Correlated Network Flows Detection

Olga Birth
Betreuer: Michael Herrmann
Hauptseminar- Innovative Internettechnologien und Mobilkommunikation SS2011
Lehrstuhl Netzarchitekturen und Netzdienste
Fakultät für Informatik, Technische Universität München
Email: birth@in.tum.de

ABSTRACT

Traffic monitoring and analysis plays an essential role in today's network security, since an unsecured network represents a grateful target to intruders. The goal of traffic analysis is, to obtain an intruders identity or to detect correlated network flows in order to allocate them to an individual. In that case, probable attacks can be answered appropriate.

Most offenders try to conceal their identity while performing attacks on their target destination. Thereby the most popoular way to stay hidden is, to link the traffic through several intermediate hosts, which had been compromised earlier. Correlated network flow detection (CNFD) would then try to detect these linked connections and reveal an attackers identity, even if the traffic is encrypted.

CNFD can also be used to detect an individuals identity in an *anonymous communication system.* Anonymous communication systems, were designed to obtain users anonymity while surfing the web. CNFD can detect senders and receivers identity and also the linkage between those two in an anonymous communication system. While traditional methods are performed by passively observing the possible connections and trying to find correlations by performing different statistical approaches [4][5][9][13], *watermarks* seem to be an elegant way to make CNFD more efficient and less expensive. Watermarking flows provide a novel approach, to reveal correlated flows. Nobody would even notice, if the traffic is been observed. Good watermarks can be inserted invisibly into the network and are more scalable than traditional passive analysis methods.

This paper is intended to give an overview of traffic analysis techniques, how they can be applied to detect correlated network flows and how watermarks can be used in this context.

Keywords

correlation, intrusion detection, watermark, flow transformation, active traffic analysis, stepping stones, anonymous communication systems

1. INTRODUCTION

Traffic analysis is the best way to keep track of all traffic that is traversing a network. If this is not performed carefully, an intruder can easily access a network and perform several attacks, without even being noticed [13]. An enemy usually knows everything about common monitoring techniques presented below, so if he/she wants to enter a network he/she always tries to stay anonymous. Besides spoofing the IP address, an intruder can obtain anonymity by using *stepping stones* [20]. Stepping stones are intermediate hosts that are used by an invader to launch an attack not from his own computer, but from compromised hosts. In addition to that, usually the traffic between the enemy and the target is encrypted. Without appropriate traffic analysis, nobody can never detect an intruder. To reveal such an attacker, it is very important to detect *similarities* between incoming and outgoing flows at the stepping stones [17]. This is also called *correlated network flow detection.*

CNFD can also be applied to anonymous communication systems. For a long time many have been convinced that with applying different transformations, a flow will become unique and so stable to correlation detection [16]. An attacker could now start applying several flow transformation techniques, in order to prevent unique network flows to be discovered [16]. This would modify a flow, that it would look completely different and could not be identified by an observer anymore. But there are still properties of flows, that cannot be erased by these transformations, like packet timing. This makes an flow, no matter how often the transformations have been applied, still unique [16].

This is where watermarking becomes important. The idea of watermarking is to uniquely identify a network flow by content-independent manipulations [4]. If two flows contain exactly the same pattern, they can be assumed to be linked. Watermarks are a new approach to traditional active monitoring techniques, because they need less computations than traditional techniques. Good watermarks are scalable, robust to packet losses and invisible [4]. This makes watermarks a good alternative to detect stepping stones and links in anonymous communication systems.

The remainder of this paper is structured as follows: the second Section is about the basics, such as traffic analysis methods, anonymous communication systems, stepping stones and different flow transformations. This should show, how traffic can be manipulated, in order to hide an individuals identity.

The second part is about traditional CNFD methods. This includes different correlation detection besides watermarking. In this work, watermarks have been picked, as a new and elegant approach to correlation detection in network flows. But there are other techniques, how correlated network flows can be detected.

Up next is a Section about watermarking, with the differ-

doi: 10.2313/NET-2011-07-2_13

ent watermarking approaches. It is aimed to provide a brief overview of the different watermarking techniques, without going into very detail.
Section 4 discusses the applications of watermarking and Section 5 concludes this topic.

2. BACKGROUND

There are several concepts that should be described first, such as diverse monitoring techniques and some term descriptions to provide the basics for this topic.
As mentioned above, there are several traffic monitoring techniques, which can basically be separated into two groups: the *router based monitoring techniques* and the *non-router based monitoring techniques* [6].
The difference between those two is simple: the former ones have the monitoring functionalities built in the routers, whereas the non-router based require further installation of hardware and software [6]. It would simply go beyond the scope to explain both techniques in detail. To understand this paper, there is no need to know the functionalities behind the router-based monitoring techniques. For further information on the router-based techniques, such as RMON or Netflow RFC see: [3][2].
The non-router based can again be separated into *active* and *passive* monitoring techniques.

2.1 Active Monitoring Techniques

To monitor traffic using active monitoring techniques, an active communication between not less than two points (sender/ recipient) is needed. For measurement issues, when using active monitoring, packets need to be inserted actively into the network. Perhaps the best known active monitoring techniques are *ping*, *traceroute* and *iperf* [6]. All techniques are dealing with availability, routes, packet inter-arrival jitter, packet delays, packet losses or bandwidth measurements [6]. They are called activity monitoring techniques, because using the ping example, the sender needs to actively send ICMP echo requests to an endpoint and waiting for the response.

2.2 Passive Monitoring Techniques

Passive monitoring, on the other hand, does not create additional traffic to the network. It simply listens to the traffic and collects information about packet rates/timings and inter-arrival timings [6]. At the end of a day, the administrators need to handle a huge amount of collected information. Packet sniffing is a good example how to perform passive monitoring. The drawback behind this monitoring technique is, that it can only be performed off-line.

Because active monitoring does inject to much overhead into the network and passive monitoring can only be done off-line, there are also combinational monitoring techniques possible, such as WREN [11] and SCNM [7].

2.3 Anonymous Communication Systems

Anonymous communication systems are designed to help people stay unrecognizable while surfing designated web sites. It is a privacy concern, when someone do not want to get profiled by a random website [16]. Tor[8] is a popular example of such anonymous communication systems, to address such privacy concerns.

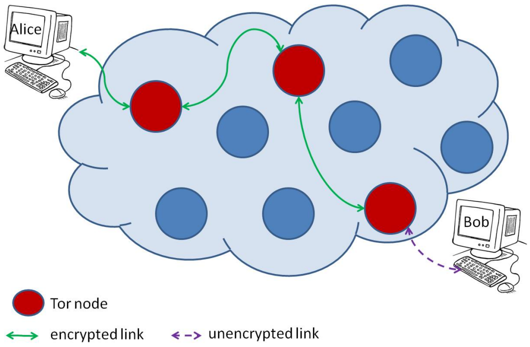


Figure 1: Anonymous Communication System (adopted by [8])

An anonymous communication system should have these three desirable features to ensure anonymity: it should provide sender anonymity, receiver anonymity and unlinkability of sender and receiver [16][12]. Sender and receiver anonymity simply means, that it cannot be identified who is communicating. Unlinkability of sender and receiver means, that even if the identity of both is known, the connection between them should be hidden [16].
Based on Tor, the functionality behind those systems should be described roughly (see Figure 1): Alice wants to communicate with Bob, but Alice wants to stay anonymous. Instead of establishing a direct connection between Alice and Bob, Alice installs an Onion proxy on her computer, which establishes a connection over three randomly chosen with Tor nodes. Between two nodes a tunnel is established using the public key of the communication node. The message travels over this tunnel encrypted. For each connection, a new random walk is chosen by the software. At no step, it can not be discovered where the traffic came from, and where it has been relayed to. Bob receives the message from Alice, but thinks that the message came from the last communicating node.

2.4 Flow Transformations

Flow transformations are applied to network flows to make them unrecognizable in order to achieve non-correlated flows. There are a few techniques, that can be applied to flows, to get rid of identifying characteristics. These techniques can widely be separated into *intra-flow transformations* and *inter-flow transformations* [16]. The former ones, are based on flow transformation within one flow without involving additional flows. The second ones produce transformation on flows by adding further unrelated flows.

2.4.1 Intra-Flow Transformation

Basically within one flow, following transformations can be applied (see Figure 2): adding chaff, packet dropping (also de-chaff) and repacketization (packet merge and fragmentation)[16]. Chaff is any cover-traffic within an anonymous system. Packet dropping can be enforced to make a flow unrecognizable. Repacketization can be done by combining packets, or by splitting a packet. Packet dropping and

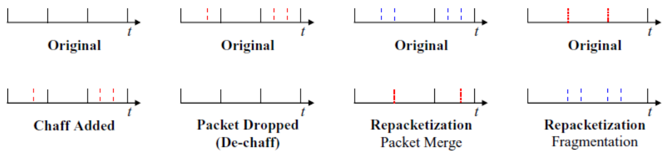


Figure 2: intra-flow transformations: adding chaff, packet dropping and repacketization (packet merging and fragmentation) (adopted by [16])

repacketization can be done intentionally, but also can happen naturally as for example by using SSH. [16].

2.4.2 *Inter-Flow Transformation*

Here, transformations are applied that include: flow mixing, flow splitting and flow merging [16]. Thereby it is important to notice, that a flow is mixed/splitted or merged with unrelated flows (see Figure 3) in contrast to intra-flow transformation where a flow was transformed within one flow without involving additional flows. As can be seen in figure 3, flow mixing mixes a random flow with unrelated flows. However flow merging combines a flow with flows that belong to the same network information flow [16].

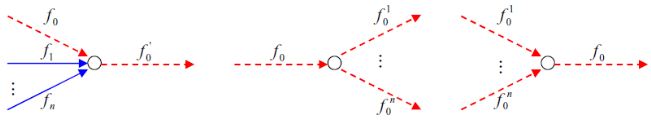
Such flow transformations occure in anonymous communication systems to change a flow to an unrelated one. As the presented flow transformation can be applied arbitrary often, it has been believed, that the produced flows are indistinguishable. However, with the use of watermarking, correlated flows, even if they are distorted like that, can be found.

2.5 Stepping Stones

Beside anonymous communication systems, stepping stones are a popular technique to conceal an atteckers identity. The idea is simple: instead of using the real computer for attacks, the attacker can connect through a sequence of intermediate hosts, which were compromised earlier. This example is from [17] and describes, how stepping stones can be applied: consider an attacker at host A, who can use SSH to login into B. B is now the stepping stone, if the attacker plans to start an attack on C, which he will do of course from B. Here comes the crucial part: the two connections between A and B, and between B and C are correlated. They are basically the same, besides the fact that they have been forwarded at the point B. This is where CNFD applies. It is searching for correlated network flows to link them together and thus to identify the attacker. Notice that SSH has been used to encrypt the traffic, so content-based analysis would not work here. Since an attacker has the authority over the stepping stone, he can apply multiple flow transformations to make the flows look different (not correlated). Watermarks can identify such flows in stepping stones.

3. CORRELATION DETECTION

The rudimentary approach to detect similarities in flows, is by comparing two flows. The procedure is described below: [21]:



a) Flow Mixing
f_0 , $f_0^{'}$ belong to the same network information flow; f_1 … f_n do not belong to the same network information flow

b) Flow Split
f_0 , $f_0^1 \dots f_0^n$ all belong to the same network information flow

c) Flow Merge
f_0 , $f_0^1 \dots f_0^n$ all belong to the same network information flow

Figure 3: inter-flow transformations: flow mixing, flow splitting and flow merging (adopted by [16])

1. Data Collection
2. Distance Function Selection
3. Flow Correlation

To detect correlated flows, information needs to be collected (e.g. arrival time using tcpdump or NetFlow [1]) about the incoming and outgoing flows. Then those arrival times need to be compared. The arriving times form a series with $\mathrm{A}_i = (\mathrm{a}_{i,1}, ..., \mathrm{a}_{i,n})$ at the input and $\mathrm{B}_i = (\mathrm{b}_{i,1}, ..., \mathrm{b}_{i,n})$ [21]. The similarities between those flows, are measured, by applying distance functions. It can be assumed, that the smaller this distance is, the more similar the packets are. This is the most important part of flow correlation detection and can be done in different ways (depending on the technology which is used to determine the similarities between two flows). The last step simply takes those flows with the minimum distance and identify them as correlated.
Generally, network flows correlation can depend on three characteristics [17]:

- host activity, which records every user login.
- And/or connection content, for example packet payload.
- And/or connection timing, that is the arrival and departure time of each packet.

CNFD can address one or more of these points.

3.1 Correlation Detection Based On Host Activity

This technique monitors the logins of an user on stepping stones. It is an passive traffic monitoring technique.
If the login information of e.g. 5 hosts is known, it is not that difficult to determine, whether there is a correlation or not. As described above, an attacker knows this problem and would try to manipulate the logins. The funny thing is, that he has the authority to do that, because every stepping stone used by an intruder, has been compromised earlier by him. As soon, as an attacker has the authority over an host, he/she can manipulate the login information.
The best known representatives for this technique are DIDS [14] and CIS [10]. Distributed Intrusion Detection System (DIDS) is the oldest approach, first published in 1991 in [14]. It is a network wide intrusion detection system with a centralized DIDS director and monitored hosts in the DIDS

domain. Each host collects information about ingoing and outgoing flows and sends this information to the DIDS director for analysis. The system keeps tracks of all movements of the users in the DIDS domain, concerning all TCP connection in this domain. Caller identification system (CIS) is aimed to authenticate an users identity. If a user logs into several hosts, each hosts asks the previous one, where the user came from and receives a list of all visited hosts. The last host in the conenction chain, knows where the user came from originally. If an attack happens to the last host, the users identity can be tracked back to the first host in the connection chain, and the attacker can be verified.

The host based approach is based upon trust of the monitored hosts. If one host is compromised, the whole idea behind host based approach fails. As mentioned before, an attacker does have authority over the hosts, so he can easily manipulate them.

There is also another technique known, which uses the host-based approach for detecting correlated networks, but this technique should not be applied because it's illegal. The US Air Force used this technique to trace intruders by breaking into the hosts the same way as the intruder did but this time backwards, applying the same techniques and methods as the intruder did. This technique is called Caller ID [19]. In contrast to DIDS and CIS, Caller ID is an active traffic analysis technique.

3.2 Correlation Detection Based On Connection Content

Connection Content, that is the payload of the each connection, is of course a good characteristic, and probable it is unique enough to identify correlated flows. But this is only possible, if the connection is not encrypted. In cases, where the connection is encrypted, the content does not reveal to much information. This approach can be neglected, as the flow is mostly encrypted. Encryption has the property, to create a completely different output, otherwise it would not be a good encryption algorithm. In addition, a good encryption algorithm creates a one-way function, that means that, given the output, it is computationally infeasible to determine the input. So, for correlation detection purpose, this approach is not helpful. Nevertheless there are techniques for correlation detection based on connection content in unencrypted traffic, like in Thumbprinting[15].

In Thumbprint a function is applied to the connection, which can distinguishes a given connection from all other ones but returns the same value over related connections. All participating hosts store this thumbprint (the unique value over a connection) and in case of an attack the stored thumbprints can be compared and related connections can be identified.

3.3 Correlation Detection Based On Connection timing

This approach is at present the most promising one. It takes the arriving and departure times of packets. The best known representatives are IPD-based [18] and ON/OFF-based [20] techniques.

The IPD-based approach takes the inter-packet arrival times of packets for correlation detection. These timings do not differ across the stepping stones [17]. In IPD-based, the timestampes of packets are measured and stored in a vectors.

Table 1: Overview Correlation Detection Approaches

	Passive	Active
Host-Based	DIDS, CIS	Caller ID
Content-Based	Thumbprinting	
Timing-Based	ON/OFF	IPD-Based

A correlation point function (CPF) compares two flows X and Y with their two timestamp vectors. If $max(CPF(X,Y))$ is greater than a threshold δ then the two flows X and Y can be considered related.

The ON/OFF-approach is based on ON and OFF periods of network traffic. The ON period starts, every time a packet appears on a network. It proceeds, until there are no packets traversing the network for at least T seconds, then the OFF period begins[20].

The reason, why ON/OFF periods are very interesting for correlation detection, is that it reveals keystroke interactions [20].

It has been discovered that keystroke inter arrivals produce always significant OFF periods. For example: 25% of interactive traffic arrives 500 msec or more appart, and 15% even 1 sec or more apart [20]. In other words: interactive traffic will always produce clear OFF periods. [17].

This approach has also an important advantage over the content based one. To detect similarities in the connection there is no need to know the content, but only the arriving and leaving time at a host. This method can thus be applied to encrypted traffic.

Of course an attacker can try to manipulate the timing by introducing delays. As described above, an attacker has the authority over stepping stones, and thus can change the timing-characteristics of packets. The result can be, that unrelated flows become suddenly related [17].

Watermarked-based techniques to detect correlated network flows are robust against those modifications on timing characteristics of packets and represent a new approach to CNFD.

4. WATERMARKS

Watermarks constitute a technique to recognize similarities in network flows, by using the timing-based approach on encrypted packets. In watermarking, a router "watermarks" a flow by adjusting the *timing-information* by applying delays of selected packets in a flow [17]. After the watermarking, the flow passes different distortions, described in 2.4, in a network. Finally, the watermarked flow arrives at a "detector", who knows the original flow and the shared secret parameters between the detector and the watermarker. The detector applies the same modification to the timing of the packets as the watermarker. If the resulting pattern is the same, the two flows can be considered correlated (see Figure 4).

Watermarking can achieve a detection rate by almost 100% and a pattern correlation by almost 0% [17]. This two rates are called true positive *tp* and false positive *fp* [17].

Compared to passive monitoring analysis, watermarking requires less computation and thus is more scalable. In passive techniques n incoming flows need to be compared with m outgoing flows to identify similarities. Therefore, $O(nm)$ computations are needed. Watermarking on the other hand only needs $O(n)$ computations and $O(1)$ for the shared key.

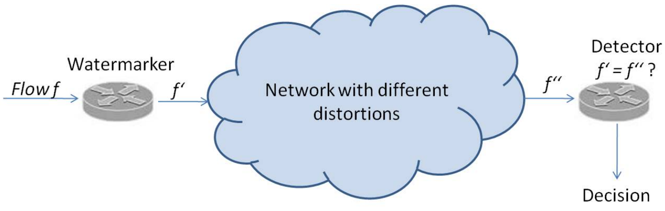


Figure 4: Network Flow Watermarking

As described above, intruders can modify the timing of packets in a flow. Here watermarks have an advantage over passive timing-based approaches, by being resistant against this kind of counter measurements by an attacker, as the decter would recognize a timing perturbation on the flow. An attacker can perhaps identify the watermarking pattern applied to the flow, but he does not know the secret parameters, and thus can't corrupt watermarks.

In the following, two techniques are described for correlation detection using watermarks: Interval-Based Watermarking Scheme and SWIRL. Both techniques work on intervals and are therefor robust to packet losses. In addition to that SWIRL is a invisible watermark because the insertion of watermark is not noticeable to outstanders. This makes SWIRL at the moment the most interesting approach in correlation detection with watermarks.
Interval based approaches divide the flow into T intervals and applie different patterns depending on the timing of the packets.

Interval-Based Watermarking Scheme

In Interval-based watermarking, the flow is devided into intervals and watermarking is done by manipulating the rate of the traffic in intervals. For watermarking, there are two options: *clearing* and *loading*. Clearing means, that an Interval I is cleared by delaying all packet from it. Loading means, that an interval is loaded by delaying all packets from the previous interval to the current.
Watermarking:
To insert Bit 0 in position i, the packets in interval I at position i are delayed and the next interval gets the packets from the previous one. To decode Bit 1 at position i, all packets from interval I at position *i-1* are delayed to the next interval (see Figure 5).
Detection:
The detector checks for existence of watermarks, as he knows the secret parameters such as the list of positions S and the interval lengths T.
Advantage:
This approach is robust to repacketization and losses.

SWIRL: Scalable Watermark that is Invisible and Resilient to packet Losses

[4]
As the title may suggest, this watermark approach can be applied to large scenarios, as it needs less computation and communication time. It is also invisible, because of small amount of distortion, that makes a multi flow attack impossible and it is resilient to packet losses [4]. **Watermarking:** First, a flow is divided into a set of intervals of length T. For

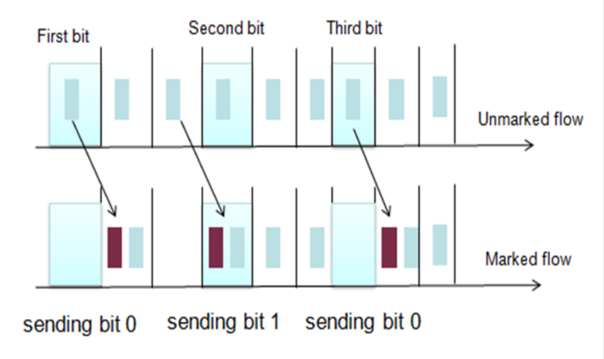


Figure 5: Interval-Based Watermarking

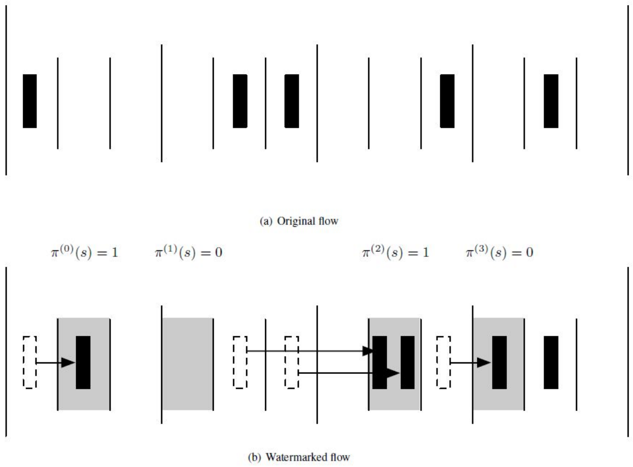
this kind of watermarking, there are two intervals needed: a mark and a base interval. It is completely irrelevant which one is the base and which one the mark intervals. As soon as they are determined, they are fixed for the whole flow. The base interval needs to come before the mark intervals, no other restrictions apply [4].
The mark interval is subdivided into r subintervals of length T/r [4]. Then the subintervals are again subdivided into m slots, which contain packets (or not) (see Figure 6) [4].
One of the secret parameters between the watermarker and the decoder is the permutation which is now applied. After applying the permutation, each packet is delayed, such that it falls into designated slots[4] (compare Figure 6). The grey slots are the result of the applied permutation. For the first subinterval that means, that all packets in the first subinterval should appear in slot 1 (that's the grey slot in subinterval 1). For the next slot, all packets should appear in 0 of subinterval 2, but there are no packets before this slot, so it remains empty. This is continued again, until all packets are in their designated slot.
Decoding:
The detector analyses the packets in the base interval, applies the permutation function and knows how the mark interval should look like. He then determines if the watermark is detected or not.
Advantage:
It could have been shown, that SWIRL can be applied to flows as short as 2 minutes with error rates in order of 0.000001 or less[4].
Table 1 compares the watermarking approaches according to robustness against losses and invisibility.

4.1 Applications

Stepping stones and anonymous communication systems are the particular applications of the presented correlation detection techniques, especially for watermarks. Following is described, how watermarks can be applied on both.

4.1.1 Anonymous Communication System

As a number of input flows enter the anonymous communication system, they are mapped to a number of output flows. But, how the flows are related is not known to outstanders. Tor is one example of such a system described in Section 2. The main objective of an attacker is to spy out, how the input and output flows are related. Watermarks in this case


(a) Original flow

(b) Watermarked flow

Figure 6: SWIRL (adopted by [4])

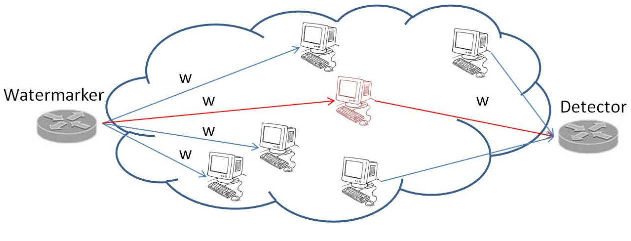


Figure 7: Stepping Stone Detection (adopted by [4])

are also called a privacy-invasive tool, because they can find out, which flows are related, because of the marks applied to incoming flows and spotted at outgoing flows.
An invader can detect such correlations by compromising an entry router in Tor (see Figure 1), then the flows are marked and detected on cooperating exit routers. Watermarking makes the attack much more efficient, since only *O(n)* instead of *O(nm)* computations (see Section 3) are needed, compared to other passive detection techniques.

4.1.2 Stepping Stones

Stepping stones were described earlier (see Section 2). The situation can be compared to the anonymous communication systems, because the incoming traffic has to be compared to the outgoing traffic. As shown in figure 10, the border routers are inserting watermarks on incoming flows, and the corresponding router is checking for watermarks on outgoing flows. Again, this can be done by passive traffic analysis but as stated before, watermarking gives a more efficient approach for detection.

5. CONCLUSION

This paper was intended to give an overview over correlation detection techniques, especially by using watermarks to detect similarities in network flows. Correlation detection is a traffic analysis method that can be applied for intrusion detection.
Correlations can be found in anonymous communication systems but also in stepping stones. The watermarking approach is based on introducing timing-delays to packets in a flow. Intrusion detection can be done in other ways, than by watermarking flows, but that is much more expensive and is very difficult to apply to large networks. Watermarking on the other hand gives a new approach on detecting correlated flows, as it is more scalable and produces less errors. By using interval-based watermarks, lower error rates are produced and they are not as vulnerable to packet droppings. The most promising one by now is SWIRL, because of its low error rates and high correlation detection. Furthermore it is invisible to attackers and can be applied to large networks.

6. REFERENCES

[1] Cisco systems inc. netflow services solutions guide.

[2] Rmon: Remote monitoring mibs.

[3] Remote monitoring, internetworking technologies handbook, 1992-2006.

[4] H. Amir and N. Borisov. Swirl: A scalable watermark to detect correlated network flows. 2011.

[5] A. Blum, D. Song, and S. Venkataraman. Detection of interactive stepping stones: Algorithms and confidence bounds. *Recent Advances in Intrusion Detection*, pages 258–277, 2004.

[6] A. Cecil. A summary of network traffic monitoring and analysis techniques. visited on May 15, 2011.

[7] A. Deb, G. J. Maria, J. Goujun, and T. Brian. An infrastructure for passive network monitoring of application data streams. *Proceedings of the 2003 Passive and Active Monitoring Workshop*, 2003.

[8] D. Dingledine, N. mathewson, and P. Syverson. Tor: The second generation onion router. *Proceedings of the 13th USENIX Security Symposium*, August 2000.

[9] D. Donoho, A. Flesia, U. Shankar, V. Paxson, J. Coit, and S. Staniford. Multiscale stepping-stone detection: detecting paris of jittered interactive streams by exploiting maximum tolerable delay. *International Symposium on Recent Advances in Intrusion Detection*, 2516:17–35, October 2002.

[10] H. Jung. Caller identification system in the internet environment. *Proceedings of 4th USENIX Security Symposium*, 1993.

[11] Z. Marcia and L. B. B. Using passive traces of application traffic in a network monitoring sytem". *IEEE Computer Society*, 2004.

[12] A. Pfitzmann and M. Waidner. Networks without user observability - design options. *Computer and Security*, 6 (2):158 – 166, 1987.

[13] J. Raymond. Traffic analysis: Protocols, attacks, design issues, and open problems. *Designing Privacy Enhancing Technologies*, pages 10–29, 2001.

[14] S. Snapp. Dids (distributed intrusion detection system) - motivation, architecture and early prototype. *Proceedings of 14th National Computer Security Conference*, 1991.

[15] S. Staniford-Chen and L. Heberlein. Holding intruders accountable on the internet. *Proceedings of the IEEE Symposium on Security and Privacy*, 1995.

[16] X. Wang, S. Chen, and S. Jajodia. Network flow watermarking attack on low-latency anonymous communication systems. *IEEE Symposium on Security and Privacy*, pages 116–130, 2007.

[17] X. Wang and D. Reeves. Robust correlation of encrypted attack traffic through stepping stones by manipulation of interpacket delays. *Proceedings of the 10th ACM conference on Computer and communication security*, page 20, 2003.

[18] X. Wang, D. Reeves, and S. Wu. Inter-packet delay-based correlatioin for tracing encrypted connections through stepping stones. *7th European Symposium on Research in Computer Security*, 2002.

[19] K. Yoda and H. Etoh. Finding a connection chain for tracing intruders. *6th European Symposium on Research in Computer Security*, 2000.

[20] Y. Zhang and V. Paxson. Detecting stepping stones. *Recent Advances in Intrusion Detection*, pages 258–277, 2004.

[21] Y. Zhu, X. Fu, B. Graham, R. Bettati, and W. Zhao. On flow correlation atttacks and countermeasures in mix networks. *Privacy Enhancing Technologies*, pages 207–225, 2005.