Twitter User Is Crawled By Shortening Services Using Public Analytics and Metadata

Deepika K

B.E. Computer Science and Engineering, Anand Institute of Higher Technology, Chennai, India
deepikakesavan13@gmail.com

Indhumathi

B.E. ComputerScience and Engineering, Anand Institute of Higher Technology, Chennai, India
indhubala9196@gmail.com

Bhuvaneshwari A

B.E. ComputerScience and Engineering, Anand Institute of Higher Technology, Chennai, India
bhuja43@gmail.com

Maheswari M

B.E. ComputerScience and Engineering, Anand Institute of Higher Technology, Chennai, India
mskkaalam@gmail.com

***Abstract*— "First, our focus on security is on the infrastructure itself. So its all about how you protect the network, the device and the application that is riding on the server". Twitter is a distinguished online social network service. It is widely used for exchanging information among friends. The characters used to send the URL in twitter is limited(i.e.,140 characters only).If it is above the limited characters, we use the URL Shortening services that provide Public Click Analytics and Twitter Metadata .Hence, Public Click Analytics is provided in an aggregated form to preserve the privacy of individual users. With the shortened URL the user information is steal by the inference attack .But it compromise the twitter user privacy with high accuracy.Our approach does not need any complicated techniques or assumptions such as script injection, phishing,Malware instruction or DNS monitoring all we need is public available information**

Keywords - NetworkSecurity,Publicclickanalytics,ShortenedUrl,Inference attack

I. INTRODUCTION

Network security is widely used in online social websites. It plays an important role for securing the privacy in the online social websites. Network security consists of policies and practices adopted to prevent and monitor unauthorized. Access, misuse, modification, or denial of a computer and network-accessibility resources. It covers a variety of computers networks, both public and private, that are used in everyday jobs; conducting transaction and communications among business, government agencies and individuals.Click analytics a special type of web analytics. It is focuses on on-site analytics. An editor of a web site uses click analytics to determine the performance of his or her particular site, with regards to where the users of the site are clicked.Metadata is data that provides information about other data . The purpose of metadata is to help users find relevant information and discover resources. Unlike the conventional history is stealed by inference attacks techniques. Our attacks only demands publicly available information provided by twitter and URL shortening services. It introduce (i) an attack to know who click on the URL updated by the target users and (ii) the attack to know which URL are clicked on by the target users.

II. RELATED WORK

A. *Background*

Implement inferences attack on public click analytic and twitter metadata.

B. *Existing Work*

In this system the twitter user's information has been attacked. In the first step we are going to convert long URL into very precise URL using the method called shortening services.The reason for converting the URL is that the characters should nt exceed more than 140 characters.In the shortening service we are going to monitor about twitter user's information for that we are using two methods (i) public click analytics.(ii) twitter metadata.

By using the shortening service it is easy for the user to use URL service frequently. And by using this the main advantage is that every user using twitter their information will be saved it cant be stealed.

C. *Problem and shortcomings*

- The periodic monitoring and matching have a limitation
- To share long URLs via tweets having length restriction.
- Most of them combine public information from
- several different data sets to infer hidden information.
- Inefficient for user privacy prevent.

III. LITERATURE REVIEW

A. *Existing work*

The existing system has implemented the following,Twitter's user-city level location based purely on the content of the user tweets, even in the absence of the geospatial cues.histories Novel timing attack method to sniff users' browsing without executing any scripts Server-based collaborative filtering systems have been very successful in e-commerce and in direct recommendation applications.The attacks allow any Web site to determine whether or not each visitor to the site has recently visited some other site (or set of sites) on the Web.The detrimental effects of browser cache/history sniffing in that neutralizes the context of phishing attacks, and detail an approach the threat by means of URL personalization.

B. *Compartive Review*

Evaluate a probabilistic framework for estimating Twitter user's city-level location based purely on the content of the user's tweets, even in the absence of any other geospatial cues[2].The existing Web timing attack methods are heavily dependent on executing client-side scripts to measure the time. However, many techniques have been proposed to block the executions of suspicious scripts recently.[5]. The detrimental effects of browser cache/history sniffing in the context of phishing attacks, and detail an approach that neutralizes the threat by means of URL personalization; we report on an implementation performing such personalization on the fly, and analyze the costs of and security properties of our proposed solution Equations.[3]The attacks allow a malicious Web site to determine whether or not the user has recently visited some other, unrelated Web page.[1]Server-based collaborative filtering systems have been very successful in e-commerce and in direct recommendation applications.[4]

IV. OVERALL SYSTEM ARCHITECTURE

Twitter is mainly used for sharing the information among people. Twitter users who want to share the long URL via tweets having length restriction. Twitter allows users to post up to 140 character tweets containing only texts. Twitter users demand URL shortening services further there reducing it. The URL shortening services provide shortened URLs' Public click analytics consisting of the number of clicks,countries, browsers and referrers of visitors which is stored in the metadata of twitter user. The browsers history of twitter user is stealed by Inference attack techniques. But it is compromise the twitter users' privacy with high accuracy.

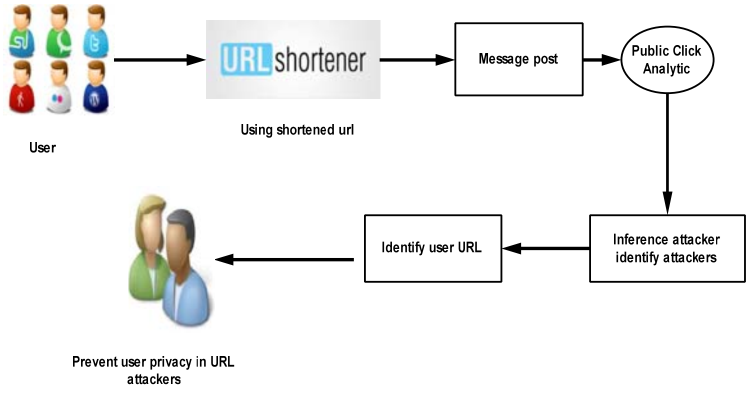

Figure (a).*System Architecture*

V. PROPOSED SYSTEM

In the proposed system it tells about now the twitter user private information has been attacked in different ways.By using the public click analytics and twitter metadata the privacy information has been stealed.

They are two types of attacks

(i) Inference attacks.

(ii) Target protection method.These are the attacks inwhich the user private information has been stealed by the attacker.

B. Modular Description

(i).URL Shortening Services.

The twitter user create own identification based on URL through web server. This web server verify to formation of shortened URL i.e., google-goo.com etc. Our attacks rely on the combination of publicly available information: click analytics from URL shortening services and metadata from Twitter. Two different attack methods

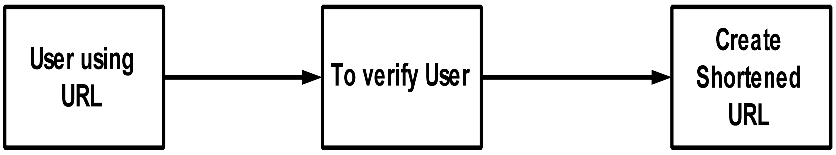

Figure 1: URL Shortening Services.

(ii)User Posting Message using shorter URL

User posting message through using shorten URL based on public click analytic. Some URL shortening services also provide click analytics about each shortened URL. Whenever a user clicks on a shortened URL, information about the user is recorded in the corresponding click analytics. The click analytics is usually made public and anyone can access it

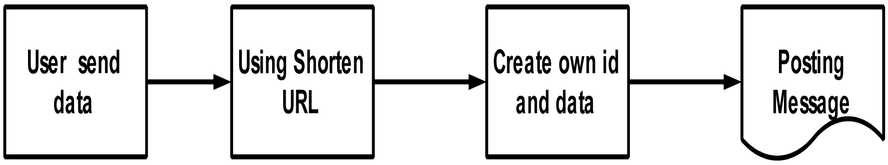

Figure 2: User Posting Message using shorter URL

(iii)Browsing history of public click analytic

To perform the first attack, we find a number of Twitter users who frequently distribute shortened URLs, and investigate the click analytics of the distributed shortened URLs and the metadata of the followers of the Twitter users

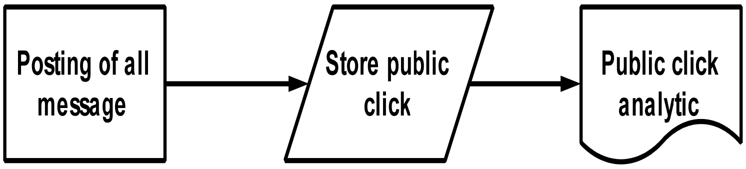

Figure 3: Browsing history of public click analytic

(iv).Inference attack to identify URL user

System attack methods for inferring whether a specific user clicked on certain shortened URLs on Twitter. Our attacks rely on the combination of publicly available information: click analytics from URL shortening services and metadata from Twitter. Two different attack methods: (i) an attack to know who click on the URLs updated by target users and (ii) an attack to know which URLs are clicked on by target users.

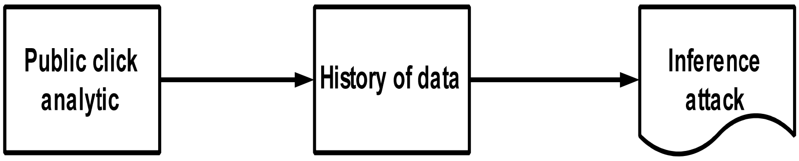

Figure4 : Inference attack to identify URL user

(v)To prevent the user privacy.

The attack system chooses a target Twitter user and extracts his or her information from Twitter. The system monitors the click analytics of all shortened URLs posted by the followings of the target user. The system compares the information about the visitor with the known information the target user. If both pieces of information match, it infers that the target user clicks on the shortened URL. Finally prevent user privacy based on inferences attack.

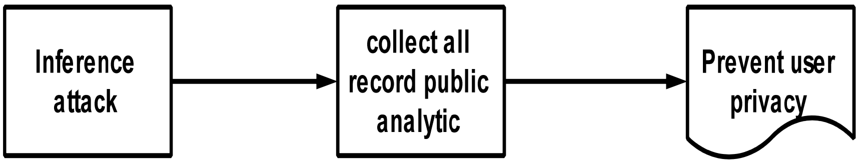

Figure 5: To prevent the user privacy.

VI. Conclusion

Inference attack collect the timeline and the favorites of the user A and check whether a tweet containing the shortened URL is exists. Twitter users include URLs in their tweets and favorite tweets with URLs only attacks to infer which shortened URLs clicked on by a target user. All the information needed in our attacks is public Information: the click analytics of URL shortening services and Twitter Meta data. To evaluate the attacks, crawled and monitored the click analytics of URL shortening services and Twitter data. check whether a target user includes the URL inferred as visited in his (re)tweets or favorites it in the near.

VII. Future Work

The future to validate the correctness of our inference. To clarify, suppose that our system infers that a Twitter user A visits a shortened URL . when they previously visit the URLs.

VIII. Applications

Efficient for specific user click on curtained shortened URL.

Easily to prevent user privacy accuracy.

Information matching process is more accuracy

References

[1] E. W. Felten and M. A. Schneider, “Timing attacks on web privacy,” in Proc. 7th ACM Conf. Comput. Comm. Secur. (CCS), 2000, pp. 25–32.

[2] M. Jakobsson and S. Stamm, “Invasive browser sniffing and countermeasures,” in Proc. 15th Int. World Wide Web Conf., 2006, pp. 523–532.

[3] S. Krishnan and F. Monrose, “Dns prefetching and its privacy implications: When good things go bad,” in Proc. 3rd USENIX Workshop Large-scale Exploits Emergent Threats, 2010, pp. 10–10.

[4] J. A. Calandrino, A. Kilzer, A. Narayanan, E. W. Felten, and V. Shmatikov, ““You might also like:” Privacy risks of collaborative filtering,” in Proc. IEEE Symp. Secur. Privacy, 2011, pp. 231–246.

[5] Z. Cheng, J. Caverlee, and K. Lee, “You are where you tweet: A content-based approach to geo-locating twitter users,” in Proc. 19th ACM Int. Conf. Inform. Knowl. Manage., 2010, pp. 759–768.

[6] A.Clover.(2002). Css visited pages disclosure.[Online]. Availability :http://sclists.org/bugtrap/2002/Feb271

[7] C.Dwork,”Differential privacy “,in Proc.33rd Int. Colloqium Automata,Languages Programm.,Springer Berlin Heidelberg,2006.pp1-12

[8] C.Jackson, A.Bortz, D.Boneh and J.C. Mitchell ,”Protecting browser state from web privacy attacks “, in Proc.15th Int. World Wide Web Conf.,2006, pp.737-744

[9] A.Chaabane, G. Acs, “You are what you like ! information leakage through users interests”, Proc 19 th Network and distributed system security Symp.,2012

[10] A.Narayanan nd v.Shamatikov,”Robust Deanonymization of Large Parse Dataset”, Proc.IEEE. Symp. Secur. Privacy,2008,pp.111-125.