PhishBook: Are you an Ed, Edd or Eddy?

Sunpreet Arora
IIIT-Delhi
New Delhi
sunpreet08053@iiitd.ac.in

Rajat Vikram Singh
IIIT-Delhi
New Delhi
rajat08044@iiitd.ac.in

Prateek Gaur
IIIT-Delhi
New Delhi
prateek08039@iiitd.ac.in

Tuhinanshu
IIIT-Delhi
New Delhi
tuhinanshu08056@iiitd.ac.in

ABSTRACT

PhishBook is a cross-network user-profiling study wherein the social networking site Facebook was mined to collect publically available information of subjects to identify their "group of friends" and gain useful insights to increase the yield of a phishing attack. This involved identifying the susceptibilities of a subject towards people of same network, gender and with mutual friends on Facebook by creating fake as well as duplicate identities and thereby duping the subjects to gain access to their hidden information. This information was then further used in structuring the phishing attacks for profiling the users based on their vulnerabilities. Overall, it was found that the human tendency is to fall for social phishing attacks within their "group of friend" and also more the closeness of nodes in the social graph, more is the probability of them falling for such attacks. Gender was found to have a significant impact on the male segment only.

General Terms

Design, experimentation, security, human factors

Keywords

Phishing, social engineering, user-profiling, email, real-world studies, personally identifiable information (PII), context aware phishing

1. INTRODUCTION

Social engineering is the act of manipulating people into performing actions or divulging confidential information, rather than by breaking in or using technical cracking techniques [1]. It relies heavily on human interaction and the fact that people are not aware of the value of the information they possess and are careless about protecting it. 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 [2]. In other words, people are tricked into revealing some personal facts which could potentially be misused for malicious purposes by the attacker.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Conference'10, Month 1–2, 2010, City, State, Country.
Copyright 2010 ACM 1-58113-000-0/00/0010…$10.00.

There are various targeted forms of phishing emerging today such as spear phishing wherein the apparent source of the email is likely to be an individual within the recipient's organization and generally a person of authority, vishing which exploits the public trust on traditional telephone systems etc. wherein different types of *context aware* [3] information is identified to be used prior to launching a phishing attack. This information can be collected in multiple ways such as using public records, websites, publically available databases, or other physical world information [1]. However, the modern day trend of social networking on the internet is one of the prime reasons which motivated us to use a different approach for data collection. Therefore, our study primarily relied on collecting such information through the social networking site Facebook (facebook.com).

The interest group of our study was restricted to the subset of people connected to our own profiles on the social network but with hidden Personally Identifiable Information (PII) to their non-friends. The type of information collected included the network to which the subject belongs, his/her gender and mutual friends. The core "groups of friends" were identified primarily based on the network to which they belonged. The subjects were then randomly chosen for our study from each "group" of friend.

The next step involved creating fake and duplicate identities on facebook taking into consideration various different heuristics such as same/different "group of friend", same/different gender and with/without mutual friends. The subjects were then added as friends of these fake identities so as to gain access to the information which they limit to being viewed by their friends only. The first level of user-profiling was done based on whether the subjects accepted or rejected these proposals to be friends. Thereafter, subject specific phishing emails were engineered to measure the susceptibility of the users who fell for being friends with these fake identities. The final step of user-profiling therefore involved differentiating between subjects who clicked on the links in the phishing email and those who didn't.

Besides user-profiling, conclusions were drawn about the general susceptibility of subjects to profiles of same/different gender, same/different "groups of friends" and having/not having mutual friends, as well as context specific phishing emails. Subjects were found to be more susceptible to social phishing attacks from within the same group of friend. Males were found to be significantly vulnerable to such attacks from opposite gender, whereas no such pattern was observed amongst the females. Also, the lesser the distance metric (hop count), the more easy it was to deceive subjects.

The section 2 of the report discusses some relevant studies done in the context of social engineering and phishing. Section 3 describes in detail the approach used in our study. Section 4 primarily aims at analyzing the data collected whereas Section 6 deals with the main conclusions drawn from the study. A new phenomenon of "Profile Honeypots" was observed wherein some identities could be intentionally created with open information by an attacker so that people automatically add them to their social network, thereby compromising on their personal information. This has been discussed in Section 5. Section 6 concludes the study, whereas Section 7 describes the limitations and challenges we faced during the course of the research study. Section 8 deals with this research study can be extended further.

2. BACKGROUND

This section presents a brief background of social phishing and the other forms of so-called *context aware phishing*. It also highlights some key points of the deception theory. The section finally concludes with some results from related empirical studies on the topic of social phishing.

2.1 Social Phishing

The phenomenon of phishing has been on the internet for quite a while. It is a form of social engineering in which an attacker masquerades as some trust-worthy party to extract sensitive and confidential information from a victim. The amount of information extracted depends on the context in which it is carried out. For example, a phisher misrepresenting as a trustworthy friend to demand money needs know much about the recipient in order to get usable results while the scenario is totally opposite in case of e-commerce frauds where (if successful) the attacker will get reasonable results without much or even any prior knowledge about the victim.

But as anti-phishing technologies are increasing, attackers are also equipping their arsenal with new technologies often referred to as *context aware phishing* [3], where the attacker gains trust of the victim by obtaining information about them like their bidding history or shopping preferences freely available on such sites or from banks, etc. There are other such attacks which take advantage of both technical and social vulnerabilities. A list of the most commonly occurring attacks and countermeasures can be found in [4].

Social Phishing is a form of context aware phishing where phishing attacks are trained by using the publicly available personal information from *social networks*. The approach is how effectively it can help in increasing the yield of phishing in general. Gartner [5] showed that about 19% of all those surveyed reported having clicked on a link in a phishing email, and 3% admitted to having given up financial and personal information [3]. All the social networking sites use *networks of friends* which can help any phisher to gain heavy amount of information from the social network data. But the *privacy* feature of these social networks poses a great problem for the attackers. Hence one of the best approaches for *anti-social phishing* could be to modify the privacy policies of such sites in order to educate the users so that they are able to protect themselves against such attacks.

2.2 Deception Theory

Masquerading in order to fraudulently retrieve information is an important approach to any phishing technique. R.C Miller [6] showed that phishers exploit the difference between the system model and users' mental model to deceive and victimize users. This concept is often referred to as *deception.* Deception is generally defined as "a message knowingly transmitted by a sender to foster a false belief or conclusion by the receiver" [7].

Communication literature suggests that many cues influence users when making trust decisions, including (1) verbal cues (e.g. language style, message content in the email); (2) non-verbal cues (e.g. timestamp of an email); and (3) contextual cues (e.g. feedback from toolbars) [8,9] and phishers spoof these cues to deceive the victims who rely on such cues. There are several methods to detect deception which have been discussed in [10].

2.3 Related Work

There are some published real-world studies that deal with the phenomenon of social phishing or overall effectiveness of the context of phishing. Jagatic et al. studied the vulnerability of a university community towards a phishing email that pretends to come from somebody in their own social network, but did not study the effectiveness of study. Sending fake emails to test a user's vulnerability has been studied one of the most frequently studied phenomenon [11,12,13]. Recently, the United States Department of Justice sent their employees fake phishing emails to test their vulnerability to phishing [13].

None of these consider the question of user-profiling on the basis of susceptibility towards social phishing attacks. In our study, we tried to invade the privacy of the user and his friends by accessing their PII to get their "groups of friends" and using this information to create fake identities and send phishing emails using deception to test the overall vulnerability. User profiling has been done on the basis of how different users react to various sets of experimental emails as well as assimilating those fake identities into their network.

A similar study has been done in Indiana [2] where a database of relationships was created by crawling several social-network websites. A phishing attack was then performed on college students aged 18-24 years old and analysis was performed on how many victims were found on respective days. However, PhishBook was limited to one social-networking site i.e. Facebook, but involved profiling of users on the basis of their reactions to a series of social phishing attacks into three different categories. There exists research to show that social phishing is effective for user-vulnerability analysis but it does not address the current Indian scenario which is also another driving force behind PhishBook.

3. METHODOLOGY

The study primarily focused on three different "groups of friends" and the subjects for the experiment were chosen randomly from each group. Then a certain set of key observations about the subjects were made by mining their social network on facebook. However, they were given different contextual treatment depending on their mutual friends, networks and gender. The next set of observations was made for user-profiling depending on the reaction of the subjects to the type of

treatments to which they were exposed to. This type of design was done to preserve the main context specific nature of social phishing whilst still being able to do a comparative analysis. This was designed as a within-subjects study simply because one subject could belong to two different "groups of friends" depending on the network he/she belonged to.

Table 1: The time frame based table showing a broad view of the methodology followed.

R	**O1**	**X1**	**O2**
R	O1	X2	O2
R	O1	X3	O2

3.1 Network Mining

Since the core interest group of the study were the people connected to our own profiles on Facebook, therefore the initial step of data collection involved gathering publically available information about them so as to gain a reasonable amount of information about their background. This step involved *targeted content mining* in the sense that information extraction was limited only to the parameters which were chosen for our study such as gender, network and mutual friends. A two-pronged approach was used for this purpose:

3.1.1 Using the Facebook API

The primary step involved developing a standalone Windows application using the Facebook API as shown in figure 1. The main aim of this technology was to get information which individuals share with other individuals on the social network. Since, the concerned individuals were limited only to our friends on the network; therefore it provided us an easy way to get some discrete content shared by them such as their date of birth, gender, network etc.

However, due to the limitations imposed by the Facebook API, sufficient desirable information couldn't be assimilated thereby initiating the need to resort to other methods for data collection.

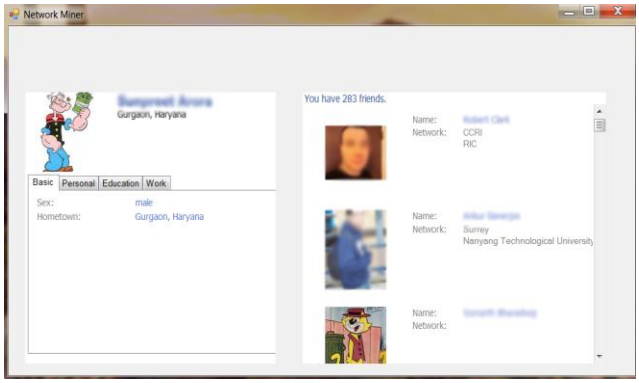


Fig 1: The Windows application used to mine the network

3.1.2 Web Content Mining

The second approach for data collection involved the traditional web content mining, wherein the social networks of the interest group individuals were taken off-the shelf and then crawled repeatedly to ascertain other specific facts about the individuals.

This provided a much better way to find correlated information between subjects and identifying the connectivity of different nodes on the network i.e. to identify nodes on the social network which are directly connected or indirectly connected on the network and also the distance measure (hop count) of these nodes.

3.2 Identifying Groups of Friends

The second step in our approach was to identify "groups of friends". The term essentially means clusters of people who are connected on the social network under a same domain. The main parameter under which this clustering was done was the different linked networks on the Facebook profile of an individual. This led to the identification of three broad networks into which most of the individuals could be placed. Note, however that one individual could belong to more than one of those identified networks. Once, this grouping was done, random sampling to select subjects from within each cluster was done. Each of those subjects' already mined information then formed the core of the next steps.

3.3 Fake Profile Creation

The earlier steps involved essentially the preprocessing steps to fine tune the heuristics for the attacks. Once sufficient background information had been gathered, we had to profile users based on their susceptibility. The motive was to compromise on the hidden user information limited to being viewable only to the friends of a particular subject on the social network. This was essentially accomplished through a bimodal strategy:

3.3.1 Duplicate Profiles

One of the strategy was to duplicate our own profiles and observe how many of the subjects still confirm being friends with them even when they are already friends with their original versions. This is one possible loophole which could be exploited by an attacker to compromise on the privacy of a mutual friend of his/her friends on the social network.

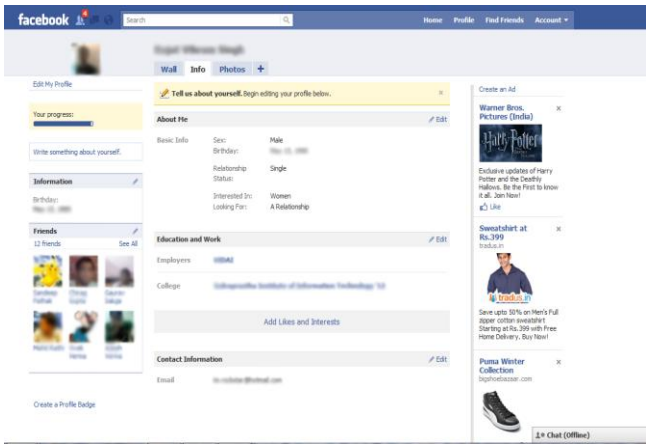


Fig 2: One of our own duplicate profiles

3.3.2 Fake Profiles

Another strategy was to create fake identities/nodes on the social network each possessing certain different set of attributes and observe how many subjects selectively choose being friends with them. This could be used as another possible way to invade the privacy of users based on their perception of trust towards certain key attributes such as mutual friends and similar network over others.

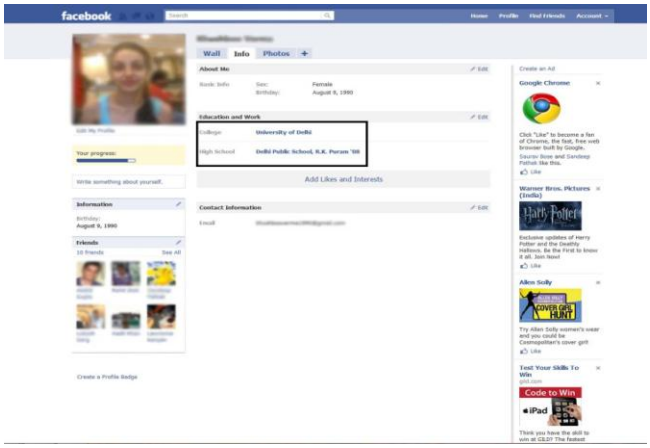


Fig 3: Fake profile containing a set of attributes common to one particular network.

3.4 Social Phishing Attacks

The second level of user profiling was done using phishing attacks on the subjects who accepted friend requests from fake profiles. The basic idea here was to first retrieve the PIIs of the subjects as much as possible and then exploit this information to do a social phishing attack on them. The partially hidden information on Facebook can reveal many details about the user and give a basic idea to the attacker about things the user can fall for. This gives the attacker additional details to plan the phishing attack.

3.4.1 Framing Phishing Mails

Phishing mails were framed keeping in mind the various factors such as the interests and the alignment of each subject towards different fields such as music, sports etc. Moreover, the relationship status of the user was also a very important factor. All this information was carefully analyzed and then a strategy was devised to deceive the subjects by exploiting this information.

For example, if the subject was single and looking for a relationship, the phishing email was sent from a fake profile of opposite sex that he/she wants to be in relationship with him/her on the social network or if the subject was more inclined towards music then a link of popular music was shared with him/her on the email.

The mails had embedded links within them to redirect the user to a different server before taking them to the original one so as to record information about them falling for the phishing email.

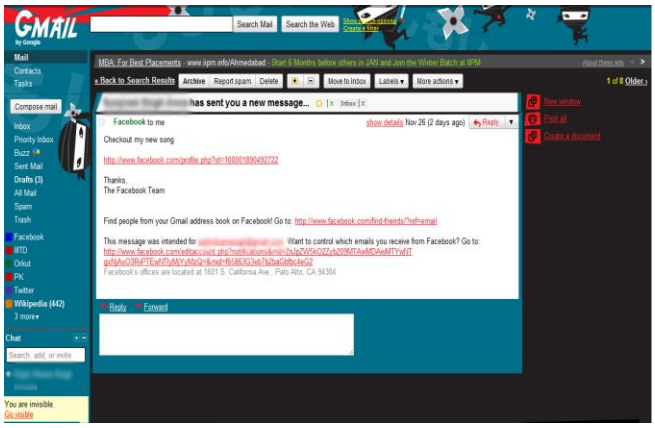


Fig 4: A sample phishing email used.

3.4.2 Storing Results

The server which was used as the redirection point of all the mails, a comma separated value file was kept there to store all the results such as IP Address, email of the user and some other information such as the operating system of the user, the time of clicking on the mail etc. This file yielded some important results which will be discussed in the next section.

4. ANALYSIS/RESULTS

The primary aim of our study was user-profiling and the data collection process already described above was aligned towards the same. We evaluated different parameters to draw conclusions about their vulnerabilities towards different types of social phishing attacks and used the results to classify users into three different categories. This section deals in entirety with the analytical process undertaken.

4.1 User Demographics

The total subjects for our study were 114. Since three major "groups of friends" were identified during the preprocessing, we focused on subjects from within one or more of these groups. Overall, there were 27 subjects in the first group, 52 subjects in the second, and 35 subjects in the third group. Since the subjects drawn were from our own social network, therefore the age range for most of them was 18-25 years. The subjects were random sampled, however under the constraint that the sample contains subjects from either gender in reasonable ratio according to each "group of friend".

4.2 Quantitative Analysis

This dealt with quantifying the results so as to be able to draw conclusions from the collected data. This was done on the basis of different parameters:

4.2.1 Based on Groups of Friends

The block level influence of social connectivity on the susceptibility of subjects to fall for social phishing attacks was the first heuristic taken into consideration. Some of the key results of the analysis were:

- Almost 71% of them tend to accept friend requests which appear to be coming from a person of the same group of friend.
- Out of these, 37% of the subjects fall for phishing emails sent to the email IDs which they use for logging onto the social networking site.
- Also 54% of the subjects accepted friend requests without being biased by the network of the sender, and 39% of this subset clicked on the links.

Based on the discussion, we can clearly see that there was a little inclination towards accepting friend requests from the entities in the same network then those not in their network, but when it comes to falling for phishing there was no stark contrast percentage of people i.e. 37-39%.

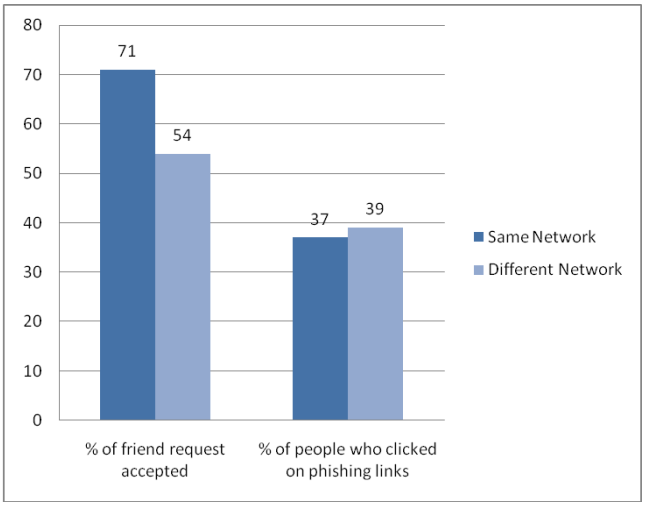


Fig 5: Bar chart showing the percentage of people who fell for different social phishing attacks based on groups of friends

4.2.2 Based on Gender

Another prominent factor studied was the reaction of subjects towards social phishing attacks from the same/different gender. The data collected was analyzed based on how subjects tend to react to social phishing attacks from the same gender and vice-versa and the following set of results were obtained:

- 40% of the subjects accepted friend requests from people of the same gender, while a whopping 82% of them accepted friend requests from people of the opposite sex. The number though was inflated by the male segment of the subjects.
- Out of the subjects who accepted friend requests of the same gender, 40% fell for the phishing attack.
- Out of the subjects who accepted friend requests from the opposite gender only 37% fell for phishing.

This discussion gives a very non-intuitive result as the subjects accepting friend requests of the opposite gender are less reactive to phishing by them while subjects of the same gender have a moderate falling tendency towards phishing.

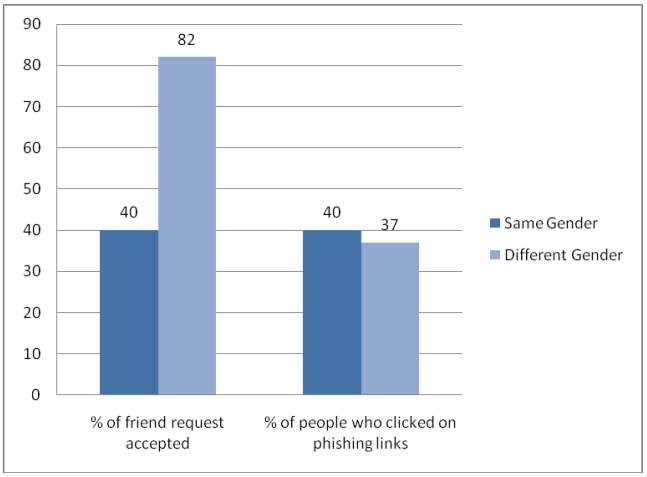


Fig 6: Bar chart showing the percentage of people who fell for different social phishing attacks based on groups of friends

4.2.3 Based on Mutual Friends

The third criteria considered was the susceptibility of subjects to attacks originating from different hops on the social network. The hops considered in our case were the first and the second ones i.e. friend and friend of a friend.

- About 73 percent of the subjects accepted friend requests having origins in the first hop whereas the percentage to only 35 percent for the second hop.
- Out of those who accepted friend requests from the first hop, about 38 percent fell for phishing attacks, whereas only a meager 15 percent were deceived at the second hop.

This clearly brings out that more closely knit two nodes are on the more social network, more the likelihood that they would be victims to social phishing attacks originating from either of them.

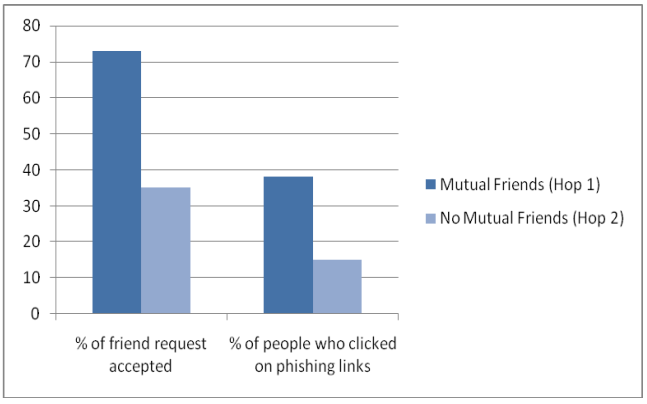


Fig 7: Bar chart showing the percentage of people who fell for different social phishing attacks based on mutual friends

4.3 User Profiling

Since our methodology followed two distinct approaches, user-profiling can be done based on whether the subject fell for a particular kind of attraction towards particular parameters of a

fake or duplicate identity created on the social network and then in turn for the phishing attack made once his email was compromised through the earlier approach. The classification basically revolves around the different combinations of giving in to these temptations:

4.3.1 The Ed

The subjects who accepted the friend requests on Facebook and then the context specific phishing email sent to them or in other words were found to be the most vulnerable to the social phishing attacks were called the Eds. About 24.5 percent of the subjects were found to fall into this category.

4.3.2 The Edd

The subjects who accepted the friend requests on Facebook, but at the same time did not fell for the phishing email were the intermediary ones and were called the Edds. About 59.5 percent of the subjects i.e. a majority of them fell into this category.

4.3.3 The Eddy

The subjects who did not give in to either of these temptations were the smartest ones, hence called the Eddys. 36 percent of subjects fell into this category.

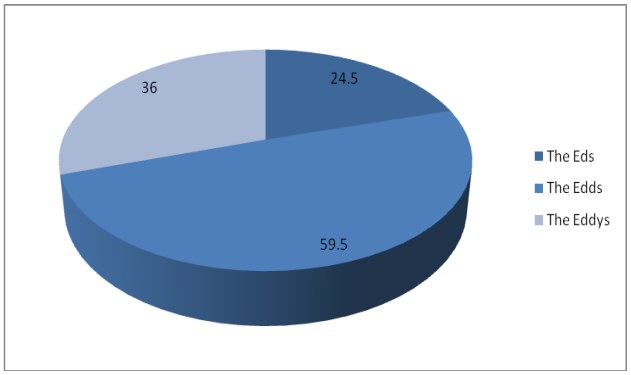


Fig 8: Pie chart showing the percentage of Eds, Edds and Eddys amongst the subjects

5. DISCUSSION

Creating fake profiles and sending phishing emails all required deception in order to acquire sensitive information from those who fall a victim for such attacks. While carrying out such experiments some popular notions were taken into consideration like people often do not suspect anything they get from their network of friends, people are often fast in reacting to mails and requests from the opposite gender, etc. Such contexts on one hand helped in improving the quality and quantity of results and on the other hand helped us in noticing some important phenomena. Instead of sending network requests at a constant rate, regular network requests were being received on an attractive female's profile even from people who were not in her network of friends but the case was totally opposite in case of most of the male profiles where even the sent requests were not being accepted. Such profiles have varied applications in various fields particularly in security which prompted us to coin a term called *Profile Honeypots*.

5.1 Profile Honeypots

The concept of *honeypots* was introduced in computing systems by Clifford Stoll in the late 80's. In the 'Cuckoo's Egg', he described the monitoring and tracking of an intruder [1,2]. In the 90's, Cheswik had implemented and deployed a real *honeypot* [3]. Lance Spitzner, a senior security architect for Sun Microsystems is the author of "Honeypots, Tracking hackers" [4]. In his book he describes honeypot and its usage as "A honeypot is a security resource whose value lies in being probed, attacked or compromised. This means that whatever we designate as a honeypot, our expectations and goals are to have the system probed, attacked, and potentially exploited. It does not matter what the resource is (a router, scripts running emulated services, a jail, an actual production system). What does matter is that the resource's value lies in its being attacked. If the system is never probed or attacked, then it has a little or no value. This is the exact opposite of most production systems, which you do not want to be probed or attacked." [4, page 40]

Similarly in our study instead of manually sending network requests we noticed that some particular profiles automatically attracted dramatically large amount of victims while other profiles were not even getting their requests approved. As Lance [4] mentioned that the essential property of honeypots should be to get probed, attacked or compromised, hence such profiles essentially serve as *honeypots* and served our purpose of attracting victims to our profile so that we can fraudulently extract their network information. This phenomenon has many applications in varied fields which require honeypots hence serve as *honeypot profiles*. The applicability can be classified into two categories- positive and negative. Some of its positive applications could be for analytical purposes where we face problems of sparse datasets, hence adding another dimension of social network can prove to be fruitful. The dataset collected can help organizations like *Perverted-Justice (PJ), etc* who suffer from dataset problems to testify against pedophiles. PJ volunteers disguise as victims in chats to catch pedophiles. But this approach has many shortfalls like small chat dataset, etc which can be easily resolved by using profile honeypots. To exploit this, institutions such as PJ can make a fake profile using victims' attributes to attract attackers which could be pedophiles in this case and hence get all the necessary details or required details to testify against them. Profile Honeypots can also be helpful for using profiles for clustering like pages. A successful methodology could be to make pages with specific attributes which can classify people into different groups. Now people who like a specific page can be divided into clusters whose centre is defined by that page. Hence it also helps in user profiling for purposes like surveying for a product or for campaigning purposes during elections where we can predict a possible winner by checking how many people are in the cluster of a particular party.

But there are various negative aspects of this context. An attacker can create a fake profile masquerading as a celebrity with a large fan following. Hence people who get spoofed by such profile honeypots will end up giving all their private information to the owner of that page. Such information can be used fraudulently against them in many ways. One of the ways

could be to perform phishing as proposed in this paper. Hence there are both positive and negative aspects of profile honeypots and their effectiveness depends on the context in which they are used.

6. CONCLUSION

We classified the users into three categories based on their susceptibility towards falling for two possible ways of privacy intrusion using the social networks. This was primarily done through *context aware phishing attack*s which are growing quite rapidly in the modern day era to deceive and victimize users in the modern day era. The wealth of information shared by people on the social networking websites provides an easy way to gain useful insights to customize phishing attacks according to each individual thereby increasing the chances of the success of those phishing attacks. This was reaffirmed by the fact that almost 38 percent of users fell for these kinds of attacks whereas traditional phishing attacks done during the pilot tests before the actual study had yielded lower success rates.

People generally fall far more easily into befriending intra-network entities on a social network than the cross-network ones. Almost 71 percent of subjects accepted the friend requests of nodes from one of their native networks whilst only 54 percent of cross network requests were accepted. Therefore, the proximity of two nodes in a social network i.e. being associated to the same network is a direct indicator of the level of trust in each other.

Gender of a person is also a decisive factor in falling for these attacks. This is especially true amongst the male gender, who would almost always accept friend requests from the opposite sex. But, amazingly the females do not show any significant affiliation towards any particular gender. The male gender also fell for phishing emails from the opposite gender far more easily than their female counterparts. Therefore, we can easily conclude that the male gender is far more prone to social phishing attacks given they manifest their origins in the opposite sex.

The level of interconnectivity within a social sub-domain or in other words “groups of friends” also yields significant influence on how users respond to friend requests as well as fall for phishing attacks. This follows a direct proportion with the proximity i.e. more closer the node, more are the chances of success of a social phishing attack originating from that node.

7. LIMITATIONS AND CHALLENGES

We focused on real world study for user classification rather than using laboratory based approach. Though real-world studies provide good dataset for both experimental and analytical purposes but we faced many challenges while conducting these. First module was analyzing facebook's social network for mining groups of friends. In this case the evolving nature of the facebook API made it difficult for us to make a customized application to solve our purpose, hence we decided to change focus on network mining rather than exploring facebook API.

Our research is only internally valid not externally and it is not a generalizable study; hence is very specific to the networks we have included. Though we have done random sampling in selection of subjects but there is a possibility of selection bias as the heuristics we followed depended on manual selection of suitable victims and there is always a chance of finding some pattern in manual efforts.

Another major challenge that we faced was in our study of phishing emails where we had to maintain victims' privacy, prevent our mails to be marked as spam or getting blacklisted and use proper deception in order to come across as a genuine sender.

In order to maintain the privacy of the recipients', we were asked not to collect any personal or sensitive information. We incorporated many ways to achieve recipients' anonymity. We shifted from collecting their email-id to assigning a hash to different recipients. Then we used different logs for different cluster of recipients and collected data in terms of “how many” instead of “who all”.

We decided to send emails just before the end semester exams as the probability of reading emails is much higher during those days and email reading behaviour is different due to increased personal pressure.

Emails sent may get marked as spam or get blocked or blacklisted by various anti-phishing tools and various browsers. To counter, a careful analysis of the detail of the mails in the inbox was done. It was found that if one clicks on show details of any mail which has been received, it contains the following fields:

- From (the sender of the mail)
- reply-to
- to (the receiver of the mail)
- date
- subject
- mailed-by
- signed-by

Among these fields, I, VI, VII are very important ones and need to be filled very carefully while sending mail through *sendmail* service otherwise the mailing sites categorize these mails as spam or a phishing mail.

As the possible targets of our study were the network of friends extracted from facebook therefore we had to make sure that friends who are separated by just 1 hop should not receive same emails format as the chances of getting caught increase. Hence for such problems we prepared multiple sets of emails for people belonging to the same network.

8. FUTURE WORK

The proposed study can be extended further to various domains by incorporating the use of other social networking sites other than facebook to improve the overall quality of phishing. It can

also be used for developing a heuristic based fake profile classifier. The classification criteria of such a tool would be to check for the attributes used by various such profiles and mark potential profile honeypots in its black list to notify the user about possible threats.

9. ACKNOWLEDGMENTS

The authors would like to acknowledge one of their colleagues Hemank Lamba for the healthy discussion with them which gave birth to the idea of deceiving people using fake profiles on a social network and the systems and network management committee of IIIT-Delhi for granting permissions to use the local resources to aid this research study. The system administrator of IIIT-Delhi also proved to be of immense help in setting up the redirection mechanism.

10. REFERENCES

[1] Goodchild, Joan (January 11, 2010). "Social Engineering: The Basics", http://www.csoonline.com/article/514063/social-engineering-the-basics

[2] Social Phishing- Tom Jagatic, Nathaniel Johnson, Markus Jakobsson, and Filippo Menczer, December 12, 2005.

[3] Markus Jakobsson. Modeling and preventing phishing attacks, In Phishing Panel at Financial Cryptography, February, 2005.

[4] Aaron Emigh, Online identity theft: Phishing technology, chokepoints and counter- measures: ITTC Report on Online Identity Theft Technology and Countermeasures, http://www.anti-phishing.org/Phishing-dhs-report.pdf, October 2005

[5] Gartner Inc, Gartner study finds significant increase in e-mail phishing attacks, http://www.gartner.com/5 about/press releases/asset 71087 11.jsp, April 2004.

[6] R. C. Miller and M. Wu, Fighting Phishing at the User Interface. O'Reilly, August 2005 In Lorrie Cranor and Simson Garfinkel (Eds.) Security and Usability: Designing Secure Systems that People Can Use.

[7] D. B. Buller and J. K. Burgoon, Interpersonal deception theory. Communication Theory, 6(3):203 – 242, 1996

[8] J. R. Carlson, J. F. George, J. K. Burgoon, M. Adkins, and C. H. White, Deception in computer-mediated communication. Group Decision and Negotiation, 13(1):5 – 28, 2004.

[9] Ponnurangam Kumaraguru, Justin Cranshaw, Alessandro Acquisti, Lorrie Cranor, Jason Hong, Mary Ann Blair, Theodore Pham, School of Phish: A Real-World Evaluation of Anti-Phishing Training

[10] A. J. Ferguson, Fostering E-Mail Security Awareness: The West Point Carronade. EDUCASE Quarterly, (1), 2005.

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

[12] New York State Office of Cyber Security & Critical Infrastructure Coordination, "Gone phishing. . . a briefing on the anti-phishing exercise initiative for new york state government. Aggregate Exercise Results for public release.", 2005.

[13] E. Spagat. Justice department hoaxes employees. News article, January 2009. http://news.yahoo.com/s/ap/20090129/ ap on go ca st pe/justice hoax.

[14] White Paper: "Honeypot, Honeynet, Honeytoken: Terminological issues2" by Fabien Pouget et al.

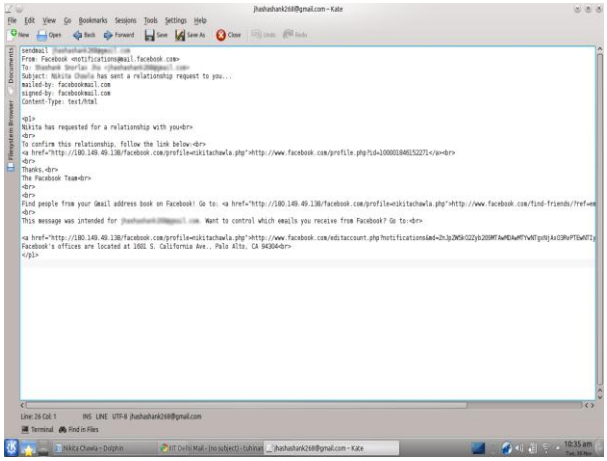
[15] C. Stoll, "Stalking the Wiley Hacker", Communications of the ACM, Vol. 31 No5. May 1988.

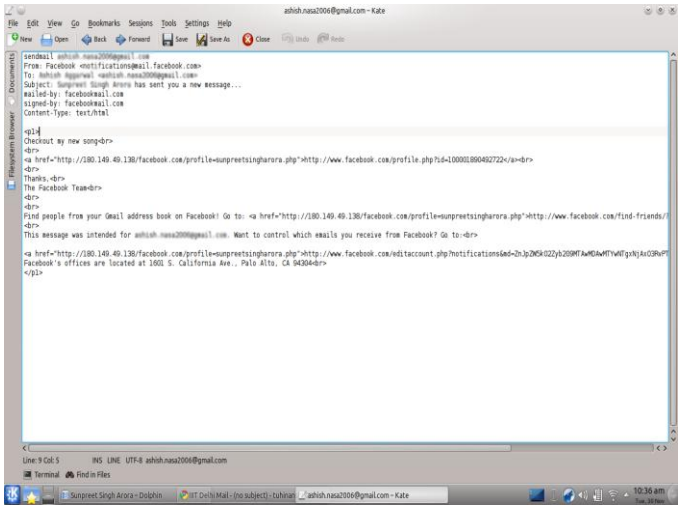
[16] B. Cheswick, "An evening with Berferd in which a cracker is lured, endured and studied", Proc Winter USENIX Conference, San Francisco, Jan 20, 1992.

[17] L. Spitzner, "Honeypots: Tracking Hackers", Addislon-Wesley, ISBN from-321-10895-7, 2002.

11. APPENDIX

Some sample screenshots of the experimental phishing emails used in the study:



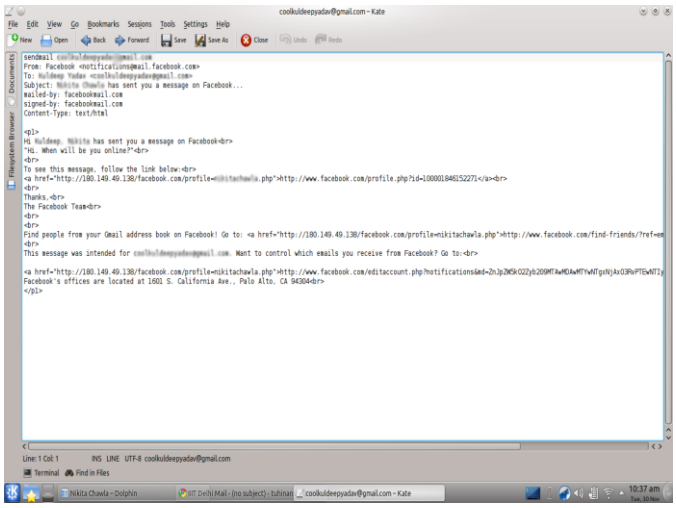

ashish.nasa2006@gmail.com - Kate
File Edit View Go Bookmarks Sessions Tools Settings Help
New Open Back Forward Save Save As Close Undo Redo
sendmail ashish.nasa2006@gmail.com
From: Facebook <notifications@mail.facebook.com>
To: Ashish Agarwal <ashish.nasa2006@gmail.com>
Subject: Sunpreet Singh Arora has sent you a new message...
mailed-by: facebookmail.com
signed-by: facebookmail.com
Content-Type: text/html
<p1>
Checkout my new song

http://www.facebook.com/profile.php?id=100001890492722

Thanks.
The Facebook Team

Find people from your Gmail address book on Facebook! Go to: <a href="http://180.149.49.138/facebook.com/profile=sunpreetsingharora.php">http://www.facebook.com/find-friends/?

This message was intended for ashish.nasa2006@gmail.com. Want to control which emails you receive from Facebook? Go to:
<a href="http://180.149.49.138/facebook.com/profile=sunpreetsingharora.php">http://www.facebook.com/editaccount.php?notifications&ad=ZnJpZW5kOZZyb209MTAwMDAwMTYwNTgxNjAxO3RvPTw
Facebook's offices are located at 1601 S. California Ave., Palo Alto, CA 94304
</p1>
Line: 9 Col: 5 INS LINE UTF-8 ashish.nasa2006@gmail.com
Terminal Find in Files
Sunpreet Singh Arora - Dolphin
IIIT Delhi Mail - (no subject) - tuhinan
ashish.nasa2006@gmail.com - Kate
10:36 am


coolkuldeepyadav@gmail.com - Kate
File Edit View Go Bookmarks Sessions Tools Settings Help
New Open Back Forward Save Save As Close Undo Redo
sendmail coolkuldeepyadav@gmail.com
From: Facebook <notifications@mail.facebook.com>
To: Kuldeep Yadav <coolkuldeepyadav@gmail.com>
Subject: Nikita Chawla has sent you a message on Facebook...
mailed-by: facebookmail.com
signed-by: facebookmail.com
Content-Type: text/html
<p1>
Hi Kuldeep, Nikita has sent you a message on Facebook
"Hi. When will be you online?"

To see this message, follow the link below:
http://www.facebook.com/profile.php?id=100001846152271

Thanks,
The Facebook Team

Find people from your Gmail address book on Facebook! Go to: <a href="http://180.149.49.138/facebook.com/profile=nikitachawla.php">http://www.facebook.com/find-friends/?ref=em

This message was intended for coolkuldeepyadav@gmail.com. Want to control which emails you receive from Facebook? Go to:
<a href="http://180.149.49.138/facebook.com/profile=nikitachawla.php">http://www.facebook.com/editaccount.php?notifications&ad=ZnJpZW5kOZZyb209MTAwMDAwMTYwNTgxNjAxO3RvPTEwNTg
Facebook's offices are located at 1601 S. California Ave., Palo Alto, CA 94304
</p1>
Line: 1 Col: 1 INS LINE UTF-8 coolkuldeepyadav@gmail.com
Terminal Find in Files
Nikita Chawla - Dolphin
IIIT Delhi Mail - (no subject) - tuhinan
coolkuldeepyadav@gmail.com - Kate
10:37 am