IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 3, March 2015.
www.ijiset.com
ISSN 2348 – 7968

Improved URL Shortening With Traffic Analysis and Spam Filter

[1]Ms. S.Rajeshwari,[2]R.Banupriya,[3]M.Manibharathi,[4]M.Samena

[1]AP/CSE,[2]UG Scholar
Department of Computer Science and Engineering
Jay Shriram Group of Institutions, Tirupur, Tamil Nadu, India

ABSTRACT

Uniform Resource Locator is the global address of documents and other resources on the website. The drawbacks of long URLs include mistyped URLs, unavailable domain names, cleanliness, SEO placement and intentionally misleading URLs. URL shortening is a service that translates long URLs into abbreviated alternatives.This service play an important role on the social networking for sharing URL on Twitter mainly due to 140 characters limit per message.The challenges in shortening URL is to detect and eliminate spammers from misusing URL Shortening sites and to maintain traffic analysis for every link.To overcome the above limitations, this approach involves 2 step verification process including Captcha and Checking with spamming URL which are predefined and traffic analysis which is done by recording each visit on local database and counting it accordingly. By proposing this technique security and privacy is achieved**.**

Keywords:URL, Spammer, Traffic Analysis

1.INTRODUCTION

This service is used for shortening the long URL into short URL. Because in twitter it restricts 140 character limit only so we can't twit to the long URL. We are going for shortening services. Spam filter aim is to detecting the spammer by using spam filter. Before going for shortening service we have to compare the link with predefined blacklist spam free domain.

Traffic analysis is used to know how many users have activated the link. BasedURL shortening services are services that use advertising and referral programsto encourage users to create and share short links by paying them a smallamount of money for every visitor their short URLs. For the user who creates theshort link, the process is similar to shortening a link with any otherURL shortening services. For instance, if a malicious user shares a link towards adrive-by download site whenever Twitter detects the threat, it can simply blockone single short URL and stop its users from visiting the malicious page, regardless of the connectivity status of the actual malicious page. If weshortening the longURL to short URL it will redirect the link to the original URL.

ISSN 2348 – 7968

2.RELATIVE WORK

2.1 WARNINGBIRD: Detecting Suspicious URLs in Twitter Stream:

Twitter is a well-known social networking and information sharing service that allows users to exchange messages of fewer than 140-character, also known as tweets, with their friends. When a user Alice updates (or sends) a tweet, this tweet will be distributed to all of her followers, who have registered Alice as one of their friends. Instead of distributing her tweets to all of her followers. Alice can send a tweet to a specific twitter user Bobby mentioning this user by including in the tweet. Unlike status updates, mentions can be sent to users who do not follow Alice. When Twitter users want to share URLs with friend's tweets, they usually use URL shortening services to reduce the length of these URLs, because tweets can only contain a restricted number of characters. Bit.ly and tinyurl.com are widely used services, and Twitter also provides its own shortening service t.co. Owing to the popularity of Twitter, malicious users often try to find a way to attack it. Most common forms of Web attacks, including spam, scam, phishing, and malware distribution attacks, have appeared on Twitter.

2.2Geographical Analysis of Spam in a URL Shortener Network:

URL Shortener services have been available for at least 10 years Such services i) take a long URL as an input, ii) offer a short URL in return, and iii) permanently redirect traffic from the short Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission. URL to the long URL. One of the first URL shortener services was to shortening URLs originated from problems observed with links in emails, where links were often re-wrapped by clients and thus rendered unclick able.

2.3Security and Privacy Implications of URL Shortening Services:

The advent of social networks and Twitter, on which messages are required to fit into 140 characters, compression of URLs becomes more and more crucial. In order to turn long URLs into shorter ones, so-called URL shortening services (USS) are offered by various entities. When users want to shorten a long URL, they submit the long URL to a shortening service that returns a short URL which typically does not exceed 30 characters.

The users then include the short URL instead of the long URL e.g. in a tweet, short message or e-mail. When others request the short URL, the shortening service automatically redirects them to the original, longer URL. An example for a short URL generated by the shortening service. The use of URL shortening service is very convenient, it imposes a number of risks for users submitting long URLs, users and machines requesting short URLs as well as the target servers hosting the resource to which a short URL points. Malicious or compromised uses may redirect users of vulnerable browsers to websites containing malware. Innocent users may submit a meant-to-be secret URL to a shortening service, but the service may leak this short URL to search engines. Also, an attacker may enumerate the short URLs generated by a USS in order to find secret URLs. Users requesting short URLs from a USS may be tracked by this USS with the help of cookies. Empirically analyze popular USS with respect to these risk. We empirically determine the most popular users are using twitter 140 character. We empirically determine the most popular uses used in over 7 million spam messages collected in the past seven years. Users want to shorten a long URL they submit long URL.

2.4 Detecting Spammers on Social Networks:

Social networking sites have become one of the main ways for users to keep track and communicate with their friends online. Sites such as Facebook, Myspace, and Twitter are consistently among the top 20 most-viewed web sites of the Internet. The tremendous increase in popularity of social networking sites allows them to collect a huge amount of personal information about the users, their friends, and their habits. Unfortunately, this wealth ofinformation, as well as the ease with which one can reach many users, also attracted the interest of malicious parties. In particular, spammers are always looking for ways to reach new victims with their unsolicited messages. This is shown by a market survey about the user perception of spam over social networks.83% of the users of social networks have received at least one unwanted friend request or message From a security point of view, social networks have unique characteristics. First, information access and interaction is based on trust. Users typically share a substantial amount of personal information with their friends. This information may be public or not. If it is not public, access to it is regulated by a network of trust. In this case, a user allows only her friends to view the information regarding herself. Unfortunately, social networking sites do not provide strong authentication mechanisms, and it is easy to impersonate a user and sneak into a person's network of trust.

2.5Detecting and Characterizing Social Campaigns:

Online social networks (OSNs) are popular collaboration and communication tools for millions of Internet users. Facebook alone boasts over 500 million users, and has recently surpassed Google as the most visited site on the Internet. As communities built out of Friends, family, and acquaintances, the public perception of OSNs is that they provide a more secure environment for online communication, free from the threats prevalent on the rest of the Internet. In fact, a study of a social auction site demonstrated that a social network could indeed provide a protective environment. Unfortunately, recent evidence shows that these trusted communities can become effective mechanisms for spreading malware and publishing attacks. Popular OSNs are increasingly becoming the target of phishing attacks launched from large botnets and OSN account credentials are already being sold online in underground forums. Using compromised or fake accounts, attackers can turn the trusted OSN environment against its users by masquerading spam messages as communications from friends and family member. Wall messages are the intuitive place to look for attempts to spread malicious content on Facebook.

2.6 Understanding and Detecting Malicious Web Advertising:

Both industry and academia have been working on this threat, typically through inspecting ads to detect their malicious content. However, malicious ads often use obfuscation and code packing techniques to evade detection. Further complicating the situation is the pervasiveness of ad syndication, a business model in which an ad network sells and resells the spaces it acquires from publishers to other ad networks and advertisers. Ad syndication significantly increases the chance of posting malicious content on a big publisher's Web site. It allows a malicious ad network to deliver ads directly to a user's browser, without the need of submitting them through the more reputable ad networks and publishers from whom it gets the ad space.

Furthermore, attackers continue to invent new, stealthy strategies for exploiting ad-delivery channels: a prominent example is leveraging a compromised publisher page to hijack user traffic. Both industry and academia have been working on this threat, typically through inspecting ads to detect their

malicious content However, malicious ads often use obfuscation and code packing techniques to evade detection. Further complicating the situation is the pervasiveness of ad syndication, business model in which an ad network sells and resells the spaces it acquires from publishers to other ad networks and advertisers. Ad syndication significantly increases the chance of posting malicious content on a big publisher's Web site. Hackers and con artists have found Web ads to be a low-cost and highly-effective means to conduct malicious and fraudulent activities.

2.7 The Utility of Tweeted URLs for Web Search:

The web, microblogging has become an increasingly popular form of blogging. One reason for its popularity is that the short post length requirement demands little time commitment for the users. But microblogging is not only used like regular blogs. Many new uses have come up that were not originally intended For example, microblogging sites like Twitter are being used to recommend popular articles in real-time, to track breaking news stories for work-related communication or for brand marketing Regardless of the type of use, one common element is the frequent link exchange that occurs through posts and the shortening of the posted links to conform to the maximum allowed post length. In this paper we look at the quality of these URLs and their value for a web search engine.

3. EXISTING SYSTEM

A URL shortening service creates compact alias URLs for longer URLs that, when visited, redirect the user to the original long URLs. While, initially, URL shortening services were used primarily for their shortening functionality, nowadays users utilize them even when there are no space limitations, as a way of beautifying "their links and tracking user clicks unfortunately, attackers found URL shortening services equally useful and started shortening links towards malicious pages. By spreading the generated short URLs instead of their original ones, attackers could evade blacklists and filtering systems looking for suspicious patterns in URLs, or simply exploit the fact that users consider short URLs as link. One special type of URL shortening services are ad-based URL shortening services, like adf.ly. These services, in addition to shortening long URLs, pay the link-shortening users a small commission every time a user clicks on their shortened link. The services generate income by exposing the visitor of short URLs to advertisements before redirecting them to the final destination. Although some of the previous work on generic URL shortenerincluded some ad based shortener in their list of studied services the ad-based URL shortening services were treated in the exact same way as the rest of the traditional URL shortener.

3.1DISADVANTAGES

- **Linkrot:**If a user, Alice, shares the link of an image of an inspiring scenery with the social network.
- **Hijacking:** The issue of link rot, hijacking can occur if an attacker manages to change the destination of short URLs and redirect the visiting users to pages under his control.
- **Obfuscation and Maliciousness:** The users may utilize a URL shortening service for beautification purposes. Can be abused by attackers, who use such services in order to hide the final, malicious destination.

4. PROPOSED SYSTEM

The URL Shortening services are used for shortening the long URL to the short URL because in twitter it restricts 140 character limit only so we can't Twitt to the long URL. So we are going for shortening services. Our main aim is to detecting the spammers by using spam filter. Before going for shortening we have to compare the link with predefined blacklist spam free domain. We can calculate how many users have activated the link so we are using traffic analysis.URL(Uniform Resource Locator) is a web addresses or host name of the web browser it will be displayed on the Above page in the address bar. This service will be shortening the long URL to the short URL if we click the short URL it will redirect the page to the long URL detecting only spammers profiles.If the profiles is detected as a spammers users add the profiles to our detection spam set. We can find whether the user is spammers or not. And get the users profile either mobile or through email. When we are sending the mail to the user if user response for our mail then we will shorten their link and give to the particular user's.

4.1ADVANTAGES

- **Length Reduction:** Reducing the long URL to the Short URL. In some social network applications such as Twitter it restricts 140 character available to type their message.
- **Beautification:** It is not uncommon for certain links to include a large number of parameters with special values and control characters.
- **Analytics:**Whenever users share a link, they may want to inspect whether the users who received the link, actually visited the page. In many cases, however, the destinationURL is not under the control of the link-sharing user, i.e., the user does not have access to web analytics or web-server logs.
- **Centralized Control:**Some social networking and micro blogging sites wrap all user-produced links with their own URLs. As millions of users were redirected to the attackers' domains.

5.CONCLUSION

In this paper, we are shortening the long URL to short URL ad-basedURLshortening services from a security and privacy perspective.Because of certain unique attributes of these services, such as the monetary incentive for clicks on shortened URLs, by using several social networks like twitter, Facebook blogs, etc. by shortening the long URL it will redirect the page to the short url.by using spam filter we can detect the spammersdomain or predefined blacklist domain.

REFERENCES

ISSN 2348 – 7968

[1] Lee, S., and Kim, J. WarningBird: Detecting Suspicious URLs in Twitter Stream. In *Proceedings of the Network and Distributed System Security Symposium (NDSS '12)* (2012).

[2] Klien, F., and Strohmaier, M. Short links under attack: geographical analysis of spam in a URL shortener network. In *HT '12* (2012).

[3] Neumann, A., Barnickel, J., and Meyer, U.Security and Privacy Implications of URL Shortening Services. In *Web 2.0 Security and Privacy Workshop (W2SP '11)* (2011).

[4] Stringhini, G., Kruegel, C., and Vigna, G.Detecting spammers on social networks. In *Proceedings of the 26th Annual Computer Security Applications Conference (ACSAC '10)* (2010).

[5] Gao, H., Hu, J., Wilson, C., Li, Z., Chen, Y., and Zhao, B. Y. Detecting and characterizing social spam campaigns. In *Proceedings of the 10th Internet Measurement Conference (IMC '10)* (2010).

[6] Li, Z., Zhang, K., Xie, Y., Yu, F., and Wang, X.Knowing Your Enemy: Understanding and DetectingMalicious Web Advertising. In *Proceedings of the 19th ACM Conference on Computer and Communications Security (CCS '12)* (2012).

[7] Kandylas, V., and Dasdan, A. The utility of tweeted URLs for web search. In *Proceedings of the 19th International World Wide Web Conference(WWW '10)* (New York, NY, USA, 2010), ACM.

[8] Nick Nikiforakis†, Federico Maggi‡, GianlucaStringhini∗, M. ZubairRafique†, WouterJoosen†, Christopher Kruegel∗, Frank Piessens†, Giovanni Vigna∗, Stefano Zanero‡†iMinds-DistriNet, KU Leuven, "Stranger Danger: Exploring the Ecosystem of Ad-based URL Shortening Services". In International World Wide Web Conference Committee (IW3C2). IW3C2 reserves the right to provide a hyperlink to the author's site if the Material is used in electronic media. WWW'14, April 7–11, 2014, Seoul, Korea. ACM 978-1-4503-2744-2/14/04. http://dx.doi.org/10.1145/2566486.2567983.