# Threshold Based Mechanism to Detect Malicious URL's in Social Networks

[1]Divya, [2]Dr. Kulvinder Singh, [3]Dr. Sanjeev Dhawan
*[1]Mtech Student, UIET , Kurukshetra University,136119 , Kurukshetra, Haryana ,India*
*[2]Assistant professor, UIET , Kurukshetra University,136119 , Kurukshetra, Haryana ,India*
*[3]Assistant professor, UIET , Kurukshetra University,136119 , Kurukshetra, Haryana ,India*
([1]simplydeevya@gmail.com, [2]kshanda@rediffmail.com, [3]rsdhawan@rediffmail.com)

***ABSTRACT:*** *In the past six years, tremendous growth in the size and popularity of social networking has fundamentally changed the way to use the Internet. As social aspects to the Internet continue to expand in both quantity and scope, security of the users of social networking sites and the data generated by them will ultimately become an unavoidable concern. Social networks includes various kinds of URL's some of them may contains harmful information that are called malicious URL's. In this paper system architecture of malicious URL detection has been presented. Various existing techniques to detect malicious URL also discussed after that a proposed mechanism threshold based scheme to detect malicious URL has been presented*

***Keywords-*** *URL, Spam, Web Crawler, filter and classifiers.*

## I. INTRODUCTION

Social networks are networks which help users to share their personal information including pictures, videos etc. with their friends also find more friends interact with themselves. These features become social network popular in todays. As interaction increases then chances of security violation also increases. In these networks URL's play an important role for users to access contents some of them are called malicious URL. In this paper main focus is on malicious URL.

1.1 Malicious data and URL: Malicious Data, short for Malicious Advertising Software (Ad-ware), is a sequence of instructions that perform malicious activity in social networks. The history of malicious programs started with "Computer Virus", a term first introduced by Cohen. It is a piece of code that replicates by attaching itself to the other executables in the system. Today, malicious code includes viruses, worms, trojans, root kits, backdoors, bots, spyware, adware, scare ware and any other program that exhibits malicious behavior. Adwares are a fast growing threat to the modern computing networks. The production of Adware has become a multi-billion dollar industry. The growth of internet, the advent of social networks and rapid multiplication of botnets has caused exponential increase in the amount of Adware. In 2010, there was a large increase in the amount of Adware spread through spam emails sent from machines that were part of botnets. McAfee Labs have reported that, there were 6 million new infections in each month [1].

Two critical elements that affect mobile use are privacy and positive user experience. The market for mobile applications is based on trust. Mobile advertising is questionable practice, such as applications that use deceptive practices adware, a negative impact on the perception of the end user privacy and user experience. Doing things like capture personal information such as email addresses, device ID, IMEI, etc. without properly notifying users and change phone settings and desktop without consent, it is annoying and unacceptable for mobile users. While most mobile ads are not malicious, however, they are undesirable for most people [2].

As malware on the Internet spreads and becomes more sophisticated, anti-malware techniques need to be improved in order to be able to identify new threats in an efficient and automatic way. Malicious web content has become one of the most effective mechanisms for cyber criminals to distribute malicious code [3].

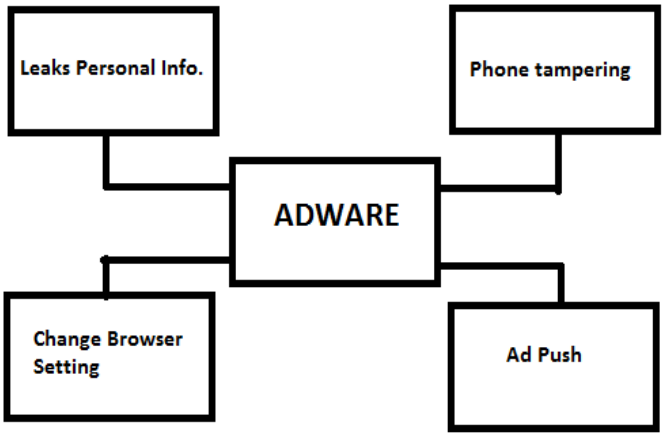

Fig. 1: Dangers of Malicious URLs and Ad-wares [3]

1.2 System Architecture: In this section we present suspicious URL and malicious add attack working steps involved in malicious add attack [4].

Malicious Ads Attack: these kinds of attacks are caused by when user click on a link presented in website then he or she will receive various kinds of attacks.

Suspicious URL: Advertisements on a website are in different forms like images, banners and may be in the form of hyperlink. When a user clicks on advertisement link shown on website then that link takes user demand to the resultant website. To detect malicious ads, it is necessary to collect URL related to advertising; these URL's are called suspicious URL [5].

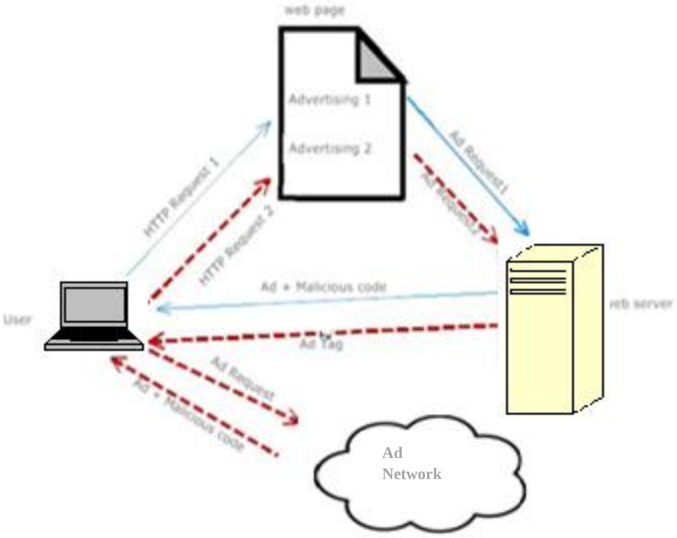

Fig. 2: Steps Involved in Malicious Advertisement Attack [5]

Above figure show steps involved in malicious advertisement attack. In first step user request for website to server then server sends requesting website to the user. When user click on website then a request message goes to the advertisement server and then server sends reply message with malicious code. This code may cause user system by different attacks. Here in this figure red line show malicious ad attack.

## II. RELATED WORK

To detect malicious add attack various researchers proposed various techniques and methods. Some of them proposed tools like Nutch Modsecurity etc. Here in this section we present literature survey of these techniques.

Li et al.[6]developed a general and fully distributed detection framework in 2012 which could be executed by every legitimate node so as to identify its malicious neighbors.

Similarly in 2011 Perez et al.[7] proposed a real time evaluation of the Twitter profiles. The evaluation performs in two steps. First, the suspicious profiles are detected based on graph, messages and behavioral features. Second, the identified suspicious profiles are scrutinized to detect malicious URLs in the messages.

J. ma et al. [8] proposed that in 2009, in which they classify lexical URL's by searching special keywords like tokens or notations attached with URL's Stringhini et al.[9] analyzed how spam's are extended means how they operated in social networks.

Wang [10] proposed graph based scheme to filter normal or malicious URL's.

Ghosh et al.[11] analyzed the link farming. Link farming is a method used for synthetically growing the significance of a profile in a given network.

In 2014 Wang [12] proposed activity-based detection scheme in which spam's are detected by evaluating low quality information of spam's.

In year 2004 Kołaczek[13] proposed mechanism to detect events which may be helpful in the detection of spam's. Author just discussed some issues and measures but not provide any technique that deals with the issues occurred in social network.

Robertson *et al.* [14] in 2010 provides an approach by which malwares are detected by using network information. This network information helps in decision making.

C. H. Gao *et al.* [15] in year 2012 proposed online spam filtering mechanism in which compare tweets with previously stored information and then find out whether it is malicious or normal and then give this information to the classifiers for classification.

## III. PROPOSED WORK

To manage with malicious status, several social network spams's discovering schemes have been proposed earlier and that can be classified into user account feature-based, friends feature-based, and message feature-based schemes. Extracting these relation features from a graph, however, requires an important amount of time and resources as a social network graph is incredible in size. However, spammers and suspicious users can easily change the shape of their messages. To contrast malicious URL we wish to propose threshold based mechanism in which:

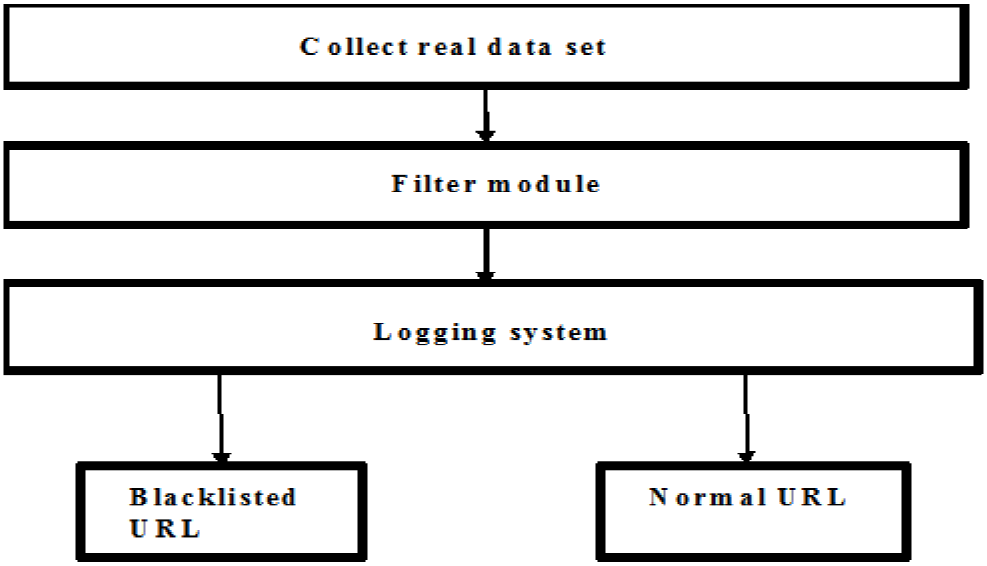

Fig. 3: Proposed Working Principle

1. Real data set collection: In this phase real data set is collected from various external sources that contain URL information.
2. Filter module: Filter module is used to filter suspicious URL by using different parameters like text, images and external links etc.
3. Logging system: In logging system complete log of process is stored also logging system stores blacklisted URL and normal URL.

## IV. CONCLUSION

Security in social network is a difficult task. A variety of attacks are occurred when a user click on malicious URL. Malicious URL could be an advertisement presented on web. In this paper suspicious URL, malicious ad on attack are discussed. Various Causes of malicious URL are discussed. After that System architecture of malicious ad attack is also presented. Next present proposed threshold based mechanism work to contrast malicious URL's from social networks.

## REFERENCES

[1] Jitendra Apte and Marina Lima Roesler, Interactive *Multimedia Advertising and Electronic Commerce on a Hypertext Network*, U.S. Patent No. 7,225,142. 29 May 2007.

[2] Ravula and Ravindar Reddy, Classification *of Malware using Reverse Engineering and Data Mining Techniques*, M.S. Dissertation, University of Akron, CS Dept., 2011.

[3] Justin Ma, Lawrence K. Saul, Stefan Savage and Geoffrey M. Voelker, Beyond Blacklists: Learning to Detect Malicious Web Sites from Suspicious URLs, *in Proceedings of 15th ACM international conference on Knowledge discovery and data mining,* 2009.

[4] Shanshan HONG, Online Advertising Alliance Based Advertisers Needs Analysis, *Proceedings of Conference on Web Based Business Management,* 2011.

[5] Zhou Li, Kehuan Zhang and Yinglian Xie, Knowing Your Enemy: Understanding and Detecting Malicious Web Advertising, *ACM conference on computer and communications security, 2012*.

[6] Yongkun Li and John C.S. Lui, On Detecting Malicious Behaviors in Interactive Networks: Algorithms and Analysis, *in Proceedings of 4th International Conference*, 2012, pp.1-10.

[7] C. Perez, M. Lemercier, B. Birregah and A. Corpel, SPOT 1.0: Scoring Suspicious Profiles on Twitter, in IEEE International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2011, pp. 377–381.

[8] Justin Ma, Lawrence K. Saul, Stefan Savage and Geoffrey M. Voelker, Identifying Suspicious URLs: An Application of Large Scale Online Learning, *in Proc. of the International Conference on Machine Learning(ICML),* 2009.

[9] G.Stringhini, C. Kruegel and G. Vigna, Detecting Spammers on Social Networks, *in proceedings of the 26th Annual Computer Security Applications Conference (ACSAC)*, 2010, pp.1-9.

[10] A. H. Wang, Detecting Spam Bots in Online Social Networking Sites: A Machine Learning Approach, *in DBSec'10: Proceedings of the 24th annual IFIP WG 11.3 working conference on Data and applications security and privacy*, Berlin, Heidelberg: Springer-Verlag, Jun. 2010, pp. 335–342.

[11] S. Ghosh, B. Viswanath, F. Kooti, N. Sharma, G. Korlam, F. Benevenuto, N. Ganguly and K. P. Gummadi, Understanding and Combating Link Farming in the Twitter Social Network, *WWW '12: Proceedings of the 21st international conference on World Wide Web,* 2012.

[12] De Wang, Analysis and Detection of Low Quality Informationin Social Networks, *in proceedings of 30th IEEE International Conference,2014,* pp. 350-354.

[13] Grzegorz Kołaczek, An Approach to Identity Theft Detection Using Social Network Analysis, *in Proceedings of 1st Asian Conference* , 2009, pp. 78-81.

[14] Michael Robertson, Yin Pan and Bo Yuan, A Social Approach to Security: Using Social Networks to Help Detect Malicious Web Content, 2010, pp. 436-441.

[15] H. Gao, Y. Chen, K. Lee, D. Palsetia and A. Choudhary, Towards Online Spam Filtering in Social Networks, *in Proc. NDSS*, 2012.