DESIGN OF CELL-COUNTING-BASED ATTACK AGAINST TOR

[1]J. SUDHEER [2] S.K.A. MANOJ
[1]M.Tech (CSE), PYDAH College Of Engineering and Technology, Visakhapatnam, AP-India,
E-mail: j.sudheerraju@gmail.com
[2]Assistant Professor, Department of CSE

ABSTRACT- *Various low-latency anonymous communication systems such as Tor and Anonymizer have been designed to provide anonymity services for users. In order to hide the communication of users, the anonymity systems pack the application data into equal-sized cells via extensive experiments on Tor. We found the size of IP packets in the Tor network can be very dynamic because a cell is an application concept and the IP layer may repack cells. Based on this, investigates cell-counting-based attack against Tor, which allows attacker to confirm anonymous communication relationship among the users quickly. In this attack, by marginally varying the number of cells in the target traffic at the malicious exit onion router, then the attacker can embed a secret signal into the variation of cell counter of the target traffic. The embedded signal carried along with the target traffic and arrive at the malicious entry onion router. An accomplice of the attacker at the malicious entry onion router will detect the embedded signal based on the received cells and confirm the communication relationship among users. We have implement this attack against Tor and our experimental data validate its effectiveness and feasibility, there are several unique features of this attack. 1st this attack is highly efficient and can confirm very short communication sessions with only tens of cells. 2nd, this attack is effective, and its detection rate approaches 100% with a very low false positive rate. 3rd, it is possible to implement the attack in a way that appears to be very difficult for honest participants to detect.*

1. INTRODUCTION

Anonymity has become a necessary and legitimate aim in many applications including anonymous web browsing, location-based services (LBSs) and E-voting. In this applications, encryption alone cannot maintain the anonymity required by participants. In the past, researchers have developed numerous anonymous communication systems. Mix techniques used for either message-based (high-latency) or flow-based (low-latency) anonymity applications. The E-mail is a typical message-based anonymity applications, which has been thoroughly investigated. The research on flow-based anonymity applications has recently received great attention in order to preserve anonymity in low-latency applications including the Web browsing and peer-to-peer file sharing.Existing traffic analysis attacks can be categorized into two groups: passive traffic analysis and active watermarking techniques. Passive traffic analysis technique will record the traffic passively and identify the similarity between the sender's outbound traffic and the receiver's inbound traffic based on statistical measures.

Because this type of attack relies on correlating the timings of messages moving through the anonymous system and does not change the traffic characteristics, it is also a passive timing attack. To improve the accuracy of attacks, the active watermarking technique has recently received much attention. The idea of this technique is to actively introduce special signals (or marks) into the sender's outbound traffic with the intention of recognizing the embedded signal at the receiver's inbound traffic.

In this paper, we focus on the active water marking technique, it has been active in the past few years. For example, proposed a flow-marking scheme based on the direct sequence spread spectrum (DSSS) technique by utilizing a pseudo-noise (PN) code. By interfering with the rate of suspect sender's traffic and marginally changing the traffic rate, then the attacker can embed a secret spread-spectrum signals into the target traffic. The embedded signals are carried along with the target traffic from the sender to receiver. Investigator can recognize the corresponding communication relationship tracing the messages despite the use of anonymous networks. In order to accurately confirm the anonymous communication relationships of users, the flow-marking scheme needs to embed a signal modulated by a relatively long length of PN code and also the signal is embedded into traffic flow rate variation.

Proposed a non blind network flow watermarking scheme called RAINBO for stepping stone detection. Their approach records traffic timing of the incoming flows and correlates them with the outgoing flows. This approach also embed watermarks into the traffic by actively delaying some packets. The watermark detection problem was formalized as detecting a known spread-spectrum signal with noise caused by network dynamics. Normalized correlation is used as the detection scheme. Their approach can classify a typical SSH connection as a stepping stone connection in 3 min. As we can see, it is hard for the flow-marking technique to deal with the short communication sessions that may only last for a few seconds.

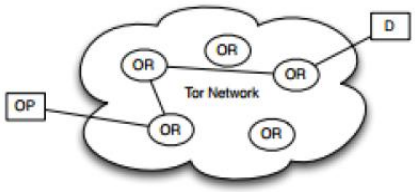
A successful attack against anonymous communication systems relies on accuracy, efficiency and detects ability of active watermarking techniques. Detect ability refers to the difficulty of detecting the embedded signal by anyone other than the attackers. Efficiency refers to the quickness of confirming anonymous communication relationships among users.

Although accuracy and/or detect ability have received great attention to the best of our knowledge, no existing work can meet all these three requirements simultaneously.

2.ALLIED WORK

The basic idea of this overlay network is to construct a circuit, which consists of onion routers (OR) that know only its predecessor and successor. User uses circuit to pass data through the Tor network anonymously. Data is wrapped in layers using symmetric cryptography and in each onion router as the data goes through; a layer is unwrapped by using a symmetric key and relayed forward. At the end of the circuit, onion router relays data to the intended destination. The destination is not required to run Tor related each onion router in the Tor network is connected to every other onion router using TLS. TLS is used to prevent possible attackers from being able to modify data, impersonate an onion router, and read the plaintext data by keeping the data secret in the connections. Tor users use onion proxy (OP) to receive directory information, create circuits in the network and manage user application connections.

Streams are multiplexed in a circuit, thus a one circuit can contain multiple TCP streams. Circuits are created preemptively by the onion proxies and rotated periodically to avoid traffic analysis. The List of available onion routers that can be chosen to the circuit are downloaded from a signed directory service. Upon creating a circuit, the user's onion proxy negotiates a symmetric key with every onion router in the circuit. Onion proxy always commands the last onion router in the circuit to extend one hop further until all intended onion routers are included.



During this creation, onion router does not care who open the circuit, but onion proxy knows the onion router. In other words, this handshake protocol is an unilateral entity authentication and provides forward secrecy. After the circuit has been established, relay cells can be sent. A cell is a fixed-size, 512 bytes; unit of c Tor is a popular overlay network for providing anonymous communication over the Internet.

Components of Tor:

Tor is a popular overlay network for providing anonymous communication over the Internet. It is an open-source project and provides anonymity service for TCP applications. As shown in Figure, there are four basic components in Tor.

1) Alice (i.e., Client): The client runs local software called *onion proxy* (*OP*) to anonymize the client data into Tor.

2) Bob (i.e., Server): It runs TCP applications such as a Web service.

3) Onion routers (ORs*):* Onion routers are special proxies that relay the application data between Alice and Bob.

In Tor, transport-layer security (TLS) connections are used for the overlay link encryption between two onion routers. This application data is packed into equal-sized cells carried through TLS connections.

4) Directory servers: They hold onion router information Such as public keys for onion router. Directory authorities hold authoritative information on onion routers and directory caches download directory information of onion routers from authorities. A list of directory authorities is hard-coded into the Tor source code for a client to download the information of onion routers and build circuits through the Tor network.

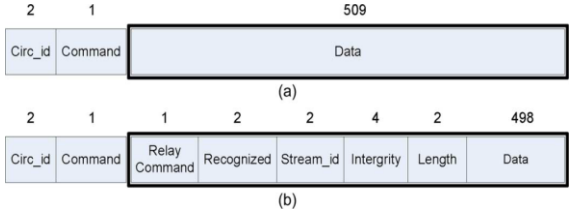


Fig 2.1 Cell format by Tor.
(a) Tor cell format.
(b) Tor relay cell format.

3.CELL-COUNTING-BASED ATTACK

In this section, we first show that the size of IP packets in the Tor network is very dynamic. Based on this finding, we then introduce the basic idea of the cell-counting-based attack and list some challenging issues related to the attack and present solutions to resolve those issues.

Dynamic IP Packet Size of Traffic over Tor

In Tor, the application data will be packed into equal-sized cells. Nonetheless, via extensive experiments over the Tor network, we found the size of IP packets transmitted over Tor is dynamic. It can be observed that the size of packets from the sender to the receiver is random over time, and a large number of packets have varied sizes, other than the cell size or maximum transmission unit (MTU) size.

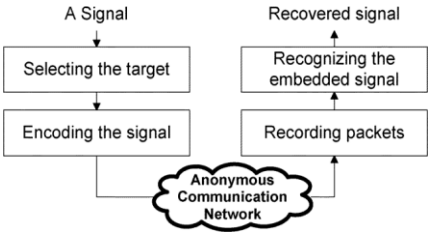


Fig 3.1 Cell-counting-based attack.

Basic Idea of Cell-Counting-Based Attack

The basic idea is as follows. An attacker at the exit onion router first selects the target traffic flow between Alice and Bob. The attacker selects a random signal, chooses an appropriate time, and changes the cell count of target traffic based on the selected random signals. In this way, the attacker is able to embed a signal into the target traffic from Bob.

The signal will be carried along with the target traffic to the entry onion router connecting to Alice. An accomplice of the attacker at the entry onion router will record the variation of the received cells and recognize the embedded signal. If the same pattern of the signal is recognized, the attacker confirms the communication relationship between Alice and Bob.

4. EXPERIMENTAL EVALUATION

We have implemented the cell-counting-based attack presented in Section III against Tor. In this section, we use real-world experiments to demonstrate the feasibility and effectiveness of this attack. All the experiments were conducted in a controlled manner, and we experimented on TCP flows generated by ourselves in order to avoid legal issues.

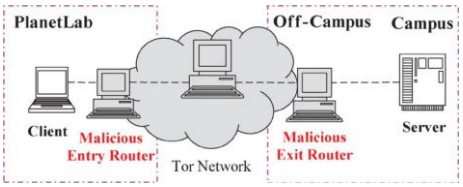


Fig 4.1 Experiment setup

A. Experiment Setup

In our experiment setting illustrated in Fig 4.1, we deployed two malicious onion routers as the Tor entry onion router and exit onion router. The entry onion router and client (Alice) located in Asia are deployed on Planet Lab. The server (Bob) is located at one university campus in North America, and the exit onion router is at an off-campus location in North America as well. All computers are on different IP address segments and connected to different Internet service providers (ISPs). Fig 4.1 shows the experiment setup.

The Tor client will intend to setup circuits through the designated malicious exit onion router and entry onion router shown in Fig4.1 The middle onion router is selected using the default routing selection algorithm released by Tor. As we stated earlier, the cell-counting-based attack intends to confirm whether the client (Alice) communicates with the server (Bob). For verification purpose, we set up a server (Bob) and download a file from the client (Alice). By using the Tor configuration file and manipulatable parameters, such as Entry Nodes, Exit Nodes, Strict Entry Nodes, and Strict Exit Nodes, we let the client choose both the malicious entry and exit onion routers along the circuit.

B. Experimental Results

To obtain the empirical property of IP packet size for the traffic within the Tor network, we downloaded a file with the size of 20M using the Tor network. Fig 4.1 shows the empirical cumulative probability function (CDF) of the IP packet size in the traffic. We know that the packets with non-MTU size are around 50%. This validates that the size of packets transmitted over the Tor is dynamic. Consequently, it also indicates that our embedded signal will be hidden in the normal traffic and hard to be detected by victims. To validate the accuracy of the cell-counting-based attack, we let the client download 30 files in our experiments.

The size of each file is around 10 MB. At the exit onion router, we generate a random signal with 100 b. When the target traffic from server Bob arrives at the exit onion router, we vary the number of cells in the circuit and embed the signal into the variation of the cell count during a short period in the target traffic. At the entry onion router, the cells in the circuit queue are recorded in the log, and the recovery mechanisms will be applied to recognize the embedded signal.

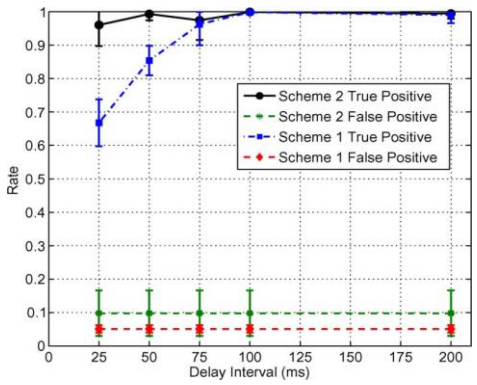


Fig 4.2 Detection rate versus delay interval (Note: The rate is for detectingone bit)

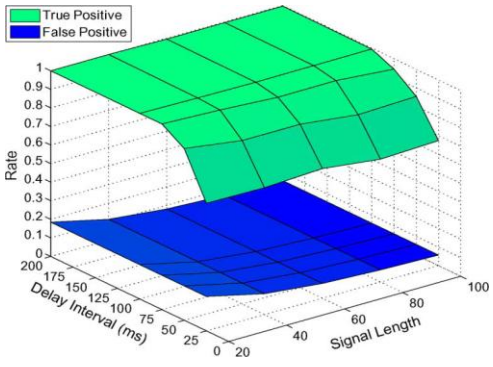


Fig 4.3 Detection rate versus delay interval and signal length with detection scheme 1 (Note: The rate is for detecting one bit)

We conduct the above experiment to evaluate the true positive and false positive by using a 100-b random signal. Fig 4.2 illustrates the correlation between the detection rate (true positive) and the delay interval for transmitting cells associated to different units of the signal.

As we can see from this figure, the detection rate will increase dramatically when the delay interval is slightly increased in two detection schemes. As expected, the detection rate of scheme 2 is higher than scheme 1 with a slightly increasing false positive rate, while the overall false positive rate for each scheme is a fixed value. When the delay interval approaches 100 ms, the detection rate of two schemes approaches 100%. All these findings validate that our investigated attack can significantly degrade the anonymity service provided by Tor.

4. CONCLUSION

In this paper, we introduced a novel cell-counting-based attack against TOR. This attack is difficult to detect and able to quickly and accurately confirm the anonymous communication relationship among users on Tor. An attacker at the malicious exit onion router slightly manipulates the transmission of cells from a target TCP stream and embeds a secret signal (a series of binary bits) into the cell counter variation of the TCP stream. An accomplice of the attacker at the entry onion router recognizes the embedded signal using our developed recovery algorithms and links the communication relationship among users. Our theoretical analysis shows that the detection rate is a monotonously increasing function with respect to the delay interval and is a monotonously decreasing function of the variance of one way transmission delay along a circuit. Via extensive real-world experiments on Tor, the effectiveness and feasibility of the attack is validated. Our data showed that this attack could drastically and quickly degrade the anonymity service that Tor provides.

6. REFERENCES

[1] N. B. Amir Houmansadr and N. Kiyavash, "RAINBOW: A robust and invisible non-blind watermark for network flows," inProc. 16th NDSS, Feb. 2009, pp. 1–13.

[2] X. Fu, Z. Ling, J. Luo, W. Yu,W. Jia, and W. Zhao, "One cell is enough to break Tor's anonymity," in Proc. Black Hat DC, Feb. 2009 [Online]. Available: http://www.blackhat.com/presentations/bh-dc-09/Fu/ BlackHat-DC-09-Fu-Break-Tors-Anonymity.pdf

[3] L. Overlier and P. Syverson, "Locating hidden servers," in Proc. IEEE S&P, May 2006, pp. 100–114.

[4] G. Danezis, R. Dingledine, and N. Mathewson, "Mixminion: Design of a type III anonymous remailer protocol," in Proc. IEEE S&P, May 2003, pp. 2–15.

[5] G. Smillie, Analogue, Digital Communication Techniques. London, U.K.: Butterworth-Heinemann, 1999

[6] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second generation onion router," in Proc. 13th USENIX Security Symp., Aug. 2004, p. 21.

[7] A. Serjantov and P. Sewell, "Passive attack analysis for connectionbased anonymity systems," in Proc. ESORICS,Oct. 2003, pp. 116–131.

[8] B. N. Levine,M. K. Reiter, C.Wang, andM. Wright, "Timing attacks in low-latency MIX systems," in Proc. FC, Feb. 2004, pp. 251–565.

[9] Y. Zhu, X. Fu, B. Graham, R. Bettati, and W. Zhao, "On flow correlation attacks and countermeasures in Mix networks," in Proc. PET,May 2004, pp. 735–742.

[10] "Anonymizer, Inc.," 2009 [Online]. Available: http://www. anonymizer.com/

[11] S. J. Murdoch and G. Danezis, "Low-cost traffic analysis of Tor," in Proc. IEEE S&P, May 2006, pp. 183–195

[12] X.Wang, S. Chen, and S. Jajodia, "Network flow watermarking attack on low-latency anonymous communication systems," in Proc. IEEE S&P, May 2007, pp. 116–130.

[13] K. Bauer, D. McCoy, D. Grunwald, T. Kohno, and D. Sicker, "Lowresource routing attacks against anonymous systems," in Proc. ACM WPES, Oct. 2007, pp. 11–20.