Attack against Anonymous Network

Akshatha Prabhu K

Abstract*— This paper focuses on the active watermarking technique, which has been active in the past few years. This paper proposes a flow-marking scheme based on the direct sequence spread spectrum technique by utilizing a pseudo-noise code. By interfering with the rate of a suspect sender's traffic and marginally changing the traffic rate, the attacker can embed a secret spread-spectrum signal into the target traffic. The embedded signal is carried along with the target traffic from the sender to the receiver, so the investigator can recognize the corresponding communication relationship, tracing the messages despite the use of anonymous networks. However, in order to accurately confirm the anonymous communication relationship 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 the traffic flow rate variation. After the signal is embedded and delay is added between cells, we gone generate Mimo graph, it gives the probability of how much data is been extracted.

Index Terms*— attack against TOR.

I. INTRODUCTION

Anonymity has become a necessary and legitimate aim in many applications, including anonymous Web browsing, location-based services (LBSs), and E-voting. In these applications, encryption alone cannot maintain the anonymity required by participants. In the past, researchers have developed numerous anonymous communication systems. Generally speaking, mix techniques can be used for either message-based (high-latency) or flow-based (low-latency) anonymity applications. E-mail is a typical message-based anonymity application [4], which has been thoroughly investigated. Research on flow-based anonymity applications has recently received great attention in order to preserve anonymity in low-latency applications, including 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 senders outbound traffic and the receivers 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. 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.

Manuscript received April, 2014.

Akshatha Prabhu K, received B.E degree in Information Science From VTU in 2010, Persuing MTECH degree in Computer Science and Engineering from VTU. Served as Lecturer under VTU from 2010 to 2012.

In our real-world experiments, the experimental results match the theoretical results well. To be specific, our attack needs only 2 s to achieve a true positive rate of almost 100% and the false positive rate of almost 0%. The attack presented in this paper is one of the first to exploit the implementation of known anonymous communication systems such as Tor by exploiting its fundamental protocol design.

There are several unique features for this attack. First, this attack is highly efficient and can quickly confirm very short anonymous communication sessions with tens of cells. Second, this attack is effective, and its detection rate approaches 100% with very low false positive rate. Third, the short and secret signal makes it difficult for others to detect the presence of the embedded signal.

II. OVERVIEW OF CELL COUNTING ATTACK

In Tor [2], the application data will be packed into equal-sized cells (e.g., 512 B). Nonetheless, via extensive experiments over the Tor network [2], we found that 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. The varied performance of onion routers may cause cells not to be promptly processed. if an onion router is overloaded, unprocessed cells will be queued. Therefore, cells will be merged at the IP layer and sent out together. Those merged cells may be split into multiple MTU-sized packets and one non-MTU-sized packet.

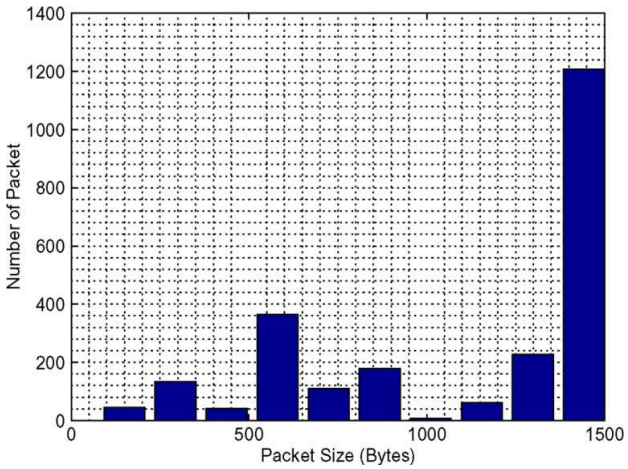

Fig.1 Number of packets versus packet size

III. LITERATURE SURVEY

There have been several recent cases in the news in which anonymous bloggers have or have not been exposed and have or have not lost jobs, etc, as a result, depending on the policy of their ISP, the interpretation of laws by various courts, and numerous other factors.

Recommendations for a technology to protect anonymous bloggers and other publishers, regardless of legal protection, would thus seem to be timely and encouraging. The Tor developers are careful, however, to warn against using Tor in critical situations: upon startup the Tor client announces, "This is Experimental software. Do not rely on it for strong anonymity." With increasing high-profile recommendations to use Tor's hidden services for applications, it is important to assess the protection they afford. Hidden servers have also been recommended for preserving the anonymity of the service offered and to resist censorship.

Recommendations for a technology to protect anonymous bloggers and other publishers, regardless of legal protection, would thus seem to be timely and encouraging. One of the major vulnerabilities for a hidden service in Tor is the server's selection of the first and last node in the communication path. To a first approximation, if an adversary can watch the edges of a Tor circuit, then she can confirm who is communicating. Wang *et al.* also investigated the feasibility of a timing-based watermarking scheme in identifying the encrypted peer-to-peer Vo IP calls. Peng *et al.* analyzed the secrecy of timing-based watermarking trace back proposed in, based on the distribution of traffic timing. This multifold-based approach intends to average the rate of multiple synchronized watermarked flows and expects to observe a unusual long silence period without packets or a unusual long period of low-rate traffic.

IV. BASIC IDEA OF CELL-COUNTING-BASED ATTACK

The packet size observed at the client shows a high probability to be random because of the performance of onion routers and Internet traffic dynamics. Motivated by this finding, we investigate a new cell-counting-based attack against Tor, which allows the attacker to confirm anonymous communication relationship among users very quickly.

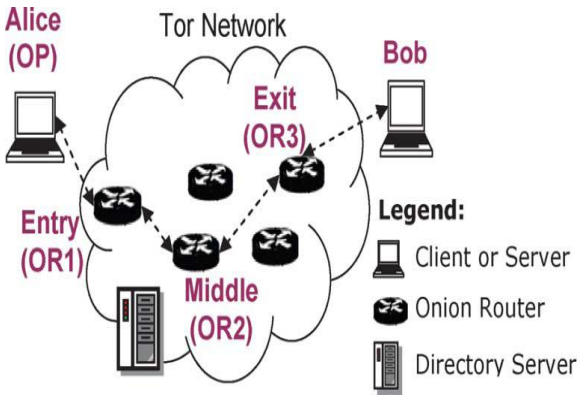

Fig.2 TOR Network

The basic idea is as follows. An attacker at the exit onion router [2] first selects the target traffic flow between Alice and Bob. The attacker then selects a random signal (e.g., a sequence of binary bits), chooses an appropriate time, and changes the cell count of target traffic based on the selected random signal. 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 [2] 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.

V. SELECTING THE TARGET

At a malicious exit onion router connected to the server Bob, the attacker will log the information, including server Bob's host IP address and port used for a given circuit, as well as the circuit ID. The attacker uses cells since those cells transmit the data stream.

VI. ENCRYPTION

Encryption [3] is the process of encoding messages or information in such a way that only authorized parties can read it. Encryption [3] doesn't prevent having but it reduces the likelihood that the hacker will be able to read the data that is encrypted. In an encryption scheme, the message or information, referred to as plaintext, is encrypted using an encryption algorithm turning it into an unreadable cipher text. This is usually done with the use of an encryption key, which specifies how the message is to be encoded. An authorized party, however, is able to decode the cipher text using Decryption algorithm that usually requires a secret decryption key.

VII. DECRYPTION

Decryption [3] is the process of encoded or encrypted text or other data and converting it back into text that you or the computer are able to read and understand. This term could be used to describe a method of un-encrypting the data manually or with un-encrypting the data using the proper codes or keys.

VIII. ENCODING THE SIGNAL

in the circuit queue are all flushed into the output buffer. Hence, the attacker can benefit from this and manipulate the number of cells flushed to the output buffer all together. In this way, the attacker can embed a secret signal (a sequence of binary bits, i.e., 0101010) into the variation of the cell count during a short period in the target traffic. Particularly, in order to encode bit 1, the attacker flushes three cells from the circuit queue. In order to encode bit 0, the attacker flushes only one cell from the circuit queue. In order to accurately manipulate the number of the cells to be flushed, the attacker needs to count the number of cells in the circuit queue. Once the number of the cells is adequate (i.e,, three cells for encoding 010 bit of the signal, and one cell for 101 bit of the signal), the attacker calls the circuit write event promptly and all the cells are flushed to the output buffer immediately. Unfortunately, due to the network congestion and delay, the cells may be combined or separated at the middle onion routers, or the network link between the onion routers.

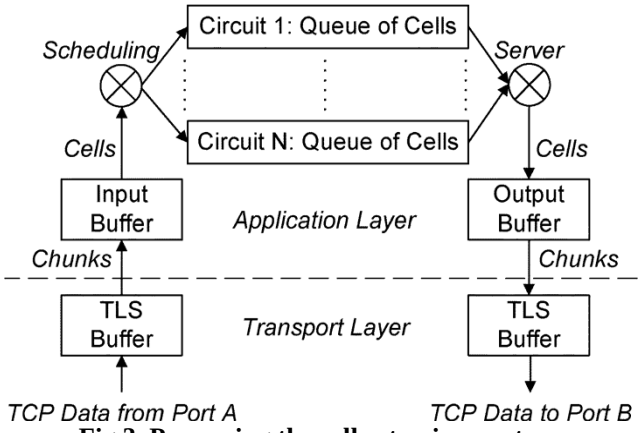

Fig.3. Processing the cells at onion routers.

IX. RECORDING PACKETS

After the signal is embedded in the target traffic in Step 2, it will be transmitted to the entry onion router along with the target traffic. An accomplice of the attacker at the entry onion router will record the received cells and related information, including Alice's host IP address and port used for a given circuit, as well as the circuit ID.

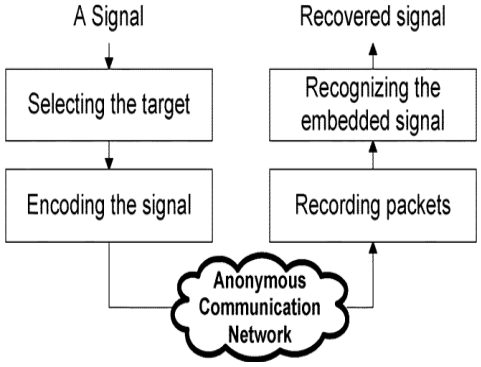

Fig.4 Cell-counting-based attack

Algorithm 1: Recovery Mechanism for Continuously Embedded Bits

Require: (a) C[1*m], an array storing the number of cell counter variation in the circuit queue at the entry router. (b)S[1*n] , an array storing the original signal bit.
1 i=0; j=0: ;
2: **while** i≤m **do**
3: **if** C[i]==S[j] **then**
4: Signal S[j] is matched.
5: **else if** C[i] < S[j] **then**
6: Signal S[j] is splitted.
7: **if** C[i] + C[i+1] == S[j] **then**
8: Signal S[j] is processed as Type I with k=1.
9: **else if** C[i] + C[i+1] > S[j] **then**
10: Signal S[j] and S[j+1] are processed as Type II with k = 1.
11: **else if** C[i] + C[i+1] < S**[j] then**
12: Find the value of k.
13: **if** C[i]+ …. +C[i+k] == S[j**] then**
14: Signal S[j] is processed as Type I with k ≥ 2 .
15: **else**
16: Signal S[j] and S[j+1] is processed as Type II With k ≥ 2 .
17: **end if**
18 i = i + k: ;
19: **end if**
20: **else if** C[i] > S[j] **then**
21: Two or more signals are combined together.
22: if C[i] == S[j] + S[j+1] **then**
23: Signal and are processed as Type II with k=1 .
24: **else if** C[i] < S[j] + S[j+1] **then**
25: Signal and are processed as Type IV With k=1 .
26: **else if** C[i] > S[j] + S[j+1] **then**
27: Find the value of k
28: **if** C[i] == S[j] +… S[j+k] **then**
29: These combined signals are processed as Type III with k ≥ 2 .
30: **else**
31: These combined signals are processed as Type IV with k ≥ 2 .
32: **end if**
33: j = j+ k:
34: **end if**
35: **end if**
36: i = i+1 ; j = j+1
37: **end while**

X. CONCLUSION

In this paper we introduced a cell-counting-based attack against Tor [2]. This attack is difficult to detect and is able to quickly and accurately confirm the anonymous communication relationship among users on Tor [2]. 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. 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 [2] provides.

ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in American English is without an "e" after the "g." Use the singular heading even if you have many acknowledgments. Avoid expressions such as "One of us (S.B.A.) would like to thank ... ." Instead, write "F. A. Author thanks ... ." Sponsor and financial support acknowledgments are placed in the unnumbered footnote on the first page.

REFERENCES

1. L. Øverlier and P. Syverson, "Locating hidden servers," in Proc. IEEE S&P, May 2006, pp. 100–114.
2. R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second generation onion router," in Proc. 13th USENIX Security Symp., Aug. 2004, p. 21.

3. Q. X. Sun, D. R. Simon, Y. Wang, W. Russell, V. N. Padmanabhan,and L. L. Qiu, "Statistical identification of encrypted Web browsing traffic," in Proc. IEEE S&P, May 2002, pp. 19–30.
4. A. Serjantov and P. Sewell, "Passive attack analysis for connection based anonymity systems," in Proc. ESORICS,Oct. 2003, pp. 116–131
5. R. Pries, W. Yu, S. Graham, and X. Fu, "On performance bottleneck of anonymous communication networks," in Proc. 22nd IEEE IPDPS, Apr. 14–28, 2008, pp. 1–11

AUTHORS PROFILE

Akshatha Prabhu k, received B.E degree in Information Science From VTU in 2010, Persuing MTECH degree in Computer Science and Engineering from VTU. Served as Lecturer under VTU from 2010 to 2012.