Active Timing Based Approach for Tracking Anonymous Peer-to-peer Network in VoIP

Karthikeyan.C, Karthikeyan.V, Jerin Sajeev.C.R, Merlin Moses.M

***Abstract*— Peer-to-peer VoIP calls are popular due to their low cost and convenience. When these calls are encrypted and anonymized the network becomes a secured one. Tracing of the anonymous VoIP call users are important and the traced information about them should be sent to the server to know how long the users are in communication.**

The key challenge in tracking encrypted VoIP calls across anonymous communication system is to identify the correlation between the VoIP flows of the caller and the callee. Since all the traffic of the peer-to-peer VoIP calls are encrypted, the best way to track anonymous VoIP calls across the internet is using the Active timing based correlation. It is done by embedding a unique watermark into the inter-packet timing domain. The analysis shows that it only takes several milliseconds time adjustment to make normal VoIP flows highly unique and the embedded delay value could be preserved across the low latency anonymizing network. In this proposal, tracking of anonymous VoIP calls across internet was successfully achieved by using active time based correlation method and the results demonstrate that tracing of anonymous peer-to-peer VoIP calls on the internet is feasible and low latency anonymizing networks are susceptible to timing attacks.

I. INTRODUCTION

In Voice over Internet Protocol (VoIP) the transmission of voice over traditional packet-switched IP networks is one of the hottest trends in telecommunications. The term Voice over Internet Protocol is typically associated with equipment that lets users dial telephone numbers and communicate with parties on the other end which is also provided with VoIP system or a traditional analog telephone. Security administrators might assume that digitized voice travels in packets; they can simply plug VoIP components into their already secured networks and get a stable and secure voice network. Unfortunately, many of the tools used to safeguard existing systems such as computer networks firewalls, network address translation (NAT), and encryption for controller area network (CAN) are not as effective as VoIP network. Although most VoIP components have counterparts in data networks, to provide better performance in VoIP applications, it must be supported by supplement ordinary network software and hardware with special VoIP components. Integrating a VoIP system into an already congested or overburdened network can be disastrous for an industries network infrastructure. Some of the challenges of introducing appropriate security measures for VoIP in an enterprise.

Communication between two users are said to be Anonymity, when the interaction between them is not being identifiable within the network. If it is not identifiable, then users are said to be Anonymous.The main objective of this project is to investigate practical techniques for the effective tracking of anonymous voice calls on the Internet and identify the weakness of some of the currently deployed anonymous communication systems.

The traced voice users are identified and send as a report to server to determine how long the users are in communication.

Manuscript received on January, 2013
Karthikeyan.C, ECE, Einstein College of Engineering, Tirunelveli.
Karthikeyan.V, ECE, Einstein College of Engineering, Tirunelveli.
Jerin Sajeev.C.R, ECE, Einstein College of Engineering, Tirunelveli.
Merlin Moses.M, ECE, Einstein College of Engineering, Tirunelveli.

II. RELATED WORK

Emanuel proposed and evaluated a detection system based on a HTTP workload model to efficiently detect VoIP calls hidden in Web traffic (2). Two attacking methods watermark attack, and complementary matching attacks on VoIP are proposed and analysed by Ge zhang (3). The main drawback of this work not considers the protection on signalling layer. Tor is the second generation Onion Router, supporting the anonymous transport of TCP streams over the Internet. Its low latency makes it very suitable for common tasks, such as web browsing, but insecure against traffic analysis attacks by a global passive adversary (7). Thamizharasi mainly focusing about how to improve privacy in peer to peer VoIP network without the involvement of third party. By creating the personal network which connects many systems through the LAN or wifi and allows users to communicate directly without any third party server in between.(1)

III. METHODOLOGY

First, we demonstrate that a previously-proposed passive, timing-based correlation scheme is vulnerable to random timing perturbation. Second, we develop a practical watermark-based correlation scheme that is much more robust in the presence of random timing perturbations. Our experimental results show that the new method consistently has a higher detection rate of active users, whether there is random timing perturbation or not. Lastly, we develop accurate models of the tradeoffs between the desired watermark correlation true positive rate (and false positive rate) and the watermark embedding parameters, as well as the defining characteristics of the random timing perturbation.
The quantitative expression of the tradeoffs is of significant practical importance in optimizing the overall correlation effectiveness under a range of conditions.

A. Connection Correlation Techniques

Tracing interactive traffic through stepping stones requires the discovery of an association between two connections at the stepping stone, such that these connections act as consecutive flows in a chain of connections.

Generally, connection characteristics which remain unchanged (*i.e.*, invariant) at the intermediate host are used to determine whether two connections are correlated. The traffic can be then traced back to the origin by linking the correlated connections together. So far, substantial research has been done on such connection correlation techniques.

Following the early work based on packet payloads, techniques based on traffic timing were proposed to deal with

encrypted connections. In these techniques, the timing characteristic of an interactive connection was assumed to be unique and preserved across stepping stones such that it can be utilized for correlating connections. Faced with the potential for an adversary to use active timing perturbation at the stepping stones in an attempt to defeat timing-based correlation, more sophisticated timing-based techniques were proposed later. They analyzed and compared the long-term behavior and packet counts of the connections, assuming that the adversary's delays are bounded.

As opposed to the above passive approaches active timing-based correlation techniques those are robust against random timing perturbation. Their method embeds unique watermarks into connection flows, by actively manipulating packet timing of the flows, such that these watermarks can be identified when correlating connections. These active approaches have shown to be very successful for traffic tracing.

B. Comparison of Tracing between PSTN Call and Voip Call

The goal of tracing anonymous VoIP is to effectively identify the caller and the callee of a particular VoIP call even if it is anonymized. Here we focused on the technical feasibility on tracing anonymous peer-to-peer VoIP calls

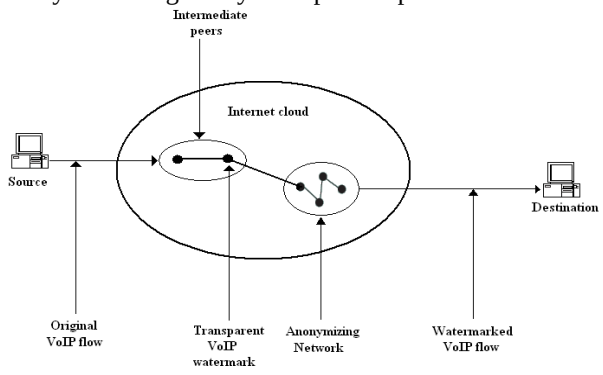

Fig. Tracking anonymous P2P VoIP calls across the internet

The telephone switch has all the call-identifying information of every call it has set up, and it is technically straightforward for the telephone switch to track any calls it has set up. In fact, existing call-identification of traditional PSTN calls is based on the signaling protocols that set up the calls.

While signaling protocol based call-identification works fine for PSTN calls, it is difficult to be applied directly to VoIP calls due to the following reasons are the signaling protocols for VoIP calls are evolving, there are currently multiple competing VoIP signaling protocols and some of them are proprietary. In order to effectively identify VoIP calls, new forms of call-identifying information is needed and no matter what signaling protocols are used to set up the call. All the call-identifying information of a managed VoIP call is available to the service provider who sets up and manages the VoIP call.

In this project we only consider how the unmanaged VoIP calls could be effectively traced. Specifically, we consider using the VoIP flow itself, rather than the VoIP signaling, to uniquely identify the caller and callee of peer-to-peer VoIP calls.

C. Timing based correlation methods

Inter-packet delay of a user is obtained from each node. The values of inter packet delays for different users is used to correlate the received packets. Based on time delay values the users has been different and related. These types of approach are called as timing based correlation. A number of timing based correlation methods have been proposed, and they can be classified into two categories, namely passive timing based correlation and active timing based correlation.

1. Passive Timing based correlation

Passive timing based approach correlates the encrypted flows based on passive comparison of their timing characteristics. It is shown to be effective when the timing characteristics of each flow are unique enough. Passive timing base correlation uses only few common VoIP packetization intervals. (i.e. 20ms or 30ms). Therefore, passively comparing the timing characteristics of VoIP flows will not be able to distinguish different VoIP flows.

2. Active Timing based correlation

Active timing based correlation method is done by embedding a unique watermark into the inter-packet timing domain of the interactive flow through deliberate timing adjustment of selected packets. Active method has the potential to differentiate flows with very similar timing characteristics. However, the method proposed can not be directly used to correlate VoIP flows due to the following three reasons.

The first one is the VoIP traffic has stringent real-time constraints, and the total end to end delay should be less than 150ms.the second reason is the inter-packet arrival time of VoIP flows is very short (i.e. 20ms or 30ms). This requires any time adjustment on VoIP packet to be very precise and small. Finally, the watermarking is based on the quantization of averaged Inter-Packet Delays (IPD), and it requires packet buffering in order to achieve the even timing adjustment over different packets. The required buffering would be too long for the real-time VoIP flows. To correlate anonymous VoIP flows with similar inter-packet timing characteristics; we use an active approach to make the VoIP flows more unique.

To address the limitations of previous work, we use a new parameter scheme that is suited for tracking anonymous VoIP traffic in real-time. The real time tracking of anonymous users is predicted with the help of inter-packet delay values which was transmitted on a packet as additional information. The key challenge in tracking anonymous VoIP calls by the active approach helps to retrieve the packet timing without destination buffering and guarantee the even time adjustment of those selected packets.

D. Estimation of inter-packet delay value in intermediate node

Randomly choose packets P1, P2, P3 and P4 from particular VoIP flow with its arrival time stamp t1, t2, t3 and t4 respectively. Group these four packets into two pairs (P1, P2) and (P3, P4).Then the inter packet delays (IPD) of these two pairs are obtained and denoted as IPD1 and IPD2.

Calculate the normalized difference between the two IPDs as IPDD = (IPD2 - IPD1) / 2 (1)

IPDD could be positive or negative .Because P1, P2, P3 and P4 are selected randomly, the distribution of IPDD is symmetric and centered around 0. To embed a binary bit '1', we deliberately delay the departure time of packets P1 and P4 for a period of time 'a'. This would effectively shift the

distribution of the original IPDD to the right by amount 'a', which means IPDD more likely to be positive than to be negative.

Similarly, binary bit '0' can be embedded by shifting the distribution of original IPDD to the left by amount 'a'. This can be achieved by deliberately delay the departure time of packets P2 and P3 for a period of time 'a'. To decode the embedded watermarking bit, we simply use the same randomly selected packets and compute the corresponding IPDD. If the IPDD is less then or equal to 0, we would get decoded bit '0'; if the IPDD is greater than 0, we would get decoded '1'.

E. Introduction of delay box in correlation node

Delay box is an ns node that should be placed in between the source and destination nodes. With delay box, packets from a UDP flow can be delayed, dropped, and forced through a bottleneck link before being passed on to the next node. A distribution can be used to specify delay, loss, and bottleneck link speed for a source to destination pair. Each flow between that source to destination pair draws from the distribution to determine its characteristics.

1. Implementation of delay box

The implementation of delay box maintains two tables. They are rule table and flow table. Entries in the rule table are added by the user in the Tcl simulation script and give an outline of how flows from a source to a destination should be treated. The fields are source, destination, delay Random Variable (in ms), and loss rate random variable.

Entries in the flow table are created internally and specify exactly how each flow should be handled. Its values are obtained by sampling from the distributions given in the rule table. The fields are source, destination, flow ID, delay, and loss. The active timing based approach is performed by embedding a delay value to the intermediate node which is linked to a delay box and this combination act as a single node which produce a delay value of about 20ms. Thus encoding is performed by adding a delay value to the correlated packets from a particular node and decoding can be performed by subtracting the known delay value at the destination. Thus active timing based approach is more efficient to trace the anonymous user since the delay value added to the correlated packets from the intermediate node is less than 20-30 ms.

IV. SIMULATION RESULTS

In order to be able to watermark any VoIP flows transparently, it is desirable to have a VoIP gateway which forwards the VoIP flows and watermarks any specified bypassing VoIP flows with specified watermarks. To embed the watermark into the inter-packet timing of a VoIP flow, some capability is needed to delay specified packet of specified flow for specified duration. The kernel of the Linux operating system will provide such capabilities.

To achieve packet delay of 20ms or 30ms, the operating system must provide a hard real-time scheduling capability. The Real Time Application Interface (RTAI) of Linux will provide a hard real-time scheduling capability. The following features of RTAI have made it an attractive platform for implementing the high precision packet delay capability. They are the hard real-time scheduling functions introduced by the RTAI coexist with all the original Linux kernel services. This makes it possible to leverage existing Linux kernel services, especially the IP stack components; from within the real-time task. The RTAI guarantees the execution time of real-time tasks regardless of the current load of non real-time tasks. The RTAI supports high precision software timer with the resolution of microseconds.

A. Simulation of peer to peer VoIP calls using NS-2

Network Simulator (NS-2) is a networking tool that allows a reliable VoIP user-level performance analysis to be carried out through simulation. Using NS-2 a network topology has been created with multiple source nodes, multiple destination nodes linked through a peer-to-peer network and the extreme end of the peer-to-peer network act as an intermediate node. The traffic analysis is viewed in this intermediate node.

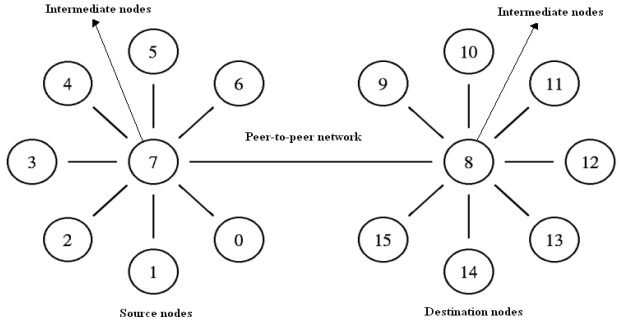

Fig. (a). A model of network topology using NS-2

Simulation of different types of voice traffic has been tried Dumbbell network topology has selected and in this VoIP traffic was defined for connected user. This network topology provides multi-source node user to multi-destination node user communication via peer-to-peer network.

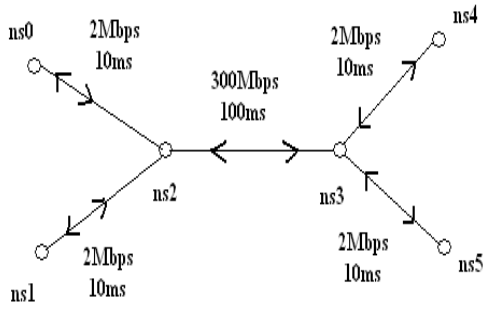

Fig. (b). Initialization of source nodes, intermediate nodes and termination of destination nodes

B. Steps involved in Simulating VoIP traffic flow

The above model of network topology is created using NS-2 by the following steps

Step1: Initialization of source nodes, intermediate nodes and termination of destination nodes

Step2: Defining nodes, links and assigning queue size

Step3: Agents and applications is performed by setup a CBR over UDP connection

Step4: Scheduling events

The Tcl script in NS-2 defines when event should occur.

The scheduler is started when running ns

Step5: The created network topology has been viewed by the network animator which shows the different voice packet transmitting among peer-to-peer nodes.

Step6: Tracing object in a simple link.

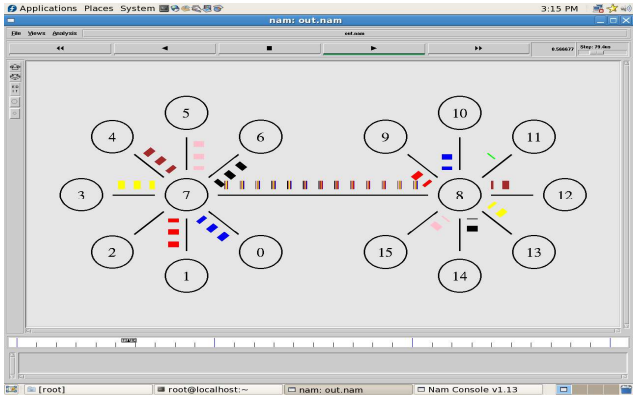

Fig.(c). Multi-user voice packets transmitting among P2P nodes

C. *Output of packet correlation technique*

The inter-packet delay of the different packets from a particular node is grouped to obtain a trace file .From the resultant trace file xgraph is plotted between received time and node. The figure 4.4 shows that the packets are transmitted from source node N0 to node N7 and also the transmission of intermediate nodes Node 7 and Node 8.

The high transmission rate of source node N0 to node N7 is shown in the figure. The high transmission rate is due to direct connections of multi-source users to the intermediate peer node N7.The transmission rate from node N7 to intermediate node N8 shows slightly decreased because all multi source users are started to transmit the packets.

After some milliseconds (i.e. 10ms) the second packet is transmitted from the same node with same characteristics as the first packet. Similarly all packets transmitted from the source node N0 will have the same characteristics with some milliseconds (i.e. 10ms) delay. Thus the graph proves that the inter-packet timing delay between each packet of single user at the intermediate node N8 is same. Thus the correlated outputs are obtained and tabulated for each source node to destination node which is communicating via the intermediate node.

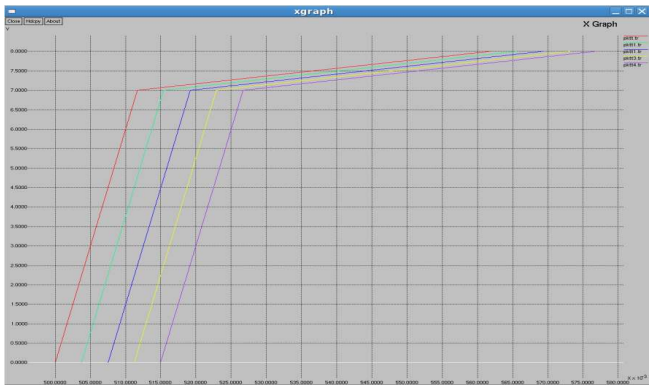

Fig.(d). X graph between simulation time and number of nodes.

Table 4.1 Correlated output obtained from Xgraph

The packets transmitted from source node N0 to destination node will have the correlated output of about 0.061848 ms between each packet at the extreme end of the intermediate node N8.

D. *Results of active timing based approach*

The active timing based approach is performed by embedding a delay value as watermarking to an intermediate node which is linked to a delay box and this combination act as a single node which produce a delay value of about 20ms.

Source Nodes	Intermediate Node	Destination Nodes	Correlated Output (ms)
N0	N8	N10	0.061848
N1	N8	N9	0.062016
N2	N8	N11	0.060352
N3	N8	N13	0.062184
N4	N8	N12	0.0662352
N5	N8	N15	0.062520
N6	N8	N14	0.062688

Each user inter packet delay has been estimated at the intermediate node of destination. The inter-packet delay value has to embed on the packet from intermediate node onwards. The embedded delay value helps the destination to identify the anonymous voice caller who is communicating for the period. To embed the delay value TCL script is written and linked to the intermediate node of the network topology. After simulating the TCL script the delay value for corresponding packets of the intermediate node are randomly generated and stored in the output delays to the file db0.out.

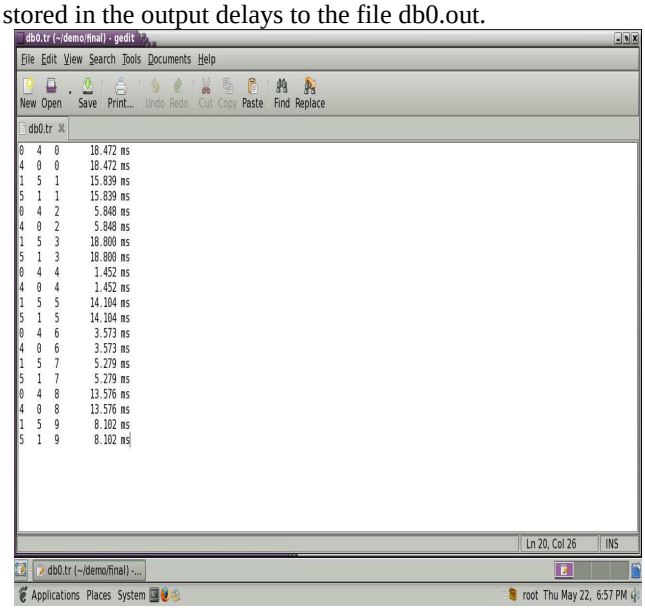

Fig.(e). Delay values of packets transmitted between intermediate node and destination node

The above figure 4.5 shows the delay value for corresponding packets transmitted between intermediate node and destination node. It also shows that the delay values are randomly generated to corresponding packets of four users. The above trace file shows that the first column indicates the transmitter node, the second column indicates the receiver node, the next column indicates packets transmitted in between transmitter node and receiver node and the final column indicates the delay values.

V. CONCLUSION

Anonymity provides high privacy and secrecy among VoIP uses. It also provides routing of their packets to the corresponding destination. Currently VoIP demands for peer-to-peer anonymous environment to route the packets directly to the destination. This helps all users free to select their corresponding connections to the other end. In anonymous peer-to-peer VoIP network, the interactions between two users are not being identifiable within the network. But how long the user was being in connection is send as a report to the server. Actually the status of the user

connection is only sent as a report to the server and the status has been framed by tracing the active connectivity of anonymous users.

To trace the active connection of users two types of methods are available. They are active time based method and passive time based method. In our proposed work, we have implemented the simulation for peer-to-peer network topology presents the selection of paths between multiple source nodes and multiple destination nodes. The simulation results of the figure 5.1 shows that the node N0 and node N7 has been selected as a correlation nodes. In this network topology, node N0 and node N7 is doing the calculation of inter-packet delay estimation and based on the inter-packet delay from each user the user is differentiated at the correlation node (peer node N7). From the correlation peer node the estimated inter-packet delay is added to the packets as an additional bit and then transmitted to the destination. Now the destination can able to understand the proper source being connected with it. From the simulated result we have observed and tabulated-table 5.1 the inter-packet delay values measured at the peer node. Hence the simulation succeeds the tracing of anonymous peer-to-peer voice call.

REFERENCES

[1] A.Thamizharasi and M.Vanitha' "Privacy and Packet Dispersion of Voice Applications in P2p Networks-VoIP", International Journal of Computer Applications (0975 – 8887),Volume 43– No.12, April 2012

[2] Ge Zhang and Simone Fischer-Hubner, "Peer-to-Peer VoIP Communications Using Anonymisation Overlay Networks" , CMS 2010, LNCS 6109, IFIP International Federation for Information Processing , pp. 130–141, 2010.

[3] Emanuel P. Freire, Artur Ziviani, and Ronaldo M. Salles, "Detecting VoIP Calls Hidden in Web Traffic", IEEE transactions on network and service management, Vol. 5, No. 4, December 2008

[4] Shiping Chen, Xinyuan Wang, and Sushil Jajodia, George Mason University. "On the Anonymity and Traceability of Peer-to-Peer VoIP Calls" IEEE Network, September/October 2006 page (32-37)

[5] Overlier, L., and Syverson. P. "Locating hidden servers". In Proceedings of the 2006 IEEE Symposium on Security and Privacy (May 2006), IEEE CS.

[6] T. Kohno, A. Broido and K. Claffy. "Remote Physical Device Fingerprinting". In Proceedings of the 2005 IEEE Symposium on Security and Privacy, IEEE, 2005.

[7] S. J. Murdoch and G. Danezis. "Low-Cost Traffic Analysis of Tor". In Proceedings of the 2005 IEEE Symposium on Security and Privacy, IEEE, 2005.

[8] X. Wang, S. Chen, and S. Jajodia, "Tracking Anonymous Peer-to-Peer VOIP Calls on the Internet," Proc. 12th ACM Conf. Comp. and Commun. Sec., Nov. 2005, Page. 81–91.

[9] S. A. Baset and H. Schulzrinne. An Analysis of the Skype Peer-to-Peer Internet Telephony Protocol. Columbia Technical Report CUCS-039-04, December 2004

[10] J. Li, M. Sung, J. Xu and L. Li. "Large Scale IP Trace back in High-Speed Internet: Practical Techniques and Theoretical Foundation". In Proceedings of the 2004 IEEE Symposium on Security and Privacy, IEEE, 2004.

[11] G. Danezis, R. Dingledine, and N. Mathewson, Mixminion: "Design of a Type Anonymous Remailer Protocol," Proc. IEEE Symp. Sec. and Privacy, May 2003, Page. 2–15.

[12] M. J. Freedman and R. Morris. "Tarzan: A Peer-to-Peer Anonymizing Network Layer". In Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS 2002), pages 193{206. ACM, November 2003

[13] Matthew Wright, Micah Adler, Brian Neil Levine, and Clay Shields. "Defending anonymous communication against passive logging attacks". In Proceedings of the 2003 IEEE Symposium on Security and Privacy, May 2003.

[14] X. Wang and D. Reeves. "Robust Correlation of Encrypted Attack Traffic Through Stepping Stones by Manipulation of Interpacket Delays". In Proceedings of the 10th ACM Conference on Computer and Communications Security (CCS 2003), pages 20-29. ACM, October 2003.

[15] Marc Rennhard and Bernhard Plattner. Introducing MorphMix: "Peer-to-Peer based Anonymous Internet Usage with Collusion Detection". In Proceedings of the Workshop on Privacy in the Electronic Society (WPES 2002), Washington, DC, USA, November 2002.