while processed in memory (where it is unencrypted) rather than when stored on disk or in transit across the network (where it is (ostensibly) encrypted).

It's interesting, but not necessarily surprising, that RAM scraping has usurped keyloggers as the most common malware functionality associated with POS compromises. One could theorize that keyloggers (most of which were common varieties such as Perfect Keylogger and Artemis) are more easily spotted than the memory-scraping code we witnessed in this data set. Or perhaps the RAM scrapers, which hook into specific processes of the POS software, simply do the job better and more efficiently.

In years past, we analyzed attack sprees that spanned multiple victims with no association with each other beyond the use of truly awful passwords. This report features almost 200 incidents, but in prior years we saw over 200 victims for one criminal group. The two biggest sprees in our 2013 dataset, one involving several franchisees of the same company, and the other affecting multiple corporations, are a bit different, and lead us to our second common scenario: the use of stolen vendor credentials. In one case the credentials stolen belonged to a point-of-sale vendor and were compromised via Zeus malware infecting the vendor's systems. The big problem among these was that the same password was used for all organizations managed by the vendor. Once it was stolen, it essentially became a default password and the attackers also gained knowledge of the customer base. Armed with this information, the familiar modus operandi of installing malicious code that captured and transmitted the desired data began.

Figure 22.
Hacking variety within POS Intrusions (n=187)

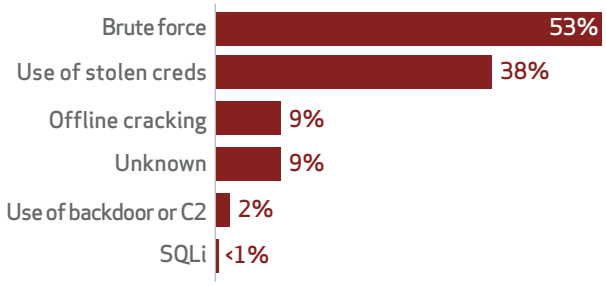
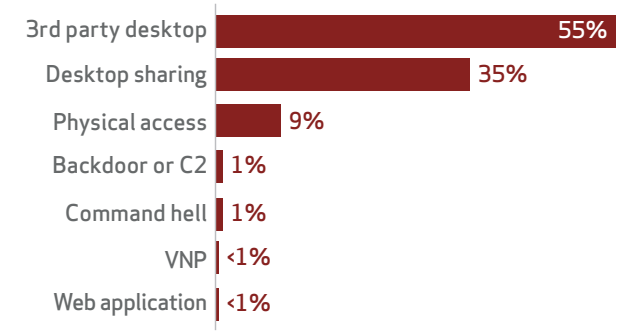
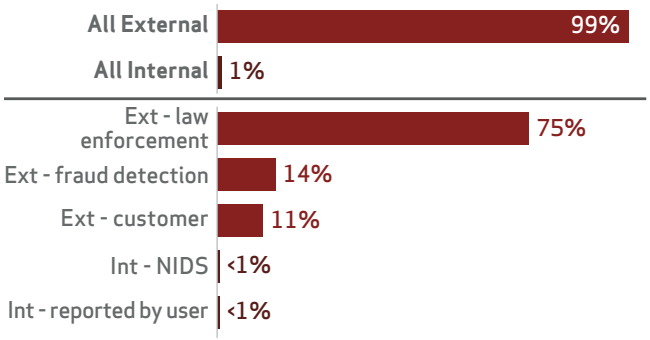


Figure 23.
Hacking vector within POS Intrusions (n=187)



While not as common as the simpler POS intrusions, our dataset does include several incidents from the first quarter of 2013 that feature a compromise at a corporate location, leading to widespread compromise of individual locations and malicious code installations across a multitude of stores. Some cases have begun with a store compromise that led to penetration of the corporate network, but the hub-and-spoke architecture allowed for efficient traversal of the network and the impact of the compromise was magnified regardless of where "device 0" was located.

Figure 24.
Top 5 discovery methods for POS Intrusions (n=197)



Regardless of how large the victim organization was or which methods were used to steal payment card information, there is another commonality shared in 99% of the cases: someone else told the victim they had suffered a breach. This is no different than in years past, and we continue to see notification by law enforcement and fraud detection as the most common discovery methods. In many cases, investigations into breaches will uncover other victims, which explains why law enforcement is the top method of discovery and the top contributor of POS intrusions in our dataset. Long story short, we're still discovering payment card breaches only after the criminals begin using their ill-gotten gains for fraud and other illicit purposes.

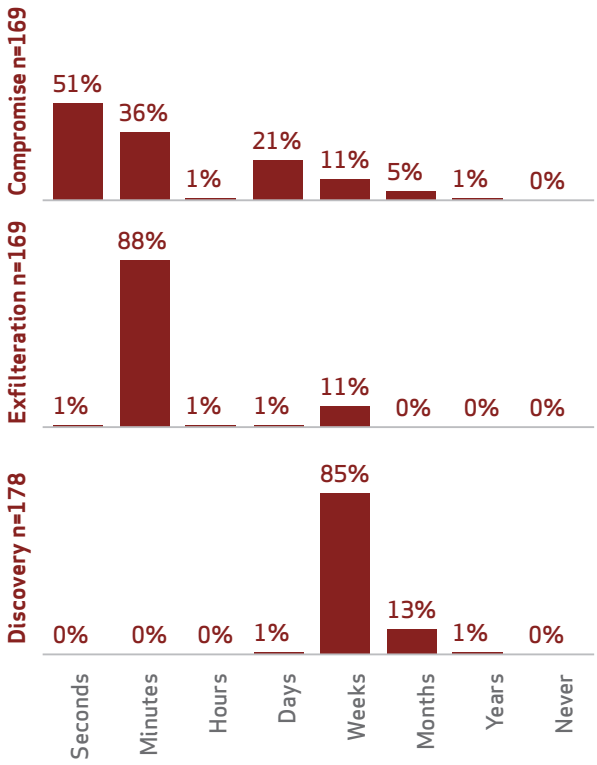
BOTNET MITIGATION: AN INCENTIVE PROBLEM

According to the NHTCU, the impact of botnets — the Swiss Army knife of cybercriminals — remained high in 2013. Furthermore, they note an apparent incentive problem when it comes to mitigating these crafty menaces. Since the impact of a botnet is often spread around the globe, federal authorities aren't always able to amass resources to fight it on a national level. While the total damage of such a botnet might be large, specific countries only deal with a small part of these damages. The initial costs for fighting such a botnet don't seem to outweigh the benefits of its takedown. Nevertheless, the NHTCU continue to fight botnets. In February of 2013, public broadcaster NOS presented findings on part of a dropzone of the so-called Pobelka botnet. After an online checking tool was made available, 500,000 people checked to see if their machines had (at some time) been infected; of that group, 23,000 self-identified as victims.

By then, the dropzone had been examined for correlations with a 2012 malware outbreak that had prompted a criminal investigation. Sixteen organizations within the vital infrastructure were informed of being infected, and relevant infected IP addresses had been communicated to the respective ISPs.

The timelines in Figure 25 reinforce both the compromise vectors and the discovery methods. Entry is often extremely quick, as one would expect when exploiting stolen or weak passwords. Most often it takes weeks to discover, and that's based entirely on when the criminals want to start cashing in on their bounty.

Figure 25.
Timespan of events within POS Intrusions



RECOMMENDED CONTROLS

FOR ALL COMPANIES

The shared vector for the major scenarios is third-party remote-access software (e.g., PCAnywhere, LogMeIn). The security of these products is not the issue here. It just happens that we often find them implemented in a very insecure manner.

Restrict remote access

Limit any remote access into POS systems by your third-party management vendor, and have serious business discussions regarding how and when they will perform their duties.

Enforce password policies

Make absolutely sure all passwords used for remote access to POS systems are not factory defaults, the name of the POS vendor, dictionary words, or otherwise weak. If a third party handles this, require (and verify) that this is done, and that they do not use the same password for other customers.

"S" is for "Sale," not "Social."

Do not browse the web, email, use social media, play games, or do anything other than POS-related activities on POS systems.

Deploy AV

Install and maintain anti-virus software on POS systems.

Bottom line: Make it difficult for miscreants to log into a device that accepts the most targeted piece of information for financially motived criminals.

FOR LARGE/MULTI-STORE COMPANIES

Larger, multi-store companies and franchises should consider a couple of additional recommendations to limit the impact of a single-location breach and prevent a mass compromise.

Debunk the flat network theory

Review the interconnectivity between stores and central locations and treat them as semi-trusted connections. Segment the POS environment from the corporate network.

Look for suspicious network activity

Monitor network traffic to and from the POS network. There should be a normalized traffic pattern, and while easier said than done, anomalous traffic must be identified and investigated.

Use two-factor authentication

Stronger passwords would cut out a huge chunk of the problem, but larger organizations should also consider multiple factors to authenticate third-party and internal users.

REFLECTIONS AFTER THE EC3'S FIRST YEAR IN OPERATION

Last year's DBIR featured an appendix from Troels Oerting, Assistant Director of the European Cybercrime Centre (EC3), discussing the plans and priorities of the newly established division of Europol. Law enforcement agencies play a critical role in this report, and it's not often we get to see them in their formative stages. Thus, we thought it would be interesting to include some reflections on EC3's first year of operations.

The job of the EC3 isn't a small one: it serves 28 European Union (EU) member states and dozens of countries, and coordinates protection for 500 million citizens, almost three-quarters of whom have Internet access. In terms of operations, the EC3 prioritizes four areas: cyber intelligence, intrusion, online fraud, and child sexual abuse. As with any new venture, much of the first year focused on building infrastructure and capabilities to fulfill these priorities. Secure network connections to EU and non-EU partners were rolled out, as well as centralized forensic analysis environments and tools.

The EC3 trained more than 100 law enforcement experts all over the EU in cyber investigation, tools, and obtaining forensic evidence. It built a new central forensic laboratory to assist member state colleagues in obtaining evidence. It distributed alerts, intelligence notifications, and threat assessments to stakeholders. Memorandums of understanding (MoUs) were signed with key private stakeholders, and a new Advisory Group consisting of experts outside the law enforcement community was established (Verizon is happy to be among them).

Trends observed by the EC3 across member states in 2013 include substantial increases in intrusions, malware, phishing, grooming, DDoS, espionage, and botnet activity. It also reports a boom in criminal infrastructure on the darknet, growth in malware affecting mobile devices, and wider distribution of malware from cloud services. In combating these trends, the EC3 has prioritized identifying criminal network operations and cases, with the potential for major and lasting impact.

WEB APP ATTACKS

AT A GLANCE

Description

Any incident in which a web application was the vector of attack. This includes exploits of code-level vulnerabilities in the application as well as thwarting authentication mechanisms.

Top industries

Information, Utilities, Manufacturing, Retail

Frequency

3,937 total incidents

490 with confirmed data disclosure

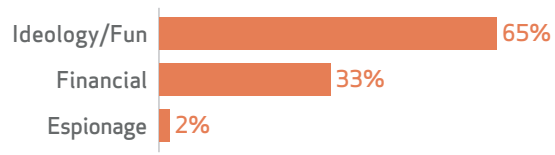
Key findings

Web applications remain the proverbial punching bag of the Internet. They're beaten in one of two ways: by exploiting a weakness in the application (typically inadequate input validation), or by using stolen credentials to impersonate a valid user. Many of the attacks in our 2013 dataset targeted off-the-shelf content management systems (e.g., Joomla!, Wordpress, or Drupal) to gain control of servers for use in DDoS campaigns.

There's no question about it – the variety and combination of techniques available to attackers make defending web applications a complex task. Regrettably, our discussion of this complexity is hampered by the level of detail provided on these incidents. Unless a forensics investigation was performed (a small subset of the overall dataset), the specific techniques utilized went largely unreported or were recorded with broad categorizations. While we have enough material to discuss web application data breaches at a high level, our ability to draw conclusions drops as we dig further into the details (which often aren't there).

Greed takes a back seat to ideology when it comes to web app attacks in the 2013 dataset. Just under two out of every three web app attacks were attributable to activist groups driven by ideology and lulz; just under one out of three came by the hand of financially motivated actors; with the small remainder linked to espionage. After some slicing and dicing we found some very distinct sub-patterns divided along these motives. The financial and ideological attacks deserve unique discussion since the treatment for each may be slightly different. While attacks perpetrated by those motivated by espionage are certainly relevant, discussion of these is taken up in the "Espionage" section.

Figure 26.
External actor motives within Web App Attacks (n=1,126)



FINANCIALLY MOTIVATED ATTACKS

Financially motivated attackers are hyper-focused on gaining access to the money, so it follows that their two primary target industries are the financial and retail industries (where data that easily converts to money is abundant and, all too often, accessible). Within the financial industry, they focus on gaining access to the user interface of the web (banking) application more so than exploiting the web application itself, because the application grants logical access to the money. This means they target user credentials and simply use the web applications protected with a single factor (password) as the conduit to their goal. These could have been included in the section on crimeware (and some did slip through cracks in the algorithm to land there), but the use of web applications as a vector of attack causes them to show up here. The tactics used by attackers are all the usual suspects: a) phishing techniques to either trick the user into supplying credentials or installing malware onto the client system, b) the old stand-by of brute force password guessing, and c) rarer cases of targeting the application through SQL injection or other application-level attacks as a means to retrieve credentials, bypass the authentication, or otherwise target the user-management system. When attribution is possible, the majority of external attackers utilizing stolen credentials somewhere along the attack chain hail from Eastern Europe.

Within the retail industry, we see a slightly different focus. The primary aim is payment card information (targeted in 95% of the incidents), which is often accessible simply by exploiting the web application. Social actions (such as phishing) are mostly non-existent, most likely because exploiting vulnerabilities inherent in web applications works plenty well enough. SQL injection was leveraged in 27 of the 34 (80%) attacks against web applications in the retail industry, followed by techniques to install and use web shells (remote file inclusion, etc.) in five of the 34.

IDEOLOGICALLY MOTIVATED ATTACKS:

Ideology represents the largest identified portion of motives for web application attacks, and the actors also tend to be the most geographically diverse. 74% focus on tried and true exploits targeting, above all else, unvalidated inputs in executed code. Nowhere is this exploited on a larger scale than Content Management Systems (CMS) such as Joomla!, Drupal, and WordPress, and even then, more in the added plugins than the core CMS code itself.

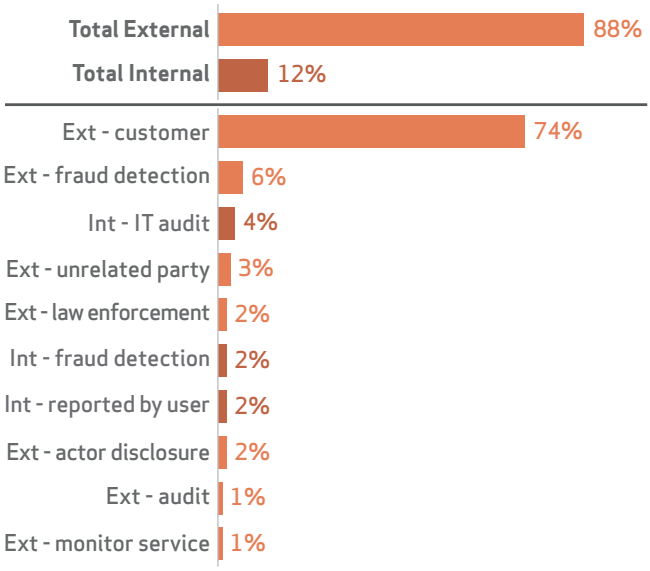
Ideological actors (whether their motivation is social, political, or just for plain fun) are less concerned about getting at the crown jewels than they are about getting a platform (in all senses of the word) to stand on. With that in mind, it's not surprising that we see two types of results from ideological attackers going after a web server: defacements to send a message or hijacking the server to attack (including by DDoS) other victims.

This focus on opportunistically owning just the web server becomes plain when looking at the assets compromised in the attack. The web server was the only asset recorded in nearly all incidents attributable to ideological motives. The actors didn't appear to be interested in pushing deeper and wider into the network. This result may be the product of simply not reporting those secondary components of the incident — so don't take this as advice to only focus on the web server — but it is logical and a point of contrast to other types of attacks in our dataset.

DISCOVERY METHODS AND TIMELINE

When the actor is financially motivated and the discovery method is recorded, we see a leading notification method that we don't see anywhere else: customers. Perhaps customers notice the fraudulent activity before anyone else, but something is definitely tipping them off before any internal mechanism. With all internal discovery methods combined, only 9% of victims discovered data breaches of their own accord.

Figure 27.
Top 10 discovery methods for financially motivated incidents within Web App Attacks (n=122)



Discovery method looks a little bleaker for activists. 99% of the notifications were external parties (primarily CSIRTs) contacting victims to let them know their hosts were involved in other attacks. This is heavily influenced by ideological attackers quietly using the platform to attack others rather than, for instance, simple defacements (which are rare in the dataset).

Even though the timeline data is a little sparse, it paints the picture of quick entry with 60% of the initial compromises occurring within minutes or less. This reflects the highly repetitive CMS exploits in this pattern; if it works, it works quickly. Just over 85% of the incidents are discovered in days or more, with about 50% taking months or longer to discover. Once discovered though, we see fairly good reaction time, with about half of the organizations taking days or less to respond and contain the incident. This is far better than the norm, which is typically weeks or longer.

COMPARING ATTACKS TO PATCH TIMEFRAMES

Wouldn't it be great to know that (quickly) patching web application vulnerabilities helps? This year we partnered with WhiteHat Security in order to combine and compare the incident data we've collected against the vulnerability assessment data they collect from tens of thousands of websites across hundreds of the most well-known organizations. After some back and forth, we decided first to break out the data by industries (because patterns emerge across industries), then we decided to compare two data points on web vulnerabilities to the incident data: the average (mean) vulnerabilities per site and median days to patch. We assumed that industries with fewer vulnerabilities and quicker patch time would be less represented in the breach data (i.e., have fewer incidents) and so we applied some good old-fashioned statistics and were admittedly let down when we didn't see the relationship[4] we were expecting.

What we found is a non-finding and the only valid conclusion to draw from this is that more work is needed to understand the relationship between web application vulnerabilities and security incidents. With a non-finding, we can only speculate on why we are seeing these results. Perhaps this is telling us that no industry is doing enough. We know three out of four web-based compromises occur in hours or less of first contact, and maybe fixing vulnerabilities in 10 days versus 70 days doesn't help all that much. Plus, the attacker only exploits one (maybe two) vulnerabilities. But a different explanation could be that our lens was focused too wide, and we could learn more by matching the high-quality WhiteHat data with specific incident data within the same sample. Whatever the causes, we do know that web application attacks occur often enough to repeat what is said in the WhiteHat Website Security Statistics Report,[5] "What's needed is more secure software, NOT more security software."

Figure 28.
Top 5 discovery methods for ideologically motivated incidents within Web App Attacks (n=775)

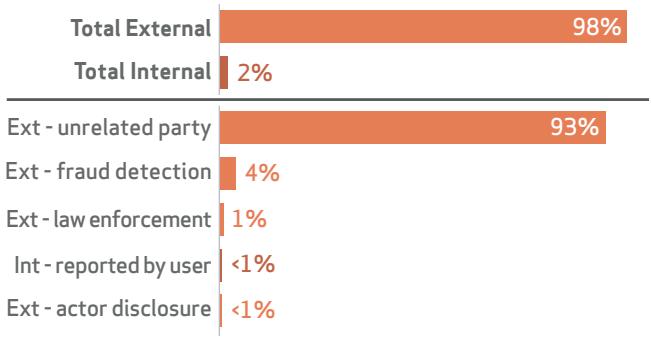
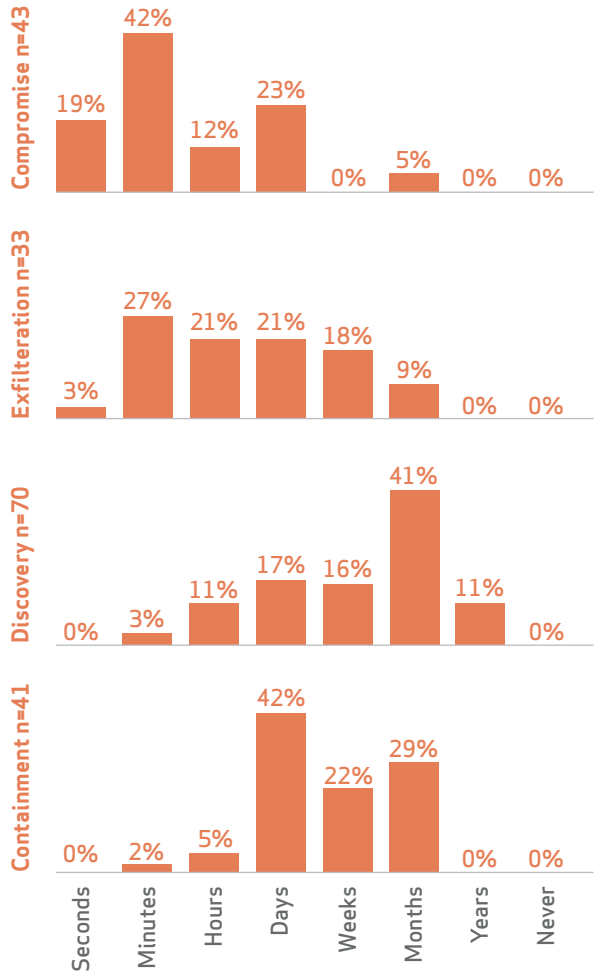


Figure 29.
Timespan of events within Web App Attacks



RECOMMENDED CONTROLS

Single-password fail

The writing's on the wall for single-factor, password-based authentication on anything Internet-facing. Even though it may draw you out of a known comfort zone, if you're defending a web application seek out alternatives to this method of identity verification. If you're a vendor in the web application space, also consider mandating alternative authentication mechanism for your customers.

Rethink CMS

And we mean "rethink" in every way. If you're committed to an active platform (Joomla!, Drupal, WordPress, etc.), then set up an automated patch process. If an automated patch process isn't viable, then develop a manual process and stick to it. This is especially true for the third-party plugins. Another way to rethink CMS is to consider a static CMS framework. Instead of executing code for every request and generating the content, static CMS will pre-generate those same pages, removing the need to execute code on the server for every request.

Validate inputs

Even though we've been tilting at this windmill for years, the advice still holds true. The best way to be sure your web application won't be exploited is to seek out and fix the vulnerabilities before the attackers do (and they will). If you don't have access to the source code and/or the developers, be sure to have something in place (e.g., a contract) to fix the problems when they're found.

Enforce lockout policies

Brute force attacks aren't the leading method in this section, but they're still worthy of mention. By instituting counter-measures, such as a slowing down the rate of repeated attempts or temporarily locking accounts with multiple failed attempts, the rate of successful brute force attempts will more than likely dissipate and disappear (although you may still be dealing with that pesky bot poking at your accounts every now and then).

Monitor outbound connections

While many web-based attacks rely heavily on the existing firewall bypass protocol (HTTP), many others change the victim's web server into a client. Critical points in the attack chain are pulling additional malware to continue the attack, exfiltrating compromised data, or attacking others on command. So unless your server has a legitimate reason to send your data to Eastern Europe or DoS'ing others is part of the business plan, try to lock down your web server's ability to do so.

INSIDER AND PRIVILEGE MISUSE

AT A GLANCE

Description

All incidents tagged with the action category of Misuse — any unapproved or malicious use of organizational resources — fall within this pattern. This is mainly insider misuse, but outsiders (due to collusion) and partners (because they are granted privileges) show up as well.

Top industries

Public, Real Estate, Administrative, Transportation, Manufacturing, Mining

Frequency

11,698 total incidents

112 with confirmed data disclosure

Key findings

Most crimes by trusted parties are perpetrated for financial or personal gain. The most noticeable shifts in the 2013 dataset, however, were an increase in insider espionage targeting internal data and trade secrets, and a broader range of tactics. We say "2013 dataset" because we do not believe the actual rate of such crimes increased significantly; we're seeing the benefit of increased visibility from insider-focused partners.

An organization's intellectual property is among its most valuable assets, frequently driving its ability to compete in the market. In many cases, organizations also have custody of vast amounts of data about the customers they serve, the employees who serve them, and the relationships they rely upon to do business. This data has value to the organization, but also to those who would seek it for their own personal benefit or myriad other reasons. For the misuse pattern, we focus on those who already have a trusted place inside the organization. Arguably, the most prominent case of internal misuse in the headlines this past year has been that of U.S. government contractor Edward Snowden. While this is an extreme example of the damage that determined insiders can inflict, it illustrates the risk that exists when an organization must place trust in individuals.

Figure 30.
Top 10 threat action varieties within Insider Misuse (n=153)

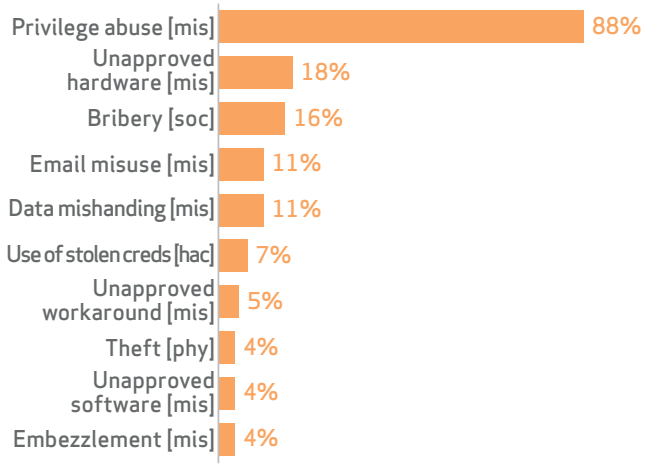
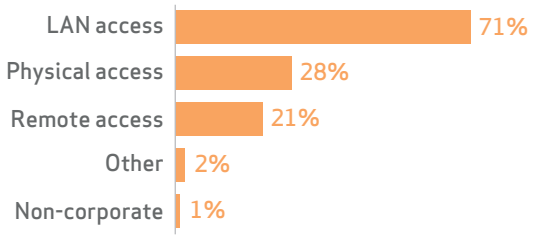


Figure 30 lists the top threat actions observed across incidents fitting the misuse pattern. Note that not all are within the misuse category [mis]; stay tuned for more on that later. Not unexpectedly, privilege abuse — taking advantage of the system access privileges granted by an employer and using them to commit nefarious acts — tops the list. We realize that encompasses a very broad range of activities, but the overall theme and lesson differ little: most insider misuse occurs within the boundaries of trust necessary to perform normal duties. That's what makes it so difficult to prevent.

Remember that action varieties in VERIS are not mutually exclusive, and it's common to see more than one in a single incident. Unapproved hardware and email misuse/data mishandling (tied) round out the top three actions in the misuse category, but they're more a function of how the data is exfiltrated rather than how it's acquired. Unapproved hardware refers to employees using devices like USB drives that are either forbidden altogether or allowed but subject to various restrictions. An employee sending intellectual property out to his or her personal address is an example of email misuse. We also reviewed cases where system administrators abused the email system, posing as another user and sending messages under that identity, with the goal of getting them fired. Data mishandling takes place when someone uses data in a manner counter to the organization's policies. For example, a call center employee who writes customer credit card numbers down on paper, or an engineer who skirts policy by taking restricted documents home to examine on a personal computer.

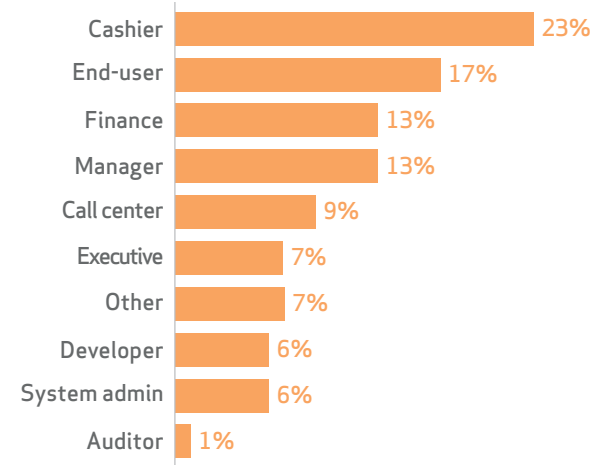
With more incidents than ever before involving trusted parties, we could more easily see how they go about acquiring the data when their own access is insufficient. In addition to abusing entrusted privileges and resources, we observed hacking techniques to elevate privileges (often by stealing others' credentials) and circumvent controls, various forms of social engineering, and the use of malware like keyloggers and backdoors. These cads have even resorted to physical theft, taking documents such as blueprints and other intellectual property, often denying availability to the original organization by taking the only copy.

Figure 31.
Vector for threat actions within Insider Misuse (n=123)



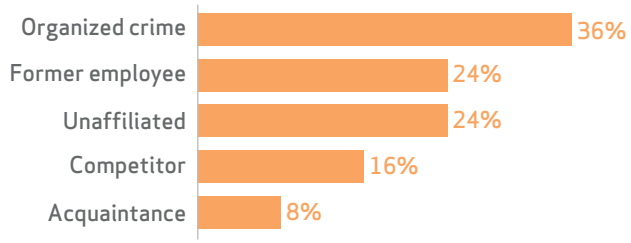
It's also worth noting that the corporate LAN was the vector in 71% of these incidents, and 28% took advantage of physical access within the corporate facility. This means the majority of employees perpetrated their acts while in the office right under the noses of coworkers, rather than hopping through proxies from the relative safety of their house. If someone wants to use these statistics to loosen up work-at-home policies and tear down cube farms in favor of more open floor plans — you have our blessing.

Figure 32.
Top 10 varieties of internal actors within Insider Misuse (n=99)



Let's take a look at the people committing these crimes. While payment chain personnel and end-users were still prominent, managers (including those in the C-suite) came in higher than in prior years. You know the type; one of those straight shooters with upper management written all over him. They often have access to trade secrets and other data of interest to the competition and, tragically, are also more likely to be exempted from following security policies because of their privileged status in the company.[6] One of those "white-collar resort prisons" won't do for their ilk.

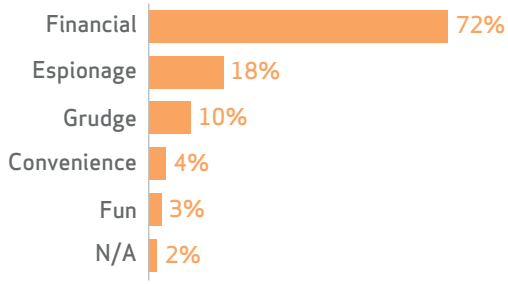
Figure 33.
Variety of external actors within Insider Misuse (n=25)



As mentioned in the beginning of this section, insiders aren't the only ones who misuse entrusted privileges and resources. Figure 33 gives an account of external and partner actors who directly or indirectly participated in incidents of misuse. Organized criminals bribe insiders to steal data for fraud schemes. Former employees exploit still active accounts or other holes known only to them. Competitors solicit intellectual property to gain business advantages. To mount a proper defense, organizations must take into account that such players are on the field.

Nearly all misuse incidents prior to 2013 centered on obtaining information to use for fraud. As Figure 34 shows, we saw more insider espionage targeting internal organizational data and trade secrets than ever before.

Figure 34.
Actor motives within Insider Misuse (n=125)

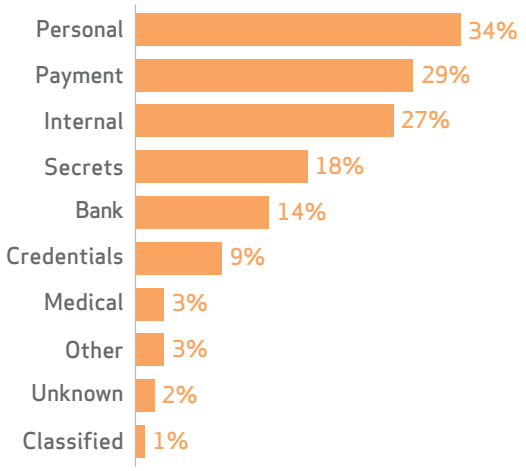


According to *The Recover Report*,[7] published by one of our DBIR contributors, Mishcon de Reya, the two most common scenarios involve perpetrators taking the data to:

- Start their own competing company (30%).
- Help secure employment with a rival (65%).

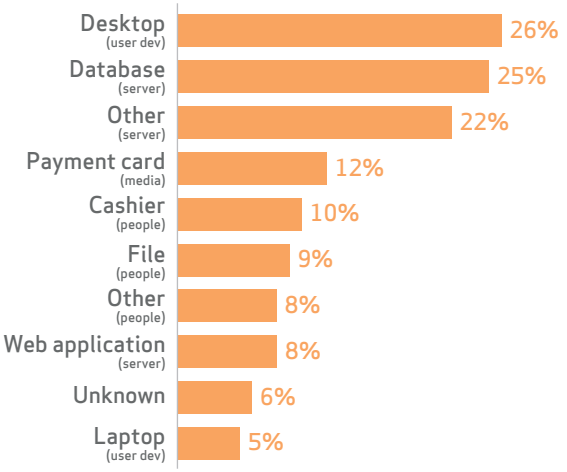
This kind of thing is certainly not new — it's largely due to the addition of more contributors who have a view into this type of activity than ever before. So, whether it's fraud-ready data sold on the quick to criminals or internal secrets eventually sold to a competitor, insider crime is still "all about the Benjamins, baby."

Figure 35.
Variety of at-risk data within Insider Misuse (n=108)



Desktops are the most frequently compromised asset in this pattern, which makes sense because desktop computers are an employee's primary interface to the rest of the network (Figure 36). Typically, this is where the data is stored, uploaded, or emailed out of the organization, or copied onto removable media. Databases and file servers, both repositories of so much valuable information, are also targeted regularly. Payment cards doesn't refer to the variety of data, but rather actual cards that were run through handheld skimming devices (or otherwise copied) in the classic "evil waiter" scenario. As far as asset ownership, we see insiders abusing corporate-owned rather than employee-owned ("BYOD") assets allowed for corporate use. However, we do see evidence they often leverage unapproved personal devices to help them get the data out of the organization (which shows up as use of unapproved hardware).

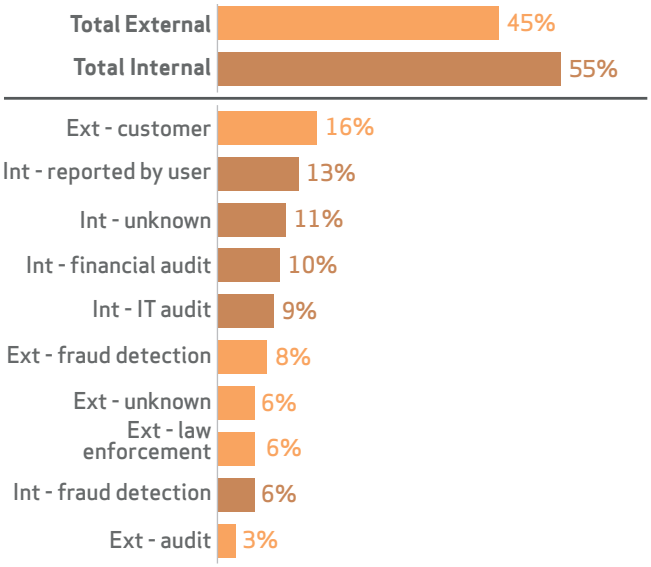
Figure 36.
Top 10 assets affected within Insider Misuse (n=142)



Discovery methods for the majority of breaches have traditionally been dominated by external signals. For insider misuse, however, internal methods (55%) are responsible for detecting more incidents than external methods (45%). The most common way organizations detected insider crimes was when employees reported them. Discoveries triggered by financial and IT audits were also very common. Reviewing the books on Monday morning is an example of the former, and a promising example of the latter is a regular process to review access for exiting employees.

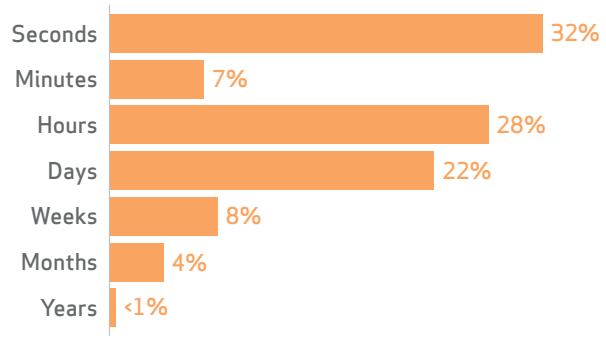
The CERT Insider Threat Center (another partner of ours) focuses research on insider breaches, and it determined that in more than 70% of the IP theft cases, insiders stole the information within 30 days of announcing their resignation.[8] On quite a few occasions, a review of the activity of outgoing employees with access to sensitive information allowed impacted organizations to detect the incident and act quickly to retrieve the information (hopefully before irreparable damage had been done).

Figure 37.
Top 10 discovery methods within Insider Misuse (n=122)



Note the discovery timeline for misuse in Figure 38 — it looks very different from the overall timeline we see for other types of incidents. The majority of the misuse incidents were detected within days (which is great), but there's also a not insignificant number of incidents (70 to be exact) that took years to discover (which ain't so great).

Figure 38.
Discovery timeline within Insider Misuse (n=1,017)



RECOMMENDED CONTROLS

The root cause of data theft and other illicit acts by trusted parties is, rather obviously, an employee breaking bad. No, not in a Walter White sense; more like a white collar crime sense (though Walter did ***SPOILER ALERT*** murder his boss and execute a forced takeover of the business, so it is rather apropos). While it's impossible to stop all rogue employees, there are some steps that can reduce the likelihood of an incident occurring, or at least increase your chances of catching it quickly.

Know your data and who has access to it

The first step in protecting your data is in knowing where it is, and who has access to it. From this, build controls to protect it and detect misuse. It won't prevent determined insiders (because they have access to it already), but there are many other benefits that warrant doing it.

Review user accounts

Having identified the positions with access to sensitive data, implement a process to review account activity when those employees give notice or have been released. Disable user accounts as soon as an employee leaves the company (and, if warranted, before that). This has proven successful in either preventing the data from leaving the organization, or in retrieving it quickly to contain the incident.

Watch for data exfiltration

In the top misuse varieties, we see actions that facilitate the data transfer out of the organization — these are excellent places to set up controls to detect this type of activity. Many data loss prevention products cover the most common actions taken to steal sensitive information, and these are certainly worth exploring.

Publish audit results

From an awareness perspective, regularly publish anonymized results of audits of access. Let employees know that there are consequences and that the policies are being enforced. This can act as a powerful deterrent to bad behavior.

PHYSICAL THEFT AND LOSS

AT A GLANCE

Description

Pretty much what it sounds like — any incident where an information asset went missing, whether through misplacement or malice.

Top industries

Healthcare, Public, Mining

Frequency

9,704 total incidents[9]

116 with confirmed data disclosure

Key findings

Loss is reported more often than theft. In a surprising finding, we discovered that assets are stolen from corporate offices more often than personal vehicles or residences. And while personal and medical information is commonly exposed, most losses/thefts are reported because of mandatory disclosure regulations rather than because of fraud.

To be honest, we debated whether or not to include a section on lost and stolen assets in this report. We decided, however, that we simply couldn't ignore the blatant fact that such incidents — while not sexy or "cyber-y" — are among the most common causes of data loss/exposure reported by organizations. This is especially apparent in industries like Healthcare, where the disclosure of all incidents that potentially expose sensitive data is mandatory. And if there's anything we know to be true about human nature, it's that losing things and stealing things seem to be inherent predispositions.

Studying the findings yielded a few interesting observations that may help inform practice, and that's where we'll focus our attention in this section. As we begin, keep in mind that we're specifically talking about information assets;[10] whatever was lost or stolen had to store, process, or transmit information in order to get our attention.

Observation #1 relates to demographics; we have evidence that every type and size of organization loses stuff and/or has stuff stolen. That may not be much of a shock, but it's at least noteworthy that this is the only incident pattern that applies across the board. Even farmers have problems with people that come and try to snatch their ~~crops~~ laptops.

Speaking of laptops, they're the most common variety of asset reported with this pattern. Incident reports — especially to CSIRTs — often don't specify the asset lost or stolen. Thus, "some kind of user device" is all we can infer and explains why "Other (user dev)" is so frequent. Beyond that, it's what you'd expect: computers, documents, and drives.

The next thing to note is the ratio of loss to theft; losing information assets happens way more than theft, by a 15-to-one difference. And that's important because it suggests the vast majority of incidents in this pattern are not due to malicious or intentional actions. Thus, the primary challenge is to a) keep employees from losing things (not gonna happen) or b) minimize the impact when they do. The smart money is on option b, though bio-implanted computing devices do hold some future promise for option a. That's about all we're going to say about loss, but theft still has a few more lessons for us.

Figure 39.
Top 10 action varieties of Theft/Loss (n=9,678)

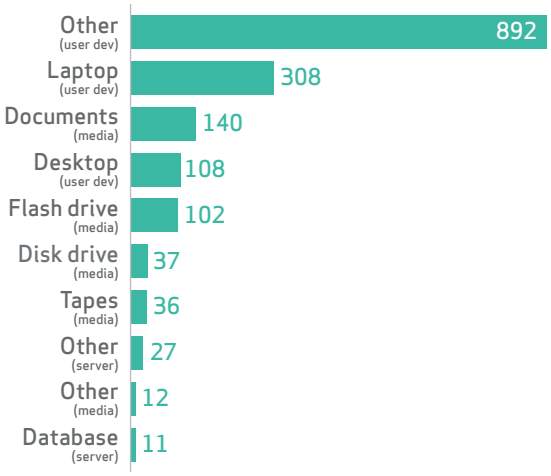
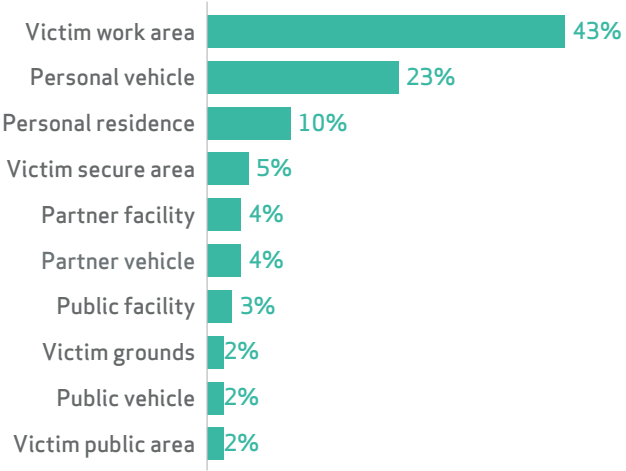


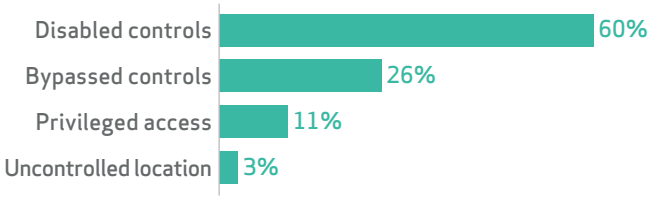
Figure 40.
Top 10 locations for theft within Theft/Loss (n=332)



We find it quite surprising that the highest proportion of thefts occur in the victim's work area, which basically refers to the main office space or cube farm (Figure 40). That suggests simply having sensitive information "behind locked doors" isn't enough; there are still a lot of people inside those locked doors.[11] Notice that thefts in internal high security areas are much less common, but still post higher than public facilities. That last bit is counterintuitive to the point of irrationality; we can't help but suspect that people whose laptops are stolen when they take a potty break at the coffee shop simply report them as "lost" to save face.

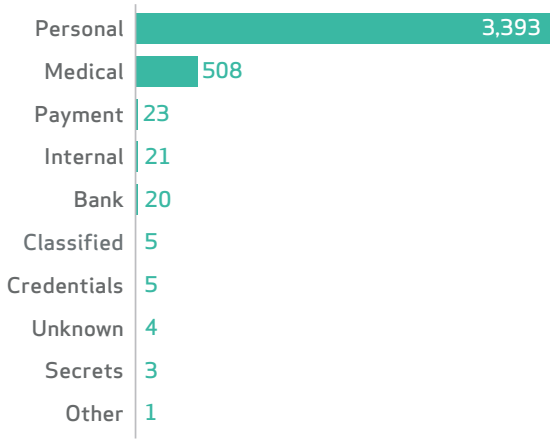
Personal residences and personal/partner/public vehicles serve as the venue for nearly 40% of thefts and remind us that devices on the go are prone to go missing.

Figure 41.
Vector of physical access within Theft/Loss (n=158)



While it's usually not known/reported exactly how actors gained physical access to these locations, over 80% of thefts where we do have that info involved disabling or bypassing controls. The remainder had access already, either because they were granted privileges or because it was a publicly accessible location.

Figure 42.
Variety of at-risk data within Theft/Loss (n=3,824)



The final set of observations covers the variety of data that was compromised or, more often, potentially exposed when assets were lost or stolen. It's worth pointing out that the primary reason most of these incidents are included is because they tripped some kind of mandatory reporting/disclosure requirement. The asset went missing, was determined to contain regulated information that is now exposed to potential unauthorized access, and therefore had to be reported. This explains the predominance of regulated data like personal or identifying information and medical records in Figure 42.

RECOMMENDED CONTROLS

The primary root cause of incidents in this pattern is carelessness of one degree or another. Accidents happen. People lose stuff. People steal stuff. And that's never going to change. But there are a few things you can do to mitigate that risk.

Encrypt devices

Considering the high frequency of lost assets, encryption is as close to a no-brainer solution as it gets for this incident pattern. Sure, the asset is still missing, but at least it will save a lot of worry, embarrassment, and potential lawsuits by simply being able to say the information within it was protected. Also, periodically checking to ensure encryption is still active is right up there too. This will come in handy when the auditor or regulator asks that dreaded question: "How do you know for sure it was encrypted?"

Keep it with you

Encourage employees to keep sensitive devices in their possession and in sight at all times. Yes, this applies to fancy client dinners and visits to the restroom. It's not a bad principle to apply to mobile devices in a corporate setting either. It may be awkward, but it's safer than leaving it in the car or unattended in a room full of strangers. If it absolutely must be left in the car, lock it in the trunk before you leave the office and don't leave it there overnight.

Back it up

Regular (and preferably automatic) backups serve a threefold purpose. They salvage weeks/months/years' worth of irrecoverable work, get you productive again on a new device with minimal down time, and help establish what data was on the device to determine if disclosure is necessary.

Lock it down

In light of the evidence that so many thefts occur in the office, cabling or otherwise securing equipment to immovable fixtures should at least be considered. The big caveat, however, is that the majority of such thefts were documents taken from the filing cabinet and mobile devices (including laptops). A more effective strategy would be to move highly sensitive or valuable assets to a separate, secure area and make sure they stay there.

BONUS - Use unappealing tech

Yes, it's unorthodox as far as recommendations go, but it might actually be an effective theft deterrent (though it will probably increase loss frequency). That shiny new MacBook Air on the passenger seat may be too tempting for anyone to resist, but only those truly dedicated crooks will risk incarceration for a 4" thick mid-90s lap brick. Or, if being the fastest hunk of junk in the galaxy is a must, perhaps there's a lucrative aftermarket for clunky laptop covers. She may not look like much, but she's got it where it counts, kid.

MISCELLANEOUS ERRORS

AT A GLANCE

Description

Incidents where unintentional actions directly compromised a security attribute of an information asset. This does not include lost devices, which is grouped with theft instead.

Top industries

Public, Administrative, Healthcare

Frequency

16,554 total incidents[12]

412 with confirmed data disclosure

Key findings

After scrutinizing 16K incidents, we've made a startling discovery — people screw up sometimes. (Nobel Prize, here we come!) The data seems to suggest that highly repetitive and mundane business processes involving sensitive info are particularly error prone. It's also noteworthy that this pattern contains more incidents caused by business partners than any other.

Nearly every incident involves some element of human error. For example, failing to apply a WordPress patch certainly leaves the application vulnerable to attack, but it doesn't directly compromise the system. Some other threat actor/action is required to do that. Without drawing that distinction, this category would be so bloated with "incidents" that it would be difficult to extract useful information.

It's also worth noting that this pattern doesn't include every incident in the Error category. Loss is a type of error, but we grouped it with theft (under Physical) in a different pattern because they share certain similarities (you no longer possess the device) and because it's often difficult to determine loss vs. theft. Please keep this in mind as you view the top actions and assets in this section.

Misdelivery (sending paper documents or emails to the wrong recipient) is the most frequently seen error resulting in data disclosure.

There are only two difficult problems in computer science: cache invalidation, naming things, and off-by-one errors. Misdelivery (sending paper documents or emails to the wrong recipient) is the most frequently seen error resulting in data disclosure. One of the more common examples is a mass mailing where the documents and envelopes are out of sync (off-by-one) and sensitive documents are sent to the wrong recipient. A mundane blunder, yes, but one that very often exposes data to unauthorized parties.

Figure 43.
Top 10 threat action varieties within Miscellaneous Errors (n=558)

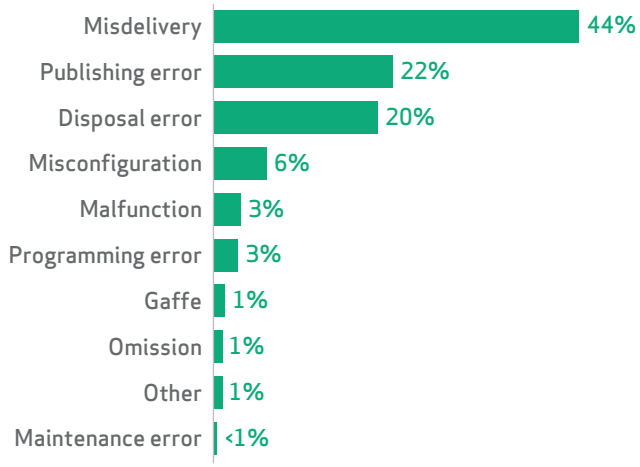
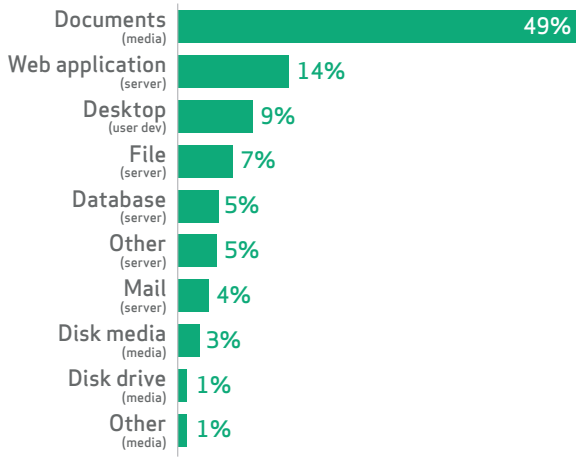


Figure 44.
Top 10 assets affected within Miscellaneous Errors (n=546)



GOVERNMENT MISDELIVERY

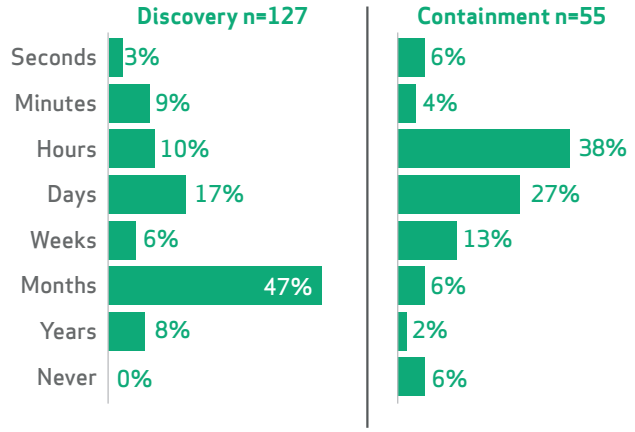
According to our sample, government organizations frequently deliver non-public information to the wrong recipient; so much so, in fact, that we had to remove it from Figure 43 so that you could see the other error varieties. Why is that number so large? The United States federal government is the largest employer in that country, and maintains a massive volume of data on both its employees and constituents, so one can expect a high number of misdelivery incidents. Public data laws and mandatory reporting of security incidents also cover government agencies. Since we have more visibility into government mistakes, it creates the impression that government mistakes happen more frequently than everyone else's, which may not be the case. This is not unlike the way we see higher numbers of overall breaches in U.S. states that have had disclosure laws on the books the longest. Case in point: even with government misdelivery removed from the results, misdelivery still dominates the list of errors resulting in exposed data.

Oxford Dictionaries declared that "selfie" was 2013's word of the year,[13] but did you know that posting content to the web and later regretting it was a meme in the corporate world too? That's right, the second most frequent error variety is publishing errors, which often involve accidentally posting non-public information to a public resource, such as the company web server. That's why web application takes the number two spot on the affected assets chart (Figure 44). Rounding out the top three in this category is disposal error, where the affected asset is thrown away without being shredded or, in the case of digital media, properly cleared of sensitive data.

Who's making all these mistakes? Well, it's almost entirely insiders, of course. End-users, sysadmins, and developers lead the pack when it comes to mucking things up, though pretty much all of us are guilty.

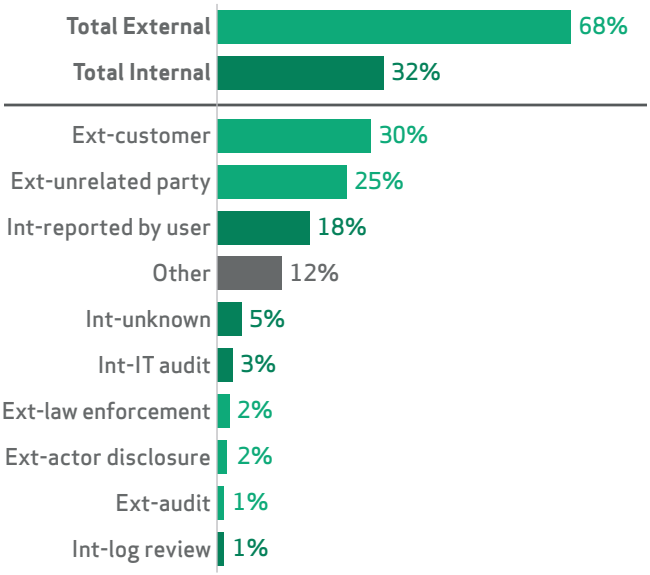
Who's making all these mistakes? Well, it's almost entirely insiders, of course. End-users, sysadmins, and developers lead the pack when it comes to mucking things up, though pretty much all of us are guilty. But the interesting thing is that there's quite a large number of incidents (70) caused by partner errors – more than any other pattern.

Figure 45.
Discovery and containment timeline within Miscellaneous Errors



Organizations only discover their own mistakes about one-third of the time. Otherwise, an external entity makes them aware of the incident, and most frequently it's the organization's own customers. You could try the "Inconceivable!" tactic when a customer calls to say they found their unprotected personal data on your website — but if you keep using that word, they'll figure out it doesn't mean what you think it means.

Figure 46.
Top 10 discovery methods for Miscellaneous Error incidents (n=148)



RECOMMENDED CONTROLS

Bob Ross, everyone's favorite painter of fluffy little clouds, once said, "We don't make mistakes, we just have happy accidents." Even still, organizations can take steps to decrease the frequency of all manner of accidents by reducing their exposure to the common error patterns that result in data disclosure.

Keep it on the DLP

Consider implementing Data Loss Prevention (DLP) software to reduce instances of sensitive documents sent by email. DLP can identify information that follows a common format, such as credit card numbers, social security numbers, or medical billing codes.

Check out the pub

Decrease the frequency of publishing errors by tightening up processes around posting documents to internal and external sites. For example, have a second reviewer approve anything getting posted to company servers, develop processes to regularly scan public web pages for non-public data, and implement a blanket prohibition against storing un-redacted documents on a file server that also has a web server running. It's amazing how easy it is for a spreadsheet to migrate over to the htmldocs folder. Make sure there's a process to test the security controls after a change — we've often seen a failure to put controls back in place result in a publishing breach.

Nail the snail mail fail whale

Say that three times really fast. When sending large postal mailings (also prone to error at higher speed and repetition), spot-check a sample to ensure that the information in the document matches the name on the envelope. Watch out for window envelopes too – sometimes that window might be too big or your content might not be centered properly, allowing sensitive information to show through. A lot of these incidents could have been prevented if someone had popped a few envelopes off the stack and inspected them before they went in the mail.

IT don't make trash

IT burns it. Any disposal or sale of information assets should be coordinated by the IT department. Educate users to think of disposing of a computer the same way they think of disposing of hazardous materials. "You can't just throw that in the trash (or sell it on eBay)! Send it to IT for proper handling." Test the disposal process by sampling devices to verify they've been sanitized properly. If a third-party handles this, ensure that contracts stipulate how to transfer, store, and dispose of data, along with roles, responsibilities, verification, and penalties for non-compliance.

IMPRISONED RESPONSE: THE FUTURE OF IR?

An example from the VCDB shows how bad things can get when document disposal goes wrong. A medical center arranged to have a vendor pick up documents and shred them prior to disposal. Apparently, an actual "pickup" truck was used, because the files ended up all over the roadside instead. "It looked like a blizzard of white paper had struck the area," according to one witness. These were old medical records with all manner of protected information. When people found them and called law enforcement, an inmate crew doing regular trash pickup in the area was sent to retrieve these sensitive documents. And that's the sound of the men working on the "cha-ching" gang.

CRIMEWARE

AT A GLANCE

Description

Any malware incident that did not fit other patterns like espionage or point-of-sale attacks. We labeled this pattern "crimeware" because the moniker accurately describes a common theme among such incidents. In reality, the pattern covers a broad swath of incidents involving malware of varied types and purposes.

Top industries

Public, Information, Utilities, Manufacturing

Frequency

12,535 total incidents

50 with confirmed data disclosure

Key findings

The primary goal is to gain control of systems as a platform for illicit uses like stealing credentials, DDoS attacks, spamming, etc. Web downloads and drive-bys are the most common infection vectors.

Many incidents in this section come from our CSIRT partners, reflecting a roll-up across many victim organizations. The level of detail tends to be lower because there was no forensic investigation or similar in-depth analysis (or the report wasn't provided to the CSIRT), leaving VERIS metrics a bit sparse. But the high number of incidents still offers some insight into day-to-day malware infections where the victim's anti-virus (AV) and intrusion prevention system (IPS) shields could not repel firepower of that magnitude.

As expected, this incident pattern consists mainly of opportunistic infections tied to organized criminals with some kind of direct or indirect financial motive (hence the title "crimeware"). Once malicious code has acquired a level of access and control of a device, the myriad possibilities to make a buck are opened up for the attacker.

In not-so-shocking news, Zeus continues to be a favorite way to make a buck with crimeware in 2013 (see sidebar for more detail). Zeus and its offspring, Citadel, primarily focus on stealing money via bank account takeovers, though they can also be used for other functions. Zitmo ("Zeus in the Mobile") also shows up in the data.[14]

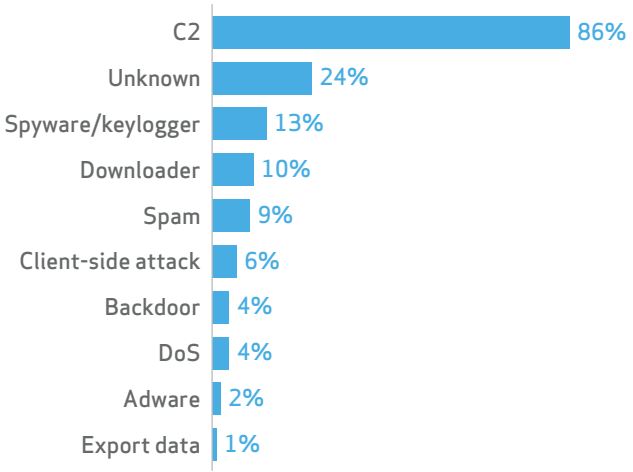
ZEUS

Zeus (sometimes called "Zbot") is sort of the cockroach of malware. It has managed to survive and even thrive despite many attempts to eradicate it. International arrests and the supposed retirement of the original author have not slowed it down, and once the source code behind it was published, other programmers could modify and extend Zeus for their own purposes, including evading antivirus software. In fact, Citadel started off as a variant of Zeus but has evolved substantially. Zeus can be used to install other malware but often grabs login and banking credentials from within browsers. Despite the efforts of many, it has continued to elude the good guys that are trying to shut it down.

This one primarily targets Android and Blackberry mobile devices for similar purposes. While Zeus serves as an example of crimeware families reported all around the world, others had a more localized presence. Nitol, for instance, was quite common among incidents reported to MyCERT of Cybersecurity Malaysia, but we have no instances of it infecting systems outside Asia. Nitol allows backdoor access and frequently causes infected systems to participate in DDoS attacks.

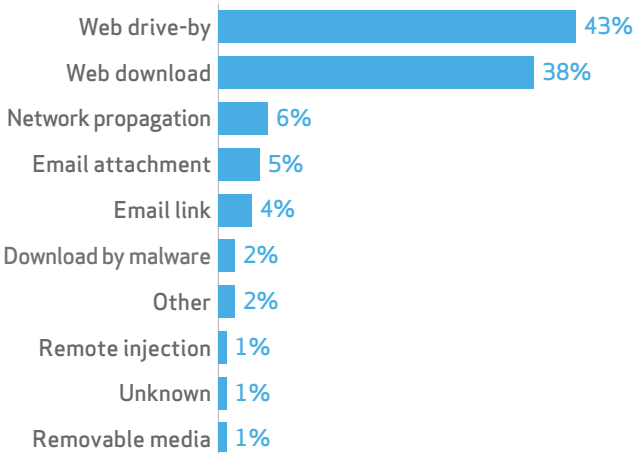
Expanding online markets, where specialists offer cybercrime-as-a-service, became a growing trend in 2013. A good example in the Netherlands was the wave of DDoS attacks on banks and specific institutions since March, 2013. So-called "booter websites" have made this type of attack available to literally anyone who wants to attack a company or institution. Naturally, a host of other malware families made appearances last year, but these two stood out to us as worthy of a brief mention.

Figure 47.
Top 10 threat action varieties within Crimeware (n=2,274)



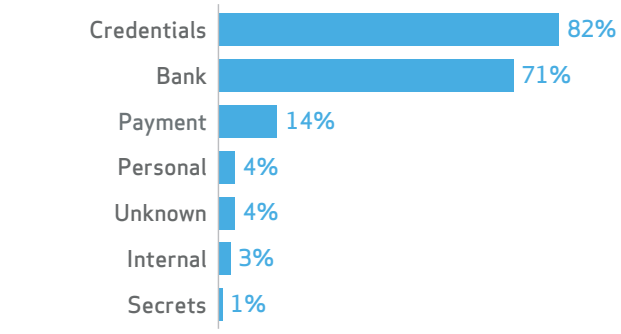
Victims don't always report malware functionality, but when they do, they prefer C2 (according to the most interesting CSIRTs in the world, at least). This makes perfect sense, as the goal is to achieve and maintain control of a device to command it to do your bidding. Whether the little compromised minions are participating in a spam botnet, stealing banking credentials, or hijacking a browser to artificially boost ad revenue, there are numerous ways to leverage compromised workstations that don't entail deeper penetration into a network.

Figure 48.
Top 10 vectors for malware actions within Crimeware (n=337)



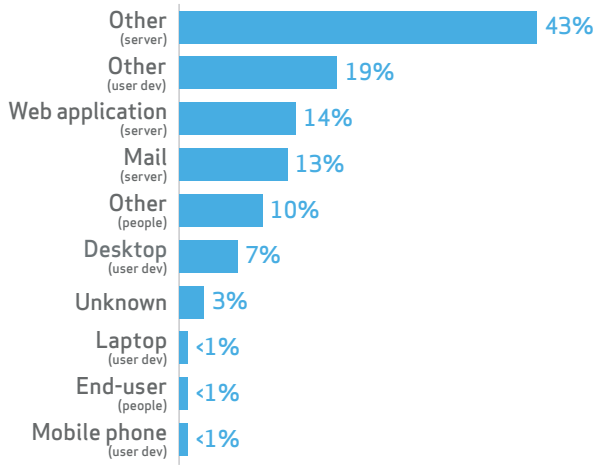
The majority of crimeware incidents start via web activity — downloads or drive-by infections from exploit kits and the like — rather than links or attachments in email.[15] Adware still shows up, though Bonzi Buddy thankfully remains extinct. For malware with a social engineering component, both scams and phishing play important roles[16] Infected assets usually weren't identified, but interestingly, those that were reported more servers than user devices. Wow. So vectors. Much families. Many incident.

Figure 49.
Variety of at-risk data within Crimeware (n=73)



Crimeware incidents are light on timeline and discovery details because the response is often to just wipe the system and get it back to work (remember, this pattern comprises a lot of one-off infections that don't fit other patterns). When known, notification by unrelated third parties (namely CSIRTs) were by far the most common way victims learned of the incident.

Figure 50.
Top 10 assets affected within Crimeware (n=1,557)



Like us, your first reaction might be "why not technologies like IDS and AV?" This reflects the role of CSIRTs as the primary provider of crimeware incidents in this dataset. The discovery method wasn't known for 99% of incidents; it's not usually within their visibility or responsibility. For all we know, CSIRTs only saw the 1% not discovered by AV or IDS. The discovery timeline in Figure 52 hints that this might, in fact, be the case. Notice the difference in N between Figure 51 and Figure 52 and how many infections are discovered within seconds — only automated detection methods would be so quick.

Figure 51.
External vs. internal discovery methods within Crimeware (n=183)

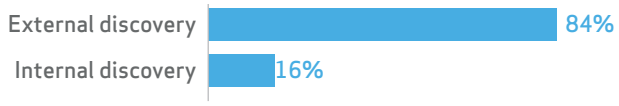
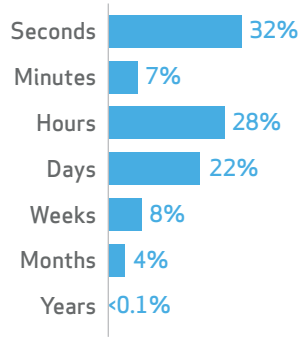


Figure 52.
Discovery timeline within Crimeware (n=1,017)



RECOMMENDED CONTROLS

These results led us to develop some specific recommendations to help keep incidents of crimeware down. The natural question to ask is "what happened to antivirus?" AV technologies play an important role in catching many types of commodity malware and preventing compromise in the first place. So this pattern reflects a reverse sort of "survivorship bias" in which we look primarily at the sorts of things AV does not do as well — or organizations that don't do AV particularly well. Nitol, in particular, infected many systems at the factory before shipping and likely before users or administrators had deployed any sort of AV. The Zeus and Citadel family has a well-deserved reputation for evolving quickly to evade signature-based detection of the sort used by many AV products.

Keep browsers up to date

Zeus frequently uses a technique called "man in the browser" that involves using browser vulnerabilities and add-on functions. Keeping browsers and plugins secure will go a long way toward reducing the impact of this sort of incident. Apply browser patches as quickly as software producers make them available.

Disable Java in the browser

Legacy apps may complicate this, but if possible, avoid using Java browser plugins, given the difficulty in sandboxing content and the history of vulnerabilities here.

Use two-factor authentication

Our results link crimeware to stolen credentials more often than any other type of data. This points to the key role of crimeware when the attack objective is to gain access to user accounts. Two-factor authentication won't prevent the theft of credentials, but it will go a long way toward preventing the fraudulent re-use of those credentials.

Change is good…except when it isn't

Consider how best to deploy system configuration change monitoring. Unlike iocane powder, many of the vectors and persistence methods used by crimeware can be easily detected by watching key indicators on systems. This goes to the general theme of improving detection and response rather than solely focusing on prevention.

Leverage threat feeds

Given the high incidence of C2 communications, using feeds of threat data that identify IP addresses and domain names used to control botnets, then matching this data against firewall or proxy logs can help accelerate detection and thus containment. We generally don't recommend using these lists for outright blocking due to possible operational issues. But malware researchers do a fine job of implementing sinkholes and reverse-engineering malware quickly to identify infrastructure used by the bad guys.

A YEAR IN THE LIFE OF THE POLISH CERT

By the CERT Polska/NASK

The Internet threat landscape in Poland is largely defined by banking Trojans — crimeware aimed at stealing users' online banking credentials. These use a combination of social engineering and software vulnerabilities to gain access to a user's computer, and subsequently to their bank account. Whenever new financial malware or attack methods surface, Polish users are often among the first hit. 2013 also saw numerous financial malware botnets using Polish Internet properties for C2 purposes (including .pl ccTLD domain names). Over 20 such botnets were taken over or disrupted by CERT Polska.

The attacker's tool of choice is a variety of web-inject malware (with Zeus/Citadel being the most popular malware family), which infects a user's machine, and then injects code into the browser whenever that user visits a banking site deemed of interest (a "Man-in-the-Browser" attack). A common theme is the use of social engineering techniques to obtain credentials. For example, attempts are made to install one-time password stealers to intercept mobile transaction authentication numbers (mTANs) used by banks to authenticate transactions. In such cases, a user is prompted to provide his mobile number, supposedly to install a new security certificate designed by the bank on his smartphone – but what, in reality, is malware used to intercept and redirect text messages to the attacker. Cruder methods to subvert two-factor authentication are also employed: fictitious bank messages are injected, notifying the recipient of an erroneous bank transfer and asking for the money to be returned - to an attacker's account. Real world events are often exploited: a recent brand merger involving the largest Polish online bank resulted in attacks forcing users to redefine lists of permanent transfers – redirecting them to an attacker's bank numbers – under the auspices of changes resulting from the merger.

However, it's not just web-inject malware at work: other tricks observed in 2013 include malware that switches bank account numbers to those of the attacker during a copy/paste operation in Microsoft Windows. It's not always about malware either: late 2013 saw large scale attacks against home routers, which had their DNS server settings subsequently reconfigured to point at rogue DNS servers. These were then used to perform Man-in-the-Middle attacks through a series of proxies, subverting SSL and two-factor authentication mechanisms by using social engineering methods similar to those described above.

PAYMENT CARD SKIMMERS

AT A GLANCE

Description

All incidents in which a skimming device was physically implanted (tampering) on an asset that reads magnetic stripe data from a payment card (e.g., ATMs, gas pumps, POS terminals, etc.).

Top industries

Finance, Retail

Frequency

130 total incidents[17]

130 with confirmed data disclosure

Key findings

There's not a ton of variation in this pattern at the VERIS level: criminal groups install skimmers on ATMs (most common) and other card swipe devices. On a more qualitative level, the skimmers are getting more realistic in appearance and more efficient at exporting data through the use of Bluetooth, cellular transmission, etc.

For a wide array of criminals ranging from highly organized crime rings to garden variety ne'er-do-wells who are turning out no good just like their mama warned them they would, skimming continues to flourish as a relatively easy way to "get rich quick." While most incidents are linked to Eastern European actors, nearly all victims of payment card skimmers in this report are U.S. organizations (the U.S. Secret Service and public disclosures being the primary sources for this data). While some don't think we should include this type of attack in the DBIR, we can't justify excluding a tried-and-true method used by criminals to steal payment card information.

Figure 53.
Origin of external actors within Card Skimmers (n=40)

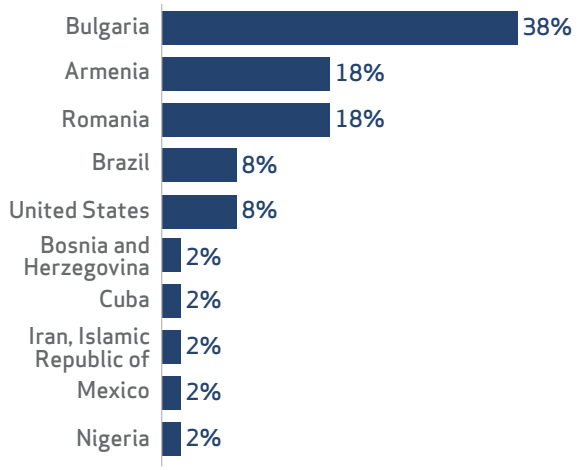
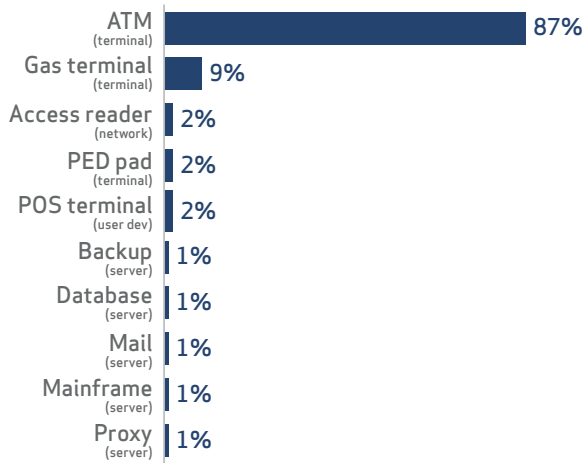


Figure 54.
Assets affected within Card Skimmers (n=537)

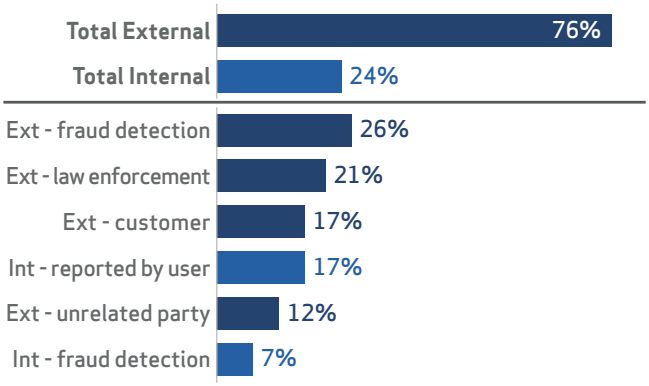


In 2013, most skimming occurred on ATMs (87%) and gas pumps (9%) due to the relative ease with which they can be approached and tampered with. Gas pump skimmers are often installed by a small group of people acting in concert. One scenario involves one or more conspirators going into the station to make a purchase and distract the cashier's attention, while a partner in crime plants the device inside the machine using a universal key.

ATM skimmers, on the other hand, are installed on the outside of the machine. While some ATM skimming devices are clunky homemade affairs that might afford an opportunity for observant customers to spot them, the design of many skimmers (both those created by the criminal and those purchased "off the shelf") can be so realistic in appearance that they are virtually invisible to the end user. In most cases they can be snapped in place in a matter of seconds and can be produced in sufficient quantities to make the attacks scalable and highly organized. This, however, has been the norm for some time and warrants only a cursory mention in this report. What has changed over time, however, are the methods by which the data is retrieved by the criminals.

In the past it was necessary for the criminal to return to the scene and physically remove the device in order to collect the stolen data. While we continue to see this, it's normally indicative of the less organized and more small time crooks out to "make some sweet moolah with Uncle Rico." They're often apprehended when retrieving the skimmed card data. In keeping with what we find with network-based attacks, the successful criminal is the one who can maintain a safe distance between themselves and the target. Therefore, the more highly skilled criminals now collect data via Bluetooth or SIM cards with remote caching and tampering alerts. Some devices actually send an SMS alert to the criminal each time the ATM is used.

Figure 55.
Discovery methods within Card Skimmers (n=42)



With subterfuge and fraud being the objectives behind skimming, it's not surprising that it's most commonly detected by a third party. Most of the time that third party is a payment card company or a customer who has noticed fraudulent activity. Other times it's a phone call from a law enforcement agency after they've arrested a gang with a trunk full of skimming devices and white plastic cards. Hanging close to that pack of external discovery methods are internal users who spot tampering and report it to management. Way to go, folks. As skimmers become more difficult to detect visually, however, we can't help but wonder if this latter scenario will become increasingly rare.

EVOLUTION OF SKIMMING

Like any technology, the tendency is to develop from bulky and slow toward streamlined and efficient. Skimming devices are no different. Many people still think of the traditional skimmer as the classic wedge – the small hand-held skimmer typically used by waitstaff to illicitly obtain mag stripe data while they had the card away from the customer. Because they were so easy to use, they became the stock-in-trade for most criminals for a long time.

On the flip side, it was often relatively easy for the good guys to pinpoint the culprit after the fraud had transpired. Common Point of Purchase (CPP) algorithms could be used to determine the restaurant responsible for the fraudulent charges. When law enforcement arrived at the restaurant they could obtain access to the receipts, and with relative ease determine that the same waiter/waitress served all the victims. The individual would then be interviewed and… well, you know the rest.

Fast-forward to the year 2000, when the first gas pump skimmer was found at a gas station in California. The skimmer was placed inside the pump, and (since it only captured track information) the criminals set up a wireless video camera 300 yards away in a weatherproof case. In this particular instance, the camera was discovered and unplugged by an investigator. Within minutes, the bad guys showed up at the gas station to see what went wrong, and were promptly taken into custody.

Eventually the risk of discovery during retrieval became too great, so more criminals started manufacturing skimmers and selling them online. This new wave of devices was Bluetooth equipped, which allowed someone to download the track and PIN data from the safety of the parking lot.

It's now possible to buy online skimming devices with built-in SIM cards that allow for remote configuration, remote uploading of data, and tampering alerts that, if triggered, will cache the data and send it out immediately, greatly reducing risk.

RECOMMENDED CONTROLS

Though some might argue, we find no obvious mistake or oversight on the part of organizations that allows skimming to succeed when it otherwise wouldn't. But there are some things that can be done to make it harder for the criminal and shorten the window of exposure.

FOR BUSINESSES

Design (or buy) tamper-resistant terminals

As the merchant, this probably isn't something you can do yourself, but be aware that certain designs are more susceptible to skimming devices than others. Many modern ATMs are designed with this in mind; choose those if possible.

Use tamper-evident controls

Do things that make it obvious (or send an alert) when tampering occurs. This may be as simple as a sticker over the door of a gas pump or more sophisticated tactics like visual anomaly monitoring on ATMs.

Watch for tampering

Regularly check terminals for signs of unauthorized tampering. Also train employees to spot skimmers and recognize suspicious behavior from individuals trying to install them. If a criminal is able to place a skimmer on one of your devices, these regular inspections will help curb the damage.

FOR CONSUMERS

Protect the PIN

When entering your PIN, cover your hand to block tiny cameras that may be recording it. You wouldn't want a ne'er-do-well getting ahold of your PIN now, would you?

Trust your gut

If something looks out of the ordinary at your ATM or gas pump, something fishy may be afoot. While criminals are increasingly sly at designing difficult-to-detect skimmers, you still might be able to notice something amiss, especially if the terminal looks different than others around it. If one of these things is not like the others, don't swipe your card!

See something, say something

If something seems out of place to you at a payment terminal, don't keep it to yourself. Be sure to tell the merchant or bank that you may have found a skimmer. Not only will you be helping them, you'll also be helping your fellow consumers.

CYBER-ESPIONAGE

AT A GLANCE

Description

Incidents in this pattern include unauthorized network or system access linked to state-affiliated actors and/or exhibiting the motive of espionage.

Top industries

Professional, Transportation, Manufacturing, Mining, Public[18]

Frequency

511 total incidents

306 with confirmed data disclosure

Key findings

Most surprising to us is the consistent, significant growth of incidents in the dataset. We knew it was pervasive, but it's a little disconcerting when it triples last year's already much-increased number. Espionage exhibits a wider variety of threat actions than any other pattern. The most evident changes from our last report include the rise of strategic web compromises and the broader geographic regions represented by both victims and actors.

Comprehensive information about "cyber"[19] espionage is really hard to come by. Organizations typically aren't required to publicly disclose breaches of internal information and trade secrets, as they are with regulated consumer data. Additionally, there's no fraud algorithm to alert victims about illicit use of such data, leaving many cases of espionage undiscovered. Most of what we know publicly about this genre of threat comes from incident responders, intelligence analysts, and malware researchers who compile and share their knowledge with the community. Thus, we're excited to have quite a few contributors from these circles, whose information has more than tripled the number of espionage incidents in this year's dataset, to 511.

Before someone concludes we're asserting a vast increase in espionage in 2013, we're quite sure countless organizations have been consistently targeted for several years. Instead, we attribute this increase primarily to our ever-expanding set of contributors conducting research in this area, along with more community information sharing that improves discovery capabilities. Like a streetlight illuminating cars parked along the street, more contributors allow us to see more cars. Unfortunately, we can also see that those cars have broken windows and stolen stereos.

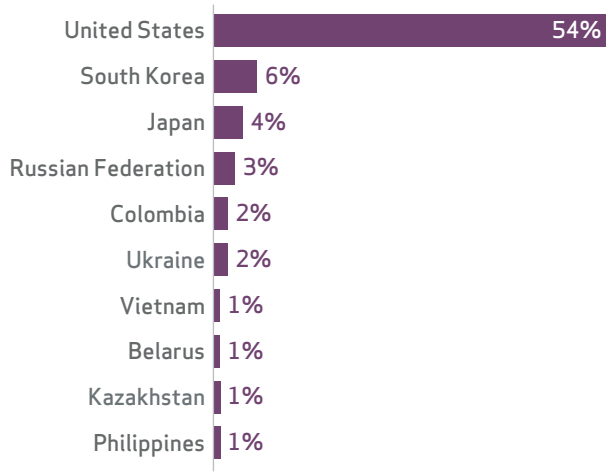
Figure 56.
Number of incidents by victim industry and size within Cyber-espionage

Industry	Total	Small	Large	Unknown
Administrative [56]	2	1	1	0
Construction [23]	1	0	0	1
Education [61]	2	1	1	0
Finance [52]	3	0	2	1
Healthcare [62]	2	1	0	1
Information [51]	11	2	2	7
Management [55]	2	1	1	0
Manufacturing [31,32,33]	81	5	17	59
Mining [21]	5	0	2	3
Professional [54]	114	11	5	98
Public [92]	133	20	19	94
Real Estate [53]	1	1	0	0
Retail [44,45]	1	0	1	0
Transportation [48,49]	5	1	3	1
Utilities [22]	8	0	1	7
Other [81]	5	5	0	0
Unknown	135	0	3	132
Total	**511**	**49**	**58**	**404**

For more information on the NAICS codes [shown above] visit:
https://www.census.gov/cgi-bin/sssd/naics/naicsrch?chart=2012

To set the tone, we need to understand the victims represented within the data. We don't claim to cover all espionage activity in 2013 — quite far from it, actually. As is evident in Figure 57, the sample is still largely (over half) U.S. based, but not as exclusively as in previous years. We expect this to continue as more global organizations join the cause. We can't help but wonder why we have no examples of Italian victims of espionage in our dataset. Our best hypothesis is that sophisticated actors remember the classic blunder of "go[ing] in against a Sicilian when death is on the line" when selecting targets (the most famous blunder, of course, is getting involved in a land war in Asia).

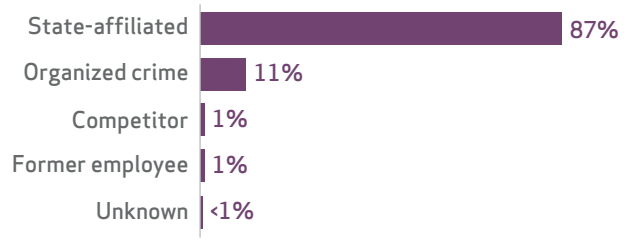
Figure 57.
Victim country within Cyber-espionage (n=470)



In addition to geographic broadening, we see a wide distribution of both sizes and types of victim organizations. Unfortunately, victim size is often not tracked, so there are a lot of unknowns here. Insofar as we can determine from the data before us, however, size doesn't seem to be a significant targeting factor. Industry, on the other hand, does: the Public, Professional, and Manufacturing sectors are more targeted by espionage than the rest of the field (which still runs a fairly wide gamut). There is little doubt that figures for the Public sector, which spans embassies, economic programs, military, and other support organizations, are boosted by our government contributors. There is also little doubt that they are a prime target for espionage. Victims within the Professional, Scientific, and Technical Services category typically deal with custom computer programming services, research and development, engineering and design, and legal practices. Many of these organizations are targeted because of the contracts and relationships they have with other organizations. For some, they can serve as both a valuable aggregation point for victim data and a trusted exfiltration point across several target organizations. Lastly, and not unexpected, Manufacturing industries are also targeted for their intellectual property, technology, and business processes.

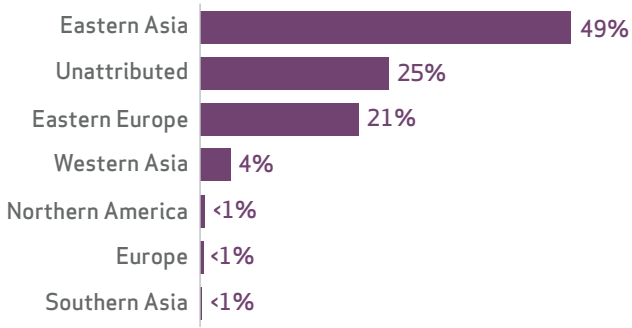
Attribution is also probabilistic in nature. Be wary of threat intelligence vendors claiming to be 100% sure an attack is X actor group from Y country with Z motives; they are "likely" incorrect. There are many methods for determining attribution — sometimes it's following the breadcrumbs left by the actors. Other times it's ruling out the alternatives using something like analysis of competing hypothesis.[20] None of these methods are perfect. It's important to carefully evaluate information to make sure one isn't suffering from some type of cognitive bias.[21] It would be more helpful if probabilistic language like Sherman Kent's "Words of Estimative Probabilities"[22] was used when describing attribution to particular countries, regions, and threat actors. With that in mind, the following would fall between "Probable" and "Almost Certain."

Figure 58.
Variety of external actors within Cyber-espionage (n=437)



As expected, most incidents in this category are attributed to state-affiliated actors. But the data also reminds us that organized criminal groups, competitors, and current[23] and former employees join in the game too. We also see that the longer game of espionage is not always the sole motive; it often exhibits a nearer-term, more direct financial element as well. An example would be a mercenary-style theft of source code or digital certificates contracted by a rival organization or other interested party.

Figure 59.
Region of external actors within Cyber-espionage (n=230)

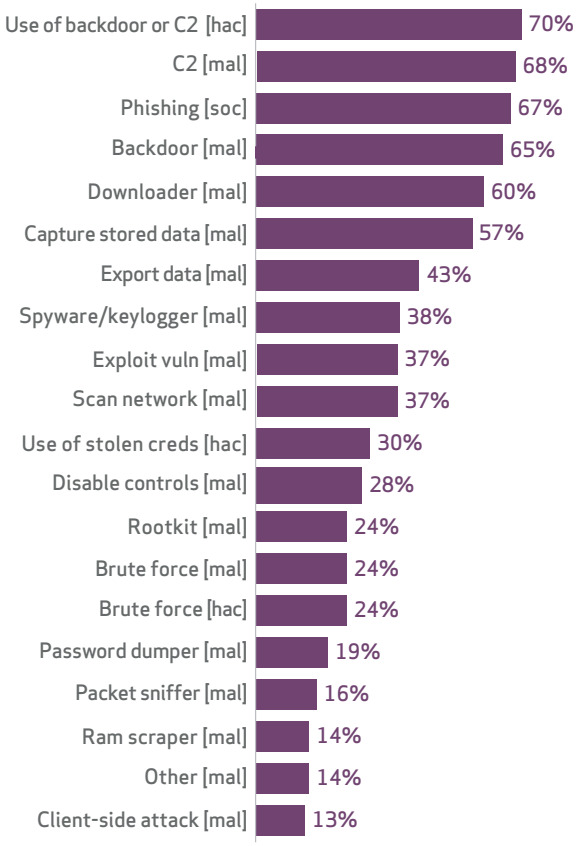


With respect to actor origin, the percentage of incidents attributed to East Asia is much less predominant in this year's dataset. Two countries in particular, the People's Republic of China and the Democratic People's Republic of Korea, represent that region. This underscores the point we made in our last report – that, despite our China-exclusive results, China definitely was not the only country conducting espionage.

The 2013 dataset shows much more activity attributed to Eastern European actors, Russian-speaking ones in particular. As before, we don't propose these are the only active regions/countries engaged in espionage. More comprehensive research into different actor groups is continually driving better detection and attribution, and we hope future versions of this report will show the fruits of those efforts. At a high level, there doesn't seem to be much difference in the industries targeted by East Asian and Eastern European groups. Chinese actors appeared to target a greater breadth of industries, but that's because there were more campaigns attributed to them.

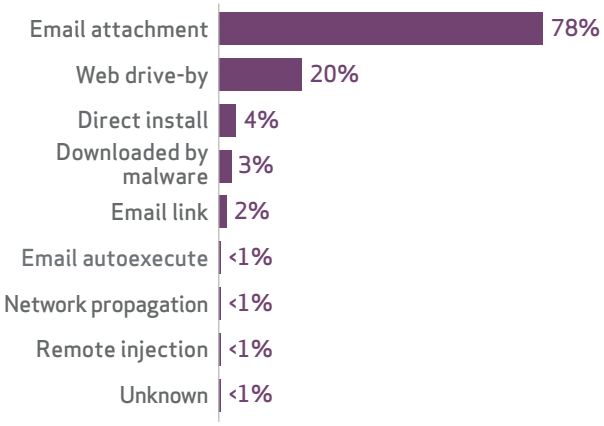
One aspect of this pattern that sets it apart from others is the wide variety of threat actions. Many of the other patterns have simpler stories with relatively few VERIS actions. Espionage breaks that mold in a big way, though the specific actions involved won't be a surprise to many readers. State-affiliated groups often deploy a wide range of tools (or tools that have wide range of capabilities), which is evident in Figure 60.

Figure 60.
Top threat action varieties within Cyber-espionage (n=426)



It's interesting that, while the array of tools is diverse, the basic methods of gaining access to a victim's environment are not. The most prolific is the old faithful: spear phishing. We (and others) have covered this ad nauseam in prior reports, but for both of you who have somehow missed it, here goes: A well-crafted and personally/professionally-relevant email is sent to a targeted user(s), prompting them to open an attachment or click a link within the message. Inevitably, they take the bait, at which point malware installs on the system, a backdoor or command channel opens, and the attacker begins a chain of actions moving toward their objective. The proportion of espionage incidents incorporating phishing is lower than our last report (it was 95%), but not because of a drop in actual frequency. This is primarily due to a big increase in the use of strategic web compromises (SWCs) as a method of gaining initial access.

Figure 61.
Vector for malware actions within Cyber-espionage (n=329)



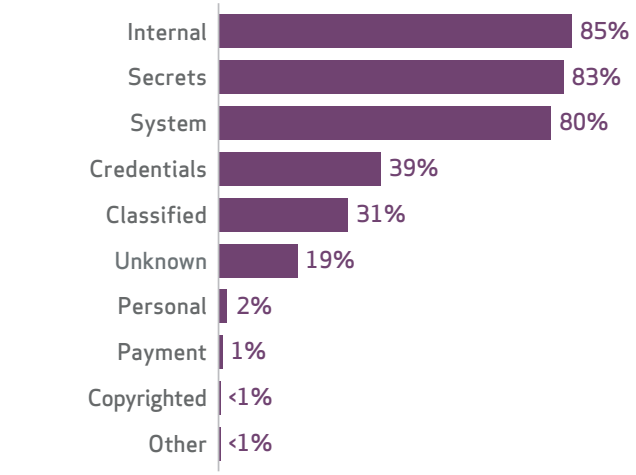
Instead of email bait, SWCs set a trap within (mostly) legitimate websites likely to be visited by the target demographic. When they visit the page, the trap is sprung, the system infected, and the rest is the same as described above. Even if detected quickly, SWCs can provide a very high reward for attackers. Furthermore, the industry has observed some maturation of the SWC technique, which assists the actors in focusing their targets and avoiding detection (see sidebar on next page for more on SWCs).

CAMPAIGN RESEARCH PUBLISHED IN 2013

The DBIR focuses on overall trends and stats related to espionage campaigns. Several of our contributors have published in-depth research on specific actors and campaigns, some examples are listed below:

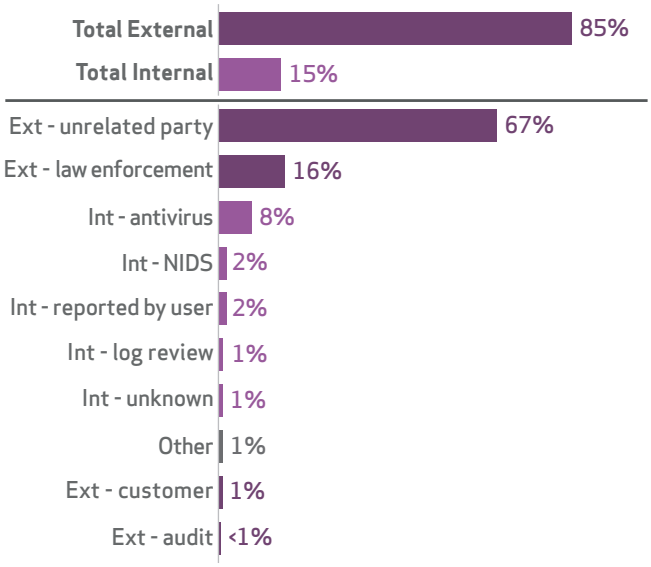
- Deputy Dog (FireEye), August-September 2013
- Ephemeral Hydra (FireEye), November 2013
- MiniDuke (Kaspersky), February 2013
- Red October (Kaspersky), May 2007-January 2013
- Sunshop (FireEye), September 2011-October 2013 (But likely ongoing)
- Troy (McAfee, part of Intel Security), January–March 2013
- Multi-campaign (U.S. Defense Security Service), "Targeting U.S. Technologies"

Figure 62.
Variety of at-risk data within Cyber-espionage (n=355)



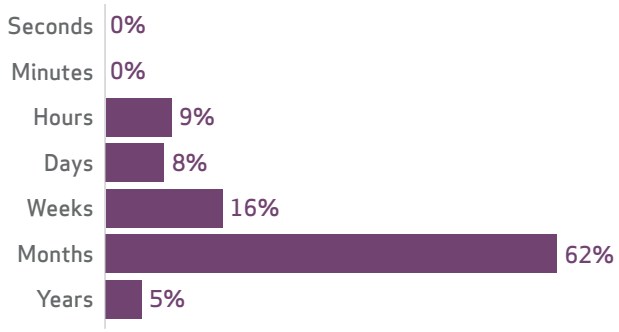
Once the phishing email or SWC has done its work, and an internal system is infected, the name of the game is moving determinedly through the network to obtain the prize. This may happen quickly, but it also may last for years. Common methods involving loading backdoors on systems to maintain access, dropping spyware/keyloggers and password dumpers to steal user credentials, and then using those credentials to elevate privileges and expand control.

Figure 63.
Top 10 discovery methods within Cyber-espionage (n=302)



Examining discovery timelines and methods for espionage incidents reveals ample room for improvement. While this information is often not known or provided (for various reasons, including the visibility and focus of our contributors), there's enough to discern the general state of affairs. It typically takes victims months or more to learn they've been breached and it's usually an outside party notifying them.

Figure 64.
Discovery timeline within Cyber-espionage (n=101)



The most common method of discovery is ad hoc notification from threat intelligence and research organizations that observe, for instance, the victim communicating with C2 infrastructure of a known threat group. While this isn't good news per se, it does suggest intelligence operations are an important tool for combating espionage.

TOOLS OF THE TRADE: STRATEGIC WEBSITE COMPROMISE

Strategic website compromises (SWCs) have proven to be an effective tactic of state-affiliated threats to infiltrate the networks of target organizations. In 2012, SWCs made their debut with the "VOHO Affair"[24] and continued in 2013 with attacks focused against the Public, Manufacturing, Professional, and Technical sectors.

SWCs leverage websites that are of critical or complementary value to an industry's line of business to distribute malware traditionally contained in spear phishing emails. Visitors are hit with a drive-by download, granting attackers access/ownership of the system. State-affiliated SWCs in 2013 exhibited three new browser-based zero-day vulnerabilities (constituting over 75% of publicly disclosed SWCs), which upped the rate of compromise per event.

So, why has the use of SWCs in espionage campaigns increased? Well, there's no doubt that attackers have realized this tactic scales well and provides reasonable assurances of ambiguity. By opting out of direct attacks like phishing, attackers effectively remove themselves from the tribulations of poor grammar, scanners, and astute users. And by leveraging zero-day exploits, they achieve higher success rate that no longer rely on carefully coerced actions.

In 2014, we'd like to predict SWCs will fade, but that seems unlikely. While there are downsides to SWCs for the attackers (high visibility and high cost to weaponize and burn a zero day), the benefits of a low-cost way to support long-term operations generally outweigh the risks.

RECOMMENDED CONTROLS

Isolating the root cause of an espionage-related breach is a bit of a snipe hunt. Sure, victims make mistakes (minor and otherwise) that are exploited in the process, but the real root issue is a determined, skillful, patient, and well-resourced adversary who will keep poking until he finds (or makes) a hole. With that in mind, let's take a closer look at the holes and other thin spots these adversaries often take advantage of.

First, we'll start with a few blocking and tackling fundamentals that you really ought to be doing regardless of whether or not you're worried about espionage. If you don't do these, all those super-advanced cybertastic APT kryptonite solutions may well be moot.

Patch ALL THE THINGS!

Exploiting browser, OS, and other third-party software (e.g., Flash and Java) vulnerabilities to infect end-user systems is a common initial step for attackers. Keeping everything up to date will make that step a lot harder to take.

Use and update anti-virus (AV)

While many proclaim AV is dead, not having it is akin to living without an immune system. It might not protect you from the dreaded zero day, but let's be honest — many espionage victims still fall to one-zero-zero days (or higher). An up to date AV (in-line and on the endpoint) can go a long way to detect anomalies in applications and find pesky shells and other malware.

Train users

Some will consider this a lost cause, but we counter with a reminder that, over the years we've done this research, users have discovered more breaches than any other internal process or technology. It's not all about prevention; arm them with the knowledge and skills they need to recognize and report potential incidents quickly.

Segment your network

Good network and role segmentation will do wonders for containing an incident, especially where actors intend to leverage access to one desktop as a stepping-stone to the entire network.

Keep good logs

Log system, network, and application activity. This will not only lay a necessary foundation for incident response, but many proactive countermeasures will benefit from it as well.

Beyond the basics, there are some specific practices that organizations concerned with state-affiliated and other determined adversaries should consider. These roughly follow critical points in the path of attack, where victims have the best chance to recognize and respond.

Break the delivery-exploitation-installation[25] chain

Users will be phished, and they will eventually click; we've got the data to prove it. Focus on implementing a solution that more completely defends against phishing, such as not relying solely on spam detection and blocklists, but also doing header analysis, pattern matching based on past detected samples, and sandbox analysis of attachments or links included.

For more mature organizations, check out the growing collection of Data Execution Prevention (DEP) and Endpoint Threat Detection and Response (ETDR) solutions. We don't promote specific products in this report, but you'll find some good options in this space by starting your search with some of our contributors.

Spot C2 and data exfiltration

Collect and/or buy threat indicator feeds. In and of themselves, they aren't intelligence, but they're certainly useful within intelligence and monitoring operations.

Monitor and filter outbound traffic for suspicious connections and potential exfiltration of data to remote hosts. In order to recognize "abnormal," you'll need to establish a good baseline of what "normal" looks like. Those indicators you collected/bought will come in handy here.

Monitor your DNS connection, among the single best sources of data within your organization. Compare these to your threat intelligence, and mine this data often.

Stop lateral movement inside the network

After gaining access, attackers will begin compromising systems across your network. ETDR, mentioned above, can help here too.

Two-factor authentication will help contain the widespread and unchallenged re-use of user accounts.

We mentioned network segmentation in the basics, but since doing it well is challenging, we'll mention it here again. Don't make it a straight shot from patient zero to a full-fledged plague.

Watch for user behavior anomalies stemming from compromised accounts.

DENIAL OF SERVICE ATTACKS

AT A GLANCE

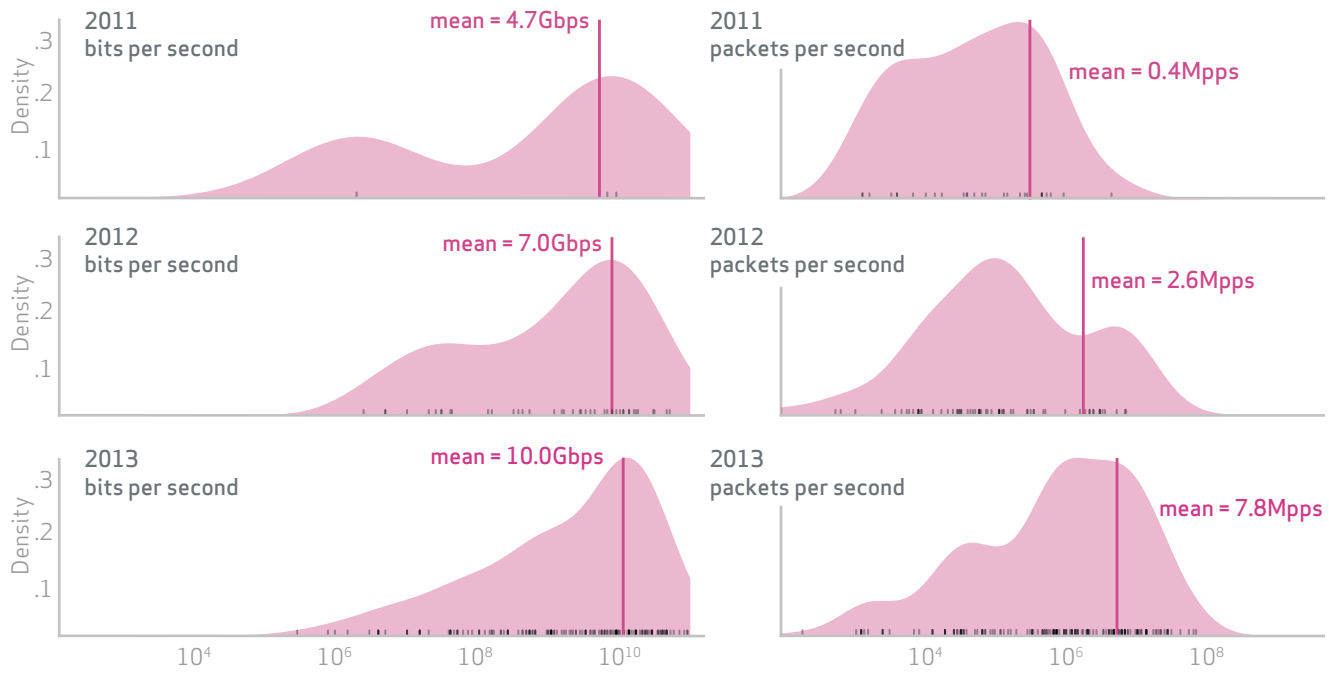
Description	Any attack intended to compromise the availability of networks and systems. Includes both network and application layer attacks.
Top industries	Finance, Retail, Professional, Information, Public
Frequency	1,187 total incidents 0 with confirmed data disclosure
Key findings	The headliner for DDoS in our 2013 dataset was the QCF campaign against the financial industry, which compromised vulnerable CMSs to create high-bandwidth attacks from hosting centers. DNS reflection attacks also became "big" but sightings of the DoS equivalent of Bigfoot (DDoS distractors covering up other nefarious activities) remain rare.

Hold on a second? DoS attacks? In a data breach report?

Knowing that these attacks are top of mind for many organizations — especially in light of late 2012/2013 events — we decided to expand the scope of the DBIR to include them. We collected quite a bit of data on the topic from multiple sources, including teams at Akamai and Verizon that spent a lot of time in the trenches fighting DDoS attacks in 2013. We could have renamed it Verizon's Security Incident Report (VSIR), but Microsoft has already laid claim to the "SIR"[26] title.

A new trend started developing in September of 2012. In the past, DOS attacks were primarily generated from compromised home computers or by willing participants. Think "your parents' desktop" system – you know, the one you're always cleaning up when you visit during the holidays. Obviously, such systems, reserved largely for normal home Internet users, have relatively small bandwidth from DSL or cable modems. The attackers then, harnessing their botnet of DoS drones, could send commands to direct attacks at a specific target. Flash forward to September 2012 and you see a different scenario altogether, along with a different method for building a better botnet. In this situation, attackers scanned for and exploited vulnerable websites and CMSs. Then they placed specific DOS attacks scripts onto these sites. The primary script used by these attackers is a customized version of a Russian kit known as Brobot or itsoknoproblembro. So what's different? Well, for starters these botnet drones aren't sitting on the Internet pipes of home broadband users.

Figure 65.
Denial of Service attack bandwidth and packet count levels 2011-2013



They're in hosted/cloud/private cloud/etc. data centers with high bandwidth pipes. The servers are also optimized for heavy traffic. Put these two together and you have the makings of what gamers might call a DoS BFG. Or, if music is your game — "these go to 11." High packet, high bandwidth attacks. In some of the Brobot attacks in the last year we saw upwards of 97 Gbps/100 Mpps attacks. These were some of the largest attacks we (and likely anyone else) have ever seen.

So who exactly was behind this new wave of DoS attacks? The simple answer is the Izz ad-Din al-Qassam Cyber Fighters (or QCF for short). The group first popped up in September 2012 with the stated goal of using DoS attacks to wreak havoc on U.S. financial institutions – part of a campaign they dubbed Operation Ababil. And they did just that; for several weeks near the end of 2012 and well into the first half of 2013, the QCF launched wave after wave of DoS attacks against U.S. banks using their powerful Brobot ion cannons (think Hoth).

All of this begs the question: Why was the QCF so determined to wage a campaign against prominent U.S. financials? If you believe the propaganda the group regularly posted to Pastebin during their attacks, then the answer to the question of what motivated them is simple: ideology. Like a broken record, the QCF repeatedly stated they would attack U.S. banks until all forms of a highly controversial and disparaging video named "Innocence of Muslims" was removed from YouTube. They even created a mathematical formula to convert the number of "likes" the video had to how long the campaign would continue.

While this was the show the QCF put on for public consumption, there were theories circulating in the security community that Operation Ababil was nothing more than a front for state-affiliated attackers based in Iran. Those theories created a VERIS dilemma about how to classify the QCF's actor variety. Are they truly hacktivists looking to get YouTube videos taken down or are they state-affiliated threat actors probing for weaknesses in the U.S. financial infrastructure at the bidding of the Iranian government? Unfortunately, the multilayer command and control infrastructure utilized in botnet creation makes it incredibly difficult to say with certainty from open sources that Iran is indeed the wizard behind the green curtain, so we ultimately decided to go with the publicly stated purpose of the actors and chalk it up to hacktivism.

While it's true that exactly what motivated the QCF isn't entirely certain, the tactics used to carry out the group's attacks are well known. Not only did the group use more traditional attacks such as UDP and SYN floods to clog up a target website's bandwidth and tie up server resources, it also carried out application-layer DoS attacks. In these low and slow attacks the QCF would send multiple HTTPS GET requests for PDF files on the target site. These types of attacks are especially frustrating: they don't require significant resources, they can be difficult to defend against, and they can be incredibly effective. The use of HTTPS is particularly problematic for mitigation because the packets are encrypted, which makes it difficult for defenders to determine junk traffic from legitimate traffic.

Speaking of DoS attacks that don't require a vast botnet to be devastating, we've observed another trend stealing the limelight recently: DNS reflection attacks. Not as clumsy or random as a botnet; these are an elegant weapon for a more civilized age. Remember the biggest DoS attack in history?[27] If not, allow us to refresh your memory. In March 2013, the anti-spam organization Spamhaus was the target of a massive and sustained DoS target that some security vendors claim spiked at nearly 300 Gbps of traffic. The key word here is spiked (and we can't emphasize that enough); the average amount of traffic hitting Spamhaus during the attack ranged anywhere from 85-120 Gbps, which still represents a sizable bombardment. The method behind generating an attack this large is DNS reflection.

So how does it work? Typically, an attacker sends a bunch of DNS queries to open DNS resolvers. The attacker forges the source address on his requests to make it look as though they originated from his desired target. The open resolvers then send their typically larger responses to the targeted address, which is quickly swamped with seemingly legitimate traffic. Hence, "reflection." Much like the low and slow attacks described above, DNS reflection doesn't require significant computing resources on the part of the attacker to produce devastating results.

DOS'ING THE MATH

If there's one thing we've learned from the attack on Spamhaus and others like it, it's the importance of understanding the numbers behind DoS attacks. Let's look at an example. Say there was a 200 Gbps attack at 25 Mpps, 200 Gbps = 2.14×10^{11} or so bps, divide that by 25,000,000 and that is about 8,500 bits per packet or just over 1k bytes per packet on average. This indicates many of the packets are pushing towards the maximum packet size most ISPs will route. We've seen attacks with a higher packet rate, but never anything close to that in bandwidth. Both attackers and defenders tend to sensationalize attacks like this. Both have motives for inflating them. Attackers want to call attention to their attacks and defenders will say, "Look, it was so large there was no way we could keep that site up and running." Or if it's a vendor, "Look how powerful our service is. We can stop all the attacks!"

That being said, data compiled by our DoS defense team shows an increase in the average size of attacks over the past three years — as shown in figure 65. In 2011, the average attack involved 4.7 Gbps of bandwidth with a 411 Kpps packet rate. Move forward to 2012 and those averages jumped to 6.7 Gbps at 2.5 Mpps. It shouldn't surprise you to learn that in 2013 the average DoS attack clocked in at 10.1 Gbps at close to 8.1 Mpps. While the QCF and its powerful arsenal likely shoulder some of the blame for this year-over-year increase, the increasing popularity of reflection attacks and the power they generate are the primary culprits.

Aside from the rise in DNS reflection attacks and converting webservers into high-powered DoS bots, not much has changed over the past few years. Sure, there are always new DoS toolkits making their way into the underground and new waves of attacks taking place, but the general principles remain the same, as do the targets. We saw many attacks directed at the financial, retail, professional services, and public sectors. What better way to inflict pain on a bank or retailer than to go after its website – something critical to customer service? And though carrying out a DoS attack isn't as difficult as it seems, results may vary. For the financially challenged attacker it's possible to download open source tools such as Low Orbit Ion Cannon (LOIC), but he'll need A LOT of friends to do the same if the attack has a chance of being successful. On the other hand, if he's got some cash to burn, the attacker can rent out a DirtJumper or Athena botnet and pummel the target of his choice for less than $10 an hour. The more enterprising (and development-minded) individual might even go so far as to write his or her own DoS script and herd a botnet together. And trust us, these three scenarios play out every day in the cybercriminal underground.

We've heard many clients and colleagues express concern about attackers using DoS attacks as a "smokescreen" to hide fraudulent automated clearing house (ACH) transfers and other illicit activity. Although there are scattered reports of this happening, hard evidence we've managed to collect doesn't indicate the rate or impact justifies the level of angst. We sometimes jokingly refer to this as the "DoS Bigfoot," not because we don't think it's real, but because we're intrigued and want to capture it on film. Data collection for the 2015 DBIR is already underway, and we invite any with shaky night vision film clips of this thing to set the record straight.

RECOMMENDED CONTROLS

Now that we've talked about the problem at length it's a good time to discuss what can be done to lessen or even prevent DoS attacks against your organization.

Let's start with the basics

Servers/services should always be turned off when not in use, patched when in use, and available only to the people who need them, especially in the case of reflection attacks.

Isolate key assets

Segregate key IP/servers from non-essential IP space. Any IP space not in active use for key servers should be announced out of a separate circuit, perhaps even purchase a small backup circuit and announce IP space. That way if it's attacked, the attack won't compromise your primary facilities/servers.

Get comfortable

Don't be shy about using your provider's anti-DDoS service. You should be able to test it quarterly without charge. Make sure that your key operations teams will react in a timely manner if there is an actual attack. Even if your provider offers "auto-mitigation," this shouldn't be an install-and-forget kind of service.

Have a plan in place

What are you going to do? Who will you call if your primary anti-DDoS doesn't work? You know what you'd do if one of your circuits or servers went down – why should this be any different?

Do the math

Know that most attacks are about the FUD numbers cited by the news media. They're above your SSL server capacity, or perhaps a few times your ingress circuit line rate. But attackers don't have infinite resources either – the biggest attack will be just over what you can manage.

Ask about capacity

Understand that all ISPs will have to, at some point, protect their general network over your company's specific traffic. Ask your anti-DDoS provider about its upstream peering capacity – if they can't get the (good and bad) traffic in no matter how much mitigation capacity they have, your good traffic will be dropped at the outside edge of their ISP's network and your call queues will light up with unhappy customers.

EVERYTHING ELSE

AT A GLANCE

Description

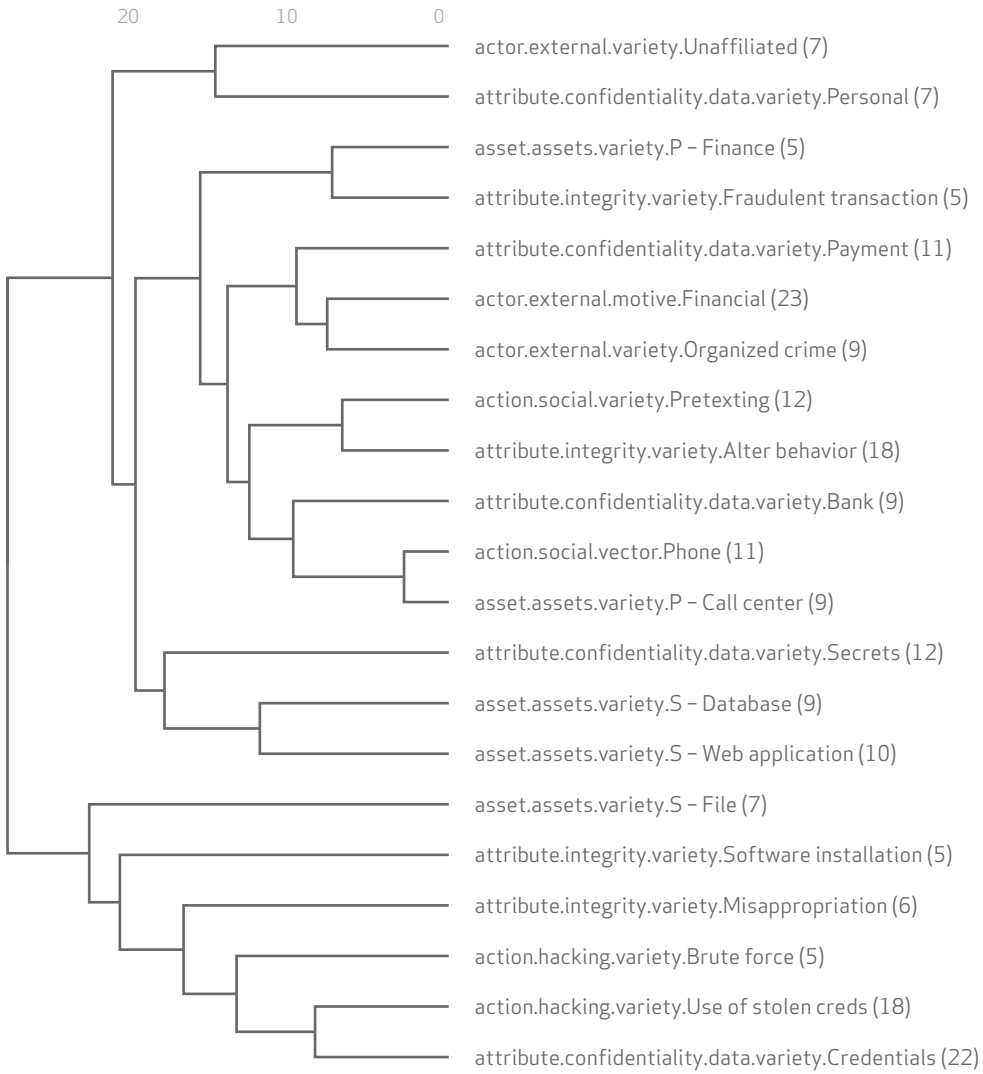
This last "pattern" isn't really a pattern at all. Instead, it covers all incidents that don't fit within the orderly confines of the other patterns. Given that, one might assume you'd never find a more wretched hive of scum and villainy than this random assortment of outcasts. But what we actually find looks nothing like the riffraff of a Mos Eisley cantina; it's almost entirely dominated by two related species of incidents.

"Why not make two more patterns, then?" you might well ask. Great question; allow us to explain.

First, let's describe what we see here among these 7,269 incidents. Actors are 99.9% external. Generic "hacking" (variety unknown), phishing, and browser-busting malware lead known threat actions, with everything else below the 1% line. Three-quarters of all incidents involved compromised web servers; the rest were unknown.

"Then why do you say they're related species?" you might counter. A little digging into the data uncovers the fact that these incidents actually represent mass attacks reported to a CSIRT. In one, thousands of servers in hosting facilities were compromised and used to host phishing sites. The other involved hundreds of servers hijacked to host malware for drive-by exploits. Nothing else was reported about the method of compromise or the phishing/malware campaigns themselves.

Figure 66.
Hierarchical clustering of VERIS enumerations for confirmed data disclosures falling outside the main incident patterns



All in all, not very informative — and therein lies the issue. It quickly becomes apparent that these incidents aren't something utterly different, but rather simply lack sufficient detail to better classify them.

A slightly more interesting view is found by narrowing in on the subset (81 incidents) that involved confirmed data disclosure.[28] We'll use the Figure 66 dendrogram to go hunting for patterns. We realize that dendrograms aren't the most inherently intuitive visualizations,[29] but their basic premise is pretty straightforward and they serve this purpose well.

The words in the dendrogram are VERIS enumerations. Enumerations are organized into clusters. Clusters are connected — directly or indirectly — by branches of varying levels. Multiple enumerations on the same cluster or low-level branch signify a close and distinguishing relationship (i.e., they commonly appear together within incidents). Enumerations and clusters separated by higher branches signify weak or infrequent relationships.

When applied to Figure 66, this results in a cluster to the far right encompassing stolen credentials and the use of those credentials to gain unauthorized access. Higher level (weaker/less frequent) clusters in that same rightmost branch hint that this pairing is sometimes seen in conjunction with brute-force attacks, unauthorized software (malware) installation, and misappropriation (illegitimate use/hijacking) of file servers. That's not exclusive, mind you, but it's a pattern recognized by the algorithm.

To the left of that, we see a strongly related cluster linking phone-based social engineering of call center employees. Including nearby clusters within that whole middle segment adds additional context around that: financially motivated, organized criminals using pretexting to steal payment information from bank call centers, and conducting fraudulent transactions.

The theft of trade secrets kind of sticks out like the extra digit on a six-fingered man (yes; the one who killed your father and must prepare to die); probably because the source knew what was taken, but not how it was taken or who took it.

The leftmost clusters appear to be generic intrusions into web servers and databases to steal personal information.

We could go deeper into this rabbit hole, but this at least gives some sense of what didn't make the cut for the other patterns. You're probably dendrogrammed out by now anyway.

YOU AND ME GO PHISHING IN THE DARK

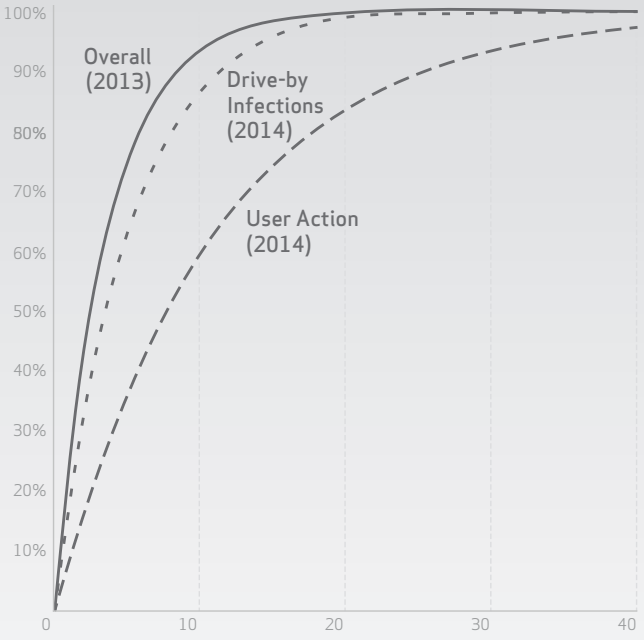
In last year's report we examined data from ThreatSim and came to the earth-shattering conclusion that phishing is an effective way to gain access to an organization. Okay, maybe that isn't news, but the revelation that a phishing campaign of only ten messages has a better-than 90% chance of getting a click was surprising to many of us.

This year we took a look at ThreatSim's data again and immediately reconfirmed the findings from last year; that even a campaign consisting of a small number of email messages has a high probability of success. However, this year we found that the overall success rate of a phishing message was slightly lower at 18%. The reason could be more awareness of phishing, or just natural variation in the samples.

We also looked at the success rate of different tactics in phishing. Are users more likely to visit a link than run an attachment? Are they more likely to click an attachment than enter their passwords in a web form? In general, it appears that about 8% of users will click an attachment and about 8% will fill in a web form. And while most users are skeptical about clicking an attachment (though not skeptical enough) they are less fearful of visiting a link in an email. 18% of users will visit a link in a phishing email. Users unfamiliar with drive-by malware might think that simply visiting a link won't result in a compromise.

Figure 67.
Success rates of phishing exercises

Figure 68.
Phishing success rates



CONCLUSION AND SUMMARY RECOMMENDATIONS

It's fascinating to study what goes wrong. But the real purpose of this research is to help you reduce the risk that these bad things will happen to you. At the end of the day, we do this work to support evidence-based risk management. We think the perspective of studying clustered incident patterns enables more tailored strategies to reduce risk, and toward that end, we did two specific things this year. First, we mapped industries to attack patterns to help answer the question, "For my industry, which threats am I most likely to face?" Second, for each pattern we made specific recommendations, including priority controls from the Critical Security Controls (CSCs) based on our collaboration with the Council on Cybersecurity. This included mapping patterns to controls.

To wrap up, let's connect the dots in one more way. Since we've used the data to map industries to incident patterns and patterns to controls, we figured we also have a decent foundation to map industries directly to recommendations for controls. Figure 69 on the next page shows which controls we think are key to the threat patterns each industry faces. And it weights each control by how often we see the patterns the control addresses in that industry. Essentially, we just did a whole bunch of multiplication to save you the trouble.

So for example, in the column for the Public sector, you'll see CSC 17 stands out as a priority. Sure, someone could have said before, "Intuitively, data loss prevention should probably be important for the Public sector." But now this report puts some hard data behind that. Misuse, theft/loss, and error constitute a strong majority of the attack patterns the public sector faces, and data loss prevention helps address all of them. The other 19 CSCs are great (and we certainly aren't saying anyone should ignore them), but this perspective is a strong argument for making sure the industry gives CSC 17 the focus and resources it deserves; specifically the sub-controls that cover full disk encryption (17.3) and detecting mis-published information (17.6). View this as evidence that can help answer the question "Based on where my industry is now (which reflects the controls already in place and the threats they often face), what should we focus on next?"

Granted, one element of this is subjective in that we and the Council's experts decided which controls would best address each threat pattern. But we based those decisions on event chains observed in our data set. And the weighting we did is based firmly on the frequency data we have for patterns and industries. So while small differences in these numbers probably aren't too meaningful, we do think there's a strong argument for taking a hard look at the controls that come out on top.

As always, we hope you find this year's report valuable, and we look forward to hearing your feedback. May the Force be with you, and have fun storming the castle!

THE COUNCIL ON CYBERSECURITY

The Council on CyberSecurity was established in 2013 as an independent, expert, not-for-profit organization with a global scope committed to the security of an open Internet. The Council is committed to the ongoing development, support, and adoption of the Critical Security Controls; to elevating the competencies of the cybersecurity workforce; and to the development of policies that lead to measurable improvements in our ability to operate safely, securely and reliably in cyberspace. For additional information, visit the Council's website.
www.counciloncybersecurity.org

Questions? Comments? Brilliant ideas?

We want to hear them. Drop us a line at dbir@verizon.com, find us on LinkedIn, or tweet @VZdbir with the hashtag #dbir.

Figure 69.
Critical security controls mapped to incident patterns. Based on recommendations given in this report.

Critical Security Controls (SANS Institute)		POS Intrusions	Web App Attacks	Insider Misuse	Physical Theft/Loss	Misc Errors	Crimeware	Card Skimmers	Cyber-espionage	DoS Attacks
Software Inventory	2.4						●		●	
Standard Configs	3.1						●			
	3.2		●				●		●	
	3.8						●			
Malware Defenses	5.1	●					●		●	
	5.2	●					●		●	
	5.6						●		●	
Secure Development	6.4		●							
	6.7		●							
	6.11		●							
Backups	8.1				●					
Skilled Staff	9.3				●					
	9.4								●	
Restricted Access	11.2	●								
	11.5	●								
	11.6	●								
Limited Admin	12.1	●		●						
	12.2			●						
	12.3	●								
	12.4	●								
	12.5	●								
Boundary defense	13.1						●		●	
	13.7	●	●				●		●	
	13.10	●								
	13.14	●								
Audit Logging	14.5	●		●						
Identity Management	16.1			●						
	16.12			●						
	16.13			●						
Data Loss Prevention	17.1				●					
	17.6			●		●				
	17.9			●		●				
Incident Response	18.1									●
	18.2									●
	18.3									●
Network Segmentation	19.4								●	●

To find out more about the SANS Institute's Critical Security Controls, visit: http://www.sans.org/critical-security-controls/

Figure 70.
Prioritization of critical security controls by industry. Based on frequency of incident patterns within each industry and recommendations for each pattern given in this report. The shading is relative to each industry.

Critical Security Controls (SANS Institute)		Accommodation [72]	Administrative [56]	Construction [23]	Education [61]	Entertainment [71]	Finance [52]	Healthcare [62]	Information [51]	Management [55]	Manufacturing [31,32,33]	Mining [21]	Other [81]	Professional [54]	Public [92]	Real Estate [53]	Retail [44,45]	Trade [42]	Transportation [48,49]	Utilities [22]
Software Inventory	2.4																			
Standard Configs	3.1																			
	3.2																			
	3.8																			
Malware Defenses	5.1																			
	5.2																			
	5.6																			
Secure Development	6.4																			
	6.7																			
	6.11																			
Backups	8.1																			
Skilled Staff	9.3																			
	9.4																			
Restricted Access	11.2																			
	11.5																			
	11.6																			
Limited Admin	12.1																			
	12.2																			
	12.3																			
	12.4																			
	12.5																			
Boundary defense	13.1																			
	13.7																			
	13.10																			
	13.14																			
Audit Logging	14.5																			
Identity Management	16.1																			
	16.12																			
	16.13																			
Data Loss Prevention	17.1																			
	17.6																			
	17.9																			
Incident Response	18.1																			
	18.2																			
	18.3																			
Network Segmentation	19.4																			

For more information on the NAICS codes [shown above] visit: https://www.census.gov/cgi-bin/sssd/naics/naicsrch?chart=2012

To find out more about the SANS Institute's Critical Security Controls, visit: http://www.sans.org/critical-security-controls/

APPENDIX A: METHODOLOGY

Based on feedback, one of the things readers value most about this report is the level of rigor and integrity we employ when collecting, analyzing, and presenting data. Knowing our readership cares about such things and consumes this information with a keen eye helps keep us honest. Detailing our methods is an important part of that honesty.

Our overall methodology remains intact and largely unchanged from previous years. With 50 organizations contributing data this year, there is no single means used to collect and record the data. Instead we employed different methods to gather and aggregate the data produced by a range of approaches by our contributors.

Once collected, all incidents included in this report were individually reviewed and converted (if necessary) into the VERIS framework to create a common, anonymous aggregate dataset. But the collection method and conversion techniques differed between contributors. In general, three basic methods (expounded below) were used to accomplish this:

1) direct recording by Verizon using VERIS

2) direct recording by contributors using VERIS

3) re-codeding using VERIS from a contributor's existing schema

All contributors received instruction to omit any information that might identify organizations or individuals involved, since such details are not necessary to create the DBIR.

Sharing and publishing incident information isn't easy, and we applaud the willingness and work of all these contributors to make this report possible. We sincerely appreciate it.

1. VERIZON'S DATA COLLECTION METHODOLOGY

The underlying methodology we used is unchanged from previous years. All results are based on first-hand evidence collected during paid external forensic investigations and related intelligence operations we conducted from 2004 through 2013. The 2013 caseload is the primary analytical focus of the report, but the entire range of data is referenced throughout. Once an investigation is completed, our analysts use case evidence, reports, and interviews to create a VERIS record of the incident(s). The record is then reviewed and validated by other members of the team to ensure reliable and consistent data.

2. METHODOLOGY FOR CONTRIBUTORS USING VERIS

Contributors using this method provided incident data to our team in VERIS format. For instance, agents of the U.S. Secret Service (USSS) used an internal VERIS-based application to record pertinent case details. Several other organizations recorded incidents directly into an application we created specifically for this purpose.[30] For a few contributors, we captured the necessary data points via interviews and requested follow-up information as necessary. Whatever the exact process of recording data, these contributors used investigative notes, reports provided by the victim or other forensic firms, and their own experience gained in handling the incident.

3. METHODOLOGY FOR CONTRIBUTORS NOT USING VERIS

Some contributors already collect and store incident data using their own framework. A good example of this is the CERT Insider Threat Database[31] compiled by the CERT Insider Threat Center at the Carnegie Mellon University Software Engineering Institute. For this and other similar data sources, we created a translation between the original schema and VERIS[32] and then re-coded incidents into valid VERIS records for import into the aggregate dataset. We worked with contributors to resolve any ambiguities or other challenges to data quality during this translation and validation process.

SECURITY INCIDENTS VERSUS DATA DISCLOSURE

The DBIR has traditionally focused exclusively on security events resulting in confirmed data disclosure[33] rather than the broader spectrum of all security incidents.[34] In the 2013 DBIR, we deviated from that tradition slightly by collecting and referencing a large number of confirmed security incidents. The 2014 DBIR breaks form altogether to gain a broader view. We chose to include these incidents to capture events such as denial of service attacks, compromises of systems without data loss, and a very large bucket of incidents where data loss was just simply unknown. While we think this change is for the better (and we hope you do too), it does mean our report on data breaches will include more than data breaches.

A WORD ABOUT SAMPLE BIAS

For years, science and statisticians debated the relationship between smoking and lung cancer. Through the 1940s and 1950s cases of epidermoid carcinoma of the lung were on the rise and medical experts sought to understand why. Individual studies starting in the 1950s would establish a correlation between smoking and lung cancer, but each of them had statistical flaws in their methodologies. These flaws were not errors or mistakes in any way; the flaws were present because the real world presented imperfect data and the researchers did the best they could to compensate for the imperfect data. R. A. Fisher (a well-respected and famous statistician, who was often shown smoking his pipe) was an outspoken opponent of those studies and would put considerable effort into dissecting and refuting the techniques and conclusions found therein. His personal beliefs were being expressed through his expertise in statistics to such a point that he even accused researchers of manipulating their data.

Finally, in 1959, Jerome Cornfield and several other researchers took a step back to conduct a meta-analysis,[35] which is analysis done by looking at the combination of several other studies (an approach Nate Silver would apply to the 2012 U.S. presidential elections with great success). They showed how the aggregate results of all the other studies provided overwhelming evidence that linked smoking with lung cancer. Even though each study was flawed in some way, they were flawed in different ways and the aggregate had a consistency that was enough to dispel any uncertainty. It would take years for this to permeate into the culture, but Cornfield's meta-analysis was the tipping point in acknowledging the health hazards of smoking.

While we believe many of the findings presented in this report to be appropriate for generalization, bias and methodological flaws undoubtedly exist. However, with 50 contributing organizations this year, we're aggregating across the different collection methods, priorities and goals of our partners. We hope this aggregation will help minimize the influence of any individual shortcomings in each of the samples and the whole of this research will be greater than the sum of its parts.

APPENDIX B:

DATA BREACHES AND IDENTITY THEFT, A CONVOLUTED ISSUE

BY THE IDENTITY THEFT RESOURCE CENTER (ITRC)

We focus primarily on the corporate side of data breaches in this report, but there is clearly an overlap into the consumer world. Because that world is very much front-and-center for ITRC, we thought it would be appropriate to have them contribute a perspective on an important aspect of breaches: consumer identity theft.

So you received a data breach notification letter. Does this mean you are now a victim of identity theft? Not necessarily (yet).

The relationship between data breaches and identity theft is trickier than you think. It's more convoluted than just these two issues being related or even correlated. There are studies touting the relationship between receiving a data breach notification and identity theft victimization, but the ITRC believes this oversimplifies the issue.

TYPES OF INFORMATION

Data breaches are becoming more commonplace and understood by the general public, due in part to publicity surrounding the many high profile incidents that occurred over the past year. As a result, consumers are faced with the fact that their personal identifying information (PII) is being left unsecured by those entrusted to protect it. Passwords, usernames, emails, credit/debit card and financial account information, and Social Security Numbers are being compromised at a staggering rate, endangering the identities of consumers nationwide.

The perceived significance of PII falls along a continuum of importance and value – as well as risk. This issue of "perception" applies to all the players in a data breach scenario – the consumer, the business entity, and of course, the "data thief." I hesitate to call the criminal an "identity thief" for we know not yet their underlying motive for stealing the PII.

LESS-SENSITIVE PII – IMPORTANCE AND VALUE

Over time, consumers have been made increasingly aware of breaches exposing passwords, usernames, and emails. These pieces of less-sensitive PII might, on the surface, appear to have no importance and/or value, and therefore represent relatively low risk of harm. Consumers use these pieces of information day-in and day-out, most likely without thinking about their value, or the measures in place to keep them secure and private. For businesses, exposure of this data does not typically even trigger the need for breach notification.

Is there risk of identity of theft? This depends on the ingenuity and level of motivation of the data thief. While it is true that thieves can use this non-PII to "socially engineer" other information about you, it does require some degree of effort.

SENSITIVE PII – IMPORTANCE AND VALUE

Most consumers readily identify credit/debit cards, and financial account information as important. When it comes to these pieces of financial information, they understand the need to protect it — there are associated risks with it being compromised. One major concern is who is responsible for any financial loss or expenses incurred as a result of this occurrence. Many consumers fear exposure of this information may result in identity theft; not realizing additional information (see Social Security number below) is necessary to take it to that level. Typically, the use of financial information is limited to various forms of financial fraud or existing account fraud.

The value of financial account information may be high, but it is usually short-lived if the consumer takes action and closes accounts quickly. This is facilitated by businesses making prompt breach notifications and alerting the consumer that they need to take proactive measures.

Social Security numbers — the gold standard of all sensitive PII — are that extra piece of information that unlocks so many doors. With this data in hand, thieves can now gain access to new credit and other benefits in the victim's name — lines of credit, government benefits, tax refunds, employment, utilities, mortgages, and even medical resources. The data thief is the one who truly appreciates the value of this information.

U.S. businesses recognize the importance of protecting these pieces of sensitive information because they know exposure of this data will trigger breach notification laws around the country in 46 states. There is a definite value to the business to implement best practices and protocols to ensure the security of this information, otherwise they will face the subsequent costs of mitigating a breach.

VICTIM IMPACT — PERSONAL COST (NON-FINANCIAL COSTS)

And while not all consumers who receive a data breach notification will become victims of identity theft, many will face the need to contact credit reporting agencies, creditors, financial institutions, health care providers, and possibly law enforcement agencies, to report that their PII has been compromised. Placing Fraud Alerts or Credit Freezes, closing credit cards and financial accounts, changing passwords and PINs, and closing or changing email accounts are just some of the possible steps one might need to take to minimize future risk of identity theft. That said, pity those who do not receive a breach notification letter for they do not yet know their information has been compromised.

Placing Fraud Alerts or Credit Freezes, closing credit cards and financial accounts, changing passwords and PINs, and closing or changing email accounts are just some of the possible steps one might need to take to minimize future risk of identity theft.

Many of these steps take a personal toll on consumers who oftentimes have no idea what steps to take – even when they are spelled out in a breach notification letter or on a company website. All they know is they are feeling angry, frustrated and confused. They are frequently in an emotional and anxious state of mind. They are trying to grasp the meaning of who is responsible for any financial losses. How much time is it going to take to make necessary calls? Who do they call? What's the number? Why is the line always busy? Can you make the call for me? Will a request for a credit report effect my credit score? How could the business let something like this happen?

UNDERSTANDING TYPES OF PII AND NEED FOR FOLLOW UP ACTION

Depending on the type of personal data compromised (sensitive or less-sensitive) in any given data breach incident, many of the associated risks to the consumer will be contingent upon on how quickly they respond to a breach notification. This is why the timeliness of the notification to the consumers is of significant importance. Laws aside, forewarned is forearmed. An aware customer/client/employee/student is one who can take proactive measures to minimize the risk of any potential harm.

Many of the associated risks to the consumer will be contingent upon on how quickly they respond to a breach notification.

With that said, it is extremely important for consumers to understand the steps they can take to keep their information as private as possible.

RECOMMENDATIONS FOR CONSUMERS

- First and foremost, never carry your Social Security card with you.
- Develop strong passwords – Don't be like the millions of others who use "12345678" or "password." Even when hashed, these passwords can easily be deciphered by data thieves.[36]
- Don't be too social on social media. Providing too much information on these very public venues provides data thieves plenty of information to go phishing at your expense. Even if you're not the hottest celebrity in town, you might be more popular than the next guy on the thief's list.
- Shred sensitive documents – what you don't shred, lock up in a secure location.
- Monitor financial statements and be on the look for any fraudulent transactions.
- Make every possible effort to guard your Protected Health Information (PHI). Minimize the number of times you provide it the doctor's office. Ask questions such as who can access it, will it be encrypted, and do they take measures to store it securely. Be your own PHI watchdog.

APPENDIX C:
LIST OF CONTRIBUTORS

CSIRTS
- CERT Insider Threat Center
- CERT Polska/NASK
- CERT-EU European Union
- CERT.PT
- Computer Emergency Response team of Ukraine (CERT-UA)
- Computer Incident Response Center Luxembourg (CIRCL), National CERT, Luxembourg
- CyberSecurity Malaysia, an agency under the Ministry of Science, Technology and Innovation (MOSTI)
- Industrial Control Systems Cyber Emergency Response Team (ICS-CERT)
- Irish Reporting and Information Security Service (IRISS-CERT)
- OpenCERT Canada
- US Computer Emergency Readiness Team (US-CERT)

CYBER CENTERS
- Centre for Cyber Security, Denmark
- Council on CyberSecurity
- Defense Security Service (DSS)
- European Cyber Crime Center (EC3)
- National Cybersecurity and Integration Center (NCCIC)
- Netherlands National Cyber Security Centre (NCSC-NL)

FORENSIC PROVIDERS
- Deloitte and Touche LLP
- G-C Partners, LLC
- Guidance Software
- S21sec
- Verizon RISK Team

INFOSEC PRODUCT AND SERVICE PROVIDERS
- Akamai
- Centripetal Networks, Inc.
- FireEye
- Kaspersky Lab
- Malicious Streams
- McAfee, part of Intel Security
- ThreatGRID, Inc.
- ThreatSim
- Verizon DoS Defense
- WhiteHat Security

ISACS
- Center for Internet Security (MS-ISAC)
- Electricity Sector Information Sharing and Analysis Center (ES-ISAC)
- Financial Services ISAC (FS-ISAC)
- Public Transit ISAC (PT-ISAC)
- Real Estate ISAC (RE-ISAC)
- Research & Education ISAC (REN-ISAC)

LAW ENFORCEMENT AGENCIES
- Australian Federal Police (AFP)
- Cybercrime Central Unit of the Guardia Civil (Spain)
- Danish National Police, NITES (National IT Investigation Section)
- Dutch Police: National High Tech Crime Unit (NHTCU)
- Policía Metropolitana (Argentina)
- Policía Nacional de Colombia
- US Secret Service

OTHER
- Anonymous contributor
- Commonwealth of Massachusetts
- Identity Theft Resource Center
- Mishcon de Reya
- VERIS Community Database (VCDB)
- Winston & Strawn

ENDNOTES

1. Yes, we'll continue to call it the Data Breach Investigations Report, even though it analyzes incidents that aren't exclusively breaches or derived through forensic investigations. Hey! Maybe we should just call it the "Data Report" because those two words are still dead-on.
2. Stay tuned, though. We may dig deeper into this topic once we're free from the pressures of publishing the main report.
3. To be fair, the biggest sprees in this year's dataset originated in both Romania and Germany
4. For the initiated, we could not reject the null-hypothesis with a p-value of 0.21 and R^2 of 0.134.
5. https://www.whitehatsec.com/resource/stats.html
6. Silowash, G., Cappelli, D., Moore, A., Trzeciak, R., Shimeall, T., & Flynn, L. (2012). Common Sense Guide to Mitigating Insider Threats. In S.E. Institute (Ed.), (4th ed., pp. 17): Carnegie Mellon University. http://www.sei.cmu.edu/library/abstracts/reports/12tr012.cfm
7. http://www.mishcon.com/assets/managed/docs/downloads/doc_2714/Mishcon_Recover_Report.pdf
8. Silowash, G., Cappelli, D., Moore, A., Trzeciak, R., Shimeall, T., & Flynn, L. (2012). Common Sense Guide to Mitigating Insider Threats. In S.E. Institute (Ed.), (4th ed., pp. 17): Carnegie Mellon University. http://www.sei.cmu.edu/library/abstracts/reports/12tr012.cfm
9. We supplemented this pattern with data from VCDB because it is a rich source of related incidents.
10. http://veriscommunity.net/doku.php?id=enumerations#assetvariety
11. Note to self: stop leaving laptop in conference room when walking down to the cafeteria.
12. We supplemented this pattern with data from VCDB because it is a rich source of related incidents.
13. http://blog.oxforddictionaries.com/2013/11/word-of-the-year-2013-winner/
14. A major NHTCU investigation into groups using mobile malware showed that in less than a year's time, five variations of mobile malware for one specific bank could be detected. Modest estimates suggest that criminals gained around €50,000 per week using this specific form of mobile malware, harvesting over 4,000 user credentials from 8,500 infected bank customers in just a few months. Mobile malware does not move the needle in our stats as we focus on organizational security incidents as opposed to consumer device compromises.
15. The two email vectors are almost certainly underrepresented based on the number of phishing actions in this data set. Not enough info was provided to discern whether those phishing incidents utilized email attachments or embedded links, and so they were marked "unknown." Therefore, the actual number of both email vectors is surely higher than shown here, but still not be enough to overtake the web vectors.
16. You can get more data on this by sending us a small fee to cover transfer costs. Just give us your bank account number and we will email you the information. Bitcoins welcome.
17. We supplemented this pattern with data from VCDB because it is a good source of related incidents.
18. Espionage is not one of the more common patterns in the Public sector, but if you look solely at data breaches, it becomes quite prominent.
19. In all honesty, some of us aren't crazy about the "cyber" thing. Be that as it may, it is increasingly THE collectively used and understood modifier for the type of attacks we discuss here. By "cyber," we're referring to computer networks, systems, and devices. We promise not to go all cyber-cyber-cyber!!! on you; we'll just use it as a way to reference this pattern.

20. http://en.wikipedia.org/wiki/Analysis_of_competing_hypotheses
21. https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/psychology-of-intelligence-analysis/
22. https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/sherman-kent-and-the-board-of-national-estimates-collected-essays/6words.html
23. See the Misuse section for analysis of insider espionage.
24. http://blogs.rsa.com/wp-content/uploads/VOHO_WP_FINAL_READY-FOR-Publication-09242012_AC.pdf
25. See Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2010). Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains.
26. http://www.microsoft.com/security/sir/default.aspx
27. ...at least until February 2014, when NTP reflection became bigger-est in history – but too late for us to edit this report beyond adding a footnote.
28. It's funny to consider that this remnant of the rejects is about equal to the total number of breaches in the 2009 DBIR.
29. For a primer on dendrograms and hierarchical clustering techniques (our method for identifying patterns in this report), see http://en.wikipedia.org/wiki/Hierarchical_clustering
30. See an example here: https://incident.veriscommunity.net/s3/example
31. http://www.cert.org/blogs/insider_threat/2011/08/the_cert_insider_threat_database.html
32. For instance, CERT has an attribute named "Motives and expectations" that maps very well to actor.internal.motive in VERIS.
33. VERIS defines data disclosure as any event resulting in confirmed compromise (unauthorized viewing or accessing) of any non-public information. Potential disclosures and other "data-at-risk" events do NOT meet this criterion, and have thus not traditionally been part of the sample set for this report.
34. VERIS defines an incident as any event that compromises a security attribute (confidentiality, integrity, availability) of an information asset.
35. Cornfield, Jerome, et al. "Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions." Journal of the National Cancer Institute 22.1 (1959): 173-203.
36. Don't believe us? Just Google 286755fad04869ca523320acce0dc6a4

ABOUT THE COVER

The "universe" of colored dots on the cover represents 4,596 incidents from the DBIR dataset, including all confirmed data breaches over the past three years and a sample of 400 Denial of Service attacks from last year. We calculated the distance between dots using a multi-dimensional scaling technique (with the Manhattan distance algorithm) with 65 VERIS fields for each incident. This required over 6 million comparisons, and the resulting distances were projected on a two-dimensional plane. The closer the dots, the more similar the incidents, meaning they share many VERIS characteristics like threat actors, actions, assets, etc. The colors represent the nine incident classification patterns discussed throughout this report (see the Table of Contents for a section detailing how these patterns were derived). Patterns in close proximity (e.g., Misuse and Error) share many VERIS characteristics, while those that are far apart (e.g., Espionage and POS Intrusions) have little in common. The tightness or looseness of dots within each pattern shows the amount of variation among incidents in that pattern. The sub-pattern clusters (overlayed points and lines) were created using 10 years of incident data (over 100,000 incidents). We generated a force-directed network graph from the frequency of VERIS fields and the relationships between them for each individual cluster.

verizonenterprise.com