The University of Texas at Austin
Center for Identity

Identity Threat Assessment and Prediction

Jim Zeiss
Razieh Nokhbeh Zaeem
K. Suzanne Barber

UTCID Report #18-06

MAY 2018

Sponsored By

HID

gemalto
security to be free

GENERALI
GLOBAL ASSISTANCE
IDENTITY PROTECTION SERVICES

IDEMIA
augmented identity

LifeLock
Guarantee Your Good Name

TransUnion®

Identity Threat Assessment and Prediction

Identity theft and related threats are increasingly common occurrences in today's world. Developing tools to help understand and counter these threats is vitally important. This paper discusses some noteworthy results obtained by our Identity Threat Assessment and Prediction (ITAP) project. We use news stories to gather raw data about incidents of identity theft, fraud, abuse, and exposure. Through these news stories, we seek to determine the methods and resources actually used to carry out these crimes; the vulnerabilities that were exploited; as well as the consequences of these incidents for the individual victims, for the organizations affected, and for the perpetrators themselves. The ITAP Model is a large and continually growing, structured repository of such information. There are currently more than 5,000 incidents captured in the model. To this body of information we apply a variety of analytical tools, collectively known as the ITAP Dashboard, that enable us to show and compare threats, losses, and trends in the identity landscape. From this analysis, we discovered notable and sometimes surprising results. A goal of this project is to be able to predict future threats, and to provide some concrete guidance for consumers, businesses, and government agencies on how to avoid them or lessen their impact.

• Security and privacy→Human and societal aspects of security and privacy.

Additional Key Words and Phrases: Identity Theft, Identity Fraud, Personally Identifiable Information (PII)

1. INTRODUCTION

Identity theft, fraud, abuse, and exposure are huge and rapidly growing problems. They reportedly affected an estimated 15.4 million persons in 2016 [Pascual et al. 2017] in the U.S. alone and have been at or near the top of the Federal Trade Commission's national ranking of consumer complaints for the 17th consecutive year. Investigating identity theft, abuse, and exposure events in detail, and thoroughly understanding them, is a first yet significantly important step toward countering these problems.

Literally every day, multiple incidents of identity theft, fraud, abuse, and exposure (hereafter *identity theft* for the sake of brevity) are reported in the news media. The Identity Threat Assessment and Prediction (ITAP) project is our endeavor focused on gathering identity theft information from such news stories, structuring this information, analyzing it, and discovering trends and characteristics. While other researchers have utilized various sources of identity theft information – such as agency data, surveys, and anecdotal reports – our approach is to extract identity theft information directly from news stories [Yang et al. 2016]. (ITAP, however, is not intrinsically limited to news stories. One could, and in the future we might, use other sources of identity theft data to populate the ITAP Model.)

The ITAP project gathers media news stories on identity theft via two distinct methods. First, we set up an RSS feed to monitor a number of websites that report on cases of identity theft. Second, we created a Google Alert, which provides a daily notification of any new website indexed by Google that reports on the subject of identity theft. The news story webpages collected through these two methods are manually winnowed down to a list of those that are reports of identity theft, fraud, abuse, and exposure incidents.

We then add the information collected from the news stories to the ITAP Model we built with the AWAREness Suite[1] application, a web-based decision support system we use to model and quantify data. The ITAP Model is a comprehensive structured collection of over fifty details about each identity theft incident. It includes the type of the incident, how and when the incident happened, the methods and resources used by the perpetrators, the vulnerabilities exploited, the types of personal information compromised, the demographics of the victims, the consequences for the victims and perpetrators, and many more features of the incident.

We apply various analytical tools to the ITAP Model to reveal useful overarching statistics and trends regarding identity theft. The ITAP Dashboard is a set of tailor-made charts and tables we

[1] http://www.awaresoftwareinc.com/products.html

have developed to explore a variety of particularly interesting aspects of the identity theft incidents, such as the resources most frequently used by perpetrators and the geographic distribution of identity theft over the U.S.

The analysis of the ITAP data leads to several interesting discoveries. Some of the noteworthy findings of this project are as follows.

1. The top five types of personally identifiable information (PII) compromised in the U.S. are name, social security number, date of birth, address, and credit card information.
2. About 17% of the incidents were non-malicious: PII was compromised, but there was no malicious intent on the part of those responsible.
3. Only 0.36% of identity theft incidents were ones that spanned the whole U.S., such as the infamous Target breach in 2013 and Equifax breach in 2017.
4. The middle class suffers most frequently from identity related crimes.
5. Emotional distress is experienced by victims of identity theft more often than other types of loss, such as financial and property loss.
6. One third of the incidents were performed solely by insiders, such as employees of companies and family members of individuals.
7. The top five most affected sectors are consumers/citizens, healthcare, government, education, and financial services.
8. Of the different age groups of victims of identity theft, fraud, abuse, and exposure, senior citizens are among the most vulnerable.

In summary, the ITAP project makes the following contributions: It gathers, models, and analyses a large number (currently about 5,400) of identity theft news stories. Although not necessarily limited to this data source, ITAP is the first project to use identity theft news stories in this manner. By modeling and analyzing the identity theft information, ITAP uncovers various interesting features and trends in the world of identity theft, fraud, abuse, and exposure.

The remainder of this paper is organized as follows. In Section 2, we discuss how the compromise and misuse of PII affects consumers. Section 3 describes the various types of laws that have been enacted to combat PII compromise and identity theft, and assesses their effectiveness. Section 4 cites work on identity theft that is related to that being done in the ITAP project, explains how ITAP differs from these other approaches, and articulates what we see as the advantages of our approach. Section 5 shows and explains numerous charts and other analytics from the ITAP Dashboard, and notes some of the most interesting or surprising results. Section 6 discusses possible future analytics that might be added to the Dashboard, and tells how these new analytics could benefit consumers. Finally, in Section 7, we sum up the paper with some concluding remarks.

2. PII COMPROMISE AND CONSUMERS

There are a variety of different ways in which a consumer's PII can be compromised. There might be a data breach at an organization that maintains personal information about one. Physical theft of one's mail, wallet or purse, laptop, or other device can result in his or her PII falling into the wrong hands. Social engineering techniques, such as e-mail phishing or telephone scams, are often successful in duping victims into revealing personal data. Old-fashioned snooping, such as overhearing another's telephone conversation, going through their trash bins, or looking over someone's shoulder at an ATM can yield sensitive personal information.

Once compromised, a person's PII can be misused in various ways. To name a few of the most common: An identity thief might put fraudulent charges on existing accounts or open new accounts in the victim's name; the accounts in question might be financial (e.g. credit cards or loans) or utility-based (e.g. electric, telephone, or cable television) [Romanosky et al. 2011]. With a few choice bits of another person's PII, a fraudster can file a tax return (and collect a refund), obtain healthcare services, or get a job in that person's name.

The potential harms to victims of identity theft and other forms of PII compromise are multifold. The general types of such harm include (i) invasion of privacy, (ii) financial loss, (iii) loss of physical or intellectual property, (iv) reputation damage, (v) effort and money spent to recover from the incident and prevent further misuse of the compromised PII, and (vi) emotional distress (over any or all of the above). Cf. [Milne 2003].

3. PII COMPROMISE AND THE LAW

There are at least four different sorts of U.S. state laws in effect that are intended to reduce, or mitigate the consequences of, PII compromise and identity theft: data breach notification laws, data disposal laws, identity theft laws, and consumer report security freeze laws.

As of May 2017, 48 U.S. states had enacted data breach notification laws (the exceptions are Alabama and South Dakota), as have the District of Columbia, Guam, Puerto Rico, and the Virgin Islands.[2] Though they differ in certain details, these laws require organizations that have suffered a data breach to inform the victims of the breach in a timely manner (but notification can be delayed if it might impede a criminal investigation).

Notification laws seem to have had a limited effect on the incidence of identity theft. One study [Romanosky et al. 2011], based on data gathered from the Federal Trade Commission (FTC) and other sources over eight years, concluded that the adoption of notification laws had reduced identity theft caused by data breaches by an average of 6.1 percent – a significant but by no means large amount. (At the time of that study, 45 states had adopted notification laws.) Another limitation is that most of the state notification laws cover electronic data only; just ten states' notification laws include tangible data.

Data disposal laws are also prevalent. At least 31 states have laws in place that require organizations (businesses and usually government agencies) to "destroy, dispose, or otherwise make personal information unreadable or undecipherable."[3] In addition, the FTC's Disposal Rule requires disposal of consumer reports or information derived from such reports.[4]

Despite the data disposal laws and the FTC rule, improper disposal incidents remain common. In March of 2016, Health IT Security reported that two of the top five healthcare data breaches in that year thus far were due to improper data disposal.[5] Data compiled by DataBreaches.net showed that 58.4% of healthcare data breaches occurring in January 2017 were due to insiders, and that four of the nine insider incidents in question were specifically the result of insider error.[6] A McAfee report from September 2015 stated that internal actors were the cause of 43% of data breaches, and that half of those were due to employee accidents [McAfee 2015].

Every U.S. state has a law regarding identity theft or impersonation. 29 states have specific restitution provisions for identity theft. Five states have forfeiture provisions for identity theft crimes. Eleven states have created identity theft passport programs to help victims from continuing identity theft.[7] Nonetheless, the incidence of identity theft and fraud in the U.S. has increased each

[2] http://www.dwt.com/statedatabreachstatutes/; https://www.databreaches.net/state-breach-notification-laws/; http://www.ncsl.org/research/telecommunications-and-information-technology/security-breach-notification-laws.aspx

[3] http://www.ncsl.org/research/telecommunications-and-information-technology/data-disposal-laws.aspx

[4] https://www.ftc.gov/tips-advice/business-center/guidance/disposing-consumer-report-information-rule-tells-how

[5] https://healthitsecurity.com/news/top-5-healthcare-data-breaches-in-2016-not-from-hacking

[6] https://www.protenus.com/blog/31-health-data-breaches-disclosed-in-january-as-hhs-fines-for-late-reporting

[7] http://www.ncsl.org/research/financial-services-and-commerce/identity-theft-state-statutes.aspx#Statutes

year since 2014. There were an estimated 15.4 million victims of these crimes in 2016 [Pascual et al. 2017].[8]

All 50 states and the District of Columbia also have laws that enable consumers to put a security freeze on their credit reports. Such a freeze prevents a consumer reporting agency from releasing a credit report or any information from the report without the consumer's consent. Thus, a security freeze is likely to thwart an identity thief from opening an unauthorized account in the consumer's name, and can help a person whose PII has been compromised monitor whether anyone has attempted to create such an account. Twenty-nine states also enable a parent or legal guardian to place a security freeze on a minor's credit report.[9]

Like the other sorts of laws enacted to counter identity theft, the effect of credit report security freeze laws has been limited. First, according to the Bureau of Justice Statistics, relatively few victims of identity theft (less than 4% in 2014) actually place a freeze on their credit reports. Second, while a security freeze can prevent the unauthorized opening of new accounts, it has no effect on a person's existing accounts. And the majority of identity fraud incidents (86% in 2014) involve the misuse or attempted misuse of existing accounts [Harrell 2015].

In sum, the various state laws in place to combat PII compromise and identity crimes are not as effective as one would wish. This is partly due to the fact that some of these laws are not consistently followed or the resources they provide for are not widely taken advantage of. It is also partly due to the fact that the scope of some of the laws is narrower than it could be.

4. RELATED WORK

Most studies that gather data and cite statistics regarding identity theft and related incidents are based on either (i) interviews of or complaints from individual victims [FTC 2016; Harrell 2015; Pascual et al. 2017; Pascual and Marchini 2016; Stuart et al. 2005;], (ii) reports from affected organizations [Gemalto 2017; Verizon 2017], or (iii) data from law enforcement agencies [Allison et al. 2005]. ITAP is one of the first projects to get its data from published news stories and other publicly available online reports of identity threat incidents [Yang et al. 2016].

ITAP's approach has a distinct advantage over these others with respect to yielding detailed information about the various ways in which identity theft is carried out. Several years ago, various authors pointed out that few details were known about the actual methods used by perpetrators of identity theft and fraud [Hoofnagle 2007; Newman and McNally 2008], and this is still largely true today. However, compared to victim- or organization-provided data, the news stories and the like used by ITAP are better sources for such specifics as the role(s) of the performer(s) who perpetrated the incident, the resources the perpetrator(s) used to help accomplish the act, the steps they undertook, and the types of PII that were compromised or misused. The ITAP Model captures these specifics and the Dashboard leverages this information. While studies based on law enforcement data often do provide more information than victims or organizations can about the perpetrators' techniques, they tend to be limited to one agency or one geographical region. ITAP, in contrast, captures incidents across the U.S. and around the world.

There is at least a small amount of research that is based on interviews with identity thieves and fraudsters [Copes and Vieratis 2009]. To be sure, such interviews can provide even greater detail than news stories about the methods these criminals use. We would urge that more such studies be done. Still, this type of research is likely to be limited by small sample sizes (59 offenders were interviewed by Copes and Vieratis) and skewed by the fact that their subjects are all perpetrators who were caught, incarcerated, and willing to be interviewed.

[8] https://www.javelinstrategy.com/press-release/identity-fraud-hits-record-high-154-million-us-victims-2016-16-percent-according-new

[9] http://www.ncsl.org/research/financial-services-and-commerce/consumer-report-security-freeze-state-statutes.aspx

More similar to ITAP in its approach, the Identity Theft Research Center (IRTC) publishes annual Data Breach Reports [ITRC 2016] whose sources are news media and notifications from state governmental agencies. However, their focus is on U.S. organizations that have suffered data breaches. Unlike ITAP, they do not maintain data concerning fraud, thefts from individual victims directly, cases of abuse, or incidents taking place in other countries. Moreover, the data the ITRC captures on the incidents they do cover is less detailed, and more unstructured, than that captured in ITAP. The ITRC even defines *breach* much more narrowly than ITAP, as an event in which an individual's name plus social security number, driver's license number, medical record, or financial record is compromised.

To sum up, compared to the other approaches discussed in this section, we believe that our method of using a large number news stories from around the globe concerning all sorts of identity threat incidents puts ITAP in a better position to gather and analyze detailed data about what these incidents involve.

5. ANALYZING THE ITAP DATA AND GETTING RESULTS

The analytics provided by the ITAP Dashboard are custom-built for showing interesting facts extracted from the ITAP Model regarding identity theft, identity fraud, and other cases in which PII is compromised. In this section, we show and explain a number of these analytics. We divide the analyses into three categories, according to what aspect of the incidents they primarily pertain: the events themselves, the victims, or the perpetrators.

5.1 Events

Amount of Non-Malicious Activity. This is the percentage of incidents in ITAP that are categorized as non-malicious. As mentioned above, a non-malicious incident is one in which PII is compromised, but without malicious intent on the part of those responsible. They are commonly caused by human error of some sort. Currently, the percentage is just over 17.4.

Digital vs. Non-Digital Theft. This pie chart shows the percentages of PII theft incidents in ITAP that were "digital", "analog", and both-digital-and-analog. A theft is considered purely digital if the resources used by the perpetrator(s) include nothing other than computers (or other digital devices), the Internet (or other computer networks), and information accessible via such networks. A theft is purely analog if it primarily involves physical actions (beyond those required to operate a digital device); e.g. breaking into an office and stealing a briefcase. An example of "both" could be a case in which the perpetrator gets someone to reveal a password over the telephone via social engineering (analog), and then uses the password on a website to access the victim's bank account information (digital). Figure 1 tells us that non-digital identity theft is more common than digital.

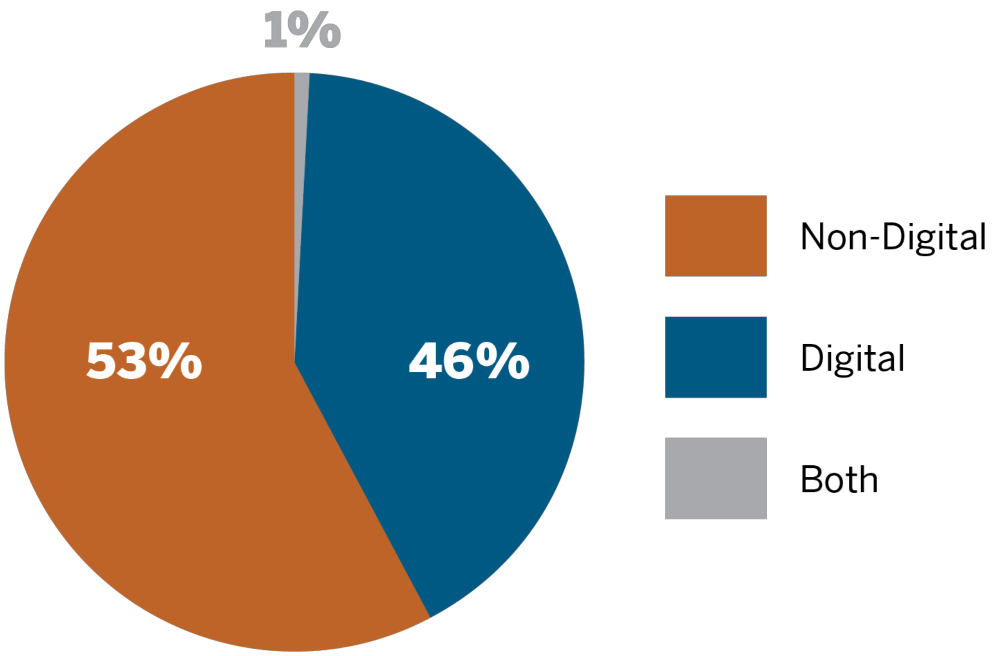


Figure 1. Percentages of digital and non-digital thefts

Market Sector. Here the user selects the number of most commonly affected market sectors s/he wants the chart to display (i.e. *Top 5*, *10*, *15,* or *All*) to see a horizontal bar chart showing the corresponding percentages of incidents associated with that sector. Figure 2 shows that the top ten sectors, in order, are: Consumer/Citizen, Healthcare & Public Health, Government Facilities, Education, Financial Services, Commercial Facilities, Defense Industrial Base, Information Technology, Law Enforcement, and Food & Agricultural.[10] Note that 90% of all incidents fall under one of the top six sectors.

[10] The sectors we consider are the Department of Homeland Security's sixteen Critical Infrastructure Sectors (https://www.dhs.gov/critical-infrastructure-sectors) and three others we find useful: Consumer/Citizen, Education, and Law Enforcement.

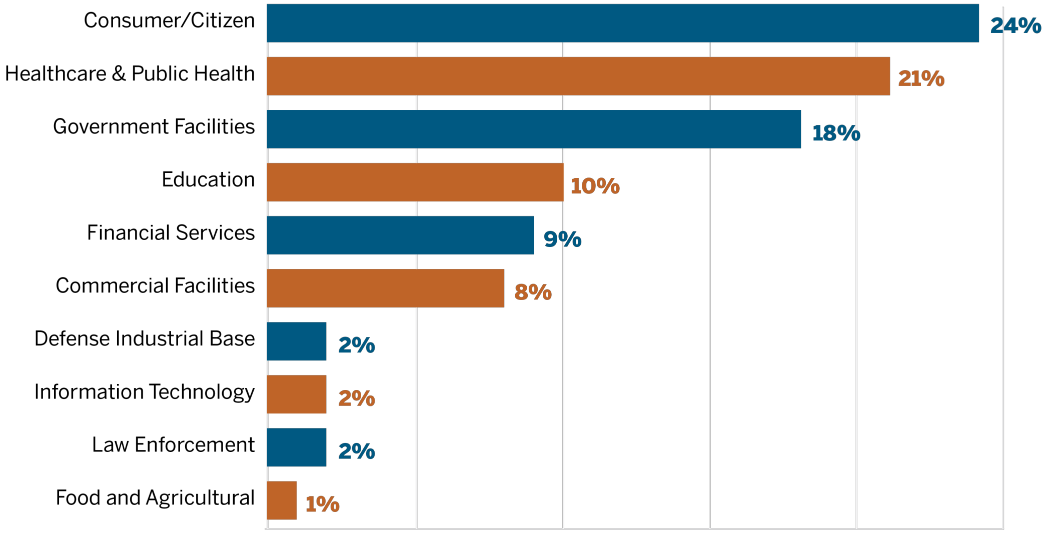


Figure 2. The top ten affected market sectors and their percentages.

National Impact of Identity Theft. This is the percentage of U.S.-based events in which PII was compromised and the incident was local to a particular city (or cities), county, state, or region. This is as opposed to incidents that have nationwide or worldwide effects. The percentage of localized incidents is currently a very high 99.64%. Thus only 0.36% of the incidents spanned the whole U.S., such as the infamous Target breach in 2013 and Equifax breach in 2017.

Incidence of PII Compromise by State. Figure 3 shows a six-color map of the U.S. The closer a state's color is to the darker brown end of the scale, the greater the number of events that have occurred in the state. In the Dashboard, users can mouse-over a given state to see the specific number of incidents from that state.

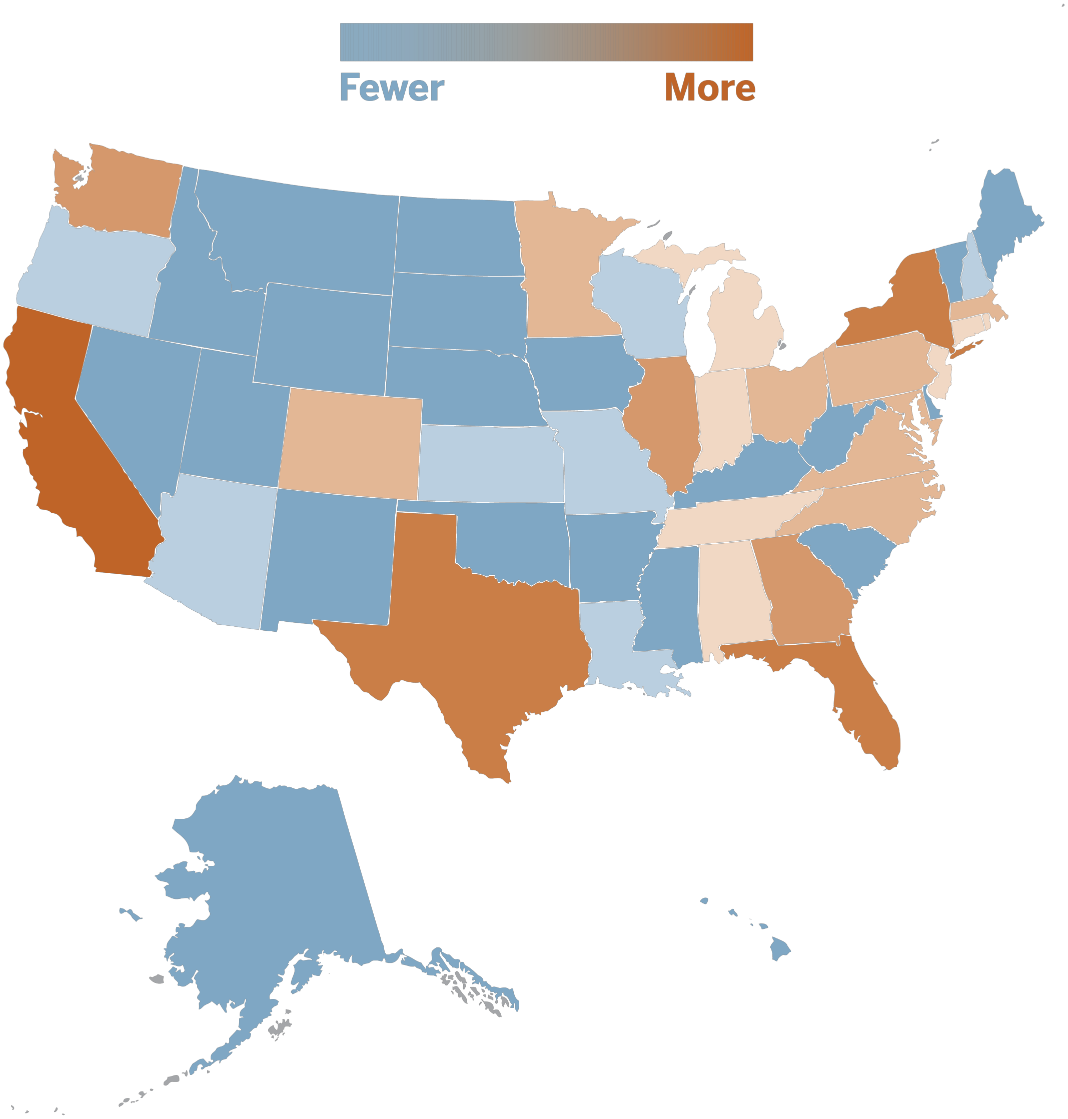


Figure 3. States' colors indicate relative incidence of PII compromise.

5.2 Victims

Age Group of Victims. This bar chart shows the percentages of incidents affecting victims of different age groups. Though adults generally were the most-affected group at 71%, seniors were specifically targeted in 21% of the events. See Figure 4.

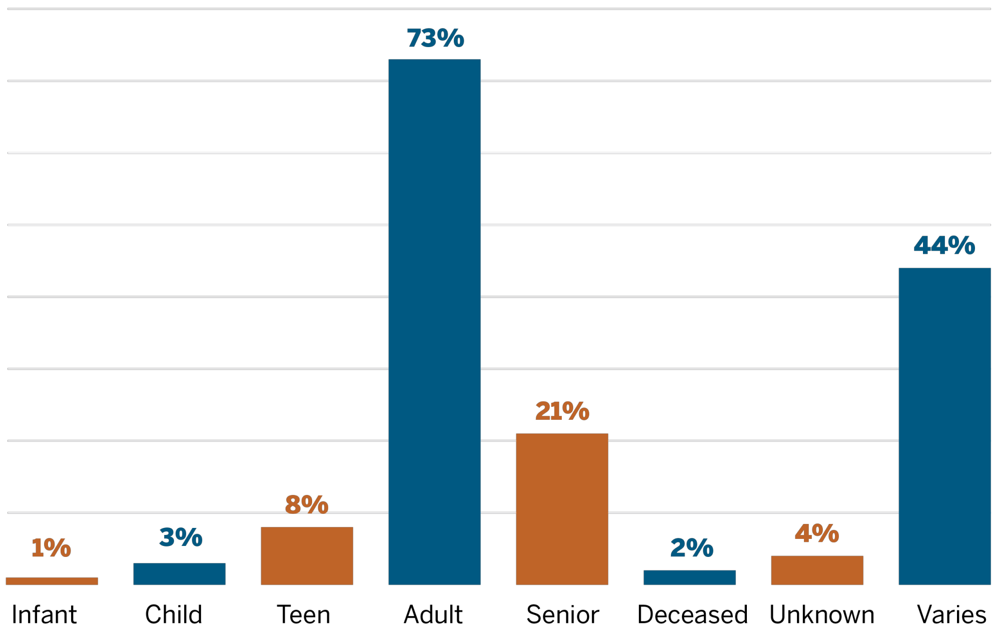


Figure 4. Age groups of victims and their percentages.

Annual Income of Victims. This bar chart shows the percentages of incidents affecting victims in various income ranges. As Figure 5 indicates, the middle class suffers most frequently from identity related threats.

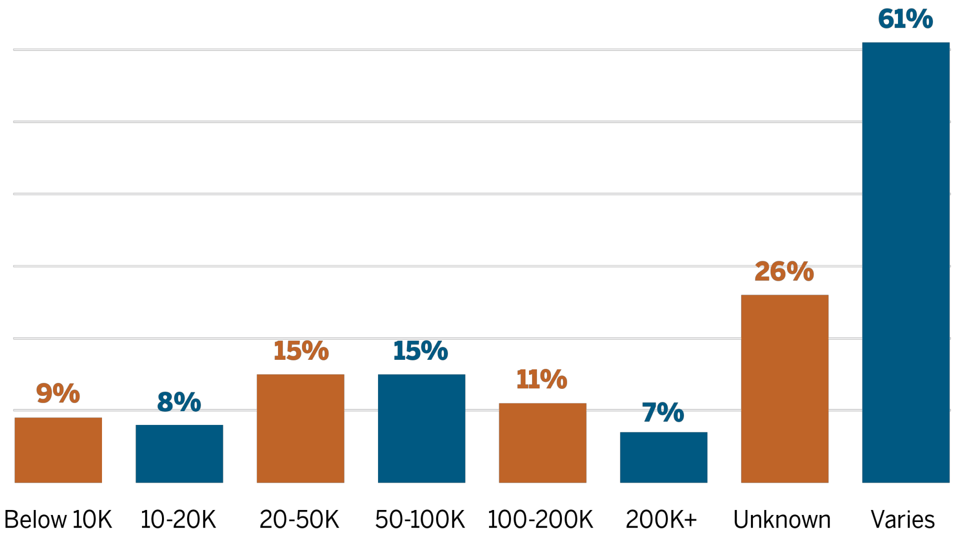


Figure 5. Distribution of the annual incomes of victims.

Education Level. This horizontal bar chart shows the percentages of incidents affecting victims of different levels of education. It turns out that the college-educated are the most often harmed group here. See Figure 6.

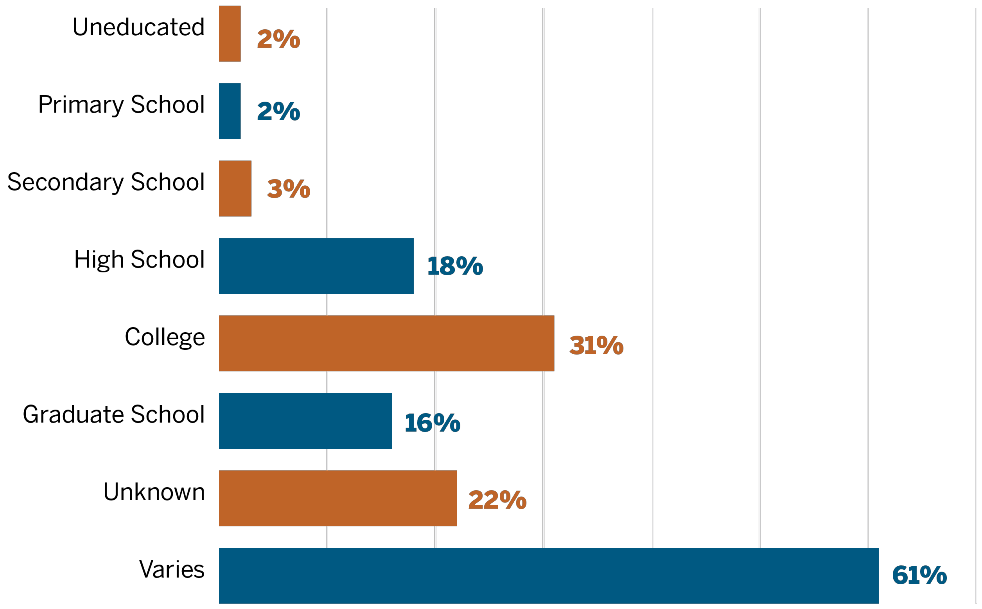


Figure 6. Education levels of victims and their percentages.

Emotional Distress. This vertical bar chart shows the distribution of incidents with respect to the level of emotional distress experienced by the victims. Figure 7 reveals that more than half of the incidents resulted in high levels of distress.

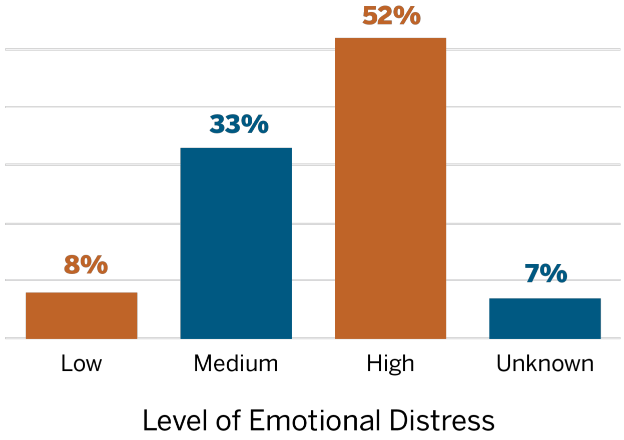


Figure 7. Percentages of victims' emotional distress levels.

Type of Loss. This horizontal bar chart displays the percentages of incidents with respect to the types of loss incurred by the victims. Figure 8 shows, notably, that emotional distress is experienced more often than other types of loss, such as financial and property loss.

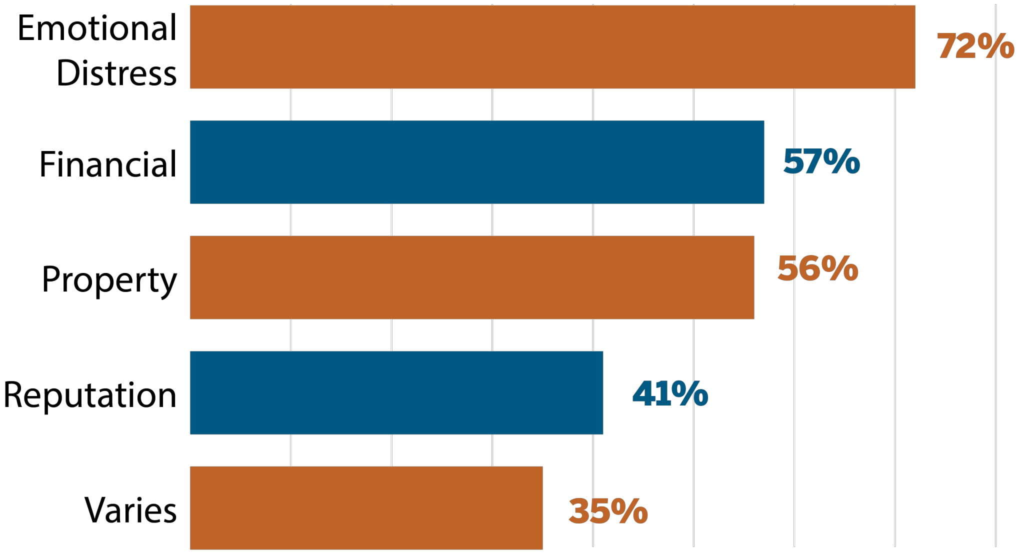


Figure 8. Types of loss incurred and their percentages.

5.3 Perpetrators

Performers. Here the user selects the number of most common types of performer s/he wants to be displayed (i.e. *Top 5*, *10*, *15*, or *All*) to see a horizontal bar chart showing the corresponding percentages of incidents associated with those performers. Figure 9 shows the current top five performers and their respective percentages. In ITAP, briefly, a *thief* is one who steals PII, a *fraudster* is one who misuses PII for personal gain, and a *hacker* is one who creates or exploits a digital or computer-based vulnerability in order to compromise identity assets.

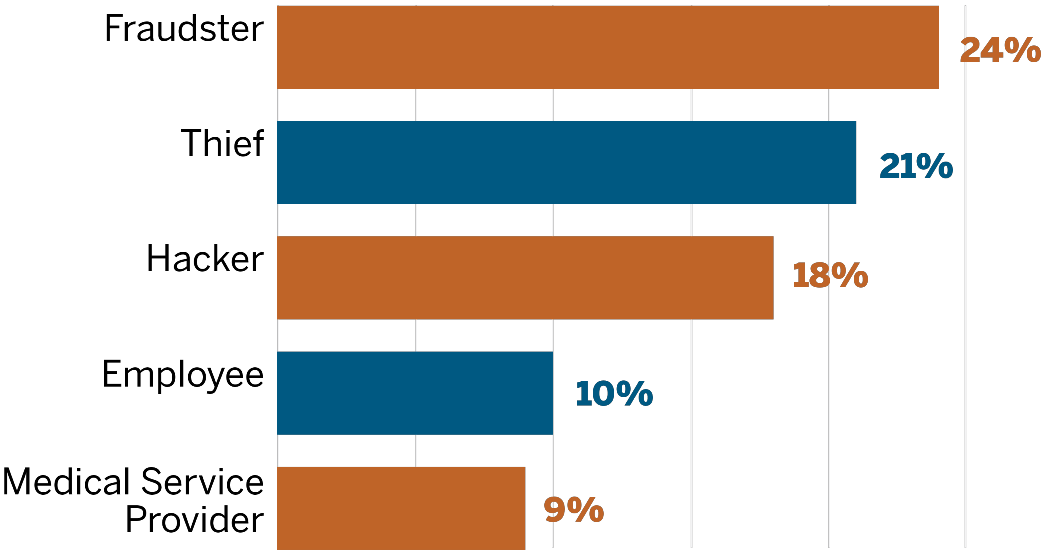


Figure 9. The top five types of perpetrators.

Resources. The user selects the number of most commonly used resources s/he wants to be displayed (i.e. *Top 5*, *10*, *15*, or *All*) to see a pie chart showing the corresponding percentages of incidents associated with those resources. (The percentages are normalized so as to total 100%

regardless of the number of resources shown.) Figure 10 tells us that the top ten resources used are: computer, database, computer network, malware, stolen credit card, and so on.

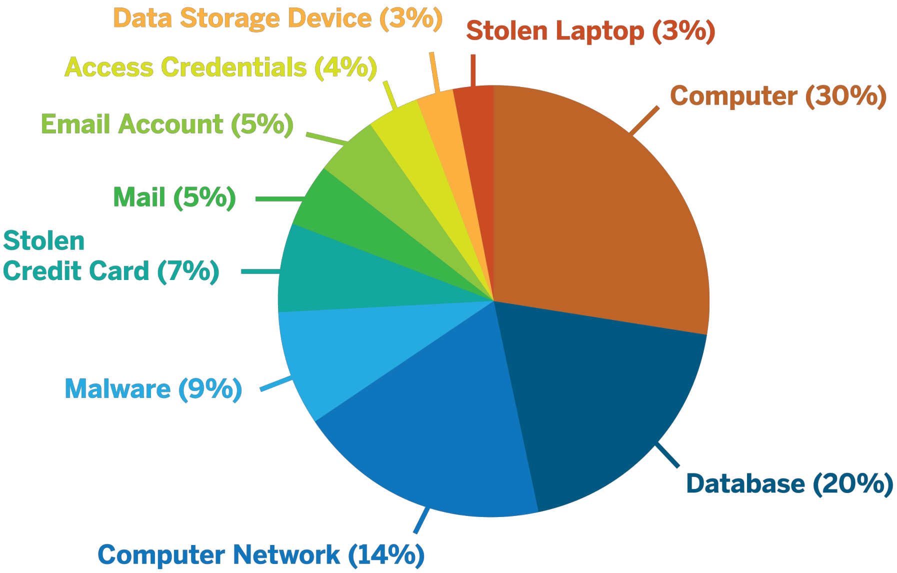


Figure 10. The top ten resources used by perpetrators.

Insider vs. Outsider Activities. This pie chart shows the respective percentages of incidents in which the perpetrator(s) were insiders, outsiders, and both insiders and outsiders. Insiders include employees of companies and family members of individuals. Figure 11 indicates that just over one-third of the events were performed solely by insiders.

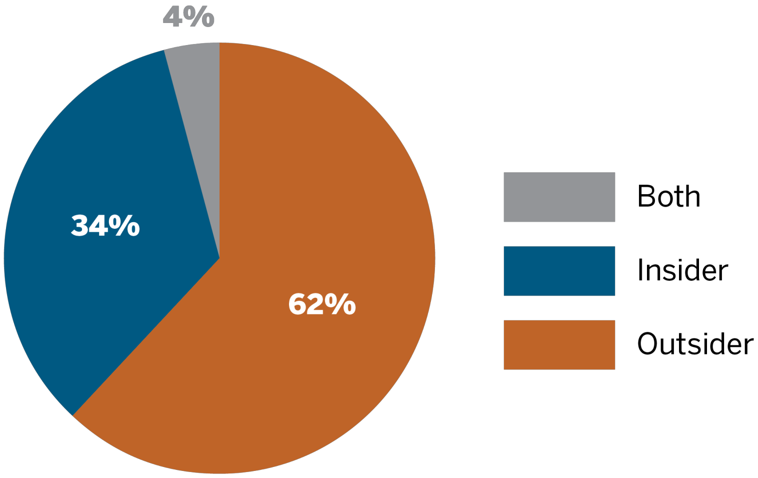


Figure 11. Percentages of PII compromise events committed by insiders vs. outsiders.

PII Compromised. The user selects the number of most commonly used types of PII s/he wants to be displayed (i.e. *Top 5*, *10*, *15*, or *All*) to see a pie chart showing the corresponding percentages of incidents that have those PII types associated with them. (Here again, the percentages are normalized so as to total 100% regardless of the number of PII types shown.) Figure 12 shows that the top five compromised PII types are: name, social security number, date of birth, address, and credit card information.

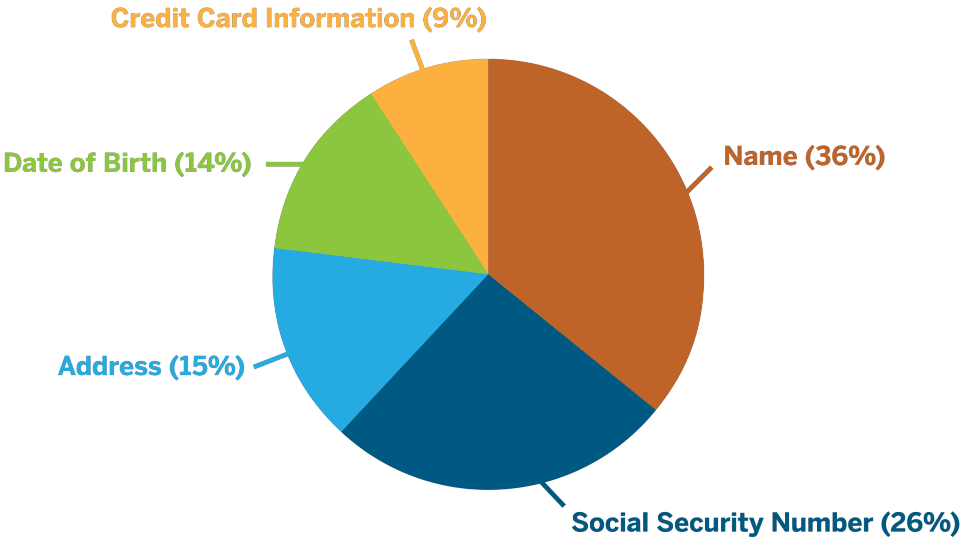
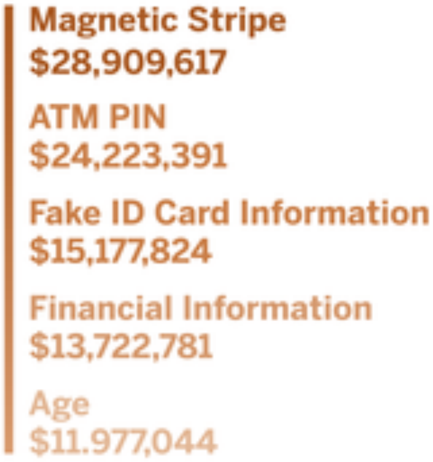


Figure 12. The top five types of PII compromised.

Financial Loss per Attribute. The user selects an item from a list of PII types and other personal attributes to see a dollar amount representing the average financial loss associated with the selected attribute. The amount is calculated as the mean amount of money lost in incidents in which that attribute was compromised. Figure 13 shows the loss amounts for the top five frequently used personal attributes.



Magnetic Stripe	$28,909,617
ATM PIN	$24,223,391
Fake ID Card Information	$15,177,824
Financial Information	$13,722,781
Age	$11,977,044

Figure 13. Average losses for cases in which the top five frequently used attributes were compromised.

6. FUTURE WORK

We envision improving the ITAP Model and extracting new results to show in the Dashboard as two major avenues of future work for this project. In this section we look at several new types of questions about identity theft that we might address and provide answers to.

By expanding upon the pieces of salient information about identity threat events that we currently capture in the ITAP Model – information that is commonly included in the news stories we use – we could answer such questions as the following.

- How frequently have various sorts of reparations been made to the victims of identity theft? (Possible reparations might include *account credited*, *settlement paid*, and *ID security service subscription offered*.)
- How frequently have various sorts of punitive measures been meted out to the perpetrators of identity theft? (Possible measures might include *arrest*, *indictment*, *conviction*, *prison sentence*, *probation*, and *fine*.)
- How do different types of organizations compare with respect to their susceptibility to PII-related data breaches? (Organization types might include *financial institution*, *retail store chain*, *government agency*, and *healthcare provider*.)
- How do the different sectors compare with respect to the types of actors who first discovered PII-compromising incidents occurring in those sectors? (Such actor-types might include *organization affected*, *individual affected*, *law enforcement agency*, and *independent investigator*.)

We might also ask and answer questions, with respect to different ways of carving up the set of identity threat events, about how the portions compare in terms of having the greatest (i.e. most severe) consequences. Here is a smattering of examples:

- For each of the four general PII types – *something you know* (e.g. a password), *something you have* (e.g. an ID card), *something you are* (e.g. fingerprints), and *something you do* (e.g. travel patterns) – what percentage of all ITAP incidents involved the compromise of PII of that type?
- For each of the three general performer-type categories – *insider-only*, *outsider-only*, and *both-insider-and-outsider* – what is the average financial loss per incident falling under that category?
- For each of the three general method-type categories – *digital*, *non-digital*, and *both-digital-and-non-digital* – what percentage of incidents in that category saw the victims experience a high level of emotional stress?
- Which sectors have the highest percentages of incidents in which two or more general types of loss were incurred? (The general types of loss we consider are *financial*, *physical property*, *intellectual property*, *emotional distress*, and *reputation damage*.)

We might also answer some more specifically time-relative questions in order to show recent trends and make predictions. Here are just a few representative examples of many such questions:

- For a given sector and year, what were the respective percentages of the incidents that involved the compromise of PII of the general type (i) something you know, (ii) something you have, (iii) something you are, and (iv) something you do?
- Cumulative over all sectors, for a given year, what were the respective percentages of the incidents that affected these numbers of victims: (i) less than 100, (ii) 100 to 1K, (iii) 1K to 10K, (iv) 10K to 100K victims, (v) 100K to 1M, (vi) 1M to 10M, and (vii) more than 10M?
- For a given year, for each sector, what was the average monetary loss amount in cases of identity fraud?
- For a given sector, what were the respective percentages of incidents during the past four years that involved physical property loss?
- Cumulative over all sectors, what were the respective percentages of incidents during the past four years that were perpetrated by insiders (outsiders, both)?
- For each of the top five sectors, what were the respective percentages of incidents during the past four years in which the perpetrators were arrested (charged, convicted, sentenced, fined)?

Having answers to questions such as the above would benefit consumers by giving them a fuller understanding of the identity threat landscape and how to navigate it. These answers might also affect their perceptions of privacy. For example, knowing about the susceptibility of retail stores to data breaches might make one less likely to use debit cards – which require the user to enter a PIN – at such establishments. Awareness of the relative severity of the negative consequences of biometric data theft might increase one's reluctance to provide one's fingerprints to a device or organization. Knowing that individuals in one's own income group have increasingly fallen victim to identity crime over the past several years might spur one to take greater precautions to monitor and protect one's privacy.

7. CONCLUSION

The main products of the ITAP (Identity Threat Assessment and Prediction) project include the continually growing ITAP Model, which consists of structured information gleaned from (currently 5,000) news stories reporting incidents involving the exposure, theft, or fraudulent use of PII (personally identifiable information). The steps taken by the perpetrators, the resources they used, the types of PII that were compromised, and other salient attributes of the incidents, the victims, and the perpetrators are captured in the model. Another main product is the ITAP Dashboard, where various charts, lists, and statistics derived from the contents of the model are displayed. The ITAP Dashboard reveals novel and sometimes surprising results. For example, one third of the incidents were performed solely by insiders, and senior citizens are particularly vulnerable to identity threats.

This paper reviewed the effects identity theft has on consumers, described the various types of laws enacted to protect consumers and organizations against such crimes, and pointed out some limitations of these laws. The paper then briefly described how the identity theft news stories are collected in ITAP, how their identity threat-related content is put into the Model, and the advantages we believe our approach has over other approaches to the study of identity theft. It explained in some detail how the modeled information is analyzed in the Dashboard and the results that were obtained. It also suggested future work for enriching our current modeling techniques and expanding our current analytics, by listing some new types of questions that ITAP could be used to answer and the benefits these answers could have for consumers.

REFERENCES

Allison, Stuart F.H., Amie M. Schuck, and Kim Michelle Lersch. 2005. Exploring the crime of identity theft: Prevalence, clearance rates, and victim/offender characteristics. *Journal of Criminal Justice*, 33 (1): 19-29.

Copes, Heith and Lynne M. Vieraitis. 2009. Understanding identity theft: offenders' accounts of their lives and crimes. *Criminal Justice Review*, 34 (3): 329-349.

Federal Trade Commission (FTC). 2017. Consumer Sentinel Network Data Book for January – December 2016. Retrieved May 1, 2017 from https://www.ftc.gov/system/files/documents/reports/consumer-sentinel-network-data-book-january-december-2016/csn_cy-2016_data_book.pdf

Gemalto. 2017. 2016: Mining for Database Gold: Findings from the 2016 Breach Level Index. Retrieved May 1, 2017 from http://breachlevelindex.com/assets/BLI-ebook-2016/Breach-Level-Index-Report-2016-Gemalto.html

Golden, Ryan and Suzanne Barber. 2014. NewsFerret: Supporting identity risk identification and analysis through news story text mining. *International Journal of Computer and Information Technology*, 3 (5): 850-859.

Harrell, Erika. 2015. Victims of Identity Theft, 2014. Bureau of Justice Statistics, U.S. Department of Justice, Office of Justice Programs. NCJ 248991. Retrieved August 29, 2016 from http://www.bjs.gov/content/pub/pdf/vit14.pdf

Hoofnagle, Chris J. 2007. Identity theft: making the known unknowns known. *Harvard Journal of Law and Technology*, 21 (1): 97-122.

Identity Theft Research Center (ITRC). 2016. ITRC Data Breach Report 2016. Retrieved May 1, 2017 from http://www.idtheftcenter.org/images/breach/ITRCBreachReport_2016.pdf

McAfee. 2015. Grand Theft Data: Data exfiltration study: Actors, tactics, and detection. Retrieved September 5, 2017 from https://www.mcafee.com/us/resources/reports/rp-data-exfiltration.pdf

Milne, George R.. 2003. How Well Do Consumers Protect Themselves from Identity Theft? *Journal of Consumer Affairs*, 37 (2): 388–402.

Milne, George R., George Pettinico, Fatima M. Hajjat, and Ereni Markos. 2017. Information Sensitivity Typology: Mapping the Degree and Type of Risk Consumers Perceive in Personal Data Sharing. *Journal of Consumer Affairs*, 51 (1): 133–161.

Milne, George R., Andrew J. Rohm, and Shalini Bahl. 2004. Consumers' Protection of Online Privacy and Identity. *Journal of Consumer Affairs*, 38 (2): 217-232.

Morris, Robert G. and Dennis R. Longmire. 2008. Media constructions of identity theft. *Journal of Criminal Justice & Popular Culture*, 15 (1):76-93.

Newman, Graeme R. and Megan M. McNally. 2005. Identity theft literature review. United States Department of Justice: National Institute of Justice. Retrieved August 29, 2016 from https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=210459

Pascual, Al and Kyle Marchini. 2016. 2016 Data Breach Fraud Impact Report. Retrieved August 29, 2016 from https://www.javelinstrategy.com/coverage-area/2016-data-breach-fraud-impact-report

Pascual, Al, Kyle Marchini, and Sarah Miller. 2017. 2017 Identity Fraud: Securing the Connected Life. Overview retrieved September 5, 2017 from https://www.javelinstrategy.com/coverage-area/2017-identity-fraud

Romanosky, Sasha, Rahul Telang, and Alessandro Acquisti. 2011. Do Data Breach Disclosure Laws Reduce Identity Theft? *Journal of Policy Analysis and Management*, 30 (2): 256-286.

Verizon. 2017. 2017 Data Breach Investigations Report. Retrieved May 1, 2017 from http://www.verizonenterprise.com/verizon-insights-lab/dbir/2017/

Yang, Yongpeng, Monisha Manoharan, and K. Suzanne Barber. 2015. Modelling and analysis of identity threat behaviors through text mining of identity theft stories. In *2014 IEEE Joint Intelligence and Security Informatics Conference (JISIC 2014)*, 184-191.

WWW.IDENTITY.UTEXAS.EDU