A Run-time Detection System for Malicious URLs in Twitter

Harshad D. Nannaware.
Dept. of Computer Engineering
BVCOE & RI, Anjeneri,
Trimbakshwar, Nashik 422 213
harshadnannaware007@gmail.com

Tushar P. Dhangar.
Dept. of Computer Engineering
BVCOE & RI, Anjeneri,
Trimbakshwar, Nashik 422 213
tushar.dhangar68@gmail.com

Aniket S. Dhanrao.
Dept. of Computer Engineering
BVCOE & RI, Anjeneri,
Trimbakshwar,
Nashik 422 213
mr.aniket1992@gmail.com

Prof. V. D. Badgujar.
Dept. of Computer Engineering
BVCOE & RI, Anjeneri, Trimbakshwar,
Nashik 422 213
badgujarvivek83@gmail.com

Abstract**—** Twitter is web based application which provide the service of sharing and mentioning the tweets and messages, which is social networking service. Twitter message could be a context information, or a URLs. Spam, scam, malicious distribution can be done easily with these URLs, spams and malicious data are usually a suspicious URLs with share another URLs in it. These URLs are known as URLs redirect chains, it is a repetition URLs. They itself contain too many URLs which is called as spam URLs. Generally there are several schemes to find such spam URLs, and they are evasion using time – based and evasion using crawling. This system is design to find- detect such spam or suspicious URLs- tweets from many other tweets, The URLS redirect chains frequently share the same URLs, and have been reused usually by attackers because they have limited resources. A Run-time Detection System for Malicious URLs in Twitter has method to show and detect the suspiciousness of the correlated URLs redirect chains, which are commonly shared URLs

Keywords*- Twitter, URL redirection, classification, spam, suspicious URL.*

*****

I. INTRODUCTION

Twitter is very popular, simple, easy to use social networking and data, photos, message distributing and share application or services provider. Said by H. Kwak, C. Lee, H. Park, and S. Moon [2] Twitter has a limit of almost 140 characters of message to share or send twitter name it as "Tweet ". Twitter is special service provider which is famous for – when someone or account user send a message that is " tweet " or update some of data that updated message is send to all the followers to that account user. It also provide a general or special mentioning "tweet", example if we want to send a tweet to specific twitter user we have to mention them by including @John. Usually a user send a message or update a photo they using a URL's , twitter allow specific length of URLs has been said by D. Antoniades, I. Polakis, G. Kontaxis, E. Athanasopoulos, S. Ioannidis, E. P. Markatos, and T. Karagiannis [3] , since twitter allows restricted no of characters.

Basically there are following services used and they are as follows Bit.ly, tinyurl.com, etc. the spam and malicious can easy attack these URLs. the most common forms of web attacks, has been observed on twitter which also includes spam, phishing , scams ,etc. attackers used shortened malicious URLs because tweets are small in length which redirect the twitter user to external attack servers.

As of December 2014, Twitter has more than 500 million users said by S. Lee and J. Kim [1], out of which more than 284 million are active users. It is difficult to find such spam and malicious message.

There are schemes are proposed such as spam and scams detections, said by Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia [6] and therefore they can be basically distinguish by two ways and they are

a. Account feature based relations and
b. Message feature based schemes.

Therefore there are following different feature of spam accounts such as Ratio of tweets, date of account created, how much followers does twitter user have. Following features such as ratio of tweets containing URLs, date on which account created and twitter graph containing relation features, all these schemes are time consuming and resource consuming. Spam is a moving target and difficult to measure. URLs redirections, dynamic behavior, the lexical features of URLs and HTML content are basically utilize features which are conventional to suspicious URL detections.

A Run-time Detection System for Malicious URLs in Twitter - chains correlation of URLS redirected have been investigated and find out in tweets.

The URLS redirect chains frequently share the same URLs, and have been reused usually by attackers because they have limited resources said by D. K. McGrath and M. Gupta and D. Canali, M. Cova, G. Vigna, and C. Kruegel [4] & [5].

A Run-time Detection System for Malicious URLs in Twitter has method to show and detect the suspiciousness of the correlated URLs redirect chains, which are commonly shared URLs. Statistical classifier is built from all the tweets which are collected randomly from timeline of twitter public user.

Suspicious URLs are showed by result of our classifier which are accurate and efficient. A Run-time Detection System for Malicious URLs in Twitter has also been presented as a real time system for classifying suspicious URLs.

II. Related Work.

Today twitter is very popular social networking service. Spam, Scam and malicious attacks can easily enter to it and make it spam and malicious tweets. Spam, scam phishing and malicious distributions are the most common attacks which observed on twitters, they are web attack forms. To classify and detect such suspicious URLs are several schemes which are proposed said by D. K. McGrath and M. Gupta [4]. There are crawlers which may execute in virtual machine honeypots and they are static and dynamic crawlers, Captured – HPC are used to find URLs. To cloak the suspicious URLs they used a number of different domain names and number different of IP addresses, simply there are many twitter accounts and URLs redirected, to make complicate to find suspicious URLs they used long redirect chains. Such suspicious URLs are observed on timeline of twitter public user. The conventional methods cannot efficiently detect these long life and repetition suspicious URLs.

It is very difficult to investigate every tweet, we can filter all the tweets that contain suspicious URLs redirect chains and separate them from the normal tweets. Therefore URLs redirect chains which used same URLs usually are separated because their resource are generally limited and reused.

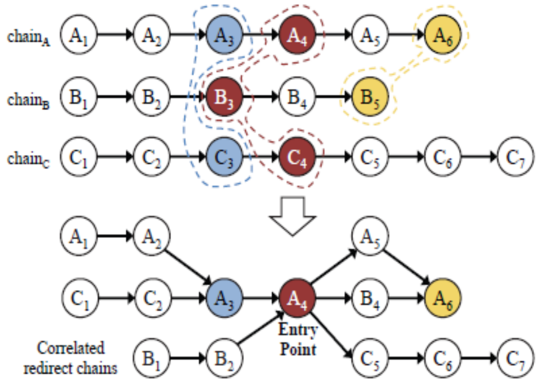

Fig : Redirect chain and their correlations

III. Proposed System

We proposed A Run-time Detection System for Malicious URLs in Twitter investigation and detection system which detect the suspicious URLs. Classification of all the URLs in every tweet is difficult and time – resource consuming on the bases of landing pages investigation. URLs redirect chains which used same URLs usually are separated because their resource are generally limited and reused.

A Run-time Detection System for Malicious URLs in Twitter has method to show and detect the suspiciousness of the correlated URLs redirect chains, which are commonly shared URLs and it can also display or show the IP address of the Attackers System.

Classification can be done by analyze the correlated URLs redirect chains and the information that tweets contain. Statistical classifier is built from all the tweets which are collected randomly from timeline of twitter public user.

IV. Mathematical Model

Over several days twitter public timeline has occurred with suspicious URLs redirect chain frequently. The classification process separate the benign – the message or tweets by active users and suspicious URLs – tweets and message updated by non-active user, are done to verify the reoccurrence of the suspicious URLs redirect chains. Let's consider the URL redirect chains for each day in and for 60 days check the average number of repeated of the extracted URL redirect chains. Let D denotes a set of days of a month , B(di) denotes a set of benign entry point URLs on di, S(di) denotes a set of entry point to suspicious URLs on di ∈ D, and A(di) denotes a set of all entry point URLs on di. For each di, j ∈ {j days later from di: 1 ≤ j ≤ 60, following equation can be computed.

$$\sum_{d_i \in \mathcal{D}} \left(\frac{|\mathbf{B}(d_i) \cap \mathbf{A}(d_{i,j})|}{|\mathbf{B}(d_i)|} \right) / |\mathcal{D}|,$$

$$\sum_{d_i \in \mathcal{D}} \left(\frac{|\mathbf{S}(d_i) \cap \mathbf{A}(d_{i,j})|}{|\mathbf{S}(d_i)|} \right) / |\mathcal{D}|,$$

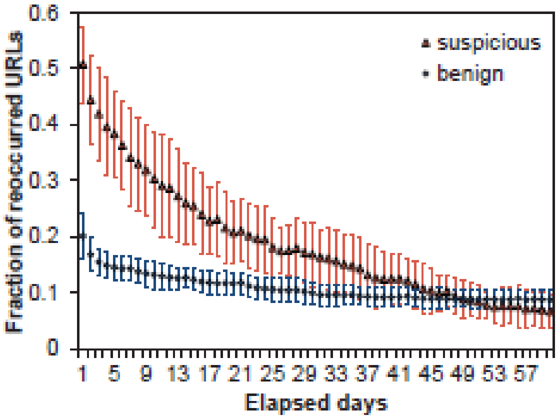

Fig. Average fraction of reoccurred URL redirect chains for 60 days.

Attackers use certain programs to manipulate these numbers because number of followers and number of friends associated with attacker's accounts are usually same. For checking such numerical similarities we use the standard deviation. Equation show that the number of followers and the number of friends of accounts of malicious entry point URLs are more similar than those of benign entry point URLs.

Similarity in the follower-friend ratio: We define the follower-friend ratio as below:

$$\frac{\min(\text{number of followers, number of friends})}{\max(\text{number of followers, number of friends})}.$$

The follower-friend ratios of attacker accounts are similar as with the number of followers and number of friends. We thus use the similarity (standard deviation) of these ratios as a feature. Above equation shows that the follower-friend ratios of accounts with malicious entry point URLs are more similar than those with benign entry point URLs.

Similarity of tweet texts: Tweeted texts containing the same URL are usually similar (e.g., retweets). Therefore, if the tweet texts associated with the same URL are different, we can assume that these tweets are related to suspicious behaviors because attackers usually want to change the appearance of malicious tweets that include the same malicious U the similarity between tweet texts as URL to evade detection. We measure the similarity between tweet texts as,

$$\sum_{t,u \in \text{a set of pairs in tweet texts}} \frac{J(t,u)}{|\text{a set of pairs in tweet texts}|},$$

Where, J (t, u) is the Jaccard index, which is a famous measure that determines the similarity between two sets t and u, and is defined as below:

$$J(t,u) = \frac{|t \cap u|}{|t \cup u|}.$$

V. METHODOLOGY

Methodology for development.
1. Data collection,
2. Features extraction
3. Training
4. Classification

1. Data collection :

There are two subcomponents in data collection and they are as follows:

a. The collection of tweets with URLs.
b. Crawling for URLs redirections.

The process of collecting tweets from timeline of twitter public account that contain URLs with their context information, component used twitter streaming APIs. The process of looking up for corresponding IP address and that follows all redirections of URLs. Redirect URLs and IP chains to the tweets are push and store in queue which are appends by crawling threads.

Conditional redirections to evade crawlers are used when our crawlers cannot find malicious landing URLs .Hence our system works independently of crawler evasion, instead of rely on features of landing URLs.

2. FEATURE EXTRACTION :

There are three subcomponents in data collection and they are as follows:

a. Grouping of identical domains.
b. Find entry point.
c. Extract feature vectors.

The process of determining the sufficient number of tweets collected has monitor by this component. Instead of individual window here we used tweet window. The w tweet is popped from queue when the tweets are more than w tweet. Firstly it check that the tweets in the queue shares the same IP address. It replace the domain names with list of domains with which is group for all the several URLs that share at least one IP address. It enables the process of detection of suspicious URLs are blacklisted which use several domain names. Then the entry point for URLs is find of every tweet. The measurement of frequency is done with every URLs that appears in tweets.

Then the most frequent URLs is showed among URLS that redirect chain in w tweet. Therefore the URLs is entry point for the redirect chain. The URLs is selected as nearest to beginning of the chain as the entry point if URLs has highest frequency. Domains which are whitelisted are not grouped with other domains and they are not called as entry point.

3. Training :

There are two subcomponents in data collection and they are as follows:

a. Retrieval of accounts
b. Training of classifier.

Offline supervised algorithm is used. The feature vector of classification is relatively new than the feature vector of training.

Twitter account status are used to label the training vectors .URLs from active account are considered benign whereas URLs from non-active accounts are considered malicious. Classifier is updated occasionally by using the labelled training vectors.

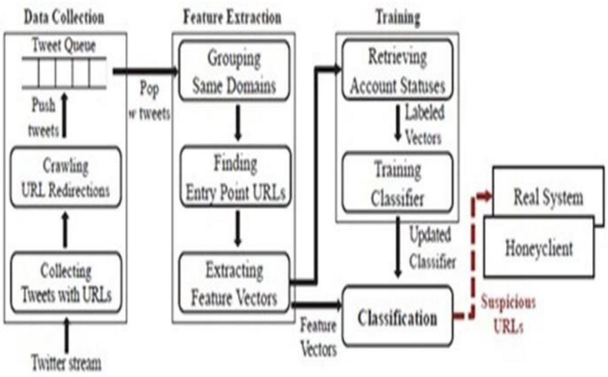

Fig: Architecture of System.

4. Classification :

The main process is to classify the suspicious URLs among the number of URLs using feature vectors. Then corresponding URLs are flagged and their tweets are context are spam or suspicious from the classifier show number of feature vectors. Thus suspicious URLs detected are send to the security experts.

VI. CONCLUSION

The process of distinguishing the normal messages and abnormal message that is the suspicious message have been done with the previous

Spam and malicious URLs detecting system. The protection against the conditional redirections cannot be done with the conventional system.

Here the main aim is to focus on the correlations of multiple redirected chains that usually share the same URLs redirect servers.

VII. REFERENCES:

[1] S. Lee and J. Kim, "WarningBird: Detecting suspicious URLs in Twitter stream," in Proc. NDSS, 2013.

[2] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social Network or a news media?" in Proc. WWW, 2010.

[3] D. Antoniades, I. Polakis, G. Kontaxis, E. Athanasopoulos, S. Ioannidis, E. P. Markatos, and T. Karagiannis, "we.b: The web of short URLs," in Proc. WWW, 2011.

[4] D. K. McGrath and M. Gupta, "Behind phishing: An examination of phisher modi operandi," in Proc. USENIX LEET, 2008.

[5] D. Canali, M. Cova, G. Vigna, and C. Kruegel, "Prophiler: A fast filter for the large-scale detection of malicious web pages," in Proc. WWW, 2011.

[6] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, "Who is tweeting on Twitter: Human, bot, or cyborg?" in Proc. ACSAC, 2010.