International Journal of Security and Its Applications
Vol. 10, No. 3 (2016), pp.129-138
http://dx.doi.org/10.14257/ijsia.2016.10.3.12

A Systematic Review of Network Flow Watermarking In Anonymity Systems

Tianbo Lu, Rui Guo, Lingling Zhao and Yang Li

School of Software Engineering，Beijing University of Posts and Telecommunications, 100876, Beijing, China
lutb@bupt.edu.cn, 123guorui123@163.com

Abstract

With the rapid development of internet technology, the connection between man and internet is closer and closer. When people is communicating with others through internet, some malicious intruders may want to eavesdrop or peep the communicators. In order to evade being watching, people use anonymous communication systems to communicate. The anonymity system can encrypt the content of communication and the identity of the communicators. But if the communicators want to know who is talking to them at the other end, they must correlate the outgoing and incoming flows to identify a host or a person. As an active traffic analysis approach, network flow watermarking technology can detect the correlation of flows, and then make the anonymous communicators accountable. While network flow watermarking technology achieves good detecting rate and low false positive rate, it could be an effective way to trace the communication connections and supervise anonymous communication. So it is a widely used way for tracing in anonymity systems. In this paper, we introduce some different schemes of network flow watermarking in anonymity systems and discuss some attacks against it. Finally, a conclusion will be given.

Keywords*: anonymous communication, network flow watermarking, security*

1. Introduction

Today, network intruders have become a big threat to the network security. They may compromise some hosts in certain communication routes to watch over the communication links. If someone is communicating with others in this compromised route, the intruder is able to get the content of the communication and even the identity information of the communicator. So for the sake of security, when people want to communicate with others without being known to malicious intruders, they can use anonymous communication systems such as Tor to communicate. This system can keep good anonymity, but the communicators at both sides may not know whether the other is believable. There are also some hidden network-based attackers who usually use a chain of hosts to relay their traffic to attack the victim. And these hosts are also called stepping stones. This process can be illustrated as following: the attacker attack the victim through some compromised hosts so that the victim can't realize who is attacking him through traditional trace-back methods, because he just know the nearest host is attacking him which actually not.

So it is important to trace anonymous communication in anonymity systems and detect stepping stones for protecting the computer security. In the past, a commonly used way was to find the correlations between incoming and outgoing flows by analyzing some patterns in them such as packet timings, sizes and counts [1,2,3] and it was apparently a passive analysis approach. This approach could be easily misled by the attacker who deliberately adjust the timing characteristics of a long flow. By this way, the attacker can

ISSN: 1738-9976 IJSIA

make some unrelated flows seem correlated and make the detecting results be false positive. Recently, as an active traffic analysis technology, watermarking is proposed to be a major approach to trace anonymous communication connections and detect stepping stones [4]. Compared to the passive ways, watermarking is more robust, because it actively embed invisible unique tags which also called watermark into the flow. If this watermark is stable and robust enough, this watermark-embedded flow can be identified at the receiver end through some detecting ways. So it could be an effective way to correlate the flows that traverse a sequence of anonymous stepping stones.

This paper gives a brief literature review of the research work on network flow watermarking technology in anonymity systems, then demonstrates some basic and major network flow watermarking schemes, and we also introduce some attack methods against network flow watermarking, finally, we discuss the conclusion and future work.

2. Literature Review

Nowadays, with the increment of network-based malicious anonymous attack, the keeping of anonymity, detecting of stepping stones and research on network flow watermarking draw a lot of attention of university labs and institutes. The following figure shows the main universities which do the research on the network flow watermarking and its application on detecting stepping stones.

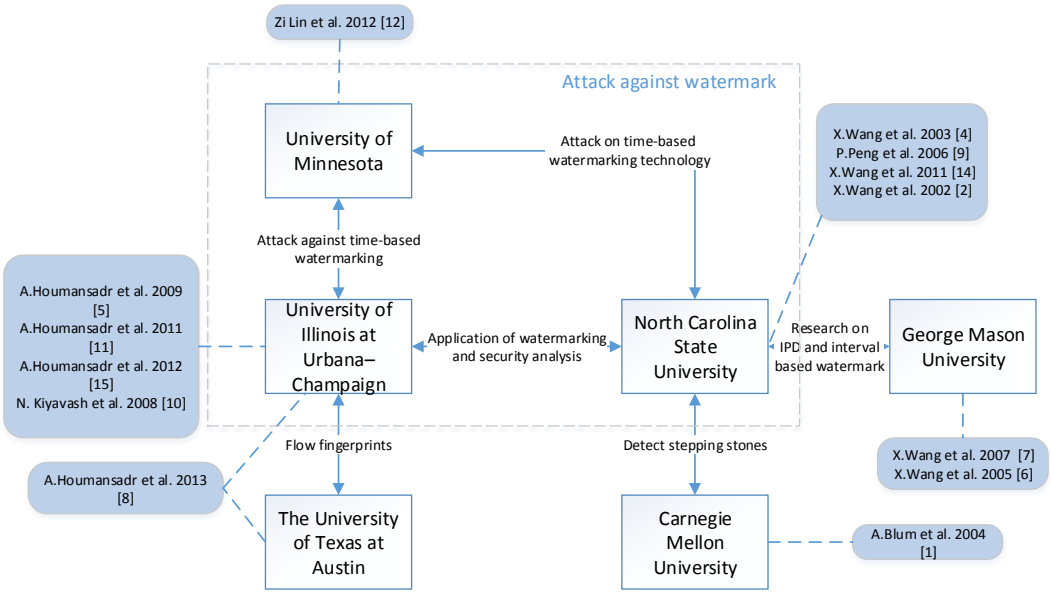


Figure 1 .Main Universities Doing Research on Network Flow Watermarking Technology in Anonymous Communication and their Relationships

As the figure shows, these universities above are all in American and are the leading forces of the study of network flow watermarking. Carnegie Mellon University (CMU) make contributions on the detecting stepping stone algorithm using the methods from Computational Learning Theory and the analysis of random walks [1]. North Carolina State University (NCSU) also study on the detection of stepping stones, but there is a little difference from CMU, they use watermarking-based active correlation analysis technology to do this job [4].George Mason University (GMU) mainly do a research on the security of traditional anonymous communication systems, and they proposed that the attacker can break the anonymity of peer to peer VoIP calls [6] and low-latency anonymous communication systems through active network flow watermarking technology. In order to do this work, GMU use inter packet-based and interval-based watermarking technology on which NCSU also do some research. University of Illinois at

Urbana–Champaign (UIUC) make great contribution on the proposal of network flow watermarking schemes to analyze the correlation between flows in anonymous communication, such as RAINBOW [5], SWIRL [11] and flow fingerprinting [8] .And the flow fingerprinting is proposed with the help of The University of Texas at Austin. UIUC also find that the previous interval-based watermarking is not secure and they propose the Multi-flow Attack [10]. University of Minnesota makes contribution on the proposal of a new attack on timing-based network flow watermarks [12] based on the accomplishment of NCSU [9] and UIUC.

3. Typical Anonymous Network Flow Watermarking Schemes

We will introduce some popular network flow watermarking schemes used to trace anonymous communication in anonymity systems and some attack methods in this section. These schemes may have some differences from each other, but we can infer that they are developing constantly based on the contributions that predecessors make. In Figure 2, we sketch the developing line of these schemes, so that we can find the relation among them, and find out whose research work contributes significantly to the developing of this technique.

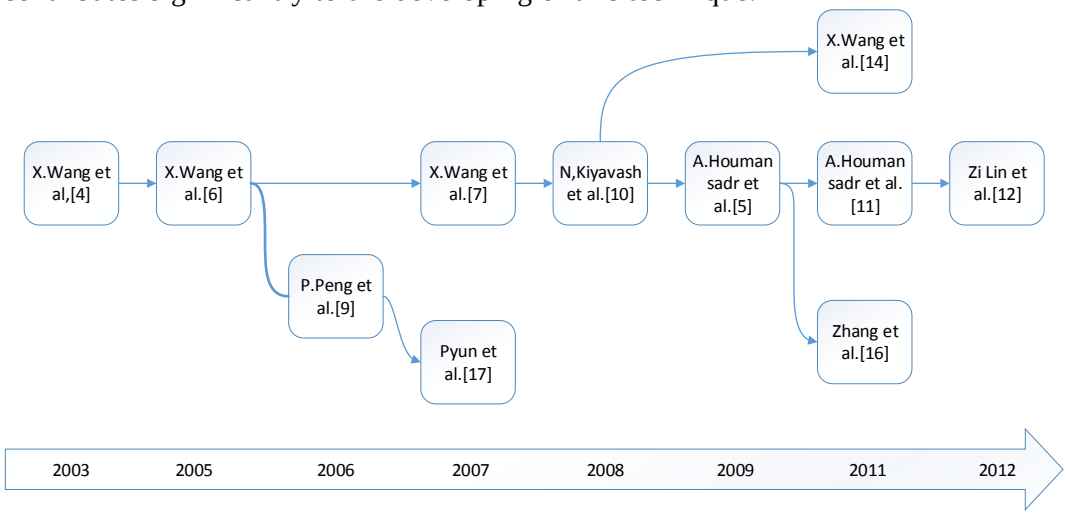


Figure 2. Developing Line of Anonymous Network Flow Watermarking Schemes

3.1 Overview

After the concept of network flow watermarking was introduced, scholars around the world has realized that this technology was a fantastic way to trace anonymous communication and they proposed many kinds of network flow watermarking schemes. These schemes are mainly choose certain characteristics which are independent of packet content to embed in the flow, and these characteristics can be detected and recovered after traversing the anonymous communication connections. These selected characteristics are also called watermarking carrier. According to the type of carrier that different watermarking schemes used, present popular watermarking schemes can be classified into three types: inter packet delay-based, interval-based and interval centroid-based. Figure 3 shows the major network flow watermarking schemes and some attack schemes against watermarking. Basically, inter packet delay-based watermarking schemes are vulnerable to timing analysis attack, and interval-based and interval centroid-based watermarking schemes are vulnerable to multi flow attack.

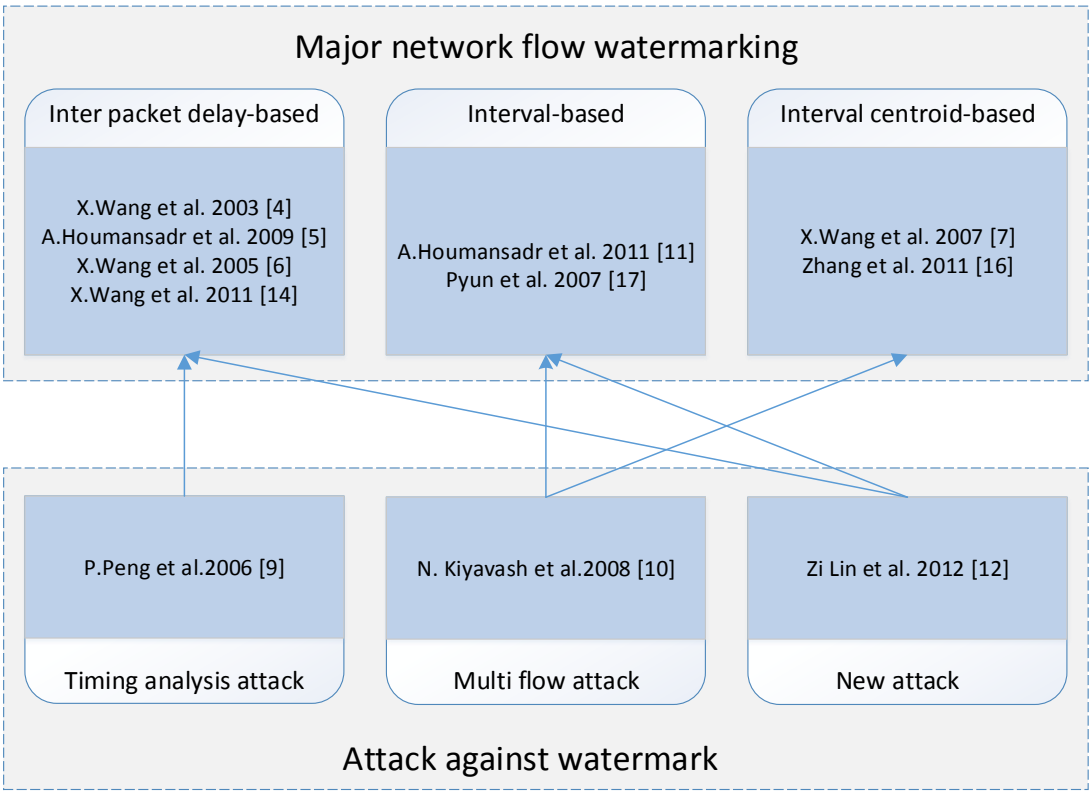


Figure 3. Major Network Flow Watermarking Schemes in Anonymous Communication

3.2 Inter Packet-Based Anonymous Network Flow Watermarking

Inter packet-based network flow watermarking choose inter packet delay (IPD) as the watermarking carrier in order to achieve better correlation analysis results. The IPD is the interval of a packet in a flow between arrive and departure time, and we can adjust some selected IPDs or the average size of IPDs to embed the watermarking information bit.

In order to solve the timing perturbations involved by attackers, Wang *et al*. [4,14] proposed a novel watermarking-based correlation scheme. They randomly choose two packet in the data flow and calculate their ipd, then they change this into a new value by using a formula they proposed. After doing this, they finish embedding the watermark at the sender end. At the receiver end, they decode the watermark to correlate the network flows. They also proposed that using the average ipd can improve the robustness against timing perturbation by the attacker. The advantages of Wang's method are that it can be used in short flows and embed more watermark information bits. But this watermark correlation approach is not as robust against non-independent random delays, and when they are calculating the average of ipds, they need a packet buffering to store certain packets of a flow. This will increase the delay of packets so that it can't be used in tracing real-time flows.

To solve this problem, Wang *et al*. [6] proposed another network flow watermarking technology in his paper whose research target is the trace of anonymous peer-to-peer VoIP calls. This technology is similar to the one above, but it doesn't quantify the ipd of data flow, it adjusts the increment parameter to change the average of a group of normalized IPD differences. So it can embed the watermark according to the change which presents the watermarking bits. Their work shows that it is feasible to trace anonymous peer-to-peer VoIP calls on the Internet and low latency anonymizing networks are not safe enough as we thought.

However, Peng *et al*. [9] found that the usual inter packet delay-based watermarking schemes were not secure enough and vulnerable to be attacked by intelligent attackers. They proposed an attack thinking that based on the analysis of packet delays between adjacent stepping stones. They thought the parameters that IPD-based watermarking schemes used are the keys to the security of it. And they proposed an algorithm to infer the important parameters. If the tracer is not carefully enough when he is choosing the watermark parameters, the malicious attackers may know these parameters. So that they can detect the watermark embedded in the flows and remove it even they can duplicate a new watermark in other normal flows. These attack methods can make the tracing of anonymous attacker meaningless, because the tracer can't detect the origin watermark or correlate the malicious activities with other benign users.

Different from the two methods above, there are some researchers use non-blind watermark to analyze the traffic correlation. Among them, A. Houmansadr is a great one. He and his colleagues proposed a non-blind watermarking scheme named "RAINBOW" [5,15] which achieves good robustness against the attack we mentioned above.

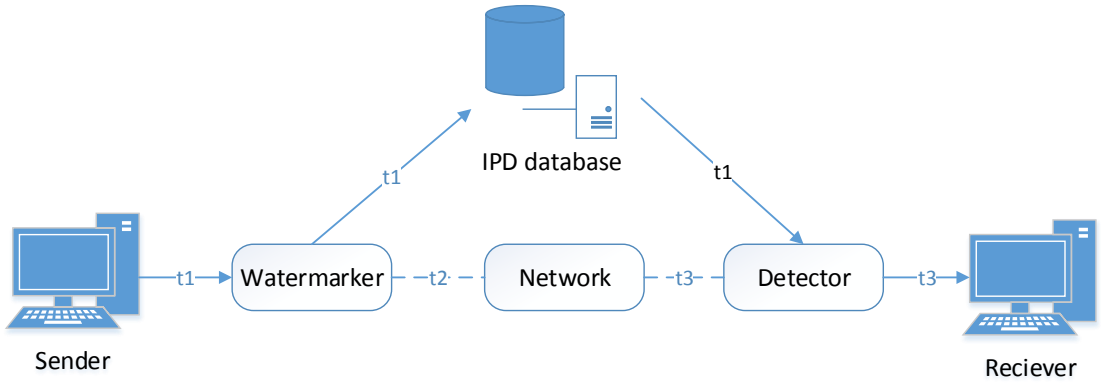


Figure 4. Model of RAINBOW Network Flow Watermarking Scheme [5]

Figure 4 sketch the model of RAINBOW network flow watermarking scheme. The IPD database is added in this scheme which is used to store the ipds. The major feature of this method is that they need the origin IPD value of flows when detecting watermarking at the receiver end. The amount of delays that RAINBOW used is far less than normal watermarking technologies because it eliminate the disturbance caused by the flow in the blind case. And it is also invisible enough to prevent being detected and removed by the attacker. Meanwhile, it has good robustness against packet loss and repacketization. But everything has two sides, because of the use of non-blind watermark, they have to use a database to record lots of IPDs, so it consumes more time in comparison between the watermark to be detected and the ones existed in the database than other blind ones.

Lin *et al*. [12] also proposed some attack schemes against RAINBOW in their paper. They use known flow attack and output-only detection attack to detect watermark embedded in RAINBOW schemes and achieve great detection rates though the watermark is long and the watermark key is unknown. After that, they utilize replay attack against RAINBOW from an isolated adversary's view to confuse the decoder and let the decoder can't correlate two flows correctly. Their work shows that a strong adversary is a big threat to the network security, so we should pay attention to this kind of attack.

3.3 Interval-Based Anonymous Network Flow Watermarking

IPD-based watermarking technology is proven an effective method in tracing traffic of an anonymous communication system, it can achieve a highly robust rate against a limited amount of perturbation involved by an attacker. However, Pyun *et al*. [17] found that that technology is not robust enough when the traffic is transformed at certain stepping stones, such as changing the packet count of the traffic flow. These transformations break the

synchronization of the packets which IPD-based watermark requires. So they proposed a new method against the timing perturbation and repacketization which uses the interval as the watermark carrier, so this method is also called Interval-based watermark (IBW). They slice the duration of each flow into fixed-length intervals which are self-synchronized, and they also adjust the packet timing to manipulate the count of packets in certain intervals. In most conventional watermarking and cryptography techniques, the encode end and decode end share a key together. But IBW make an assumption that some watermarking parameters are pre-distributed such as a stochastic offset, an interval length, an interval selection function, and a binary watermark. Firstly, the encode end encodes a selected flow with specific watermark w1. Secondly, at the decode end, they decode a watermarked flow and get another watermark w2. Thirdly, they compare w1 with w2 to analyze the correlation between two flows.

In the embedding process of IBW, two elementary operations are important, and they are load and clear. Assuming that there are two adjacent intervals I1 and I2, the load operation delay all the packets in interval I0 to I1, and the clear operation delay all the packets in interval I1 to I2. When embedding a watermarking bit '0', IBW load the packets in interval I1 and clear the packets in interval I2. When embedding a watermark bit '1', the method first clear I2 and then load I2. By this way, IBW can effectively accomplish the embedding of watermark.

These processes make this new watermark effective against repacketization and perturbation of the traffic timing and achieve high detection rates and low false positive rates. But this kind of watermark is not robust against the interference of chaff packets and is vulnerable to multi flow attack (MFA) [10]. This kind of attack can be used in almost all network flow watermarking applications that are interval-based or interval centroid-based. Especially, for interval-based watermarking systems, MFA can detect the watermarks embedded in flows and even remove, modify and rebuild the watermarks. To defeat this attack, different flows should be watermarked by multiple time interval assignments.

With the purpose of overcoming traditional IBW's weakness, A. Houmansadr *et al.* [11,15] proposed another watermarking technology named "SWIRL" that is also interval-based.

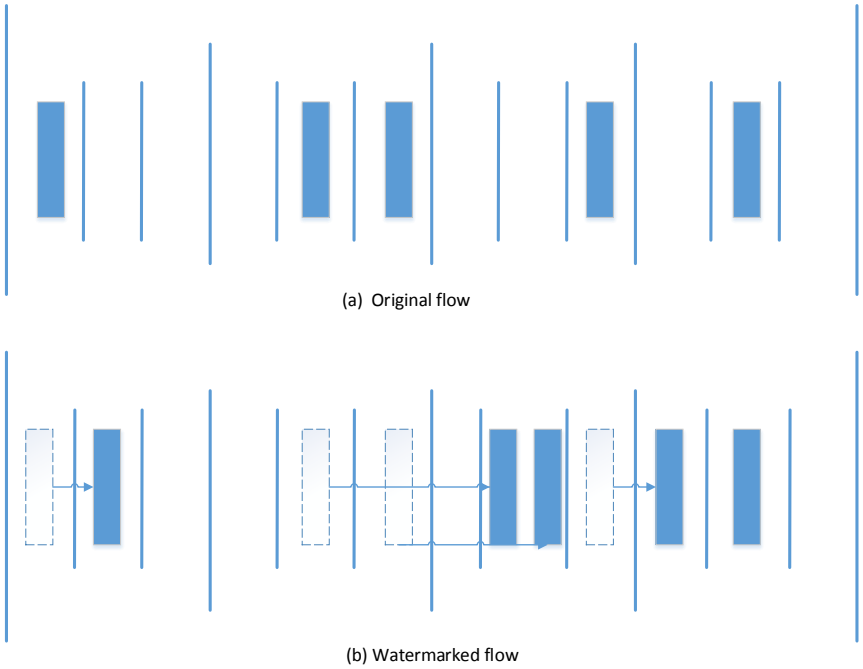


Figure 5. Delaying Packets to Insert the Watermark (4 Subintervals, 3 Slots in each Subinterval) [11]

SWIRL uses a novel method to defend against multi-flow attack, packet loss and network jitter. As Figure 5 shows, a selected mark interval is split into four subintervals, and each subinterval is divided into three slots. A slot is selected in each subinterval and then each packet is delayed so that the selected slot is also delayed. Because the selecting process is controlled by a random watermark parameter, the pattern in the mark interval is distinct. In brief, its pattern is chosen on the base of the characteristics of the flow being marked, so every flow is marked by a different pattern. SWIRL also introduces small delay to the traffic flows which enable it to be used in real world and it is proven to be a practical way to defense where previous traffic analysis methods would not be appropriate. But in some cases, the watermark embedded in SWIRL can be effective detected by malicious attacker, and the attacker is able to transfer watermarks in one flow to another innocent flows. These attack is essential in some anonymous communication systems, such as Tor. [12].

3.4 Interval Centroid-Based Anonymous Network Flow Watermarking

People always think that an anonymous communication system can achieve good anonymity by transforming flows such as traffic padding, adding bogus packets, flow mixing, flow splitting and flow merging which are described in Figure 6 and Figure 7. But Wang *et al*. [7] proposed a novel watermarking scheme named "Interval Centroid Based Watermarking"(ICBW). This scheme embed a watermark bit through adjusting the timing offset of certain packets in two intervals. And it can make any long enough flow identifiable even though it is transformed by those methods mentioned above. They have achieved good results in attacking low-latency anonymous communication system through this watermarking technology. Their work demonstrates that present anonymity may be broken down at any time and the attacker always advance the development of anonymous technology.

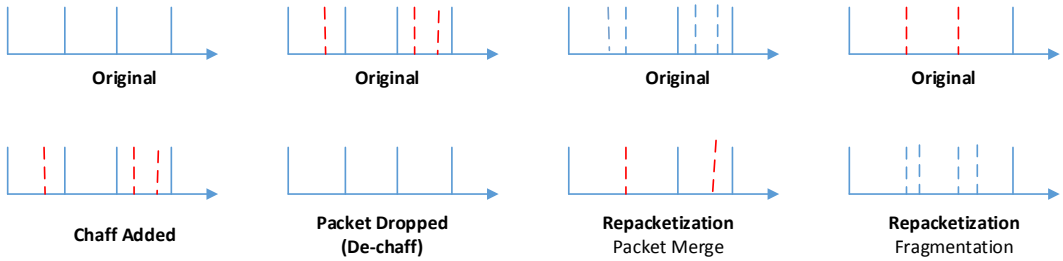


Figure 6. Intra-Flow Transformations [7]

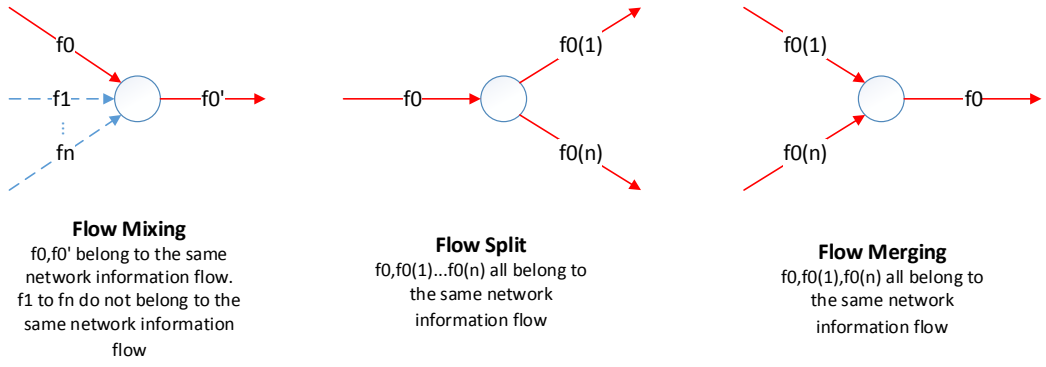


Figure 7. Inter-Flow Transformation [7]

However, this scheme is also vulnerable to multi flow attack [10] which means that the watermarking information that embedded in flows can be detected by the attacker who use MFA.

Zhang *et al*. [16] also do some research on ICBW. They found that some existing watermarking schemes cannot effectively trace the interactive traffic which most anonymous communication systems produced such as Web browsing, instant communication and remote login. So they proposed an anonymous traceback technique based on ICBW which can adapt to the trace of interactive traffic. In order to improve the robustness of the watermark, they choose a group of intervals to be the watermarking carrier. The watermarking carrier based on interval centroid is independent of certain flows, so it is stable, and it can trace both interactive and non-interactive traffic flows, hence it is universally adapted. Otherwise, the random selection of interval section and various distribution methods make it robust against MFA, meanwhile ensure the anonymity of trace back.

4. Conclusion

Communicating through anonymity system can protect people's privacy, because the traffic in an anonymous link is encrypted and can't be traced. Network flow watermarking technology plays an important role in this system. By actively embedding the unique and robust enough watermarks into the flows at the sender end, the receiver can detect these watermarks and confirm the identity of sender. And this technology can be applied to detect stepping stones which are used by malicious attackers to evade traceback. So this is an important technique in network security. This article gives a review of network flow watermarking technologies including current network flow watermarking research situation of some famous universities, popular network flow watermarking schemes at present and some attacks against them.

It is seen that related technologies are developing rapidly, and every scheme has its strength and weakness. The threat from malicious attackers will never be eliminated, so researchers are still facing great challenge in enhancing the robustness and invisibility of these watermarking schemes. We can infer that with the development and gradual maturity of technologies, the network flow watermarking will be more universal and effective in protecting network security.

Acknowledgements

This work is supported by the following programs: the National Natural Science Found ation of China under Grant No.61170273; 2010 Information Security Program of China N ational Development and Reform Commission with the title "Testing Usability and Securi ty of Network Service Software".

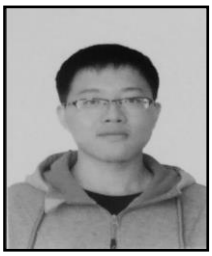
References

[1] Blum, Avrim, D. Song, and S. Venkataraman. "Detection of Interactive Stepping Stones: Algorithms and Confidence Bounds," Recent Advances in Intrusion Detection 3224**(2004)**.pp.258-277.
[2] Wang, Xinyuan, D. S. Reeves, and S. F. Wu. "Inter-Packet Delay Based Correlation for Tracing Encrypted Connections through Stepping Stones," European Symposium on Research in Computer Security 2502**(2002)**.pp. 244--263.
[3] Y. Zhang and V. Paxson. "Detecting stepping stones," USENIX Security Symposium, pages 171–184, Berkeley, CA, USA, Aug. **(2000)**. USENIX Association.
[4] X. Wang and D. S. Reeves, "Robust correlation of encrypted attack traffic through stepping stones by manipulation of interpacket delays," in ACM Conference on Computer and Communications Security. New York, NY, USA: ACM, **(2003)**, pp. 20–29.
[5] A. Houmansadr, N. Kiyavash, and N. Borisov, "RAINBOW: A Robust And Invisible Non-Blind Watermark for Network Flows," in Network and Distributed System Security Symposium, Feb. **(2009)**.

[6] X. Wang, S. Chen, and S. Jajodia, “Tracking anonymous peer-to-peer VoIP calls on the Internet,” in ACM Conference on Computer and Communications Security. New York, NY, USA: ACM, Nov. **(2005)**, pp. 81–91.
[7] X. Wang, S. Chen, and S. Jajodia, “Network flow watermarking attack on low-latency anonymous communication systems,” in IEEE Symposium on Security and Privacy, **(2007)**, pp. 116–130.
[8] A. Houmansadr, and N. Borisov. “The Need for Flow Fingerprints to Link Correlated Network Flows,” Privacy Enhancing Technologies. Springer Berlin Heidelberg, **(2013)**, pp. 205-224.
[9] P. Peng, P. Ning, and D. S. Reeves, “On the secrecy of timing-based active watermarking trace-back techniques,” in IEEE Symposium on Security and Privacy. IEEE Computer Society Press, May **(2006)**, pp. 334–349.
[10] N. Kiyavash, A. Houmansadr, and N. Borisov, “Multi-flow attacks against network flow watermarking schemes,” in USENIX Security Symposium. Berkeley, CA, USA: USENIX Association, **(2008)**.
[11] A. Houmansadr and N. Borisov, “SWIRL: A scalable watermark to detect correlated network flows,” in Proceedings of the Network and Distributed System Security Symposium, NDSS’11, **(2011)**.
[12] Zi Lin, and N. Hopper. "New attacks on timing-based network flow watermarks." Proceedings of the 21st USENIX conference on Security symposium USENIX Association, **(2012)**, pp.20-20.
[13] He Ting, and L. Tong. "Detecting Encrypted Stepping-Stone Connections." IEEE Transactions on Signal Processing 55.5 **(2007)**. pp. 1612-1623.
[14] Xinyuan Wang, and D. S. Reeves. "Robust Correlation of Encrypted Attack Traffic through Stepping Stones by Flow Watermarking." IEEE Transactions on Dependable & Secure Computing 8.3 **(2011)** pp. 434-449.
[15] A. Houmansadr. "Design, analysis, and implementation of effective network flow watermarking schemes." Dissertations & Theses - Gradworks **(2012)**.
[16] Zhang, Luo, *et al*. "Interval Centroid Based Flow Watermarking Technique for Anonymous Communication Traceback." Journal of Software 22.10**(2011)** pp. 2358-2371.
[17] Pyun, Young June, *et al*. "Tracing Traffic through Intermediate Hosts that Repacketize Flows." INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE IEEE, (**2007)**. pp. 634-642.

Authors

Tian-Bo Lu, He was born in Guizhou Province, China, 1977. He is an Associate professor in School of Software Engineering, Beijing University of Posts and Telecommunications, China. His technical interests include information security and computer network.



Rui Guo, He was born in Shaanxi Province, China, 1993. He is a graduate student in School of Software Engineering, Beijing University of Posts and Telecommunications, China. His technical interests include information and network security, anonymous communication.

Ling-Ling Zhao, She is a graduate student in School of Software Engineering, Beijing University of Posts and Telecommunications, China. Her technical interests include Cyber-Physical System and P2P network.

Yang Li, He was born in Hunan Province, China, 1978. He is a PhD and his technical interests include information security, distributed computing and P2P network.