Abstract-- Online impersonation attacks occur on a daily basis and are a rising concern in the area of online privacy and security. Methods are analyzed in regard to their ability to detect impersonation behavior based on a user’s web browsing history information. Reviewing previous research on impersonation detection shows that web browsing history, in conjunction with other methods is a viable medium for verification. Impersonation attack evasion methods are discussed and recommendations are provided for securing online browsing information.
Internet Browsing Behavior Profile and its Application in Detecting Impersonation

An individual’s browser information stored on their computer can often reveal quite a bit about that person’s life. This data is stored in the form of cookies, cache, web pages accessed, and saved form information. If given even a cursory look into someone’s browsing history and online internet activity one can discover a wealth of information about that person including:

- Spending habits
- Geographical location
- Education history
- Work history
- Social Contacts
- Hobbies and interests
- Physical Appearance
- Age and date of birth

While access to some of this information can vary based on a user’s online profile and use of online services, for a huge number of people this information can be easily accessed and possibly manipulated by a motivated individual. In fact, current statistics put the number of Facebook accounts around 1 billion and the current number of twitter accounts around 500 million (Smith, 2013). This massive increase in use of online services has led to an incursion of a number of inevitable problems including spam, viruses, worms, phishing, and spyware. Many of these issues and problems are aided by, involved with, or used in conjunction with online impersonation.

While the average person may not be aware of the mechanics and issues behind impersonation, they likely encounter situations where it is happening every day in their online life. Receiving a fraudulent email from an otherwise reputable institution asking for your personal information, following a fake novelty twitter account modeled after a celebrity, and getting a Facebook message about a new product you should try are all examples of online impersonation. While some of these situations are used merely for entertainment and advertising purposes, the issue of online impersonation can cause serious financial damage and emotional harm to businesses and individuals when utilized by persons with malicious intent. One of the more obvious examples includes the Mante Te’o impersonation involved with college football. For this reason, many websites and online services are adopting policies to address the issue of impersonation; which, more often than not, do not tolerate the practice (Share, 2009).

Online impersonation involves using knowledge of a person’s personal life, web browsing behavior, and any other available information in order to imitate or copy the actions of another person. This behavior has been declared a crime in a small handful of States in the U.S. and is slowly working its way toward becoming a federally punishable offense (Kawa, 2013). While frequently employed for its comedic uses in the form of online parody, impersonation has more frequently become associated with the defamation of someone’s character, fraud, harassment, or attempts to obtain private and secure information. With this knowledge of how impersonation is being used and its forthcoming legal status in the future it is
important to understand how to differentiate between the victim of impersonation and the perpetrator responsible for the attack.

An impersonation attack is carried out quite easily in a fashion that most internet users are capable of performing themselves. A user simply has to use social engineering techniques and investigation tactics in order to determine information that is relevant to a particular user. Using this information the attacker can create an online profile that mimics the user they are trying to impersonate. Posing as the user; the attacker is able to promote advertisements, mine personal information, initiate malware downloads, and publish inaccurate or inflammatory information about an individual. Upon assembling all of this information into a profile, the attacker is able to establish a basis for the creation of a fake identity.

The perpetrator of an impersonation attack likely has a few different goals in mind including defamation, social engineering, advertising and phishing attempts (Reddy & Kumar, 2012). In order to carry out these activities, attackers will engage in very specific behavior in regard to their web browsing history. Just like the user profiles the impersonators themselves are trying to imitate, the browser behavior profile created by the attacker will also leave clues as to what they are up to and that they are in fact not who they are representing online. The profile of the attacker is comprised of a handful of information that can differentiate them from the user they are impersonating. This information includes analyzing the websites being visited by the attacker. More specifically, it is useful to compare the domain names of visited websites as well as information in regard to websites that provide real world services in specific geographical locations. An attacker frequenting websites that provide real world services located in New York City while impersonating a user based on the West Coast is more likely to be detected as fraudulent.

Furthermore, in developing the profile of a possible impersonator it is important to gather information on all websites and domain names in an effort to create an accurate online fingerprint. Upon gathering this data and comparing it with the legitimate user, information should be flagged if it displays “significant changes in a user’s behavior profile” (Erbacher, Prakash, Claar & Couraud, 2007). In similar studies conducted on masquerade detection techniques, researchers found that by monitoring specific behaviors and actions about user interaction in GUI systems they were able to achieve up a 96% detection rate between original users and their impersonators” (Bhukya, Kommuru & Negi, 2007).

While it is not possible to use web browser history to determine the mouse movement, mouse clicks, and keyboard strikes that were recorded in this study, the same adaptation of the methodology could be applied to web browser analysis. Taking into account a number of different types of information gathering statistics and comparing them between the impersonator and legitimate user could provide a solid foundation for impersonation detection. A sample of these statistics could include information such as:
- Quantity of websites visited in a given time frame
- Log in/Log out activity on social media websites
- Frequency of visitation to a single website
- Time spent on a single website
- Number of subdomains accessed on a particular website

Even if an impersonator is primarily visiting the same websites as their victim with these statistics it could be possible to find significant variations in how the same website was accessed and utilized. In addition, comparing the amount of activity on the website being used as the impersonation platform could reveal additional information in regard to message posting history and frequency and type of communication between contacts. Using these methods in conjunction with each other may not be a quick method because it does require a significant amount of information gathering and analysis. However, its potential and applicability to impersonation detection is important to consider.

Another method of impersonation detection in regard to a user's web browser profile involves looking at search queries on major search engines and profile investigation on social media websites. An attacker looking to impersonate a particular user will utilize broad web searches specifying the victim's name as the keyword. This method would be an attempt to gain as much information about a person's online presence as quickly as possible. While the average user might perform this type of search every once in a while out of mere curiosity, these searches would be a primary and frequently used information gathering technique to be used in phishing attempts and other social engineering attacks. In analyzing browser history, the number and frequency of these “ego-searches” performed could significantly correlate with a user looking to masquerade as someone else.

Upon performing a generic web search of their victim, the attacker is likely to engage in information gathering on the online platform on which they will be impersonating. Using the popular online communication platform Facebook, an attacker could identify their victim’s location, interests, list of friends, and a variety of other information based on the victim’s security settings. This is important to understand because the attacker is likely to use this information to create their mimic account with as much in common with their victim as possible. The key identifier of the impersonator is how quickly they are going to try to get their profile to match their victims. The victim has likely taken at least a few months of time to build up their profile, all of which the impersonator has stolen and put together very likely in a matter of a few hours. Identifying factors of this behavior in an impersonator’s web browser history would involve spam contacting of the victim’s social contacts, joining the victim’s online social groups, and installing any apps the victim is using. Viewing the attacker’s history will reveal a substantial amount of activity, in a small amount of time, on a very specific website. Identifying this unusual behavior could be an indication that a particular user is engaging in impersonation. (Kontaxis, Polakis, Ioannidis, and Markatos, 2011)
There are many instances in which the impersonator is a physical person performing the attack and gathering information manually. On the other hand, it is quite common for impersonation attacks to be carried out via botnets. Botnets are described as "networks of compromised computers with malicious code which are remotely controlled and which are used for starting distributed denial of service (DDoS) attacks, sending enormous number of e-mails and other sorts of attacks" (Stankovic & Simic, 2009).

Often time these botnets are used to carry out impersonation attacks through unattended or unsuspecting computers and servers. Through the use of botnets, detecting impersonation can be become even more difficult due to the fact that the physical medium being used to perform the attack likely belongs to an innocent and unsuspecting victim. These botnets are more commonly used to send phishing emails in which the impersonator is likely posing as a member or a large well known financial institution (e.g. PayPal, Amazon, and EBay). The message included in these emails is usually designed to entice the user to click a link provided in the email in order to claim a prize or verify account information for security reasons. If the link is clicked, the user is taken to a fake site that resembles its authentic counterpart where any information provided by the unsuspecting user is taken and used for malicious purposes (Steel & Lu, 2008).

These types of phishing emails are common and users likely encounter them on almost a daily basis. In order to identify the source of the impersonation it is important to identify the botnet which is being used to carry out the attack. According to Steel and Lu

"There are four key points at which an investigation into impersonation schemes can be launched- during the withdrawal of the stolen funds, from the website hosting the impersonation pages, from the email source, or from the initial site construction (2008)."

Gathering this information together in an investigation of an impersonation scheme it could be possible to track down and detect multiple attackers utilizing botnets. Investigation into the common email source domain as well as whatever site is hosting the fake material could provide a direct link to the individuals responsible for the attacks. If an inspection of their browser history and comparison with the above tracking techniques reveals similarities then an impersonator has more than likely been identified.

Another method for determining the presence of impersonation through a user's web browser is visited in a study conducted by Utah State University. The method, referred to as compression, is focused on analyzing the compression of image files online. The basic idea involves a user uploading some original images to a social media or personal website. In order for an impersonator to achieve duplication of the user uploading the pictures they have to download the pictures to their personal computer and upload them to the duplicate account profile they have
created. Because the impersonator has downloaded a unique picture file and uploaded for their purposes it is likely that the picture has been compressed at a different ratio than what was used by the originator of the picture (Erbacher, Prakash, Claar & Couraud, 2007). If an impersonator's browsing history indicated that they uploaded any pictures to the internet, this uploaded picture could be compared to other similar pictures on the internet utilizing a search engine capable of performing an image search. If the same unique picture is displayed for two different accounts providing the same service (such as a Facebook), a comparison of the image ratios could be step in differentiating the impersonator from the originator of the image.

Going beyond the basic visual inspection of browser history to determine if fraudulent activity exists, there are a number of tools and programs that can be utilized for a more in depth analysis of browser history, cookies, cache, and even deleted information. These programs emphasize the importance of interoperability scanning between multiple browsers due to the fact that there are multiple browsers commonly used for online access. Complementing this interoperability between browsers is the ability of programs to examine common search engine host and paths names as well as how URL's are encoded. These features allow for the program to interpret and decipher any language that might be encoded. Once this information has been analyzed it can be used to determine any evidence of online impersonation existing (Oh, Lee & Lee, 2011).

Using web browsing history analysis programs it is possible to uncover deleted history information due to holes in the "delete history" functions in a majority of web browsers. This is based on the fact that many of these web browser features delete log files created by their own programs but do not necessarily delete the files and information downloaded to separate locations associated with the browsing information. Despite the use of a number of tools that promote complete removal of browsing information such as Window Washer, Privacy Expert, Secure Clean, Evidence Eliminator, and Cyber Scrub it is still possible to recover certain information about user history (Geiger, 2005). While this information may be somewhat discouraging to the average citizen it is encouraging to security professionals in determining where unlawful activities are being performed. Employing a utility such as Web Browser Forensic Analyzer (WEFA) has proved to be the most useful in determining web history browsing between multiple services, including information and log files that need to be extracted or have been encoded or deleted (Geiger, 2005). Analyzing a potential suspect using WEFA could provide information in regard to the impersonation of a particular individual.

In conjunction with web browsing analysis programs an investigator is able to utilize web browser fingerprinting methods to confirm the existence of impersonation. In a study performed by Keiji Takeda, browser information was gathered between different users in regard to plugins, system fonts, screen size, and the HTTP accept header. Using this information they were able to determine that because a web server is able to identify a specific web browser it could be possible
to narrow down identification clues such as geographical location and the physical
device being used (Takeda, 2012). Comparing this information with the online
profile suspected of being impersonated it is possible to make a reasonable
determination about the motivations of the actions being performed and whether or
not they seem to indicate impersonation.

Expanding on this idea, Peter Eckersley was able to determine that the
particular configuration and manner in which a web browser is used is quite unique.
Taking samples of how different users interacted on a specific website an algorithm
was created for identifying unique web fingerprinting. Their research revealed that
their methods for tracking unique users were very effective with a less than 1%
false positive rate (Eckersley, 2010). Employing this methodology in combination
with other context clues it is possible to identify impersonation. This could be
useful in order to detect individuals who have already been suspected of performing
malicious activity. Looking at the questionable online activity and comparing it with
the suspect’s web browser footprint could provide a way for indicating if
impersonation is occurring.

Further research suggests that contextual information has a significant
bearing in determining the motivations of an internet user (Jakobsson & Stamm,
2006). This information is important to distinguish because it goes further than
merely identifying what websites a user has visited. It attempts to establish the
reasoning and actions performed when visiting a particular website. In web
browser forensic analysis it is not enough to know where someone has been but
also why they were there and what they did. In a real world, physical scenario this
information seems obvious and crucial to obtain in order to establish an individual's
motivations. The case is the same in regard to online activity.

A cursory inspection of browser history can reveal the websites a user has
visited giving an investigator a general idea of what websites a user has been
accessing. While this information is important the point of focus now must be on
the context of the user’s visit. A legitimate user may access an online banking
website in order to transfer funds between accounts, look at statements, or pay a
bill. On the other hand, someone looking to impersonate or perform illegitimate
activities would visit the same banking website in order to scout out the intricate
details of the site. Thieves perform the same functions before carrying out a
 robbery. The visit the place they are going to raid in order to gather information on
the best way to carry out their heist. The idea is the same with an impersonator
wishing to carry out a phishing attack. In order to fool the victim into giving up
information the attacker's fake website must look genuine and authentic.

In order to achieve this look of authenticity the perpetrator would have to
spend their time analyzing the websites their victims frequent. This is where the
differentiation between a legitimate user and an impersonator can be established in
regard to their web browser history information. An impersonator would likely
visit financial websites in order to gather screen shots of the site and its sub
domains to use in the creation of a fake site. While performing these actions the impersonator would specifically visit homepages and log-in sub-sites in order to replicate accurate depictions of the web pages. When visiting specific financial websites, regular users have a legitimate purpose for accessing the available content. This means they will likely log on to the website, engage in the tasks they need to complete, and then leave the website. In contrast, an impersonator is not likely to try to log in to obtain any services from the web site they are visiting. Examining these details and performing a detailed forensic analysis on web browsing history can reveal accurate motivations of what someone is trying to accomplish online.

Upon identifying that impersonation is very commonplace when interacting online it is important to point out methodology and best practices in order to detect impersonation of an internet user. While complete prevention may be unable to obtain, knowing common detection measures is useful in quickly identifying impersonation attempts so that they may be handled before getting out of control. Some best practices for minimizing impersonation attempts include:

- Enabling enhanced security and privacy settings on social media websites
- Limiting personal information provided on social websites
- Enabling 2 step verification
- Utilize private or incognito features provided by major browsers
- Perform online transactions with secure connections (HTTPS vs. HTTP)

Employing these techniques will decrease the possibility of being impersonated by an attacker. However, these steps will not be helpful in determining if a user is already being impersonated. In order to establish whether or not these events is occurring an individual can take a number of steps to make this assumption. One of these steps involves a user performing a text based search online to find out where information about them is being listed online. This step will identify foremost if impersonation is actually occurring as well as the extent to which it is happening. Another form of exposure involves employing the same search methodology and using a personal image to search the internet for impersonation. Using an image from a commonly used online profile in a web search is capable of revealing many of the different locations that similar images are being hosted from. This ability is constantly being redesigned and narrowed down to more accurately identify images based on metadata and previously performed algorithms (Grauman, 2010).

The information presented in this paper has identified the methods involved with determining if and how impersonation attacks occur via web browser history. Impersonation in an online context occurs on a daily basis. In terms of legality, internet impersonation is an activity that examines the use of how and for what purpose the originator’s profile has been intended for. While it may be used for advertising and comedic uses, its vitality as a form of identity theft and online security calamity is far more destructive. Methods have been identified in order to
differentiate the web browsing habits of an intrusive impersonator and a casual user. These methods involve identifying frequency of visits, context of the website visit, compression analysis, phishing activity, and use of web browsing forensic tools.

However, using browser history as a primary means of detecting impersonation has many pitfalls. Due to private browsing features and the ability of a user to have complete control over what information they want to keep, it is unlikely that individuals performing impersonation attacks will keep a log of their online activities readily available. What is more likely to happen is that attackers will utilize browser features that disable history keeping as well as delete any information that was logged in their most recent browsing session. Understanding the security risks involved with web browsers tracking and holding on to user data, certain companies have implemented browser designs that make sure that session activity is not logged. One of the more well-known web browsers designed with anonymity in mind is the Tor browser. The Tor browser works by relaying any online communication through thousands of networks in order to conceal user location and identity (Tor).

In order to successfully detect and deter online impersonators, analysis of web browsing history should be used on conjunction with a variety of investigation tools. These tools include forensic web browser analyzers, Computer Online Forensic Evidence Extractor (COFEE), Metasploit, and The Sleuth Kit. Employing these tactics is step in the right direction in impersonation detection. Additionally, it should be noted that many of the tools, strategies, and methods in regard to detection were utilized in controlled environments under very specific conditions. This is important to understand because the legality of employing these techniques can vary based on the agreement set forth by the individuals and organization involved. Upon identifying these forms of destructive impersonation users are encouraged to enforce security settings in online environments, encourage awareness of their personal online profile presence, and be careful about the information they post in an online environment.
References


*Tor: Overview*. (n.d.). Retrieved from [https://www.torproject.org/about/overview](https://www.torproject.org/about/overview)