WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Bi-Directional Bursting Defence against Website Fingerprinting Attack

Selvakarthikeyan S

Dr MGR Educational and Research Institute, Maduravoyal, Chennai- 600095, India

ABSTRACT

The use of network traffic analysis in diverse applications to protect or endanger people, information, and systems is becoming more common. Website fingerprinting is a type of passive traffic analysis attack that puts the privacy of web users at risk. It's a collection of approaches for deducing patterns from a stream of network packets created as a user navigates between websites. To safeguard their privacy, Internet users (such as online activists or journalists) may seek to conceal their identity and online activity. This is usually accomplished through the use of an anonymity network. These anonymity networks, such as Tor (The Onion Router), provide layers of data encryption, making traffic analysis techniques difficult to use. Despite the fact that different defenses have been offered to prevent this passive attack, fresh attacks have proven the ineffectiveness and/or impracticality of such defenses. We present a novel defense technique to combat website fingerprinting attacks in this paper. The suggested defense obfuscates original website traffic patterns by deforming packet sequences and destroying traffic flow dependency features used by attackers to identify websites using double sampling and mathematical optimization approaches. We compare our defense to state-of-the-art studies and demonstrate its efficacy with minimum overhead and zero-delay transmission to real-world traffic.

Introduction

Many research over the last decade have focused on user privacy on the internet. Security and privacy technologies are increasingly being utilized to secure users' identities as the number of applications and ways to access information grows. SSH, SSL/TLS, VPN, and IP Sec are examples of these technologies. An attacker's ability to identify the web sites visited by a user is one aspect of web privacy. To secure the content accessed, private browsing and proxy tunneling are frequently employed. Network identity, on the other hand, may not be sufficiently protected. A user (for example, an activist or journalist) may seek to remain anonymous or circumvent active online restrictions on their freedom. Passive traffic analysis of network packets when a user sees a website, according to recent studies, can thwart these privacy defenses. The Website Fingerprinting assault, which is most commonly utilized in attack situations by a passive adversary who is considered to have access to the victim's network, is known as this. An adversary seeks to identify a client's web browsing habits by passively listening to network traffic between the client and a server in an attack scenario. The client's website is identified or predicted using traffic analysis, which is done using several statistical methods. Clients frequently use proxies or low-latency anonymity network services such as Tor to remove deterministic identification features such as destination IP and webpage content (The Onion Router). These services encrypt and disguise network packets destined for a certain destination. Attackers use network traffic from numerous websites to learn the parameters of statistical models using machine learning techniques. Such models can be used to categorize network traffic observations.

Various defenses have been developed in the literature to combat the website fingerprinting attack. The battle between attackers and defenders has been changing over time. On the one hand, the attacker collects encrypted packets sent between the client and server, extracts patterns and features, and analyzes data using machine learning techniques in an attempt to determine the destination website a user is attempting to access. Defenders, on the other hand, have been inventing numerous methods to prevent such attacks by concealing and morphing network packets meant for a certain website (such as Tor). Existing defenses in the literature attempt to counter such assaults by morphing the (source) website distribution to look to come from a different (target) website distribution in order to confuse the machine learning classifier. The goal of such defenses is to change packet length, timing, and successive sequences of packets in a given way (i.e., client to server and vice versa). We offer bi-morphing in this research, a novel website fingerprinting defense that thwarts fingerprinting attacks by taking into account bi-directional dependence between consecutive sequences of packets traveling in different directions. The suggested defense system uses bi-directional statistical sampling and optimization techniques to conceal website patterns and achieve low bandwidth overhead and zero-delay delivery to actual traffic. This is the first study that we are aware of that uses a size and time (double) concurrent sampling strategy.

1.1. Proposed system

BIMORPHING is a new defense against the passive traffic fingerprinting attack. Designing an effective defense that prevents attackers from extracting knowledge from encrypted traffic while minimizing bandwidth and time overhead is one of the issues that any defense system faces. BIMORPHING

uses optimum size and time sampling, as well as bi-directional dependence, to achieve the smallest feasible bandwidth overhead. The defense accomplishes a zero-delay packet trans-mission by sending extra dummy packets in gaps between real packets that are transmitted immediately. The user's identity is deterministically hidden thanks to a combination of encryption and proxy server. Furthermore, because anonymity networks use several proxies between the user and the destination server, it is more difficult to identify the destination IP. Anonymity networks like Tor, for example, hide information about their users by providing low latency anonymization and pipeline randomization. Techniques for webpage fingerprinting have been proposed. A supervised learning technique is used, in which a collection of features from traffic flow at the user's end are collected. Packet length, direction (i.e., uplink from client to server or downlink from server to client), and time are all factors to consider. The authors aggregate successive packets in addition to employing packet length histograms

1.2.Proposed system advantages

- Use link-state area hierarchy for topology
- Exchange route summaries at area borders
- Use Time-stamps Update numbering & counters

2.System design

2.1 Input design

One of the most crucial phases of the system design is input design. Input design is the process of planning and designing the input received by the system in order to obtain the necessary information from the user while avoiding the information that is not required. The goal of the input design is to guarantee that the input is as accurate as possible while simultaneously being accessible and understandable to the user. The input design is a component of the overall system design that necessitates meticulous consideration. If the data entering the system is wrong, the errors will be amplified by the processing and output. During input design, the following goals are taken into account:

.The objectives considered during input design are:

- Nature of input processing.
- Flexibility and thoroughness of validation rules.
- Handling of properties within the input documents.
- Screen design to ensure accuracy and efficiency of the input relationship with files.
- Careful design of the input also involves attention to error handling, controls, batching and validation procedures.

Input design characteristics can either ensure the system's stability and produce accurate results, or they can result in the creation of incorrect data.

.

2.2 Output design

The most essential and immediate source of information for the user is computer output. Efficient, comprehensible output design should strengthen the system's user relationships and aid decision-making. The physical copy from the printer is a common form of output. The output devices to consider are determined by a variety of parameters, including device compatibility with the system, reaction time requirements, estimated print quality, and the quantity of copies required. Unpredictably, all nodes in the network may leave or fail.

The continuous measurement data is partitioned into time slots, with a source block referring to the quantity of data generated in one time slot on a node. Clearly, the size of the node cache storage determines how many time slots of data can be cached.

A synchronization packet (also known as the timing reference signal) comes before the first active sample on each line and after the last active sample on each line (and before the start of the horizontal blanking region).

The master files, transaction files, and computer applications are all listed in a systems flowchart. Input data is gathered and arranged into groupings of data that are related. After that, the proper input media for processing is chosen. The output devices to consider are determined by a variety of parameters, including device compatibility with the system, reaction time requirements, estimated print quality, and the quantity of copies required. Unpredictably, all nodes in the network may leave or fail.**.**

3.Conclusion

We introduced the BIMORPHING defense to prevent encrypted traffic fingerprinting attacks, which combines size and time sampling with bi-directional dependence, ensures low bandwidth overhead by mathematical optimization, and incurs zero latency for genuine packets transferred between client and server. We examined the defense against passive attacks and compared it to state-of-the-art solutions to demonstrate the efficiency of the suggested methodology. The encouraging results, low bandwidth overhead, and zero latency for genuine packets provide a new perspective for a more realistic website fingerprinting defense

Reference

[1] E. Zheleva and L. Getoor, "To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles," in Proceedings of the 18th international conference on World wide web. ACM, 2009, pp. 531–540.

[2] C. R. Davis, IPSec: Securing VPNs. McGraw-Hill Professional, 2001.

[3] M. Liberatore and B. N. Levine, "Inferring the source of encrypted http connections," in Proceedings of the 13th ACM conference on Computer and communications security. ACM, 2006, pp. 255–263.

[4] D. Herrmann, R. Wendolsky, and H. Federrath, "Website fingerprinting: attacking popular privacy enhancing technologies with the multinomial naïve-bayes classifier," in Proceedings of the 2009 ACM workshop on Cloud computing security. ACM, 2009, pp. 31– 42.

[5] X. Cai, X. C. Zhang, B. Joshi, and R. Johnson, "Touching from a distance: Website fingerprinting attacks and defenses," in Proceedings of the 2012 ACM conference on Computer and communications security. ACM, 2012, pp. 605–616.

[6] T. Wang, X. Cai, R. Nithyanand, R. Johnson, and I. Goldberg, "Effective attacks and provable defenses for website fingerprinting," in Proc. 23th USENIX Security Symposium (USENIX), 2014.

[7] X. Cai, R. Nithyanand, T. Wang, R. Johnson, and I. Goldberg, "A systematic approach to developing and evaluating website fingerprinting defenses," in Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security. ACM, 2014, pp. 227–238.

[8] J. Hayes and G. Danezis, "k-fingerprinting: A robust scalable website fingerprinting technique," in 25th USENIX Security Symposium (USENIX Security 16). Austin, TX: USENIX Association, 2016, pp. 1187–1203. [Online]. Available: https://www.usenix.org/conference/usenixsecurity16/ technical-sessions/presentation/hayes.

[9] K. Al-Naami, S. Chandra, A. Mustafa, L. Khan, Z. Lin, K. Hamlen, and B. Thuraisingham, "Adaptive encrypted traffic fingerprinting with bi-directional dependence," in Proceedings of the 32Nd Annual Conference on Computer Security Applications, ser. ACSAC '16. ACM, 2016, pp. 177–188. [Online]. Available: http://doi.acm.org/10.1145/2991079.2991123.

[10] A. Panchenko, F. Lanze, A. Zinnen, M. Henze, J. Pennekamp, K. Wehrle, and T. Engel, "Website fingerprinting at internet scale," in Proceedings of the 23rd Internet Society (ISOC) Network and Distributed System Security Symposium (NDSS 2016), 2016, to appear.

[11] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The secondgeneration onion router," DTIC Document, Tech. Rep., 2004. [12] K. P. Dyer, S. E. Coull, T. Ristenpart, and T. Shrimpton, "Peek-aboo, i still see you: Why efficient traffic analysis countermeasures fail," in Security and Privacy (SP), 2012 IEEE Symposium on. IEEE, 2012, pp. 332–346.

[13] C. V. Wright, S. E. Coull, and F. Monrose, "Traffic morphing: An efficient defense against statistical traffic analysis," in In Proceedings of the 16th Network and Distributed Security Symposium. IEEE, 2009, pp. 237–250.

[14] M. Juarez, M. Imani, M. Perry, C. Diaz, and M. Wright, Toward an Efficient Website Fingerprinting Defense. Cham: Springer International Publishing, 2016, pp. 27–46.

[15] T. Wang and I. Goldberg, "Walkie-talkie: An efficient defense against passive website fingerprinting attacks," in 26th USENIX Security Symposium (USENIX Security 17). Vancouver, BC: USENIX Association, 2017, pp. 1375– 1390. [Online]. Available: https://www.usenix.org/conference/ usenixsecurity17/technical-sessions/presentation/wang-tao