National College of Ireland

Towards a Conceptual Model for Mitigating against Social Engineering on the Online Social Network (OSN)

MSc Internship
Cyber Security

Olabode Ololade
x17127599

School of Computing
National College of Ireland

Supervisor: Rohan Singla

National College of Ireland
Project Submission Sheet – 2017/2018
School of Computing

Student Name:	Olabode Ololade
Student ID:	X17127599
Programme:	Cyber Security
Year:	2018
Module:	Academic Internship
Lecturer:	Rohan Singla
Submission Due Date:	13/08/2018
Project Title:	Towards Conceptual Model for Mitigating against Social Engineering on the Online Social Media (OSN)
Word Count:	7626

I hereby certify that the information contained in this thesis is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are encouraged to use the Harvard Referencing Standard supplied by the Library. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature:	
Date:	14th September 2018

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. **You must ensure that you retain a HARD COPY of ALL projects**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specifically requested.
3. Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Towards a Conceptual Model for Mitigating against Social Engineering on the Online Social Network (OSN)

Olabode Ololade
x17127599
MSc Internship in Cyber Security

14th September 2018

Abstract

Defeating social engineering in recent times has proven to be quite herculean as it most times defeats traditional security measures. The security of Online Social Network (OSN) users typically depends finally on the OSN users as they are the creators of the desired content they upload, post, update and display on the OSN. Daily, a large amount of Sensitive Personal Information (SPI) or Personal Identifiable Information is sent through the Online Social Network (OSN) and as a result of this, some users are susceptible to social engineering attacks based on PII they post. In this research, we have developed a novel solution called the Social Engineering Notification System (SENS) to improve the Security and Privacy Settings of the Online Social Network (OSN). SENS has an embedded indicator that determines the risk measurement from (Low risk, medium risk, and high risk) and also displays susceptible keywords to further guide the users based on the number of susceptible keywords and certain potential PII with high value. This solution can be replicated and integrated into any social media platform to enable OSN users to have more control and monitoring ability on their potential post before going live on the OSN. By so doing, SENS will minimize the high volume of Sensitive Personal Information (SPI) or Personal Identifiable Information (PII) OSN users post on their profile. With SENS we believe there is a great potential to bridge the gap between the OSN users inability to vet their post/updates for sensitive personal information (SPI) hence improving privacy and mitigating against the feasibility of social engineering attack on OSN users.

1 Introduction

1.1 Project Background and Motivation

A massive amount of information is held by the Online Social Network (OSN) and it is considered one of the biggest repositories of information asset not just for individuals but also organizations. One major issue with the OSN is the issue of users sharing sensitive personal information which social engineers capitalize on to launch several types of attack using various deceptive mechanism. According to In 2011, Dimensional Research, ran a

survey involving 850 Information Technology and Security professional in the United States and seven other countries which over 45 percent of these professionals had been victims of a social engineering attack. The report further stated the average cost of a social engineering incident averages between $25,000 and $100,000. Although Online Social Network (OSN) has its good sides as a great tool for communication and particularly information dissemination but nevertheless has its downside as confidential informations are sometimes shared and can lead to compromise of individuals. Many of the currently used mitigation strategies from traditional security measures have not worked effectively due to social engineering utilization of deceptive mechanisms to manipulate the human mind. Intelligence gathered from the Online Social Network (OSN) also known as Social Media according to [1], are one of the leading propagators of cyber-attacks in the world. They even specifically described how hackers use Twitter to socially engineer employees to click on a deceptive link which at some point enabled access to Pentagon systems. According to [2], $1 billion has been stolen in a total of 100 banks and 30 countries within 2 years because of social engineering attacks. Based on OSN behavioral patterns of individual, fake accounts can be created, and specially crafted link of interest can be advertised or portrayed to an individual in order to click or follow hence compromising an asset of the OSN user. Policies, user management, and user awareness have not proven to be fully effective either due to fact social engineering attacks evolve regularly. Fan et al. (2014)) explained how social engineers leverage on the display of sensitive personal information to build a word list which afterwards they attempt to crack using brute force or dictionary attack. The attacks are getting complex, hence the need for a more solutions to mitigate against it. Looking at the few statistics and many more on the Internet, it is evident the need for more effective solutions to address the issues of social engineering. In this research, we have developed a conceptual model we call the Social Engineering Notification System (SENS) that could enable Online Social Network (OSN) users to vet their update before it goes live based on the susceptible keywords and expressions. The vulnerable keyword are based on Personal Identifiable Information (PII) or Sensitive Personal Information (SPI). Our major focus is on Personal Identifiable Information as the theme for checking the post of OSN users.

The Paper is structured around six sections as follows: Section 1 addresses the specification of the project which includes the research question, purpose, and research variable utilized in the paper. We have also given an overview of Online Social Network (OSN), and Social Engineering. Section 2 gives details of related works on Online Social Network and Social engineering. Section 3 gives details of the project design and methodology utilized in this research. Section 4 describes the implementation on how we carried out the development of the model. Section 5 is based on evaluation, some case studies and discussion as to the result we have gotten from running our SENS model. In section 6, we conclude our research and propose future works regarding the potential of (SENS).

[1] https://www.nytimes.com/2017/05/28/technology/hackers-hide-cyberattacks-in-social-media-posts.html

[2] https://opendatasecurity.io/the-most-famous-cases-of-social-engineering/

1.2 Project requirement specification

1.2.1 Research Question

Based on the issues of social engineering, we ask this question - Can we reduce the amount of Personal Identifiable Information (PII) or Sensitive Personal Information that Online Social Network (OSN) users post on the OSN platforms?

1.2.2 Purpose

The purpose of this research is to alert Online Social Network (OSN) users on how vulnerable their post could be to a social engineering attack. Although Online Social Network has become a major form of communication but also poses a threat to Personal Identifiable Information (PII) or Sensitive Personal Information (SPI) leakages. OSN users knowingly and sometimes unknowingly reveal a lot of this SPI such as telephone/mobile numbers, emails, and geographical location hence exposing themselves to social engineers who could use this SPI for cyber-attacks and sometimes even physical attacks Gross and Acquisti (2011). Krishnamurthy and Wills (2009) discussed how Personal Identifiable Information can be accessed by third parties through referrer headers and cookies. According to Williams et al. (2009), Online privacy management on the social platform is relatively new and requires more in-depth solutions. They further explained how the exposure of Personal Identifiable Information to an unauthorized person can lead to Identity theft.Algarni et al. (2013) further buttressed the dire need for more solutions to solve the susceptibility of human behavior that leads to social engineering. We believe SENS is a step in the right direction in mitigating against the end result of social engineering like Identity theft, credit card fraud and many more as Algarni et al. (2013) mentioned.

1.3 Research Variable

Mitigating against the risk of social engineering is the goal of this research and in so doing we have developed a novel solution called Social Engineering Notification System (SENS). We believe if OSN users are notified on potential threat in posting SPI, they will take more cautious of posting their PIIs on the Online Social Network (OSN). In the development of SENS, we have used Windows Apache MySQL PHP (WAMP) as our webserver, PHP, HTML, CSS, Bootstrap and JavaScript to build the model.

1.4 Social Engineering

Social engineering is a security attack that aims to take advantage of human behavioral weaknesses to meet the desired outcome of an attacker Mitnick and Simon (2011). An individual who performs a social engineering attack is called a social engineer. Social engineers make use of meeting other users on the OSN to create trust and afterward take advantage of the trust by launching an attack. According to Algarni et al. (2013), social engineering poses a real-life threat to enterprise, governments, individuals and on the Online social network (OSN). They further discussed how these threats evolve from technical to exploiting human vulnerabilities through deception. Social engineering is one of the lowest cost-effective types of attack which makes it one of the most utilized attack strategy Mataracioglu and Ozkan (2011). The flow below depicts how a social engineer could utilize several Sensitive Personal Information from OSN users to craft various attacks.

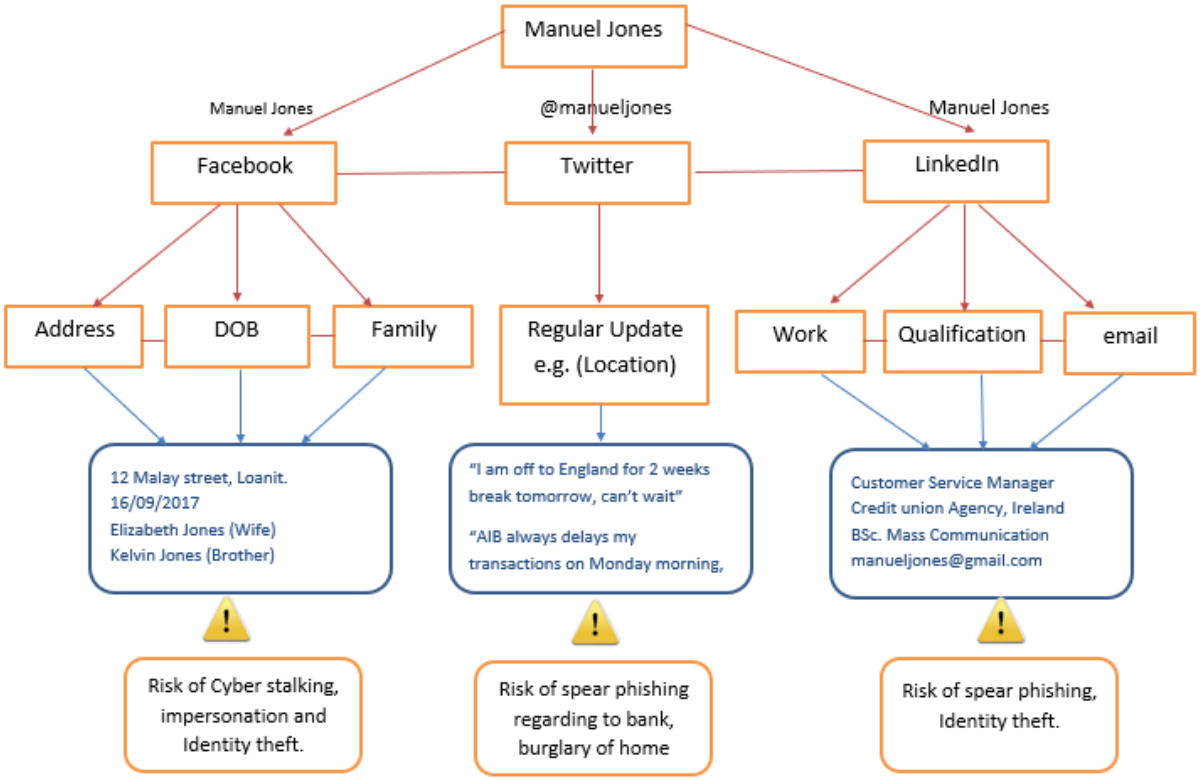

Figure 1: Social Engineering on OSN

1.5 Online Social Network (OSN)

In recent times, communication and dissemination of information have taken a different turn from the traditional methods of a post, discussion boards, and newspaper. According to Williams et al. (2009), Online Social Network (OSN) are services based on the web platform that gives opportunities for individuals to create private or public profile and connect with friends or people with similar interest within the Online Social network (OSN). The image above depicts how social engineering can make use of Personal Identifiable Information to launch an attack.It shows how certain PII's released by an individual can be collaborated by a social engineer to launch several attacks. OSN now influences how its users collaborate, communicate and influences livelihood. Kaplan and Haenlein (2010) define OSN as a mobile or web-based application on the Internet that enables the formation and interchange of content created and generated by its users. Social media has exponentially increased digital communication in general and leaves an audit trail of who did what, when, under what circumstance, and with whom Gleave et al. (2009). Behavioral patterns can now be deduced by OSN and used to influence how products can be sold, services can be offered and even now used as a tool to drive political campaign as we had seen recently in the Cambridge Analytica. Some examples of the popular Online Social Networks (OSN) are Facebook, Twitter, and Instagram.

1.6 Programming Languages

1.6.1 HTML

This stands for Hypertext Text Markup Language and is typically used in building web pages. They are the building blocks of web pages and send documents from a web server to a web browser. HTML also works with other languages to create interactive/dynamic web pages and according to (Flanagan, 2011), HTML, CSS, and JavaScript are the three core technologies in the World Wide Web.

1.6.2 CSS

Cascading Style Sheet (CSS) basically creates a presentation and formatting of content by adding extra properties such as colors, fonts, and layout to improve the aesthetic of a web page (Flanagan, 2011). They basically work in synergy with the Hypertext Markup Language.

1.6.3 Bootstrap

This is an open source framework utilized in designing websites and web applications. Some of the features integrated into bootstrap are CSS and HTML Predefined structures. They are typically used in the development of front-end interfaces.

1.6.4 JavaScript

JavaScript is a programming language with features that make up dynamic, interactive websites and application. JavaScript is one of the triads of web technologies according to Flanagan (2011).

1.6.5 PHP

This is used for server-side scripting basically implore for development of websites. They can be typically embedded into HTML codes and can be made use of by many other web platforms. Interpretation and executions are typically functions performed by the web server in form of several components such as images and different types of data. PHP also have a Command Line Interface (CLI) and a Graphical User Interface (GUI).

1.6.6 MySQL

This is an open source relational database. They are utilized by top Online Social Network (OSN) such as Facebook, Twitter and even the Google platform.

1.7 Personal Identification Information (PII) or Sensitive Personal Information (SPI)

This is defined as specific features or attributes pertaining to the identity of an individual either directly or indirectly. Most of which typically includes home addresses, mobile phone number, social security number and other information Chen and Rea Jr (2004). Alim et al. (2011) further identified some other SPI such as photos, names, emails, date of birth and how they can be used by attackers for identity theft and impersonation.

With the increasing amount of Internet usage and the processing of electronic entities, it is paramount Personal Identifiable Information or Sensitive Personal Information is protected to ensure the privacy and safety of Internet users in general. Quite a lot of measures have been taken to ensure the protection of PII/SPI from compromise, an example of which is Security Assertion Markup language (SAML) which provides secure communication using Digital signature. Krishnamurthy and Wills (2009) explained in their paper how it is possible to combine PII or SPI gotten from Online Social Network (OSN) with other information to craft a social engineering attack. They further discussed how PII may be leaked to third-party applications as a result of the cookies in the HTTP headers which third-party applications typically have access to. Our next section explores the relationship between social engineering and Online Social Network (OSN) pointing out some key strategies on how related works have handled this attack. We look inward why most of the related works have not been able to handle the social engineering threats effectively and further explain why our novel idea of SENS have the potential of being a better solution.

2 Related Work

Lately forecast of the human behavior through Online Social Network (OSN) has been utilized for various activities, example, Politics, Marketing, relationship, achievement, work fulfillment among numerous activities. A large portion of the predictions is mostly to enhance user experience and associating activities they are analyzed for. Sensitive Personal Information (SPI) has been shown through Online Social Network (OSN) most times to reflect an individuals private life Wilcox et al. (2014), which in turns makes them vulnerable to social engineering attacks. In this research, we developed SENS that could enhance the chances of users in becoming less susceptible by notifying them before their post is published. Now, OSNs only warn users in their privacy policy on how they are responsible for what they post on the Online Social Network (OSN). This is a good way to inform users to avoid posting Sensitive Personal Information (SPI) about themselves, however, the reverse has mostly been the case as a lot of users do not read the privacy policy. Statistics buttress the point a lot of social media users do not read the privacy policy on the social network website. Lee and Lee (2002), stated the need for implementing a technical and social solution in synergy to predicting behavioral patterns on the Online Social Network (OSN). Our argument points to the fact we believe OSN users should be notified and more informed of how vulnerable their posts could be to social engineering. We believe it is better if posts are analyzed based on certain Personal Identifiable Information (PII) and behavioral patterns of the social media users to allow users to vet their post before it goes live on the OSN. Since social engineers typically look out and crawl Online Social Networks (OSN) to gain sensitive personal information, it will be harder for them to have sensitive information if Online Social Network (OSN) users do not post about the sensitive personal information on the Online Social Network (OSN). With SENS, Social Media Managers will be able to alert to OSN users on susceptibility and risk ratio before users post their information Online.

The related works of this paper are further organized in the following way:

- How Social Engineering leverages on the Online Social Network (OSN)? A case study of the Big Five Personality Traits.

- Social engineering on Online Social Network (OSN) and how it can affect an Enterprise,
- Access Control, Privacy Settings and how it affects Social Engineering
- Spear-Phishing as Social Engineering Attack Vector.

2.1 How Social Engineering Leverages on the Online Social Network (OSN)? A Case Study of the Big Five Personality Traits

In the paper of Golbeck et al. (2011), they indicated how the enormous five personality traits (Openness, extraversion, neuroticism, agreeableness, and conscientiousness) can be anticipated in view of type of post they communicate on Twitter. The aim of their paper was to check whether OSN could predict the inclinations of the user by investigating a relationship between behavioral patterns and other features considering their personality. They completed a personality test considering information gathered from the profile of the user. They utilized ZeroR and Gaussian algorithms in predicting the Accuracy of the big five personality utilizing machine learning concepts. They collected 2000 tweets from each user and made an API that administered 45 questions considering the database of the top five personalities in view of psychology. The outcomes from the research of Golbeck et al. (2011) in predicting personality from Twitter was to show how digital footprints of OSN users can be predicted based various psychological personality traits. They further discussed how these traits can be used to drive online preferences and improve experience of users. However Nurse (2015) further iterated on how oversharing PII leaves a users digital footprints on the cyberspace that can be used against an OSN user by a social engineer. This paper improved knowledge due to the fact they leveraged on five core identity qualities on Twitter users. It additionally created an understanding of how the profile features associate and reflect user's behavior on the OSN. Our paper uses a few factors on the personality trait to foresee the potential vulnerability of OSN users are to a Social engineering attack in view of the behavioral pattern depicted in their post. The five-personality trait of this human central nature are reflected on the OSN however the vulnerabilities that accompany it are exponential on the OSN due to the borderless information dissemination on the Online social network. These attributes are the building blocks of human vulnerability considering their traits on the OSN platforms and can be leveraged on by social engineers. One noteworthy issue in the works of Golbeck et al. (2011) is the issue with incorrectly spelled and condensed words by Twitter users which could have influenced the accuracy of their predictions. Our research focuses towards notifying a social media user on susceptible sensitive keywords in the light of ensuring they indeed want to allow their post on the live feed. Emotions affect a noteworthy part in social engineering and even affect further a far more profound part of how OSN users utilize social media regarding what they post at a specific time. Earlier research by Choudhury and Counts (2012) further buttresses the research of Golbeck et al. (2011) on how OSN users mostly post based on emotions caring less of the potential threat their digital footprints leave for social engineers. In this section, we have outlined the Big Five Personality Psychological Trait which defines the baseline of what human nature is built on and how social engineering leverages on human personality in applying its deceptive mechanisms. In the next section, we briefly look into how social engineering can further

pose a huge threat to organizations solely based on the behavioral pattern of a Twitter user.

2.2 Social Engineering on Twitter and how it can affect an Enterprise

In this section, we examine how the vulnerabilities of OSN users can influence an enterprise and furthermore shed more light on the reasons why our Social Engineering Notification System (SENS) model will be a good addition to the Online privacy and security settings of the OSN. According to Wilcox et al. (2014), Online enterprise social engineering has to do with employees manipulation into doing a certain task in place of a social engineer such as revealing an enterprise sensitive information. To mitigate against social engineering attack considering users pattern of the post, we bridge the gap in mitigating against the vulnerability of social engineering attack on an enterprise. This could be because of an OSN post from an organization's staff that uncovers a specific pattern to social engineers that could compromise the Confidentiality, Integrity, and Availability of product or services offered by an enterprise. Wilcox et al. (2014) discussed in their research on how social network in professional workplaces represent a threat to individuals, organization, and innovation. They discussed how employees of organization could be tricked by social engineering techniques hence releasing enterprise information resources. Based on their research paper, cyber espionage is done through information gathered by competitors employees from an Insider threat based on their OSN behavior to give an organization a competitive advantage over the other. They additionally proposed a solution revolving around policy implementation and employee awareness, however, this isn't effective enough as people need to see exact outcomes and evidence on how they could be vulnerable to it. Our solution, Social Engineering Notification System (SENS) acts as an additional layer of avoidance from sharing Personal Identifiable Information (PII) as releasing SPI of an enterprise could lead to a colossal business impact. In this section, we examined the impact of vulnerable OSN post and how social engineers gain sensitive information of an organization. One contention on the most proficient method to understand this issue of social engineer gathering information without approval sparks an argument on whether Access Control can resolve this issue or not.

2.3 Access Control, Privacy Settings and how it affects Social Engineering

Despite how critical Online security is on the OSN and access control measures that have been developed over time, Online social network users still fall victim of social engineering attacks. Pwint Oo (2013) explained how Privacy is a major issue on the OSN platform. A noteworthy test of OSN is typically how OSNs like Twitter and Facebook considers each follower or connection a friend paying little mind to trust Shehab et al. (2012), however, the trust of such connections and followers can't be evaluated by OSN connection Boyd and Ellison (2007). This influences social engineers to approach potential Personal Identifiable information of their target due to simple access to their profile. This dependably prompts the inquiry with regards to the principal motivation behind why OSN users disclose sensitive personal information on the Online social network platform which Krasnova et al. (2010) disclosed was due to creating friendship and relationship building. Gritzalis et al. (2014) likewise denoted the reason behind taking an interest in OSN by a

user would typically be for networking professionally and as initially specified by Krasnova et al. (2010) to likewise be for communication with other OSN users. They emphasized how Open Source Intelligence (OSINT) plays a strong role in predicting patterns of social media users. Randazzo et al. (2007) additionally discussed how OSINT can be utilized to enhance security against an insider threat by predicting vulnerable behavioral conduct that could have a terrible effect. This was in accordance with the contention on the "danger of Social engineering through Social Media Enterprise" by Wilcox et al. (2014). For a social engineering attack to happen on an OSN user, the social engineer requires a level of access to have the capacity to view post and profile of the potential target; however, most OSN platforms have a privacy setting that can either enable or disable the profile of a user from the public. By enabling the privacy setting, a social media user can either grant or deny access to people to view their profile and the information they share. This Online privacy security feature is a step in the right direction but cannot totally prevent a social engineer from gaining access to a n OSN users profile. The research by Misra and Such (2017) proposed a Personalized Agent for Access Control in Social Media decision maker that relies on the interpersonal bond between OSN users and the information displayed to suggest who they allow or deny access to their personal profile information based on community network detection. To prove this, they performed an experiment by building an application utilizing Facebook Query Language and Facebook Graph API for access control decision making by participants. The solution they came up with is quite good as it could help OSN users make informed decisions regarding whom to follow but does not completely solve the issue of social engineering. The research done by Bilge et al. (2009) shows that social engineers have grown smarter on how they bypass access control strategies and mechanisms by creating or cloning an identity of a mutual friend and sending a request to follow or have access to a social media user. The attackers most times study interpersonal relationships and eventually spoof the identity of a friend who hardly post at the social engineering target but still has an active account on OSN. The potential victim sees a known username and probably has forgotten they already follow each other and unfortunately grants access to the spoofed account. They also added how the attackers craft special social engineering messages like "Dear friends, I forgot the secret password of my old account, I am recreating my friend list. If you don't mind including me once more!" This procedure has a high success rate and means the social engineer presently has full access to the target in the wake of being confirmed and accepted as a friend and can monitor OSN post again and plan an attack. Even though access control is a decent measure to mitigate against social engineering on the OSN, the research paper by Bilge et al. (2009) likewise proves that research of "PACMAN" proposed by Misra and Such (2017) is not a silver bullet in mitigating against social engineering exploits on the Online Social Network. In this section, we discussed access control mechanisms that have been implemented by some researchers, but we also argue that the access control feature on the privacy setting does not effectively solve the social engineering attack.

2.4 Spear-Phishing as Social Engineering Attack Vector

In this section, we consider one major attack vector OSN users are vulnerable to and the impacts of this attack vectors. For a social engineer to successfully execute an attack, they leverage on Twitter and other social media sites to study their target and gather in-depth information Khonji et al. (2013). They further discussed how social engineering

techniques are now used by hackers to convey malicious links by also leveraging on the OSN. In 2013, over 3 billion Yahoo accounts were compromised as a result of a single spear phishing email accidentally opened by an employee of the company who further clicked the link attached to the email [3]. Mohebzada et al. (2012) characterized Phishing as an event that happens where a potential target has a message conveyed to him/her that imitates an approved source. Spear-Phishing is typically a decoy to trick a potential victim in releasing Personal Identifiable Information, for example, passwords, credit/debit card details and among numerous different personal details. The borderless flexibility, monetary profits and the easy transfer of phishing make it a standout as one of the most used techniques for social engineers Jagatic et al. (2007). Being susceptible to this attack is typically dependent on the amount of sensitive personal information an OSN user post on the social media platform. For example, if a social media user continually posts about their failure about a disappointing banking experience on a particular day and within a timeframe, there could be a plausibility this behavior can be utilized by a social engineer to craft a phishing attack as a Direct Message (DM) coordinated to the target hence getting sensitive personal information from the user. These messages are well crafted and typically come with attractive offers or bonuses as to regard what the social media users had expressed interest in based on the information posted on the OSN platform. Based on the research by Grazioli (2004), they gathered social engineers have since identified how many OSN users are attracted to offers and have since utilized this weakness against users in manipulating them into giving sensitive information such as credit or debit card details hence losing lots of money. Iacovos and Sasse (2012), even discussed further how attackers utilize attractive offers to lure Internet users to go to fake websites hence extracting Personal Identifiable Information such as financial details and many more. Our research centers around using OSN users post to notify them based on certain identifiable keywords of how vulnerable they are to social engineering attack who mostly utilize sensitive personal information retrieved from potential targets to launch a Spear-Phishing attack. Our argument further supports the motion of **?** and Herley (2010) that education awareness to Internet users alone is not a silver bullet to solving this issue but we must also explain furthermore the full attack strategy utilized by giving a step by step procedural guide to the users. As indicated by [4] review demonstrates social engineering tops the rundown of the highest technique for hacking. In this section, we have characterized Spear-Phishing as an attack vector for social building and how it exploits information gathered.

Our observations also did make us understand the reason why attackers have recorded high success in stealing sensitive personal information from OSN users is due to the gap in knowledge and awareness from the user hence the exploitation and the success rate. We believe a holistic approach should be taken to bridge the gap in knowledge between vendors and OSN users.

3 Methodology

The Waterfall or linear-sequential life cycle model methodology is the model we have utilized in the development of our model to proof the concept of SENS. According to Howcroft and Carroll (2000) Waterfall methodology involves a set of steps that span

[3]https://www.cybersecuritymastersdegree.org/2017/11/top-5-social-engineering-attacks-of-all-time/

[4]http://www.computerweekly.com/news/4500272941/Social-building is-top-hacking-strategy

the process of development with a slight level of repetitions between each level. This methodology is considered the traditional model used in the IT industry, quite sequential and is best used when there is are minimal uncertainties in the development Sommerville (2010). We have chosen this model in our development of SENS as Waterfall is linear, easily measurable since we already have a pre-defined scope as to how SENS was to be developed and the functions it should execute. Below are the steps involved in Waterfall methodology.The image below depicts the Waterfall model.

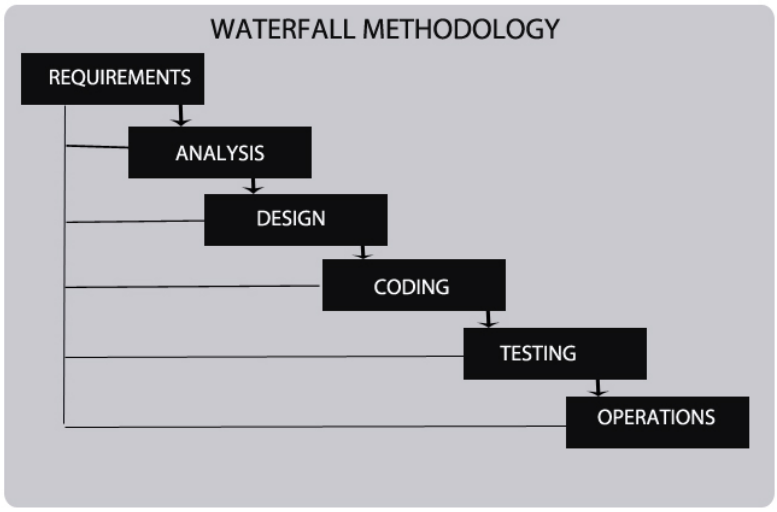

Figure 2: The Waterfall model

3.1 Requirements

Since the purpose of the SENS model is as a security feature to mitigate against social engineering attacks, one major requirement is the development of an OSN website to integrate SENS to prove the concept. The minimum requirement to ensure SENS model can be used locally are based on our initial proposals, we understood what was needed to be done and the necessary further steps to take to achieve SENS.

3.2 Analysis

At this phase, we had strategized on how to build the SENS model and how we may achieve the proof of concept. According to the research of Schmidt et al. (2001), lack of commitment and system requirement are the two-major reason software product have high failure rates. We have carried out a risk analysis on the best approach to take in order to prove the concept of SENS. After this phase, the goals of SENS was clearer. After the analysis, we were able to create a document as to the objective of SENS, and risk register to classify the potential risk in the approach we had taken.

3.3 Design

We created our database and how the web pages in our OSN model interact with each other. To also ease usability, better navigation and aesthetic, we have made use of graphic design tool (Adobe Photoshop) to design some icons, we also made use cascading stylesheet, bootstrap, and jQuery to endure the graphical user interface is appealing to the intended audience.

3.4 Coding

3.4.1 Client-Side Coding

In building the OSN website structure, we have used Hypertext Markup Language (HTML) based on the World Wide Web Consortium (W3C). In creating our OSN model to prove the SENS concept, we have used HTML as the building blocks for our web pages. We have made use of the WAMP server as our local server in developing SENS since its easy usability, no downtime and give full access over the configuration of the setup. We made use of the MySQL database. We also made use of a sublime text editor and notepad ++ as my integrated development environment (IDE). We have made use of the bootstrap library to make our OSN website more responsive at the front-end. We have utilized bootstrap because it is easily customizable and due to how fast it is to build the front-end of our OSN model as a result of its predefined framework. We have made use of JavaScript to do some client-side validation, improve the performance of the website and also to integrate SENS. We also embedded custom Script to define rules/policies for all the forms on the OSN website. We have also used jQuery libraries to simplify our website client-side document traversing, event handling and some animation.

3.4.2 Server-Side Coding

We have used PHP. To manage sessions, we have also made use of the PHP session function. I made use of UTF-8-character encoding scheme that allows you to conveniently store a wide variety of special characters in MySQL database. We have also used MySQLi with strongly typed parameterized queries to avoid injections. We utilized a custom script to define form validation error messages and to ascertain regular expressions for my email pattern. MySQLi opens a connection to your database using the MySQLi library. We have made use MD5 hashing algorithm for storing password. This to mitigate against brute force attacks. We have made use of PHP Session function to store information about a visitor from one web page visit to the next. We have used the session last activity of a user to determine the last time he/she was active on the system.

3.5 Testing

In testing the SENS model, we have done security checks based on the OWASP top 10 to ensure it is not susceptible to cross-site scripting (XSS), SQL injection and can portray the SENS functionality without issues. We have also resolved and debugged all discovered bugs in this phase to ensure the model works as initially proposed.

3.6 Operations

To execute the SENS model, there are two ways we have made use of. Through the localhost using WAMP as a server. Once we start the WAMP server, SENS can be ran by inputting `http://localhost/social/index.php` as the URL in your browser. We have also hosted SENS model on the internet where you need input `http://www.trifectang.com/social/register.php` This requires your computer system is connected to the internet.

4 Implementation

The implementation of SENS is divided into two phases which is the design of the OSN model and phase two which is the integration of SENS conceptual model into the OSN. To proof the SENS concept, we have developed an OSN with some functionalities. We have made use of WAMP as our web server due to its robustness and simplicity. We have utilized a MySQL database to create tables for users, post, trends, notifications, messages, friend request and comments respectively. The tables are necessary for the social network site to enable cohesion from a user and the basic social platform functionality. We have also used Hypertext Markup Language to define web pages structures. We have created Index folder which is the main page of the social network and enables navigation to other pages on the web. Some of the features we have implemented on the home page are for the profile picture and name of the user to display once logged in. Once a user is logged in successfully, the users name is retrieved from the database table called users. We also display the number of post from the user and likes if any available from the table users. To retrieve the number of trending topics based on the frequency of how they have been posted about, Accounts can also be closed as we have used PHP and HTML to enable users to close their accounts if they want to. Once they go to the setting page of the website, they can close their account. Once they do this, the session automatically destroys and goes back to the (register.php) Sign In/Up form. We have created the form using HTML and ordered using CSS. Users can also delete their post and likewise update their passwords if they desire to do so. We have created the comment frame to enable users post and update information on the SENS social media. Once the user is signed in and has a session, the comment frame is functional for posting into the live feed else we take users back to register which is the Sign In/Sign-up form. They are mostly utilized to remove HTML tags, delete white spaces, update post, check if the user who posted, have their account closed, and specify timeframe. We have utilized the user table to validate users, retrieve the number of usernames, post, names, profile picture and friend request from the database, We have also made use of Ajax to enable certain web pages to load and get up to date asynchronously while undergoing an interchange data with the web server without disrupting any process. We have utilized handlers such as login handler to store and validate sessions, login informations, from the client side and the server side. We also used the register handler for validating information from the users that intend to login by comparing informations provided and what is being stored on the database to check if they are already existing or not. There is also a setting handler for enabling information update such as name and password. Accounts can also be closed from here. We have also developed a configuration file to enable database connection to communicate with our client and server side. A folder called asset is where we have created our CSS, JavaScript and Images folder. This content of this folder basically controls the user interface of

the social network website. The CSS folder contains bootstrap libraries which we have used for some of the typography, forms, buttons, navigation and some component of the graphical user interface. We have utilized CSS for styling and basically to define font size, font type, font color, height, width, border, and padding.The figure below is a flow chart that describes how SENS work.

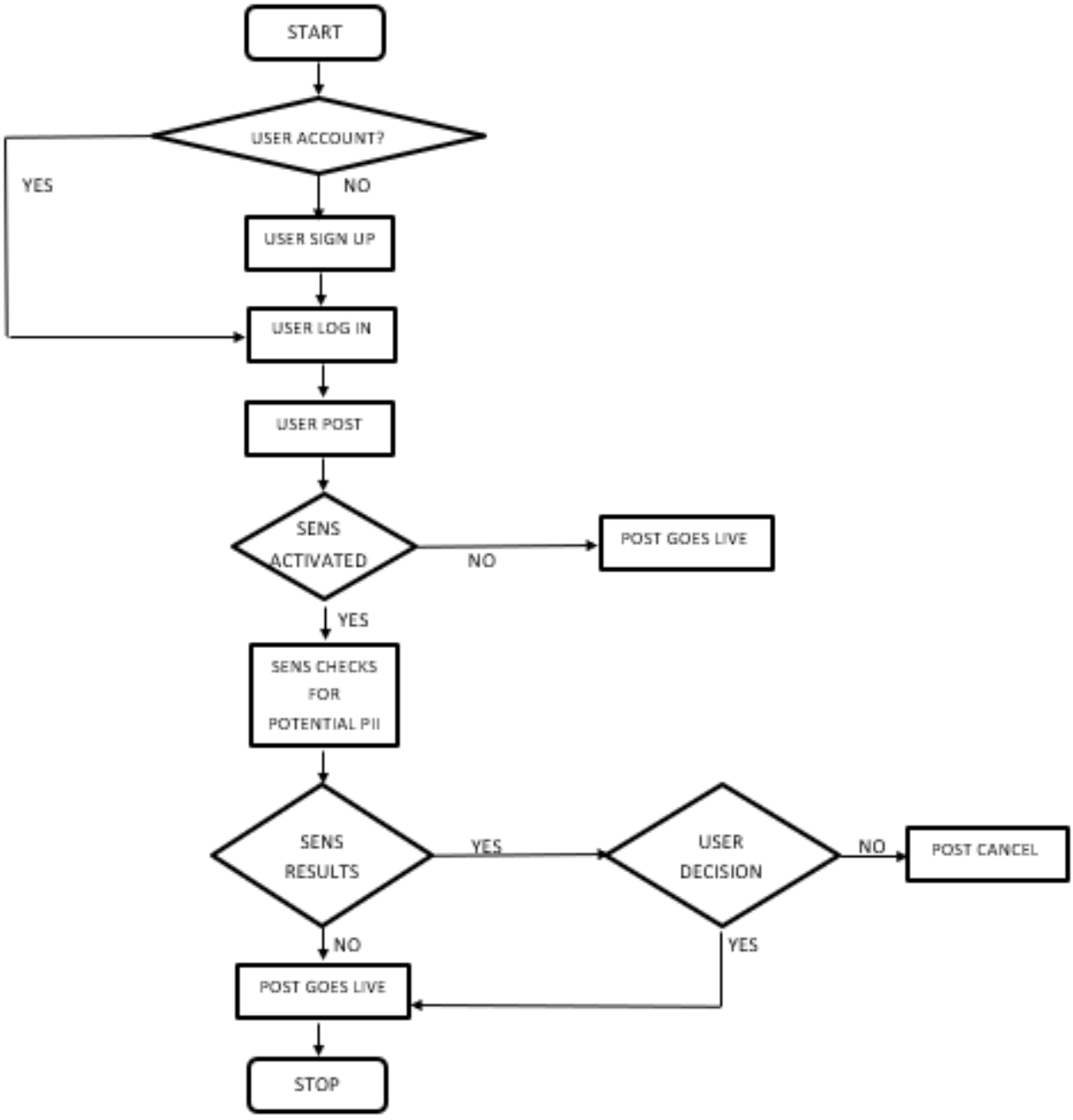

Figure 3: Flow chart of SENS

4.1 Integration of SENS

The main language we have employed to integrate SENS into the social network model is JavaScript. We have chosen JavaScript due to its functionality and flexibility and also

due to the interaction between the post and the privacy settings. JavaScript is quite known for its web form validation and some of these concepts are what we have utilized in developing the SENS feature. We have created a Boolean variable in the settings of the social network model to enable or disable SENS. We have created a column for security in the users table of the database. This could have 1 as a value or 0 at a value both representing SENS enablement or SENS disablement in the Privacy settings page. When users attempt to post a Potential Personal Identifiable Information (PII), we make use of a JavaScript custom alert function to warn the users on the PII or potential PII they are about to post. The first major check we had enabled SENS to be able to check is for emails found in an OSN users post using regular expressions function in JavaScript. Once a user inputs an email and clicks on a post, it triggers a click event. We have equally used the regular expression function for checking if a credit or debit card such as Mastercard, American Express card, discover card or Visa card has been posted by the user. It does this by checking the numerical format and combination of the various debit or credit cards respectively. If SENS spots any of the card numbers in the post, the click event in JavaScript is triggered immediately. Regular expression function was able to determine the type of credit card being utilized due to the fact each card issuer basically has a range of numbers on their card, which is denoted by the first 4 digits. We have created a list of text and expressions linked with Potential Personal Identifiable Information (PII) or Sensitive Personal Information (SPI) based on the most common PII's used from the European Commission. There is a possibility an OSN user might be releasing information on his post that could be a building block for a social engineer to capitalize on. This are some of the associated susceptible words we have used in building the model. (Bank, Account number, Birthday, Age , Email, Address, House address, Home, Gender, Mother's maiden name, Mother's first name, Father's first name, Child's name, Son's name, Pet's name, Daughter's name, Best holiday, Debit card, Credit card, Car license number, Car Registration number, Post code, Social security number, Middle name, My current location, Marital status, married, Marital status is single, Divorced, salary, Earnings, home phone number, personnel medical, financial information, place of birth, biometric records, Genotype, blood type, mobile phone number, international passport, religious preference, legal status, bank statement, invoice, drivers license, Date of birth, Paypal, Gay, Lesbian, Transgender, Homophobic, Mortgage, Loan, Job, Profession, Bank card, Career, Password)

5 Evaluation

5.1 SENS functionality and Accuracy

As a result of our research, our novel solution was able to notify us of all the hardcoded susceptible keywords and also detect Personal Identifiable Information such as credit card details and emails.

5.2 Discussion

Based on the 58 keywords and 65 expressions we have hard-coded, we have carried out an intensive test to check if our novel solution SENS does notify OSN users on potential susceptibility based on the purpose of our research. The risk ratings of SENS also conform to increase based on the number of potential Personal Identifiable Information (PII).

SENS only functions once it is enabled from the privacy settings page and does not function once it is disabled respectively. If the SENS functionality is disabled from the security settings, the SENS functionality and icons are automatically removed from the OSN. The corresponding dates on all post update automatically. SENS meter which depicts the risk rating of the potential susceptibility of the OSN users also works based on requirement. SEN meter increases by one bar for each susceptible keyword included by an OSN user in their post. The SENS risk meter also has priority keywords based on value in determining risk level. For the model, we have prioritized credit card details, email address, home address, password, date of birth and bank details due to its high importance. This keywords and details take precedence over other potential PII and are always considered high risk regardless of the keyword count.

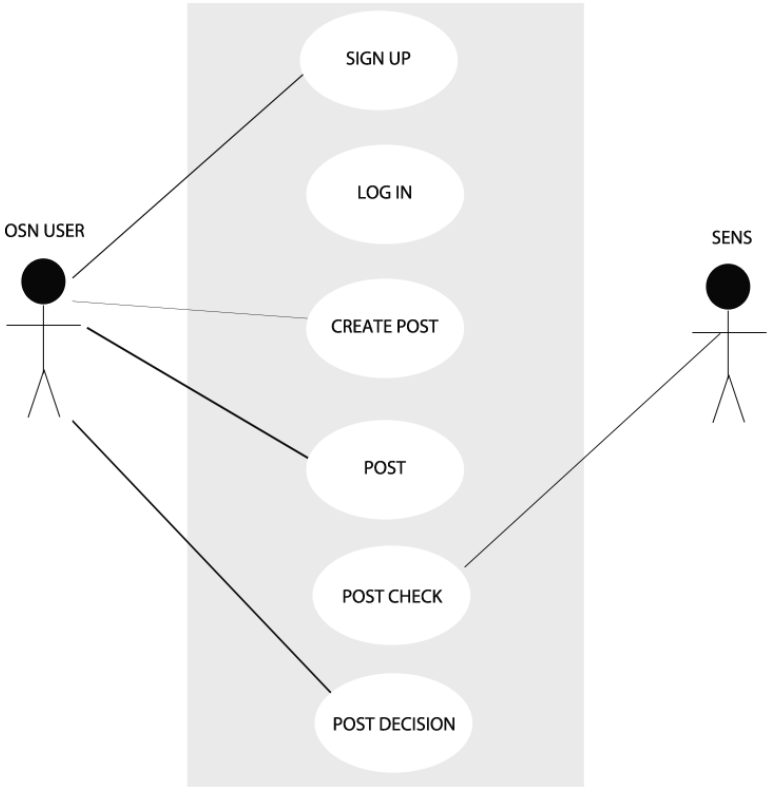

Figure 4: Use Case for SENS

5.3 Use and Test Case Testing

We carried out a Use and Test Case among 10 users, we did this by deploying SENS in five different systems. We had explained the objective of SENS and handed them the configuration guide containing all the susceptible keywords in relation to Personal Identifiable Informations. We also had the users create an account in order to enable the test. The figure below depicts the registration page used by each participant.

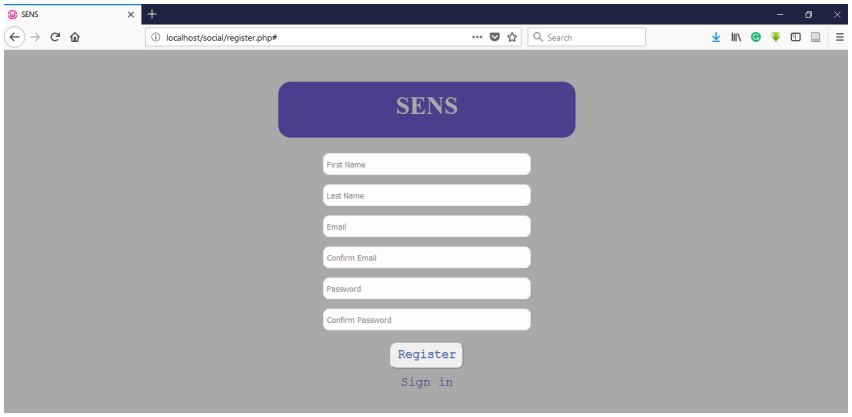

Figure 5: Account creation for SENS

After account creation, each user went next to the Log in page. They used their email address and password initially used to register to login. The next figures depict the registration page they utilized and the home page of SENS.

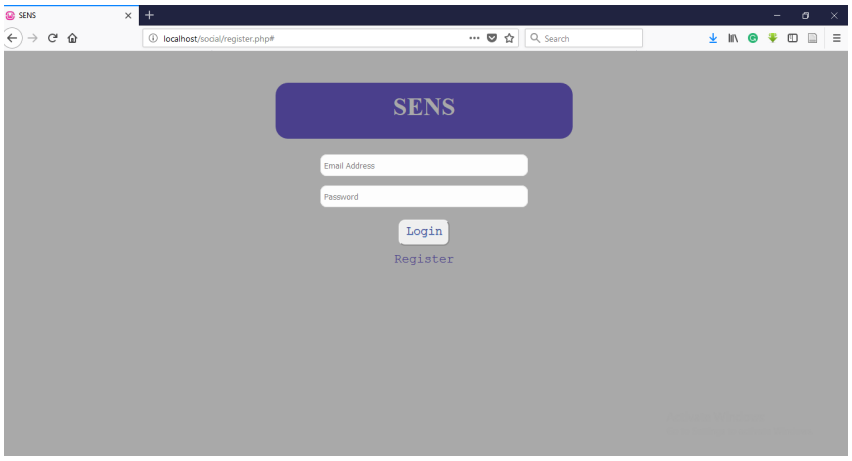

Figure 6: Login Page for SENS

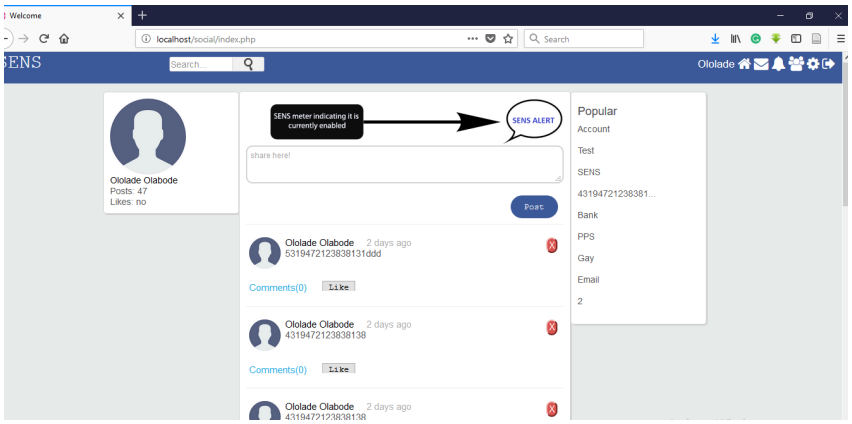

Figure 7: Home Page for SENS

The next step the participants took was to check SENS was enabled in the Privacy and security settings of the social platform before posting. The image below depicts the security page and where they had check if SENS was either enabled or disabled.

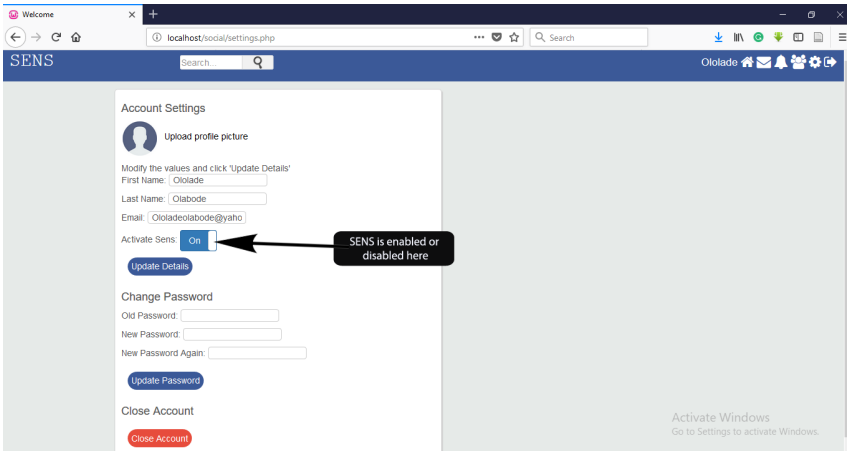

Figure 8: Privacy Page for SENS

After this was confirmed, the actors went on to attempt to post susceptible keywords from a pool of over 100 susceptible keywords that SENS checks for based on Personal Identifiable Information (PII). They had also taken into cognizance the response time of SENS while posting and it was gathered SENS provided results immediately there was an attempt to post and update. The figure below depicts the result of the output from a post.

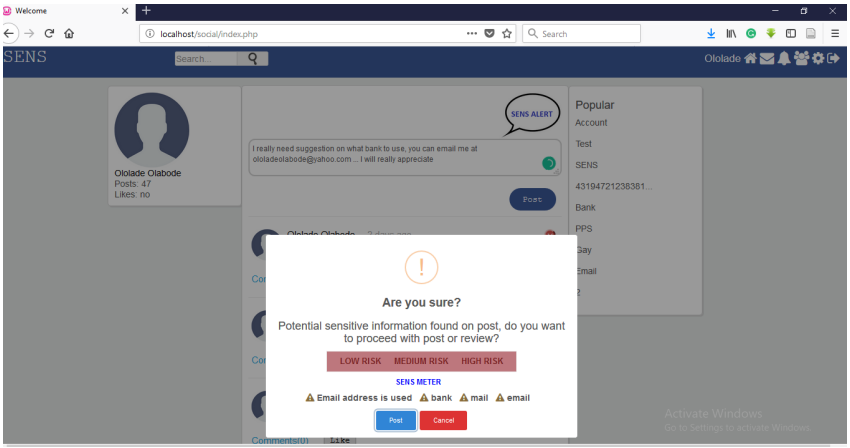

Figure 9: Output from SENS

Our participants have tested SENS on five different systems and on Mozilla, Chrome, and Internet explorer browsers also to see the flexibility of SENS. The essence of both test we carried out was to ensure the objectives and purpose of SENS was met based on our predefined scope. Through the test carried out, we were able to discover a few bugs that were immediately corrected which afterwards we did a rerun of the test case based on the initial anomalies discovered and all functioned according to the scope and objective of SENS. At the end of the test, SENS was seen to check for susceptibility to

social engineering based on vulnerable keywords and exposure of Personal Identifiable Information PII.

6 Conclusion and Future Work

Based on the research we have carried out, we believe since the basic feature required by a social engineer to execute an attack are typically Personal Identifiable Information, if we mitigate against PII or SPI dissemination from the root communicator which are the users, it will be difficult from social engineers to gain information hence almost impossible for them to launch an attack. In as much as the OSN is of great benefit to its users, we consider it to be a double-edged sword. In our research, we have been able to show the potential of SENS to notify users on likely vulnerability based on Personal Identifiable Information (PII) and categorize them into various risk level using the SENS meter. We also understand based on related works, the more a user shares sensitive personal information, the more susceptible they are to Identity theft, burglary and spear phishing attacks. One major limitation I encountered in developing SENS is the fact we could not get people to sign up to use the social platform. This would have enabled us further to integrate existing machine learning algorithm to check behavioral pattern over a period.

6.1 SENS as a Plug-in for Social Media Online Privacy setting

We look forward to SENS being integrated into the Online privacy settings in the Online Social Network (OSN). We believe SENS will cut down the release of Sensitive Personal Information (SPI) on the OSN. We believe with the use of machine learning, we can create an algorithm for SENS to be more contextual. This will improve SENS by minimizing the false positives in the discovery of SPI before notifying the OSN user. To succeed in this, we will have to make use of a live social network website to study the behavioral pattern of 200 participants over a period of two weeks with each user posting at least 10 times daily. The results of this will help train the model to enable it to be more contextual in notifying users based on behavioral pattern and susceptibility over a certain period. Security is paramount, and we believe with SENS we can mitigate against social engineering attacks.

6.2 SENS as a Web browser extension

We foresee a future for SENS being a plug-in with compatibility with any browsers which would be able to identify Sensitive Personal Information being attempted to be sent while a user utilizes the web browser. SENS web browser extension will be able to notify the browser users regardless of any website they are on as far as it is on the browser in real time, the algorithm will be able to alert users to have a double check to ensure they indeed want to post or input a potential Personal Identifiable Information. The role of the plug in will be as a threat control factor to mitigate against social engineering. SENS wont only warn you, but also give more information about the potential threats the PII can be used for. We have noticed based on human behavior, we throw caution sometimes and input our PII in websites due to the freebies or offers in general and most time do not have a clue what their PIIs are used for. According to Finklea (2017), social engineers penetrate the digital world to steal PII and use in identity theft, credit card fraud and intellectual property fraud. SENS would be able study the users behavioral pattern on

that browser and report anomalies to you when seen. The technology we believe will enable this will be using Artificial Intelligence (AI).

6.3 SENS as a Service for Enterprise

We look towards SENS being able to be offered to small, medium and large-scale organizations as a service, where they could have access to it on-demand and be able to tailor it to notify them of susceptibilities to threat of releasing specific Sensitive Personal Information (SPI) based on their business model. We believe a business should be able to use SENS as a Service to prevent exposure of Sensitive Business Information to the reach of an attacker.

Acknowledgment

Firstly, I will like to God for strength and his grace while conducting this research. I will also like to thank my supervisor Rohan Singla for his support during this research and also the entire Cyber Security department. Finally, I will like to thank my family and friends for their support during this sojourn.

References

Algarni, A., Xu, Y., Chan, T. and Tian, Y.-C. (2013). Social engineering in social networking sites: Affect-based model, *8th International Conference for Internet Technology and Secured Transactions (ICITST-2013)*, pp. 508–515.

Alim, S., Neagu, D. and Ridley, M. (2011). Axioms for vulnerability measurement of online social network profiles, *International Conference on Information Society (i-Society 2011)*, pp. 241–247.

Bilge, L., Strufe, T., Balzarotti, D. and Kirda, E. (2009). All Your Contacts Are Belong to Us: Automated Identity Theft Attacks on Social Networks, *Web Security* p. 10.

Boyd, D. M. and Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship, *Journal of Computer-Mediated Communication* **13**(1): 210–230.
URL: *http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2007.00393.x/abstract*

Chen, K. and Rea Jr, A. I. (2004). Protecting personal information online: A survey of user privacy concerns and control techniques, *Journal of Computer Information Systems* **44**(4): 85–92.

Choudhury, M. D. and Counts, S. (2012). The Nature of Emotional Expression in Social Media: Measurement, Inference and Utility, p. 10.

Fan, J., Han, F. and Liu, H. (2014). Challenges of big data analysis, *National science review* **1**(2): 293–314.

Finklea, K. (2017). Specialist in Domestic Security March 10, 2017, *Dark Web* p. 19.

Flanagan, D. (2011). JavaScript: The Definitive Guide, p. 1098.

Gleave, E., Welser, H. T., Lento, T. M. and Smith, M. A. (2009). A conceptual and operational definition of'social role'in online community, *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on*, IEEE, pp. 1–11.

Golbeck, J., Robles, C., Edmondson, M. and Turner, K. (2011). Predicting personality from twitter, *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, IEEE, pp. 149–156.

Grazioli, S. (2004). Where did they go wrong? an analysis of the failure of knowledgeable internet consumers to detect deception over the internet, *Group Decision and Negotiation* **13**(2): 149–172.

Gritzalis, D., Kandias, M., Stavrou, V. and Mitrou, L. (2014). History of information: the case of privacy and security in social media, *Proc. of the History of Information Conference*, pp. 283–310.

Gross, R. and Acquisti, A. (2011). Information Revelation and Privacy in Online Social Networks (The Facebook case), p. 11.

Herley, C. (2010). So Long, And No Thanks for the Externalities: The Rational Rejection of Security Advice by Users, p. 12.

Howcroft, D. and Carroll, J. (2000). A Proposed Methodology for Web Development, p. 8.

Iacovos, K. and Sasse, M. A. (2012). Security education against phishing: A modest proposal for a major rethink, *IEEE Security & Privacy* **10**(2): 24–32.

Jagatic, T. N., Johnson, N. A., Jakobsson, M. and Menczer, F. (2007). Social phishing, *Communications of the ACM* **50**(10): 94–100.

Kaplan, A. M. and Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media, *Business Horizons* **53**(1): 59–68.
URL: *http://linkinghub.elsevier.com/retrieve/pii/S0007681309001232*

Khonji, M., Iraqi, Y. and Jones, A. (2013). Phishing Detection: A Literature Survey, *IEEE Communications Surveys & Tutorials* **15**(4): 2091–2121.
URL: *http://ieeexplore.ieee.org/document/6497928/*

Krasnova, H., Koroleva, K. and Veltri, N. F. (2010). Investigation Of the Network Construction Behavior on Social Network Sites, p. 21.

Krishnamurthy, B. and Wills, C. E. (2009). On the Leakage of Personally Identiable Information Via Online Social Networks, p. 6.

Lee, J. and Lee, Y. (2002). A holistic model of computer abuse within organizations, *Information Management & Computer Security* **10**(2): 57–63.
URL: *https://www.emeraldinsight.com/doi/abs/10.1108/09685220210424104*

Mataracioglu, T. and Ozkan, S. (2011). User Awareness Measurement Through Social Engineering, p. 7.

Misra, G. and Such, J. M. (2017). PACMAN: Personal Agent for Access Control in Social Media, *IEEE Internet Computing* **21**(6): 18–26.
URL: *http://ieeexplore.ieee.org/document/8114620/*

Mitnick, K. D. and Simon, W. L. (2011). *The art of deception: Controlling the human element of security*, John Wiley & Sons.

Mohebzada, J. G., Zarka, A. E., Bhojani, A. H. and Darwish, A. (2012). Phishing in a university community: Two large scale phishing experiments, *2012 International Conference on Innovations in Information Technology (IIT)*, pp. 249–254.

Nurse, J. R. C. (2015). Exploring the Risks to Identity Security and Privacy in Cyberspace, *XRDS: Crossroads, The ACM Magazine for Students* **21**(3): 42–47.
URL: *http://dl.acm.org/citation.cfm?doid=2752549.2730912*

Pwint Oo, S. H. (2013). Intelligent Access Control Policies for Social Network Site, *International Journal of Computer Science and Information Technology* **5**(3): 183–190.
URL: *http://www.airccse.org/journal/jcsit/5313ijcsit15.pdf*

Randazzo, M. R., Keeney, M. and Kowalski, E. (2007). Insider Threat Study: Illicit Cyber Activity in the Banking and Finance Sector, p. 37.

Schmidt, R., Lyytinen, K., Keil, M. and Cule, P. (2001). Identifying Software Project Risks: An International Delphi Study, *Journal of Management Information Systems* **17**(4): 5–36.
URL: *https://www.tandfonline.com/doi/full/10.1080/07421222.2001.11045662*

Shehab, M., Squicciarini, A., Ahn, G.-J. and Kokkinou, I. (2012). Access control for online social networks third party applications, *Computers & Security* **31**(8): 897–911.
URL: *http://linkinghub.elsevier.com/retrieve/pii/S0167404812001186*

Sommerville, I. (2010). Software Engineering, p. 790.

Wilcox, H., Bhattacharya, M. and Islam, R. (2014). Social engineering through social media: an investigation on enterprise security, *International Conference on Applications and Techniques in Information Security*, Springer, pp. 243–255.

Williams, K., Boyd, A., Densten, S., Chin, R., Diamond, D. and Morgenthaler, C. (2009). Social Networking Privacy Behaviors and Risks, *Pace University, USA*.