Privacy preserving sequential padding in web-based applications

Divyabharathi.P, Sathiyavathi.R

Abstract—These Encrypted traffic of many popular Web applications may actually contain disclose highly sensitive data, and lead to serious breaches of user privacy. Specifically, when searching for unique patterns exhibited in packets’ sizes and/or timing, a hacker can potentially find an application’s internal state transitions and the users’ inputs. A solution for preventing this type of side channel attack is padding packets such that each packet size will no longer map to the unique input. One extreme cases to pad all packets to the identical size, namely, maximizing. We design many heuristic algorithms for solving these PPTP problems in polynomial time with acceptable overhead.

Index Terms— l-Diversity, k-In distinguishability, PPTP, PPDP, Side-Channel Leak, Traffic Padding, Web Application.

1 INTRODUCTION

Network security consists of the provisions, policies adopted by a network administrator for preventing and monitoring illegal access, misuse, modification, or denial of computer network and network accessible resources. Network security involves in authorization of accessing the data in a network, which is controlled by network administrator. And it covers a variety of computer networks, such as public and private, which are used in everyday job conducting transactions and communications in between businesses, government agencies and individuals. These networks can be private, such as within a single company, and others might be open to public accesses. Network security is involved in enterprises, organizations, and other type of institutions. Web-based applications are gaining popularity and they require less client-side resources, they are easier to deliver and maintain. Other hand, Web applications may also present new security and privacy challenges, because the un-trusted Internet has an integral component of such applications for carrying the continuous interaction between users and servers. Many high profile Web applications might have sensitive user inputs, they may leaking the input from encrypted traffic due to side-channel attacks.

In our proposed system, we first present a model of the PPTP issue based on the mapping to PPDP, which formally characterizes the interaction between users and Web applications, the observation made by eavesdroppers, the privacy requirement, and the overhead of padding we present a formal PPTP model encompassing the privacy requirements, padding costs, and padding methods. And then formulate PPTP problems under different scenarios, analyze their complexity, and design efficient heuristic algorithms. Next, we design several heuristic algorithms for solving the PPTP problems in polynomial time with acceptable overhead. Finally, we demonstrate the effectiveness and efficiency of our algorithms by both analytical and experimental evaluation. The remainder of the paper is prepared as follows. In Section II, the related work which is surveyed. In Section III, solution methodologies will be described. In Section IV, performance analysis and finally, we conclude the paper in Section V.

2 RELATED WORK

Kehuan Zhang proposed a suite of new techniques for automatic detection, quantification of side-channel leaks in web applications. This approach, called Side buster, it automatically analyses the application’s source code to detect its side channels then performs a rerun test to assess the amount of...
information disclosed through channels. Side busting had been designed to work on event-driven applications and can effectively handle the AJAX GUI widgets which we are using in most web applications.[1] Michael Backes developed a framework for derivation of formal guarantees which is against to the traffic side-channels. And he presented a model which captures important characteristics of web traffic, defined measures of security, and also provided an assembly kit for countermeasures and their security guarantees, and showed that security guarantees are preserved on lower levels of the protocol stack.[2] George Dean Bissias at el proposed a straight forward traffic analysis attack against encrypted HTTP streams that is surprisingly effective in identifying the source of the traffic. An attacker twitches by creating a profile of the statistical characteristics of web request including displaced packets of encrypted streams and information disclosed through channels. Side buster had been designed to work on event-driven applications and can effectively handle the AJAX GUI widgets which we are using in most web applications.[1] Michael Backes developed a framework for derivation of formal guarantees which is against to the traffic side-channels. And he presented a model which captures important characteristics of web traffic, defined measures of security, and also provided an assembly kit for countermeasures and their security guarantees, and showed that security guarantees are preserved on lower levels of the protocol stack.[2] George Dean Bissias at el proposed a straight forward traffic analysis attack against encrypted HTTP streams that is surprisingly effective in identifying the source of the traffic. An attacker twitches by creating a profile of the statistical characteristics of web request including displaced packets of encrypted streams and information disclosed through channels.

Adversary will create the account for users to send files over internet or else user must register their details such as username, and Email id, password, into the server. Then server will create the account for users to send files over internet with more security.

### 3 METHODOLOGIES

#### 3.1 USER INTERFACE DESIGN

To connect with server user should give their username and password then, they can able to connect to the server. If in the case user already exits directly they can login into the server or else user must register their details such as username, and Email id, password, into the server. Then server will create the account for users to send files over internet with more security.

#### 3.2 KEY STROKE

In this module, the data is given by customer requests goes to server. Keystroke logging, often referred to as key logging or keyboard capturing, is the action of recording (or logging) the keys struck on a keyboard, characteristically in a covert manner so that the person using the keyboard is unaware that their actions are being monitored. Each and every packet exhibit same size and different type of patterns.

#### 3.3 MAPPING PPTP TO PPDP

In this model alone may lead to excessive data distortion or insufficient protection Privacy-preserving data publishing (PPDP) provides methods and tools for publishing useful information while preserving data privacy. The k-anonymity model divides the table into anonymized groups and then generalizes the quasi-identifier such that at least k individuals in the table will share the same generalized, and hence a linking attack using this quasi-identifier will fail.

#### 3.4 PACKET ORDERING

In this model received packets are going to sort using

---

**Mail:** sathyakrish2723@gmail.com


**ISSN:** 2229-5518

**Author:** Divyabharathi.P is currently pursuing masters degree program in Information Technology in Sathyabama University, India, PH-9966493368. E-mail: pothanidilu@gmail.com

**Address:** Sathyavathi.R is currently working as a assistant professor in department of Information Technology in Sathyabama University, India. E-mail:sathyakrish2723@gmail.com
SVSD diversity algorithm. It first sorts the actions in non-increasing order of their weight values, then among the actions with same weight, and sorts them in a predefined order based on their flow-vectors.

4 TECHNIQUE USED

In this proposed system we are mainly using SVSD Simple Algorithm. The main intention of presenting the SVSD Simple algorithm is to show that, when applying k-in distinguishability to PPTP problems, an algorithm may sometimes be devised in a very forthright way, and yet attain a dramatic reduction in costs when compared to existing approaches. Basically, the SVSD Simple algorithm attempts to minimize the cardinality of padding groups in the SVSD case.

SVSD Algorithm:
Input: a vector-action-weight set VAW, the privacy property l;
Output: the partition PVAW of VAW;
Method:
1. Let PVAW = ;
2. Let S be the sequence of VAW in a non-increasing order of its W;
3. If(pr(S[1]; S) > 1 l )
4. Return;
5. Sort elements in S with same weight value in non-increasing order of its V;
6. While (S 6= ;)
7. Let P_ = fS[i] : i 2 [1; _]g, P_+ = fS[i] : i 2 [+ 1; jSj]g;
8. Let _ 2 [l; jSj] be the smallest value such that:
pr(S[1]; P_ ) _ 1 l and
(pr(S[+ 1]; P_+) _ 1 l or P_+ _ ;)
9. Create partition P_ on PVAW;
10. S = P_+;
11. Return PVAW;

5 PERFORMANCE EVALUATION

In this section sorted packets are received by server. Here we have presented algorithms to determine the amount of padding for each and every flow given in the vector action set. First, gather information about possible action-sequences and corresponding vector-sequences in the application. Second, feed the vector action sets into our algorithms to calculate the desired amount of padding. Third, implement the padding according to the calculated sizes.

6 CONCLUSION

We have proposed a formal model for quantifying the amount of privacy protection provided by traffic padding solutions. Our algorithms have confirmed the performance of our solutions to be superior to existing ones in terms of communication and computation overhead.

FUTURE ENHANCEMENT

Traffic morphing is proposed to alleviate the threats by traffic examining on packet sizes and sequences over network. Although proposed system morphs classes of traffic to be indistinguishable, traffic morphing pads or splits packets on the fly which may degrade application’s performance. Quasi-identifiers are pieces of data that are not of themselves matchless identifiers, but they are sufficiently well correlated with an entity that they could be combined with other quasi-identifiers to create a unique identifier. In proposed warn about potential privacy breaches being enabled by publication of large volumes of government and business data containing quasi-identifiers. As an example, neither gender nor postal codes uniquely identify an individual, but the combination of all is sufficient to identify 87% of individuals a country. A search engine will no longer provide auto-suggestion feature once the query string beats a certain length.

AKNOELEDGEMENT

This research was supported by Sathyabama University, Chennai.

REFERENCES

1. TITLE: Sidebuster: Automated Detection and Quantification of Side-Channel Leaks in Web Application Development.
AUTHOR: Kehuan Zhang, Zhou Li, Rui Wang, XiaoFeng Wang
2. TITLE: Preventing Side-Channel Leaks in Web Traffic: A Formal Approach
AUTHOR: Michael Backes
3. TITLE: Privacy Vulnerabilities in Encrypted HTTP Streams
AUTHOR: George Dean Bissias, Marc Liberatore, David Jensen
YEAR: 2005
4. TITLE: Privacy-Preserving Payload-Based Correlation for Accurate Malicious Traffic Detection
AUTHOR: Janak J. Parekh, Salvatore J. Stolfo, Ke Wang
YEAR: 2007
5. TITLE: Traffic Morphing: An Efficient Defense Against Statistical Traffic Analysis
AUTHOR: Charles V. Wright, Fabian Monrose
6. TITLE: Background Knowledge-Resistant Traffic Padding For Preserving User Privacy in Web-Based Applications
AUTHOR: Wen Ming Liu, Lingyu Wang, Kui Reny, Mourad Debbabi
7. TITLE: Models of Privacy Preserving Traffic Tunneling
AUTHOR: David M. Nicol, Nabil Sear
8. TITLE: KIPDA: k-Indistinguishable Privacy-preserving Data Aggregation in Wireless Sensor Networks
AUTHOR: Michael M. Groat, Wenbo He, Stephanie Forrest
   AUTHOR: Zhiguo Wan, KuiRen, and Ming Gu
10. TITLE: Anomalous Payload-based Worm Detection and Signature Generation