A

Project Report On

SOCIAL ENGINEERING ATTACK BY PROFILE HIJACKING AND USING HUMANS AS BOTNETS

Submitted By

Mayank Jain	**B8058578**
Rahul Patil	**B8058598**
Anjali Shira	**B8058622**

Under the guidance of

Prof. S.C. Dharmadhikari

In partial fulfilment of

Bachelor of Engineering

[B. E. Information Engineering]

[May 2012]

AT

Department of Information Technology

Pune Institute of Computer Technology

Dhankawadi, Pune 411043

Affiliated to

University of Pune

Pune Institute of Computer Technology

Department of Information Technology
Dhankawadi, Pune 411043

CERTIFICATE

This is certify that the Dissertation entitled **"Social Engineering Attack By Profile Hijacking and Using Humans as botnets",** submitted by **Mr. Mayank Jain** is a record of bonafide work carried out by him, in the partial fulfilment of the requirement for the award of Degree of Bachelor of Engineering (Information Technology) at Pune Institute of Computer Technology, Pune under the University of Pune. This work is done during year 2011-2012, under our guidance.

_____

(Prof. S.C. Dharmadhikari)

Project Guide

_____ _____-

Prof. Emmanual M. Dr. P. T. Kulkarni

HOD, IT Department **Principal PICT**

Examination:

Examiner _____

Date:

Acknowledgements

I am profoundly grateful to **Prof. S.C. Dharmadhikari** for his expert guidance and continuous encouragement throughout to see that this project rights its target since its commencement to its completion.

I would like to express deepest appreciation towards **Dr. P. T. Kulkarni**, Principal PICT, Pune, **Prof. Emmanual M.** HOD Information Technology Department and **Prof. Manish R. Khodaskar** (Project Coordinator) whose invaluable guidance supported me in completing this project.

I am particularly grateful to **Dr. Navin Kabra** (Punetech) who allows me to work in the company.

At last I must express my sincere heartfelt gratitude to all the staff members of Information Technology Department who helped me directly or indirectly during this course of work.

Mayank Jain
Rahul Patil
Anjali Shira

CERTIFICATE

This is to certify that the project report entitled

Social Engineering Attack By Profile Hijacking and Using Humans as botnets

Submitted by

Mayank Jain B8058578
Rahul Patil B8058598
Anjali Shira B8058622

is a bonafide work carried out by them with the Sponsorship from ————————- under the supervision of Mr. ..... and has been completed successfully .

(Mr. .....)
(Designation)
External Guide

Place : Pune
Date:

Contents

List of Figures

Chapter 1

Introduction

1.1 Need

1.1.1 Social Engineering

Social Engineering is the act of manipulating a person to take an action that may or may not be in the targets best interest. This may include obtaining information, gaining access, or getting the target to take certain action[5]

1.1.2 Current Scneario

Businesses spend a significant portion of their annual information technology budgets on high-tech computer security. But the firewalls, vaults, bunkers, locks and biometrics those dollars buy can be pierced by attackers targeting untrained, uninformed or unmonitored users. Humans are the weakest link in any security system, according to KL-based organizers of the Hackers Halted Asia Pacific 2009 conference.[6]

The chief minister of Malacca, Datuk Seri Haji Mohd Ali Bin Mohd Rus- tam, said there is no perfect system in the world. Even if you have the best security devices and softwareyour organization still relies on humanswho are the weakest link in any security system. Public education and awareness is essential.

Figure 1.1 is a Data Security Breach Statistics of 2008 revealing that Malicious Insider and Careless/Untrained Insider is a bigger threat than an outside cracker.

Also, in the age of Social Networking Sites like Facebook, Twitter, Linkedin, Google+ information of companies internal hierarchy structure, employees,

personal information is readily available online for the information gathering phase to breach companies security perimeter by various social engineering ways like phishing links, Trojans, Backdoors, Password guessing, breaking security questions etc.

Is it possible to break into the circle of friends network on social networking sites and make them reveal certain information about them which can be used further to infiltrate the security perimeter of a company?

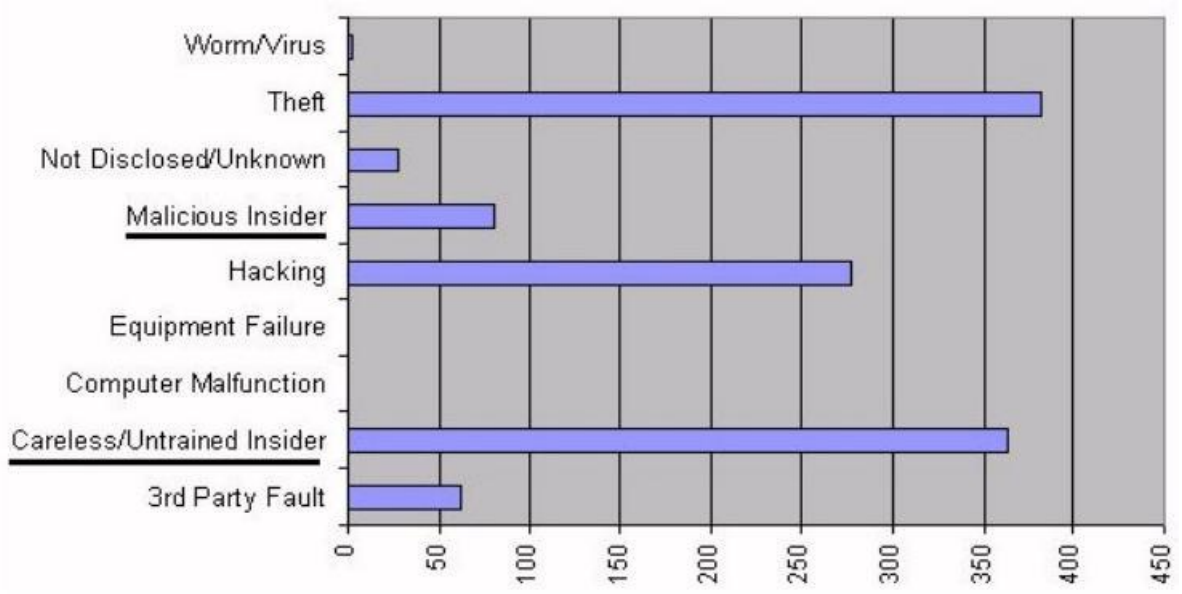

Figure 1.1: Figure 1.1: Data Security Breach Statistics of 2008

Our project is designed to answer this Question and provide a Proof of Con- cept that Yes, it can be done and without even raising any suspicion on the target for a long time.

1.2 Basic Concept

In this project we will use two main concepts:

1.2.1 Profile Hijacking

In Profile Hijacking, we will impersonate the target(s) profile on Social Net- working Sites like Facebook and use their identity to infiltrate the network of other targets friend to fit in the group. Profile Hijacking is important, to use them to reveal certain information without their knowledge.[2]

1.2.2 Using Humans as Botnets

We studied many AI based Chat Bots that are available on the web and we found that Using Chat Bots to chat with humans for Social Engineering Attack is not feasible as Humans detect that the other person is not a human but a program and hence the attack fails even before it is launched.[4]

Since developing a fully convincing AI based Chat Bot is not possible consid- ering the Scope, we will use another human to do the talking with our target while we sit in between, watch and modify the conversation towards the conversation which would reveal certain information which we are interested in.[1]

1.3 Application

Some of the best tools for fighting social engineering attacks are security awareness training and social engineering testing. The effectiveness of these controls will vary based on the quality of their

implementation, including follow-up and retraining.

Social engineering testing, by its very nature, can be difficult to conduct without third-party assistance. One option is to engage an information secu- rity organization to conduct testing. The testing can uncover areas in which an organization is most vulnerable so that risk can be assessed and mitigation strategies can be formulated and implemented.

Rolling social engineering testing into a larger security penetration engage- ment can reduce the cost of the social engineering component, says Jim Patterson, director of consulting for Rapid7.[6]

Main Application of this Project is to develop a tool that will aid in Do- ing Social Engineering Testing on Companies Employees, also it will provide as a live demonstration to employees under training on how social engineer- ing can be done and how by being cautious one can prevent serious damage not only to the company but also to their private life.

Chapter 2

Literature Survey

2.1 Towards Automating Social Engineering Using Social Networking Sites

2.1.1 Summary

In this paper, we saw the use of artificial intelligence to create a chat bot to do the automatic social engineering attack via social networking sites. Experiments were done on a small group of people on facebook.

2.1.2 Advantages

- Automatic tool, minimal input required from the user.
- If perfected, could be the best way to automate social engineering attack.

2.1.3 Disadvantages

- Fails to convince that the chat bot is a real human.
- Has no input of real world information.
- Chat algorithm is hard coded using regular expressions.
- No real world testing because of ethical issues.
- Building a real world chat bot is out of scope.

2.2 All Your Contacts Are Belong to Us: Automated Identity Theft Attacks on Social Networks

2.2.1 Summary

Presents a novel concept of profile hijacking and cloning, which laid the foundation for the proposed idea.

2.2.2 Advantages

- Works with all social networking sites which has no authentication mechanism of who you say is who you are in real world.
- Very Easy to do, as most of the data required is already available.

2.2.3 Disadvantages

- Does not give us information which is not available directly.
- Impersonating or Faking a human identity is a punishable act.
- Real world testing has ethical issues.

2.3 Honeybot, Your Man in the Middle for Automated Social Engineering

2.3.1 Summary

This paper presents the basic idea we can do man in the middle attack on two users. The testing was done on public IRCs but was suggested that it can be done by using profile hijacking on a social networking sites like Facebook.

2.3.2 Advantages

- Results of success is much higher than by using artificial intelligence via a chat bot.
- Easy to build compared to building artificial intelligence.

2.3.3 Disadvantages

- Required gender conversion in IRCs chat when impersonating a different gender.
- Testing was done on IRCs which is not targetted information gathering.

2.4 Eight Friends Are Enough Social Graph Approximation via Public Listings

2.4.1 Summary

Presents a novel idea of Publically available data of users on social networking sites. Also presents the idea of Dominating Sets. It was also mentioned that the facebook friends public data policy keeps changing, before it was 10 friends than it became 8 and now again its back to 10.

2.4.2 Advantages

- Avoids detection by Facebook.
- Presents a novel idea of using Dominating Sets.

2.4.3 Disadvantages

- Is not succesful in the current scenario where users are security concious.
- Has very low sucess rate as per our experiments.

Chapter 3

Proposed Work

3.1 Problem Statement

To develop a tool which can be used for automated social engineering attack on users of social networking sites using profile hijacking and man in the middle attack. Also test the tool via modified version of Turing test at basic conversation level

3.2 Passive Information Gathering Module

3.2.1 Social Networks

A social networking service is an online service, platform, or site that focuses on building and reflecting of social networks or social relations among people, who, for example, share interests and/or activities and people with similar or somewhat similar interests, backgrounds and/or activities make their own communities.

A social network service consists of a representation of each user (often a profile), his/her social links, and a variety of additional services.

According to ComScore, Facebook has the largest number of Social Network User shares (up to end of November 2011)[7]

Worldwide	Unique Visitors	Percentage
Facebook.com	792,999	55.1%
Twitter.com	167,903	11.7%
LinkedIn.com	94,823	6.6%
Google Plus	66,756	4.6%
MySpace	61,037	4.2%
Others	255,539	17.8%
Total	1,438,877	100%

And here are some statistics about Facebook (Alexa estimates, as of 20/04/2012) :-

- daily page views : 8.4 billion

- daily visitors : 650 million
- views per visitor : 12.9
- site rank : 2nd
- traffic fraction : 0.45% 1 in 220 of all web traffic

This shows that there is a lot of user generated information that is available to tap into. Facebook has its own API but we cannot use that to extract data about anyone. So we thought of writing a module which can do Web Scraping.

3.2.2 Web Scraping

Web scraping (also called web harvesting or web data extraction) is a computer software technique of extracting information from websites. Usually, such software programs simulate human exploration of the World Wide Web by either implementing low-level Hypertext Transfer Protocol (HTTP), or embedding a fully-fledged web browser, such as Internet Explorer or Mozilla Firefox.[8]

3.2.3 Features

This module will have the ability to scrape contents of facebook profile users page. It'll basically focus on the friends list of a user. The data generated from this module will help in focusing on who are the victims friends, what are the statistics of them etc.

Figure 3.1 is a screenshot of a user and his friends list

Figure 3.1: Friends List on Facebook

This module should be able to extract

- Full Name

- Gender : Male/Female/NA
- username : Username/NA
- Number of Friends : Total/NA
- Unique UserID

And in addition it should be able to extract each user's friends detail in a BFS Manner. And it can continue this as long as we want it to.

3.3 Facebook Chat Attack Module

3.3.1 Extensible Messaging and Presence Protocol

Extensible Messaging and Presence Protocol (XMPP) is an open-standard communications protocol for message-oriented middleware based on XML (Extensible Markup Language). The protocol was originally named Jabber, and was developed by the Jabber open-source community in 1999 for near-real-time, extensible instant messaging (IM), presence information, and contact list maintenance. Designed to be extensible, the protocol today also finds application in VoIP and file transfer signaling.[9]

Facebook uses XMPP (Extensible Messaging and Presence Protocol) for chatting.

3.3.2 Attack as Example

Following is an example.
Lets say we have two marks (targets) Alice (Human) - aliceHuman and Bob (Human) - bobHuman

We make two profiles both run by bots. The highlight is bothbots hijack the profilesof marks namely alice and bob.

So aliceBots profile is similar to aliceHumans profile and bobBots profile is simlar to bobBots profile. aliceBotsends a friend request to bobHuman. bobBotsends a friend request to aliceHuman. So now, aliceHuman is a friend of bobBot and bobHuman is a friend of aliceBot. Also, bobBot and aliceBot can exchange information outside of facebook to each other.

So now lets say, bobHuman and aliceHuman are online. Our bots start the conversation. Whatever is being passed to one of the Bot by one human, it is passed on to the other Bot and in turn passed on to the other human, i.e. two humans are having conversation through two bots but they think that they are talking to a human since the conversation sounds like a human (which it is).

After some amount of bonding between them our bots start modification by using injecting questions inside the conversations

AliceHuman - bobBot : Hey how was the movie yesterday?
botBot - aliceBot : Hey how was the movie yesterday?and hey btw whats your fav color?
aliceBot - BobHuman : Hey how was the movie yesterday? and hey btw whats your fav color?

BobHuman - aliceBot : movie was great, and its blue btw, whats yours?
aliceBot - bobBot : movie was great,you know my fav color is blue, what is yours?
bobBot - AliceHuman : movie was great, you know my fav color is blue, what is yours?

AliceHuman - bobBot : mine is pink :)
botBot - aliceBot : mine is pink :)
aliceBot - BobHuman : mine is pink :)

Notice how the conversation has been altered to make it unclear that no one asked each other about favorite color and yet both of them told us their favorite color. We got hold of two marks information without raising suspicion.

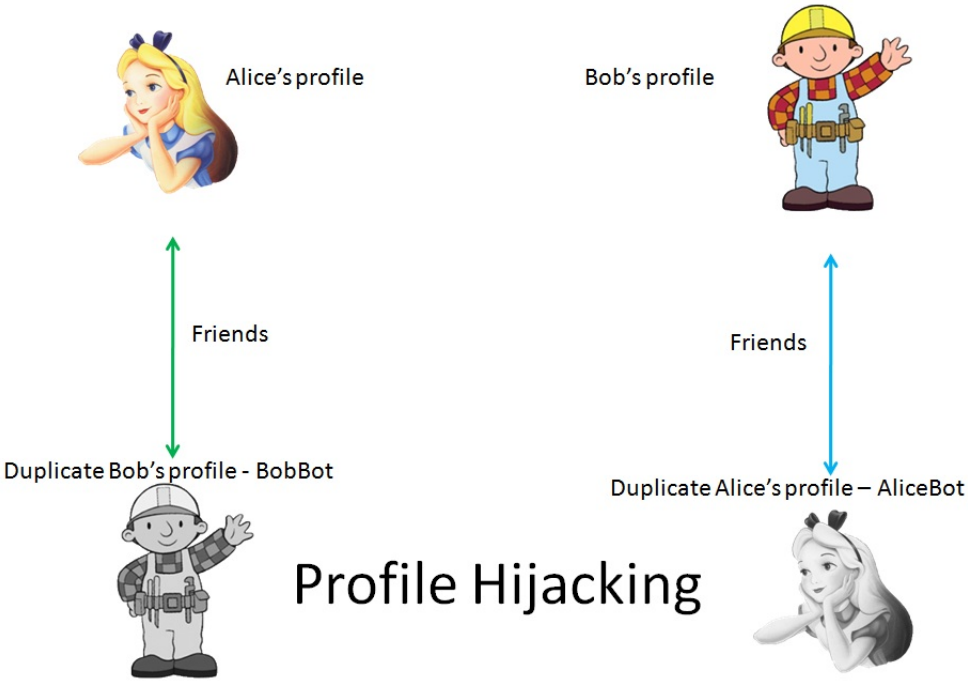

Figure 3.2: Example of Profile Hijacking

3.3.3 Constraints

We will develop a POC (Proof of Concept) only. To modify the conversation we will use simple regex based parsing to match and substitute it with matching strings. There will be no Artificial Inteligence or Machine Learning involved as that is out of scope at this level. Though it can be done and will make the modification more scalable and justified.

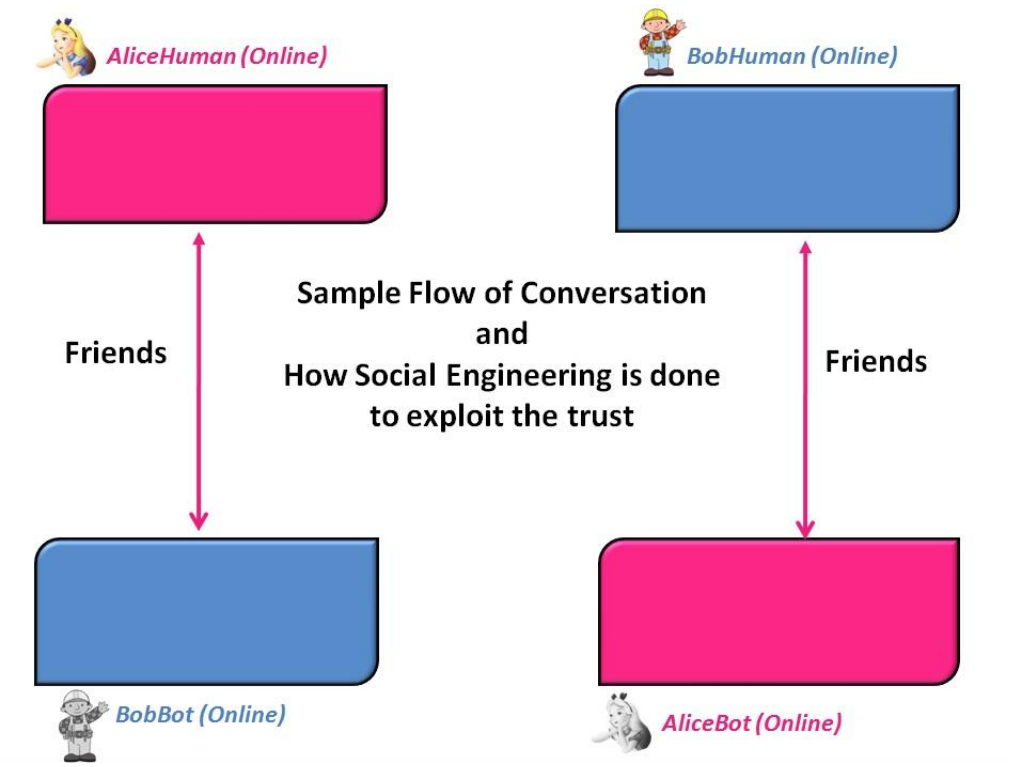

Figure 3.3: Example of Conversation Modification - 1

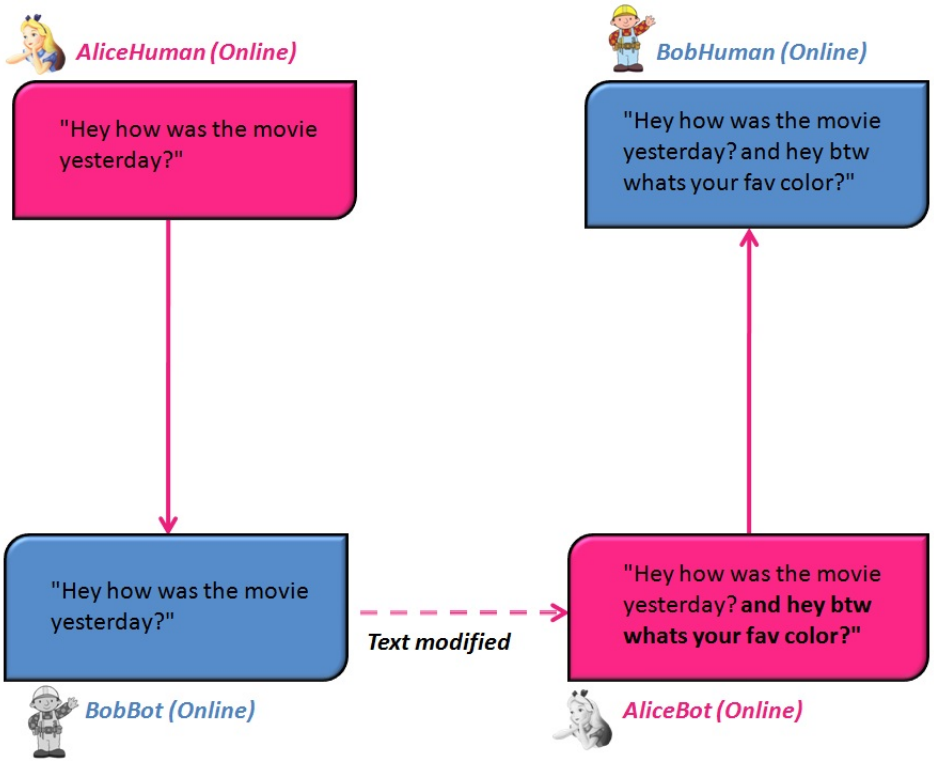

Figure 3.4: Example of Conversation Modification - 2

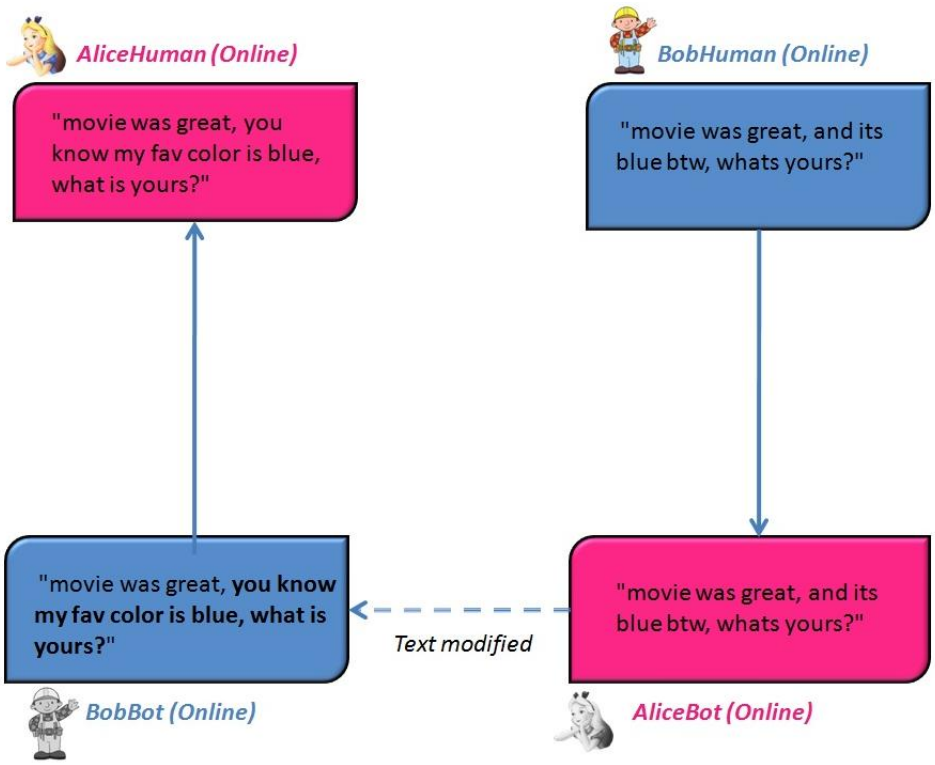

Figure 3.5: Example of Conversation Modification - 3

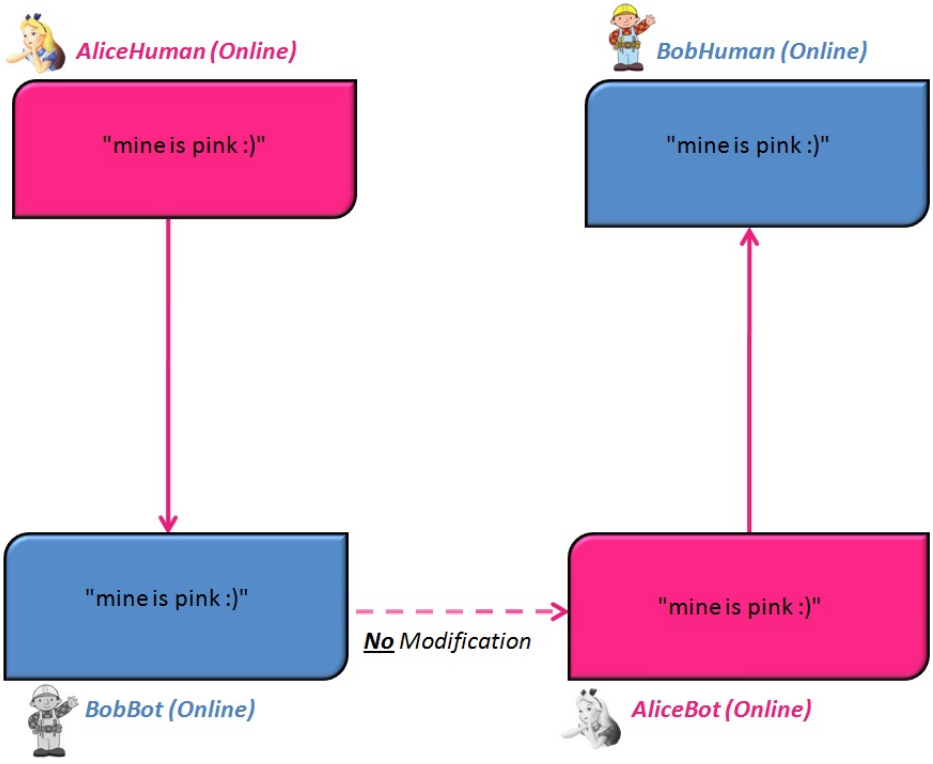

Figure 3.6: Example of Conversation Modification - 4

Chapter 4

Research Methodology

4.1 Passive Information Gathering Module

4.1.1 HTTP

To implement Web Scraping module for Information Gathering we need to understand the HTTP protocol. HTTP (The Hypertext Transfer Protocol) is an application protocol for distributed, collaborative, hypermedia information systems. HTTP is the foundation of data communication for the World Wide Web.

Request methods

HTTP defines nine methods (sometimes referred to as ”verbs”) indicating the desired action to be performed on the identified resource. What this resource represents, whether pre-existing data or data that is generated dynamically, depends on the implementation of the server. Often, the resource corresponds to a file or the output of an executable residing on the server.

Following are the main request methods we will be dealing with :-

POST : Submits data to be processed (e.g., from an HTML form) to the identified resource. The data is included in the body of the request. This may result in the creation of a new resource or the updates of existing resources or both.

GET : Requests a representation of the specified resource. Requests using GET should only retrieve data and should have no other effect. (This is also true of some other HTTP methods.) The W3C has published guidance principles on this distinction, saying, ”Web application design should be informed by the above principles, but also by the relevant limitations.”

4.1.2 urllib2

Since we will be using Python 2.x for our project, we will be using the standard library for doing the HTTP requests to be made. The urllib2 module defines functions and classes which help in opening URLs (mostly HTTP) in a complex world basic and digest authentication, redirections, cookies and more.[10]

urllib2.urlopen(url[, data][, timeout])

- Open the URL url, which can be either a string or a Request object.

4.1.3 Regular Expressions

A Regular Expression is the term used to describe a codified method of searching invented, or defined, by the American mathematician Stephen Kleene.[11]

References

[1] *Honeybot, Your Man in the Middle for Automated Social Engineering*;Tobias Lauinger, Veikko Pankakoski, Davide Balzarotti, Engin Kirda; EURECOM Sophia-Antipolis, France. LEET10 Proceedings of the 3rd USENIX conference on Large-scale exploits and emergent threats.

[2] *All Your Contacts Are Belong to Us: Automated Identity Theft Attacks on Social Networks*; Leyla Bilge, Thorsten Strufe, Davide Balzarotti, Engin Kirda EURECOM Sophia Antipolis, France. WWW 09 Proceedings of the 18th international conference on World Wide Web.

[3] *Eight Friends Are Enough Social Graph Approximation via Public Listings*; Joseph Bonneau, Jonathan Anderson, Frank Stajano, Ross Anderson. SNS 09: Proceedings of the Second ACM EuroSys Workshop on Social Network Systems.

[4] *Towards Automating Social Engineering Using Social Networking Sites*;Huber, M.; Kowalski, S.; Nohlberg, M.; Tjoa, S.; Computational Science and Engineering, 2009. CSE 09. International Conference.

[5] Social Engineering : The art of Human Hacking

[6] Link Name
`http://goliath.ecnext.com/coms2/gi_0199-7186209/ The-human-element-the-weakest`

[7] Market Share of Social Network Sites
`http://techcrunch.com/2011/12/22/googlesplus`

[8] Web Scraping
`http://en.wikipedia.org/wiki/Web_scraping`

[9] Extensible Messaging and Presence Protocol
`http://en.wikipedia.org/wiki/Extensible_Messaging_and_Presence_Protocol`

[10] Urllib2 Python Library
`http://docs.python.org/library/urllib2.html`

[11] Regular Expression
`http://www.zytrax.com/tech/web/regex.html`