Detection and Analysis of Malicious URLs

Thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Technology
In
Computer Science and Engineering
By

B Murali
(Roll No: 213CS2156)

under the guidance of
Prof. Sanjay Kumar Jena

Department of Computer Science and Engineering
National Institute of Technology, Rourkela
Rourkela-769 008, Odisha, India
June, 2015.

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India.

Declaration

I certify that

- I have complied with all the benchmark and criteria set by NIT Rourkela Ethical code of conduct.
- The work done in this project is carried out by me under the supervision of my mentor.
- This project has not been submitted to any other institute other than NIT Rourkela.
- I have given due credit and references for any figure, data, table which was being used to carry out this project.

B MURALI
National Institute of Technology
Rourkela-769008

Place: NIT,Rourkela-769008
Date: 01 - 06 - 2015

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India.

Certificate

This is to certify that the work in the thesis entitled ***"Detection and Analysis of Malicious URLs"*** submitted by ***B Murali*** is a record of an original research work carried out by him under our supervision and guidance in partial fulfilment of the requirements for the award of the degree of Master of Technology in Computer Science and Engineering, National Institute of Technology, Rourkela. Neither this thesis nor any part of it has been submitted for any degree or academic award elsewhere.

Prof. Sanjay Kumar Jena
Assistant Professor
Department of CSE
National Institute of Technology
Rourkela-769008

Place: NIT, Rourkela-769008
Date: 01 - 06 - 2015

Acknowledgment

First of all, I would like to express my deep sense of respect and gratitude towards my supervisor Prof. Sanjay Kumar Jena, who has been the guiding force behind this work. I want to thank him for introducing me to the field of Social Networks and giving me the opportunity to work under him. His undivided faith in this topic and ability to bring out the best of analytical and practical skills in people has been invaluable in tough periods. Without his invaluable advice and assistance it would not have been possible for me to complete this thesis. I am greatly indebted to him for his constant encouragement and invaluable advice in every aspect of my academic life. I consider it my good fortune to have got an opportunity to work with such a wonderful person.

I wish to thank all faculty members and secretarial staff , Phd Scholar specially to Jintendra Kumar Rout, Soubhagya Sankar Barpanda of the CSE Department for their sympathetic cooperation.

During my studies at N.I.T. Rourkela, I made many friends. I would like to thank them all, for all the great moments I had with them.

When I look back at my accomplishments in life, I can see a clear trace of my family's concerns and devotion everywhere. My dearest mother, whom I owe everything I have achieved and whatever I have become; my beloved father, for always believing in me and inspiring me to dream big even at the toughest moments of my life; and my sister; who were always my silent support during all the hardships of this endeavour and beyond.

B Murali

Abstract

Social Network Sites(SNS) is the soul of the Internet. It has become a global phenomenon with enormous social as well as economic importance within a few years of their launch. Because of larger user space SNS has become popular day by day. Information exploitation popularity in SNS has attracted not only novice users but also spammers. In SNS spammers are using evolving technology and they safely trading their illegal activities by phishing through e-mails, Social Reverse Engineering(SRE), by posting some incite messages. The novice users often becomes victim to these malicious activity which impacts them both socially and economically. The study show that because of this illegal activity the SNS organisers and users are loosing $2 million for three months. In this thesis we exploited the security gap that many popular SNS services like Twitter, Facebook do not provide to its users. We have collected a large scale of long URLs and short URLs from multiple sources of SNS which are checked against malicious and non-malicious detectors and we analyse their features to classify the URLs. Our result shows that Naïve Bayes classifier performs better then other classifier algorithm with accuracy 95.4%.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter gives the overview of the thesis. Description of Social Network Services. Describes the security issues in SNS. Common attacks in SNS and the motivation of my research.

1.1 Social Network Sites(SNS)

From ancient time man is called as a social animal. From his beginning man has maintained a social relation with nature, animals and with a fellow human being. It was this social relationship that helps him to have a close relationship in the universe with one another. In modern times with increase in population the SNSs has become an easy and a much efficient platform in maintaining social relationships. Online Social Network sites like Facebook, Twitter, Linkedln, MySpace or Google+ has become popular sites in Internet platform. They have attracted of all ages from technicians to novice users. In the wide area sphere like research, industries, business, working Office, news media, organization, entrepreneurship SNS have become a daily practice in use. Mostly SNS have mainly used for information sharing and to express on common interest views example political view.

1.1.1 Definition of SNS

SNS are basically a web-based application which usually allows the individual to construct the semi-public data [16] with in a closed system, articulate a list of users to whom an individual can connect, share information, express their common view in a common platform. SNS allow the individual to meet the strangers of common interest, and view and traverse the list of one's individual connection. Though SNS vary from one to the other in terms of their nomenclature in connection and service provision to its users, the basic principality is to share information. The following Table 1 show the most popular SNS in brief.

1.2 Security Issues in SNSs

For the past few decades the popularity in SNS [17] such as Facebook, Twitter, Google+ has increased rapidly. Though it has attracted all age groups, but the youngster has outset among the other groups. SNSs has become an important communication platform in social life, with increasing security concern over a period of time. Some of the security breaches may be like viral marketing, network, structural attacks malware attacks. Some of them are explained briefly as follow.

- **Privacy Breach Attack** : SNS allow the individual to construct their SNS network with semi-public data like date-of-birth, current address, photo, videos. Such ready available personal information can mark for privacy Breaches.
- **Breaches by Service Provider** : This readily available data may be used by the one's service provider for advertisement purpose to benefit them in multi-ways. As such the data may fall into the hands of untrustworthy person.
- **Breach form Third Party** : To have more functionality the user in the SNSs may use the trusted third party application. To use such application the user must have to accept or compromise some privacy issues by accepting theirs term and conditions.

1.2.1 Viral Marketing

Because SNS has been formed by the real people, they are an easy target for the viral marketing. The public perception that the information shared with their friend are of trust messages. These perceptions have benefited spammers to attack the users by employing the SRE trick to enhance the viral management effectiveness. Some of the viral management issues are outlined as below.

SNS service providers			
Service Provider Name	Description	Date Launch	Registered Users
Flickr	Photo sharing, commenting, photography related networking, worldwide	February 2004	32,000,0007
Facebook	General: photos, videos, blogs, apps.	February 2004	1,280,000,000
Haboo	General for teens. Over 31 communities worldwide. Chat room and user profiles	August 2000	268,000,000
lbibo	Talent based social networking site that allows to promote one's self and also discover new talent	2007	3,500,000
Istagram	A photo and video sharing site.	October 2010 AS	150,000,000
LinkedIn	Business and professional networking	May 2003	200,000,000
Myspace	General	August 2003	30,000,000
Google+	General	28 June 2011 Open to people 13 and older	120,000,000

Table 1.1: Social Networking Services

Spam in SNSs

Spam in SNSs are of two types mainly.

- **Context-aware Spamming** : Context-aware Spamming is the advantage taken by the spammers by click-through mechanism. Here the spammers analyze the target users behavior and post the malicious content in a obfuscate manner. Here the spammers incite the users, and draw attention to click their post message. When the user clicks the obfuscate message he get trapped of spamming .

- **By Broadcasting** : Here the user may not have any particular target, but broadcast the abusive content on the SNSs. Here the spammer may use the sexy photos and seduce content to attract the users .

- **Phishing Attack** : Phishing Attack Figure 1.1 is usually employed by the attackers to steal the credential attacks by the spammers. This is one of the most popular.

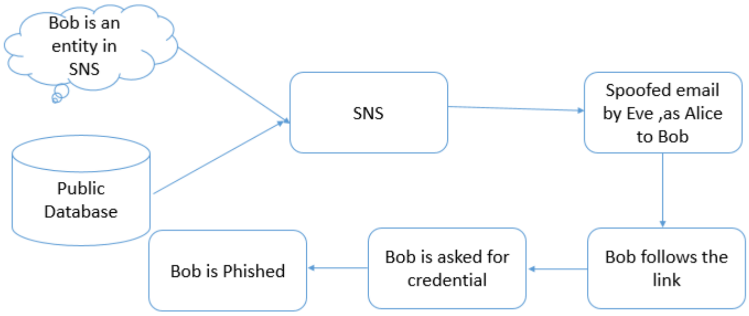


Figure 1.1: Phishing Attack

Viral Marketing method in SNSs. Novice persons in the SNSs usually attack by the spammers.

- **Social Reverse Engineering(SRE)** :SRE is usually used as a bait for the users so that the users by themselves get into the trap of the attackers. This may usually does by the attackers by posting popular post and by e-mails.

1.2.2 Network Structural Attack

A change in the in the Network structure by the malicious attackers [33] may result a serious threat in SNSs if the SNSs in mainly based on membership. One such type attack is Sybil Attack.

•**Sybil Attack** is a general Figure 1.2 form attack on malicious attack where the malicious attacker creates a number of fake identities which outvote the genuine identities. This attack may set aside the intended purpose of the SNSs.

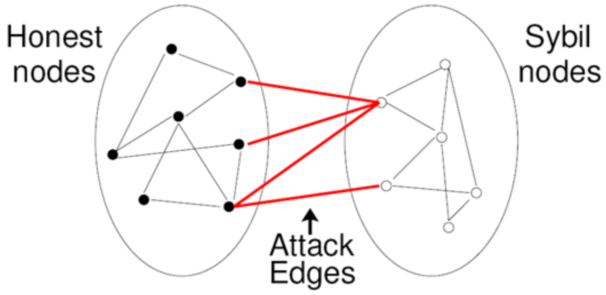


Figure 1.2: Sybli Attack

1.2.3 Malware

Malware is a concrete term used to refer to any form of hostile or intrusion software like computer virus, worm, trojan horses, spyware, adware and other malicious program. Due to present of large information the attackers are also employed to trade malware content in SNSs. It can spread in SNSs by profile, interaction, a third party application. Koobface is the first successful worm the attacker spread in SNSs. Koobface infection allows attacker to get into your personal information like your banking information, passwords, or other personal details. It is considered a security risk and should be removed from the network.

Koobface [30] is an worm traded on Facebook , it work as follow.

- Register and activates a Facebook account by using the Gmail-account
- Join random Facebook groups
- Post in the wall of the friends which coantians koobface component

The following Table 1.2 shows the impact factor of the different attacks

Brief about SNSs Attack				
Measure Attack	Phising	Sybil	Malware	Spamming
Difficult,	Yes	Hard	Hard	easy
Server Defence Effectiveness	No	Yes	Yes but limited effectiveness	Yes
User Defence Effectiveness	Yes	No	Yes but limited effectiveness	No
Threat to User	High	Medium	High	Low

Table 1.2: Impact of various attacks in SNSs

Common attacks in SNSs	
Attacks	Description
ID Theft	The presence of semi-public data available helps the attackers to theft ID of the target SNS users
Profile Cloning	The users in the SNSs generally trust the other profiles with same interest area . Thus the attackers will impersonate the genuine user are by creating clone profile.
Secondary Data Collection	Here the attackers collect the data available in SNSs and extract the useful information like SSN number, particularly such information where the attacker can access the target particular users information.
Communication Tracking	The movement of data in SNSs arises the concern of communication privacy . Here the malicious SNSs provider or malicious member with the appropriate set of privilege can be able track the communication. It is very difficult to identify such attack.
Defamation and Ballot Stuffing	These attacks are generally target the particular member to defame their reputation by posting them obscene content or text.
Friend-in-the-middle Man Attack	Here the attackers hijack the HTTP session in the network layer , where he retrieves the cookies and the them, then start attack them by impersonating as a genuine user
Friend Injection	Here the attackers the network session and then add themselves as an friend and penetrates into the closed group.
Application Injection	Third party application such as online games , popular SNS or hiding malicious hiding a malicious application without any activity visible to the user is possible. The attacker will install application which automatically collect the personal data of the users.

Table 1.3: Common attacks in SNSs

1.3 Common Attacks in SNSs

The following Table1.3 shows the common attacks in SNSs. Some of the attacks measures from easy to hard.

1.4 Research Motivation and Aim

1.4.1 Research Motivation

Application of advanced information technology in web service has offered many services to its users. The services may range from chatrooms, sending text, information sharing, and multimedia sharing. The above mentioned services are built in one place as an SNS, hence its usage and popularity increasing rapidly. With the growing number of SNSs services in internet correspondingly there is a huge amount of information disseminated. This large amount of information is usually used in research to analyze the user behavior like social relation, information trading. Mostly some of the popular SNSs like Facebook, Google+, Myspace, Twitter have attracted a large number of users. As the increasing number of users, the SNSs services started providing various functionality like charting rooms, sharing information, advertisement etc. Hence SNSs has seen as an economic growth platform for some of the entrepreneurship.

Unfortunately, SNSs have attracted the attackers and started attacking on SNSs by a phishing attack, inject malicious code, and launching drive-by-download attacks. These malicious attacks have led to serious privacy crime and economic crime. Once the user click through this malicious content the user will take to the other users without their notice, which embedded with malicious code which steals the personal information and credential of the user, by this they fall into the victims of the malicious attackers The SNS begin its era in a slow phased manner and emerged as a soul of the internet. It has fascinated all groups of ages, and marked as a part of the social community. The researches for the past decades has seen SNS as a huge information reposit. From then onward the SNS services are trying to provide the best security features to their users. Anyhow the spammers, phishers have resisted to these security techniques and evolve to the new techniques those cannot be identified by the traditional technique (blacklisting). One of the most prevailing technique that spammer are using in the SNS is a short URL technique.

Short URLs Services

From the beginning of the short URL services the use of short URLs had become a norm in SNSs where generally character limitation exists (Twitter has 140 character limit). The usage of the short URLs has resulted in a space reduction methodology in the SNSs. The following draws the attention of the space reduction Figure1.3 reduction by short URLs.

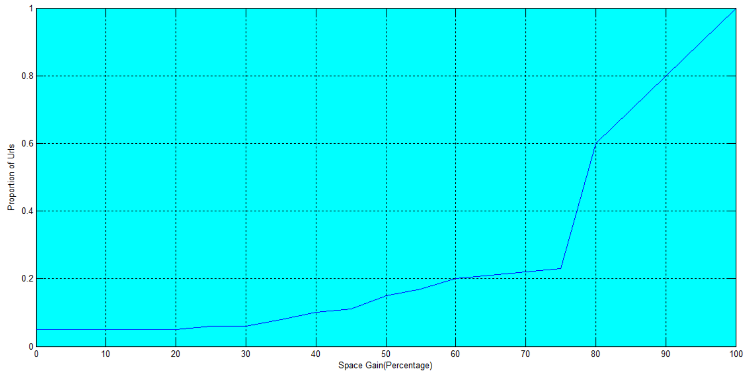


Figure 1.3: Space Reduction by shorting long URLs

1.5 Advantage and Disadvantage of Short URLs

The short URL services take the long URLs as input and give corresponding short URLs with a unique Hashtag, generally append at last. For example, the long URL *https://pypi.python.org/pypi/python-graph* is given to the any short URLs services as bit.ly it returns the short URLs as *https://bit.ly/xxxxx*. Though short URL services resulted in space, reducing methodology in SNSs but it has resulted in a security breach like cybercrime. The resulted short URLs may be malicious or benign. The malicious short URLs are obfuscate in nature and cannot be identified by traditional methods (blacklisting). The multiple redirection of short URL has made it very difficult to identify the real malicious URLs.

Chapter 2

Literature Survey

2.1 Literature Review

With the limitation of the text characters in SNSs for example, in Twitter(140 character) has made the SNSs users to use shorter services. The following Table 2.1 are some of the short URL services that are popular in use. The above are shortened services Table 2.1 takes input a long URL and it returns the short URL This short URL that is generate will redirect to the same long URL but the short URL looks random and obfuscate in nature. The character limit of the SNS has laid immense usage of short URL services. As stated Figure 1.3 though URL shorten has reduced the space in textbooks of the SNS. Due to obfuscate usage these services it leads to security breaches in SNS.

Information Security is of growing interest of policy makers as society become more dependent on secure communication. Andreson and Moore [30] in their research work have briefly explained about the security concern in economic perspective how these malicious content have impact the economic issue. Despite this large malicious activity, information about the malicious content and the losses done by such crimes have largely remained hidden from public crime. The reasons may be as follow. First fear of negative impact on public which arises if incident are openly discard and discussed Second some argues that disclose of information about the incident actually aid attackers more than it helps defenders. Ransbitham [8] has observe that vulnerability in open source software are more frequently exploited by attackers in open software than in closed software.

Yet there are also some clear benefit to public disclosure of malicious incident. First the criminal already know how to attack on SNSs by disclosure the security incident

Short URLs Domains	
Domain	Site Tittle
Tinyurl.com	Tiny.URL
Bit.do	Bit.do
Indkin	LinkendIn
qr.ae	Quora
adf.ly	Adf.ly
goo.gl	Google
is.gd	Is.gd
u.ub	Adf.ly
tweez.me	Tweez.me
tr.im	Tr.in

Table 2.1: Feature of the short URLs

the ‘defenders’ can know how the attacker incident can work and an effective counter measure can employed to not to happened in future.

Secondly, information threat is an key barrier to optimal security investment . Better measurement to optimal security investment and impact of incident can help organisation to better to tightened their security features from the frequent impact of the incident.

Thirdly, bring incident to publicity will help the defender to quickly identify the attack impact measures and find some loop hole.

Meanwhile, firm have undertaken a no of collaborative efforts to impose security without disclosing result publicity. Google operates a large blacklist [12] of malicious website and URLs which marked as spam, malware, worm without revealing the information about the infected website.

The spammer uses number of techniques to find any vulnerabilities in the website. The way they can employ is by scanner. Scanner is the technique where the spammers scan the other website and such for some loop holes and inject some rootkits [26]. A rootkit is a stealth type of software , typically malicious design, to hide the existence of certain process from normal method of detection and help to find some loopholes from which they access the privilege of the server. By applying rootkits they compromise the other end software. Then they exploit the machine for their own purpose, and sell to the third party. If rootkit is not added the criminal adds few websites and does phishing attack.

Hu Huaping, Weijianli [19] have analyse how the IM(Instant Message) has become one of the popular online communication tools among the users in SNS. The added functionality and increasing popularity has attracted the attackers, by sending worm, malicious content from current list of content .

Federio Maggi *et al.* [24] has exclusive research on short URL for two years and analyse the security threat and explain the countermeasure. They said that users are seldom exposed, while browsing to threat spread via short URL or atleast no more than they are exposed to the same thread spread via long URLs.

Neils Provos and Dean McName *et al.* [25]have explain how drive-by-download malicious URLs link has exploited the threads in SNS.

Zhan *et al.* [27] has analyse that the many phishing sites are design by some modification of popular site. He analyse the phising site by content based method like lexical analyse, HTML content, domain name, Google page ranking.

Fette *et al.* [36] has analyse the feature of the email by HTML tag, number of links, by using the SVM. Bergholz *et al.* did the same kind of work like he use email bodies, weblink , keyword properties by using Markov chain training. Abe-Nimen *et al.* has used 43 most popular keyword.

Ma *et al.* [6] has done the similar work but he added more specific feature like host machine features like IP address, WHOIS, domain name, regional location and he classified by machine learning classifier by Bayesian and SVM.

Morse *et al.* [7] has research on the highly connected nodes. From that he analyse that highly connected nodes are responsible for the spreading of worm through out of the network.

Koobface which attack the facebook and Myspace primarily by sending that contains like malicious website and leverage various SNSs information.

Thomas *et al.* [31] has contributed by collection more than 213 thousands users who compromised on Koobface, such web sites are taken more than four days to blacklisted and only 27% of URLs are detected as malicious and stated 81% users clicked on Koobface spam.

Stringhini *et al.* [29] used decoy profile approach for collecting social information from SNS feature and they are used to identify spamming in social network without considering the malicious link or content posted. They concluded that spammer would send numerous friend request but they were not a real-life friends of others.

Jin *et al.* [20] used three types of features like image content text, social network features, used to characters the user profile and their behaviour.

2.2 Analysis of Malicious Long URLs

The pervious study has conducted to understand how the attackers has been URL to spread their spam in the SNS, many previous study has contributed how the spam and phishing is spread through email, and they concluded that the attacker have heavily used URLs to spread the bad Content. Benevenuto *et al.* [9] collected a large scale of URLs of nearly 2 billion and identify the features of the URLs which detect spam on Twitter. And after the manual labelling of features they classified and achieved 70% accuracy. They have founded that a fraction of tweets particularly of some hot topics contains more number of URLs then other. This clearly highlights to what extend the attackers has been using. Though the research has came up the most efficient popular blacklisting in detecting spam but it has been observed that their evaluation technique has not suitable for the detecting spam in the Twitter when the user employs the short URLs technique for the particular long URL as they are in obfuscate in nature. By using this services the attacker has taken more advantage in trading their illicit content. In such case in both Twitter and Facebook by using such services, the spammer has complicated the process of detecting ny applying the multiple chain re-direction. Wang *et al.* [34]. in research they used the Click rate measure as a feature and concluded that the rate of spam in Twitter is (0.13%) is higher amount of spam that spread through spam e-mails. Thomas *et al.* [32] have develop an automatic system named Monarch through which they have classify the URL submitted to any web services as malicious or non-malicious in real time . Their classification is based on the landing page properties on the webpage like html tag, hosting infrastructure, pop-ups , plugins, cookies, content .Thought it may be the most effective method which is based on the content search it reduces the performance issues. Their classification is based on the landing page properties on the webpage like html tag, hosting infrastructure, pop-ups, plugins, cookies, content. Thought it may be the most effective method which is based on the content search it reduces the performance issues. Lee in his contribution has reveal that spammer has employ the multiple redirection methods to a malware/phishing website. As his contribution he created a system that would not fall as prey to conditional redirection. In his classification he has used both the conditional redirection with the accuracy of 86.3% was attained

2.3 Analysis of Malicious Short URLs

Reduction of the space reduction in the SNSs and their heavy usage of short URLs services , therefore, there must be a comparative level of understand level of acceptance of the short URL services in SNSs. Kandylas *et al.* [21] research confirmed that the attackers have used short URLs to trade their illicit work and they found that the malicious short URL by clicked based method. And they concluded that the usage of the short URLs in trading malicious URLs is more then by long URL. They also found that duration of the short URLs is less than the long URLs by which they evade the security check. The dataset they have collected by crawling the webpage , they dataset mainly consist of two domain short URLs and reveal that 50% of the short URLs exceeds 100 days. They have analyse that the usage of the short URL services is because of the space reduction. concentrated mainly on the malicious short URLs in the emails and highlighted the privacy and security concern by the short URLs service over the SNSs. They found a lot of private user information traces associated with short URLs and observed a low spam detection rate for 16 shortening services they analysed. For a particular short URL domain they found that 57% of them are bit.ly. Chhabra *et al.* [14] has also found that the attackers has employ short URL not only for the space reduction but also to spread phishing attack in twitter. They also found that the attackers employs the short URLs technique to hide their malicious link.

Chapter 3

Research Contribution

In first part of our research contribution, we have collected a large scale of URLs from SNS like Facebook [1], Twitter [2] by using their secure APIs [13] which are provided by them to their registered users. These API allows us to collect data from Public Domain, and allows the users by using search method followed by keyword as an input. The major study finding are

- The analysed malicious URLs heavily referred to many SNSs.
- They are short lived and registered with unpopular domains which are also short lived.
- We have analyse that their exit a large number of communication Propagation that Malicious illegal trade activity in SNSs.

In second part of our research contribution we have analyse the behaviour of the URLs by there feature which enables them to classification.

- These features enable to detect the behaviour of the URLs like benign or malicious.
- They show the security breaches the present in Facebook and Twitter.

Finally, after identifying the security breaches in SNSs through URLs we have classified the large scale of URLs by classifier algorithms .

- Unlike the existing previous studies our research contribution doesn't depends only on Bitly but on wide range of short URLs services like Bitly [3], Twitter [2], Owly etc.

- We have not only rely on Online Detectors to identify the malicious character of the URLs but we also used the lexical standard feature to detect the Malicious character of the URLs.

- Our classification work efficiently when we employ the combination of features like click, life time of the URL, Domain life time, References. And our results shows that Navies Bayesian has out perform the best classification result.

Chapter 4

Solution Approach

4.1 Exploitation of URLs in SNS

As the popularity of SNS are marking high in the internet world, it is not only used for information sharing, but also used for the trading of the illicit content [30] [35] on SNS which varies from Worm, virus, drive-by- download, kbooface, phishing etc. Our research contribution is mainly focused on the URLs. URLs are used as bait for the users to trap the users where they end with the reveal of the sensitivity and credential information. It has been observed that a well experienced community on the internet spread the malicious content the URLs in SNS. The attackers [10] employs many different techniques like short URLs technique where usually a long URL is given to the available short URL services and result back the equivalent short URL with unique hash tag.The attacker may also employ the malicious URLs [12] by creating the malicious URLs which is very similar to the popular existing URLs. Our research mainly focused on the URLs which contained both short and long URLs, which involves wide range of security breaches ranging from malicious, worm, virus, malware, phishing [28]. Because the usage of URL is the main obfuscate method that the attackers employ to hide from the scanner checkers and invades the target.

4.2 Suspicious Long URLs Exploitation

4.2.1 Long URLs Data Collection

Our long URL data collection contains the URLs [12] , where some of the URLs land the user the to the intended landing page and few others redirects from one page to the other page. Such a way the attackers employ multiple redirection technique to evades the security scanners. These multiple redirection had made the security observers very difficult to detect the attackers. In some instances the attackers has used the some popular long URLs and modified them to trade their malicious URLs. In some cases they have used the unpopular domain and get registered their URLs [11] whose domain info do not have with the security observer. The other technique the attackers employ is that they registered themselves with the domain for a short period of life span which make them to evade without any security checking. We have also concentrated on lexical features to detect the malicious of Long URLs, we have used these feature form the standard red tag words, these words are the already available tag words which are used to detect the suspicious features of the URLs. In our research, we have used nearly 13 features which are as below.

Lexical Feature [18]help us with analysis to detect the URLs which are different from the benign and help to mark them as malicious. For example, *www.amazon.com* is a benign but the attackers employ the obfuscate method in which the spammers use the most popular URLs illicit manner like *www.amazon.com.phishy.xxxxxx*.

Similarly the most effective traditional method is crawling webpage [18] and searching of all malicious content like hyperlink, spamming emails. This method results in high accuracy, but it is much tedious and time consuming when the user employs dynamic changing content technique in web services.

4.3 Suspicious Short URLs Exploitation

4.3.1 Short URLs Data Collection

Because of the evolution of the emerging technology, the attackers have used the most efficient methods to trade their illicit content on the SNSs, one of the such technology is the usage of the short URLs services. Due to the characters limit in social networks (for example 140 characters in Twitter) the users used the social network services as a space, reducing technology in SNSs. Due to the increasing popularity in usage of the short URLs services which comes by obfuscate behavior in SNS it has not only attracted the genuine users but also the attackers to use them as a safe trading methodology for their malicious data trading.

We have collected a large Figure4.1 set of URLs from the SNSs like Facebook and Twitter [23] by their APIs which are freely provided for their registered users, from where, we have separated the Long and short URLs of different domains like Twitter, Facebook, Owly, Tiny URL, Bitly.

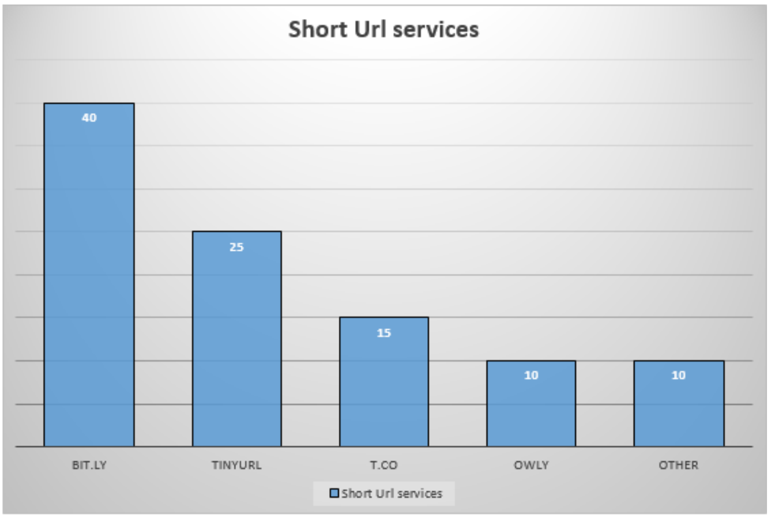


Figure 4.1: Popular short URL services collection

For collected short URLs Figure 4.1 we have analyzed them by their features Table 4.1 which are provided by their respective services. The following are some of the features that they have provided for analysis their short URLs some of them are listed below. These features help us to analyze the behavior of the short URLs, for example the countries feature [22] help us to analyze the referrer of the short URL from different countries. Click feature [34] used to analyze why the malicious URLs have a low number of click.info used to analyze the link creation and creation time. Referring domain help to analyze the domain referring click traffic.

We have collected a large scale of URLs which consist of both the long and short URL. Then we used the available Malicious detectors like VirusTotal [4], Bitly, Phish-

Features of the short URLs	
Features	Description
expand	expansion of short URL to Long URL
info	basic information of the URLs
lookup	page title
shotern	for the long URL return short URL
linkedit	link meta info form the user history
domain	to know the domain of the URL
clicks	click count of the requested URLs
countries	countries reffering click traffic
encoder	the user who encode the user
reffers	return about page referring click traffic
refferbydomain	click group by referring short URL domain
history	created before, created after information
prodoamin	gives information about the domain of the short URL
doaminclick	click on the particular domain
domainshiorterncount	specfied tracking domain

Table 4.1: Feature of the short URLs

tank, Google Safe Browsing, and we marked as a malicious and non-malicious as '1' and '0'. and we use above features for classification.

Chapter 5

Experiment and Detection of URLs

This chapter gives the overview of the Experiment and analysis of our research work.

5.1 Proposed Model

The following proposed Figure 5.1 model briefly explains the approach that we have followed in our research. It starts with the data collection model where the data are collected from the Facebook, and Twitter, followed by extraction of URLs and their features, and by labelling, followed by the feature extraction and then by classification and the result of the classification algorithm. The following section briefly explains each and every module in detail.

5.2 Data Collection Approach

Tweets

Tweets are the short message services provided by the Twitters to their users to share information. These tweets may be Public visible and also privately according to the option set by the users. Because of the limitation of space (140 character in Twitter) the users are forced to use short URLs. Though it really reduces the space, it has many disadvantages. Because of the small and attractive of the short URLs nature it attracts the novice users to go around it and fall into spammers, phisher traps.

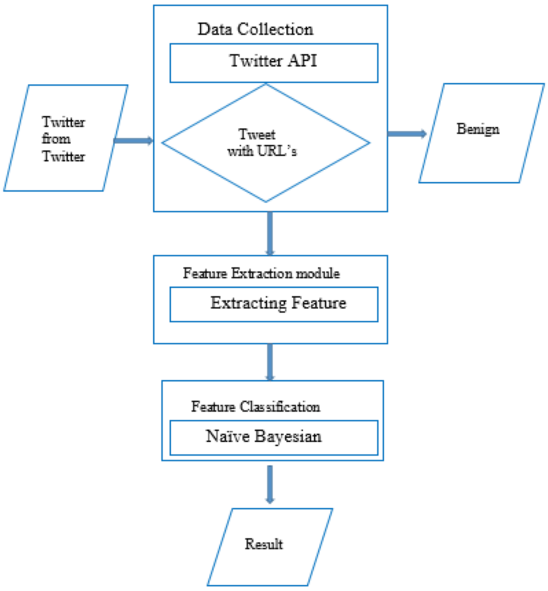


Figure 5.1: Proposed Model

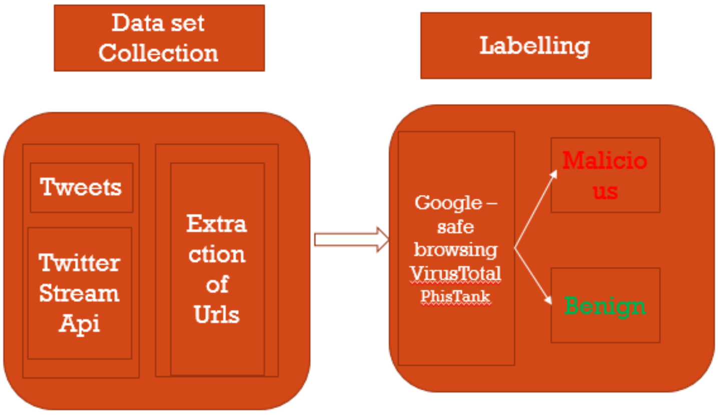
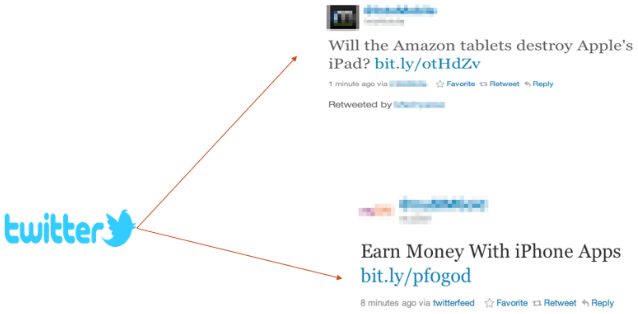


Figure 5.2: Data Collection

Twitter API

Twitter provides its own authenticated API [2] to the registered user. The users get the authenticated keys, which allows the user to access the public visible Tweets. Based on the search method and search key words, it provides the relevant message to the user. The following subsection explains briefly about data collection, extraction of short URLs, feature extraction and classification.



Will the Amazon tablets destroy Apple's iPad? bit.ly/otHdZv

Earn Money With iPhone Apps bit.ly/pfogod

Figure 5.3: Example of short URLs in Twitter

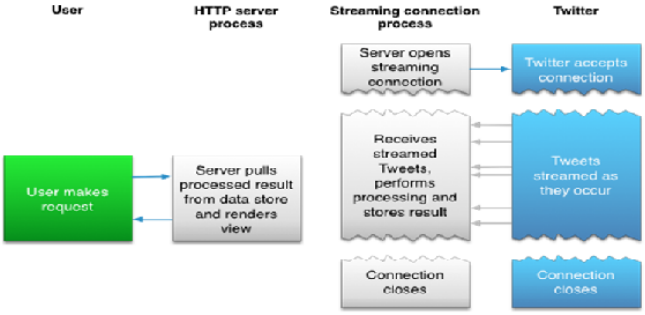
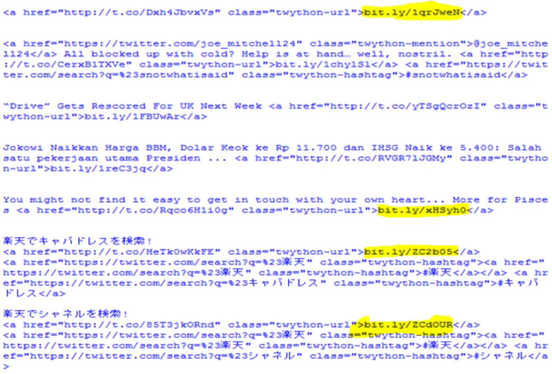

User
HTTP server process
Streaming connection process
Twitter
Server opens streaming connection
Twitter accepts connection
User makes request
Server pulls processed result from data store and renders view
Receives streamed Tweets, performs processing and stores result
Tweets streamed as they occur
Connection closes
Connection closes

Figure 5.4: Data Collection by Twitter API

Data Collection by Twitter API

The REST API 5.3 provide programmatic access to read and write Twitter data. Author a new Tweet, read author profiles and follower data, and more. The REST API identifies Twitter applications and users using OAuth; responses are available in JSON. After collecting the tweets from the public domain from tweets from the twitter by applying streaming methods like on current trends we have extracted the required data from tweets which contains URLs for example as below for bit.ly URLs short. We



```
<a href="http://t.co/Dxh4JbvxVs" class="twython-url">bit.ly/1qrJweN</a>

<a href="https://twitter.com/joe_mitchell24" class="twython-mention">@joe_mitche
ll24</a> All blocked up with cold? Help is at hand… well, nostril. <a href="http
://t.co/CerxB1TXVe" class="twython-url">bit.ly/1chylSl</a> <a href="https://twit
ter.com/search?q=%23snotwhatisaid" class="twython-hashtag">#snotwhatisaid</a>

“Drive” Gets Rescored For UK Next Week <a href="http://t.co/yTSgQcrOzI" class="t
wython-url">bit.ly/1FBUwAr</a>

Jokowi Naikkan Harga BBM, Dolar Keok ke Rp 11.700 dan IHSG Naik ke 5.400: Salah
satu pekerjaan utama Presiden ... <a href="http://t.co/RVGR71JGMy" class="twytho
n-url">bit.ly/1reC3jq</a>

You might not find it easy to get in touch with your own heart... More for Pisce
s <a href="http://t.co/Rqco6H1i0g" class="twython-url">bit.ly/xHSyh0</a>

楽天でキャバドレスを検索！
<a href="http://t.co/HeTk0wKkFE" class="twython-url">bit.ly/ZC2b05</a>
<a href="https://twitter.com/search?q=%23楽天" class="twython-hashtag"><a href="
https://twitter.com/search?q=%23楽天" class="twython-hashtag">#楽天</a></a> <a hr
ef="https://twitter.com/search?q=%23キャバドレス" class="twython-hashtag">#キャバ
ドレス</a>

楽天でシャネルを検索！
<a href="http://t.co/85T3jkORnd" class="twython-url">bit.ly/ZCdOUR</a>
<a href="https://twitter.com/search?q=%23楽天" class="twython-hashtag"><a href="
https://twitter.com/search?q=%23楽天" class="twython-hashtag">#楽天</a></a> <a hr
ef="https://twitter.com/search?q=%23シャネル" class="twython-hashtag">#シャネル</a
>
```

Figure 5.5: Tweets from Twitter

Ranking Table	
URLs Service Provider	Count
Bit.ly	139253
t.co	85321
Tiny.com	62478
Goo.gl	10122
ow.ly	8092
others	2300

Table 5.1: Ranking Table

have followed followed the same approach for the Facebook. But we have observed that accessing of the data from the Facebook on public domain is very restrictive. We have applied this method to collect both and short URLs. The following Table 5.1 shows the example of the short URLs that we have collected from the twitter.

Our total collection of URLs of nearly 312000 URLs compromising of about both long and short URLs. The following Ranking Tables give the of URLs collection.

The following ranking Table 5.1 gives us the information that the most number of users both genuine and attackers have heavily depended on the short URLs services for both spaces reduce methodology and for trading illicit content on the Social Networks. From the ranking Table we can conform that the most number of users has heavily used on bit.ly which has outset the other URL services which followed by the twitter short URL services t.co, followed by the Tiny.URLs, goo.gl, followed ow.ly and other(which constitute few of other shorter URL services). The above Table 5.1 show the number of short URLs that we have been collected from November 2014 to February 2015. From the above Table, we have been seen that Bit.ly has generated more no of short URLs followed by t.co and Tiny.com.

5.3 Detection Methodology

We have used wide range of online detection methodology Figure5.6 to detect whether the collected URLs are malicious or non-malicious . The following fig explain the method we have employ to detect the URLs are malicious or benign . Virus Total [14]

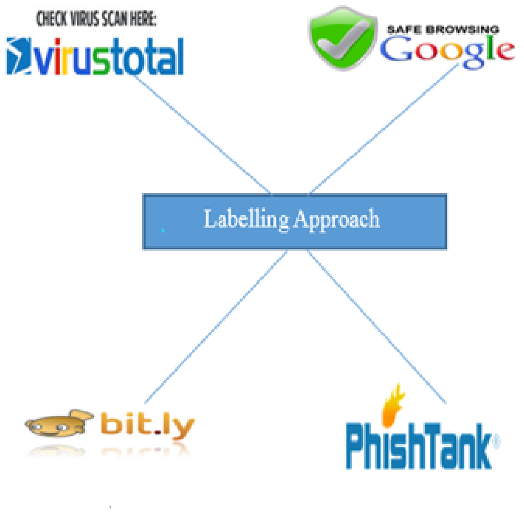


Figure 5.6: Labelling Approach

stores all the analyses it performs, this allows users to search for reports given an MD5, SHA1, SHA256 or URL. Search responses return the latest scan performed on the resource of interest. VirusTotal also allows you to search through the comments that users post on files and URLs, inspect our passive DNS data and retrieve threat intelligence de-tails regarding domains and IP addresses. Learn more about searching with Virus Total.

PhishTank Phishing is a fraudulent attempt usually employ by the attackers through email, to steal one's personal information for their benefit. The best to protect from such type of email is to know the behavior of the emails. Phishing emails generally appear that they are delivered by the well-know organizations and they personally entice one individual to theft their personal information such as credit card number, social security number, account number, or password. Most often such mails are received from the sender where the receiver does not have any account with them. Often the email of the one individual data is sold to the attackers by the third party. One should keep in mind that the legitimate organizer do not ask for the personal credential through insecure email.

- Generally there are three types of e-mail

Generic greeting Phishing email are sent in bulk. The Phish promotors usually buy the credential information from the third party and they the phishing emails in bulk. To incite the users in no time they use the “Generic Bank Customer “ so they don’t have to type all the customer name and they these phish mail [14] one-by-one. If one donot see their name they do not get any doubt.

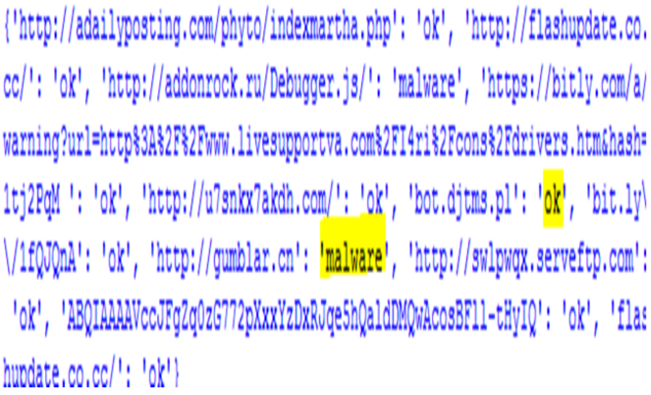
Produced connection. Regardless of the possibility that a connection has a name you perceive some place in it, it doesn’t mean it connections to the genuine association. Move your mouse over the connection and check whether it coordinates what shows up in the email. On the off chance that there is a discrepency, don’t tap on the connection. Additionally, sites where it is safe to enter individual data start with “https” — the “s” remains for secure. In the event that you don’t see “https” don’t contain.

Demands individual data. The purpose of sending phishing email is to deceive you into giving your own data. On the off chance that you get an email asking for your own data, it is presumably a phishing attempt. Feeling of criticality. Web culprits need you to give your own data now. They do this by making you think something has happened that obliges you to act quick. The speedier they get your data, the quicker they can proceed onward to another casualty.

Google Safe Browsing [5] Figure 5.7 is an online detector service available that provides the registered users API which enable the user to scan against the URLs. In return it checks the whether the query URLs are malicious and non-malicious against their frequently updated list of URLs . It check against the suspected phishing, malware and unwanted software.

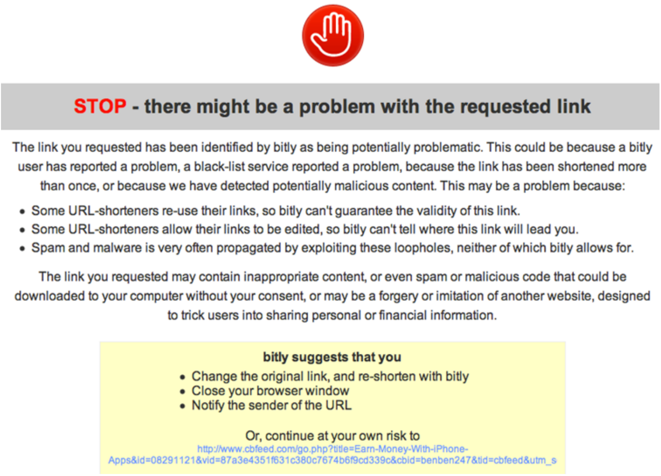
The following are the services that the Google Safe browsing provides us.

- Warns the user before click on the web link by highlighting as malicious or non-malicious.
- Prevents by posting the phishing URLs for on one’s site.
- Check against the malicious, phishing, worm against their updated list of domains.



```
{'http://adailyposting.com/phyto/indexmartha.php': 'ok', 'http://flashupdate.co.
cc/': 'ok', 'http://addonrock.ru/Debugger.js/': 'malware', 'https://bitly.com/a/
warning?url=http%3A%2F%2Fwww.livesupportva.com%2FI4ri%2Fcons%2Fdrivers.htm&hash=
1tj2PqM ': 'ok', 'http://u7snkx7akdh.com/': 'ok', 'bot.djtms.pl': 'ok', 'bit.ly\
\/1fQJQnA': 'ok', 'http://gumblar.cn': 'malware', 'http://swlpwgx.serveftp.com':
 'ok', 'ABQIAAAAVccJFgZq0zG772pXxxYzDxRJqe5hQaldDMQwAcosBFl1-tHyIQ': 'ok', 'flas
hupdate.co.cc/': 'ok'}
```

Figure 5.7: Google Safebrowsing Snippet



STOP - there might be a problem with the requested link

The link you requested has been identified by bitly as being potentially problematic. This could be because a bitly user has reported a problem, a black-list service reported a problem, because the link has been shortened more than once, or because we have detected potentially malicious content. This may be a problem because:

- Some URL-shorteners re-use their links, so bitly can't guarantee the validity of this link.
- Some URL-shorteners allow their links to be edited, so bitly can't tell where this link will lead you.
- Spam and malware is very often propagated by exploiting these loopholes, neither of which bitly allows for.

The link you requested may contain inappropriate content, or even spam or malicious code that could be downloaded to your computer without your consent, or may be a forgery or imitation of another website, designed to trick users into sharing personal or financial information.

bitly suggests that you

- Change the original link, and re-shorten with bitly
- Close your browser window
- Notify the sender of the URL

Or, continue at your own risk to
http://www.cbfeed.com/go.php?title=Earn-Money-With-iPhone-Apps&id=08291121&vid=87a3e4351f631c380c7674b6f9cd339c&cbid=benben247&tid=cbfeed&utm_s

Figure 5.8: Bit.ly Warning Page

Bit.ly is an short URL services. As bit.ly [3] does not apply any restrict to its user in service providing it can only raise the warning page as shown below. We have used above following online detection method to label our collected URLs which consist of both long and short URLs as malicious and non-malicious. We also checked the URLs against the public blacklists including McAfee, SiteAdvisor, URIBL, SURBL against constantly updated list of phishing and malware.

Chapter 6

Analysis and Feature Selection of URLs

6.1 Data Analysis

6.1.1 Creator Analysis

To know the behaviour of the attackers it is very important to know from where the attackers has posted the URLs in the SNS. The following are the listed below information from where the attackers have posted the URLs.

- twitterfeed
- tweetdeck Api
- roflquiz
- tweetmeme
- flocktome
- therealtwitter

6.2 Data Collection Approach

The twitter feed is utility allowed to feed content to Twitter to its registered user.

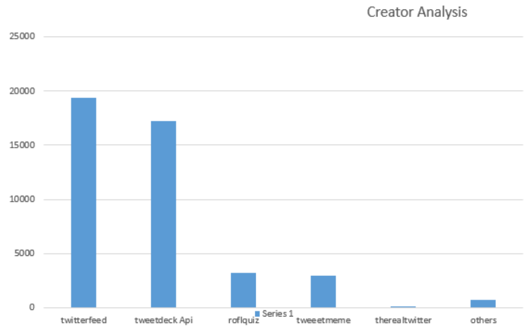


Figure 6.1: Creator Analysis

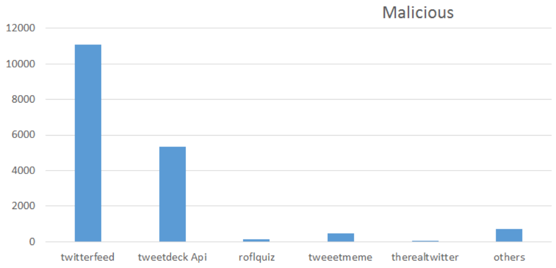


Figure 6.2: Malicious Creator Analysis

Tweetdeck API provides services to the Tweetdeck, an open social media dashboard application for management of Twitter.

Penguinfb is a service provided exclusively to send status updates from the Facebook.

Roflquiz is a website offering funny Twitter quizzes.

Tweetmeme is the popular organization which keeps track of popular links on Twitter.

The below Figure briefs that Tweetdeck with 19345 outset the remaining followed by Tweetdeck API, Roflquiz, Tweetmeme. We have conducted our analysis over nearly 52000 URLs.

We have analyzed these URLs against the online detectors to find whether they are malicious or non-malicious the following is the Figure 6.2 and Figure 6.1 help us to analyze the impact of the services to spread malicious on the SNS.

6.3 Data Collection Approach

6.3.1 Analysis of URLs

As stated earlier, we have extracted the URLs from the tweets [23] which contains both the short and long URL, and we also stated that the URLs are labelled by the online detectors.

6.3.2 Short URLs Analysis

Our dataset contains short URLs, thus it is very important to analyze the feature which helps to analyze the whether the URLs is malicious or non-malicious. The dataset contain the short URLs from the domain bit.ly [3], t.co [2], ow.ly, ad.ly, Goo.gl [5], Tiny.com etc.

The following are the basic general information that every short URL services provides to its registered users.

Basic information it gives the following attributes information. They are as follow

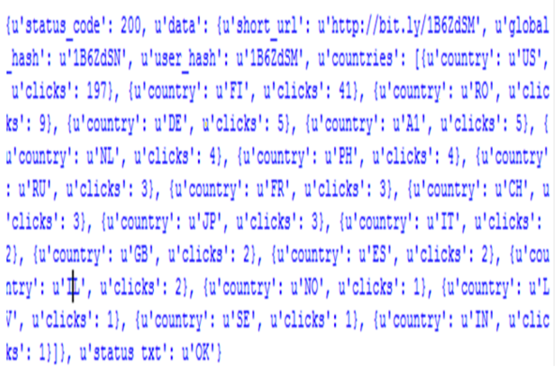
- **id :** it gives the details of the short URLs.
- **longurl :** it gives information of the long URLs to which short URL it points.
- **title :** it gives the title of the web page.
- **creator :** creator of the short URLs.

Click Source Analysis : In additional to the analysis of the short URLs by above feature we ,the most important among the short URL provided is the click source analysis [34]. There are two types of the clicksource services that most of the short URL services provides the are

- Country Source Analysis.
- Referrer Source Analysis.

Country Source Analysis The country source [22] click analysis Figure 6.3 will help us to analyse the from which country a there is a spread of the spam and legitimate URLs. During our data collection by the respective short URL services.

The following Figure6.4 shows the collection of the data URLs in the and the country clicks [15] on the spam URLs of sample 32133 URLs.



```
{u'status_code': 200, u'data': {u'short_url': u'http://bit.ly/1B6ZdSM', u'global
_hash': u'1B6ZdSN', u'user_hash': u'1B6ZdSM', u'countries': [{u'country': u'US',
 u'clicks': 197}, {u'country': u'FI', u'clicks': 41}, {u'country': u'RO', u'clic
ks': 9}, {u'country': u'DE', u'clicks': 5}, {u'country': u'A1', u'clicks': 5}, {
u'country': u'NL', u'clicks': 4}, {u'country': u'PH', u'clicks': 4}, {u'country'
: u'RU', u'clicks': 3}, {u'country': u'FR', u'clicks': 3}, {u'country': u'CH', u
'clicks': 3}, {u'country': u'JP', u'clicks': 3}, {u'country': u'IT', u'clicks':
2}, {u'country': u'GB', u'clicks': 2}, {u'country': u'ES', u'clicks': 2}, {u'cou
ntry': u'IL', u'clicks': 2}, {u'country': u'NO', u'clicks': 1}, {u'country': u'L
V', u'clicks': 1}, {u'country': u'SE', u'clicks': 1}, {u'country': u'IN', u'clic
ks': 1}]}, u'status txt': u'OK'}
```

Figure 6.3: Geographical Analysis

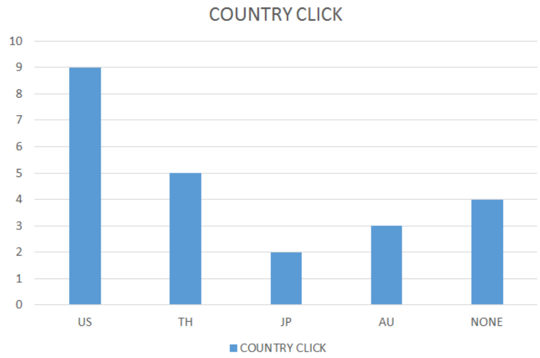


Figure 6.4: Total number of clicks by each country

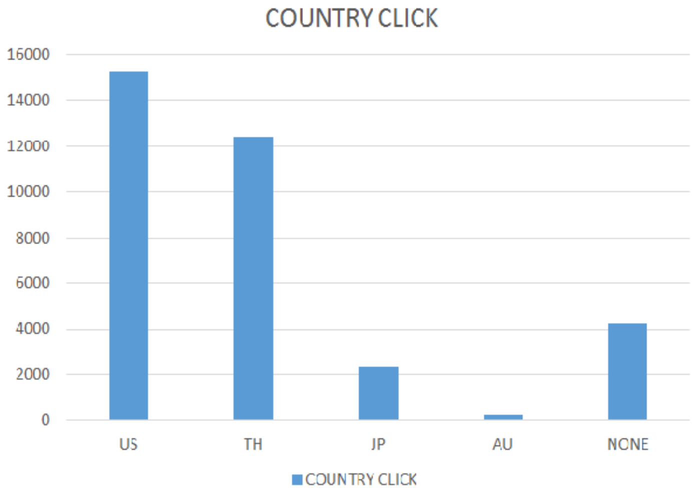


Figure 6.5: Click on Malicious URLs

References of Short URls	
direct	Click
http://twitter.com	392112
http://twitter.com/home	23244
http://www.yputube.com/watch	89322
http://www.facebook.com	32435
domain	to know the domain of the URL
Clicks	click count of the requested URLs
countries	countries reffering click traffic
encoder	the user who encode the user
reffers	return about page referring click traffic
refferbydomain	click group by referring short url domain
history	created before ,created after information
prodoamin	gives information about the domain of the short URL
doaminclick	click on the particular domain
domainshiorterncount	specfied tracking domain

Table 6.1: Feature Provided by the Short URL services

The above Figure6.5 gives us the info about the number of URLs that each country has click counts and on both the spam and legitimate URLs. It help us to analyse the from which country the both long and short URLs are spreading more. From the above it is clear that the US to in spreading both the short and long url which consist of legitimate and spam followed by the japan and thailand. Hence from these we can conclude that these 3 top most countries US, JP, TH top the list in spreading spam URL in SNSs.

Referrer source analysis is the webpage to which the short URLs is referred. It help us to analyse that the URLs that are provided by the short URLs are not only used in the SNSs but also in the email. But the spammers has used many of the SNS like Twitter and Facebook[1] etc.

The following feature Table6.1 show the number of the referrers of spam URLs based on links.

6.3.3 Domain Analysis

Domain names are used to identify the source of the IP addresses. For example www.nitrkl.ac.in whose Domain name is *nitrkl.ac.in* which consist of many IP address

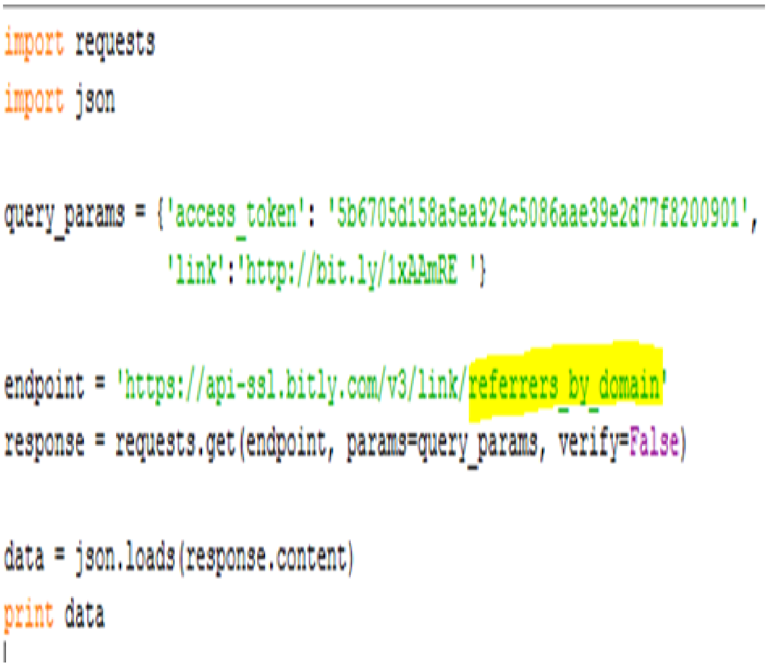
. Domain names are used to identify the particular webpage. Every domain has a suffix that indicates which Top Level Domain(TLD) is belongs to Domain names are used to identify one or more IP addresses. For example, the domain name microsoft.com represents about a dozen IP addresses. Domain names are used in URLs to identify particular Web pages. For example, in the URL *http://www.pcwebopedia.com/index.html*, the domain name is *pcwebopedia.com*. Every domain name has a suffix that indicates which (TLD) it belongs to. For example

- gov - Government agencies
- edu - Educational institutions
- org - Organizations (nonprofit)
- mil - Military
- com - commercial business
- net - Network organizations
- ca - Canada
- th - Thailand

Because the Internet is based on IP addresses, not domain names, every web server requires a Domain Name System (DNS) server to translate domain names into IP addresses.

Because of the availability of the Domains in the market at a cheaper or lesser rate the attackers are using the Domain as the vehicle to carry out their illicit work. Hence in our research work we have analyse the Domain of the most of the spam URLs. From this, we have analyse that the attacker are just concentrating on most of the spam URLs form the few concentrating domains Figure6.6 only. The following Tables 6.1 shows the few top URLs that the spammers are employ.

Domain age Most of the spammers have registered that their domain which has lives for short time. This showed that the malicious domain is the short lives domain.



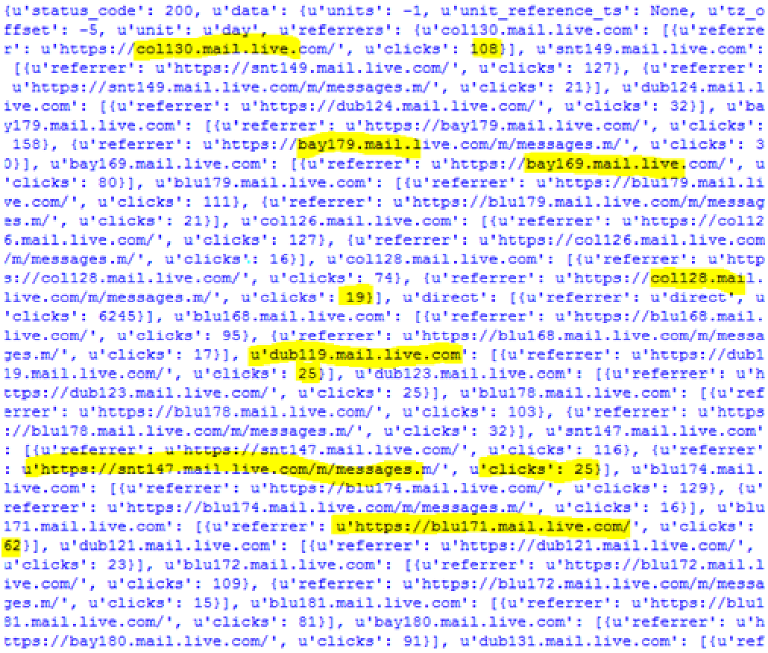
```
import requests
import json

query_params = {'access_token': '5b6705d158a5ea924c5086aae39e2d77f8200901',
                'link':'http://bit.ly/1xAAmRE '}

endpoint = 'https://api-ssl.bitly.com/v3/link/referrers_by_domain'
response = requests.get(endpoint, params=query_params, verify=False)

data = json.loads(response.content)
print data
```

Figure 6.6: Snippet for Domain Analysis



```
{u'status_code': 200, u'data': {u'units': -1, u'unit_reference_ts': None, u'tz_o
ffset': -5, u'unit': u'day', u'referrers': {u'col130.mail.live.com': [{u'referre
r': u'https://col130.mail.live.com/', u'clicks': 108}], u'snt149.mail.live.com':
 [{u'referrer': u'https://snt149.mail.live.com/', u'clicks': 127}, {u'referrer':
 u'https://snt149.mail.live.com/m/messages.m/', u'clicks': 21}], u'dub124.mail.l
ive.com': [{u'referrer': u'https://dub124.mail.live.com/', u'clicks': 32}], u'ba
y179.mail.live.com': [{u'referrer': u'https://bay179.mail.live.com/', u'clicks':
 158}, {u'referrer': u'https://bay179.mail.live.com/m/messages.m/', u'clicks': 3
0}], u'bay169.mail.live.com': [{u'referrer': u'https://bay169.mail.live.com/', u
'clicks': 80}], u'blu179.mail.live.com': [{u'referrer': u'https://blu179.mail.li
ve.com/', u'clicks': 111}, {u'referrer': u'https://blu179.mail.live.com/m/messag
es.m/', u'clicks': 21}], u'col126.mail.live.com': [{u'referrer': u'https://col12
6.mail.live.com/', u'clicks': 127}, {u'referrer': u'https://col126.mail.live.com
/m/messages.m/', u'clicks': 16}], u'col128.mail.live.com': [{u'referrer': u'http
s://col128.mail.live.com/', u'clicks': 74}, {u'referrer': u'https://col128.mail.
live.com/m/messages.m/', u'clicks': 19}], u'direct': [{u'referrer': u'direct', u
'clicks': 6245}], u'blu168.mail.live.com': [{u'referrer': u'https://blu168.mail.
live.com/', u'clicks': 95}, {u'referrer': u'https://blu168.mail.live.com/m/messa
ges.m/', u'clicks': 17}], u'dub119.mail.live.com': [{u'referrer': u'https://dub1
19.mail.live.com/', u'clicks': 25}], u'dub123.mail.live.com': [{u'referrer': u'h
ttps://dub123.mail.live.com/', u'clicks': 25}], u'blu178.mail.live.com': [{u'ref
errer': u'https://blu178.mail.live.com/', u'clicks': 103}, {u'referrer': u'https
://blu178.mail.live.com/m/messages.m/', u'clicks': 32}], u'snt147.mail.live.com'
: [{u'referrer': u'https://snt147.mail.live.com/', u'clicks': 116}, {u'referrer'
: u'https://snt147.mail.live.com/m/messages.m/', u'clicks': 25}], u'blu174.mail.
live.com': [{u'referrer': u'https://blu174.mail.live.com/', u'clicks': 129}, {u'
referrer': u'https://blu174.mail.live.com/m/messages.m/', u'clicks': 16}], u'blu
171.mail.live.com': [{u'referrer': u'https://blu171.mail.live.com/', u'clicks':
62}], u'dub121.mail.live.com': [{u'referrer': u'https://dub121.mail.live.com/',
u'clicks': 23}], u'blu172.mail.live.com': [{u'referrer': u'https://blu172.mail.l
ive.com/', u'clicks': 109}, {u'referrer': u'https://blu172.mail.live.com/m/messa
ges.m/', u'clicks': 15}], u'blu181.mail.live.com': [{u'referrer': u'https://blu1
81.mail.live.com/', u'clicks': 81}], u'bay180.mail.live.com': [{u'referrer': u'h
ttps://bay180.mail.live.com/', u'clicks': 91}], u'dub131.mail.live.com': [{u'ref
```

Figure 6.7: Collection for Domain Analysis

Domain Analysis		
URL	Doamin	count
http://rnqqvdxji.myftp.com	myftp.com	1213
http://ogqumk.serveftp.com	serveftp.com	1000
http://tfdeqkrfj.servettp.com	serveftp.com	1324
http://fmvks.serveftp.com	serveftp.com	4343
http://crjstqrmeb.myftp.com	myftp.com	4754
http://aapyczwshw.myftp.com	myftp.com	2134
http://jrkgaxeust.myftp.com	myftp.com	5634
http://fkwwoqdr.myftp.com	myftp.com	2324
http://fyeyck.myftp.com	myftp.com	1043
http://fyeyck.myftp.com	myftp.com	5356
http://aapyczwshw.myftp.com	myftp.com	2134
http://ozpojs.myftp.com	myftp.com	4355
http://aapyczwshw.myftp.com	myftp.com	4355

Table 6.2: Domain Analysis

The spammers has used these techniques to evades the security check . When we have used the APG anti-virus to check the domain we have analyse that the domain is no more in active state.

From the Table 6.2 we can conclude that the attackers are mainly dependent on few domains to trade their illicit malicious content on the SNSs. We have also concluded that these domains URLs are short lived. By these adopted mechanism we can say that the there is community who does these specific work in a secure manner such that within less time they are able to achieve their target in less time.

6.4 Analysis of Long URLs

Many research work has done for the analysis of the features of the long URLs which help them to detect the malicious and non-malicious URLs. It is fact that the short URLs by themselves does not include the physical properties of the URLs hence it is much needed them to convert to long URL to study some of the available features Some of the features that we used to analyse their behaviour as follow.

Though crawling of the web page is the most effective method to evaluate the URLs as malicious and non-malicious but it turned to a very difficult as it consume more time.

Long URLs Analysis		
Feature Description	Malicious(%)	Non-malicious(%)
Username in URLs	0.33	0.08
Password in URLs	0.23	0.00
’_ín Querypart 3.45	10.43	
Digit[0-9] in the host	30.06	3.11
’ín URL	2.02	9.03

Table 6.3: KeyWord features of Long URL

Therefore, we adopted to know the available URL feature behaviour to analyse them. Some of the feature used in our research as below.

Host based features: These features help us to analyse the features of the URLs based on the hostname portion of the URLs. They help us to approximate the to decide whether the URL are malicious or not. They are able to identify us to to know ‘owns’ the URLs and form which geographical location they are managed. Some of the features are explained below.

- **WHOIS information :** This feature gives us the information about the domain name registered dates, there registration etc. So if a set of malicious domains are registered by the some individual we have concluded such registered as an malicious URLs.

- **Location :** This refers to the hosts geography, IP address prefix, and autonomous system (AS) number.

- **Members in blacklists :** To mark some of the domain as a malicious or not we have used third party submitted blacklist against the marked list of domains.

- **Red Flag Key words :** This red flag word Table 6.3 are the standard word which are used by the research in determining the URLs are malicious or not. We have used some of the most popular feature in determining the maliciousness of the URLs.

Chapter 7

Results

7.1 Approach for Classicication

Classification is considered as an light weight operation for analyse the URLs whether it is a malicious or non-malicious. Since though the crawling webpage method is considered as an most effective method in reality it come with the time as a cost. In our research we have used most popular Machine tool 'Weka' as an classifier for our dataset.

7.1.1 Machine Learning Classification

Machine learning classifier act as a teacher, which predict probability instance class based on the predefine features. Weka is a free software package with various classifier and we use Naïve Bayes Classifier. In this section we discuss how we organise our dataset to classify using the 'Weka software'.

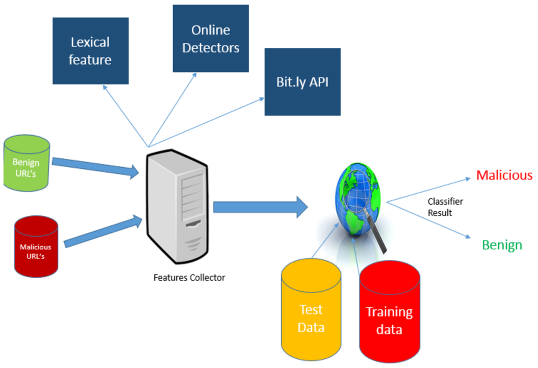


Figure 7.1: Classification Model

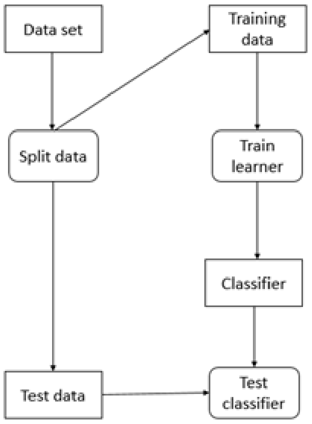


Figure 7.2: Machine Learning Classifier

Now we describe the way of classification of malicious users. Initially dataset is divided into training dataset (80%), testing dataset (20%). In order to assess the most efficient mechanism to detect malicious accounts, we inspected various machine learning algorithm. All below classifiers are the standard classifiers and widely used in solving problems.

7.1.2 Supervised Learning Algorithms

The following is the detail description about the classifiers.

Naïve Bayes Classifier

Naïve based classifiers is based on the probability and based on applying Bayes theorem with strong independence assumption. The descriptive term for the above probability model is “independent feature model”.

Naïve Bayes classifier assumes that particular class feature presence or absence is unrelated to the other class feature presence or absence. In this classifier, we have a hypothesis that the given data belongs to the related class. Precise nature of the probability model, in supervised learning settings we can train naïve Bayes classifier very efficiently. In many practical applications, it uses maximum likelihood for parameter estimation. In many complex real world situations, naïve Bayes classifier works well. The advantage of naïve Bayes classifier is that for estimate the parameters it require only the small amount of training data.

Naïve Bayes probabilistic model

The probability model is a conditional model over a dependent class variable with limited number of outcomes means classes, conditions on the feature variables F1 to Fn.

$$P(c/F1, F2....Fn) \tag{7.1}$$

If the value of n is large, basing a model is infeasible. Then we reformulating the model then it feasible or tractable.

$$P(c/F1, F2....Fn) = ((p(c)p(F1...Fn))/c)/(p(F1.....Fn)) \tag{7.2}$$

The above equation can be written plain English as follows
posterior=(prior*likelihood)/evidence

In reality we are only concentrating on numerator, because denominator not depending on the class c and values of features F.

Decision Tree Classifier

Decision tree most popular classifier which generates a tree like structure feature names corresponding to internal nodes feature values corresponding to branches, and class labels corresponding to leaf nodes. In this each node represents the test on the attributes i.e. decisions of the attribute. If the attribute is satisfies the required condition based on that it divide the data. Tree display the relationships among attributes are there in the training data set. Decision tree is predictive model that uses a set of binary rules applied to calculate the target value. Constructing the decision tree is done by selecting the attributes that splits the training data in proper class i.e. legitimate and malicious classes. Decision trees implemented based on the information gain. Which is based on the entropy. If the entropy is low then the set is homogeneity of type and if entropy is zero then the set is contains only one type of data. Once identified splitting attribute then rest of the training data are pushing down the tree i.e. data that is satisfying the splitting criteria are thrown into the “true” side of the tree. While, if the data is not satisfies the required criteria are thrown into the “false” side of the tree. The above process is repeated until the each node in the tree contains data of the same class, at that it store the class label. During the classification, it predicts the class of an unknown data based on criteria defined over the node, starting from the root node. If the attribute in the data satisfies the condition then the classifier follows the “YES” class. If not satisfies then it follows the “NO” class. It checks the each criteria in the right path until reaching the leaf nodes.

Evaluation Metric From the confusion matrix table 7.1 Accuracy (A) and F-measure are the metrics which are used for the evaluation of the classifier performance. F- Measure is defined in terms of Recall (R) and Precision (P). If evaluation metrics having higher value, then the classifier is best suitable for data set. The evaluation metrics described effectively by confusion matrix.

Table 7.1: Confusion Matrix

	Malicious	Legitimate
Malicious	TP	FN
Legitimate	FP	TN

True Positive (TP): Malicious samples are labelled as a Malicious.
False Negative(FN): Malicious samples are labelled as a benign(non-malicious).
False Positive (FP): Benign samples are labelled as a Malicious.
True Negative (TN): Benign samples are labelled as a Malicious.

$$P = \frac{TP}{(TP + FP)} \tag{7.3}$$

$$R = \frac{TP}{(TP + FN)} \tag{7.4}$$

$$F - Measure = \frac{2 * (P * R)}{(P + R)} \tag{7.5}$$

$$A = \frac{(TP + TN)}{TP + FN + FP + TN} \tag{7.6}$$

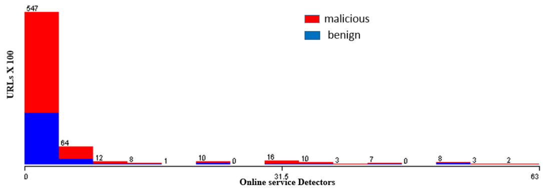


Figure 7.3: URLs vs Online Detectors

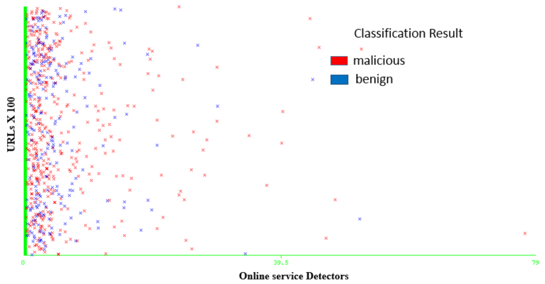


Figure 7.4: Classify URLs vs Online Detectors

7.2 Our Experimental Results

Our experimental result is briefly explained as below

URLs vs. Online Service Detectors

The Figure7.3 and Figure7.4 URLs against Online Service Detectors. We have gone through many online service detectors like Virus Total, Sucuri, Bit.ly, Phish Tank .Totally we went through 63 Online Detectors Services. From the Figure 7.3and Figure 7.4we can see that top 3 detection services (Fortinet, Kaspersky, Netcraft) are able to detect most of the URLs as malicious and non-malicious.

URLs vs. Clicks

The Figure7.5 and Figure7.6 shows Number of clicks against URLs. We can see

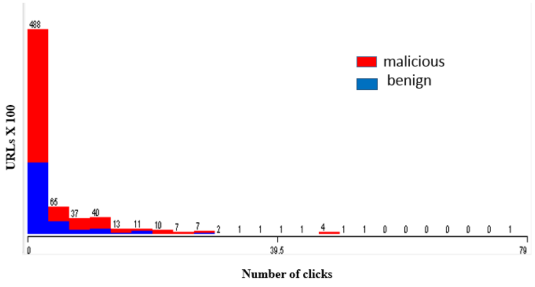


Figure 7.5: URLs vs. Clicks

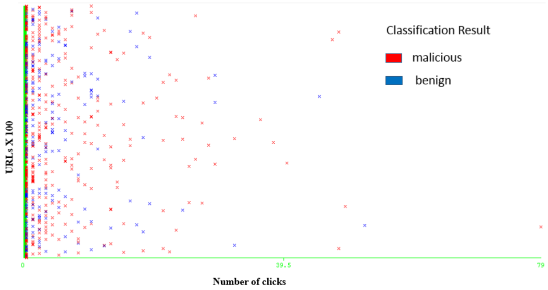


Figure 7.6: Classify URL vs. Clicks

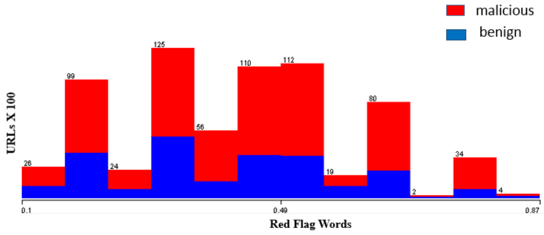


Figure 7.7: URLs vs. Geographical Analysis

from the Figure 7.5 and Figure 7.6 most of the malicious URLs are in between 0-7 clicks only. The reasons behind for the low clicks may be the spam filters have detected early these malicious URLs and blocked them.

The Figure 7.7 and Figure 7.8 shows URLs against Red Flag Words(for example the URLs may contain Freemoney, Congrates, Winner, Please Fill). Here each Red Flag Words are given a standard value Ranging from 0-1 in percentage.

7.2.1 Classification Result

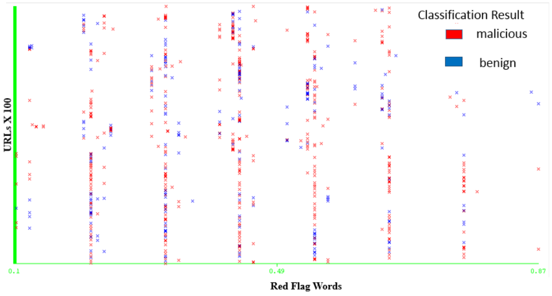


Figure 7.8: Classify URLs vs. Geographical

Table 7.2: Comparison of Classifiers

Evalution Metric	**Naive Bayesian**	**Decision Tree**	**K-Star**
Accuracy	**95.41%**	89.9%	90.4%
F-Measure(Malicious)	**81.0%**	64.4%	76.3%
F-Measure(Legitimate)	**95.5%**	93.4%	94.0%
True Positive Rate	88.2%	80.9%	79.0%
False Positive Rate	93.6%	90.1%	93.0%
Positive Predictive Rate	74.9%	53.5%	73.7%
Negative Predictive Rate	97.3%	97.1%	94.8%

From the above Table 7.2.1, we can see that Naïve Bayes outperforms the other classification algorithm with an accuracy of 95.41%.

7.3 Conclusion

Our research includes the large-scale experimental study on detection of spam URLs which include both long and short Urls. We have nearly collected of about 1,20000 URLs which consist of nearly 44.8% are spam URLs. Our collection doesn't confirm to only specific services doamin ,but our collection includes URLs from various domain like Ow.ly,t.co ,Bit.ly ,goo.gl etc. Our experiment includes 9 standard features of both short and long URLs. In our classification Naïve Bayes Classification resulted as an optimum classification with 95.41% accuracy. We have successfully shown the security gap that existed in many of the Social Networks.

7.4 Limitation

To find out the some features of the URLs we dependent on online sources like Virus Total, Phish Tank and Bit.ly. We cannot guarantee about the credibility of these services. In future they may change their evaluation method to detect the URLs malicious or not.

Bibliography

[1] FacebookDeveloper.`https://developers.facebook.com/docs/apps/`.

[2] TwitterDeveloper.`http://https://dev.twitter.com/`.

[3] BitlyDeveloper.`http://dev.bitly.com/index.html`.

[4] VirusTotalDeveloper.`https://www.virustotal.com/en/community/`.

[5] GoogleSafeBrowsing.`https://developers.google.com/safe-browsing/`.

[6] Daron Acemoglu, Munther A Dahleh, Ilan Lobel, and Asuman Ozdaglar. Bayesian learning in social networks. *The Review of Economic Studies*, 78(4):1201–1236, 2011.

[7] Daron Acemoglu, Asuman Ozdaglar, and Ali ParandehGheibi. Spread of (mis) information in social networks. *Games and Economic Behavior*, 70(2):194–227, 2010.

[8] Ross Anderson and Tyler Moore. The economics of information security. *Science*, 314(5799):610–613, 2006.

[9] Fabrício Benevenuto, Tiago Rodrigues, Meeyoung Cha, and Virgílio Almeida. Characterizing user behavior in online social networks. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pages 49–62. ACM, 2009.

[10] Andre Bergholz, Jeong Ho Chang, Gerhard Paaß, Frank Reichartz, and Siehyun Strobel. Improved phishing detection using model-based features. In *CEAS*, 2008.

[11] Leyla Bilge, Engin Kirda, Christopher Kruegel, and Marco Balduzzi. Exposure: Finding malicious domains using passive dns analysis. In *NDSS*, 2011.

[12] Aaron Blum, Brad Wardman, Thamar Solorio, and Gary Warner. Lexical feature based phishing url detection using online learning. In *Proceedings of the 3rd ACM Workshop on Artificial Intelligence and Security*, pages 54–60. ACM, 2010.

[13] Chia-Mei Chen, DJ Guan, and Qun-Kai Su. Feature set identification for detecting suspicious urls using bayesian classification in social networks. *Information Sciences*, 289:133–147, 2014.

[14] Sidharth Chhabra, Anupama Aggarwal, Fabricio Benevenuto, and Ponnurangam Kumaraguru. Phi. sh$ocial: the phishing landscape through short urls. In *Proceedings of the 8th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference*, pages 92–101. ACM, 2011.

[15] Hyunsang Choi, Bin B Zhu, and Heejo Lee. Detecting malicious web links and identifying their attack types. In *Proceedings of the 2nd USENIX conference on Web application development*, pages 11–11. USENIX Association, 2011.

[16] Nicole B Ellison et al. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, 2007.

[17] Ralph Gross and Alessandro Acquisti. Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, pages 71–80. ACM, 2005.

[18] DJ Guan, Chia-Mei Chen, and Jia-Bin Lin. Anomaly based malicious url detection in instant messaging. In *Proceedings of the joint workshop on information security (JWIS)*, 2009.

[19] Huaping Hu and Jianli Wei. Instant messaging worms propagation simulation and countermeasures. *Wuhan University Journal of Natural Sciences*, 12(1):95–100, 2007.

[20] Emily M Jin, Michelle Girvan, and Mark EJ Newman. Structure of growing social networks. *Physical review E*, 64(4):046132, 2001.

[21] Vasileios Kandylas and Ali Dasdan. The utility of tweeted urls for web search. In *proceedings of the 19th international conference on World wide web*, pages 1127–1128. ACM, 2010.

[22] Florian Klien and Markus Strohmaier. Short links under attack: geographical analysis of spam in a url shortener network. In *proceedings of the 23rd ACM conference on Hypertext and social media*, pages 83–88. ACM, 2012.

[23] Rui Li, Kin Hou Lei, Ravi Khadiwala, and KC-C Chang. Tedas: A twitter-based event detection and analysis system. In *Data engineering (icde), 2012 ieee 28th international conference on*, pages 1273–1276. IEEE, 2012.

[24] Federico Maggi, Alessandro Frossi, Stefano Zanero, Gianluca Stringhini, Brett Stone-Gross, Christopher Kruegel, and Giovanni Vigna. Two years of short urls internet measurement: Security threats and countermeasures. In *proceedings of the 22nd international conference on World Wide Web*, pages 861–872. International World Wide Web Conferences Steering Committee, 2013.

[25] Niels Provos, Dean McNamee, Panayiotis Mavrommatis, Ke Wang, Nagendra Modadugu, et al. The ghost in the browser analysis of web-based malware. In *Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets*, pages 4–4, 2007.

[26] Sam Ransbotham and Gerald C Kane. Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in wikipedia. *MIS Quarterly-Management Information Systems*, 35(3):613, 2011.

[27] Lisa Singh and Justin Zhan. Measuring topological anonymity in social networks. In *Granular Computing, 2007. GRC 2007. IEEE International Conference on*, pages 770–770. IEEE, 2007.

[28] Pravin Soni, Shamal Firake, and BB Meshram. A phishing analysis of web based systems. In *Proceedings of the 2011 International Conference on Communication, Computing & Security*, pages 527–530. ACM, 2011.

[29] Gianluca Stringhini, Christopher Kruegel, and Giovanni Vigna. Detecting spammers on social networks. In *Proceedings of the 26th Annual Computer Security Applications Conference*, pages 1–9. ACM, 2010.

[30] Brian K Tanner, Gary Warner, Henry Stern, and Scott Olechowski. Koobface: The evolution of the social botnet. In *eCrime Researchers Summit (eCrime), 2010*, pages 1–10. IEEE, 2010.

[31] Kurt Thomas. The koobface botnet and the rise of social malware. 2010.

[32] Kurt Thomas, Chris Grier, and David M Nicol. unfriendly: Multi-party privacy risks in social networks. In *Privacy Enhancing Technologies*, pages 236–252. Springer, 2010.

[33] Bimal Viswanath, Ansley Post, Krishna P Gummadi, and Alan Mislove. An analysis of social network-based sybil defenses. *ACM SIGCOMM Computer Communication Review*, 41(4):363–374, 2011.

[34] De Wang, Shamkant B Navathe, Ling Liu, Danesh Irani, Acar Tamersoy, and Calton Pu. Click traffic analysis of short url spam on twitter. In *Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom), 2013 9th International Conference Conference on*, pages 250–259. IEEE, 2013.

[35] David Weiss. The security implications of url shortening services, 2009.

[36] Peng Yali and Yu Min. Research of intrusion detection technology and its formal modeling. *International Journal of Information Technology and Computer Science (IJITCS)*, 1(1):33, 2009.

Dissemination

1. Murali B, Jintendra Rout, Prof. Sanjay Jena. *“Detection and Analysis of Suspicious URLs in Social Networks”* Paper Presented at *“International Conference on Communication, Information and Computing Technology (ICCICT-2015)” to be published in‘‘International Journal of Advance Foundation and Research in Computer”(IJAFRC)(2015)(Accepted).*