ijcrr
Vol 04 issue 16
Category: Research
Received on:01/07/12
Revised on:09/07/12
Accepted on:15/07/12

TRAFFIC ANALYSIS ATTACKS ON ANONYMITY NETWORKS

Vetrivendan. R

Department of Computer Science and Engineering, As-Salam College of Engineering and Technology, Thirumangalakudi, Tamilnadu

E-mail of Corresponding Author: vetriascet@gmail.com

ABSTRACT

In this paper, we focus on a particular class of traffic analysis attacks, flow correlation attacks, by which an adversary attempts to analyze the network traffic and correlate the traffic of a flow over an input link with that over an output link. Two classes of correlation methods are considered, namely time-domain methods and frequency-domain methods. Based on our threat model and known strategies in existing mix networks, we perform extensive experiments to analyze the performance of mixes. We find that all but a few batching strategies fail against flow-correlation attacks, allowing the adversary to either identify ingress or egress points of a flow or to reconstruct the path used by the flow. Counter intuitively, some batching strategies are actually detrimental against attacks.

Keywords: Traffic analysis, flow-correlation attack, counter intuitively, detrimental attacks.

INTRODUCTION

As the Internet is increasingly used in all aspects of daily life, the realization has emerged that privacy and confidentiality are important requirements for the success of many applications. It has been shown that, in many situations, encryption alone cannot provide the level of confidentiality required by users, since traffic analysis can easily uncover information about the participants in a distributed application.

Anonymity Network

The anonymity of a system can be passively attacked by an observer in two ways, either through inspection of payload or headers of the exchanged data packets, or, when encryption is used, through traffic analysis.

Traffic analysis is typically countered by the use of intermediary nodes, whose role is to perturb the traffic flow and thus confuse an external observer. Such intermediaries (often called mixes) delay and reroute exchanged messages, reorder them, pad their size, or perform other operations. Chum proposed such a mix network to handle mail traffic.

The original Chum mix network operates on entire mail messages at a time and therefore does not need to pay particular attention to latency added by the mixes. Increasingly, the data exchanged exceed by far the capacity of mixes, for example, in file-sharing applications.

In conjunction with source routing at the sender, this allows for very efficient network-level implementations of mix networks. Mixes are also being used in applications where low latency is relevant, for example, voice-over-IP or video streaming.

Two classes of correlation methods are considered, namely time-domain methods and frequency-domain methods. In the time domain, for example, statistical information about rate distributions is collected and used to identify the traffic dependency.

Similarly, in the frequency domain, we identify traffic similarities by comparing the Fourier spectra of timing data. Our experiments indicate that mixes with many currently used batching strategies are weak against flow-correlation attacks, in the sense that attackers can easily determine the path taken by a protected flow.

Related Work

The idea of anonymous communication in 1981. Since then, researchers have applied the idea to different applications such as message-based e-mail and flow-based low-latency communications, and they have developed new defense techniques as more attacks have been proposed. For anonymous e-mail applications.

Using relay servers, called mixes, which encrypt and reroute messages. An encrypted message is analogous to an onion constructed by a sender, who sends the onion to the first mix:

1. Using its private key, the first mix peels off the first layer, which is encrypted using the public key of the first mix.
2. Inside the first layer is the second mix's address and the rest of the onion, which is encrypted with the second mix's public key.
3. After getting the second mix's address, the first mix forwards the peeled onion to the second mix. This process repeats all the way to the receiver.
4. The core part of the onion is the receiver's address and the real message to be sent to the receiver by the last mix.

Terminology

Enhancement of Mix Network

In this module the sender of a message attaches the receiver address to a packet and encrypts it using the mix's public key. Upon receiving a packet, a mix decodes the packet. Different from an ordinary router, a mix usually will not relay the received packet immediately. Rather, it collects several packets and then sends them out in a batch.

Building Batching Strategies

Building batching strategies are designed to prevent not only simple timing analysis attacks, but also powerful trickle attacks. We will evaluate each of these strategies to send the packets.

Traffic Flow Correlation

Traffic flow-correlation used to the adversary either to correlate senders and receivers directly or to greatly reduce the searching time for such a correlation in a mix network.Objective is to correlate an incoming flow to an output link at a Mix. Also find the Flow-correlation attack.

Detection Metrics

This module used to analyze the detection rate of the traffic attacks. Use detection rate, the probability that the adversary correctly correlates flows into and out of a mix, defined as the measure of success for the attack. We will show that, given a sufficient amount of data, known mix strategies fail.

METHODS

An encrypted message is analogous to an onion constructed by a sender, who sends the onion to the first mix:

1. Using its private key, the first mix peels off the first

Layer, which is encrypted using the public key of the first mix.

2. Inside the first layer is the second mix's address and

the rest of the onion, which is encrypted with the second mix's public key.

3. After getting the second mix's address, the first mix forwards the peeled onion to the second mix. This process repeats all the way to the receiver.

4. The core part of the onion is the receiver's address and the real message to be sent to the receiver by the last mix.

Mix Network

Different from an ordinary router, a mix usually will not relay the received packet immediately. A mix network, such as Onion Routing network or Tor Network, consists of multiple mixes that are interconnected by a network. Building Batching Strategies Any of the batching strategies can be implemented in two ways:

Link-Based Batching With this method, each output link has a separate queue. A newly arrived packet is put into a queue depending on its destination (and hence the link associated with the queue). Once a batch is ready from a particular queue (per the batching strategy), the packets are taken out of the queue and transmitted over the corresponding link.

Mix-Based Batching

In this way, the entire mix has only one queue. The selected batching strategy is applied to this queue. That is, once a batch is ready (per the batching strategy), the packets are taken out the queue and transmitted over links based on the packets' destination. Each of these two methods has its own advantages and disadvantages. The control of link-based batching is distributed inside the mix and hence may have good efficiency. On the other hand, mix-based batching uses only one queue and hence is easier to manage. We consider both methods.

Threat Model : The adversary uses a classical timing analysis attack, which we summarize as follows: The Mix network topology and the general Mix strategies are known to the adversary. This is a natural assumption for many overlay mix networks.

Traffic Flow Correlation

Recall that the adversary's objective is to correlate an incoming flow to an output link at a Mix. We call this flow correlation. This flow-correlation attack is harmful in a variety of situations. For example, in the single-mix, the adversary can discover whom sender (say, S1) is talking to (R1 or R2 in this case) by correlating the output traffic at the Mix to S1's traffic despite cross traffic from S2 or other senders.

Performance Evaluation

Mixes in Networks with Packet Losses Fig.1 & Fig.2

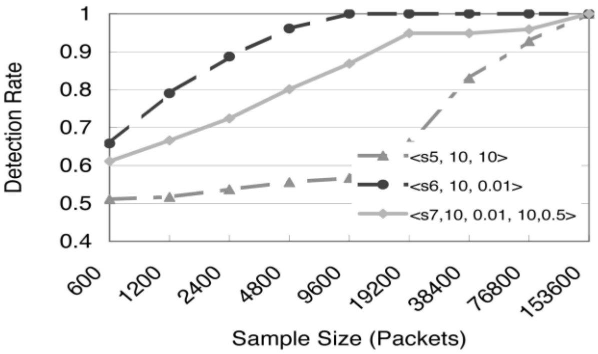

Fig. 1

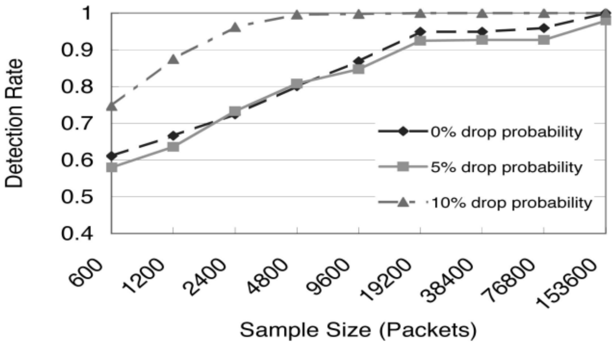

Fig.2

Experiment setup of mix network (six layers) – Fig.4

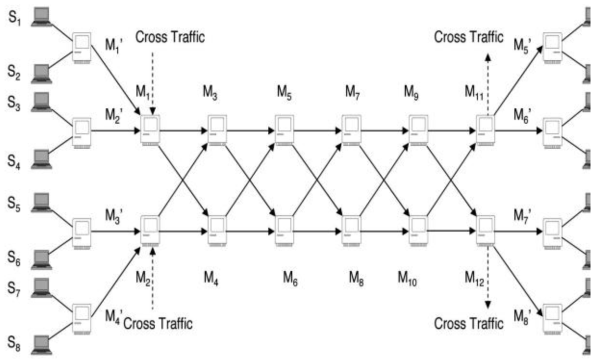

Fig.3

Fig. 2 shows the detection rate for emulated networks using the detection method based on mutual information. We can observe that flow-correlation attacks approach a 100 percent detection rate when the sample size is sufficiently large.

Fig. 3 shows the detection rate when the network is dropping packets. The mix strategy used in this set of experiments is <s7; 10; 0:01; 10; 0:5>, that is, a timed dynamic-pool Mix with pool size 10, batch size 10, batch interval 10 msec, and forwarding probability 0.5. Based on the results, we make the following observations:

1. The detection rate still approaches 100 percent when the sample size is sufficiently large.
2. The results for small drop rates (5 percent or less) appear to be no different than for no packet drops at all. As expected, for larger drop rates (more than 5 percent) the detection rate is higher than for lower drop rates. The reason for this is that a large number of packet drops makes the timing footprint of the TCP dynamics more obvious.

Fig. 4 shows the network setup in this experiment. The center part of the topologies used in experiments is the mix cascade of different number of layers. Each sender on the left side has four flows traversing the mix network.

We arrange paths of traffic flows so that each link in the cascade has some number of traffic flows. To simulate the cross traffic in the mix network, four larger aggregates of flows are added to the mix network. According to the self-similar nature of the network traffic, the high-volume cross traffic is Pareto distributed.

CONCLUSIONS

The analyzed mix networks in terms of their effectiveness in providing anonymity and quality-of-service. Various methods used in mix networks were considered: seven different packet batching strategies and two implementation schemes, namely the link-based batching scheme and mix based batching scheme. Found that mix networks that use traditional batching strategies, regardless of the implementation scheme, are vulnerable under flow-correlation attacks. By using statistical analysis, an adversary can accurately determine the output link used by traffic that comes to an input flow of a mix. The detection rate can be as high as 100 percent as long as enough data are available. This is true even if heavy cross traffic exists. The data collected in this paper should give designers guidelines for the development and operation of mix networks.

The failure of traditional mix batching strategies directly leads us to the formulation of anew packet control method for mixes in order to overcome their vulnerability to flow correlation attacks. Appropriate output control can achieve a guaranteed low detection rate while maintaining high throughput for normal payload traffic. Our claim is validated by extensive performance data collected from experiments.

REFERENCES

1. D. Chum, “Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms,” Comm. ACM, vol. 24, no. 2, pp. 84-90, Feb.1981.
2. A. Saratov and G. Danzig, “Towards an Information Theoretic Metric for

Anonymity," Proc. Privacy Enhancing Technologies Workshop (PET '02), R. Dingle dine and P. Ryerson, eds., pp. 41-53, Apr. 2002.

3. C. Dı´az, S. Says, J. Claessens, and B. Preneel, "Towards Measuring Anonymity," Proc. Privacy Enhancing Technologies Workshop (PET '02), R. Dingle dine and P. Ryerson, eds., pp. 54-68, Apr. 2002.
4. Y. Zhu and R. Bettati, "Anonymity vs. Information Leakage in Anonymity Systems," Proc. 25th IEEE Int'l Conf. Distributed Computing Systems (ICDCS '05), pp. 514-524, 2005.
5. O.R.D. Achives, "Link Padding and the Intersection ttack,"http://archives.seul.org/or/dev, 2002. [6] P.F. Ryerson, D.M. Goldschlag, and M.G. Reed, "Anonymous Connections and Onion Routing," Proc. IEEE Symp. Security and Privacy, pp. 44-54, 1997.
6. R. Dingle dine, N. Mathewson, and P. Ryerson, "Tor: The Second-Generation Onion Router," Proc. 13th USENIX Security Symp, pp. 303-320, Aug. 2004.
7. M.K. Reiter and A.D. Rubin, "Crowds: Anonymity for Web Transactions," ACM Trans. Information and System Security, vol. 1, no. 1, pp. 66-92, 1998.
8. K. Suh, D.R. Figueiredo, J. Kurose, and D. Towsley, "Characterizing and Detecting Skype-Relayed Traffic," Proc. 25th IEEE INFOCOM '06, pp. 1-12, Apr. 2006.
9. Y.J. Pyun, Y.H. Park, X. Wang, D.S. Reeves, and P. Ning, "Tracing Traffic through Intermediate Hosts that Repacketize Flows," Proc. 26th IEEE INFOCOM '07, pp. 634-642, May 2007.