
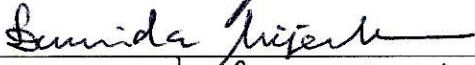
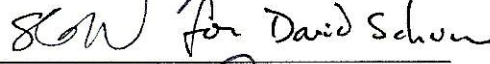
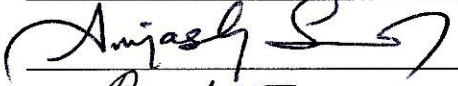
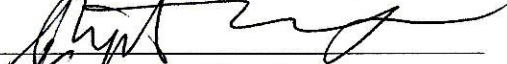
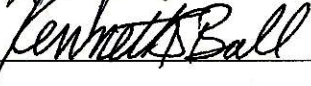


PROFILING, TRACKING, AND MONETIZING:  
AN ANALYSIS OF INTERNET AND ONLINE SOCIAL NETWORK CONCERNS

by

Jason W. Clark  
A Dissertation  
Submitted to the  
Graduate Faculty  
of  
George Mason University  
In Partial fulfillment of  
The Requirements for the Degree  
of  
Doctor of Philosophy  
Information Technology

Committee:

	Dr. Damon McCoy, Dissertation Director
	Dr. Duminda Wijesekera, Dissertation Co-Director
 for David Schum	Dr. David Schum, Committee Member
	Dr. Avinash Srinivasan, Committee Member
	Dr. Stephen G. Nash, Senior Associate Dean
	Dr. Kenneth S. Ball, Dean, Volgenau School of Engineering

Date: 7/2/14 Summer 2014  
George Mason University  
Fairfax, VA

Profiling, Tracking, and Monetizing:  
An Analysis of Internet and Online Social Network Concerns

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Jason W. Clark  
Master of Science  
Rensselaer Polytechnic Institute, 2002  
Bachelor of Science  
Syracuse University, 2001

Director: Dr. Damon McCoy, Professor  
Department of Computer Science

Summer 2014  
George Mason University  
Fairfax, VA

Copyright © 2014 by Jason W. Clark  
All Rights Reserved

## **Dedication**

This is dedicated to my wife, Sarah, daughter, Addison (Addie), and dog, Otto, who sacrificed many days so I could work on my research and write this dissertation. Also, to baby Brady, I can't wait for your arrival.

## **Acknowledgments**

I would like to thank my advisor, Dr. Damon McCoy, for allowing me to work so closely with him on some very interesting research. I would also like to thank my co-dissertation director Dr. Duminda Wijesekera, who mentored me from the very first day I enrolled in the Ph.D program. Also, I would like to thank the rest of my dissertation committee, Dr. David Schum and Dr. Avinash Srinivasan for their guidance and insightful comments. In addition, I would like to thank Dr. Jeremy Allnutt, who supported me until his much deserved retirement. Special thanks to my initial advisor Dr. Angelos Stavrou who urged me to publish. A big thank you goes out to all of the talented researchers I had the opportunity to collaborate with: Chris, Peter, and Neha. I also wanted to thank past work colleagues, especially Jim, Domingo, and Greg who provided much guidance, support, and insight. I thank my new work colleagues at CMU-SEI, especially Randy, Bill, and Michael for their support. I also want to thank Dr. Dwayne Phillips who has proofread every single one of my published papers, he is a tremendous asset and I really appreciate his assistance. Of course, I want to thank my friends and family especially my mother. Lastly, I wanted to thank Trey, Mike, Page, and Jon for keeping the music flowing through many endless nights of writing.

# Table of Contents

	Page
List of Tables . . . . .	viii
List of Figures . . . . .	ix
Abstract . . . . .	xi
1 Introduction . . . . .	1
1.1 Problem statement . . . . .	1
1.2 Thesis statement . . . . .	3
2 Background . . . . .	4
2.1 Profiling and tracking Internet visitors . . . . .	4
2.2 Correlating a persona to a person . . . . .	7
2.3 Determining the effect of multiple attacks on privacy-preserving technology users . . . . .	9
2.4 Limitations of privacy preserving technologies . . . . .	10
2.5 Survey scams . . . . .	12
2.6 Understanding users' perceptions of privacy and value of information stored in their accounts . . . . .	13
3 Literature Review . . . . .	16
3.1 Summary . . . . .	16
3.1.1 Profiling and tracking . . . . .	16
3.1.2 Breaching and protecting a privacy preserving technology . . . . .	17
3.1.3 Online social networks . . . . .	19
3.1.4 Cyber crime . . . . .	23
3.1.5 Privacy and security of password and images on the cloud . . . . .	27
4 Profiling and Tracking (Re-identification) Attack Experiments . . . . .	32
4.1 <b>TR</b> acking and <b>PR</b> ofiling Internet visitors via website analytics - TRAP . . . . .	32
4.1.1 Participant's role . . . . .	33
4.1.2 Lead researcher's role . . . . .	34
4.2 Results . . . . .	36
4.2.1 Cumulative entropy . . . . .	38
4.2.2 Weighted entropy . . . . .	39

4.3	Correlating a Persona to a Person . . . . .	40
4.3.1	Participant recruitment . . . . .	41
4.3.2	Participant’s role . . . . .	41
4.3.3	Lead researcher’s role . . . . .	42
4.4	Results . . . . .	44
4.5	Determining the effect of multiple attacks on privacy preserving technology users . . . . .	47
4.5.1	Network monitoring . . . . .	50
4.5.2	Phishing . . . . .	51
4.5.3	Online social networks . . . . .	51
4.6	Results . . . . .	52
4.7	Attacking a privacy preserving technology system . . . . .	56
4.7.1	Participant’s role . . . . .	58
4.7.2	Technical prerequisites and traffic analysis . . . . .	58
4.8	Results . . . . .	60
5	Monetizing Online Social Network Users via Survey Scams . . . . .	64
5.1	MyPageKeeper spam feed . . . . .	65
5.2	Infiltration . . . . .	66
5.3	Ad network and affiliate ID extraction . . . . .	67
5.4	Prevalence . . . . .	67
5.5	Carbon dating . . . . .	68
5.6	Revenue estimation . . . . .	69
5.7	Content locking and clickjacking . . . . .	70
6	Understanding User Perceptions of Privacy and Value of Information Stored in their Accounts . . . . .	71
6.1	Participant recruitment . . . . .	71
6.2	Surveys . . . . .	73
6.3	Password scan via Cloudsweeper . . . . .	73
6.4	Image scan via Gmail Image Extractor . . . . .	75
6.5	Approach . . . . .	76
6.6	Passwords . . . . .	76
6.7	Images . . . . .	78
6.8	Exit interview . . . . .	81
7	Defense Mechanisms . . . . .	84
7.1	Profiling and tracking related defenses . . . . .	84
7.1.1	Privacy preserving technologies (Anonymization) . . . . .	84
7.1.2	Privacy enhancing browser extensions . . . . .	86

7.2	Security awareness training defenses . . . . .	87
7.2.1	Outbrief . . . . .	87
7.2.2	Phishing awareness . . . . .	88
7.3	Cyber crime and survey scam defenses . . . . .	89
7.4	Protection of passwords and images on the cloud . . . . .	90
8	Conclusions . . . . .	92
8.1	Experimental limitations . . . . .	94
8.2	Future research . . . . .	95
A	Appendix I . . . . .	97
	Bibliography . . . . .	101



## List of Tables

Table	Page
4.1 Assumptions made about the participants . . . . .	34
4.2 Data collected from website #1 during experiment . . . . .	35
4.3 Data collected from website #2 during experiment . . . . .	35
4.4 Data collected from website, fan page, and application during experiment . . . . .	43
4.5 User agent string . . . . .	47
4.6 Demographics of the survey respondents . . . . .	49
4.7 Survey results associated with phishing knowledge and privacy on the Internet . . .	53
4.8 Survey results associated with PPT usage . . . . .	54
5.1 Summary of crawled and manually collected data . . . . .	65
5.2 Summary of the prevalence of the affiliates calculated using two different methods: 1) Initial URL 2) Landing (Land) . . . . .	66
5.3 Extraction of affiliate ID, offer ID, and Wireshark capture of URL . . . . .	67
5.4 Offers taken from CPAlead as seen in June 2013 . . . . .	69
5.5 Offer payouts for June 2013 . . . . .	70
6.1 Demographics of our 30 participants . . . . .	72
6.2 Responses from exit interview . . . . .	82
A.1 Operating System Usage . . . . .	97
A.2 Browser Usage . . . . .	97
A.3 Search Engine Usage . . . . .	97
A.4 Search Resolution Usage . . . . .	98
A.5 ISP Usage . . . . .	99
A.6 City Location Usage . . . . .	100

## List of Figures

Figure	Page
2.1 Showing the flow of money between the different entities . . . . .	13
4.1 Visual representation of the collected persona data, truth files, and link files for website #1 and website #2 . . . . .	33
4.2 Visual representation of the success rate for our guesses. . . . .	37
4.3 Calculated entropy for each network persona feature. . . . .	39
4.4 Depicts the adversary using tracking code and analytics to correlate a persona to a person . . . . .	41
4.5 Accuracy of guesses displayed by ID overlaid by profile name . . . . .	46
4.6 Shows an adversary implementing a variety of attack phases that when used collectively can identify a PPT user . . . . .	48
4.7 Security awareness training displayed as a pie chart . . . . .	55
4.8 Summary of the experiment including key players and entities . . . . .	60
6.1 Location of email access among participants . . . . .	72
6.2 Different email accounts among participants . . . . .	73
6.3 Cloudsweeper interface . . . . .	74
6.4 Cloudsweeper account theft audit . . . . .	74
6.5 Cloudsweeper cleartext password audit . . . . .	75
6.6 Password memorization tools and techniques . . . . .	77
6.7 Frequency of participants who decided to redact their passwords . . . . .	78
6.8 Frequency of participants who decided to encrypt their passwords . . . . .	78
6.9 Breakdown of who took the images not meant for the public . . . . .	79
6.10 Intended audience for the image not meant for the public . . . . .	79
6.11 Destination of image not meant for the public . . . . .	80
6.12 Threat (adversary) responsible for sharing and uploading images . . . . .	80
6.13 Types of methods used by participants to protect their images from being uploaded and shared without their consent . . . . .	81
6.14 Exit interview anecdote quotes . . . . .	83

7.1 Security awareness training displayed as bar graph . . . . .	88
--	----

## **Abstract**

PROFILING, TRACKING, AND MONETIZING:  
AN ANALYSIS OF INTERNET AND ONLINE SOCIAL NETWORK CONCERNS

Jason W. Clark, PhD

George Mason University, 2014

Dissertation Director: Dr. Damon McCoy

This dissertation explores concerns facing Internet, specifically Online Social Network, users. The attacks we discuss can lead to identity theft, biased and tailored website content delivery, geolocation threats, monetization, and an overall lack of privacy. We introduce a profiling and tracking attack that correlates a users online persona that is captured from seemingly innocuous website traffic (e.g., operating system, search engine, browser, time spent on website, etc.) with that of the same users real Facebook profile through analytics captured from a custom Facebook Fan Page. We show how an adversary might identify the personally identifiable information of the user given only their online persona.

The protection of ones identity is paramount especially for users working in the intelligence community. As a result, these organizations are currently employing privacy preserving technologies as part of their standard network defenses to anonymize their outbound traffic. Our results show that while network-level anonymity systems are better at protecting end-user privacy than having no privacy preserving technology, they are unable to thwart de-anonymization attacks aimed at applications and private data of end-users. We demonstrate and substantiate our claims using a targeted experiment using actual scenarios of real-world users who are relying on a privacy preserving technology.

To this end, we execute multiple attacks associated with network monitoring, phishing, and Online Social Networks. We also discuss how a user can be monetized through an attack vector such as spam. Spam is a profit-fueled enterprise, and cybercriminals are focusing more of their efforts at growing Online Social Networks. One of the common methods of monetizing Online Social Network spam is to entice users to click on links promising free gift cards and iPads. However, these links actually lead to ad networks that bombard users with surveys in an attempt to collect personal and contact information that are then sold to other marketers. To date, we lack a solid understanding of this enterprises full structure. We examined the survey scam process to determine the affiliates that are behind this lucrative scam by performing an analysis of five months of Facebook spam data. We provide the first empirical study and analysis of survey scams and demonstrate how to determine which ad networks are sponsoring the spam.

Next, we focus on why people act in an insecure way when specifically handling their passwords and personal images. We believe this is a major problem as seen in sextortion-related cases. By using a combination of well-known human-computer interaction methods such as surveys and exit interviews, combined with custom software, we show that study participants act differently if they visually see the threat associated with their security behavior. We analyze responses from 30 Craigslist participants via a set of three surveys and an exit interview. Furthermore, we analyze the results of Cloudsweeper which is designed to scan Google Mail accounts and report any cleartext passwords, their associated monetary value, and provides the option to allow for such passwords to be encrypted and redacted. Additionally, we introduce for the first time the Google Image Extractor which is designed to extract selected images from the participants Google Mail account and provide the opportunity for users to delete their images seamlessly. Our contributions will help determine if there is a need for applications such as Cloudsweeper and, the Google Image Extractor or if an overhaul of the traditional password management strategy is necessary. All of this research highlights the importance of education on prevalent attack vectors for compromising client systems and violating user privacy. We show the extent to which information made freely available on the Internet, can negatively impact the organization and users. Upon completion of the experiments, we compiled the results and presented it as security awareness briefings.

# Chapter 1: Introduction

## 1.1 Problem statement

The exposure of information on the Internet poses a genuine threat to the privacy and financial well being of everyday users.<sup>1</sup> Indeed, as we show, there are numerous ways in which an adversary can profile, track, and monetize an Online Social Network (OSN) user. The privacy threats we quantify and analyze in this dissertation can be employed for identity and financial theft, biased and tailored website content delivery, geolocation determination, and an overall lack of protection of information that can lead to exposing the real identify of a user. We investigate in depth how different entities including ad networks, affiliates, website owners, and Internet Service Providers (ISP) can collect information from website visitors through the actions of the website, programs that are transmitted through the Internet, and direct user input through forms associated with survey scams. We have collected information from domain name system (DNS) traffic, multiple website sessions, geolocation data, persistent cookies, and performed a rigorous investigation and analysis of OSN spam feeds. The information in aggregate is a concern that can lead to a victim being profiled, tracked, identified, and monetized.<sup>2</sup>

Furthermore, organizations and their employees often have a need to remain anonymous while on the Internet. This need is typically predicated on the fact that undermining of the organization or its employees privacy and anonymity would have a severe negative impact on the mission of such an organization. To this end, organizations often implement and configure privacy preserving technologies (PPT) to help protect their identity and that of their users while traversing the Internet. We define a PPT as a privacy service or technique that attempts to hide the users' IP address, location, and additional information, including but not limited to their research tasks as they access

---

<sup>1</sup>The terms user, visitor, human, and victim are often used interchangeably.

<sup>2</sup>We define the term monetized to mean converting the actions (e.g., pay-per-click, survey scams) into a form of currency.

websites on the Internet. Throughout this dissertation we will often refer to the *Case-Study Privacy Preserving Technology - Organization* (CSPPT or CSPPTO).

The privacy implications of data gathered when users access websites on the Internet, especially under the supposed “protection” of a PPT, need to be examined and addressed more closely. The exposure of information on the Internet poses a genuine threat to the privacy of users.

The growing user bases of OSN sites has become an increasingly lucrative target for profit motivated cyber criminals, such as the “koobface gang” that targeted Facebook and Twitter with large-scale spam campaigns [1]. These spam campaigns lure users to click on enticing posts, such as “free giveaway” offers, for gift cards and iPads. However, once the user clicks on one of these links they are often instructed to complete a survey prior to receiving their free gift card, iPad or being able to view the advertised video clip. These “spamvertised” links on Facebook are being monetized by directing users to specialized ad networks that are known as *Cost Per Action* (CPA) or *lead generation* affiliate based ad networks that pay their affiliates a commission for every “survey” that a user completes.<sup>3</sup> These “surveys” in reality are crafted by clever advertisers and are merely focused at having the user install some profit-generating browser toolbar or rapidly getting the user’s contact information so that they can contact them with follow-up offers and finally presenting the user with “limited time discounted” subscriptions to dating sites and magazines. To date, we lack a solid understanding of this enterprise’s full-structure.

If security and privacy becomes too difficult or time consuming, people will just default back to poor security practices because it is more convenient to do so. At the crux of this study is the question: *Does it make a difference to people if the threat of mishandling their passwords and images is personalized and visually represented to them?*

---

<sup>3</sup>In these types of ad networks an advertiser only pays for the ad when the desired action has occurred. This action can range from the visitor installing a paid browser toolbar, purchasing some product, or providing their contact and personal information.

## 1.2 Thesis statement

This dissertation will address the inherent concerns associated with accessing and providing information on seemingly innocuous websites and through purposely or accidentally providing information via survey scams. We hypothesize that the majority of Internet visitors are at risk of being profiled, tracked, monetized, and led to reveal their true identity with a high level of confidence by an adversary who uses a variety of website analytic tools and “free gift card” surveys. Essentially, users are at risk because of the actions they perform on the Internet including allowing potentially malicious applications, completing forms, and clicking links that are associated with survey scams.

Given this, the general problem we are trying to solve is multi-faceted: First, determine whether our hypothesis is a valid concern by introducing our threat model and conducting experiments with real participants, real data (spam) feeds, or both. Second, website visitors need to be informed on profiling, tracking, and monetizing risks. Third, contrary to popular belief, organizations and website visitors can’t solely rely on a PPT to protect their identity. Fourth, we investigate the survey scam process and attempt to uncover the affiliates (often referred to as sponsors) that are driving this profitable criminal activity. Fifth, much has been made of measures that Internet users can employ to protect their passwords and images. One question that we find lacking in these discussions is, do users believe their passwords and images are worth protecting? More specifically, we are curious to know if users make security versus convenience decisions rationally or are they simply unaware that there is a security risk. We also consider both the savviness and effort of people. Our premise is that generally speaking people like security but the effort extended for security and privacy really matters. Sixth, we wish to better understand whether it makes a difference to people if the threat of mishandling their passwords and images is personalized and “visually” represented to them. The notion of visualizing the threat is an untapped research area which we will explore in this dissertation. Seventh, provide defense mechanisms to help mitigate the profiling, tracking, and monetizing attacks we outline in this dissertation.



## **Chapter 2: Background**

This chapter provides background of the different aspects of the research we have undertaken. It begins with a discussion on the privacy implications of profiling and tracking Internet visitors. This chapter includes an experiment to attempt to break the anonymity of real users at an organization who rely on a privacy preserving technology to protect their privacy while performing research. The next section discusses how an adversary could advance the profiling and tracking attacks to identify a unique person via their OSN profile. This is followed by a discussion on how phishing and network monitoring related attacks could break the anonymity of a user behind a privacy preserving technology. Next, we turn our attention to discussing a method via the use of survey scams to monetize an unsuspecting OSN user. We also discuss our research regarding our investigation of visualizing the threat as it relates to users' storing their passwords and images in the clear on the cloud (e.g., Google Mail). A number of previous studies look at the relationship between users and their (often poor) security practices.

### **2.1 Profiling and tracking Internet visitors**

The privacy implications of data gathered when users access websites on the Internet needs to be examined and addressed more closely. As defined by Clarke, "Privacy is the interest that individuals have in sustaining a personal space, free from interference by other people and organizations" [2]. In today's digital world, it is anyone's guess as to how pervasive these security and privacy concerns will become for Internet users. These privacy concerns include identity and financial theft, biased and tailored website content delivery, and compromised true geographic location.

Although the legitimate aspects of tracking and profiling such as law enforcement tracking criminals, parents and guardians monitoring children, etc., cannot be ignored, our research highlights not only the extent to which website sessions can be tracked, correlated, and profiled, but also the ease

with which it can be done. When used constructively, website owners can recognize returning visitors through the use of profiling [3]. On the contrary, when used for malicious reasons, the same profile may allow an adversary to track a user between websites and to identify visitors who use a privacy preserving technology (PPT) [3]. There are various ways by which information about a website visitor's activities can be collected without consent, including software downloads, search engine queries, e-commerce, e-mail, spam, and Internet Relay Chat (IRC) [4].

Eckersley suggests that the most common way to track web browsers is via HTTP cookies [5]. As a result, there has been a growing awareness among website visitors that HTTP cookies represent a serious threat to their privacy, which is leading them to block, limit, and even periodically delete cookies. However, privacy is not the only issue when it comes to tracking and profiling website visitors. A deeper concern lies with website owners purposely and dynamically changing their websites' contents based on visitors' previous browsing interests, history, geolocation, etc [5].

We also discuss websites that attempt to present different content based on the personas of the visitors. For this purpose, we consider the idea of fingerprinting visitors' systems and behaviors including web browser fingerprinting [5]. We believe that a user is bound to open multiple browsers without any bias; therefore, profiling and tracking based solely on web browsers will fall short and hence to ameliorate we implement the correlation of visitors across multiple websites.

Furthermore, we argue that users are extremely unlikely to change some of their persona features even in a large time window that spans months or even years. Such features include the operating system (OS), search engine preference, and Internet service provider (ISP) to name a few. Therefore, we propose a method of profiling and tracking website sessions based on entropy computed by multiple features extracted purely on the information website visitors expose through their network connection. Our hypothesis is that we can profile and track a majority of visitors on our custom website with a high level of confidence by using multiple persona features that tend to remain constant over large window of time.

Atterer et al. investigated how the detailed tracking of user interaction can be monitored using standard web technologies, such as tracking website visitors based on their proficiency for completing basic website tasks (e.g., filling out a form) [6]. Similarly, we utilized a brief questionnaire

in our experiment to obtain a baseline of users' proficiency in answering questions about website content utilizing the variable *time spent* by website visitors as part of their profiles. We repeated this process on a second website with a nearly identical questionnaire.

This particular experiment does not attempt to determine whether a persona is associated with an identifiable person. Rather, we are interested in tracking and profiling persona features to determine whether we have seen a certain profile in the past. The European Union's privacy directive defines an "identifiable person" as "one who can be identified, directly or indirectly, by reference to an identification number or to one or more factors specific to 'his physical, physiological, mental, economic, cultural, or social identity [4]. It has been shown that three common pieces of data —zip code, birth date, and gender —can be used in combination to determine the identity of a unique person in 87% of the United States' population [7].

Our motivation for conducting this research lies in the fact that threats associated with tracking and exposing website visitors by their personas has not been addressed to the extent it should have. Consequently, it has failed to draw significant attention from the research community. An entity could acquire personally identifiable information (PII) from a website visitor through both direct and indirect information-gathering techniques. Third parties can also expose an individual's profile for financial, informational, and other type of gain. By luring a visitor to input complementary information across multiple browser sessions over a variety of different websites, an entity can garner and connect piecewise information to create and accrue a complete persona. In summary, our contributions in this area of research:

1. Define a privacy threat model by describing a wide range of features leaked over the Internet that can be used to track and profile website visitors.
2. Evaluate experimentally the impact that identified threats can have on profiling and tracking website visitors using actual human subjects.
3. Quantify the entropy of each of the computed features for each measured persona to determine its privacy impact.
4. Provide defense mechanisms to combat the threats outlined in our threat model.

## 2.2 Correlating a persona to a person

Recently, Online Social Networks (OSNs) have been gaining popularity. Two thirds of the world's Internet population visits a social network site weekly, and the time spent on these sites accounts for more than 10% of all Internet time, with this percentage growing three times faster than the rate of the overall Internet growth [8]. OSNs contain sensitive and personal data of hundreds of millions of people and are integrated into millions of websites. A number of empirical studies on OSNs highlight challenges to the security and privacy of social network users and their data.

OSNs, including but not limited to Facebook, LinkedIn, and Twitter, have been increasing in popularity in every year since their inception. Specifically, Facebook currently has more than 800 million active users and more than 50% of those active users log on to Facebook in a given day [8]. Clearly, the usage statistics backup the fact that Facebook is surging in popularity. Given the popularity of Facebook, there is also a direct relationship with the amount of private data that is posted on the profiles of the 800 million active users. As a result, OSNs are critical with respect to security and especially privacy due to their very large user base [9].

As part of their offerings, OSNs allow their users to connect by means of various link types. These networks allow people to list details about themselves that are relevant to the nature of the network. Since Facebook is a general-user OSN, the individual users often list their favorite activities, books, and movies [10]. The display of the OSN user's private information allows a social network application provider a unique opportunity: direct use of this information which could be useful for advertisers for direct marketing [10].

As we will show in Chapter 7 (Defense Mechanisms) there are at least some privacy protections in place to prevent these efforts. However, there is still the opportunity for social network data mining, that is, the discovery of information and relationships from social network data.

The privacy concerns of OSN users can be classified into two categories: privacy after data release and private information leakage [10]. Heatherly et al. discuss the problem of sanitizing a social network to prevent inference of social network data and then examine the effectiveness of these approaches on a real-world data set. Their solution is to sanitize both details and link details.

There exists multiple ways to determine and track an online persona of a website visitor. One

such way is based on browser fingerprinting [5]. Eckersley et al. investigated the real-world effectiveness of browser fingerprinting algorithms. In addition, numerous other researchers have conducted research in digital fingerprinting.

However, we aim to use profiling and tracking tactics to correlate to an identifiable person as was previously defined. Determining a visitor's identity based only on their persona has major privacy ramifications such as targeted advertisements, cyber-stalking related crimes, tailored and biased content changing websites, identity and financial theft, etc.

In this particular experiment, we leverage the capabilities and the ease in which website analytics and Facebook applications can be used to determine the true identity of an OSN user. To this extent, we created a custom website complete with website analytics.

In addition, we created a custom Facebook Fan Page that is linked from the aforementioned website. We created a Facebook application that links from the custom webpage and has profiling and tracking methods implemented. Our goal is to determine the success rate of relating the analytics taken from the original website with that of our custom Facebook Fan Page and Facebook application.

We hypothesize that visitors are at risk of being profiled, tracked, and led to reveal their true identity by an adversary who uses only basic website analytic tools. This is because of the actions the visitor performs on the Internet, including allowing potentially malicious Facebook applications as well as displaying private information on their Facebook profile.

Given this, the general problem we are trying to solve is twofold: First, determine whether our hypothesis is a valid concern by introducing our threat model and running an experiment with real participants. Second, provide defense mechanisms to help prevent the profiling and tracking attacks we outline in the rest of the dissertation.

## **2.3 Determining the effect of multiple attacks on privacy-preserving technology users**

Organizations and their employees often have a need to remain anonymous while on the Internet. This need is typically predicated on the fact that undermining of the organization or its employees privacy and anonymity would have a severe negative impact on the mission of such an organization. These privacy concerns include identity and financial theft, biased and tailored website content delivery, and compromised true geographic location. To this end, organizations often implement and configure privacy preserving technologies (PPT) to help protect their identity and that of their users while traversing the Internet. We define a PPT as a privacy service or technique that attempts to hide the user's IP address, location, and additional information including but not limited to their research tasks as they access websites on the Internet.

It is worth differentiating between two closely related concepts: privacy and secrecy. According to Warren et al. privacy and secrecy both involve boundaries and the denial of access to others; however, they differ in the moral content of the behavior that is concealed [11]. While we use the term privacy throughout the dissertation, one could argue that we are really dealing with a problem of secrecy because an organization is already by the definition of privacy not concerned about its privacy as it is not an individual. As we will show later, the organization is concerned about secrecy (e.g., research tasks).

The privacy implications of data gathered when users access websites on the Internet especially under the supposed "protection" of a PPT need to be examined and addressed more closely. The exposure of information on the Internet poses a genuine threat to the privacy of users. There are numerous ways in which an online entity can profile and track a website visitor. We propose to investigate how an adversary can collect information from website visitors through the actions of the user, the website itself, or programs that are transmitted through the Internet.

For this experiment, we configured a network monitoring application to monitor network traffic coming from the case-study privacy preserving technology organization's (CSPPTO) users. The network monitoring application is set to create a daily report showing network activity going to

websites including OSNs, personal email, and other attributable websites. We define attributable as search queries that can yield the website visitors true origin such as the address and phone number of a local pizza place in the visitors true geographic location. The categorization of the network traffic was entirely based on DNS names and IP addresses as opposed to intercepting network traffic in a manner that would allow us (e.g., authors) to examine the content of a participant's email or social network profile.

Meanwhile, we created an initial 20-question survey that gathers answers to questions associated with demographic information, Internet behavior, and overall feelings toward privacy and security on the Internet. The objective of the survey is to determine the privacy awareness of users who utilize a real-world implementation of a PPT. The first phase of the experiment is to compare the results of the survey with the true network traffic captured by the network monitoring application. The rationale for collecting demographic information is to determine if gender, age, education, and Internet experience play a role when it comes to privacy and anonymity best practices. We compared the results of the survey with real network traffic that was collected on the PPT.

Next, we transition to the second phase of the experiment namely phishing. The objectives of the phishing campaign phase included crafting a phishing email based on previously completed tradecraft and reconnaissance of the target organization. Furthermore, we included a link within the phishing email to a Facebook application that has profiling and tracking enabled. We leveraged the Facebook application design from our previous research study [12]. We offered security awareness training and discussed defense mechanisms to help prevent against the attacks we discuss in our threat model.

## **2.4 Limitations of privacy preserving technologies**

The Internet offers a low-cost, low-risk, and high-value of return intelligence gathering and archival system. With the advent of online social networking and information aggregation the Internet has become an information highway storing copious amounts of information that is typically unintended for the attacker, but is freely available and relatively easy to discover for a user.

Often users of computer systems perform operations on the Internet, such as searching for information, which they would like to keep anonymous. One popular way to help protect a users identity against a possible attacker is to implement a privacy preserving technology (PPT). The field of anonymous communications and specifically, PPT started in 1981 with David Chaum's Mix-net [13]. In general, a PPT is a system or set of systems that is designed to protect a users identity while they are using a computer system that can access the network. Therefore, a PPT is a tool that attempts to make activity on the Internet untraceable.

As described by Pang et al., the first goal of a PPT is aimed at preventing the true identities of specific hosts from being leaked such that an audit trail of user activity cannot be formed [14]. The second goal is to prevent the true identities of internal hosts from being leaked such that a map of supported services can be constructed [14]. The third goal is to prevent the leakage of specific security practices within the publishing organizations network [14].

When implemented correctly, a PPT will access the Internet on the users behalf and in theory protect personally identifiable information (PII) by hiding the source computers identifying information along the way. The goals of any PPT are to first hide the structural information about the network on which traces are collected and second to prevent the assembly of behavioral profiles for users on that network, such as the sites that they browse [15]. The ultimate goal of Internet anonymization is to allow a host to communicate with a non-participating server in such a manner that nobody can determine the users identity.

Simply implementing a PPT will not necessarily protect a users identity. However, it is safe to say that a PPT system is better than no system at all. As we will describe in Chapter 7, a user will have to be careful and cognizant of what websites they visit and what actions they perform even when protected by a PPT. Often users of PPT gain a false sense of security as they are under the impression that they are completely protected and anonymous. As we will show, this assumption could not be further from the truth.

There exists a plethora of different PPTs including the popular Anonymizer<sup>1</sup> and Tor<sup>2</sup> that are available on the market today. As a result, these PPTs are predominantly built with the goal of

---

<sup>1</sup><https://www.anonymizer.com/>

<sup>2</sup><https://www.torproject.org/>



protecting the network layer of the OSI model in mind. While this does provide a certain level of anonymity and identity protection compared to no system at all, it does leave itself vulnerable to a variety of different styles of attacks.

## **2.5 Survey scams**

Social networking is continuing to reinvent how users consume, view, and share information as millions of users are now using some form of social media [16]. However, as social media platforms become more prevalent, security threats, attacks, and malware continue to grow as cyber-criminals take advantage of the implied trust relationships inherent in social networking [16]. The focus of our experiment will be on the attack vector that Orebaugh identifies as malicious content - wall posts, tweets, and other social media public message mediums that may contain enticing post, such as free gift cards, links to videos of significant events, false advertisements for applications, fake security issues, and celebrity gossip [16]. For example, a common attack is the survey scam that lures users to complete a survey prior to receiving a free gift card [16]. Each survey that is completed earns the scammer affiliate advertising money, and the victim never actually receive the intended gift.

Survey scams are typically found on OSNs like Facebook [17]. Most often they come in the form of wall posts with an embedded link. The cyber-criminal is attempting to lure the victim into clicking on the link with the false promise of a free giveaway. Once the users click on the link they are often re-directed to rogue Facebook applications or pages embedded with malicious URLs [17]. Users are often re-directed again to ad tracking-sites. These sites track the number of web page visits, which serve as additional revenue for cyber-criminals [17].

At the core of the OSN spam ecosystem are the CPA affiliate-based ad networks that handle the task of monetizing the visitors generated by these spam-based abusive advertising channels. These ad networks' presence frees the spammer from needing to deal with monetizing their victims and in turn allows the spammers the ability to specialize in generating more effective spam campaigns. The affiliate program is an efficient organizational model that decreases the risk to both parties and allows for greater flexibility and innovation [18].

As with most ad networks, these CPA ad networks are simply intermediaries in this scheme that

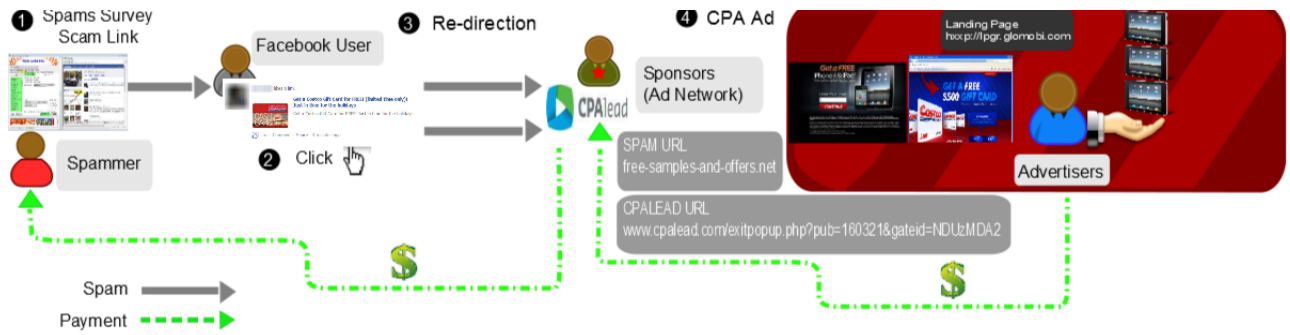


Figure 2.1: Showing the flow of money between the different entities

in turn line up advertisers that are responsible for creating and hosting the actual “surveys.” These advertisers pay the ad network for each successfully performed action, and finally the ad network pays a fraction of this revenue to the affiliates that originally attracted the user to their ad network. Figure 2.1 depicts the typical chain of events and flow of money involved in these survey scams.

## 2.6 Understanding users’ perceptions of privacy and value of information stored in their accounts

According to a September 2008 survey from the Pew Research Institute, nearly 69% of Americans use cloud computing services (such as webmail and online data backup sites) [17]. It is well-known that users often save passwords, images, and other personally identifiable information in the cloud. However, the question arises: why do some people store their passwords in their email? Here are few reasons we identified through survey responses and exit-interview discussions with our participants. 1) They can’t remember all of them so they need somewhere to track them. 2) They underestimate the value of their passwords. 3) The password reset is sent to their email account and they don’t delete the email or change that particular password. 4) They underestimate the threat that someone might steal their password and the damage it might cause if their password is stolen. This is probably most important due to a few dynamics (e.g., being one out of many and wondering why they would be attacked in the first place).

According to Florencio and Herley [19], the average person (tracked by an installed client)

appears to have about seven passwords that are shared, and five of them have been reused at least once within three days of installing the client. Given this many people forget their passwords at an increasing rate because they simply have too many to remember. This leads to the person resetting passwords fairly often for sites that are used irregularly in a lot of cases. Many people need long periods of time too delete old accounts that they don't use and to remove duplicate passwords [19].

In essence, this can be thought of as trimming their attack space. Once an email account is compromised, if it has been used for the primary account on other websites, an attacker can potentially reset the other sites passwords as well. From an adversarys perspective, finding answers to security questions such as “what is your mothers maiden name” is trivial in nature. Another crucial point is that an adversary with access to a persons email address in conjunction with the fact that most people use repeat passwords can often gain access to websites such as Facebook.

There are defense mechanisms available such as two-factor authentication, but that technology has not been fully understood and incorporated by the general public. If the adversary obtains credentials for a shared or linked log in, the end user is going to be *pOwned*. One may argue that a person's Google Mail (Gmail) password is quite valuable because it is often used to log into a plethora of related sites.

The notion of protecting images is not as understood and researched as is the password counterpart. However, with many recent news stories associated with sextortion [2], [1], [3] it is a real threat that the general public needs to recognize. This type of breach raises the risk of sextortion, which occurs when a person has obtained sexually explicit photos of a user and threatens to post the picture online unless he or she receives cash or often even more sexually explicit pictures from the victim.

According to Snyder and Kanich [20], there exists an adversary whose top priority is too make money. The primary adversaries are economically motivated cyber-criminals and opportunistic eavesdroppers. In this threat model, it is assumed that the adversary has complete control over the cloud storage account in question (e.g., stole password or other authenticated device) [20]. Furthermore, the threat model assumes that the attacker has bypassed additional authentication methods like RSA or smart-phone based security. Lastly, it is assumed that the adversary is focused on data

at rest and not in-motion [20].

Additionally, there is another type of adversary to possibly include family, friends, Facebook “friends”, strangers, ex-spouses, enemies, Government, and self (through unintentional and accidental means) who pose the threat of stealing and uploading personal images that were not intended to be shared with the public.

Typically the goal of this type of adversary (sans self) may be to embarrass, incriminate, and extort money from the victim. Related to this is sextortion, a term previously mentioned and is created to describe a form of sexual exploitation that employs non-physical forms of coercion to force sexual favors from the victim [21]. Online this may occur when a child, or teenager, meets a stranger online, through social media or a chat site. The attacker gains the adversary’s trust and convinces him or her to send a lewd photo [21]. The adversary then uses that image to blackmail the victim for more compromised photos, threatening to send the image to the victims parents or friends or publish it online for the world to see unless the victim provides something racier [21].

The consequences that can result in the unintended sharing or posting of an image can be devastating to include potential loss of a job, loss of friends, financial cost, legal trouble, and embarrassment. This threat can impact many different types of people especially children and celebrities. For example, in 2011, Christopher Chaney, a Florida man, was charged with hacking into the emails of Christina Aguilera, Mila Kunis, and up to fifty other celebrities. The case gained the attention of the Federal Bureau of Investigation (FBI) when nude pictures from Scarlett Johansson’s smartphone were posted on the Internet [22]. Though not confirmed, it is speculated that Chaney was able to hack into mobile phones by guessing the users passwords [22].

## Chapter 3: Literature Review

### 3.1 Summary

Based on a review of current literature, there exists publications that highlight the various methods of threatening a user's privacy and safety while on the Internet. Most of the literature ties various aspects of seemingly unrelated information to uncover a user's real identity. Additionally, many papers focus on the increasing popularity of OSN as a fertile research ground into both attacking and defending against threats to a user's privacy. We also found that many security researchers are studying PPT attacks that focus on traffic analysis, network-layer, and application-layer based attacks.

Since cyber-crime is a very large field, we explored a sample of papers associated with spam infrastructure, revenue measurements, phishing, malware, underground forums, fraud, and effectiveness of interventions. The papers are grouped into the following categories: profiling and tracking, privacy preserving technologies, online social networks, cyber-crime, and privacy and security of passwords on the cloud.

#### 3.1.1 Profiling and tracking

Part of our research was motivated by Eckersley's web browser fingerprinting [5]. We also drew inspiration from Atterer et al. [6]. We were interested in looking at how website visitor tracking interacts with web applications. Whether directly or indirectly, virtually all popular websites gather data about the visitors' system [23]. Most website visitors are unaware of the various pieces of personal information that, once combined, can be used to determine their identities and be disseminated to parties other than the websites they visited directly [23].

Assuming the visitors believe in a fundamental right to privacy, the question is raised over who should be able to decide what information about them is being recorded, by whom, and for

what reason. Unfortunately, this notion of “informed consent” is rarely available as an option, as discussed by Krishnamurthy [24]. Even if website visitors follow traditional privacy guidelines and block the use of cookies, JavaScript, and many of the other privacy-exposing features, they still are at risk of privacy fingerprinting [23].

It has long been known that many types of technological devices possess subtle but measurable variations that allow them to be “fingerprinted” [5]. According to Eckersley, a website visitor seeking to avoid being tracked must initially pass two tests: (1) find appropriate settings that allow websites to use cookies for necessary user interface features, but present other less-welcome methods of tracking and (2) learn about each type of supercookie and disable them [5].

Only a few visitors will pass these first two tests; those who do will face a difficult third test in the form of the aforementioned fingerprinting [5]. As a tracking mechanism for use against people who limit cookies, fingerprinting also has the insidious property of being much harder for investigators to detect as compared to supercookie methods [5]. This is because fingerprinting leaves no persistent evidence of tagging on the user’s computer [5].

As Cooper et al. suggest, information about location, both in real-time and at permanent locations, requires special attention due to the consequences for both privacy and physical safety that may flow from its disclosure [25]. A new class of threats related to behavioral profiling and tracking has recently emerged. Examples of this kind of profiling include, but are not limited to, Amazon book recommendations, Google’s tailored search results, advertisements, etc.

Such profiling and tracking can lead to discovery of a visitor’s real identity and location beyond a PPT. Although website visitors often believe that implementing a PPT will protect their identity on the Web, attacks such as those as described in [26], [27], [28], [29], and [30] can break anonymization by using network analysis methods (e.g., watermarking, browser caching, and parsing).

### **3.1.2 Breaching and protecting a privacy preserving technology**

Next, we explore the current literature associated with attacks that target network anonymization which is a common defense mechanism employed by privacy conscious organizations and their users. Almost all PPTs are vulnerable to a variety of different types of attacks. Specifically, we

describe current research that is being conducted to attack PPTs. The literature review considers traffic analysis, application, and network-based PPT attacks.

Scheer et al. [28] discuss the idea that network encryption, both at the packet and session layer, is used widely for securing private data. They use simulation and an analytical model to examine the impact on user experience via a scheme that masks the behavior of real traffic by embedding it in synthetic and encrypted cover traffic [28].

Coull et al. [31] attempt to solve the problem of publishing data that can potentially leak sensitive information about the publishing organization. Coull et al. introduce the problem of publishing data by suggesting that it is imperative that trace and log data be made publicly available for verification and comparison of results [31]. Coull et al. conduct an analysis and create techniques to infer sensitive information from these network traces [31]. The study, as performed by the authors, demonstrates that there are more substantial forms of information leakage that inherently compromise current anonymization methodologies [31].

Wang describes the concept of watermarking the network traffic in order to break the PPT [26]. The watermarking aspect in this case can potentially allow an adversary to influence the traffic thus allowing them to discern the identity of the user [26]. Wang describes how they were able to successfully penetrate the Anonymizer, Inc. “Total Net Shield, the ultimate solution in online identity protection” which is almost exactly the same PPT that is used as a case study in our experiments.

Coull et al. improve on previous research done on reconstructing web browsing activities from anonymized packet-level traces. This is accomplished by accounting for real-world challenges such as browser caching and session parsing [15]. This research evaluated the effectiveness of the author’s techniques by identifying the front pages of the fifty most popular websites on the Internet.

Hopper et al. present two attacks on low-latency anonymizing network systems (ANS) schemes [27].<sup>1</sup> The first attack allows a pair of colluding websites to predict, based on local timing information and with no additional resources, whether two connections from the same Tor exit node are using the same circuit with high confidence [27]. The second attack requires more resources but allows a malicious website to gain several bits of information about a client each time a user visits

---

<sup>1</sup>ANS is used interchangeably with PPT as we previously defined in this dissertation.

the site [27]. The authors evaluate both of their attacks against the Tor network and the MultiProxy (proxy) aggregator service [27].

Coull et al. primary concern is to evaluate the efficacy of network data anonymization techniques with respect to the privacy that they afford [32]. Specifically, the authors are trying to make the network flow uniquely identifiable even after it has gone through the ANS. They are also considering techniques for evaluating the anonymity of network data and to simulate behavior of an adversary whose goal is to de-anonymize the objects (host or web pages) [32].

Jung et al. describe the design and implementation of a system called Privacy Oracle that is capable of finding application leaks [33]. The authors attempt to solve the problem of application leaks by using black box differential testing [33]. The creators of Privacy Oracle are most interested in what information was leaked, when it is exposed, and who can receive it [33].

Scott et al. perform new techniques which address the problem of application web security. They describe their solution to address the problem of application layer web security by describing a scalable structuring mechanism. This mechanism facilitates the abstraction of security policies from web applications developed in heterogeneous multi-platform environments [34]. The second aspect to their solution is to represent a tool which assists programmers developing secure applications which are resilient to a wide range of common attacks. Finally, the third aspect is to report results and experiences arising from the implementation of these techniques [34].

Wondracek et al. introduce a novel de-anonymization attack that exploits group membership information that is available on social networking sites [9]. This paper illustrates the tactic of using a malicious website to launch a de-anonymization attack to learn the identity of the visitors. The advantages of this type of attack are that it has a low cost and has the ability to affect a plethora of users [9].

### **3.1.3 Online social networks**

Unfortunately, from the user's perspective, threats of private information leakage increase along with the growth of OSNs [35]. Furthermore, current research examines the privacy diffusion on the Internet via hidden transactions [36]. This can potentially allow for a few websites to be able to



construct a profile of an individual [36].

Users assume that the information they provide to one OSN will be kept within the boundaries of that particular OSN. The danger of this implicit assumption is that many users don't realize how privacy information can be revealed to other dissociated OSNs [35]. Essentially, the information disclosed at one website could be combined with the information at other OSN websites.

Irani et al. define the term "online social footprint" to be the resulting combination of the information revealed by multiple social networking websites. The main contributions provided by Irani et al. were to measure the size of users' online social fingerprint and to investigate the ease it is to reconstruct a users online social footprint [35].

As indicated by Egele et al., there is currently a non-satisfactory privacy situation regarding third-party Facebook applications [37]. Current access controls for Facebook applications are too coarse grained and even non-existent for legacy applications [37]. For some Facebook applications to work, they must "request data" to be transferred from the OSN user's profile. Often this request is done behind the scenes without the user even realizing it. In some cases, the users quickly click "Allow" to let this request occur [37]. Forcing applications to make profile data requests explicit to the user and funnel such requests through client-side proxies is the goal of the "PoX" solution discussed in Chapter 7 [37].

These third-party applications further escalate the privacy concerns as OSN user data are shared with these applications [38]. When considering Facebook applications, there is typically minimal control over what user information these applications are allowed to access. Even more concerning is the fact that these applications are not hosted on the actual OSN (e.g., Facebook). This makes it difficult to police the data being leaked from the application after the data are shared with the application [38]. This usually occurs due to user error such as the user inadvertently allowing the application or in other cases it is due to vulnerabilities in the application itself.

Sing et al. are concerned with protecting users' private information from leaks by third-party applications and as a result present a mechanism called "XBook" to control not only what the third-party application can access but also what these applications can do with the data they are allowed to access [38].

Additionally, there has been research that attempts to correlate a photo taken of a particular individual with that of the photo used in their Facebook profile [39]. In this research, Acquisti et al. studied the consequences and implications of the convergence of three technologies namely face recognition, cloud computing, and OSNs. We are often able to acquire a Facebook photo that can be used in concert with the research of Acquisti et al. to assist the attacker in achieving their desired objective. In another research experiment, Acquisti et al. show how information readily found on a Facebook profile such as location and date of birth can be exploited to predict that persons social security number [40]. The inferences were made possible by the social security administration's death master file and the wide spread accessibility of personal information on profiles of OSNs [40].

Additional research on de-anonymizing OSN users has been conducted by Krishnamurthy et al. where they show that it is possible for third parties to link personally identifiable information, which is leaked by the OSN, with user actions both inside and outside the realm of the OSN [41]. Given the amount of available personally identifiable information, OSNs have the ideal properties to become attack platforms. Makridakis et al. define the term "Antisocial Networks" that refer to distributed systems based on social networking web sites which can be exploited to carry out network attacks [42].

Another key aspect associated with profiling and tracking based attacks is phishing. Phishing is a crucial aspect of the experiment outlined in Section 4.5. By definition, phishing is a form of social engineering in which an attacker attempts to fraudulently acquire sensitive information from a victim by impersonating a trustworthy third party [43]. Jagatic et al. completed an experiment to quantify, in an ethical manner, how reliable social context would increase the success rate of a phishing attack [43]. This was accomplished by mining information about relationships and common interests provided on a growing number of OSNs such as Facebook. This allowed the phisher to harvest large amounts of reliable social network information.

Many OSN users tend to be overly revealing when publishing personal information while on an OSN [44]. Furthermore, there is information that exists that a user cannot control, and may not even be aware of [44]. Given this, Balduzzi et al. introduced an attack that targets the OSN user

action of finding friends specifically by email address. The results provided by Balduzzi et al. show that given an initial sample of 10.4 million email addresses, the authors were able to automatically identify 1.2 million user profiles associated with these addresses [44].

Chew et al. point out three distinct areas where the hyper-linked world of social networking sites can compromise user privacy [45]. They include a lack of control over activity streams, unwelcome linkage, and de-anonymization through merging of social graphs [45]. Heatherly et al. introduce an attack that explores how to launch inference attacks using released social networking data to predict undisclosed private information about OSN individual users [10].

In the paper, “Imagined Communities: Awareness, Information Sharing, and Privacy on Facebook” Acquisti et al. survey a sample number of Facebook members at a college and compare the findings to the information retrieved from the network itself [46]. The hypothesis set forth by Acquisti et al. is that OSNs raise privacy and security concerns. However, it is unclear as to how different demographics or behaviors play a role in impacting the privacy and security of OSN users. Therefore, Acquisti et al. compare attitudes versus behavior and find that individual privacy concerns are a weak predictor of membership.

Similarly, it was determined that so-called privacy concerned individuals join the OSN and reveal great amounts of personal information that they may not even be aware of [46]. The results of the research performed by Acquisti et al. showed that the Facebook privacy concerns stated in the survey were actually quite high; however, the privacy concerns that were actually captured were mixed but were on the low scale [46]. The reasons stated by Acquisti et al. were due to peer pressure and unawareness of the true visibility of their profiles.

Dhamija et al. provide the first empirical evidence about which malicious strategies are successful in deceiving general users [47]. To this end, Dhamija et al. developed a large set of phishing attacks and hypothesized why these attacks might work [47]. Most phishing attacks focus on either a fake email or a fake website to lure unsuspecting users to. Dhamija et al. primarily focused on answering the question of what makes a website credible [47]. However, they added an interesting twist and researched what made a bogus site credible [47]. The key findings of this paper were that good phishing websites fooled 90% of the participants and that existing anti-phishing browsing

cues are ineffective [47]. They found that participants proved vulnerable across the board to phishing attacks and that neither age, sex, experience, nor hours of computer use mattered when it came to likelihood of falling victim for a spoofed website [47].

Moody et al. point out the use of the Internet and networking technologies continues to rise [48]. Alongside the benefits that are derived from the use of these technologies, numerous threats continue to emerge [48]. The study performed by Moody et al. examines several variables within the message characteristics, personality traits, and Internet-related experience to determine an individual's susceptibility for phishing attacks [48]. The results of an ethical phishing experiment show that message characteristics and Internet-related variables are better predictors of whether an individual will be more susceptible to phishing attacks [48].

Dodge et al. performed an experiment that considers the benefits of training as it relates to phishing susceptibility [49]. The authors report on a recent phishing experiment where the effects of training were evaluated [49]. Additionally, Dodge et al. gathered demographics data to explore the susceptibility of given groups. The results indicated that over short periods of time (e.g., 10 days), there is no significant difference in susceptibility based on training. However, over longer periods of time (e.g., 63 days), training does contribute significantly to the reduction in susceptibility. These findings are in-line with what we concluded based on our training exercises. Therefore, it is recommended that the CSPPTO and other organizations balance the importance of reducing susceptibility of security threats with the increased time and organization efforts involved with providing mandatory training resources [49].

### **3.1.4 Cyber crime**

With the growing digital economy, it comes as no surprise that criminal activities in digital business have lead to a digital underground economy [50]. Holz et al. performed a study of an active underground economy that trades stolen digital credentials [50]. Specifically, Holz et al. report on measurements of the actual kind and amount of data that is stolen by attackers from compromised machines [50]. Essentially, they directly observe the goods that can be traded at an underground market [50]. The key contributions made by Holz et al. are to investigate keylogging attacks based

on dropzones and provide a detailed analysis of the collected data giving a first-hand insight in the underground economy of Internet criminals [50]. Their method can be generalized to many other forms of credential-stealing attacks, such as phishing as we previously discussed.

Franklin et al. studied an active underground economy which specializes in the commodization of activities such as credit card fraud, identity theft, spamming, phishing, online credential theft, and the sale of compromised hosts [51]. Internet miscreants of all sources have banded together and established a bustling underground economy [51]. This economy operates on public IRC channels and actively flaunts the laws of nations and the rights of individuals [51].

Underground forums, where participants exchange information on abusive tactics and engage in the sale of illegal goods and services, are a form of online social networks [52]. The work of Motoyama et al. characterize six different underground forums: BlackhatWorld, Carders, HackSector, HackElite, Freehack, and L33tCrew [52]. We leveraged some of these forums in an effort to gain intelligence and assistance for infiltrating the affiliate programs we were studying (e.g., interview preparation, contacts, etc.) which will be detailed in Chapter 5. Of interest in the work of Motoyama et al. is the means by which a member of these groups first are admitted, how they might be elevated (e.g., status upgrade) within the group, and in some cases what they might have done to be banned from the forum altogether [52].

At the core of the modern spam ecosystem are the “affiliate programs” that provide retail content (e.g., storefront templates and site code) as well as back-end services (e.g., payment processing, fulfillment, and customer support) to a set of client affiliates [53]. One key question often asked is how much money and from who do spammers make money? The work of Kanich et al. represent the most comprehensive attempt to answer these questions to date [53]. The results show that while the profit that a spam-advertised pharmacy makes is substantial, with annual revenue in the many tens of millions of dollars, it indeed falls vastly short of the annual expenditures on technical anti-spam solutions [53].

Despite being a nuisance and a waste of costly resources, spam is used as a delivery mechanism for many criminal scams and large-scale compromises [54]. In the case of survey scams, the URL

links that are posted to Facebook typically originate from spam. In the study of the Cutwail botnet (spam generator) performed by Stone et al., most of the content included pornography, online pharmacies, phishing, money mule recruitment, and malware [54]. This is not surprising as these are many of the key areas in the cyber-crime world [54]. It is not until recently, have we seen a proliferation of survey scams [54].

Kanich et al. present a methodology for measuring the conversion rate of spam using a parasitic infiltration of an existing botnet's infrastructure [55]. For nearly a half billion spam e-mails, Kanich et al. identified the number that were successfully delivered, the number that pass through popular anti-spam filters, the number that elicit user's visits to the advertised sites, and the number of "sales" and "infections" produced [55]. Similar to the conversion of affiliates within the scope of survey scams, there exists three basic parameters of the spam value proposition: the cost to send spam, offset by the "conversion rate," and the marginal profit per sale [55]. The results of their study showed that after 26 days, and almost 350 million e-mail messages, only 28 sales resulted [55]. While it is true that this reflects a very low conversion rate, the fact remains that this was still a lucrative endeavor for the cyber-criminals. After taking the conversion rates and extrapolating this to include an increase in worker bots, the daily revenue increases to \$7,000 to \$9,000 [55].

Thomas et al. examine the abuse of online social networks at the hands of spammers through the lens of the tools, techniques, and support infrastructure they rely on [56]. However, Thomas et al. perform their analysis of suspended Twitter accounts while we focus predominantly on a spam feed collecting traffic from Facebook. There are many parallels between our study and the work of Thomas et al. For example, we both utilize our dataset to characterize the behavior and lifetime of spam accounts, the campaigns they execute, and the wide-spread abuse of legitimate Web services such as URL shortners and free-web hosting [56]. One limitation that both of our studies suffer from is the diverse array of social network spam and its evasive nature makes it difficult to obtain a comprehensive source of ground truth for measurement [56]. Given this limitation, Thomas et al. identified two prominent affiliate programs used by spammers namely Clickbank and Amazon [56]. The Twitter spam marketplace relies on a multitude of services that include popular URL shortners (e.g., bit.ly), free web hosting, and legitimate affiliate programs like Amazon and

illegitimate programs like Clickbank, Assetiz, and other account sellers [56]. Thomas et al. found that the current marketplace for Twitter spam uses a diverse set of spamming techniques, including a variety of strategies for creating Twitter accounts, generating spam URLs, and distributing spam [56].

Anderson et al. suggest that there are many types of losses associated with cyber-crime namely direct losses and indirect loss [57]. Direct loss is the monetary equivalent of losses, damages, or other suffering felt by the victim as a consequence of cyber crime (e.g., money withdrawn, time and effort to reset credentials, etc.) [57]. Indirect losses are the monetary equivalent of the losses and opportunity costs imposed by society by the fact that certain cyber-crime is carried out, no matter whether successful or not and independent of specific instances of that cyber-crime [57]. Indirect costs generally cannot be attributed to individual victims. Examples include losses of trust, missed business opportunities, and PC cleanup [57]. Lastly, there are defense costs which are the monetary equivalent of prevention efforts. They include security products, services, fraud detection, and law enforcement [57]. The summation of these three types of costs is known as the cost to society [57].

Spammers use questionable search engine optimization (SEO) techniques to promote their spam links into top search results [58]. Wang et al. focus on one prevalent type of spam namely, redirection spam whereby one can identify spam pages by third-party domains to which these pages redirect traffic [58]. They ask a fundamental question, “who are the middlemen who indirectly sell spammers service to sites like orbitz.com?” They answer the question posed by “following the money” by virtue of monitoring the resulting HTTP traffic [58].

Profit lies at the heart of the modern malware ecosystem [59]. Grier et al. provide a brief overview of services and techniques that fuel the underground economy. The key areas identified by Grier et al. are Spamming, Information Theft, Clickfraud, Browser Hijacking, Fake Software, Proxies (Hosting), and Droppers [59]. We have seen evidence of all these services and techniques sans droppers in our research study detailed in Chapter 5.

The reason why cyber-criminals continue to use the same social-engineering ploys over and over again is because they work. Human beings, in the majority of instances, seem to be vulnerable to very simple ploys aimed at compromising themselves. This susceptibility allows cyber-criminals

to plant Trojan horse programs, backdoor access tools, and consistently continue to be successful in their widespread campaigns to carry out criminal campaigns. Social networks, such as Facebook, are considered “low-hanging fruit” for these criminals because of the sheer number of users and the overall lack of technical security [60].

### **3.1.5 Privacy and security of password and images on the cloud**

There are many areas of research related to our investigation of visualizing the threat as it pertains to users’ storing their passwords and images in the clear on the cloud (e.g., Google Mail). A number of previous studies looked at the relationship between users and their (often poor) security practices.

Odom et al. [61] conducted a study to look at how people value and form attachments to virtual possessions (e.g., books, photos, music, and movies) with the goal of comparing similarities and differences with their material things. This research is related to our investigation as to how people value their photos. One difference is Odom et al. used a participant pool of teenagers while our participant pool was over a wider age spectrum.

At the heart of our research investigation is the tradeoff between privacy and convenience. Engleman performed a laboratory experiment to study the privacy tradeoff offered by Facebook Connect: disclosing Facebook profile data to third-party websites for the convenience of logging in without creating separate accounts [62]. Engleman observed that most users understood the trade-off between privacy and convenience, meaning those that cared more about convenience than privacy used Facebook Connect [62].

The work of Acquisti et al. [63] showed that empirical and theoretical research suggests that consumers often lack enough information to make privacy-sensitive decisions. Furthermore, these same consumers given sufficient information, are likely to trade off long-term privacy for short-term benefits [63].

The work of Gaw et al. [64] looks at password management strategies for online accounts. In their study, they had a participant pool of 49 undergraduates and quantified how many passwords they had and how often they reused these passwords [64]. Furthermore, over time, password reuse rates increased because people accumulated more accounts but did not create more passwords [64].



While these participants want to protect financial data and personal communication, reusing passwords simply made the passwords easy to manage [64]. The participants in this study visualized threats from human attackers, particularly viewing those close to them as the most motivated and able attackers [64].

Both Grawemeyer et al. [65] and Hayashi et al. [66] investigated real password use in the context of daily life. Grawemeyer presented the results of an empirical study where participants completed a password diary over seven days, followed by debrief interviews to gain further knowledge and understanding of user behavior [65]. The research conducted by Grawemeyer et al. investigated how people create, use and manage a plethora of passwords in the reality of their daily lives [65]. Hayashi et al. focused their research on overall password usage in daily life meaning they did not examine only a specific computer, web site, or organization [66]. Through their diary study, Hayashi et al. collected 1,500 password events which illustrated how participants used passwords in their everyday lives [66].

Howe et al. [67] looked at the psychology and behavior of the home computer user regarding password requirements. Several other studies including Shay et al. [68], Aquisti et al. [46], and Gross et al. [69] looked at privacy concerns and overall attitudes regarding account management especially in online social networks (e.g. Facebook).

Rader et al. found that non-expert computer users regularly need to make security-relevant decisions; however, these decisions tend not to be particularly good or sophisticated [70]. Nevertheless, their choices are not random. Where do these non-expert users find information for their decisions? They argue that much of this information comes from stories they hear from other people [70]. Rader et al. found that most people have learned lessons from stories about security incidents informally from family and friends [70].

Walrave et al. present a study that serves two purposes. First, they explore how adolescents and adults approach the disclosure of personal information and the application of privacy settings on social network sites (SNS) [71]. Second, they investigate whether the factors that predict these two privacy-management strategies differ for adolescents and adults [71]. To achieve the goals of this study, an online survey was conducted among a sample of 1484 SNS users ranging in age from 10

to 65 years [71]. In addition to gender and age, Rader et al. investigated the following predictors: frequency of and motives for SNS use, trust in other users, peer influence and concerns related to privacy and contact risks [71].

Das et al. showed that despite an impressive effort at raising the general populace's security sensitivity, the awareness of, motivation to use, and knowledge of how to use security and privacy tools, much security advice is ignored and many security tools remain underutilized [72]. Part of the problem may be that we do not yet understand the social processes underlying peoples decisions to 1) disseminate information about security and privacy and 2) actually modify their security behaviors (e.g., adopt a new security tool or practice) [72]. To that end, Das et al. report on a retrospective interview study examining the role of social influence or, our ability to affect the behaviors and perceptions of others with our own words and actions in peoples decisions to change their security behaviors, as well as the nature of and reasons for their discussions about security [72].

Shay et al. [68] investigated user attitudes and behaviors regarding text-based passwords. They captured users' opinions about the new, stronger policy requirements [68]. Furthermore, they performed an entropy analysis and discussed how their findings relate to NIST recommendations for creating a password policy [68].

Egelman et al. [73], performed an experiment to determine if password (strength/weak) meters influenced users' password selections when they were forced to change their real passwords. They performed a follow-up experiment to test the creation of passwords for an unimportant account (as deemed by the participant). They found that the meters made no observable difference [73]. That is to say participants simply reused weak passwords that they used to protect similar low-risk accounts [73]

Acquisti et al. [46] looked for underlying demographic and behavioral differences between the communities of Facebook members and non-members with the objective of comparing a member's stated attitudes with actual behavior [46]. The research found that privacy concerned individuals join the network and reveal great amounts of personal information. Some manage their privacy concerns by trusting their ability to control the information they provide and the external access to it. There are many misconceptions about the online communities size and composition and, about

the visibility of member's profiles [46].

Gross et al. [69] studied patterns of information revelation in online social networks and their privacy implications and evaluated privacy settings as well as the amount of information users disclosed. The work highlights potential attacks on various aspects of users privacy, and Gross et al. showed that only a minimal percentage of users changed their highly permeable privacy patterns [69].

Ion et al. [74] looked at privacy concerns for users in the space of consumer cloud storage. To explore user's privacy practices and expectations on the cloud, Ion conducted 36 semi-structured, in-depth interview studies and designed an online survey to confirm the interview conclusions. Additionally, the work of Zhao et al. [75] performed a password-based security analysis and designed and implemented a cloud-based storage-free browser-based password manager.

There were numerous studies that offered potential solutions and recommendations to the password use problem that impact many people regardless of demographics, computer training, experience, education etc. For example, the work of Kelley et al. [76] used the concept of a nutrition label as their inspiration for implementing a similar label for privacy. Kelley et al. were able to offer the user the ability to find information, clearly articulate the difference between privacy policies and control over one's information, and reduce the simple time-based costs of users needing to read privacy policies [76].

The work of Klasnja et al., presented results from an exploratory study that examined how users from the general public understand Wi-Fi and the associated privacy risks, as they expose their names, email addresses, and ZIP codes [77]. The two main contributions of this work are: 1) examination of how users from the general public understand and deal with privacy threats associated with Wi-Fi use and 2) introduction of end-user awareness tools and infrastructural improvements, that seem to hold promise for addressing privacy and security problems with Wi-Fi use [77].

Synder and Kanich, [20], designed a system called "Cloudsweeper" which gives users the opportunity to "lock-up" sensitive, unexpected, and rarely used information to mitigate the risks of cloud storage accounts without sacrificing the benefits of clouds storage or computation. Synder and Kanich conclude that by focusing on automated strategies for finding sensitive information they

can level the playing field with cybercriminals, thereby forcing determined attackers to conduct manual inspection of stolen accounts, which is unlikely to be profitable [20].

There is a large body of past work in the field of password management, password behavior, user attitudes and privacy concerns, and the overall lack of success with passwords as a security mechanism. Our study adds insight into the concerns, consequences, and the behavior of users for both passwords and images.

Unlike previous studies, our investigation is focused on: 1) understanding the behavior and actions humans taken when protecting the privacy of both passwords and images that are stored in the cloud 2) understanding whether human users will behave differently if they can visually see the threat, and 3) identifying the most likely adversary's according to our participants.

## **Chapter 4: Profiling and Tracking (Re-identification) Attack Experiments**

In this section, we describe the first phase of our research associated with re-identification of users through profiling and tracking. Specifically, we will discuss the source of our participants, datasets, tools, experimental methodology, and results. The limitations of our methodology will become evident and include the fact that our participant pool and datasets were quite lacking. Another limitation was the amount of manual work that is potentially error prone. We also discuss the potential ethical issues associated with the types of experiments we conducted. Some organizations and their users may feel that purposely trying to collect and discover system data is a violation. Therefore, we gained authorization from the George Mason University (GMU) Institutional Review Board (IRB) prior to conducting any experiments. The process required us to take an online ethics training course. In addition, an application package including an application form, participation informed consent forms, recruitment poster, recruitment email, and IRB instrumentation documents was submitted.

### **4.1 TRacking and Profiling Internet visitors via website analytics - TRAP**

In the first of our experiments, we determine accuracy for profiling and tracking website visitors by the features (e.g., operating system, browser, search engine) they use to access websites.<sup>1</sup> To the best of our ability, we have tried to make the experiments as realistic as possible.

The participants involved with this experiment knew in advance that the assignment was associated with Internet privacy. However, they were led to believe the experiment was more about

---

<sup>1</sup>This section is based on the unpublished joint work of Jason W. Clark, Angelos Stavrou, and Avinash Srinivasan.

navigating websites and finding, accessing, and allowing Facebook applications. We stated up-front that absolutely no private information would be shared outside of the authors nor would any information that could be used to identify any specific participant be published.

To accomplish our experiment, we collected traffic generated by real visitors accessing our two custom websites.<sup>2</sup> The goal of the experiment was to acquire and determine the usefulness of information that we collected during the exchange between real visitors and the web portal. This included capturing information from website analytics, cookies, and DNS traffic. Using the collected information, our objective was to identify the visitor who accessed both of our custom websites.

#### 4.1.1 Participant's role

The participant provided their partial IP address and generated a random number directly from the website. The combination of random number and IP address is known as the link file. Participants submitted their link files (via a feedback form located on the website) to an email address that the authors did not access until the conclusion of the experiment.

After submitting the link file, each participant completed a feedback form posted on the website, which they submitted to an email inbox. This form, known as the truth file, recorded all the actual features used by the participant, including the IP address and random number, as shown in Figure 4.1.

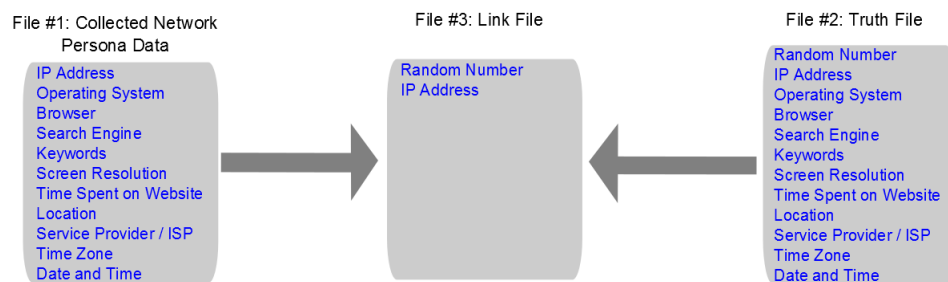


Figure 4.1: Visual representation of the collected persona data, truth files, and link files for website #1 and website #2

<sup>2</sup>The custom websites are located at [www.jamigen.com](http://www.jamigen.com) and [www.nufzek.com](http://www.nufzek.com). Unfortunately, there is no guarantee that these websites will be available and maintained in the future.

The rationale for collecting the IP address was to tie the network persona data we collected with what the participant provided in the truth file. The IP address was something that would likely not change during the course of the assignment. For example, using a variable such as “time spent on the site” would be problematic because the inconsistency of this variable as the visitor accesses each of the custom websites. Once the link file and truth file for each website were submitted, the participant’s role in the experiment was complete.

#### 4.1.2 Lead researcher’s role

We advertised our experiment using fliers on the George Mason University campus in Fairfax, VA. Presumably, the majority of our participants were young, educated, and between the ages of 18 and 25 years old. To encourage participation, we held a random drawing for \$50.00 that was open to those participants who completed our assignment. Prior to and during the experiment, we made several assumptions regarding the participants as shown in Table 4.1.

Table 4.1: Assumptions made about the participants

---



---

The participants were deceived to some extent.
The participant uses many of the same features on each of the custom websites.
The participants accurately and truthfully complete the assignment.
All of the participants followed the exact same steps in order.
All of the participants used their own computer system and Internet connection.
The participant uses the same computer system throughout the experiment.
The analytics are working correctly.
The participants were not monitored during the experiment.
The participants were to act as they normally do on the Internet.

---

We began our role in the experiment by analyzing the collected data (CD) from both websites. Our analysis required us to match features including operating system, browser, search engine, keywords, screen resolution, ISP, and location from both Jamigen and Nufzek. We used the match to determine which visitors ( $x$ ,  $y$ ) to the Jamigen (J) and Nufzek (N) websites are the same. This is expressed in equation (4.1) below.

$$x == y \text{ if } \rightarrow J_{CD}^x = N_{CD}^y \tag{4.1}$$

Once we determine which visitors we believe to be the same across the two websites, we extract from the collected data the IP address for visitor  $J^x$ . Next, we consult the link file of Jamigen and look up the IP address to determine the random number that correlates to the IP address for visitor J. We repeat the procedure by confirming the IP address for visitor  $N^y$ . If the random numbers from  $J^x_{CD}$  and  $N^y_{CD}$  are the same, we are confident in our determination that these visitors are the same person.

However, there is a chance that multiple visitors utilized the same IP address. To ensure certainty, we consult the truth file for Jamigen and Nufzek for these random numbers. If the data provided in the truth file are consistent with the collected data within files #1 for Jamigen and Nufzek we know we have a correct match. We repeat this entire process for all the visitor records. Table 4.2 and Table 4.3 show an example of how we use collected data (file #1), link file (file #2), and truth file (file #3) to profile and determine which visitors are the same.

Table 4.2: Data collected from website #1 during experiment

Visitor#1	Visitor#2	Visitor#3
Windows 7	Windows XP	Windows XP
Firefox	Internet Explorer	Firefox
Google	Yahoo	Google
Jamigen Airlines	Jamigen Planes	Jamigen airplanes
1680x1050	1200x1200	1680x1050
12 minutes	18 minutes	9 minutes
Verizon FIOS	Cox	Comcast
Alexandria, VA	Fairfax, VA	Washington, DC

Table 4.3: Data collected from website #2 during experiment

Visitor#4	Visitor#5	Visitor#6
Windows XP	Windows 7	Windows XP
Firefox	Firefox	Firefox
Yahoo	Google	Google
Nufzek railroad	Nufzek trains	Nufzek trains
1200x1200	1680x1050	1680x1050
16 minutes	14 minutes	11 minutes
Cox	Verizon FIOS	Comcast
Fairfax, VA	Alexandria, VA	Washington, DC



We placed all the collected data into a database. Using the collected data from fictitious visitor #1, we ran a variety of queries against the database. These queries returned that visitors' #1 and #5 were likely to be the same person because they shared many of the same features. Similarly, visitors #2 and #4 were likely the same person. Finally, we stated that visitors #3 and #6 were likely the same person because they share many of the same features. This was an unrealistic example given the small number of visitors and the large amount of persona data that they shared, but it helps explain our experimental design.

## 4.2 Results

We had collected and stored 62 records created by the 31 participants that accessed both of our websites. We present our results as record couples (x,y) where x = the visitor ID on Jamigen and y = the visitor ID on Nufzek.

First, 50 out of the 62 persona records matched almost exactly. The remaining six record couples (12 individual records) seemed to match on some but not all the key persona data.<sup>3</sup> These record couples are (8, 38), (11, 49), (16, 44), (18, 54), (19, 55), and (28, 40).

Couple (8, 38) matched on operating system, screen resolution, location, date, and time. However, they did not match on a few crucial items, such as browser and search engine. The record for visitor #8 suggested that the browser utilized was Internet Explorer 9.0, while the browser used for visitor #38 was Mozilla Firefox 6.02. In addition, the web analytics captured suggested that visitor #8 typed *www.jamigen.com* directly into the browser without using any search engine. The visitor record for #38 suggested that Google was the utilized search engine.

Record couple (11, 49) matched on nearly every persona feature except for two key variables, namely location and ISP. This may have been caused by completing the task from two different locations, incorporating a PPT, or for other unknown reasons. We committed to this answer although our confidence was low.

Next, we considered record couple (16, 44), which matched on every feature except for browser. The browser utilized by visitor #16 was Internet Explorer version 8, while visitor #44 used Internet

---

<sup>3</sup>The collected record data can be found at [mason.gmu.edu/~jc1arks/data.pdf](http://mason.gmu.edu/~jc1arks/data.pdf)

Explorer version 9. It is likely that this visitor upgraded their browser during the assignment.

After analyzing couple (18, 54) we saw that the visitors used two different browsers. In this case, the browser change was from Internet Explorer to Google Chrome. This is conceivable as many visitors utilize multiple browsers.

The next record couple we analyzed was (19, 55), and most of their features matched with the exception of screen resolution. We thought it unlikely that a visitor would alter their screen resolution between tasks. The record for visitor #19 did not utilize any keyword searches, and we found it odd that visitor #55 did.

The record couple we felt least confident about was (28, 40); although their records matched on the operating system (OS), browser, and search engine, there were other mismatched items. For example, their locations and ISP's were different. As we saw with couple (11, 49) there are sound reasons why the location and ISP could be different. However, the keyword searches of visitor #28 and visitor #40 also differed.

We were correct on 50 of 62 total records. This calculates to an accuracy rating of 81%. The initial results were promising because we could profile and track the majority of participants as they accessed each of our websites during the experiment. Figure 4.2 shows a graphical representation of the success rate for our guesses.

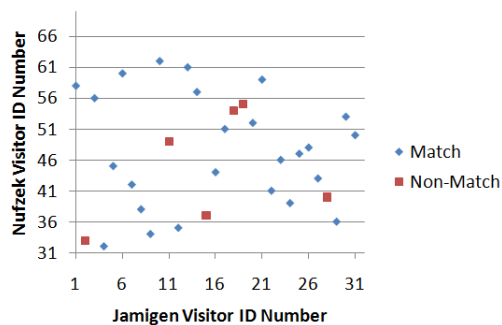


Figure 4.2: Visual representation of the success rate for our guesses.

### 4.2.1 Cumulative entropy

Next, we analyzed the results and considered cumulative entropy and how it can be used to learn more about a given visitor. There is a mathematical quantity which allows us to measure how close a fact comes to revealing somebody's identity uniquely [78]. Entropy can be thought of as a generalization of the number of different possibilities there are for a random variable measured in bits [78]. For example, if there are two possibilities, there is one bit of entropy; if there are four possibilities, there are two bits of entropy, etc. Adding one more bit of entropy doubles the number of possibilities [78]. Using equation (4.2), we calculated the threshold entropy needed to identify a visitor to our website [78].

$$S = \log_2 P_r(n/N) \quad (4.2)$$

Where  $n$  is the number of unique visitors and  $N$  is the number of total visitors. In our experiments  $n=1$  and  $N=31$ . Using equation (4.2) we calculate the entropy to be 4.95 bits of information. When we learn a new fact about a visitor, that fact reduces the entropy by a certain amount which can be computed using equation (4.3) below.

$$\Delta S = -\log_2 P_r(X=x) \quad (4.3)$$

Using equation (4.3) we calculate the entropy for the persona features that we collected. We begin by considering OS; during our experiment we learned there were five distinct operating systems used by the participants. As an example, we input the values into equation (4.3) with  $X=OS$  and  $x=Windows\ XP$ .

$$OS = \Delta S = -\log_2 P_r(OS = Windows\ XP) = -\log_2(1/5) = 2.32 \quad (4.4)$$

We repeat the process for each of the other persona features we collected and the results are shown in Figure 4.3. Given we know the visitor operating system (OS) and browser combination, we have a cumulative entropy of  $2.32 + 3.58 = 5.90$  which is greater than the threshold entropy

calculated in equation (4.2). Hence, we know that a given persona correlates to a given visitor ID. If we take other combinations, such as OS and search engine, we have a cumulative entropy of  $2.32 + 2.0 = 4.32$  which is less than the threshold entropy of 4.95 as shown in equation (4.2). Hence, we need to utilize additional features such that the cumulative entropy is greater than or equal to the threshold entropy calculated in equation (4.2). Therefore, we add an extra feature such as ISP for an additional 3.91 to exceed the threshold entropy.

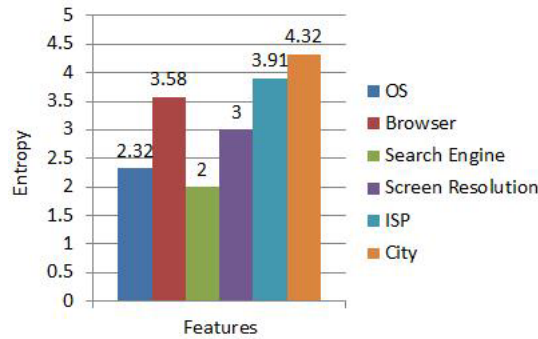


Figure 4.3: Calculated entropy for each network persona feature.

## 4.2.2 Weighted entropy

Since different features are used more frequently than others, we extend the cumulative entropy calculation. Specifically, cumulative entropy fails to consider the fact that certain features are dependent on one another. For example, if we know that the operating system is Linux, we can be almost certain that the browser is not Internet Explorer. Therefore, the entropy that we gain or lose with the browser depends on the operating system we already know.

We calculated a weighted entropy using the same aforementioned equations. We analyzed the visitor data in order to determine which features were the most utilized. In our analysis, we did not consider specific versions of operating system (e.g., service packs) or browser versions (e.g., Internet Explorer version 8 or Internet Explorer version 9). The breakdown of feature type, frequency, and entropy are shown in Appendix A.

### 4.3 Correlating a Persona to a Person

In our second experiment<sup>4</sup>, we test the hypothesis that the majority of visitors can be profiled, tracked, and led to reveal their identity by an adversary mostly due to the actions the visitor performs on the Internet and the information they freely display on their Facebook profiles.<sup>5</sup> To this end, we correlated a visitor's online persona with their true identity at least according to their Facebook profile. We made a conscious effort to make our experiments as realistic as possible. To accomplish this, we collected traffic generated by real participants accessing our custom websites, Facebook Fan Page, and Facebook application

Figure 4.4 shows a visual depiction of how an adversary can profile, track, and identify unsuspecting users simply by capturing basic network information as they visit seemingly innocuous websites owned by the adversary. Of note are the three collectors: 1) Website collector, 2) Fan page collector, and 3) Facebook application collector. Both the website collector and the fan page collector capture the following information: IP address, operating system, browser, search engine, keyword, screen resolution, ISP, and location. The Facebook application collector is responsible for capturing the same information as the other collectors but also captures the user's Facebook ID, name, and profile data. The goal of our second experiment was to acquire and determine a true persona that can be correlated to a given Facebook profile. This included capturing information from website analytics, session cookies, DNS traffic, and overall network data in order to extract the necessary Facebook information.

---

<sup>4</sup>This experiment is based on the work of Jason W. Clark as published in [12].

<sup>5</sup>While the terms are the same, this particular experiment is a slight deviation and includes a different set of participants when compared to experiment #1.

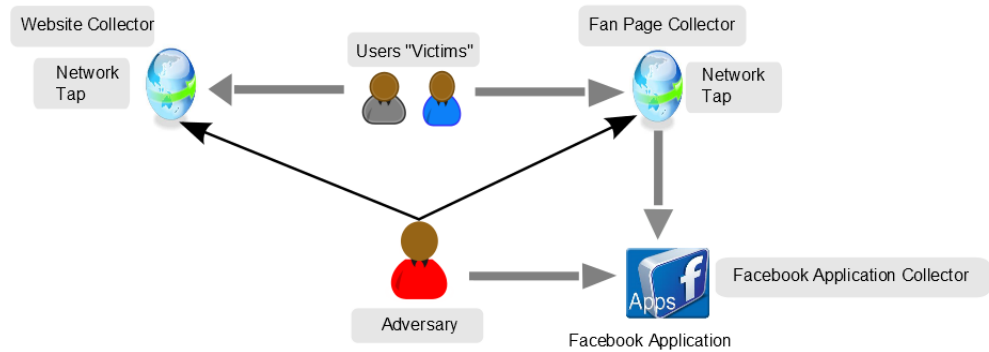


Figure 4.4: Depicts the adversary using tracking code and analytics to correlate a persona to a person

### 4.3.1 Participant recruitment

From January to March 2012, we obtained 25 participants to complete our experiment. We initially utilized Amazon Mechanical Turk and created human intelligence tasks (HITS) based on our assignment.<sup>6</sup>

We decided to post the assignment with the caveat that it was an Internet privacy experiment. However, we still incorporated some level of deception as the participants were not specifically told why they were visiting the website and Fan Page. Furthermore, the participants likely did not understand the ramifications of allowing the Facebook application.

We advertised our experiment using fliers on the George Mason University campus in Fairfax, VA. However, we only had a few responses and so we opened it up by advertising the experiment via <http://www.craigslist.org> again stating that it was an Internet privacy experiment. To encourage participation, we held another random drawing for \$50.00 that was open to those participants who completed the assignment.

### 4.3.2 Participant's role

To accomplish the objective, the participants were to complete an assignment described below. The participant used their own personal Facebook account to complete this experiment. Using their preferred computer system and following their typical behavior when on the Internet, the

<sup>6</sup>We had a strong response, but soon realized that we were violating the Amazon Mechanical Turk privacy policy. Specifically, we were violating the portion of the policy that stated HIT “workers” should not perform tasks that will give away their true identity. We regret not following the policy close enough and immediately pulled the HIT as a result.

participants were provided a link to access the custom website.

Their second objective was to complete a form displayed on the website. On the form, the participants, were asked whether or not they had private browsing enabled. Next, the participant was asked to reveal the Facebook profile that they would be using. The form would be emailed to an inbox that was not accessed by any of the authors until the completion of the experiment. The form containing the participants answers and accompanying name of the Facebook profile is known as the truth file.

In addition, the participants, were asked to access a Facebook Fan Page associated with this experiment. The participants were informed that a Facebook Fan Page has been created and it can be linked directly from the initial website the participants accessed. Next, the participants were required to access a Facebook Fan Page using their Facebook account. Lastly, the participant is directed to a Facebook application. The Facebook application, if allowed by the participant, will track the profile of their Facebook account. The participant's role in the experiment concludes once they access the custom websites, Fan Page, and application.

### **4.3.3 Lead researcher's role**

At the conclusion of the experiment, we review the analytics that are uniquely created for each participant that accessed our custom website. Next, we look at the analytics from the Facebook Fan Page that is provided to see if we can correlate the visitor from the custom website with the same visitor who accessed the custom created Facebook Fan Page. At this point, we collect the data and input into a database. Next, using our custom application, we obtain similar analytical data along with the crown jewel the actual Facebook profile. To help clarify the experiment, consider the following example.<sup>7</sup>

---

<sup>7</sup>The data are displayed as operating system, browser, search engine, screen resolution, time, ISP, and location.

### Website #1

Collected Data for Visitor #1 = (Windows 7, Firefox, Google, 1680x1050, 9 minutes 1 sec, Comcast, Fairfax, VA)

### Facebook Fan Page

Collected Data for Visitor #1 = (Windows 7, Firefox, Google, 1680x1050, 11 minutes 14 sec, Comcast, Fairfax, VA)

### Facebook Application

Collected Data for Visitor #1 = (Windows 7, Firefox, Google, 1680x1050, 10 minutes 5 sec, Comcast, Fairfax, VA) and Profile of Visitor #1 = (This includes all of the information provided by the visitor on their profile (e.g., name, profile picture, location, date of birth, pictures, relationships, friends, etc.)).

Using the combination of the analytics captured from the custom website Fan Page and correlating that to the profile uncovered from the application, we hypothesize that we will be able to uncover the majority of the visitors. Table 4.4 displays the data collected for each particular aspect of the experiment, namely, the custom website, Facebook Fan Page, and the Facebook application.

Table 4.4: Data collected from website, fan page, and application during experiment

Website#1	Fan Page#2	Application#3
Windows 7	Windows 7	Windows 7
Firefox	Firefox	Firefox
Google	Google	Google
1680x1050	1680x1050	1680x1050
Comcast	Comcast	Comcast
Fairfax, VA	Fairfax, VA	Fairfax, VA
No Profile Info	No Profile Info	Profile Info

Recall that the data were collected by the authors in the form of analytics and provided to the authors (for later use) by the participants in the way of emailed forms. Figure 4.1 depicts how the forms are interconnected for verification of the authors' guesses.



## 4.4 Results

By the completion of the experiment, we had collected data from 25 participants who completed the assignment associated with our experiment. Of the 25 participants, we were able to accurately determine the Facebook profile name of 16 participants as verified by the truth file.<sup>8</sup>

We were wrong about the remaining 9 participants. This yielded a success rate of 64%. We present our results as record couples  $(x, y, z)$  where  $x$  = the visitor ID on the custom website,  $y$  = the visitor ID on the Facebook Fan Page, and  $z$  = visitors first name on the Facebook Profile.

The record couples that we were correct on included: (1, 42, Leslie), (2, 29, Jasdeep), (4, 33, Siva), (5, 37, Gina), (8, 40, Tarun), (9, 50, Animesh), (10, 41, Alicia), (12, 27, Banu), (13, 30, Stalin), (14, 43, Buxer), (15, 44, Gregory), (17, 28, Ali), (20, 45, Flores), (22, 49, Malcolm), (23, 31, Xavier), and (24, 35, Johnny).

One minor and perhaps interesting observation we made is the (Facebook profile) names are not popular names in the United States. This seems to indicate visitors perhaps coming from foreign countries which we later corroborate based on the GeoIP information that we collected. Given the collected data associated with the profile, we were able to determine the participants picture, Facebook ID, IP address, and other computer-based features. For example, the data for the first record couple are:

*Name (Leslie), Profile Picture (Cat), Profile ID (ends in 40), IP Address (ends in .55), Browser (Opera), and OS (Linux).*

Next, we considered the record couples that we were wrong about according to the truth file. We define wrong to be either that we were not able to acquire a corresponding Facebook profile at all or the Facebook profile guess that we made was incorrect.

We begin by discussing the two record couples that we did not correlate to any profile namely (7, 38) and (21, 47). There are a plethora of reasons that we might not have been able to capture the Facebook profile. For instance, there may have been a technical issue associated with our analytics and Facebook application. Another viable reason is that the visitor removed or changed their profile after completing this experiment. The most likely scenario is that the participant did not complete

---

<sup>8</sup>The collected record data can be found at [mason.gmu.edu/~jclarks/data2.pdf](http://mason.gmu.edu/~jclarks/data2.pdf)

the entire assignment. Perhaps the participant accessed the custom website and the Facebook Fan Page but for possible privacy and security reasons decided not to access the Facebook application which would have given us the profile information we were seeking.

Next, we grouped together three sets of groups that all had similar analytic information. This made it difficult to determine which persona was associated with what Facebook profile. Simply put, we had to do our best to parse through similar data and make our guesses. Unfortunately, since the data points were essentially the same, our guesses ended up being incorrect. Specifically, data records (6, 46, Sanjay), (11, 26, Jozef), and (19, 39, Deepak) were interchanged. The correct record couple should have been (6, 46, Jozef), (11, 26, Deepak), and (19, 39, Sanjay) according to the truth file.

Next, we considered record couple (16, 48, Jasmine), which matched on all features with the exception of search engine. The search engine utilized by visitor #16 was Yahoo, while visitor #48 used Google. As a result, when we compared this to the analytics taken from the Facebook application data associated with our guess of Jasmine, they had reverted back to Yahoo. We utilized the logic that the participant started with Yahoo, tried Google, and reverted back to Yahoo. However, when we consulted the truth file, record couple (16, 48) did not link to Jasmine.

We completed a similar analysis for record couple (18, 36, Matthias), which matched on all features except for search engine. The search engine utilized by visitor #18 was Google, while visitor #36 utilized Yahoo. The captured analytics according to the Facebook application stated that the search engine used was Google. When we consulted the truth file we saw that the name was Matt. A logical explanation is that Matt is a nickname for Matthias, and that we were actually correct in our guess. However, as indicated in the assignment, the participants were asked to be specific and conscientious about stating the exact profile name as it appears. We decided to err on the side of the caution and not give credit for this guess.

The record couple we felt least confident about was (3, 34, Marcy); although their records matched on some features, others did not match at all. For example, their locations and corresponding ISP's were different. However, it is possible that the visitor started the assignment at work and finished at home. In the end, we guessed that (3, 34) was Marcy when the truth file revealed it

was Arianna. Similarly, we considered record couple (25, 32) to be Arianna, but it turned out it was really Marcy. Figure 4.5 shows a graphic representation of the success rate for our guesses. In summary, we were correct on 16 of 25 total records. This calculates to a grade of 64%. The results were encouraging because we could profile, track, and obtain the profiles of nearly two-thirds of the participants in our experiment.

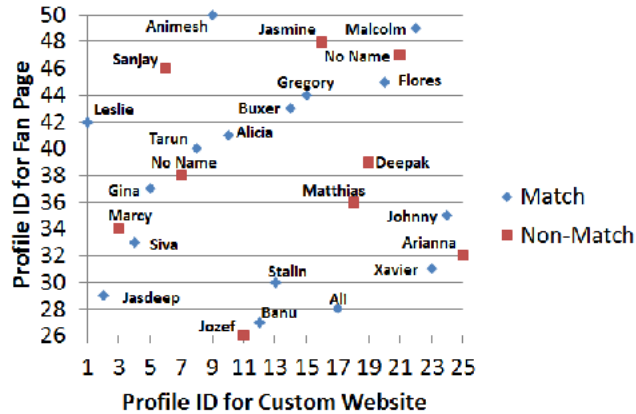


Figure 4.5: Accuracy of guesses displayed by ID overlaid by profile name

The information that we were able to acquire as a result of this experiment was the name of the Facebook profile, their picture (where applicable), Facebook Profile ID, IP address, Browser, and Operating System.

Going more in depth, we were able to utilize the MaxMind database and specifically the Perl API and the commercial GeoIP City, GeoIP ISP, and GeoIP Organization databases to effectively gather even more data about the location of the participant. Specifically, we were able to identify the hostname (e.g., shown as IP address), Country Code (e.g., US), Country Name (e.g., United States), Region Name (e.g., Minnesota), City (e.g., Deluth), Latitude (e.g., 46.8144), Longitude (e.g., -92.1269), ISP (e.g., Charter Communications), Metro Code (e.g., 676), and Area Code (e.g., 218) of each participant. All of this collected can be stored, shared, or sold by the attacker for a variety of economic, political, or personal reasons.

Next, we were able to acquire the user agent string and get more detailed information regarding the visitor as they accessed our Fan Page. For example, we acquired the following user agent string of a participant: *Mozilla/5.0 (Windows NT 5.1) AppleWebKit/535.16, (KHTML, like Gecko)*

*Chrome/18.0.1003.1 Safari/535.16*. Using this particular string we were able to get a more detailed snapshot of the computer used by the participant. The detailed meaning of the aforementioned user agent string is shown in Table 4.5.

Table 4.5: User agent string

Mozilla	Used for historical reasons
5.0	Mozilla Version
Windows NT 5.1	XP OS
AppleWebKit	Set of classes to display web content
535.16	Web Kit build
KHTML	KDE Open Source HTML layout engine
Chrome	Browser is Chrome
18.0.1003.1	Chrome Version
Safari	Based on Safari
535.16	Safari build

## 4.5 Determining the effect of multiple attacks on privacy preserving technology users

Our third experiment<sup>9</sup>, attempts to understand the behavior and motivations of users when they are accessing the Internet where protecting privacy and anonymity is crucial. Figure 4.6 depicts an adversary who implements three phases namely network monitoring, phishing attacks, and online social network applications that when used collectively can determine the identity of a user behind a privacy preserving technology (PPT).

To begin, we organized a team to perform the experiment consisting of the research manager, security manager, and researchers who conducted and implemented the attacks. We administered an initial survey to the case study privacy preserving technology organization (CSPPTO) users to gain an understanding of how they valued their anonymity and privacy. Furthermore, the survey allowed us to gauge whether or not the CSPPTO users were aware of best practices in protecting their anonymity and privacy. The survey was complemented by analysis of data captured from

<sup>9</sup>This experiment is based on the work of Jason W. Clark as published in [79].

the CSPPTO users. All the data from the survey, captured network traffic, phishing emails, questionnaire, and Facebook application was collected, analyzed, and written into a final report that was briefed to the CSPPTO managers. Based on the results, tailored security awareness training was provided to all CSPPTO users. However, we still continued to monitor the network traffic to attributable websites to determine if it has been reduced or stopped by virtue of the (successful) security awareness training.

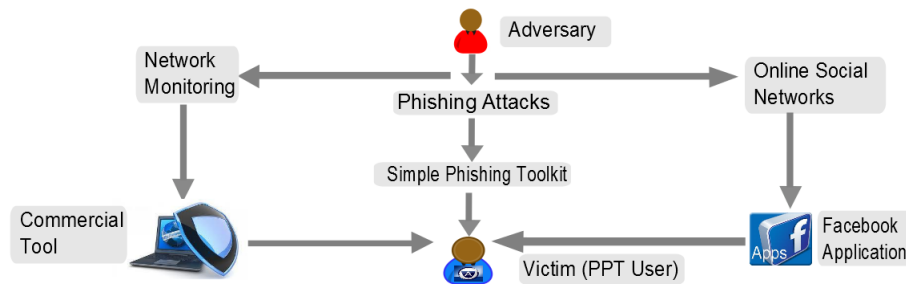


Figure 4.6: Shows an adversary implementing a variety of attack phases that when used collectively can identify a PPT user

At the start of the experiment, we captured network traffic and generated a report showing the websites accessed by the users via the PPT. As a necessary prerequisite for luring the CSPPTO users, we devised a phishing campaign targeted at the CSPPTO users. We investigated the frequency and type of information the CSPPTO users placed on increasingly popular OSNs, such as Facebook. It should be pointed out that these attacks were very much interconnected and did not run in sequence but rather in parallel.

The adversary (e.g., one of the authors) viewed OSNs, such as LinkedIn and Facebook, to find employees of the targeted organization. To add to the legitimacy of the phishing email, we also scoured the Internet looking for a sample of writing of the sender. Upon completion of the survey and subsequent attacks, we provided security awareness training that discussed possible defenses against the attacks outlined in our threat model. At the conclusion of the training, we continued to monitor the network traffic in an effort to see if the CSPPTO users refrained from accessing websites that would allow an adversary to discern their true identity, location, and relevant task work.

The participants were all employees of the case study organization and were users of the PPT

offered by the organization. We had 160 participants that were recruited by the CSPPTO and provided to one of the authors since they were the users who would most likely be impacted by the threats defined in our threat model. The adversary would most likely only acquire a subset of this particular group of users.

Table 4.6 displays the demographics of the survey participants including gender, age range, field of study, education, and Internet usage. The purpose of this table is to determine whether or not we can draw any conclusions about the likelihood of the user falling victim to our attacks based on demographics.

Table 4.6: Demographics of the survey respondents

Survey Results Demographics		
Demographic	Value	Number of Observations (Percent)
Gender	Male	89 (56%)
	Female	71 (44%)
Age Range	18 to 30	92 (58%)
	31 to 40	51 (32%)
	41 to 50	11 (7%)
	51 to 60	3 (2%)
	61 to 70	2 (1%)
	71 and older	1 (1%)
Field	Computer Related	118 (74%)
	Non-Computer Related	42 (26%)
Education	Doctorate (Ph.D)	38 (24%)
	Masters	41 (26%)
	Bachelors	81 (50%)
Internet Usage Per Week	0 to 40	121 (76%)
	41 to 80	29(18%)
	81 to 120	9 (6%)
	121 and higher	1 (less than 1%)

The survey questionnaire contained 20 questions, a consent section, demographic section, usage questions, privacy questions, anonymizer-related questions, and Facebook-related questions. We asked the participants to answer all questions as truthfully as possible. Only participants affiliated with the CSPPTO were allowed to take the survey and it was anonymous. We used a non-technical approach by way of using a locked dropbox where the participants could drop the survey. Upon

completion of the survey, we compared the survey responses with the results of our attacks. Specifically, we compared the networking monitoring that we collected to determine if the responses of the survey were accurate as to what the users were truly doing on the PPT.

To complete our experiment, we unleashed the “Kitchen Sink” attack in three different but overlapping vectors. The vectors are as follows: Network Monitoring, Phishing, and Online Social Networks.

#### **4.5.1 Network monitoring**

In order to invoke this attack, we used a commercial network monitoring tool. The network monitoring tool we selected provides real-time situational awareness, security event correlation, and vulnerability assessments. Using this tool, one of the authors created a daily report showing the network monitoring access of the CSPPTO users. We could have easily used an open source monitoring tool such as Wireshark, but the commercial tool gave us the ability to capture different types of network traffic and generate reports. With the assistance of the CSPPTO system administrators, we configured the commercial tool to monitor network traffic coming from the PPT. The network monitoring tool is set to create a monthly report showing network monitoring activity to online social networks.

As an adversary might do, we utilized the tool to capture all the network activity. Thus, we tried to determine what websites users were accessing and whether they might yield interesting data from an adversary’s point of view. Additionally, we took the results of the network monitoring we captured and determined whether it conformed to the results of the initial survey. The rationale for doing that was to see if users were honest about their behavior and whether or not they were aware that they might be easily giving up information they did not intend to. We captured the initial network traffic prior to the training. After that, we gave the security awareness training and re-monitored the network traffic. We only analyzed the network traffic that was captured on the weekdays since that is when the CSPPTO users were expected to be working on their research tasks.

## 4.5.2 Phishing

Phishing is a common first step used by an adversary trying to penetrate and acquire sensitive information from a targeted organization. Essentially, phishing is a form of social engineering in which an adversary attempts to fraudulently acquire sensitive information from a victim [43]. This practice is usually done by directing or luring users to fraudulent websites, tricking them into clicking on attachments, or both. Therefore, we decided to conduct an experiment in which we designed an internal phishing attack on the CSPPTO users. The purpose was to study whether CSPPTO users would be coerced into revealing information via a well-crafted phishing attack. Additionally, we also wanted to assess the level of awareness of the CSPPTO research staff when it comes to protecting their privacy.

To assist with our phishing campaign, we utilized an open source tool called the “Simple Phishing Toolkit” (SPT).<sup>10</sup> To create a phishing campaign, we modified an existing SPT template and systematically selected an individual target for the attack. The SPT allows us to embed links directing to the website of our choosing.

As part of the initial phishing email, we attached a document that provided an overview of the anonymity and privacy form and a brief questionnaire for the CSPPTO researchers to complete and return. The objective of the questionnaire is determine the rate of success of the phishing campaign. Presumably, any CSPPTO users who willingly provide answers to the questionnaire would likely provide information about the research task. We waited for the questionnaire to be sent back to us. Upon receiving the questionnaire, we collected the results of the number of people that responded to the questions. After a period of three days, we sent a follow-up phishing email from the SPT.

## 4.5.3 Online social networks

The online social network attack phase had two goals from the adversaries point of view. First, to determine whether the CSPPTO users have OSN accounts. Second, to determine if those that do access their online social network account while on a privacy preserving technology attempt to protect their personally identifiable information. The experiment was designed to test the hypothesis

---

<sup>10</sup>[www.sptoolkit.com](http://www.sptoolkit.com)



that the majority of visitors can be led to reveal their identity by an adversary mostly due to the actions the visitor performs on the Internet and the information they display on their Facebook profiles [80].

## 4.6 Results

In the initial survey<sup>11</sup>, we asked about users' knowledge of phishing and feelings toward organizational and individual privacy concerns. Additionally, we asked whether users have a Facebook account and if they regularly post personally identifiable information to their profile. Finally, we asked whether or not users alter their privacy and geographical settings, allow Facebook applications, and are aware that Facebook can reveal personally identifiable information (PII). The survey responses were either Yes / No or based on a 5-point Likert scale, which included the following acceptable responses: Extremely Important (5), Very Important (4), Neutral (3), Somewhat Important (2), and Not Important at All (1). The number of users who answered with a given response and the corresponding percentages are shown in Table 4.7.

---

<sup>11</sup>The entire set of surveys can be found at [mason.gmu.edu/~jc1arks/survey\\_kitchen.pdf](http://mason.gmu.edu/~jc1arks/survey_kitchen.pdf)

Table 4.7: Survey results associated with phishing knowledge and privacy on the Internet

Survey Results Privacy		
Individual Privacy	5	60 (37.5%)
	4	69 (43%)
	3	16 (10%)
	2	11 (7%)
	1	4 (2.5%)
Organization's Privacy	5	55 (34%)
	4	65 (40%)
	3	20 (13%)
	2	9 (6%)
	1	11 (7%)
Facebook Account	Yes	157 (98%)
	No	3 (2%)
Purposely Post Inaccurate PII	Yes	42 (26%)
	No	118 (74%)
Change Privacy Settings	Yes	43 (27%)
	No	117(73%)
Change Geographical Settings	Yes	63 (39%)
	No	97 (61%)
Allow Applications	Yes	50 (31%)
	No	110 (69%)

Table 4.8 displays the responses associated with the participants use of a PPT to search various categories of websites. We asked the users to state how often they access Facebook, Twitter, personal email, and search for local restaurants, sports, or both while under the protection of a PPT. The survey responses were based on a 5-point Likert scale, which included the following responses: All the time (5), Almost Always (4), Often (3), Sometimes (2), and Never (1).

Table 4.8: Survey results associated with PPT usage

PPT Usage		
Local Restaurant	5	5 (3%)
	4	19 (12%)
	3	43 (27%)
	2	42 (27%)
	1	51 (32%)
Local Sports	5	6 (4%)
	4	19 (12%)
	3	30 (19%)
	2	38 (24%)
	1	67 (42%)
Facebook on PPT	5	7 (4%)
	4	19 (12%)
	3	32 (20%)
	2	38 (24%)
	1	64 (40%)
Twitter on PPT	5	4 (3%)
	4	9 (6%)
	3	15 (9%)
	2	38 (24%)
	1	94 (59%)
Personal Email on PPT	5	21 (13%)
	4	39 (24%)
	3	14 (9%)
	2	20(13%)
	1	66 (41%)

In regards to the phishing phase of the “Kitchen Sink” attack, 92 of the 160 participants (58%) sent back the questionnaire with validated information. We found no correlation between gender, age, field of study, or Internet usage when it came to falling victim to our phishing experiment. This correlation validated what was found in other phishing experiments such as the work completed by Acquisti et al. [46] and Dhamija et al. [47].

Additionally, 6 victims replied to the sender of the initial phishing email asking if the content of the email was valid. We of course responded to this inquiry, assuring them that everything was legitimate and safe. Another 31 participants sent the email directly to the CSPPTO internal phishing monitoring email inbox. Of the 160 participants, we were able to determine the Facebook profile name of 34 participants for a success rate of 21% which we later confirmed to be real employees

at the CSPPTO. This was a much lower success rate compared to the results of our previous study [12]. This may be due to the smaller participant pool initially used. The remaining 126 participants did not click on the Facebook application. One interesting result is that there was one person who did not send back the questionnaire but did click on the Facebook application.

In Chapter 7, we discuss the details of the security awareness training; however, we wanted to share the results of the post-training survey here in the results section. After the training on phishing awareness was given, a survey was provided to all the participants to help determine the effectiveness of education as it relates to defending against phishing attacks. Of the 160 participants, 96 responded to the post training survey. For a variety of reasons, not all of the participants either attended the training or wished to fill out the post survey.

Figure 4.7 shows the responses to the question of “What was your understanding of phishing **before** the security awareness training?” and “What was your understanding of phishing **after** the security awareness training?” In conclusion, the phishing experiment was successful and, by the number of responses showed that our phishing email, was quite convincing, that people are just replying to emails without attempting to analyze their authenticity, or both.

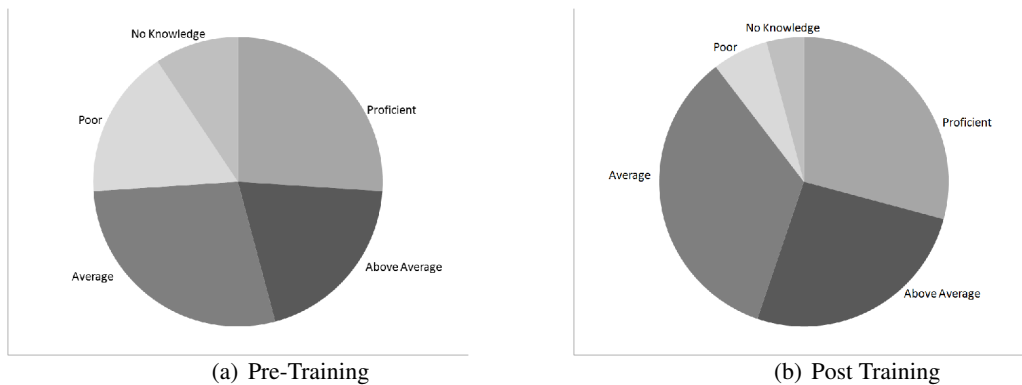


Figure 4.7: Security awareness training displayed as a pie chart

## 4.7 Attacking a privacy preserving technology system

The focus of this experiment<sup>12</sup> is the analysis of the actual anonymity that a PPT provides. We assume that this system is going to be used to protect the identity of a user or an organization in a real-world setting using operational scenarios. The actual setup employed a widely used variation of the PPT offered by Anonymizer, Inc. To gather experimental evidence, we developed a case-study based on a typical PPT deployment in a large United States based organization. We call this version the case-study privacy preserving technology or CSPPT. The CSPPT represents a typical implementation of a PPT that is available today.

The purpose of this experiment is to introduce the shortcomings of the CSPPT by conducting attacks that focus on the actions of the users. We were approached by the management team at the case-study agency about the security and privacy of the CSPPT. Specifically, the CSPPT management was interested in determining whether the users were performing actions that could lead to their identity being discerned. The goal of this experiment is to allow defenses to be developed that can help reduce the likelihood that attacks against PPT are successful.

The main argument that we make is that organizations and users can't solely rely on a PPT to protect their identity. Rather, we warn users to be cognizant of their actions when on a PPT and provide users recommendations and defenses to help protect against the attack vectors we outline. This experiment is significant because it shows that users behind a PPT are not as protected as they need to be. By completing this experiment, we show which attacks are the most successful and more importantly why they are successful.

The concept of researching PPT and their ability to protect a client's identity is something that has been studied for the last decade. Also, the concept of basic attacks by analyzing web traffic and attempting to perform application-layer attacks against systems and users is not original either.

However, the merging of the two concepts is a new idea. This experiment focuses on explaining how the security of PPT, such as Anonymizer, Tor, and the privacy preserving technology implemented at the case study organization, can be "broken" by performing a variety of different and often unsophisticated attacks. It is this new outlook that makes the research original and worthwhile.

---

<sup>12</sup>This experiment is based on joint work between Jason W. Clark and Angelos Stavrou as published in [81].

Furthermore, our research is original in the sense that it takes a hard look at all the essential ingredients of a specific attack used against a real-world PPT. We will conclude the research by offering recommendations to users of PPTs as to best defend against the attacks that we outline. While it is true that we are predominantly focusing on CSPPT, our research will be able to scale to many other PPTs. The reason for this is because they all have the potential to suffer from the same kinds of attacks that we describe.

The primary goal of this experiment is to determine the extent to which we could use application and user-based attacks to discern the true identities of the CSPPTO users. Also, we wanted to show that network-level anonymization is not enough and that there needs to be better education to promote better anonymization [81]. To accomplish this goal, we created documents that were used to both track and infect the users. We also examined the impact of users performing attributable search queries while on the PPT. We relied on analysis of traffic that we captured to find key points of interest.

The experiment for our case study lasted for two months from January 2011 to March 2011. There were 20 users that were included in the experiment, and they were asked to conduct all research within the confines of the PPT offered by the case study organization. Prior to the start of the experiment, we created and hosted an inconspicuous website and ensured that both the non-malicious and malicious files were tested before being uploaded. Furthermore, we ensured that file logging and tracking was enabled to allow us to capture statistics from the web server. The website utilized various sources for content and was approved by the research manager as being enticing enough to attract the researchers.

The next step was to ensure that the website could actually be found via a simple search. By using Google search engine optimization and similar search ranking tools, we found that our website could in fact be located by virtue of simple search logic. A major aspect of the proposed experimental design was to determine the attack vector. For the purposes of this experiment, we focused exclusively on Adobe PDF.

The design of our experiment focuses on three major aspects, namely impact, stealth, and depth. The impact aspect measures how many users actually accessed the website and clicked on the files.

The stealth aspect indicates how many users observed and notified management about the website and malicious file. Lastly, the depth aspect quantifies the quality of the malicious file as it relates to the amount and value of personally identifiable information (PII) that a potential adversary could discern [81].

#### **4.7.1 Participant's role**

The users were specifically asked to complete a research task that focuses on the commercial airline industry. This fictitious research task is like any other task that the CSPPTO users would typically perform. The research manager provided the users with a list of keywords to search on as part of the task. To protect the anonymity of the research tasks being done at the case-study agency, the actual and specific keywords have been sanitized and modified to appear more generic. However, to give the reader a frame of reference, we mostly considered keywords associated with the airline industry including: *Airports, Flight Times, Pricing, Mileage Rewards, Safety and Security, and Jobs*. The assumption here is that the keywords are general enough to be known by a potential adversary as part of their targeting and profiling. Therefore, simply protecting the keywords is not a viable security measure.

The necessary prerequisite for this experiment was to design and build the aforementioned website. We researched each of the keywords independently and extracted content from other related websites to build into our own custom website. Since the goal of the experiment is not a demonstration in website design, we developed a basic HTML website and hosted it on a third-party hosting site. The website corresponded to the different topics associated with the keyword list and, using simple logging techniques, the website can log the actions of the users. The files were uploaded to the “documents” section of the website. The files included documents associated with the keywords.

#### **4.7.2 Technical prerequisites and traffic analysis**

Each file was randomly selected to fall into one of three categories. The first category was that the files were original and completely clean. The second category was that the files were embedded with a JavaScript (`app.launch`) function that would automatically open a new webpage in the user's

browser. The file would open a webpage called `www.website.com/3490` where 3490 is a random number that can be easily identified in the data logs. The third category was that the file itself was embedded with malicious code via the Metasploit Framework.<sup>13</sup> If the malicious code is opened in a vulnerable version of Adobe Reader on a vulnerable Windows operating system, the user's machine could possibly become compromised. Once compromised, it would be trivial for the adversary to de-anonymize the user.

In the final case, as was described earlier, we created a custom website that is meant to entice the CSPPTO users. The website content is associated with a list of keywords that a typical CSPPTO researcher would use as their search query. Included in this website are a series of files that serve different purposes. The files have the ability to track the users and to also infect the users. The files, if downloaded (and transferred) to their client machine, will potentially give the adversary the ability to discern the user's true identity. It is for this reason that the success rate of the adversary is the highest. The complexity of obtaining traffic from a custom website is also low because there is no dependence on intrusion prevention systems (IPS) or similar services.

To substantiate our claims, we also attempted to gather and analyze the traffic that was generated by the researchers while on CSPPTO. We ran the Wireshark packet analyzer to accomplish this. The generated Wireshark .pcap was then run against a tool called tcpextract and used to carve headers and footers from the traffic.<sup>14</sup> This extraction yielded several web images, cookies, files, and portions of web activity. Next, by using a custom-developed Perl script, we were able to take the destination IP address that was captured and in most cases identify the hostname, internet registries (e.g., apnic, arin, ripe), autonomous system (AS) number, the country where the servers are located, the routing information, and the date the website was registered.

We took all this information and created a database so that we could perform queries on the data that we found. The database included over 5,000 records as a result of running the experiment and capturing the traffic. Figure 4.8 shows a visual summary of the experiment including the relationship between key players and entities.

---

<sup>13</sup><http://www.metasploit.com/>

<sup>14</sup>[tcpextract.sourceforge.net](http://tcpextract.sourceforge.net)



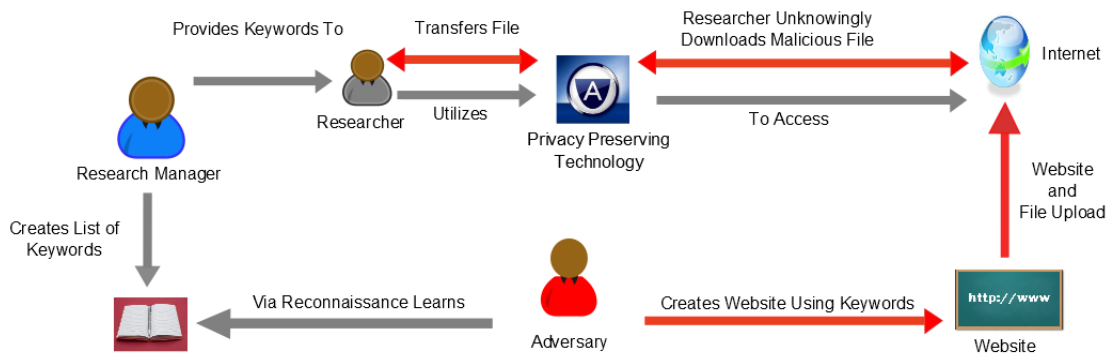


Figure 4.8: Summary of the experiment including key players and entities

## 4.8 Results

Upon analysis of the traffic, we were able to identify a variety of data points that could have been used to identify the CSPPTO and its users. First, we saw several connections being made from the real (un anonymized) IP address of the CSPPTO. The source of the IP address was likely a result of the researchers utilizing remote desktop (RDP) from their client (un anonymized) system to access CSPPTO.

Furthermore, we identified almost daily requests to Facebook. This was a concern because of the attributable nature of this particular social networking website. We began to analyze traffic cookies and ultimately uncovered two distinct email addresses via the Firefox profiles.<sup>15</sup> Next, we logged on to Facebook and did a “Friend Finder” search based on the two emails that we uncovered. The results of our search yielded the names of two different people who were later shown to have been part of the research group at CSPPTO.

Next, we discuss the concept of Facebook profile ID. For example, when you login to Facebook (assuming you have not replaced ID with username) you will see something similar to the following whereby the *xyz* correlate to real numbers:

*http://www.facebook.com/profile.php?id=604xyz*

The traffic analysis also uncovered the researchers accessing Amazon, Continental Airlines, and several other United States based companies. Furthermore, there were other attributable websites

<sup>15</sup>Note: It was necessary for us to download SQLite database browser in order to extract this information from Firefox.

such as LinkedIn and Twitter that were being regularly accessed by the CSPPTO users.

By using the tcpextract tool, we were able to uncover portions of email that were created from Google Mail (Gmail) and Thunderbird. While it was difficult to view the email in its entirety, it appeared to have a subject associated with the Transportation Security Administration (TSA).

By capturing the traffic for such a prolonged amount of time, we were able to find various patterns associated with the time of day users log in to CSPPTO, the types of sites (both attributable and non-attributable), and the routing of the traffic as it makes its way to these sites. We compared this traffic to that of non-anonymized traffic acquired from George Mason University (GMU) and found that it was quite easy to tell which traffic patterns were associated with what entity namely GMU or CSPPTO.

In addition to capturing the traffic and utilizing tcpextract, we also made use of a tool called network miner.<sup>16</sup> This tool allowed us to import a Wireshark .pcap file and mine the traffic. The tool quickly places the traffic into categories such as hosts, frames, files, images, messages, credentials, sessions, DNS, parameters, keywords, cleartext, and anomalies. The main results that were acquired from this tool were the search criteria and queries used by the researchers. For example, we found the following search criteria used by the researchers “petrochemical production process” which is an acceptable search query that relates to the given research task. However, we also came across a search query for “Fairfax County School Closings” that is clearly more attributable to the CSPPTO location.

We saw the anonymized IP address 207.195.x.x that was identified as Global TAC, LLC. This log entry showed up several times throughout the experiment on any given day. We identified that several of the files from the website were downloaded by this IP address. Shortly thereafter, we could see that our random page was identified in the log files but this time with the NAT IP address which was confirmed to be that of the target organization. This was repeatable as we saw new IP addresses being related to the CSPPTO IP address.

Also, most of this activity appeared during the hours of 8:00 a.m. to 5:00 p.m. EST, the time-frame most United States based organizations are open for business. In our logs, we were able to

---

<sup>16</sup><http://www.netresec.com>

clearly see the web browser and the operating system of the user who opened a certain page in our website. This information could be extended to other de-anonymization projects such as [17]. The adversary could also use this information to customize particular Metasploit exploits.

In addition to the logs supplied by our web server, we also utilized tracking statistics. This allowed us to have a separate tracking mechanism in place, and we were able to identify on several occasions that a user was in fact logged into our website. We also noticed that both the CSPPTO anonymized IP address and the case study organization were logged in at the same time. Since the traffic to our website is light, it became quite clear that there was some relationship between the two sets of IP addresses that are simultaneously logged in. This was further supported by the fact that both users logged out of our website at nearly identical times.

Within the logs, we were able to determine whether or not we have seen this person before and, in some cases, determine the first time they visited and the last time they visited. In addition, we were able to identify the system hardware used and the browser, browser language, operating system, and screen resolution. Next, we could uncover the ISP and NAT IP address of the CSPPTO organization. Another piece of data that we saw in the logs was the hostname of the CSPPTO. This hostname appeared to be fictitious as it was the name of a children's game. This information was then looked up via nslookup and came back with an IP address that is tied to the CSPPTO web proxy. Finally, we identified the connecting city, state, country, time zone, latitude and longitude.

One particular user was paying close enough attention and identified a suspicious "tracking code icon" that was intentionally displayed on the site. However, as far as we can tell, this was the only time that anyone from the CSPPTO identified the website or the experiment as suspicious.

Also, we were able to see that on certain days, due to what the IP address was on that given day, that Google Location was able to identify the location of the CSPPTO organization within about 20 miles. However, on other days, the Google Locator was not able to identify the CSPPTO location. These results show that at least one CSPPTO IP address was compromised. This was likely due to attributable search queries being performed. It is not immediately clear whether or not it was a CSPPTO user who was responsible for the attributable search queries.

Lastly, and most damaging, one user downloaded a malicious file, transferred it to their vulnerable Windows XP computer, and opened it in a vulnerable version of Adobe Reader. By utilizing the Metasploit multi-handler listener, we were able to insert the Metasploit payload and open a shell to the CSPPTO user's system. At this point, we could easily have dropped a keylogger, taken a screen capture of the page, run system commands, and quite easily harvested information that would result in determining the user's true identity.

## Chapter 5: Monetizing Online Social Network Users via Survey Scams

In this chapter, we discuss the second phase of our research associated with how an adversary can monetize unsuspecting users. Specifically, we focus on a type of scam known as the survey scam. Cyber-criminals have financial incentives to implement scams. One such scam is known as the survey scam whereby a person will be asked to complete a survey with the (almost always) fake promises of free merchandise (usually in the form of giftcards) [82]. Traditionally, a survey form is provided to the unsuspecting victim via a plethora of different means such as URL redirection, email spam, and posts to an online social network (e.g., Facebook Wall). When a potential victim completes the survey form, they are often asked to provide personally identifiable information which is collected and re-sold by the affiliate (often referred to as sponsor) behind the scam. Furthermore, the victim will almost never get to the end of a survey and will be re-directed through a seemingly never-ending labyrinth of free-offers, advertisements, and forms to complete before receiving their falsely promised iPad. Today we lack insight into the most basic characteristics of this particular scam. How many affiliates are behind the survey scams? What are the most popular giveaways? What is the impact in terms of new spam offers (email and on Facebook wall) of survey scam victims? What, if anything, can be done to discourage cyber-criminals from using survey scams as a means to monetize users? The avidity to answer the questions posed empirically is the major motivating factor of our current and future research.<sup>1</sup>

Our exploration of the connection of ad networks and their sponsoring of affiliates that engage in Facebook spamming is based on two sources of data. The primary source of data for our investigation is a feed of Facebook spam detected by the MyPageKeeper Facebook application and crawling data [83]. Our secondary source is data from infiltrating ad networks that were identified as sponsoring abusive affiliates.

---

<sup>1</sup>This research is based on joint work between Jason W. Clark and Damon McCoy as published in [82].

Here, we describe the source of our datasets, and our approach to infiltrating the affiliate networks. The limitations of our methodology will become evident and include the fact that our data feed came from one source. Another limitation was the amount of manual work that was potentially error prone. Another difficulty was the ability to infiltrate the affiliate networks especially those hosted in foreign countries.

## 5.1 MyPageKeeper spam feed

From December 2012 until April 2013, we successfully crawled 1708 Facebook spam URLs that were identified by the MyPageKeeper Facebook application, which 2.2 million Facebook users located around the world have installed to protect their Facebook profiles from spam [83]. Of these, 16% (283) were survey scams, 50% (862) were broken links, 27% (458) were false positives, and 6% (105) were non-survey scams. This means that 73% (283 of the 388 working spam URLs) were survey scams. An inherent limitation of our study is the fact that our spam feed might contain a bias based on the algorithm used to detect spam posting and the user base that has installed MyPageKeeper. However, previous studies on the quality of spam feeds shows that user-based spam feeds as opposed to spam trap-based feeds tend to provide good coverage [84].

We then crawled these URLs using a webcrawler that is capable of following HTTP and Javascript redirection chains [85]. This crawling produced screen shots of the page that we visually inspected to identify CPA ad network offers and that we manually interacted with to identify which ad networks were sponsoring each URL.<sup>2</sup>

Table 5.1 illustrates the total number of Analyzed URLs, Spam URLs, Landing URLs, Ad Networks, and Publisher ID.<sup>3</sup>

Table 5.1: Summary of crawled and manually collected data

Total	Analyzed	Landing	Ad Network	Publisher ID
1708	129	93	32	77

<sup>2</sup>Note that many of the initial spam URLs are URL-shortening services and lead to a smaller number of unique landing pages.

<sup>3</sup>To reiterate, the analyzed URLs and associated landing pages, ad network, publisher ID, and offer ID are far less than the total URLs we crawled because we only analyzed URLs with unique landing domain names or that directly redirected to an ad networks offer page.

## 5.2 Infiltration

We attempt to register as an affiliate at each ad network that we encountered in order to gain additional insight into the ad networks.<sup>4</sup> Out of the 32 ad networks, we were able to successfully join as an affiliate at 18. Table 5.2 shows the names of the ad networks we encountered and whether or not we were able to register successfully.<sup>5</sup> It should be mentioned that there are at least three unique ways to measure/calculate prevalence. The first method is to count all the raw URLs that we encountered. The second method is to count all the unique landing pages, and the third method is to count all the publisher IDs. Each method has its own bias and drawback.

By infiltrating these ad networks, we were able to obtain two important pieces of ground truth information: (1) links provided to their affiliates, which allowed us to extract affiliate and offer IDs as demonstrated below, and (2) it also allowed us to understand how affiliate IDs are assigned, which we use in the results section to estimate the age of spammer’s affiliate accounts.<sup>6</sup>

Table 5.2: Summary of the prevalence of the affiliates calculated using two different methods: 1) Initial URL 2) Landing (Land)

<i>URL</i>	<i>Init : Land</i>	<i>URL</i>	<i>Init : Land</i>
007CPA	2 : 2	Fileice	2 : 2
A4D	4 : 3	Forestview	2 : 2
Ad.fly*	3 : 3	Gurumedia	2 : 2
Adjal*	1 : 1	LifeStreet Media*	6 : 6
AdscendMedia*	2 : 2	Lyris	1 : 1
AdvertMarketing	2 : 2	MaxBounty*	1 : 1
Adworkmedia	2 : 2	Obey.my*	2 : 1
Altervista*	1 : 1	PulsePoint*	2 : 2
Amung.us*	35 : 33	Rapleaf*	1 : 1
Aweber	1 : 1	ViralUrl*	2 : 2
Bodis*	1 : 1	W4	1 : 1
ClickBanner	13 : 4	Whitefire	1 : 1
Clicksor*	6 : 1	YeahMobi*	1 : 1
CPALead*	15 : 12	Zoosk	1 : 1
Escalatenetwork	1 : 1		:

<sup>4</sup>Often, the ad networks required that we provide an explanation of our marketing methods, including our plans for driving traffic to the affiliate and a required upfront “interview.”

<sup>5</sup>Sometimes language barriers were the primary reason we were unsuccessful at infiltrating the ad network, as was the case with ClickBanner, which is based in Greece. Other times the ad network was essentially closed and by invitation only.

<sup>6</sup>The \*indicates that we successfully joined this ad network at least once.

### 5.3 Ad network and affiliate ID extraction

We manually interacted with these 129 unique landing pages and recorded network traffic traces. Analysis of these network traffic traces allowed us to identify the sponsoring ad network in most cases and, in addition, we were able to identify what is called the publisher or affiliate ID that belongs to an account that the spammer registered with the ad network.

In order for an affiliate to get credited with a completed survey, the ad network provides their affiliates with a URL that in most cases includes the affiliate’s ID number and the offer ID. Table 5.3 provides some examples of the initial URL, the parsed affiliate and offer id, sponsor name, and the full ad network URL. Furthermore, for the programs we were able to infiltrate, we verified that our methods of identifying the ad network and extracting the affiliate and offer IDs were correct.

Table 5.3: Extraction of affiliate ID, offer ID, and Wireshark capture of URL

Initial URL	Aff. ID	Offer ID	Ad Network URL
bit.ly/PJEPk4	2862	14885	network.clickbanner.gr/z/14885/CD2862/
claimafreeiphone5	117373	5055	mb01.com/lnk.asp?o=5055&c=918273&a=117373
bit.ly/TygM3T	1292	2199	track.007cpa.com/aff_c?offer_id=2199&aff_id=1292
bit.ly/SjEthA	3326	699	track.adjal.com/aff_c?offer_id=699&aff_id=3326
getappleipad2	13495	?	adscendmedia.com/gwjs.php?aff=13495&prf=15206
SportsRewards.me	454369	?	a4dtrk.com/?a=454369&c=17648
bit.ly/R9v5ft	1643	7424	adworkmedia.com/go.php?camp=3079&pub=1643&id=7424
goo.gl/1tkTrj	5654	809FA6	jmp.realtraq.net/aff_c?offer_id=8096&aff_id=5654

Next, we present some of our initial results on prevalence of ad networks, carbon dating, and revenue generation as it relates to survey scams.

### 5.4 Prevalence

As previously shown, Table 5.2 includes two metrics, both unique spam URLs and landing pages, to estimate the prevalence of ad networks. Using either of these metrics, Amung.us ranks first and CPAlead is second. Both of our methods have different limitations, such as link shortener URLs are over counted in the case of unique spam URLs and landing pages are under counted if the landing page is the offer page instead of an intermediate page. Given these limitations, using the first metric of 129 unique spam URLs shows that over 50% of these URLs were traced back to four ad networks:



Amung.us, CPAlead, ClickBanner, and LifeStreet Media.

In short, prevalence of ad networks that sponsor Facebook spam is difficult to measure with a high degree of confidence given our limited dataset and limitation of our methods to infer prevalence.

## 5.5 Carbon dating

In a previous study by Kanich et al. [86], they were able to estimate the revenue generated by illicit pharmacy affiliate programs using the insight that order identification (IDs) were sequentially allocated for each new order. We make use of a similar insight that affiliate IDs appear to be allocated sequentially in five of the ad networks identified. If our sequential affiliate ID allocation hypothesis is correct, we use it along with some minimal ground truth data to “carbon date” (estimate the age of) affiliate IDs we extracted from the spam URLs.<sup>7</sup>

Via ad network infiltrations, we were able to obtain ground truth data of the age of two or more affiliate IDs spread out over time for eight ad networks. We use this information to estimate the rate of affiliate ID allocation based on the increase in affiliate IDs we were allocated and the time that elapsed between registrations. If we assume that the rate of affiliate account creations is somewhat stable in the past, we can use the measured account registration rate to carbon date the affiliate IDs we have extracted.

We use the MaxBounty ad network as a concrete example of how our carbon dating methods works. When we initially infiltrated this ad network on 1-30-2013, we were assigned an ID of 123929 and when we joined MaxBounty for the second time on 4-22-2013 we received 129103 as our ID. To calculate the rate of affiliates joining the program, we take 82 days, which is the time that elapsed between our first and second time joining MaxBounty, and the difference in ID assigned, which is 5174, and this results in a rate of 63 affiliates per day joining MaxBounty.

We observed MaxBounty affiliate ID 117373 in the scam feed which is 6556 less than our ID issued on 1-30-2013. Thus, the estimate from our carbon dating method is that the 117373 was

---

<sup>7</sup>In the case of CPAlead, we confirmed by rapidly creating two affiliate accounts that we were allocated sequential affiliate IDs. Additionally, we have never observed an affiliate ID in the wild that was greater than one we were allocated within the corresponding time periods.

issued on approximately 10-18-2012. To reinforce our results, we joined MaxBounty for a third time on 6-21-2013 and received 132424 as our ID. Therefore, we take 142 days which is the time elapsed between our first and third time. The 132424 ID is 8495 less than the initial ID of 123929 which results in a rate of 60 affiliates per day joining MaxBounty. This rate would still put the carbon dating of the ID 117373 we observed in the scam feed to be within the week of 10-18-2012.

We acknowledge that our carbon dating method can only provide an estimate of an affiliate account's age. However, given this limitation, we find that for the eight ad networks that we can compute their rate of affiliate account registrations the average spammer affiliate account age was approximately nine months old upon first observing this affiliate ID. There are two possible reasons that the majority of the spammer's affiliate accounts are old: (1) Spammers age their account before using them to avoid suspicion from the ad networks that are generating survey completions by spamming. (2) The ad networks are not doing a good job of detecting misbehaving affiliates that are engaging in abusive spamming activity.

## 5.6 Revenue estimation

Ad networks allow affiliates to select which offer to direct users to from a wide number of offers, and each one of these offers has a payout that is the amount of money paid to the affiliate for each successfully completed offer. Additionally, some ad networks provide the average conversion rate and Expected Payout per Click (EPC). Figure 5.4 shows that CPAlead offered a survey with the name "Airline Survey" with a payout of \$1.24, an EPC of \$0.02, and a conversion rate of 2%.

Table 5.4: Offers taken from CPAlead as seen in June 2013

Name	Payout	EPC	Conversion	Type
Coke or Pepsi?	\$1.20	\$0.04	3%	Survey
Best Buy Spree	\$1.20	Unavail	Unavail	Survey
Airline Survey	\$1.24	\$0.02	2%	Survey

Table 5.5 shows the minimum, maximum, average, and median payouts of survey offers from three ad networks.

Table 5.5: Offer payouts for June 2013

Ad Network	Minimum	Maximum	Median	Average
Adscend	\$0.11	\$11.90	\$0.63	\$1.33
CPAlead	\$0.03	\$34.00	\$1.05	\$3.52
MaxBounty	\$0.60	\$3.75	\$2.50	\$2.33

## 5.7 Content locking and clickjacking

Cyber-criminals have implemented a useful method for enticing potential victims into clicking on (YouTube) videos which takes them to a third-party webpage pretending to show them a video but instead helps the scammers spread their link. This method, known as content locking with elements of clickjacking, typically depicts a video of sexual content. During our analysis of the .pcap(s), 147 out of 1708 URLs were content locking, clickjacking related YouTube videos all coming from CPAlead. In each case, there was an enticing video, and when we (acting as the unsuspecting user) clicked on the play button we were immediately given a choice of three surveys to complete. There was a never-ending supply of surveys and free offers, and we were never able to actually see the content of the video.

The purpose of scams is typically to lead the victim to online surveys (which earn the scammers affiliate commission) or to trick the victim into handing over personal information such as their cellphone number which will then be subscribed to a premium rate service [87]. One day the scammers will be using links purporting to be videos of giant snakes eating zookeepers, the next it might be a sexual explicit video of an underage teenager with a title such as “*Shocking 17-year-old public high school antics*” [87]. The content locking URLs utilize a piece of code aptly named *locker.php* along with a publisher ID and a gate ID (e.g. offer ID) that is used as the method of doing accounting in a federated domain environment.

## **Chapter 6: Understanding User Perceptions of Privacy and Value of Information Stored in their Accounts**

In the last phase of our research<sup>1</sup>, we aim to get at the heart as to why users behave in a certain manner as it results to their security and privacy actions. Specifically, we investigate users behavior with respect to storing and securing both passwords and images within their Google Mail account. We use data collected in the form of three separate but related surveys, along with the responses taken from an audio recorded exit interview. These data are combined with the results of running both custom applications we designed namely the Cloudsweeper and the Gmail Image Extractor. These data when combined allows us to answer questions associated with risk tolerance for password and images.

### **6.1 Participant recruitment**

We recruited 30 participants from <http://www.craigslist.org>. We selected Craigslist based on popularity, local coverage, and to get a more diverse sample compared to posting fliers around campus. We offered \$60.00 USD in compensation for the participants time and effort. The requirements were that the participants needed to be at least 18 years old, regular Google Mail users as determined by both age and number of emails, and willingness to sign the required consent forms. Table 6.1 shows some of the demographics of our 30 participants.

---

<sup>1</sup>This section is based on the unpublished joint work of Jason W. Clark, Damon McCoy, Chris Kanich, and Peter Snyder.

Table 6.1: Demographics of our 30 participants

Demographic	Number of Participants
Male	11
Female	19
18 to 24 years of age	16
25 to 34 years of age	9
35 to 44 years of age	4
45 to 54 years of age	1
African American	15
Asian	6
Caucasian	6
Hispanic	2
Other	1
Left blank	1
High School degree or equivalency	3
Some college but no degree	9
Associate degree	2
Bachelor degree	12
Graduate degree	3

The survey results showed that the average time spent per day on the Internet was 7.6 hours, with the maximum time being 15 hours, and the minimum time 2 hours. Figure 6.1 shows the frequency as it pertains to the location where the participants typically access their email.

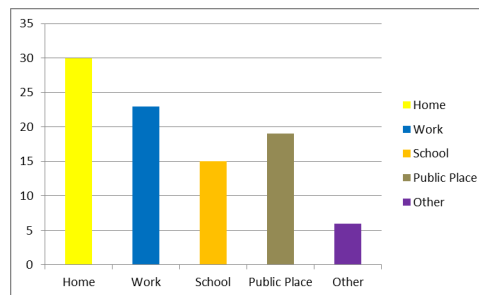


Figure 6.1: Location of email access among participants

The majority of the participants we interviewed had multiple email addresses. The most widely used email address was Google Mail, although that result may be biased, since Google Mail is a requirement to participate in the study. Among our participants, the second most used email was Yahoo Mail. Figure 6.2 shows the breakdown of the most popular Internet based emails among our participants.

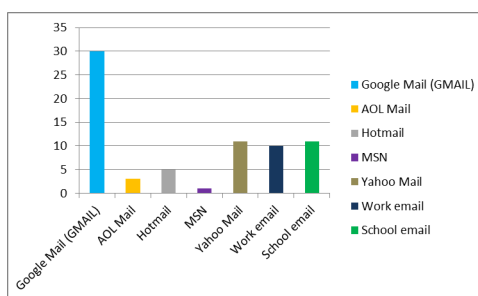


Figure 6.2: Different email accounts among participants

## 6.2 Surveys

All of the surveys were given through SurveyMonkey<sup>2</sup>, an online survey service that uses advanced technology for Internet security to protect user data and responses. We developed three surveys that were presented to the participants at different stages during the experiment. The initial survey consisted of 55 questions, with a mix of yes or no, multiple choice, Likert scale and short answers. The initial survey consisted of demographic, password, and image protection behavior related questions. The second survey consisted of 21 questions, with a mix of yes or no, multiple choice, Likert scale, and short answers. The second survey consisted of both the use of Cloudsweeper and specific password related questions. The third survey consisted of 23 questions, with a mix of yes or no, multiple choice, Likert scale, and short answers. The third survey consisted of both the use of the Gmail Image Extractor and specific image related questions. The surveys can be found at the following URL's.<sup>3 4 5</sup>

## 6.3 Password scan via Cloudsweeper

Cloudsweeper is designed to withstand full compromise of a user's cloud service credentials while presenting a straightforward interface [20]. The primary goal of Cloudsweeper is to protect the user in the event of a complete account compromise [20]. The implementation of Cloudsweeper is comprised of three components: a user interaction and authentication front end, a message translation

<sup>2</sup><https://www.surveymonkey.com>

<sup>3</sup><https://www.surveymonkey.com/s/CloudsweeperInitial1>

<sup>4</sup><https://www.surveymonkey.com/s/CloudsweeperPasswords>

<sup>5</sup><https://www.surveymonkey.com/s/CloudsweeperImages>

and encryption engine, and a cloud storage communication back end [20]. The main contribution of Cloudsweeper is a data-centric strategy for protecting high value information. For more details regarding the design and implementation of Cloudsweeper the reader is urged to reference Snyder and Kanich [20]. Cloudsweeper can be found at: <https://cloudsweeper.cs.uic.edu/>.

As shown in Figure 6.3, the user is presented with a webpage that essentially gives them three different options 1) Account Theft Audit 2) Cleartext Password Audit, and 3) Decrypt Messages.



Figure 6.3: Cloudsweeper interface

The account theft audit tool can help a user understand just how much a cybercriminal could access were they to take over the users email account. The tool will scan a Gmail account and give the user a visualization of the accounts hackers could take over if they get access to the users email account. Figure 6.4 shows a scan of a Google Mail account first established in 2005.

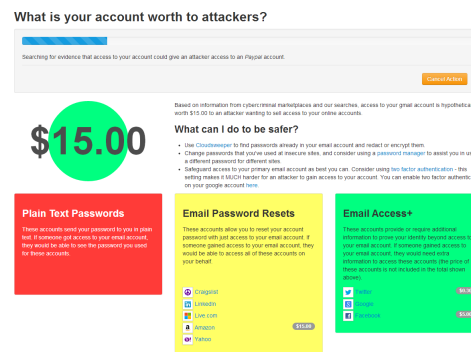


Figure 6.4: Cloudsweeper account theft audit

The cleartext password audit protects the user in the case of an attacker who gains access to their account, an attacker who can eavesdrop on the Internet session while reading email, or even someone who borrows a laptop without permission and begins to dig around [20]. Figure 6.5 shows the Cloudsweeper system programmatically looking through an email of one of the authors, to find

plain text passwords in the same way a hacker or spy might. The user is then presented with a list of found passwords that they can, optionally, redact from the account or encrypt [20].

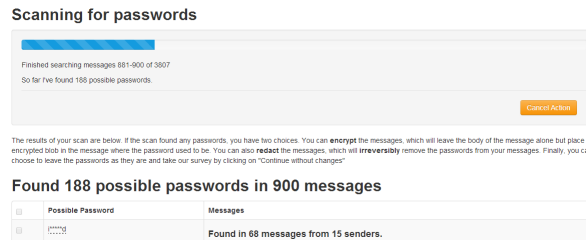


Figure 6.5: Cloudsweeper cleartext password audit

## 6.4 Image scan via Gmail Image Extractor

The Gmail Image Extractor (GIE) works by first connecting to the users Gmail account using IMAP over SSL. The users Gmail account is then searched for any message, sent or received, that includes an attachment. This is done by using the Gmail extended search IMAP extension (the X-GM-RAW command with the argument has: attachment). The returned collection of messages is downloaded in full from Gmail, 10 messages at a time. Each message attachment is extracted from the body of the message and checked to see if its MIME type is that of a well-known image type (image/jpeg, image/png, or image/gif). Each found image is then saved to the users disk, and an internal mapping is made between each saved image and the email message that contained the image.

Once all image attachments have been saved to disk, the user is prompted to delete any images they wish to remove from their local disk. The GIE then finds which images have been deleted from the disk, and re-downloads their corresponding message from Gmail. The local copy of the re-downloaded message is then edited to remove the relevant attachment. The GIE then replaces the original version of the message in the users Gmail account with the newly edited version.

Since the IMAP protocol does not directly support editing of message bodies on the server, the GIE achieves the same effect through the following steps: 1) The original message is deleted from Gmail. 2) The edited message is inserted into the user's Gmail account. 3) The edited messages metadata is updated to match the metadata of the original message (both the standard IMAP flags and Gmails label data). This process is repeated for every message that contains an attachment



deleted by the user, until the total set of image attachments contained in the users Gmail account is reduced to match the set of images remaining on the users local disk.

## **6.5 Approach**

Each of the 30 participants from Craigslist arrived on the campus of George Mason University and was taken to a private and secured office room. They were given access to a laptop with the aforementioned software installed and the links to the surveys pre-loaded. The software was loaded using a VMWare virtual machine. Thus, before and after each participant arrived we used the snapshot feature to roll-back to a known good state. This eliminated any possibility of residual information being compromised between participants.

The participant's first task was to complete the preliminary demographics-based survey. Next, they completed a second survey associated with their password protection behavior. Upon completion of this survey, the participants ran the Cloudsweeper application twice. The first time Cloudsweeper scanned their Google Mail accounts for cleartext passwords. The second time it reported on the monetary value of the aggregate cleartext passwords in their Google Mail account. At this point the participants could have redacted and encrypted their passwords.

Next, the participants completed a third survey associated with how they protect and share their images. Upon completing this survey, the participants ran the GIE that was designed to extract images that were found in their Google Mail accounts. At this point, the participants could have deleted their images if they so desired. Lastly, the participants completed a brief exit interview where we discussed the results of the Cloudsweeper and Gmail Image Extractor scans and discussed related questions that were not found on any of the surveys.

## **6.6 Passwords**

Figure 6.6 shows how users generally store, save, and remember their passwords. Of note is the fact that many users rely on memorization which has been shown to be a very weak method of securing

one’s password. Similarly, most users do not use a dedicated password manager such as KeePass.<sup>6</sup> From our exit interview discussions, the most cited reason for not relying on technology such as password databases or even two-factor authentication are the technical challenges and the learning curve. In regards to the password database, some users were concerned that if the password to their database vault was cracked, all of their passwords would be compromised.

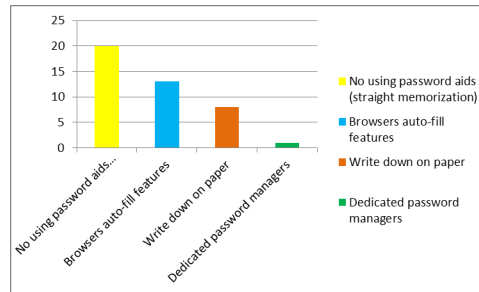


Figure 6.6: Password memorization tools and techniques

Of interest, is that 47% of our participant pool admitted to sharing their password with another person. This shows that even well-known security measures (e.g., not sharing your password) are sometimes not followed. The rationale for sharing the password with another person was often for convenience or for someone to “quickly” do “something” on their behalf. Additionally, many participants admitted to sharing their password with a parent. However, when asked just a few minutes later, who they thought were the most likely person they know to be a security risk, many mentioned their parents, the same ones that have their password.

Of the participants that we interviewed, 70% mentioned that they do not intentionally email passwords to themselves. On average, the participants answered that they currently had 6 passwords, with the minimum being 1, and the maximum being 25. However, when asked how many passwords did Cloudsweeper find in your mailbox that you didn’t realize were there, the results ranged from 0 to 272 passwords.<sup>7</sup>

Recall Cloudsweeper gives the option for the participants to redact and encrypt their passwords. Given the sheer amount of cleartext passwords that the participants on average found, we hypothesized the participants would take advantage of the opportunity to protect themselves. However, we

<sup>6</sup><http://keepass.info/>

<sup>7</sup>Some of these were likely password reset emails or other passwords the participants received in cleartext.

were quite wrong, as 67% of our participants did not redact any passwords. Of those who did redact passwords, Figure 6.7 shows a breakdown of the types of websites for which they redacted passwords. Furthermore, of the messages that were redacted, 55% were messages that the participant stated that they deliberately sent, 27% were messages that were automated, and 46% were both.

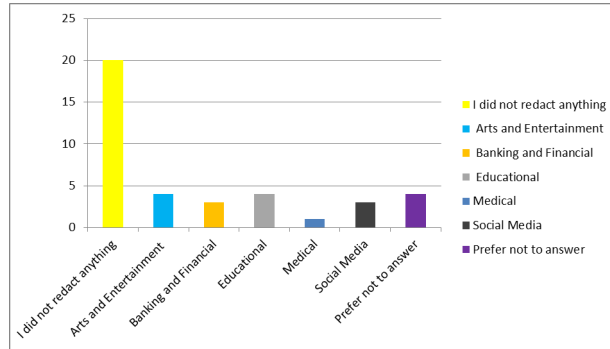


Figure 6.7: Frequency of participants who decided to redact their passwords

Regarding encrypted passwords, our results show that 60% of participants did not encrypt any passwords. Of those that did encrypt passwords, Figure 6.8 shows what websites the participants encrypted passwords for. Of the messages the participants encrypted, 18% were messages deliberately sent, 18% were messages that were automated, and 63% were both.

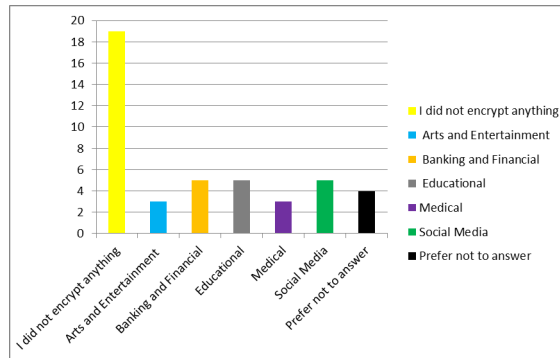


Figure 6.8: Frequency of participants who decided to encrypt their passwords

## 6.7 Images

Now we turn our attention to images, 53% of the participants that had images saved in their email, that they did not realize were there, and that they would not have wanted made public. Of those that

stated they saw images that they did not want made public, there was a wide frequency regarding who recorded the image in question. Figure 6.9 shows a breakdown of who took the images that the participant stated they did not want made public.

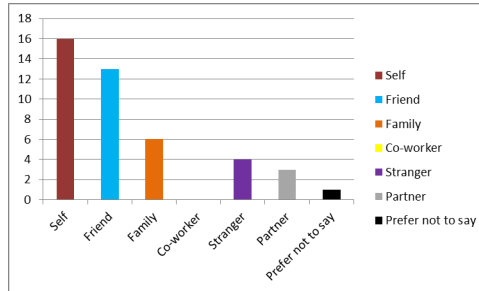


Figure 6.9: Breakdown of who took the images not meant for the public

When the participant was asked who the images were primarily of that they did not want made public, the majority said self, followed by friend, and family. Figure 6.10 shows the intended audience for the images that were not meant for public viewing.

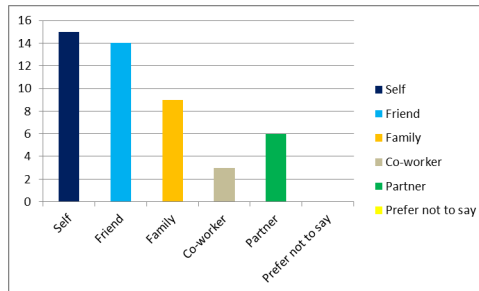


Figure 6.10: Intended audience for the image not meant for the public

The results show that 46% of the participants stated that they did upload the images that they did not want made public. This helps support the conclusion that at least some percentage of the participants care more about convenience and the fun factor of sharing, uploading, and being involved with social media websites as opposed to protecting the privacy and security of their personal images. Specifically, the participants stated that they uploaded them to various different websites. Figure 6.11, displays the destination website (location) of the image not meant for public sharing. The majority were sent to Google Mail and to a personal device such as a computer, phone, and tablet.

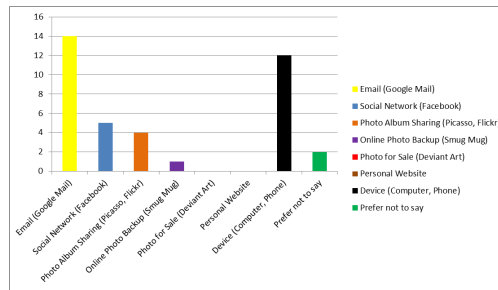


Figure 6.11: Destination of image not meant for the public

The participants responded that they thought the most likely threat (source) of images being uploaded, viewed, and shared with the public was self-inflicted through unintentional or accidental means. Figure 6.12 shows the breakdown of the most concerning adversary according to our participants.

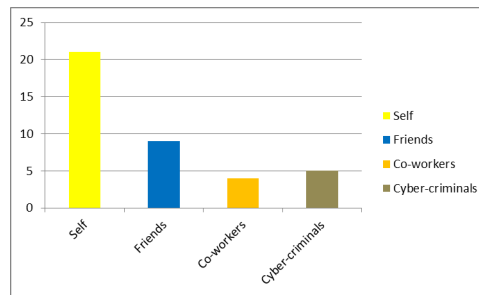


Figure 6.12: Threat (adversary) responsible for sharing and uploading images

The most cited consequence regarding if the image was made public was embarrassment (95%), followed by incrimination (22%), and then extortion (18%). As was the case with the passwords counterparts and in light of these consequences, the majority (57%) of the participants did not delete the images when given the opportunity.

In light of the consequences, we found that most participants rely on less-technical means such as asking someone not to share or deleting. Only a small percentage of the participants mentioned that they redact or use a technology such as snapchat. Figure 6.13, shows the frequency and type of image protection most often used by the participants.

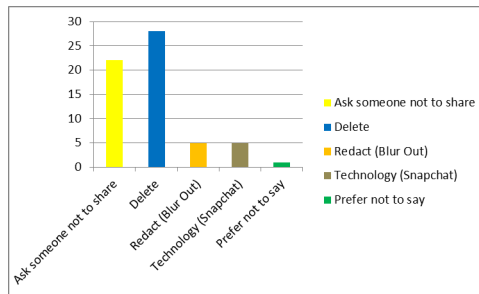


Figure 6.13: Types of methods used by participants to protect their images from being uploaded and shared without their consent

## 6.8 Exit interview

Near the conclusion of the study, each participant consented to an exit interview (discussion) where they were asked their thoughts on a variety of different security and privacy topics associated with the management of passwords and images. Table 6.2 below shows the questions that were asked, along with a categorical answer, and the percentage of those participants who answered in that manner. Figure 6.14 shows a “mind map” that gives an example of the types of responses that we received.

Table 6.2: Responses from exit interview

Question	Answer Type	Participant	Percentage
Have you ever read an email providers privacy policy?	Yes	7	23%
	No	23	77%
Have you been or do you know anyone that has ever been a victim of identity theft?	Yes, it was me	6	20%
	Yes, it was someone else I know	11	37%
	No	13	43%
If your account were hacked, do you think they could reset passwords to other accounts and answer security questions?	Yes, the majority of them	10	33%
	Yes, but just some of them	7	23%
	No	8	27%
	I don't know	5	17%
If you could visually see the passwords that would be known to a cyber-criminal would that make a difference to you?	Yes, a lot of difference	7	23%
	Yes, some difference	10	33%
	No	13	44%
Can you describe an example of an image that you did not want made public?	Personal and embarrassing	12	40%
	Picture of someone else	6	20%
	Sensitive document	6	20%
	All were publicly acceptable	11	37%
Do you feel comfortable sharing images that you don't want made public? If so, with whom?	Yes, with everyone	4	13%
	Yes, only family and friends	9	30%
	No	7	23%
	It depends	10	33%
Who do you see as the primary threat (sources) when it comes to acquiring and uploading your passwords and images?	Cyber-criminals	18	60%
	Someone I know or that knows me	4	13%
	Government	1	3%
	Self (accidental/unintentional)	9	30%
What consequences do you see if an unintended photo was made public?	Embarrassment	15	50%
	Loss of relationship	2	7%
	Loss of job	1	3%
	None	4	13%
Have you or anyone you know been a victim of image stealing, uploading, or extortion?	Yes, it was me	2	7%
	Yes, it was someone else I know	6	20%
	No	22	73%
What do you think can be done (legally) if someone steals or misuses your photos?	Contact the website	6	20%
	Contact the adversary (if known)	7	23%
	Contact Legal Counsel	6	20%
	Nothing	11	37%

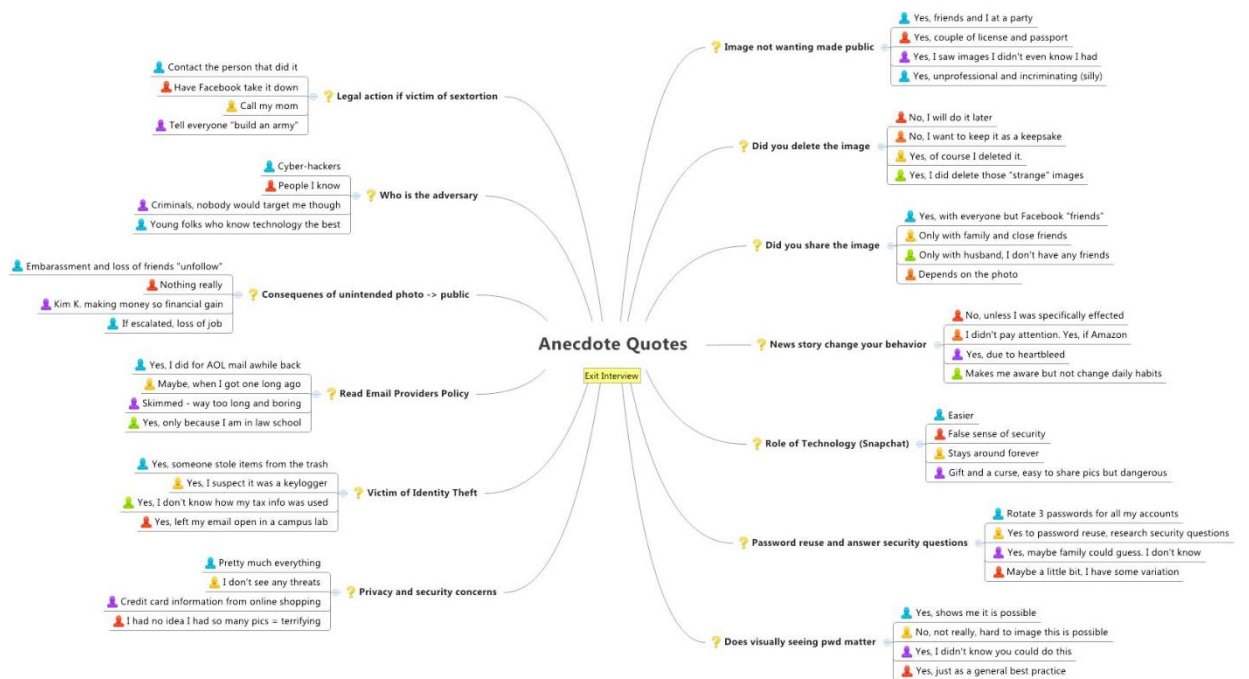


Figure 6.14: Exit interview anecdote quotes



## **Chapter 7: Defense Mechanisms**

In this section, we describe a variety of defense mechanisms that attempt to mitigate the actions of the adversary as described in each of our experiments. Given that many of our experiments were related, it should not come as a surprise that many of the defense mechanisms we offer will overlap. However, in an effort to organize the defense mechanisms, we tried to categorize them based on the experiment where applicable.

### **7.1 Profiling and tracking related defenses**

What can be done to help prevent against profiling and tracking of website visitors? In this section, we attempt to answer this question by proposing that website visitors alter the features they use to access websites. For example, if a visitor prefers Internet Explorer version 8, it may behoove them to occasionally use Firefox. Next, we suggest being cognizant about what private data the website visitor is providing to the website. Similarly, a website visitor should be aware of any security issues with the website ranging from a mismanaged certificate to websites that have tracking enabled.

#### **7.1.1 Privacy preserving technologies (Anonymization)**

There needs to be better defenses associated with privacy preserving technology users. A privacy preserving system is designed to protect a user's identity while they are using a computer system that can access the Internet. In general, it is a solution that attempts to make activity on the Internet untraceable with three broad goals [14]. In one popular scheme, anonymized users are mixed with "regular" worldwide consumers. The behaviors of all users (both anonymized and regular) are mixed together in an effort to try and prevent behavioral patterns from being noticed. The regular users are diverting attention away from the anonymized users by virtue of the search queries and traffic that they are generating.

However, given that anonymization will, at best, mask only a website visitor's IP address, location, and potentially the corresponding ISP, this is not a complete solution. This is because anonymization may not be able to defend against the geolocation threats addressed earlier because the majority of anonymization schemes such as Anonymizer utilize an IP-rotator scheme with many customers using the same IP address. If enough visitors over a long enough time continuously search for the same topics related to their true locale, Google ignores all of the IP address assumptions that were made as part of the anonymization scheme. Google basically assumes that the data associated with the search results are the most accurate and updates the location that they give away in *Google Location*. Since all users in a given block of IP addresses, say 95% of the time, are always looking for information about this locale, it is a very good assumption that the users are located there regardless of what the other IP addresses and anonymization think. As was pointed out by Eckersley, sometimes technologies intended to enhance user privacy make fingerprinting easier [5]. Therefore, it is imperative to find additional defenses that are less effected by this paradox.

The Google Location feature has a major impact on operation security associated with visitors even if they are using a PPT such as Anonymizer or Tor. Google has massive amounts of technology and tools to be able to analyze traffic and find correlations between the user's search criteria and the user's likely locations. As previously mentioned, visitors forget or are unaware that they are performing a search that should not be done while on a PPT. As a result, Google is often able to correlate these searches and show the user's real location.

Now suppose that the anonymized IP address is 12.34.56.78 and it shows that the location is anonymized to be in Madrid, Spain. Now let us assume that a few visitors make the mistake searching for local restaurants in their real home city of Denver, CO. Google is then able to identify this behavior and given enough searches and traffic will assume with a high likelihood that this IP address is not associated with Madrid but instead with that of Denver. The visitor's behavior is being used by Google to find the "most likely" location from which this visitor and corresponding IP address is coming [81]. Given enough of this data, the Google Locator actually becomes quite accurate thus reducing the effectiveness of the privacy preserving technology [81].

We recommend that visitors complete all of their searches using Google with a different country

code. For example, a visitor could utilize: `www.google.com.gh` whereby the “gh” is the country code for Ghana. Still another possible defense would be to set the Google Location to be a random place in the country. This concept is commonly referred to as “artificial pinning” and is just another small measure that could potentially help to prevent the identity of that particular IP address from being known. At the time of this writing, it was not possible to change the location from outside of the United States assuming that you were using `www.google.com(.en)` as your main search engine. Also, in some cases it is not even possible to change this setting or it may not appear for a plethora of reasons that are outside the scope of this research.

### 7.1.2 Privacy enhancing browser extensions

There exists a plethora of privacy enhancing browser extensions which can also be used to hide some of the information and features leaked by the visitor. Below, we have presented a few of the more popular and useful extensions we know.

1. **Adblock Plus:** Provides a means for removing online advertising and blocking well known malware domains.<sup>1</sup>
2. **BetterPrivacy:** Attempts to protect against lasting cookies that are hard to delete.<sup>2</sup>
3. **Lightbeam formerly Collusion:** Allows user to see all of the third parties that are tracking their movement across the web in real-time.<sup>3</sup>
4. **Garlik:** Service that carries out frequent scans of a wide range of sources for evidence of misuse.<sup>4</sup>
5. **NoScript:** Blocks JavaScript, Java, Flash, and other plugins except for trusted sites.<sup>5</sup>

---

<sup>1</sup><https://adblockplus.org>

<sup>2</sup><https://addons.mozilla.org/en-US/firefox/addon/betterprivacy/>

<sup>3</sup><http://www.mozilla.org/en-US/lightbeam/>

<sup>4</sup>[www.garlik.com](http://www.garlik.com)

<sup>5</sup><https://addons.mozilla.org/en-US/firefox/addon/noscript/>

In regards to the OSN (e.g., Facebook Application) attack phase, there exists several general defense mechanisms that the CSPPTO user could utilize to help prevent against the Facebook application attack [12]. First, the visitor should be cognizant of the information that they place on their Facebook profile. Several papers including [46], [88], [89], and [90] have described the privacy concerns associated with Facebook and a few defenses we are aware of are discussed during the rest of this section.

Luo et al. present an architecture known as “FaceCloak” that enforces user privacy on OSN by shielding a user’s personal information from the site and from other users that were not explicitly authorized by the user [91]. In the context of our attacks, even if the victim allows the malicious application, at least their online social fingerprint could be shielded.

## **7.2 Security awareness training defenses**

### **7.2.1 Outbrief**

In this section, we discuss the security awareness training that we offered to all of the CSPPTO participants. First, the training began by explaining the experiment conducted and described in Section 2.3. Specifically, we discussed the results of the experiment and the importance of attending and participating in the security awareness training that we developed. We began the security awareness training by explaining to the CSPPTO users the dangers and drawbacks of performing personal and attributable search websites on the Internet. In this part of the security awareness training, we covered the fact that search engines such as Google are often able to correlate search queries with a user’s true location. We provided a list of attributable search queries and websites that were verboten. This included the previously discussed categories, such as personal email, personal banking, and online social networks, etc. In Figure 7.1 we see a noticeable decrease in the amount of network traffic to attributable websites.

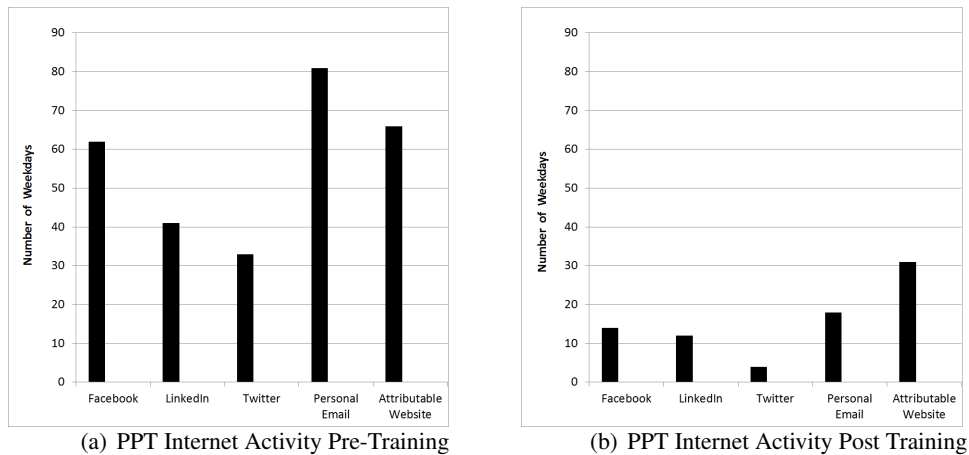


Figure 7.1: Security awareness training displayed as bar graph

## 7.2.2 Phishing awareness

Next, we shift our attention to the phishing aspect of the experiment as described in Section 2.3. We explained to the users the importance of recognizing a phishing attempt as the first layer of defense in defending against a phishing attack. Given the significant number of users that still clicked on the phishing link, this suggested to us that greater awareness and educational seminars are indeed warranted. To this end, we introduced the Simple Phishing Toolkit and demonstrated the ease with which adversaries can create legitimate looking phishing emails. Furthermore, we led an open discussion to explain what the obvious signs were that the email, websites, or both were forged. The underlying theme of our security awareness training was that when it comes to phishing, the educated user is often in the best position to protect against an adversary who implements a phishing attack.

We defended our phishing exercise by suggesting that sending periodic internal phishing emails is a good way to increase awareness from employees and to improve the overall security posture of the CSPPTO. We predict that if enough phishing campaigns are sent over a long period of time, the users will eventually learn to combat phishing attacks. The security awareness training devoted to improving susceptibility to phishing attacks was beneficial for the participants. At the conclusion of the training, we polled the attendees and they claimed a definite increase in understanding and awareness of phishing. However, we wanted to confirm whether our training was truly a success. To

accomplish this, we continued to monitor the network traffic of the CSPPTO users with the objective being to compare the true actions of the users with the responses they indicated on the survey.

Aside from the previously described security awareness training, there exists several defense mechanisms that the CSPPTO users could implement. The users should be vigilant enough not to access and allow the unfamiliar Facebook application in the first place. One defense mechanism to consider is “SafeBook” which was specifically designed to prevent privacy violations by an adversary, malicious users, and OSN providers [92]. Essentially, “SafeBook” is mainly characterized by a decentralized architecture relying on cooperation among peers in order to prevent potential privacy violations due to a centralized architecture [92]. Sirivianos et al. created a system called “FaceTrust” that verifies online personas in an efficient manner [93]. This particular system was designed to assess the credibility of statements made by online users.

### **7.3 Cyber crime and survey scam defenses**

In light of the growing epidemic of survey scams, we offer a few defenses that may alleviate some of the concerns we outline. Virtually all Internet scams make use of a domain name resolution as a critical part of their execution [94]. For example, a spam email containing the URL <http://toppills.com/> can only be monetized if the user both clicks on the link and the domain name can be correctly resolved to the Web site being advertised [94]. Domain names are central to attracting Web traffic and thus are used in large-scale Internet scams as well [94]. Email spam, blog spam, search spam, phishing, drive-by downloads, click frauds, and so on all generally require their victims to resolve domain names provided by scams [94].

Liu et al. conclude that local interventions on a registry level are likely to be ineffective when it comes to domain registration [94]. To have an impact on spammer’s domain registration, these interventions have to be extended to a global scale by ICANN [94]. Establishing stricter registration policies (e.g., by requiring a photo ID) on a global scale would raise the cost and leave less options to scammers according to Liu et al. [94].

Of note, is that Facebook and the state of Washington have filed separate lawsuits against West Virginia-based Adscend Media LLC, alleging the company was responsible for spreading malware

through Facebook and for stealing personal information from users of the social networking site [95]. The Attorney General’s office included a series of screen shots demonstrating the alleged infractions starting with a Facebook post that attempts to lure users to click on a link to see an erotic video [95]. Clicking the link takes the user to a non-Facebook page where they are asked to click a “Like” icon. The complaint describes the use of content locking and suggests that affiliates were paid each time the user took one of the surveys on the bait pages [95].

According to an Internet security specialist<sup>6</sup>, users can protect themselves from alleged scams such as this by installing browser add-ons such as the no-script plug in but unfortunately there are no scripts that can stop a user from entering personal details in a form as a result of social engineering [95]. A potential defense is to demonetize the spam ecosystem by intervening on these deceptive ad networks that are sponsoring the majority of Facebook spam seen in our study.

## **7.4 Protection of passwords and images on the cloud**

Given the seemingly everyday news stories about recent hacks, information leaks, and privacy related consequences it is imperative that the general public become knowledgeable about how to best protect their passwords and images. We recommend that users master basic security and password management practices such as eliminating password reuse, using a strong password, and not sharing passwords with anyone. Additionally, through the use of Cloudsweeper, users can get snapshot of the passwords that are vulnerable to the types of attacks that we discussed in this dissertation. Similarly, when applicable, users should encrypt and redact their passwords to effectively reduce their attack space. It is recommended that users incorporate password databases to help manage their passwords.

Regarding images, it is imperative that users use caution when sharing and uploading pictures. The best protection may be to reduce or eliminate taking compromising photos, however, this is usually unrealistic. Therefore, we have introduced the Google Image Extractor, that gives the user a chance to easily view all of their photos they have within their Google Mail account. One of the main benefits of the Google Image Extractor is the ability for a user to immediately view and then

---

<sup>6</sup>Marcel van den Berg of Team Cymru, a nonprofit that specializes in Internet security.

delete any images that they don't wish to have any longer. We recognize that the Google Image Extactor has its limitations, especially, seeing as though it currently only works for Google Mail. In general, we once again highly recommend security awareness when it comes to protecting ones passwords and images.



## Chapter 8: Conclusions

Through our research and experiments to date, we have highlighted the ease which adversaries and websites can be correlated with the aim of exposing the online persona of a user accessing the Internet. Moreover, we have demonstrated how common features can be employed to profile the majority of a fixed set of website visitors using our defined threat model. In our extended experiments, we explored the ease in which an adversary can create a custom webpage, Facebook Fan Page, and a malicious Facebook application that has tracking enabled.

In addition, we have analyzed existing attacks and introduced new methods of measuring the information leaks across websites. We also provided scenarios on how website owners can utilize their visitor's personas to change website content. One requirement of the profiling and tracking attacks from an adversary's point of view is to lure the visitor to access the personal websites using phishing and social engineering techniques. The adversary simply has to take the time to create and setup the analytics. From there, all that is required is a phishing attack which has been shown to be quite successful.

Additionally, we have investigated how cyber-criminals have implemented survey scams to take advantage of the growing number of users joining online social networks. In particular, we examined a continuous feed of spam URLs. The URLs were crawled and analyzed to determine the different affiliate programs and the offers they were executing on unsuspecting users.

We presented an empirical study of Facebook spam revealing that 73% of the working spam URLs in our Facebook spam feed were monetized via survey scams sponsored by ad networks. Based on our analysis of 129 unique spam URLs, over 50% of these URLs were traced back to four ad networks: Amung.us, CPAlead, ClickBanner, and LifeStreet Media. Furthermore, we infiltrated each of these affiliates to gather first-hand intelligence on their operations and means to monetize online social network users. We presented a carbon dating method that can estimate the age of a spammer's affiliate ID and showed they are on average nine months old. We showed how prolific

this problem is by measuring the amount of spam that originates in a (test) user's email account and Facebook wall. Our results provide a potential point to demonetize the spam ecosystem by intervening on these deceptive ad networks that are sponsoring the majority of Facebook spam seen in our study.

The research we conducted associated with *understanding users' perceptions of privacy and value of information stored in their accounts* makes several contributions. To our knowledge this is the first study that investigated how a subset of users from the general public behave if the threat of mishandling the protection of their passwords and images was made "visually" apparent to them. We found that the majority of participants seemed to value convenience over privacy when it came to how they were storing their passwords on the cloud (e.g., Google Mail). Similarly, we showed that many participants accepted the risk of leaving their images vulnerable to being uploaded and shared with unintended entities. Our study provided a glimpse into what participants from the general public view as the likely adversary and the consequences that can arise.

Next, we used a previously developed application called Cloudsweeper, which allows a person to scan their Google Mail in search of cleartext passwords. The Cloudsweeper application provides an opportunity to redact or encrypt the passwords to improve security. Furthermore, the Cloudsweeper application has the ability to place a monetary value on the passwords of our participants. This is the first time the Cloudsweeper application was specifically used to measure how participants behave if they visually see the threat caused by their poor password security behavior.

We designed a custom application called the Gmail Image Extractor, which allows a person to quickly and safely scan their Google Mail account for images. The images are then saved to a local folder where the person has the opportunity to delete them which in-turn automatically deletes them from their Google Mail. We feel this contribution will be quite useful to the general public because it will give them an efficient way to view all of their photos and remove those that are no longer needed and could potentially cause a threat to the person. We hope to collaborate with Google to allow for our software to be built-in or as an add-on option to the existing Google Mail interface.

Our study combined a set of three surveys, the use of two custom applications namely Cloudsweeper and the Gmail Image Extractor, and an on-site exit interview to better understand the participants

behavior and concerns when it comes to protecting passwords and images. We conclude that there is a well-known risk to people storing their passwords and images in the clear. While this is unlikely to stop, we believe our analysis along with the custom applications we developed can better protect users from being victims of the threats we have outlined.

The main conclusion from this research is the plethora of different ways a user can disclose their identity when accessing websites. Some of them are well known like protecting your social security number, limiting the amount of data placed on social networking sites, and limiting use of cookies. However, we show commonly used features such as the operating system, browser, search engine, the time one spends on a website, and other features can be used to profile the majority of website visitors. The users were not educated about these vectors for violating user privacy.

The public and specifically those users on OSNs have to be aware that their personal information provided in their profile is of interest to a wide variety of adversaries. These adversaries will do whatever it takes to acquire private information of OSN users.

## **8.1 Experimental limitations**

The experiments we designed are not without their limitations. This section serves to explain some limitations as well as shed some light on why we designed the experiment in the manner that we did. The major limitation of our work is the low number of participants that we had for the majority of the experiments. We had a difficult time gathering participants for these assignments. In the future, we will plan for a better recruitment process. A related limitation is that the majority of participants were of the same “type” meaning they were presumably young, educated, and enrolled as college students.

Next, we had no real way to monitor the participants to ensure that they were completing the assignments accurately and truthfully. Specifically, our critics might argue that we simply just created a “new” problem if the participants utilized fake or fictitious profiles and claimed them to be their own. While the creation of fake Facebook accounts goes against Facebook policy, it is a well known fact that fake Facebook accounts do exist.

The biggest limitation of our survey scam experiment and analysis was that our data feed came

from one source. Granted the source had different types of spam and scams, they were all of the same type namely social networking (e.g., Facebook) related. This inherent limitation of our study is the fact that our spam feed might contain a bias based on the algorithm used to detect spam posting and the user base that has installed MyPageKeeper.

Another major limitation of our study was the amount of manual work that was required. The manual effort was not only time consuming but also potentially error prone. We tried to alleviate this concern by verifying our work and analyzing just a few hundred URLs on a given day. This research study would definitely benefit from more automation. Another difficulty that we encountered was the ability to infiltrate the affiliate networks. This was especially troublesome with respect to the affiliates that were hosted in foreign (e.g., non-US countries). We tried our best to join as members, but certain affiliates (e.g., Clickbank.gr) simply would not allow us in. Additionally, our carbon dating method can only provide an estimate of an affiliate account's age.

Another limitation with respect to the user perception and behavior study is that it only considered participants with Google Mail accounts. There exists research that shows the sophistication level of the participants based on their preferred choice of web mail. Additionally, the experiment only lasted for three months between April and June 2014. A longer study with more participants may yield different results. Furthermore, there are biases based on the survey responses themselves. For example, it is often the case that there is over-reporting when it comes to technical and computer related savviness.

## **8.2 Future research**

In future work, we hope to determine the extent to which a multi-user system with multiple users logged in can be used to profile and track visitors. One of the main objectives of future work would be to complete a more rigorous statistical and mathematical treatment of our data. For example, it would be useful to see exactly what percentage of visitors utilized a certain combination of features and how that statistically compared to other visitors' records. We could potentially investigate and determine the easiest and most difficult visitors to profile and track. An overall increased level of statistical analysis would benefit subsequent iterations of this experiment.

Furthermore, we would like to develop an automated algorithm which makes use of the entropy calculations in order to correlate re-identify users. This algorithm could be evaluated on a large scale. In general, our experiments were on a limited scale and so future research would call for a larger scale survey. Similarly, the associated experiments with breaching a PPT were small and so extending our methodology to consider different anonymization schemes and different applications would be warranted. The survey scam experiment could certainly be improved with more automation.

We plan to focus on developing the Cloudsweeper and Google Mail Image Extractor to support other email services such as Yahoo Mail and Outlook (Exchange). We feel that our custom applications can potentially be used by organizations to help protect their overall security posture.

The highest priority for future research is to actually take steps to combat the survey scam threat by targeting the affiliate programs where it hurts them the most—namely in their wallet. Future research could provide methods to improve the defenses offered throughout this dissertation and more importantly find ways to help educate the users on how to best defend themselves against the adversary described in our threat model. While not a revelation on its own, we once more showed that users and their activities remain the weakest link in any non-trivial security scheme. Therefore, we are convinced that security awareness training is of paramount importance and should be treated as such.

## Appendix A: Additional Data for Profiling and Tracking (Re-identification) Experiments

Table A.1: Operating System Usage

OS	Frequency	Entropy
Windows 7	15	1.05
Windows Vista	4	2.95
Windows XP	6	2.37
MAC OSX	5	2.63
Linux	1	4.95

Table A.2: Browser Usage

Browser	Frequency	Entropy
Firefox	11	1.49
Internet Explorer	1	1.49
Chrome	10	1.78

Table A.3: Search Engine Usage

Search Engine	Frequency	Entropy
Google	23	0.43
Yahoo	2	3.95
Bing	1	4.95
No Referrer	5	2.63

Table A.4: Search Resolution Usage

Screen Resolution	Frequency	Entropy
1366x768	9	1.78
1440x900	1	4.95
1920x1200	6	2.37
1280x800	7	2.15
1024x768	1	4.95
1600x900	2	3.95
1311x737	1	4.95
1280x1024	4	2.95

Table A.5: ISP Usage

ISP	Frequency	Entropy
Boston	1	4.95
Bresco	2	3.95
Charter	1	4.95
Clearwire	2	3.95
Click	1	4.95
Comcast	4	2.95
Cox	5	2.63
GMU	2	3.95
PSINet	2	3.95
PSU	2	3.95
Road Runner	2	3.95
Shadyside Hospital	1	4.95
Sunflower	2	3.95
Vanderbilt	1	4.95
Verizon	3	3.37



Table A.6: City Location Usage

Location (City)	Frequency	Entropy
Annandale	2	3.95
Apex	2	3.95
Ashburn	1	4.95
Baltimore	3	3.37
Boston	1	4.95
Bremerton	1	4.95
Burke	2	3.95
Columbus	1	4.95
Fairfax	3	3.37
Gaithersburg	1	4.95
Herndon	1	4.95
Houston	1	4.95
Lawrence	2	3.95
Los Angeles	2	3.95
Madison	1	4.95
Nashville	1	4.95
Pittsburgh	1	4.95
San Diego	1	4.95
Seattle	1	4.95
State College	2	3.95
Tacoma	1	4.95

## Bibliography

- [1] K. Thomas and D. M. Nicol, “The koobface botnet and the rise of social malware,” in *MALWARE 2010*, 2010.
- [2] R. C. P. <http://www.rogerclarke.com>.
- [3] K. Mowery, D. Bogenreif, S. Yilek, and H. Shacham, “Fingerprinting information in javascript implementations,” in *Proceedings of Web*, vol. 2.
- [4] A. P. <http://www.privacy.gov.au>.
- [5] P. Eckersley, “How unique is your web browser?” in *Privacy Enhancing Technologies*. Springer, 2010, pp. 1–18.
- [6] R. Atterer, M. Wnuk, and A. Schmidt, “Knowing the user’s every move: user activity tracking for website usability evaluation and implicit interaction,” in *Proceedings of the 15th international conference on World Wide Web*. ACM, 2006, pp. 203–212.
- [7] L. Sweeney, “k-anonymity: A model for protecting privacy,” *International Journal on Uncertainty Fuzziness and Knowledgebased Systems*, vol. 10, no. 5, pp. 557–570, 2002.
- [8] F. <http://www.facebook.com/press/info.php?statistics>.
- [9] G. Wondracek, T. Holz, E. Kirda, and C. Kruegel, “A practical attack to de-anonymize social network users,” in *Security and Privacy (SP), 2010 IEEE Symposium on*. IEEE, 2010, pp. 223–238.
- [10] R. Heatherly, M. Kantarcioglu, B. Thuraisingham, and J. Lindamood, “Preventing private information inference attacks on social networks,” 2009.

- [11] C. Warren and B. Laslett, "Privacy and secrecy: A conceptual comparison," *Journal of Social Issues*, vol. 33, no. 3, pp. 43–51, 1977.
- [12] J. W. Clark, "Correlating a persona to a person," in *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Conference on Social Computing (SocialCom)*. IEEE, 2012, pp. 851–859.
- [13] D. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM*, vol. 24, no. 2, pp. 84–90, 1981.
- [14] R. Pang, M. Allman, V. Paxson, and J. Lee, "The devil and packet trace anonymization," *ACM SIGCOMM Computer Communication Review*, vol. 36, no. 1, pp. 29–38, 2006.
- [15] S. Coull, M. Collins, C. Wright, F. Monroe, and M. Reiter, "On web browsing privacy in anonymized netflows," in *Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium*. USENIX Association, 2007, p. 23.
- [16] A. Orebaugh, "Social media malware."
- [17] "Survey scams aimed at social network netizens. <http://about-threats.trendmicro.com>."
- [18] K. Levchenko, A. Pitsillidis *et al.*, "Click trajectories: End-to-end analysis of the spam value chain," in *IEEE Symposium on Security and Privacy*, ser. SP '11. Washington, DC, USA: IEEE Computer Society, 2011. [Online]. Available: <http://dx.doi.org/10.1109/SP.2011.24>
- [19] D. Florencio and C. Herley, "A large-scale study of web password habits," in *Proceedings of the 16th international conference on World Wide Web*. ACM, 2007, pp. 657–666.
- [20] P. Snyder and C. Kanich, "Cloudsweeper: enabling data-centric document management for secure cloud archives," in *Proceedings of the 2013 ACM workshop on Cloud computing security workshop*. ACM, 2013, pp. 47–54.
- [21] R. A. Javier, J. Dillon, C. DaBreo, and J. De Mucci, "Bullying and its consequences: In search of solutions part ii," *Journal of Social Distress and the Homeless*, vol. 22, no. 2, pp. 59–72, 2013.

- [22] L. E. Gomez-Martin, “Smartphone usage and the need for consumer privacy laws,” 2012.
- [23] B. Krishnamurthy, D. Malandrino, and C. Wills, “Measuring privacy loss and the impact of privacy protection in web browsing,” in *Proceedings of the 3rd symposium on Usable privacy and security*. ACM, 2007, pp. 52–63.
- [24] B. Krishnamurthy, “I know what you will do next summer,” *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 5, pp. 65–70, 2010.
- [25] A. Cooper, D. Mulligan, H. Schulzrinne, and E. Wilde, “Challenges for the location-aware web,” 2010.
- [26] X. Wang, S. Chen, and S. Jajodia, “Network flow watermarking attack on low-latency anonymous communication systems,” in *Security and Privacy, 2007. SP’07. IEEE Symposium on*. IEEE, 2007, pp. 116–130.
- [27] N. Hopper, E. Vasserman, and E. Chan-Tin, “How much anonymity does network latency leak?” *ACM Transactions on Information and System Security (TISSEC)*, vol. 13, no. 2, p. 13, 2010.
- [28] N. Schear and D. Nicol, “Performance analysis of real traffic carried with encrypted cover flows,” in *Proceedings of the 22nd Workshop on Principles of Advanced and Distributed Simulation*. IEEE Computer Society, 2008, pp. 80–87.
- [29] G. Bissias, M. Liberatore, D. Jensen, and B. Levine, “Privacy vulnerabilities in encrypted http streams,” in *Privacy Enhancing Technologies*. Springer, 2006, pp. 1–11.
- [30] M. Liberatore and B. Levine, “Inferring the source of encrypted http connections,” in *Proceedings of the 13th ACM conference on Computer and communications security*. ACM, 2006, pp. 255–263.
- [31] S. Coull, C. Wright, F. Monrose, M. Collins, M. Reiter *et al.*, “Playing devils advocate: Inferring sensitive information from anonymized network traces,” in *Proceedings of the Network and Distributed System Security Symposium*, 2007, pp. 35–47.

- [32] S. Coull, C. Wright, A. Keromytis, F. Monrose, and M. Reiter, “Taming the devil: Techniques for evaluating anonymized network data,” in *Network and Distributed System Security Symposium (NDSS)*, 2008.
- [33] J. Jung, A. Sheth, B. Greenstein, D. Wetherall, G. Maganis, and T. Kohno, “Privacy oracle: a system for finding application leaks with black box differential testing,” in *Proceedings of the 15th ACM conference on Computer and communications security*. ACM, 2008, pp. 279–288.
- [34] D. Scott and R. Sharp, “Abstracting application-level web security,” in *Proceedings of the 11th international conference on World Wide Web*. ACM, 2002, pp. 396–407.
- [35] D. Irani, S. Webb, K. Li, and C. Pu, “Large online social footprints—an emerging threat,” in *Computational Science and Engineering, 2009. CSE’09. International Conference on*, vol. 3. IEEE, 2009, pp. 271–276.
- [36] B. Krishnamurthy and C. Wills, “Generating a privacy footprint on the internet,” in *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*. ACM, 2006, pp. 65–70.
- [37] M. Egele, A. Moser, C. Kruegel, and E. Kirda, “Pox: Protecting users from malicious facebook applications,” in *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2011 IEEE International Conference on*. IEEE, 2011, pp. 288–294.
- [38] K. Singh, S. Bhola, and W. Lee, “xbook: Redesigning privacy control in social networking platforms,” in *Proceedings of the 18th conference on USENIX security symposium*. USENIX Association, 2009, pp. 249–266.
- [39] A. Acquisti, “Faces of facebook privacy in the age of augmented reality. <http://www.heinz.cmu.edu>.”
- [40] A. Acquisti and R. Gross, “Predicting social security numbers from public data,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 27, pp. 10 975–10 980, 2009.

- [41] B. Krishnamurthy and C. Wills, “On the leakage of personally identifiable information via online social networks,” in *Proceedings of the 2nd ACM workshop on Online social networks*. ACM, 2009, pp. 7–12.
- [42] A. Makridakis, E. Athanasopoulos, S. Antonatos, D. Antoniadis, S. Ioannidis, and E. Markatos, “Designing malicious applications in social networks,” In *IEEE Network Special Issue on Online Social Networks*, 2010.
- [43] T. Jagatic, N. Johnson, M. Jakobsson, and F. Menczer, “Social phishing,” *Communications of the ACM*, vol. 50, no. 10, pp. 94–100, 2007.
- [44] M. Balduzzi, C. Platzer, T. Holz, E. Kirda, D. Balzarotti, and C. Kruegel, “Abusing social networks for automated user profiling,” in *Recent Advances in Intrusion Detection*. Springer, 2011, pp. 422–441.
- [45] M. Chew, D. Balfanz, and B. Laurie, “(under) mining privacy in social networks,” 2008.
- [46] A. Acquisti and R. Gross, “Imagined communities: Awareness, information sharing, and privacy on the facebook,” in *Privacy Enhancing Technologies*. Springer, 2006, pp. 36–58.
- [47] R. Dhamija, J. Tygar, and M. Hearst, “Why phishing works,” in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2006, pp. 581–590.
- [48] G. Moody, D. Galletta, J. Walker, and B. Dunn, “Which phish get caught? an exploratory study of individual susceptibility to phishing,” 2011.
- [49] R. Dodge, K. Coronges, and E. Rovira, “Empirical benefits of training to phishing susceptibility,” *Information Security and Privacy Research*, pp. 457–464, 2012.
- [50] T. Holz, M. Engelberth, and F. Freiling, “Learning more about the underground economy: A case-study of keyloggers and dropzones,” *Computer Security–ESORICS 2009*, pp. 1–18, 2009.
- [51] J. Franklin, V. Paxson, A. Perrig, and S. Savage, “An inquiry into the nature and causes of the wealth of internet miscreants,” in *ACM Conference on Computer and Communications Security (CCS)*, 2007.

- [52] M. Motoyama, D. McCoy, K. Levchenko, S. Savage, and G. Voelker, “An analysis of underground forums,” in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*. ACM, 2011, pp. 71–80.
- [53] C. Kanich, N. Weaver, D. McCoy, T. Halvorson, C. Kreibich, K. Levchenko, V. Paxson, G. Voelker, and S. Savage, “Show me the money: characterizing spam-advertised revenue,” in *USENIX Security Symposium*, 2011.
- [54] B. Stone-Gross, T. Holz *et al.*, “The underground economy of spam: A botmaster’s perspective of coordinating large-scale spam campaigns,” in *USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)*, 2011.
- [55] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. Voelker, V. Paxson, and S. Savage, “Spamalytics: An empirical analysis of spam marketing conversion,” in *Proceedings of the 15th ACM conference on Computer and communications security*, 2008, pp. 3–14.
- [56] K. Thomas, C. Grier, D. Song, and V. Paxson, “Suspended accounts in retrospect: An analysis of twitter spam,” in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*. ACM, 2011, pp. 243–258.
- [57] R. Anderson, C. Barton, R. Böhme, R. Clayton, M. van Eeten, M. Levi, T. Moore, and S. Savage, “Measuring the cost of cybercrime,” 2012.
- [58] Y. Wang, M. Ma, Y. Niu, and H. Chen, “Spam double-funnel: Connecting web spammers with advertisers,” in *Proc. of the 16th International Conference World Wide Web (WWW)*, 2007.
- [59] C. Grier, L. Ballard, J. Caballero, N. Chachra, C. Dietrich, K. Levchenko, P. Mavrommatis, D. McCoy, A. Nappa, A. Pitsillidis *et al.*, “Manufacturing compromise: The emergence of exploit-as-a-service,” 2012.
- [60] P. Ferguson, “Observations on emerging threats,” in *Proceedings of the 5th USENIX conference on Large-Scale Exploits and Emergent Threats*. USENIX Association, 2012, pp. 4–4.

- [61] W. Odom, J. Zimmerman, and J. Forlizzi, “Teenagers and their virtual possessions: design opportunities and issues,” in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2011, pp. 1491–1500.
- [62] S. Egelman, “My profile is my password, verify me!: the privacy/convenience tradeoff of facebook connect,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 2369–2378.
- [63] A. Acquisti and J. Grossklags, “Privacy and rationality in individual decision making,” *Security & Privacy, IEEE*, vol. 3, no. 1, pp. 26–33, 2005.
- [64] S. Gaw and E. W. Felten, “Password management strategies for online accounts,” in *Proceedings of the second symposium on Usable privacy and security*. ACM, 2006, pp. 44–55.
- [65] B. Grawemeyer and H. Johnson, “Using and managing multiple passwords: A week to a view,” *Interacting with Computers*, vol. 23, no. 3, pp. 256–267, 2011.
- [66] E. Hayashi and J. Hong, “A diary study of password usage in daily life,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2011, pp. 2627–2630.
- [67] A. E. Howe, I. Ray, M. Roberts, M. Urbanska, and Z. Byrne, “The psychology of security for the home computer user,” in *Security and Privacy (SP), 2012 IEEE Symposium on*. IEEE, 2012, pp. 209–223.
- [68] R. Shay, S. Komanduri, P. G. Kelley, P. G. Leon, M. L. Mazurek, L. Bauer, N. Christin, and L. F. Cranor, “Encountering stronger password requirements: user attitudes and behaviors,” in *Proceedings of the Sixth Symposium on Usable Privacy and Security*. ACM, 2010, p. 2.
- [69] R. Gross and A. Acquisti, “Information revelation and privacy in online social networks,” in *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*. ACM, 2005, pp. 71–80.
- [70] E. Rader, R. Wash, and B. Brooks, “Stories as informal lessons about security,” in *Proceedings of the Eighth Symposium on Usable Privacy and Security*. ACM, 2012, p. 6.



- [71] M. Walrave, I. Vanwesenbeeck, and W. Heirman, "Connecting and protecting? comparing predictors of self-disclosure and privacy settings use between adolescents and adults." *Cyberpsychology*, vol. 6, no. 1, 2012.
- [72] S. e. a. Das, "The effect of social influence on security sensitivity," *Symposium on Usable Privacy and Security - SOUPS*, 2014.
- [73] S. Egelman, A. Sotirakopoulos, I. Muslukhov, K. Beznosov, and C. Herley, "Does my password go up to eleven?: the impact of password meters on password selection," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 2379–2388.
- [74] I. Ion, N. Sachdeva, P. Kumaraguru, and S. Čapkun, "Home is safer than the cloud!: privacy concerns for consumer cloud storage," in *Proceedings of the Seventh Symposium on Usable Privacy and Security*. ACM, 2011, p. 13.
- [75] R. Zhao and C. Yue, "All your browser-saved passwords could belong to us: A security analysis and a cloud-based new design," in *Proceedings of the third ACM conference on Data and application security and privacy*. ACM, 2013, pp. 333–340.
- [76] P. G. Kelley, J. Bresee, L. F. Cranor, and R. W. Reeder, "A nutrition label for privacy," in *Proceedings of the 5th Symposium on Usable Privacy and Security*. ACM, 2009, p. 4.
- [77] P. Klasnja, S. Consolvo, J. Jung, B. M. Greenstein, L. LeGrand, P. Powledge, and D. Wetherall, "When i am on wi-fi, i am fearless: privacy concerns & practices in eeryday wi-fi use," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009, pp. 1993–2002.
- [78] P. Eckersley, "A primer on information theory and privacy," *Electronic Frontier Foundation*, 2010.
- [79] J. W. Clark, "Everything but the kitchen sink: determining the effect of multiple attacks on privacy preserving technology users," in *Secure IT Systems*. Springer, 2012, pp. 199–214.

- [80] J. Clark, "Correlating a persona to a person," in *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Conference on Social Computing (SocialCom)*. IEEE, 2012, pp. 851–859.
- [81] J. Clark and A. Stavrou, "Breaching & protecting an anonymizing network system," in *6th Annual Symposium on Information Assurance*, 2011, p. 32.
- [82] J. W. Clark and D. McCoy, "There are no free ipads: An analysis of survey scams as a business," in *Presented as part of the 6th USENIX Workshop on Large-Scale Exploits and Emergent Threats*. ACM, 2013.
- [83] M. Rahman, T. Huang, H. Madhyastha, and M. Faloutsos, "Frappe: detecting malicious facebook applications," in *Proceedings of the 8th international conference on Emerging networking experiments and technologies*. ACM, 2012, pp. 313–324.
- [84] A. Pitsillidis, C. Kanich *et al.*, "Taster's choice: a comparative analysis of spam feeds," in *IMC '12*, 2012. [Online]. Available: <http://doi.acm.org/10.1145/2398776.2398821>
- [85] C. Kanich, N. Chachra *et al.*, "No plan survives contact: Experience with cybercrime measurement," *Proc. of 4th USENIX CSET*, 2011.
- [86] C. Kanich, N. Weaver *et al.*, "Show Me the Money: Characterizing Spam-advertised Revenue," in *Proceedings of the USENIX Security Symposium*, San Francisco, CA, August 2011.
- [87] "Public high school antics. <http://www.nakedsecurity.sophos.com>."
- [88] H. R. Lipford, A. Besmer, and J. Watson, "Understanding privacy settings in facebook with an audience view," in *Proceedings of the 1st Conference on Usability, Psychology, and Security*. USENIX Association Berkeley, CA, USA, 2008, pp. 1–8.
- [89] H. Jones and J. H. Soltren, "Facebook: Threats to privacy," *Project MAC: MIT Project on Mathematics and Computing*, vol. 1, 2005.



## **BIOGRAPHY**

Jason W. Clark graduated from Schalmont High School, Schenectady, New York, in 1997. He received his Bachelor of Science from Syracuse University in 2001. He received his Master of Science in Information Technology from Rensselaer Polytechnic Institute (RPI) in 2002. He received a second Master of Science in Computer Forensics from George Mason University in 2011. He is currently employed as a researcher at the Software Engineering Institute (SEI) part of Carnegie Mellon University out of the Arlington, Virginia office.