Journal – ICON (Integrating Concepts) ISSN 2456-6071

Analytical Data Processing Model to Reveal the Hidden Traffic Patterns in MANETs

Deva Prasad D[#1], P.Murugan[*2]
[#]Assistant Professor, Dept of CSE, Sivaji College of Engineering and Technology,
[*]Associate Professor, Dept of CSE, PSN Engineering College

Abstract*: A Mobile Ad hoc Network (MANET) is a continuously self-configuring, infrastructure-less network of mobile devices connected without wires. The primary challenge in building a MANET is equipping each device to continuously maintain the information required to properly route traffic. MANET is vulnerable to the attacks such as Malicious code, Repudiation in application layer, session hijacking in transport layer etc. STARS is the first Statistical Traffic Analysis approach. STARS cannot globally monitor the traffic across the entire network region. Hence, in our proposed work, sensors (signal detectors) are deployed around some particular mobile nodes to track their movements and eavesdrop all of their traffic. These sensors may even move accordingly. This method is called as variant of STARS or the Generalized STARS (GSTARS). To perform GSTARS, the adversaries only need to monitor the nodes beside the boundaries of the supernodes. The traffic inside each supernode can be ignored, since it will not affect the inter-region traffic patterns. In addition, GSTARS does not need the signal detectors to be able to precisely locate the signal source. GSTARS uses DSR protocol for route exploration. *Dynamic Source Routing Protocol (DSR) creates a route on **demand using source routing protocol. This protocol floods a route request message in the network to establish a route and it consists of two procedures: Route Discovery and Route Maintenance. The advantage of this protocol are 1) Aware of existence of alternative paths that helps to find another path in case of node or link failure. 2) It avoids routing loops and 3) less maintenance overhead cost as it an on-demand routing protocol.**

***Index Terms*—STARS, Dynamic Source Routing Protocol, Mobile Adhoc Network, Route Discovery, Route Maintenance.**

I. INTRODUCTION

Mobile ad hoc network (MANET) is a continuously self-configuring, infrastructure-less network of mobile devices connected without wires. Each device in a MANET is free to move independently in any direction, and will therefore change its links to other devices frequently. Each must forward traffic unrelated to its own use, and therefore be a router. The primary challenge in building a MANET is equipping each device to continuously maintain the information required to properly route traffic. Such networks may operate by themselves or may be connected to the larger Internet. They may contain one or multiple and different transceivers between nodes. This results in a highly dynamic, autonomous topology.

MANETs are a kind of Wireless ad hoc network that usually has a routable networking environment on top of a Link Layer ad hoc network. MANETs consist of a peer-to-peer, self-forming, self-healing network in contrast to a mesh network has a central controller (to determine, optimize, and distribute the routing table). MANETs circa 2000-2015 typically communicate at radio frequencies (30 MHz - 5 GHz).

A router is a networking device, commonly specialized hardware, that forwards data packets between computer networks. This creates an overlay internetwork, as a router is connected to two or more data lines from different networks. When a data packet comes in one of the lines, the router reads the address information in the packet to determine its ultimate destination. Then, using information in its routing table or routing policy, it directs the packet to the next network on its journey. Routers perform the "traffic directing" functions on the Internet. A data packet is typically forwarded from one router to another through the networksthat constitutes the internetwork until it reaches its destination node.

In computer networking, Point-to-Point Protocol (PPP) is a data linkprotocol used to establish a direct connection between two nodes. It can provide connection authentication, transmission encryption (using ECP, RFC 1968), and compression.

PPP is used over many types of physical networks including serial cable, phone line, trunk line, cellular telephone, specialized radio links, and fiber optic links such as SONET. PPP is also used over Internet access connections. Internet service providers (ISPs) have used PPP for customer dial-up access to the Internet, since IP packets cannot be transmitted over a modem line on their own, without some data link protocol. PPP is commonly used as a data link layer protocol for connection over synchronous and asynchronous circuits.

Passive network security attacks are in the nature of monitoring, or eavesdropping of transmissions of many types. The goal of this attack or the hacker doing the attack is to gain information or the information that is being transmitted in the message to gain a edge on the other party.

An active attack on a communications system is one in which the attacker changes the communication. He may create, forge, alter, replace, block or reroute messages. This contrasts with a passive attack in which the attacker only eavesdrops; A packet analyzer (also known as a network analyzer, protocol analyzer or packet sniffer, or for particular types of networks, an Ethernet sniffer or wireless sniffer) is a computer program or a piece of computer hardware that can intercept and log traffic passing over a digital network or part of a network. As data streams flow across the network, the sniffer captures each packet and, if needed, decodes the packet's raw data, showing the values of various fields in the packet, and analyzes its content according to the appropriate RFC or other specifications.

STARS includes two major steps: 1) Construct point-to-point traffic matrices using the time-slicing technique, and then derive the end-to-end traffic matrix with a set of traffic filtering rules; and 2) Apply a heuristic approach to identify the actual source and destination nodes, and then correlate the source nodes with their corresponding destinations. The contribution of STARS is twofold: 1) To the best of our knowledge, STARS is the first statistical traffic analysis approach that considers the salient characteristics of MANETs: the broadcasting, ad hoc, and mobile nature; and 2) most of the previous approaches are partial attacks in the sense that they either only try to identify the source (or destination) nodes or to find out the corresponding destination (source) nodes for given particular source (destination) nodes.

II. Related Work

ANODR (ANonymousOn Demand Routing) has already been devised as the anonymous routing scheme that is compliant with the design principles. ANODR[1] is more efficient and effective for routing in MANET.

A novel anonymous ondemand routing protocol, termed MASK, accomplishes both MAC-layer and network-layer communications without disclosing real IDs of the participating nodes under a rather strong adversary model. MASK offers the anonymity of senders, receivers, and sender-receiver relationships in addition to node unlocatability and untrackability and end-to-end flow untraceability. It is also resistant to a wide range of attacks. Moreover, MASK preserves the high routing efficiency as compared to previous proposals[2].

On-demand Lightweight Anonymous Routing (OLAR) scheme, applying the secret sharing scheme is based on the properties of polynomial interpolation. OLAR is an identity-free routing scheme, which provides source and destination anonymity, end-to-end communication relation anonymity, as well as route anonymity. In addition, this scheme highly decreases the overhead of data transmission, while making packets more untraceable compared to the previous solutions[3].

Another novel distributed routing protocol[4] which guarantees security, anonymity and high reliability of the established route in a hostile environment, such as ad hoc wireless network, by encrypting routing packet header and abstaining from using unreliable intermediate node. The major objective of this protocol is to allow trustworthy intermediate nodes to participate in the path construction protocol without jeopardizing the anonymity of the communicating nodes.

ARM: Anonymous Routing Protocol for Mobile Ad hoc Networks[5] for wired networks such as the Internet often cannot be applied to mobile ad hoc networks (MANETs). It provides anonymity in a stronger adversary model.

Anonymous Connections and Onion Routing developed anonymous connections that are strongly resistant to both eavesdropping and traffic analysis. Onion routing's[6] anonymous connections are bidirectional and near real-time, and can be used anywhere a socket connection can be used. Any identifying information must be in the data stream carried over an anonymous connection. An onion is a data structure that is treated as the destination address by onion routers; thus, it is used to establish an anonymous connection. Onions themselves appear differently to each onion router as well as to network observers. The same goes for data carried over the connections they establish. Proxy aware applications, such as web browsing and e-mail, require no modification to use onion routing, and do so through a series of proxies.

Traffic Analysis: Protocols, Attacks, Design Issues and Open Problems [9] performs traffic analysis and expose the most important protocols, attacks and design issues. Afterwards, we propose directions for further research. As we are mostly interested in efficient and practical Internet based protocols, most of the emphasis is placed on mix based constructions.

Network FlowWatermarking Attack on Low-Latency Anonymous Communication Systems[11] fundamental limitations of flow transformations in achieving anonymity, and hence, it shows that flow transformations do not necessarily provide the level of anonymity people have expected or believed. By injecting unique watermark into the inter-packet timing domain of a packet flow, we are able to make any sufficiently long flow uniquely identifiable even if 1) it is disguised by substantial amount of cover traffic, 2) it is mixed or merged with a number of other flows, 3) it is split into a number subflows, 4) there is a substantial portion of packets dropped, and 5) it is perturbed in timing due to either natural network delay jitter or deliberate timing perturbation.

Crowds[12]: Anonymity for Web Transactions introduced a system called Crowds for protecting users' anonymity on the world- wide-web. Crowds, named for the notion of blending into a crowd", operates by grouping users into a large and geographically diverse group (crowd) that collectively issues requests on behalf of its members. Web servers are unable to learn the true source of a request because it is equally likely to have originated from any member of the crowd, and even collaborating crowd members cannot distinguish the originator of a request from a member who is merely forwarding the request on behalf of another.

The Predecessor Attack[13]: An Analysis of a Threat to Anonymous Communications Systems

uses that when a particular initiator continues communication with a particular responder across path reformations, existing protocols are subject to the attack. This result to place an upper bound on how long existing protocols, including Crowds, Onion Routing, Hordes, Web Mixes, and DC-Net, can maintain anonymity in the face of the attacks described. This provides a basis for comparing these protocols against each other.

Statistical Disclosure Attacks Traffic Conformation[15] in Open Environmentsis implemented the improvement over the previously known disclosure attack is presented that allows, using statistical methods, to effectively deanonymize users of a mix system. Furthermore the statistical disclosure attack is computationally efficient, and the conditions for it to be possible and accurate are much better understood. The new attack can be generalized easily to a variety of anonymity systems beyond mix networks.

Two-sided Statistical Disclosure Attack[17] introduced a new traffic analysis attack: the Two-sided Statistical Disclosure Attack, that tries to uncover the receivers of messages sent through an anonymizing network supporting anonymous replies. An abstract model of an anonymity system is used with users that reply to messages. Based on this model, a linear approximation is used describing the likely receivers of sent messages.

Perfect Matching Disclosure Attacks[18] presented a user behavior model that, to the best of our knowledge, is the least restrictive scheme considered so far. Second, it has been developed the Perfect Matching Disclosure Attack, an efficient attack based on graph theory that operates without any assumption on user behavior. The attack is highly effective when de-anonymizing mixing rounds because it considers all users in a round at once, rather than single users iteratively. Furthermore, the extracted sender-receiver relationships can be used to enhance user profile estimations.

Traffic Inference in Anonymous MANETs [22], a novel traffic inference algorithm, called TIA, which enables a passive global adversary to accurately infer the traffic pattern in an anonymous MANET without compromising any node. As the first work of its kind, TIA works on existing on-demand anonymous MANET routing protocols. Detailed simulations show that TIA can infer the traffic pattern with an accuracy as high as 95%.

A statistical traffic pattern discovery system (STARS)[23] aimed to derive the source/destination probability distribution, i.e., the probability for each node to be a message source/destination, and the end-to-end link probability distribution, i.e., the probability for each pair of nodes to be an end-to-end communication pair. To achieve its goals, STARS includes two major steps: 1) Construct point-to-point traffic matrices using the time-slicing technique, and then derive the end-to-end traffic matrix with a set of traffic filtering rules; and 2) Apply a heuristic approach to identify the actual source and destination nodes, and then correlate the source nodes with their corresponding destinations. The contribution of STARS is twofold: 1) To the best of our knowledge, STARS is the first statistical traffic analysis approach that considers the salient characteristics of MANETs: the broadcasting, ad hoc, and mobile nature; and 2) most of the previous approaches are partial attacks in the sense that they either only try to identify the source (or destination) nodes or to find out the corresponding destination (source) nodes for given particular source (destination) nodes. STARS is a complete attacking system that first identifies all source and destination nodes and then determines their relationship.

III. GSTARS

A. Proposed Work

In the existing system the adversaries can globally monitor the traffic across the entire network region. This assumption is conservative from the network users’ point of view. Usually, it is difficult for the attackers to perform such a global traffic detection. However, even though the adversaries are not able to monitor the entire network, they can monitor several parts of the network simultaneously. For example, an attacker can deploy sensors (signal detectors) around some particular mobile nodes to track their movements and eavesdrop all of their traffic. These sensors may even move accordingly. With the restricted capabilities, the attacker can take advantage of STARS to perform traffic analysis as follows: 1. divide the entire network into multiple region geographically; 2. deploy sensors along the boundaries of each region to monitor the cross-component traffic; 3. treat each region as a supernode and use STARS to figure out the sources, destinations, and end-to-end communication relations; and 4. analyze the traffic even when nodes are close to each other by treating the close nodes as a supernode. This method is called as variant of STARS or the Generalized STARS (GSTARS). To perform GSTARS, the adversaries only need to monitor the nodes beside the boundaries of the supernodes. The traffic inside each supernode can be ignored, since it will not affect the inter-region traffic patterns. In addition, GSTARS does not need the signal detectors to be able to precisely locate the signal source. They are only required to determine which supernode (region) the signals are sent from. Moreover, in STARS, the actual receiver of a point-to-point transmission is not identifiable among all the potential receivers within the sender’s transmitting range. This inaccuracy can be mitigated in GSTARS because most potential receivers of a packet will be contained within one or a few supernodes.

B. Algorithm

1. Compute end-to-end traffic matrix , R.
2. Obtain the probability distribution vector D.
3. Modify point-to-point matrix by eliminating the traffic sent by node i and destination probability distribution vector D$^-$
4. Subtract D$^-$ from D resulting in L_{s-d}
5. And this determines the optimized routing.

C. End-to-end Traffic Matrix

- *Capture all the traffic* within a period of time partitioned into a sequence of time intervals ¢*t1, ¢t2, ..., ¢tK to form a sequence of traffic* matrices W*j1£K=(W1, W2, ..., WK).*
- *Each traffic matrix* W*e = (we(i; j))N£N records the traffic captured in time* interval ¢*te, where N is the size of the network, e=1, 2, ..., K,* and
- *we(i; j) is the point-to-point traffic volume captured from* node *i to node j.*
- *we(i; j):pkt denotes the* set of all packets contributing to *we(i; j).*
- Given a sequence of point-to-point traffic matrices W*j1£K, an end-to-end traffic matrix* R = (*r(i; j))N£N can be derived, where r(i; j) is the* accumulative traffic volume from node *i to node j.*
- *The algorithm to derive point-to-point (Accumulative) traffic matrix is*

```
F (w|$_{1xk}$ )
R=W$_1$
for e=1 to k-1
        R=α(R,W$_{e+1}$) +W$_{e+1}$
return β(R)
α(R,W$_{e+1}$)
for j, k=1 to N
    for each x ε W$_{e+1}$ (j, k).pkt
    for i=1 to N
    If Э yεr (i, j).pkt that y.hop<H
    andx.time-y.time<T
    Create z with z.size=y.size
    z.time=y.time
    z.hop=y.hop+1
    r (i,j).pkt= r(i,j).pkt U {z}
    r (i,j)=r(i,j) +z.size

return R
β(R)
fori,j=1 to N
    for each p ε r(i,j).pkt
    ifp.hop<h
    r(i,j).pkt= r(i,j).pkt- { p }
    r(i,j)=r(i,j)-p.siz

Return R
```

D. Probability Distribution VectorCalculation

To S-Vectors (S_0, S_1,...,S_f) and D-Vectors (D_0, D_1,...,D_f)

Each element S_n i in S_n stands for the probability that node i is a source, and d_n i in D_n stands for the probability that node i is a destination. The ith row (r(i; 1) ...r(i;N)) in the matrix R is a vector of the traffic from node i to every node in the network. If we multiply this vector by D0 (inner product), the resultant is

$$s'i = \sum_{j=1}^{N} r(i\,,j)\; X\; d^{0}\mathrm{j}$$

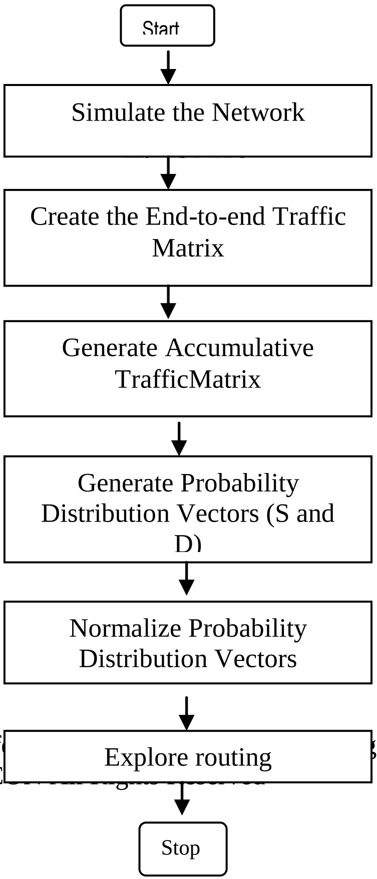
$$d_i{}' = \sum_{j=1}^{N} r(j,i)\; X\; s'j$$

$$(D)^{I+1} = (R.R^{\,T}).\; D^{\,T}$$

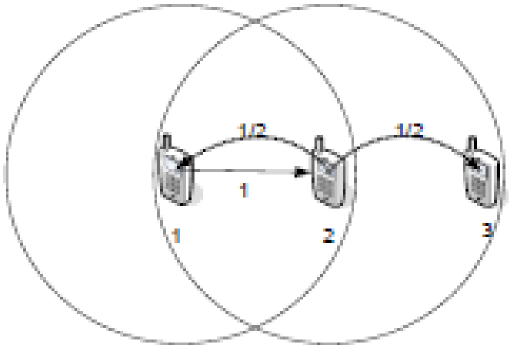
IV. Workflow of GSTARS

A. Architectural Flow Diagram

The hidden traffic patterns in a MANET communication system, STARS includes two major steps. First, it uses the captured traffic to construct a sequence of point-to-point traffic matrices and then derives the end-to-end traffic matrix. Second, further analyzing the end-to-end traffic matrix, it calculates the probability for each node to be a source/destination (the source/destination probability distribution) and that for each pair of node to be an end-to-end communication link (the end-to-end link.


Start
Simulate the Network
Create the End-to-end Traffic Matrix
Generate Accumulative TrafficMatrix
Generate Probability Distribution Vectors (S and D)
Normalize Probability Distribution Vectors
Explore routing
Stop

B. A Simple Ad hoc Network



C. End-to-end Traffic Matrix

$$W_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\; W_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0.5 & 1 & 0.5 \\ 0 & 0 & 0 \end{bmatrix}$$

$$R = \begin{bmatrix} 0 & 1 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0 & 0 & 0 \end{bmatrix}$$

D. Probability Distribution Vector Calculation

S-Vector = $(0.77,0.23,0)^{T}$

D-Vector = (0.08,0.55,O.37

S* = $(0.6,0.4,0)^{T}$

D* = $(0.2,0.4,0.4)^{T}$

D_n=$(0.5,0,0.4)^{T}$

E. Probability Reduction Vector Calculation

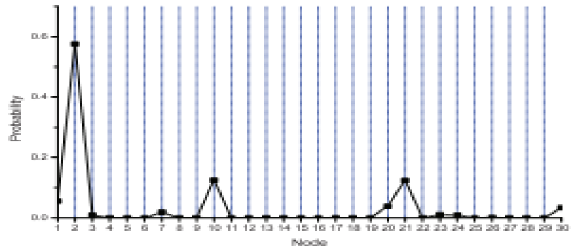
L’1 $= D_0 - D_n$ = (-0.42,0.55,-0.13)^T

V. EXPERIMENTAL RESULTS

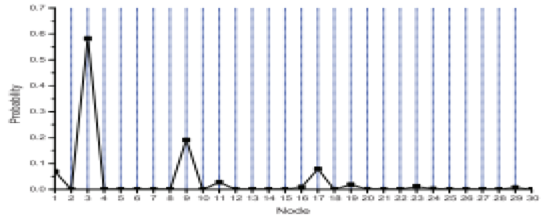
The experiment is conducted as a two steps: Simulation and evaluation. First, the MANET is simulated with 30 mobile nodes and sensors and base stations are identified. Routing strategies are explored. Then according to the routing, the traffic matrix is constructed for point-point traffic with large number of probabilities. And then the accumulative matrix is evaluated. Then it calculates the S and D vectors which are the probability distribution vectors of source and destination respectively. Finally the Probability distribution vectors are optimized by normalization. This gives the routing exploration as a generalized one.

Evaluation

The probability distributions produced by GSTARS are good indicators of actual traffic patterns, i.e., actual sources, destinations, and end-to-end links. To measure the performance the top k nodes and links with highest probabilities are selected.



(a) Source Probability Distribution in (S1).



(b) Destination Probability Distribution in (S2).

III. Conclusion

In this paper, a novel GSTARS for MANETS has been proposed. A simulated MANETs’explores the hidden traffic and evaluation shows that nodes with highest probabilities are automatically selected and regulates the traffic with high performance.

In future the GSTARS can be designed in such a way that it explores the hidden traffic and in addition it may handle the active and passive attacks of the Mobile Adhoc networks.

References

[1] J. Kong, X. Hong, and M. Gerla, “An Identity-Free and On-Demand Routing Scheme against Anonymity Threats in Mobile Ad Hoc Networks,” IEEE Trans. Mobile Computing, vol. 6, no. 8, pp. 888-902, Aug. 2007.

[2] Y. Zhang, W. Liu, W. Lou, and Y. Fang, “MASK: Anonymous On-Demand Routing in Mobile Ad Hoc Networks,” IEEE Trans.Wireless Comm., vol. 5, no. 9, pp. 2376-2385, Sept. 2006.

[3] A. Boukerche, K. El-Khatib, L. Xu, and L. Korba, “SDAR: A Secure Distributed

Anonymous Routing Protocol for Wireless and Mobile Ad Hoc Networks," Proc. IEEE 29th Ann. Int'l Conf. Local Computer Networks (LCN '04), pp. 618-624, 2004.

[4] S. Seys and B. Preneel, "ARM: Anonymous Routing Protocol for Mobile Ad Hoc Networks," Proc. IEEE 20th Int'l Conf. Advanced Information Networking and Applications Workshops (AINA Workshops '06), pp. 133-137, 2006.

[5] R. Shokri, M. Yabandeh, and N. Yazdani, "Anonymous Routing in MANET Using Random Identifiers," Proc. Sixth Int'l Conf. Networking (ICN '07), p. 2, 2007.

[6] R. Song, L. Korba, and G. Yee, "AnonDSR: Efficient Anonymous Dynamic Source Routing for Mobile Ad-Hoc Networks," Proc. Third ACM Workshop Security of Ad Hoc and Sensor Networks (SASN '05), pp. 33-42, 2005.

[7] M. Reed, P. Syverson, and D. Goldschlag, "Anonymous Connections and Onion Routing," IEEE J. Selected Areas in Comm., vol. 16, no. 4, pp. 482-494, May 2002.

[8] D. Chaum, "Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms," Comm. ACM, vol. 24, no. 2, pp. 84-88, 1981.

[9] J. Raymond, "Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems," Proc. Int'l Workshop Designing Privacy Enhancing Technologies: Design Issues in Anonymity and Unobservability, pp. 10-29, 2001.

[10]W. Dai, "Two Attacks against a PipeNet-Like Protocol Once Used by the Freedom Service," http://weidai.com/freedomattacks. txt, 2013.

[11]X. Wang, S. Chen, and S. Jajodia, "Network Flow Watermarking Attack on Low-Latency Anonymous Communication Systems," Proc. IEEE Symp. Security and Privacy, pp. 116-130, 2007.

[12] M. Reiter and A. Rubin, "Crowds: Anonymity for Web Transactions," ACM Trans. Information and System Security, vol. 1, no. 1, pp. 66-92, 1998.

[13] M. Wright, M. Adler, B. Levine, and C. Shields, "The Predecessor Attack: An Analysis of a Threat to Anonymous Communications Systems," ACM Trans. Information and System Security, vol. 7, no. 4, pp. 489-522, 2004.

[14] D. Figueiredo, P. Nain, and D. Towsley, "On the Analysis of the Predecessor Attack on Anonymity Systems," technical report, Computer Science, pp. 04-65, 2004.

[15] G. Danezis, "Statistical Disclosure Attacks: Traffic Confirmation in Open Environments," Proc. Security and Privacy in the Age of Uncertainty (SEC '03), vol. 122, pp. 421-426, 2003.

[16] G. Danezis and A. Serjantov, "Statistical Disclosure or Intersection Attacks on Anonymity Systems," Proc. Sixth Information Hiding Workshop (IH '04), pp. 293-308, 2004.

[17] G. Danezis, C. Diaz, and C. Troncoso, "Two-Sided Statistical Disclosure Attack," Proc.

Seventh Int'l Conf. Privacy Enhancing Technologies, pp. 30-44, 2007.

[18] C. Troncoso, B. Gierlichs, B. Preneel, and I. Verbauwhede, "Perfect Matching Disclosure Attacks," Proc. Eighth Int'l Symp. Privacy Enhancing Technologies, pp. 2-23, 2008.