# ON TRAFFIC ANALYSIS ATTACKS AND COUNTERMEASURES

A Dissertation

by

XINWEN FU

Submitted to Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2005

Major Subject: Computer Engineering

# ON TRAFFIC ANALYSIS ATTACKS AND COUNTERMEASURES

A Dissertation

by

XINWEN FU

Submitted to Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

| | |
|---|---|
| Co-Chairs of Committee, | Wei Zhao |
| | Riccardo Bettati |
| Committee Members, | Narasimha Reddy |
| | Dmitri Loguinov |
| Head of Department, | Valerie E. Taylor |

December 2005

Major Subject: Computer Engineering

# ABSTRACT

On Traffic Analysis Attacks and Countermeasures. (December 2005)

Xinwen Fu, B.S., Xi'an Jiaotong University;

M.S., University of Science and Technology of China

Co-Chairs of Advisory Committee: Dr. Wei Zhao
Dr. Riccardo Bettati

Security and privacy have gained more and more attention with the rapid growth and public acceptance of the Internet as a means of communication and information dissemination. Security and privacy of a computing or network system may be compromised by a variety of well-crafted attacks.

In this dissertation, we address issues related to security and privacy in computer network systems. Specifically, we model and analyze a special group of network attacks, known as *traffic analysis attacks*, and develop and evaluate their countermeasures. Traffic analysis attacks aim to derive critical information by analyzing traffic over a network. We focus our study on two classes of traffic analysis attacks: link-load analysis attacks and flow-connectivity analysis attacks.

Our research has made the following conclusions:

1. We have found that an adversary may effectively discover link load by passively analyzing selected statistics of packet inter-arrival times of traffic flows on a network link. This is true even if some commonly used countermeasures (e.g., link padding) have been deployed. We proposed an alternative effective

countermeasure to counter this passive traffic analysis attack. Our extensive experimental results indicated this to be an effective approach.

2. Our newly proposed countermeasure may not be effective against active traffic analysis attacks, which an adversary may also use to discover the link load. We developed methodologies in countering these kinds of active attacks.
3. To detect the connectivity of a flow, an adversary may embed a recognizable pattern of marks into traffic flows by interference. We have proposed new countermeasures based on the digital filtering technology. Experimental results have demonstrated the effectiveness of our method.

From our research, it is obvious that traffic analysis attacks present a serious challenge to the design of a secured computer network system. It is the objective of this study to develop robust but cost-effective solutions to counter link-load analysis attacks and flow-connectivity analysis attacks. It is our belief that our methodology can provide a solid foundation for studying the entire spectrum of traffic analysis attacks and their countermeasures.

To my parents, my parents-in-law, my wife, my sisters, and my brother

# ACKNOWLEDGMENTS

It has been almost six years since I started my graduate studies at Texas A&M University. When I sit back and think about the people who have influenced and helped me to complete this dissertation, I am overwhelmed!

First and foremost, I would like to thank my advisor, Dr. Wei Zhao, for recruiting me as a student from China about six years ago. It was an extraordinary piece of good fortune to have the opportunity to work with him. He has been an ideal advisor in every aspect, both in terms of technical advice on my research and in terms of professional advice. I am greatly indebted to him for his great concern, strong encouragement, and unceasing support in matters inside and outside of academics. My choice of a career path has been greatly influenced by him and I hope I can live up to his high standards.

I benefited greatly from the technical and career advice provided by my co-advisor Dr. Riccardo Bettati. Dr. Bettati has been a co-principal investigator on all the projects discussed in this dissertation. I am grateful to him for serving on my dissertation committee and for the contributions he has made to my research. He has given me extremely useful and incisive comments on my research and paper writing and advised me how to view a system problem and ask the right questions.

I would like to thank Dr. A. L. Narasimha Reddy for his involvement and guidance as a member of my dissertation committee. I admire his deep insight in the area of network systems. His insightful suggestions and constructive criticism of my work have inspired me with excellent results.

Appreciation also goes to Dr. Dmitri Loguinov for his involvement and his guidance as a member of my dissertation committee. Interaction with him in and out of his classes on networking has greatly helped me to make progress in my research. I learned a lot about the hot topics in the reseach of networking when I took his class of “Special Topics on Networking”.

I thank Dr. Nitin Vaidya of the University of Illinois at Urbana-Champaign for his involvement and guidance as an ex-member of my dissertation committee. He offered me many valuable suggestions on my career path development. It now proves that his suggestions are right and the best for me. The discussion during our group seminars has greatly helped me to start my work on security and networking.

It is a great pleasure to thank and acknowledge many professors and researchers with whom I have worked. Among them, I am especially grateful to Dr. Miaolan Zhang and Dr. Xianglin Li of the Graduate School of University of Science and Technology of China, my graduate advisors, for their guidance and endless help. Also, I would like to thank Dr. Guofang Tu of the Graduate School of University of Science and Technology of China for his guidance on my research and strong support. I would like to thank Dr. Lawrence Petersen, Dr. Bart Childs, and Dr. Don Friesen for their suggestions and support on my teaching and career path choice.

My fellow graduate students in the real-time and security system group have made countless hours in the office enjoyable. I have benefited greatly from working with Dr. Yong Guan, Dr. Dong Xuan, Dr. Shu Jiang, Bryan Graham, Hongyun Xu, and Ye Zhu on anonymity and information assurance; Dr. Bo Sun, Dr. Bin Lu, Dr. Jianfeng Cai, Dr.

Jian Chen, and Benjamin Collins on wireless security; and Dr. Guangtong Cao on distributed systems. I have also had many helpful discussions with faculty members: Mr. Willis Marti, Dr. Marina Vannucci, Dr. Udo Pooch, Dr. Hank Walker, and Dr. Dianxiang Xu, and students: Dr. Byung-kyu Choi, Liliana Grigoriu, Linli He, Shengquan Wang, Zhibin Mai, Jianjia Wu, Dr. Nan Ni, Dr. Sangig Rho, Dr. Jianwen Yin, Dr. Yong Xiong, Dr. Yu Zhang, Guobin He, Chunhua Tang, Weimin Zhang, Dr. Di Wu, Dr. Hao Yu, Wei Yu, Dan Cheng, and Nan Zhang.

I would like to express appreciation to Ms. Elena Catalena, Ms. Sherry Escalante, Ms. Ms. Patricia M Rudkin, Ms. Kathy Flores, Ms. Susan Spears, Ms. Sandra Morse, and Ms. Larisa Archer. In particular, I am deeply grateful to Ms. Archer for kindly helping me read through this whole dissertation. Ms. Larisa Archer has been revising most of my writing during the last few years and I have learned a lot from her on both written and spoken English skills.

I owe a special debt of gratitude to my parents and family. They have been, more than anyone else, the reason I have been able to get this far. I would like to express my heartfelt appreciation to my wife, Shan Tang, for her selfless love and support that has made me want to excel. She has encouraged and helped me in many ways during my studies. Words cannot express my gratitude to my parents and parents-in-law, Yumei Hou, Kede Fu, Zhenhua Luo, and Shengqi Tang; and to my sisters and brother, Wenhua Fu, Wangdi Fu, and Wenge Fu. They all have given me their support and love from across the seas. They instilled in me the value of hard work and taught me how to overcome life’s disappointments.

TABLE OF CONTENTS

Page

Page

Page

# LIST OF FIGURES

Page

Page

# LIST OF TABLES

Page

# 1. INTRODUCTION

## 1.1. Overview of Network Security and Attacks

As the use of computer networks, especially the Internet, has become widespread, the concept of computer security and privacy has expanded to denote issues pertaining to the networked use of computers and their resources. Thousands of successful break-ins over the years illustrate that we need to evaluate security of systems systematically, study and understand attack techniques, and then develop corresponding countermeasures.

Security attacks can be classified into two classes: *active* and *passive*. In active attacks, an attacker changes, hence targeting damages directly. For example, network messages can be altered, and web service can be denied by active attacks. In passive attacks, an attacker simply listens or eavesdrops without actively changing anything. Examples of passive attacks in daily life and compute domain are shoulders-surfing (i.e., observing displays and keystrokes over someone's shoulder), telephone tapping, and wireless sniffing. Obviously, it is more difficult to detect passive attacks than active ones.

To deal with these attacks, a variety of countermeasures have been developed, and they can be classified as either offensive or defensive. By offensive countermeasures, we mean that defenders act on attacks (aimed at the sensitive targets), and intend to actively stop operations of the adversary. For example, the widely used free-source intrusion detection and response system *snort* [1] is an example of the offensive countermeasure.

---

This dissertation follows the style of *IEEE Transactions on Computers*.

It can monitor live traffic on a network, use a rule engine to identify patterns within traffic, and take actions such as filtering out the identified attacking traffic. When referring to *defensive* countermeasures, we mean that defenders hide the sensitive target and passively deny or limit an enemy's ability to access it. Encryption is an example of a defensive countermeasure.

In this dissertation, we model and analyze a special class of attacks, namely traffic analysis attacks, which can be either active or passive, and develop and evaluate countermeasures to them.

## 1.2. Overview of Traffic Analysis Attacks

Traffic analysis attacks are aimed at deriving critical information by analyzing the statistics of traffic flows. For example, in a military communication network, by intercepting traffic using sniffing tools and monitoring pattern change of link load, an adversary may uncover the location of command centers, determine the state of alertness of various units, and/or detect covert information flows to or from apparently non-involved parties.

Traffic analysis attacks challenge the design of traditional systems where encryption is typically used as the main method for protecting security and privacy. However, it is obvious that encryption cannot protect many other important characteristics of traffic which may be mission critical and require protection.

We consider two types of traffic analysis attacks. The first is *link-load analysis attack* that aims at discovering traffic rate on a network link. The second is *flow-connectivity attack* that intends to discover flow connectivity between two hosts. Link

load and flow connectivity in many applications are considered mission critical information and should be protected. We will develop and evaluate countermeasures against these two types of attacks from revealing link load and flow connectivity.

## 1.3. Summary of This Dissertation Research

As mentioned earlier, we are dealing with traffic analysis attacks and their countermeasures. In particular, we focus on link-load analysis attacks and connectivity analysis attacks, as defined above.

For each kind of attack, we assume that an attacker may use either passive or active means to launch attacks. We formally propose adversary models and define performance metrics based on which effectiveness and efficiency of attacks and countermeasures can be evaluated. We will systematically analyze existing countermeasures and identify their weaknesses. Based on our studies of the existing solutions, we will develop and evaluate new countermeasures aimed at overcoming problems associated with current solutions.

In this dissertation research, we have obtained the following promising results:

1. An adversary may effectively discover link load by passively analyzing statistics of packet inter-arrival times of traffic flows passing through a network link. This is true even if some commonly used countermeasures such as link padding have been deployed. We proposed an alternative effective countermeasure to counter this kind of passive link-load analysis attack.
2. Our newly proposed countermeasure may not be effective against active traffic analysis attacks, which an adversary may also use to discover the link load. We propose to develop methodologies in countering these kinds of active attacks.

3. To detect the connectivity of a flow, an adversary may invoke active attacks over the flow traffic and embed a recognizable pattern of marks into the flow traffic. We proposed new countermeasures based on digital filtering technology. Initial results have demonstrated the effectiveness of our method.

In dealing with both link-load analysis attacks and connectivity analysis attacks, we measure the degree of system security by *detection rate*, which is defined as the probability that an adversary discovers the target information. For a link-load analysis attack, the needed information is the link load while for a connectivity analysis attack; the target information is the flow connectivity between hosts. We derive analytical formulas of detection rates for both attacks.

Furthermore, we carry out extensive experiments on both local and wide area networks for the two classes of attacks and their countermeasures. The objective of these experiments is to validate the correctness of our theoretical analysis.

In summary, in this dissertation, we model and analyze a few traffic analysis attacks, and develop and evaluate corresponding countermeasures. We hope that our methodology will provide a solid foundation for studying the entire spectrum of traffic analysis attacks and their countermeasures.

## 1.4. The Organization of This Dissertation

The rest of this dissertation is organized as follows. In Section 2, we review related work. In Section 3, we discuss passive link-load analysis attacks and their countermeasures. In Section 4, we discuss active link-load analysis attacks and their countermeasures. In

Section 5, we discuss flow-connectivity analysis attacks and the filter-based countermeasures. We summarize this dissertation research in Section 6.

# 2. RELATED WORK

## 2.1. Related Work for Link-Load Analysis Attacks

Generally speaking, perfect secrecy theory by Shannon [2] is the foundation for the ideal countermeasure system against statistical analysis attacks. Researchers have proposed and analyzed various countermeasures and intended to realize or approximate the perfect secrecy model in one form or another.

Baran [3] suggests adding *dummy* (fraudulent) traffic to conceal the true amount of traffic. A survey of countermeasures for traffic analysis is given in [4]. To mask the frequency, length and origin-destination patterns of an end-to-end communication, dummy messages are used to pad the traffic to a predefined pattern. It is evident that such a predefined pattern is sufficient but not necessary based on the perfect secrecy theory.

Newman-Wolfe and Venkatraman [5][6][7] give a mathematical framework to optimize the bandwidth usage while preventing traffic analysis of the end-to-end traffic rates. Timmerman [8] proposes an adaptive traffic hiding model to reduce the overhead caused by traffic padding, in which the link padding rate is reduced along with the decrease of real traffic rate. Kung, Cheng, Tan, and Bradner [9] use a similar approach, denoted as on demand traffic padding, which will generate dummy traffic only when the payload traffic is present. These approaches render large-scale variations in traffic rates still observable.

Researchers at Texas A&M University have developed NetCamo [10], which provides the end-to-end prevention of traffic analysis while guaranteeing QoS (the worst case delay of message flows) in time constraint communication networks. Song, Wagner, and Tian [11] analyze how SSH 1 and SSH 2 can leak user passwords under a passive traffic analysis attack and propose using traffic padding to counter the attack.

The above mentioned previous studies focus on applications utilizing the commonly used traffic padding approach to counter traffic analysis attacks. In this dissertation, we will evaluate the security level a traffic padding approach can provide and design new effective countermeasures.

## 2.2. Related Work for Connectivity Analysis Attacks

To protect the anonymity of email transmissions, Chaum [12] proposes the use of Mixes. Many researchers suggest using CIT padding between a user and the first mix [13]. Raymond in [14] gives an informal survey of several *ad hoc* traffic analysis attacks on systems providing anonymous services. One of his conclusions is that traffic padding is essential to achieving communication anonymity. Back, Möller, and Stiglic [15] list many possible attacks in Freedom anonymous communication system [16]. Danezis, Dingledine, and Mathewson [17] give a list of attacks to anonymity systems. Most of those attacks are only briefly discussed and lack systematic analysis. Freedman and Morris developed Tarzan [18], which provides anonymity in a peer-to-peer environment by using link padding to counter possible attacks.

The concept of continuous-time mix is introduced by Danezis in [19]. He proves that the optimal mix strategy that maximizes anonymity is the *Exponential Mix*, i.e. a Stop-and-Go Mix that delays packets individually according to an exponential distribution.

Kong and Hong [20] develop an anonymity protocol for wireless ad-hoc networks. When Alice tries to communicate with Bob, by broadcasting encrypted route discovery messages recursively, she can find a route to Bob, who responds to the request through the reverse path. Thus an anonymity path is built from Alice to Bob. The authors and other researchers also mention using broadcast MAC addresses to achieve more protection. But the whole protocol is still susceptible to the flow marking attack.

Sun, Simon, Wang, and Russell [21], and Hintz [22] give quantitative performance analysis for an anonymous web server that uses encryption and packet header mangling such as in a NAT proxy. The analysis takes advantage of the fact that a number of HTTP features, such as the number and size of objects, can be used as signatures to identify web pages with some accuracy. Unless the web anonymizer addresses this issue, these signatures are visible to the adversary. Serjantov and Sewell [23] analyze the possibility of a lone flow along an input link of a mix in peer-to-peer anonymity systems. If the rate of this lone input flow is approximately equal to the rate of a flow out of the mix, this pair of input and outflow flows is correlated.

Guan, Fu, Bettati, and Zhao [24] define an entropy-based metric to evaluate the anonymity degree of an anonymity system. They consider attacks where compromised nodes cooperate to correlate packets passing those nodes and find the sender of a packet. Many works in the related literature study this kind of packet-level attack. In this

dissertation, we consider flow-level attacks which may be more flexible and dangerous. We also study their countermeasures.

To find if a party (Bob) is communicating with another party (Alice) an adversary may measure the similarity between Bob's outbound traffic and Alice's inbound traffic. Zhu, Fu, Graham, Bettati, and Zhao [25] propose using mutual information for the similarity measurement. Levine, Reiter, Wang, and Wright [26] use cross correlation to measure similarity between flows. Both papers consider passive traffic analysis attacks against flow connectivity, while we discuss a class of active traffic analysis attacks which may be more flexible and dangerous.

Serjantov and Peter [23] and some other researchers mention very briefly that an adversary may introduce a *spike* into traffic to find the communication relationship between users, but without any in-depth study of how to introduce spikes, what kind of spike should be introduced, or how to recognize the spike. We generalize this kind of attack in wired and wireless networks and build a complete framework to answer the above questions.

# 3. PASSIVE LINK-LOAD ANALYSIS ATTACKS AND COUNTERMEASURES

In this section, we investigate the effectiveness of link padding, which can be used as a countermeasure against passive link load analysis attacks. We first present the network model, padding mechanism, and adversary strategy. We then develop a theoretical model and derive closed-form formulae for detection rates. Finally, we validate results from the theoretical model by experiments.

## 3.1. Models

In this section, we present the model of the network in our study and then formally define the model of the adversary, which uses statistical pattern recognition strategies for traffic analysis attacks.

### 3.1.1. Network Model

In this work, we assume that the network consists of *protected subnets*, which are interconnected by *unprotected networks*. It is assumed that traffic within protected subnets is shielded from observers. Unprotected networks can be either public networks (e.g., the Internet) or networks that are deployed over an easily accessible broadcast medium. These networks are accessible to observation by third-parties and are, therefore, open to traffic analysis. This model captures a variety of situations, ranging from battleship convoys (where the large-scale shipboard networks are protected and the inter-ship communication is wireless) to communicating Personal Digital Assistants (PDAs) (where the protected networks consist of single nodes).

Figure 1 illustrates the setup of the network in this study. Two security gateways $GW_A$ and $GW_B$ are placed at the two boundaries of the unprotected network and provide the link padding necessary to prevent traffic analysis of the payload traffic exchanged between the protected subnets A and B.

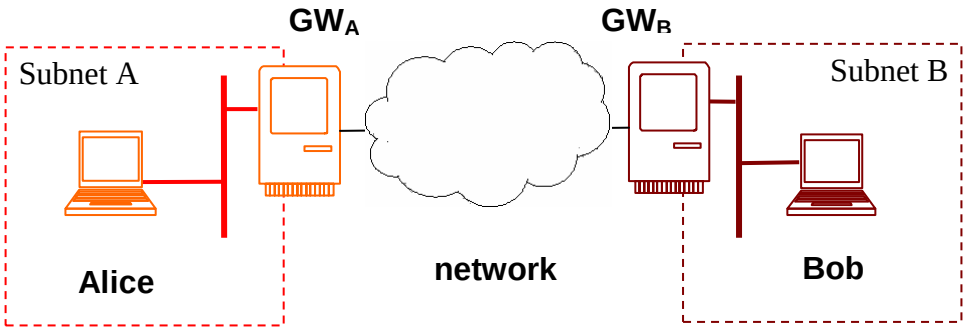



Figure 1. Network Model

Note that the gateways can be realized as either stand-alone boxes, modules on routers or switches, software additions to network stacks, or device drivers at the end hosts[1]. In this section, we assume that gateways are stand-alone boxes. Nevertheless, the analysis in this section should also be valid for other implementations. To simplify the discussion, the communication is one-way from Subnet *A* to Subnet *B*. Consequently, $GW_A$ and $GW_B$ are also called *sender gateway* and *receiver gateway*, respectively.

[1] This is the case when end users of onion-routing-like anonymity systems choose to use link padding.

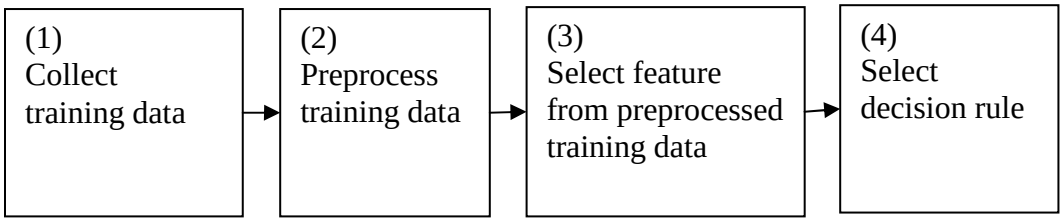
### 3.1.2. Adversary Model

The goal of the adversary is to perform traffic analysis and infer critical characteristics of the payload traffic exchanged between protected subnets over the unprotected network. We assume that the adversary is only interested in the *payload traffic rate*, that is, the rate at which payload traffic is exchanged between protected networks. As mentioned earlier, the traffic rate is an important piece of information in many mission-critical communication applications [14]. Specifically, we assume that there is a set of discrete payload traffic rates $\{\omega_1, \ldots, \omega_m\}$. The rate of payload traffic from the sender will be one of those rates at a given time. Consequently, the objective of the adversary is to identify at which of these rates the payload is being sent.

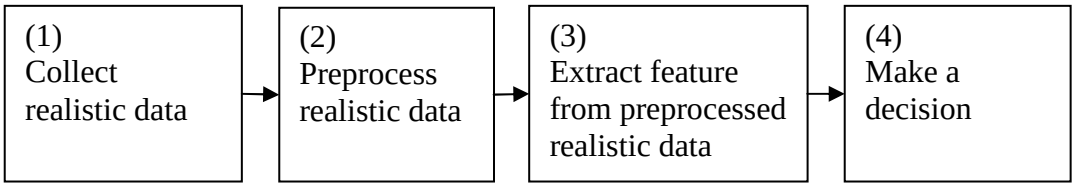
We assume that to identify at which rate the payload is being transmitted, the adversary limits himself to passive attacks, i.e., observations of the traffic. In addition, the adversary's access to the system is limited to the unprotected networks. The protected subnets and hosts within are not accessible. Neither is the link padding infrastructure. This means that, in Figure 1, the adversary can only tap somewhere between gateways $GW_A$ and $GW_B$.

We also assume that the adversary has complete knowledge about the gateway machines and the countermeasure algorithms used for preventing traffic analysis. For example, the adversary can simulate the whole system, including the gateway machines, to obtain *a priori* knowledge about traffic behavior. In many studies on information security, it is a convention that we make worst case assumptions, such as this one.

Based on these assumptions, the adversary can deploy a strategy based on Bayes decision theory [27]. The entire attack strategy consists of two phases: the offline training phase and the online classification phase as shown in Figure 2. We will describe these below.




(a) Offline Training Phase




(b) Online Recognition Phase

Figure 2. Flow Marking Attack Framework

The offline training phase in Figure 2 (a) can be decomposed into the following steps:

1. Collecting training data: The adversary reconstructs the entire link padding system and collects timing information at different payload traffic rates.

2. Preprocessing training data: From packet timing information, the adversary derives the *Packet Inter-Arrival Time* (PIAT) of the traffic.
3. Selecting feature from preprocessed training data: The adversary selects a statistical *feature* of the *Packet Inter-Arrival Time* (PIAT) that will be used for traffic rate classification. Possible features we study in this section are sample mean, sample variance, and sample entropy. The adversary then derives the *Probability Density Functions* (PDF) of the selected statistical feature. As histograms are usually too coarse for the distribution estimation, we assume that the adversary uses the Gaussian kernel estimator of PDF [28], which is effective in our problem domain.
4. Selecting decision rule: Based on the PDFs of statistical features for different payload traffic rates, Bayes decision rules are derived. Recall that possible payload traffic rates are $\omega_1$, …, $\omega_m$. The Bayes decision rule can be stated as follows:

*The sample represented by feature "s" corresponds to payload rate $\omega_i$ if*

$$\forall j \in [1, m], P(\omega_i | s) \geq P(\omega_j | s) \tag{3.1}$$

*That is,*

$$f(s|\omega_i)P(\omega_i) \geq f(s|\omega_j)P(\omega_j) \tag{3.2}$$

*where $f(s|\omega_i)$ is the PDF of feature s conditioned on payload traffic rate $\omega_i$, $P(\omega_i)$ is the a priori probability that the payload traffic is sent at rate $\omega_i$, and $P(\omega_i|s)$ is the post priori*

*probability that the payload traffic is sent at rate $\omega_i$ when the collected sample has the measured feature s.*

Once the adversary completes its training phase, she can perform the classification at run time. We assume the adversary uses some means to tap the network between gateways $GW_A$ and $GW_B$. In particular, when she wants to determine the current payload rate, the adversary collects a sample of packet inter-arrival times. She calculates the value of the statistical feature from the collected sample, and then uses the Bayes decision rules derived in the training phase to match the collected sample to one of the previously defined payload traffic rates.

## 3.2. Overview of Countermeasures to Link-Load Analysis Attacks

We now discuss mechanisms that can be used as a countermeasure for traffic analysis attacks.

One way to counter the traffic analysis attacks is to "pad" the payload traffic, that is, to properly insert "dummy" packets in the payload traffic stream so that the real payload status is camouflaged. Link padding algorithms can be implemented in various ways over the two gateways in Figure 1. The most common method is to use a timer to control the packet transmission. It works as follows: On $GW_A$, incoming payload packets from the sender are placed in a queue. An interrupt-driven timer is set up on $GW_A$. When the timer fires, the interrupt processing routine checks whether there is a payload packet in the queue. If there is a payload packet, one is removed from the queue and transmitted to $GW_B$. Otherwise, a dummy packet is transmitted to $GW_B$.

We need to make a few remarks before we proceed further. In this section, we assume that packet contents are perfectly encrypted (e.g., by IPSec with appropriate options) and are, thus, non-observable. In particular, the adversary cannot distinguish between payload packets and “dummy” packets used for padding.

It is obvious from the implementation described above, the only tunable parameter is the time interval between timer interrupts. The choice of this parameter discriminates different padding approaches. A system is said to have a *constant interval timer* (CIT) if the timer is a periodic one, i.e., the interval between two consecutive timer interrupts is constant. This is the most common method used for padding. On the other hand, a system is said to have a *variable interval timer* (VIT) if the interval between two consecutive timer interrupts is a random variable that satisfies some distribution. As we will see in the later part of this section, CIT and VIT systems may perform significantly differently in preventing traffic analysis attacks.

We assume that all packets have a constant size. Thus, observing the packet size will not provide any useful information to the adversary. The only information available for the adversary to observe and analyze is the timing of packets.

## 3.3. Performance Metric and Analysis

### 3.3.1. Detection Rate as Performance Metric

Given the models described in the previous section, we would like to evaluate the system security in terms of detection rate. *Detection rate* is defined as the probability that the adversary can correctly identify the payload traffic rate. In this section, we derive the

closed-form formulae for detection rates when the adversary uses sample mean, sample variance, or sample entropy, as the statistical feature, respectively. Our formulae will be approximate due to the complexity of the problem. Nevertheless, these formulae do correctly reflect the impact of various system parameters, including the type of padded traffic, sample size, and statistical feature used. These relationships are extremely useful in design of a link padding system so that the overall detection rate can be minimized. In the next section, we will see that experimental data matches well to the performance predicted by our approximation formulae.

We will focus our discussion on systems with only two payload traffic rates, namely the low traffic rate and the high traffic rate and assume that both traffic rates occur with equal probability. It is trivial to extend this work to complicated cases.

Figure 3 shows the PDFs of the statistical features conditioned on two alternative payload traffic rates $\omega_l$ *and* $\omega_h$. Let *d* be the solution of the equation

$$f(s|\omega_l) = f(s|\omega_h) \tag{3.3}$$

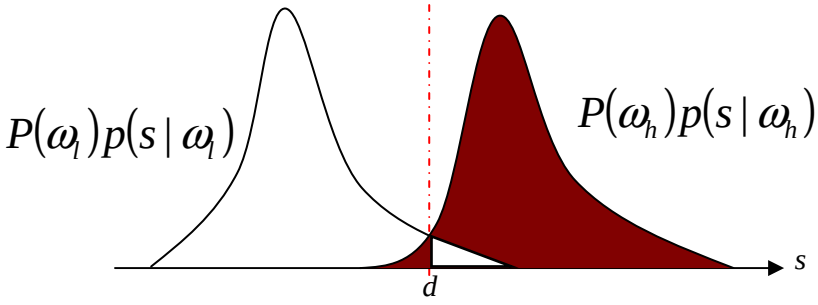



Figure 3. Bayes Decision Making for the Case of Two Payload Traffic Rates

We assume that there is a unique solution to the equation. Consequently, the Bayes decision rule now becomes: *If* $s \leq d$*, the payload traffic rate is* $\omega_l$*. Otherwise, the rate is* $\omega_h$. The error rate for the Bayes decision rule can be calculated as follows:

$$\varepsilon = P(\omega_l)\int_{d}^{\infty} f(s|\omega_l)ds + P(\omega_h)\int_{-\infty}^{d} f(s|\omega_h)ds \tag{3.4}$$

The detection rate is then given by

$$v = 1-\varepsilon \tag{3.5}$$

$$= P(\omega_l)\int_{-\infty}^{d} f(s|\omega_l)ds + P(\omega_h)\int_{d}^{\infty} f(s|\omega_h)ds \tag{3.6}$$

While numerical methods can be applied to calculate the detection rates, (e.g. with the use of (3.2)), our goal here is to derive close-form formulae that can reveal the relationship between the detection rate and other system parameters.

### 3.3.2. Derivation of Detection Rate

#### 3.3.2.1. Decomposition of Packet Inter-Arrival Time

Recall that the adversary collects a sample of PIATs at run time in order to perform the classification. Thus, to derive the detection rate, we need to formally model the PIAT. For a given system, let random variable $X$ be the PIAT. $X$ can be considered as the sum of three other random variables:

$$X = T + \delta_{gw} + \delta_{net} \tag{3.7}$$

where $T$ is the designed interval of two consecutive timer interrupts for the timer, and $\delta_{gw}$ and $\delta_{net}$ reflect the noise added by disturbance in the gateway system and by congestion in the network, respectively.

Note that $T$ is defined by the link padding policy. $T$ should be constant for CIT link padding but follows a specific distribution for VIT link padding.

$\delta_{gw}$ is caused by a number of factors, which may impact the accuracy of the timer's interrupt. First, the context switching from other running process to the timer's interrupt routine may take indeterminate time. Furthermore, a timer interrupt may be temporally blocked due to other activities. For example, if a payload packet from the sender is arriving at the network interface card of the gateway, the network interface card would

generate an interrupt request, which can block all the processes including the (scheduled) timer interrupt[2]. Thus, the timer's interrupts may be subtly but randomly delayed by incoming payload packets. *This implies that the padded traffic's PIAT may be correlated with the payload traffic*.

$\delta_{net}$ captures the disturbance on the padded traffic's PIAT caused by crossover traffic at routers and switches. Clearly, $\delta_{net}$ depends on the position at which the adversary collects its sample. If the collection is made right at the output of the sender gateway, this noise may be ignored. However, if the adversary collects its sample far away from the sender gateway, the noise level can be high as crossover traffic may significantly interfere with the padded traffic.

In this section, we assume that $T$, $\delta_{gw}$ and $\delta_{net}$ are normally distributed. These assumptions simplify analysis without loss of generality and will be validated by our experiments in Section 3.4.1. Specifically,

$$T \sim N\left(\tau, \sigma_T^2\right) \tag{3.8}$$

where $\tau$ is the mean of the timer timeout interval, and $\sigma_T^2 = 0$ in the case of CIT link padding. And

$$\delta_{net} \sim N\left(0, \sigma_{net}^2\right) \tag{3.9}$$

where $\sigma_{net}^2 = 0$ when the adversary observes the padded traffic at a position next to the sender's gateway $GW_A$. Similarly

[2] For TimeSys Linux [29] used in our experiments, this request proceeds before the incoming packet reaches the IP layer [30]. From that instant on, the network subsystem in the kernel becomes preemptive. Other high priority tasks such as the timer interrupt routine can then proceed as scheduled.

$$\delta_{gw} \sim N\left(0, \sigma_{gw}^2\right) \tag{3.10}$$

As $\delta_{gw}$ may be correlated to the payload traffic, we denote $\sigma^2_{gw,l}$ and $\sigma^2_{gw,h}$ as the variances of $\delta_{gw}$ when the payload traffic rate is low and high, respectively. Consequently, we denote $X_l$ and $X_h$ as random variables when the payload traffic rate is low and high, respectively. Thus,

$$X_l \sim N\left(\mu, \sigma_l^2\right) \tag{3.11}$$

where $\mu = \tau$ ($\tau$ is the mean of the timer timeout interval) and

$$\sigma_l^2 = \sigma_T^2 + \sigma_{net}^2 + \sigma_{gw,l}^2 \tag{3.12}$$

Similarly,

$$X_h \sim N\left(\mu, \sigma_h^2\right) \tag{3.13}$$

where $\mu = \tau$ and

$$\sigma_h^2 = \sigma_T^2 + \sigma_{net}^2 + \sigma_{gw,h}^2 \tag{3.14}$$

Here we assume that $X_l$ and $X_h$ have the same mean. This assumption will be validated by our experiments later.

For the convenience of the discussion in the rest of this section, we need to introduce a ratio defined as follows:

$$r = \frac{\sigma_h^2}{\sigma_l^2} = \frac{\sigma_T^2 + \sigma_{net}^2 + \sigma_{gw,h}^2}{\sigma_T^2 + \sigma_{net}^2 + \sigma_{gw,l}^2} \tag{3.15}$$

where $\sigma_T^2$, $\sigma_{net}^2$, $\sigma_{gw,l}^2$,and $\sigma_{gw,h}^2$ are defined in (3.8), (3.9), (3.12), and (3.14), respectively. The usage of $r$ will become clear when we derive the formulae for detection rates for three different statistical features, namely, sample mean, sample variance, and sample entropy.

3.3.2.2. Detection Rate Formula for Sample Mean

Let $\{X_1, X_2, \ldots, X_n\}$ be a random sample of packet inter-arrival times. The *sample mean* is the average of the elements in the sample:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{3.16}$$

Note that sample mean $\overline{X}$ is a random variable and an unbiased estimation of $X$'s mean $\mu$.

The following theorem provides a closed-form formula for estimation of detection rate when the adversary uses sample mean as the feature statistic.

Theorem 3.1: Using sample mean as the classification feature gives rise to an estimated detection rate

$$v_{\overline{X}} \approx 1 - \frac{1}{\sqrt{2\left(1/\sqrt{r} + \sqrt{r}\right)}} \tag{3.17}$$

where $r$ is defined in (3.15).

The proof of Theorem 3.1 can be found in the first part of Section 3.5.1. From Theorem 3.1, the following observations can be made:

1. The detection rate in (3.17) is independent on sample size *n*. That is, when sample mean is used as feature statistic, changing the sample size has no impact on detection rates.
2. As shown in the second part of Section 3.5.1, the detection rate $v_{\bar{X}}$ is an increasing function of *r*, where $r > 1$. That is, the smaller *r*, the lower the corresponding detection rate. When $r = 1$, the detection rate reaches 50%. Note that for such a system with two possible payload traffic rates, the detection rate for the adversary is lower-bounded at 50% corresponding to random guessing. In reality, $r = 1$ may occur when $\sigma_T^2$ is sufficiently large in comparison with $\sigma^2_{gw}$. This corresponds to the case when the VIT padding is used.

3.3.2.3. Detection Rate Formula for Sample Variance

Let $\{X_1, X_2, \ldots, X_n\}$ be a sample of size *n* from the distribution of random variable *X*. The *sample variance Y* is defined as follows

$$Y = \frac{\sum_{i=1}^{n} (X_i - m)^2}{n-1} \tag{3.18}$$

Note that sample variance *Y* is a random variable, and an unbiased estimation of *X*'s variance [31].

Recall that $\sigma_h^2$ is the variance of padded traffic's PIAT conditioned on the high payload traffic rate and $\sigma_l^2$ the variance of padded traffic's PIAT conditioned on the low payload traffic rate. $\sigma_h^2$ is slightly larger than $\sigma_l^2$, which is validated by our experiments in Section 3.4.1. Based on these observations, the following theorem provides a closed-form formula for estimation of detection rate when the adversary uses sample variance as the feature statistic.

Theorem 3.2**:** Using sample variance as the classification feature gives rise to an estimated detection rate

$$v_Y \approx \max\left(1-\frac{C_Y}{n-1}, 0.5\right) \tag{3.19}$$

where *C* is calculated as follows:

$$C_Y = \frac{1}{2\left(1-\frac{1}{r-1}\log r\right)^2} + \frac{1}{2\left(\frac{r}{r-1}\log r - 1\right)^2} \tag{3.20}$$

and *r* is defined in (3.15).

The proof of Theorem 3.2 can be found in the first part of Section 3.5.2. From Theorem 3.2, the following observations can be made:

1. The detection rate $v_Y$ is an increasing function in terms of sample size *n*. When $n\to\infty$, the detection rate is 100%. This means that if the payload traffic lasts for a long time at one rate, either low or high, and the adversary obtains such a sample,

the adversary may detect the payload traffic rate by using sample variance as a statistical feature.

2. As shown in the second part of Section 3.5.2, the detection rate $v_Y$ is an increasing function of $r$ defined in (3.15), where $r \geq 1$. That is, the smaller $r$, the lower the corresponding detection rate. When $r = 1$, the detection rate is 50%. This corresponds to the case of VIT padding with sufficiently large $\sigma_T^2$. This suggests that although the adversary may use a large sample size to detect the payload rate by sample variance, using a VIT padding with a large interval variance can make such an attack impossible, since no payload traffic can last very long at a fixed rate in practice, and the adversary cannot obtain a sufficiently large sample.

3.3.2.4. Detection Rate Formula for Sample Entropy

While there are many empirical entropy estimators available, it is generally difficult to obtain those estimators' PDFs. In this work, we take advantage of the relationship between entropy and variance of a normal distribution in order to describe sample entropy's effectiveness as the feature statistic. We will then use an empirical robust histogram-based entropy estimator for our analysis. The following theorem provides a closed-form formula for estimation of detection rate when the adversary uses sample entropy as the feature statistic.

Theorem 3.3**:** Using sample entropy as the classification feature gives rise to an estimated detection rate

$$v_{\tilde{H}} \approx \max\left(1-\frac{C_H}{n}, 0.5\right) \tag{3.21}$$

where $C_H$ is calculated as follows:

$$C_{\tilde{H}} = \frac{1}{2\left(\log\left(\frac{r}{r-1}\log r\right)\right)^2} + \frac{1}{2\left(\log\left(\frac{r-1}{\log r}\right)\right)^2} \tag{3.22}$$

and r is defined in (3.15).

The proof of Theorem 3.3 can be found in the first part of Section 3.5.3. From Theorem 3.3 we can make a similar set of observations to that of the case of sample variance.

1. Detection rate $v_{\tilde{H}}$ is an increasing function in terms of sample size *n*. That is, when the adversary obtains a larger sample, the detection rate will approach 100%.
2. As shown in the second part of Section 3.5.3, the detection rate $v_{\tilde{H}}$ is an increasing function of *r* defined in (3.15), where $r \geq 1$. When $r = 1$, the detection rate reaches 50%. In reality, this may occur when $\sigma_T^2$ is sufficiently large. This corresponds to the case when VIT padding with sufficiently large $\sigma_T^2$ is used.

From statistical knowledge, we know sample variance is very sensitive to outliers[3]. In order for empirical estimation of sample entropy to be robust against outliers, we use

[3] An outlier is an observation that lies an abnormal distance from other values in the sample of the padded traffic PIAT.

the method developed in [32]: First, we create a histogram of the PIAT sample for a given bin size (say, $\Delta h$). Then, according to [32], the differential entropy estimator of a random variable $X$'s continuous distribution is given by

$$\tilde{H} \approx \sum_i \frac{k_i}{n} \log \frac{k_i}{n} + \log \Delta h \tag{3.23}$$

where $n$ is the sample size, $k_i$ is the number of sample points in the $i^{th}$ bin, and $\Delta h$ is the histogram's bin size. If a constant bin size is used throughout the experiment, term $log\Delta h$ in (3.23) is a constant and hence does not influence the recognition result. It can, therefore, be discarded, and the entropy estimation formula then simplifies to the following:

$$\tilde{H} \approx \sum_i \frac{k_i}{n} \log \frac{k_i}{n} \tag{3.24}$$

Note that $\Delta h$, the histogram's bin size, plays an important role in the entropy estimation. As the bin size approaches positive infinity, all the estimated entropy approaches zero. In Theorem 3.4, we provide formulae of the optimal bin size by minimizing the mean square error (MSE) of the entropy estimation. Please see Section 3.5.4 for the proof.

Theorem 3.4**:** The optimal bin size $\Delta h$ for the histogram-based entropy estimator can be calculated as follows:

$$\Delta x = \frac{6\sigma}{I} \tag{3.25}$$

where $\sigma$ is the standard deviation of the underlying distribution and the number of bin $I$ can be calculated as follows:

$$(I-1)I^2 = 3n \tag{3.26}$$

This entropy estimator is robust in the sense that it is based on probability weighted sum. Generally, the probability that outliers occur is small. Thus, the probability weight reduces the noise's impact on the entropy estimation. Moreover, from the discussion in [32] and our experiments, we found that this histogram-based entropy estimator matches the value predicted by Theorem 3.3.

## 3.4. Evaluations

In this section, we report results on evaluating system security in terms of detection rate. The evaluations will be based on both theoretical analysis (from the previous section) and experiments.

In our experiments, we let the adversary use a high-performance network analyzer, such as Agilent's J6841A [33], to dump the padded traffic for traffic analysis. A series of experiments were carried out. In terms of padded traffic type, we measure both systems with CIT and VIT padding. In terms of experimental environments, we consider the following cases: a) a laboratory environment, b) a campus network, and c) a wide area network.

$GW_A$ and $GW_B$ in Figure 1 are installed with TimeSys Linux/Real-Time [29]. Both CIT and VIT paddings use a timer with interrupt interval mean equal to 10ms. The payload has two rate states: 10 packets per second (pps) and 40pps. We assume both rates occur in equal probability, that is, $P(\omega_l)=P(\omega_h)=50\%$.

### 3.4.1. Experiments in a Laboratory Environment

The advantage of performing the experiments in a laboratory environment is that we can control the cross traffic over the network. The disadvantage is that the generated cross traffic may not have the same characteristics of those in a real network. Nevertheless, our experiment setup is shown in Figure 4.

The two gateways are connected by a Marconi ESR-5000 enterprise switch router [34]. Subnet C is connected to the router as the cross traffic (noise) generator while the cross traffic receiver is located in Subnet D. Note that the cross traffic shares the outgoing link of the router, creating a situation where the cross traffic has an impact on the padded traffic. We run several cases of experiments as described below.

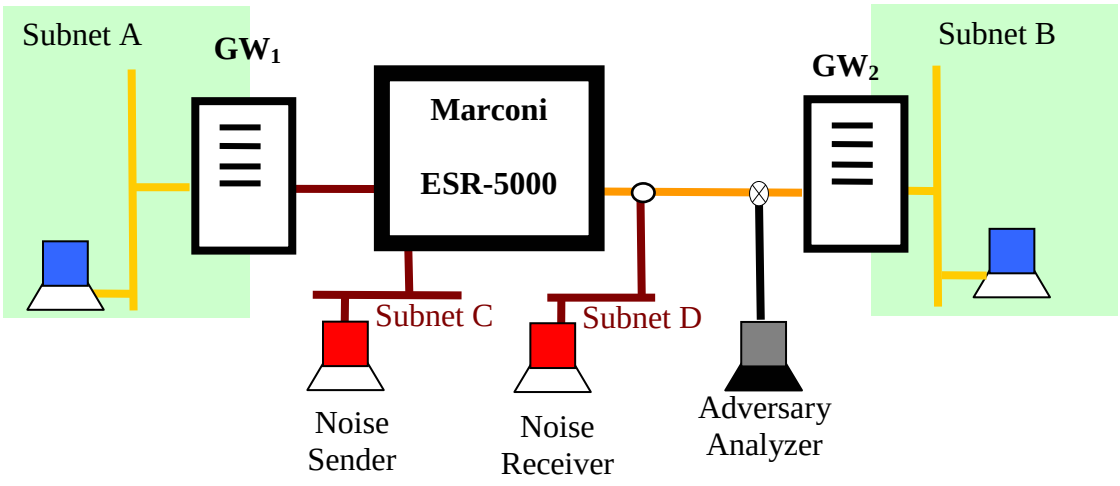



Figure 4. Experiment Setup in Laboratory for Passive Attacks

3.4.1.1. The Case of Zero Cross Traffic

For the case of no cross traffic, the workstation in subnet C does not transmit, and the router only deals with the padded traffic from $GW_A$. That is, $\sigma_{net}^2$ in (3.15) is 0. Hence, the variance ratio $r$ becomes

$$r = \frac{\sigma_T^2 + \sigma_{gw,h}^2}{\sigma_T^2 + \sigma_{gw,l}^2} \tag{3.27}$$

This situation is the best case for the adversary as he can observe traffic with minimum disturbance. Hence, this is the worst case for the administrator that wants to prevent traffic analysis attacks.

3.4.1.1.1. CIT Link Padding

First, we analyze systems that use CIT link padding. That is, $\sigma_T^2$ is 0. Hence, (3.27) is further simplified as follows:

$$r = \frac{\sigma_{gw,h}^2}{\sigma_{gw,l}^2} \tag{3.28}$$

From the theorems in Section 3.3.2, we see that the detection rate is a function of sample size *n* and ratio *r*.

Figure 5 (a) shows the distributions of padded traffic's PIAT under low rate (10pps) and high rate (40pps) payload traffic. We have the following observations:

1. The two distributions are almost bell-shaped. This (partially) validates our assumption that the padded traffic's PIAT has a normal distribution.
2. The means of padded traffic's PIAT under different rates of payload traffic are virtually identical. This is also consistent with the assumption made in 3.3.2.1.
3. The two distributions are slightly different. The variance of padded traffic's PIAT conditioned on the high-rate payload traffic, $\sigma^2_{gw,h}$ in (3.14) is slightly larger than the variance of padded traffic's PIAT conditioned on the low-rate payload traffic, $\sigma^2_{gw,l}$ in (3.12). This implies

$$r = \frac{\sigma_{gw,h}^2}{\sigma_{gw,l}^2} > 1 \tag{3.29}$$

Figure 5 (b) shows both empirical and theoretical curves of detection rate for different feature statistics. We have the following observations:

1. The empirical detection rate curves coincide well with their theoretical counterparts. This validates our theoretical model. The empirical curve of sample variance is a little lower than the theoretical one because sample variance is very sensitive to outliers in the data.
2. The detection rate of sample mean is almost *50*%. Sample mean is not an effective feature for the adversary.
3. On the other hand, as the sample size increases, detection rates for both sample variance and sample entropy increase as predicted by our Theorem 3.2 and Theorem 3.3. At sample size of 1,000, both features achieve the detection rate of almost *100*%. This means that CIT padding fails if the adversary uses sample variance or sample entropy as feature statistic. Generally speaking, sample entropy performs empirically better than sample variance in terms of detection rate.

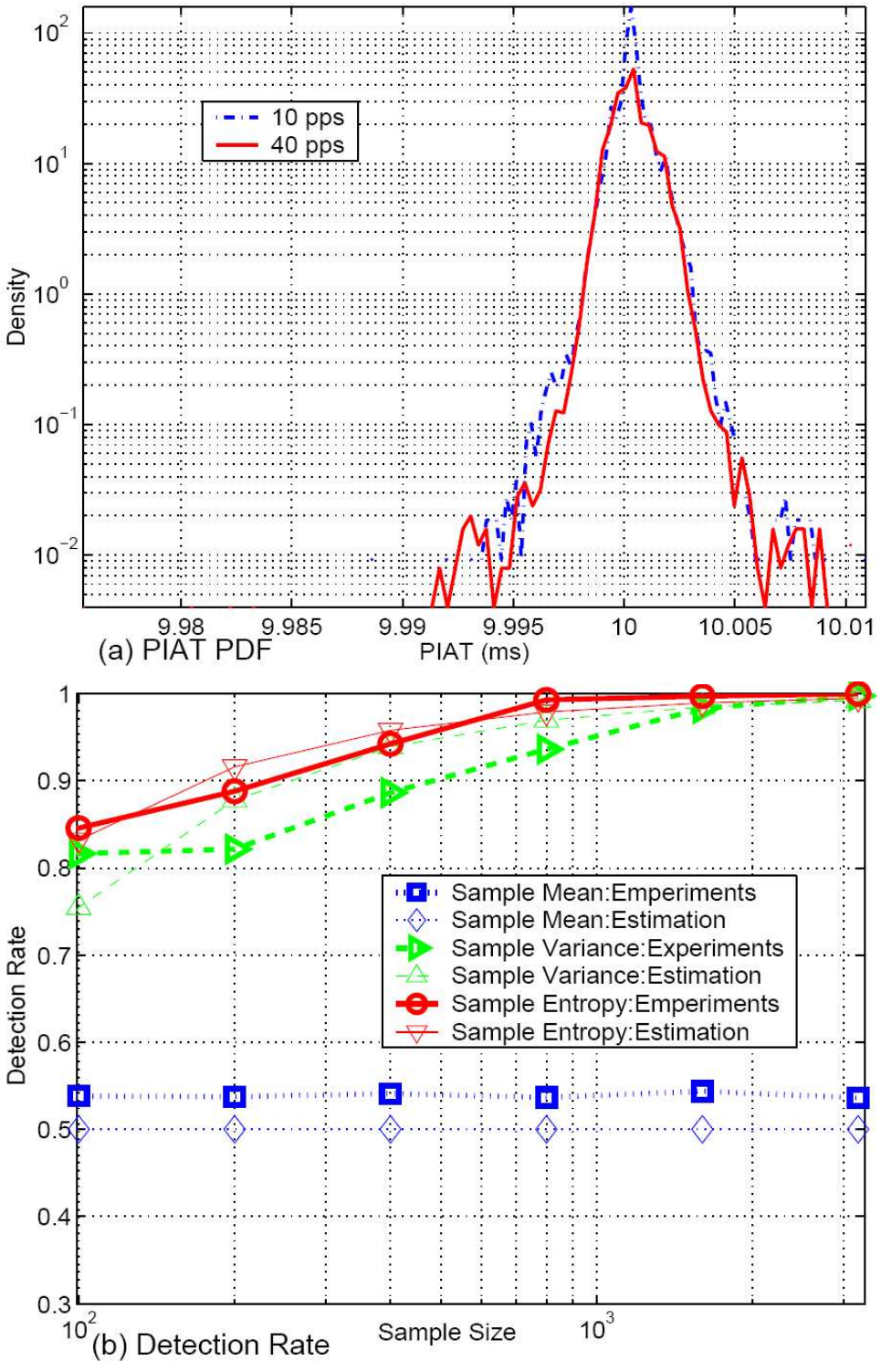



Figure 5. CIT Padding with Zero Cross Traffic

3.4.1.1.2. VIT Link Padding

Recall from (3.27), the variance ratio $r$ in (3.15) is given by

$$r = \frac{\sigma_T^2 + \sigma_{gw,h}^2}{\sigma_T^2 + \sigma_{gw,l}^2}$$

where $\sigma_T^2 \geq 0$ since we are using VIT padding.

Theorems in 3.3.2 show that when $r$ approaches *1*, the detection rates approach 50% for all the three feature statistics. We note that for CIT padding, the value of $r$decreases with increasing values of $\sigma_T^2$. Figure 6 (a) displays the empirical curves of detection rate in terms of $\sigma_T$ for a fixed sample size of *2,000*. We can see that when $\sigma_T$ increases, the detection rate quickly drops and approaches 50%, as expected. Clearly, a system with VIT padding performs better (i.e., with lower detection rate) than one with CIT padding.

In any case, as shown in (3.17) and (3.21), when the size of sample increases, the detection rate increases as well. An interesting question is: How large does a sample have to be in order for the adversary to have sufficiently high probability in making a correct detection? Let $n(p)$ be the sample size that can achieve a detection rate of $p$ percent. Figure 6 (b) provides the theoretical curve of $n(99\%)$ vs. $\sigma_T$. We can see that with a reasonable value of $\sigma_T$, the sample size needs to be extremely large in order to achieve a 99% detection rate. For example, when the timer interval standard deviation $\sigma_T$ $=1$ms, to achieve the detection rate of 99%, the sample size has to be greater than $10^{11}$. It is virtually impossible for an attacker to obtain such a large sample. This clearly shows the effectiveness of VIT padding.

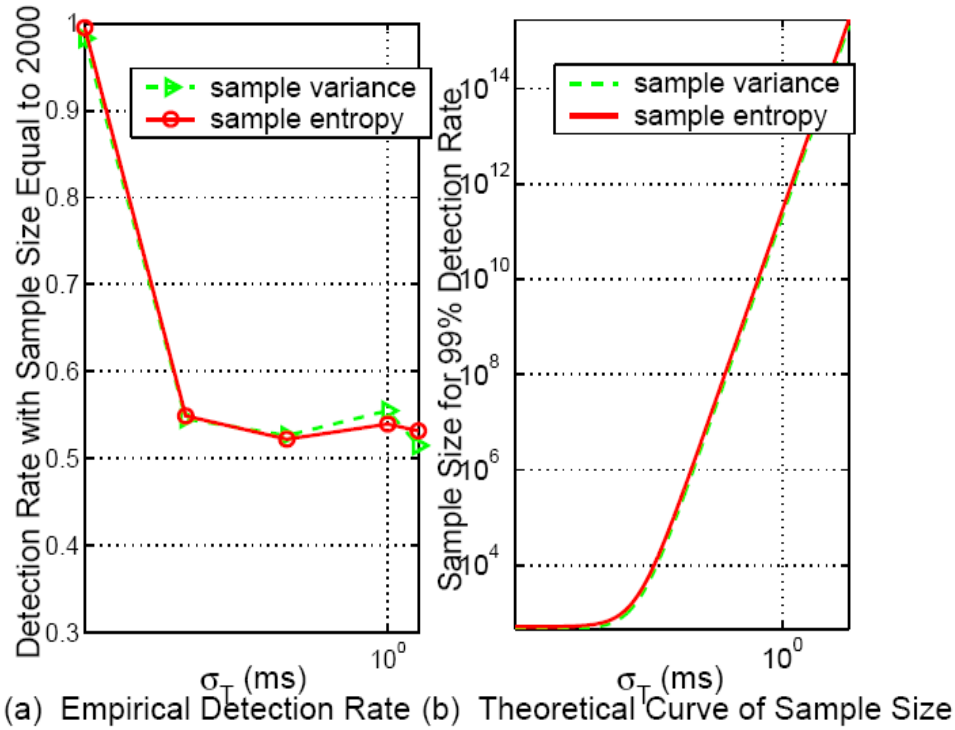



Figure 6. VIT Padding - Detection Rate vs. Sample Size

3.4.1.2. The Case of Non-Zero Cross Traffic

Recall that the case of zero-cross traffic is the best case for the adversary. As VIT has shown to be effective in the case of zero cross traffic, we will no longer have to consider systems with VIT padding here since VIT has been shown to be effective even for the adversary's best-case scenario (zero cross traffic with a line tap very near the sender gateway). We thus concentrate on the system with CIT padding. In a system with cross

traffic, $\sigma_{net}^2$ in (3.15) may no longer be zero. As for CIT padding, where $\sigma_T^2=0$, the variance ratio in (16) now becomes

$$r = \frac{\sigma_{net}^2 + \sigma_{gw,h}^2}{\sigma_{net}^2 + \sigma_{gw,l}^2} \tag{3.30}$$

We observe that $r$ decreases with increasing $\sigma_{net}^2$, resulting in a low detection rate for all feature statistics. Thus, the bigger $\sigma_{net}^2$, the smaller the detection rate.

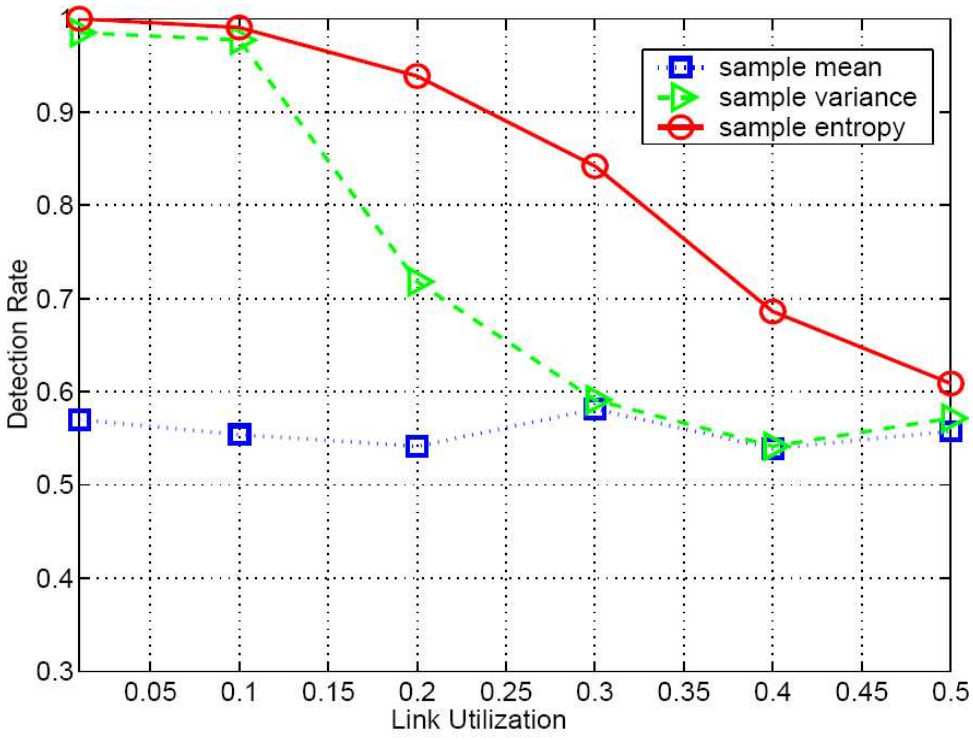



Figure 7. Empirical Detection Rate with Cross Traffic in Laboratory

In the experiments described here, cross traffic generated from within subnet C causes router congestion, which in turn affects the observation by the adversary. Figure 7 shows how the detection rate is impacted by the amount of cross traffic. We can make the following observations:

1. Note that the PIAT for the padded traffic remains at 10ms. Hence, the amount of cross traffic is directly proportional to the utilization of the link shared between Subnet B and Subnet D. The data shows that as the link utilization increases, the detection rate of sample entropy and sample variance decreases. Intuitively, this is because the crossover traffic between Subnet C and Subnet D interferes with the padded traffic between $GW_A$ and $GW_B$, and $\sigma^2_{net}$ increases with the shared link's utilization. The sample mean's detection rate remains low, as expected.
2. We observe that sample entropy results in a better detection rate than sample variance does. It can be perceived that, with the increase of the shared link's utilization, outliers have more chance of occurring. Sample variance is much more sensitive to outliers and, hence, it has a low detection rate.
3. Even with the link utilization of *40*%, sample entropy can still have a detection rate of about *70*%, implying that CIT padding may still not be as effective in this kind of situation.

### 3.4.2. Experiments over Campus and Wide Area Networks

Figure 8 shows the setup for the experiments discussed in this subsection. Figure 8 (a) is a setup for experiments over the Texas A&M Campus Network. That is, the padded traffic goes through Texas A&M campus network before it reaches the receiver's gateway. Figure 8 (b) is a setup for experiments over the Internet between Ohio State University and Texas A&M University. Here, the sender workstation and the sender gateway are located at Ohio State University. The padded traffic goes through the Internet and arrives at Texas A&M University, where the receiver gateway and the receiver's workstation are located. In both cases, the observation point of the adversary is located directly in front of the receiver gateway and thus maximally far from the sender. We note that in this case, the path from the sender's workstation to the receiver's workstation spans over 15 routers.

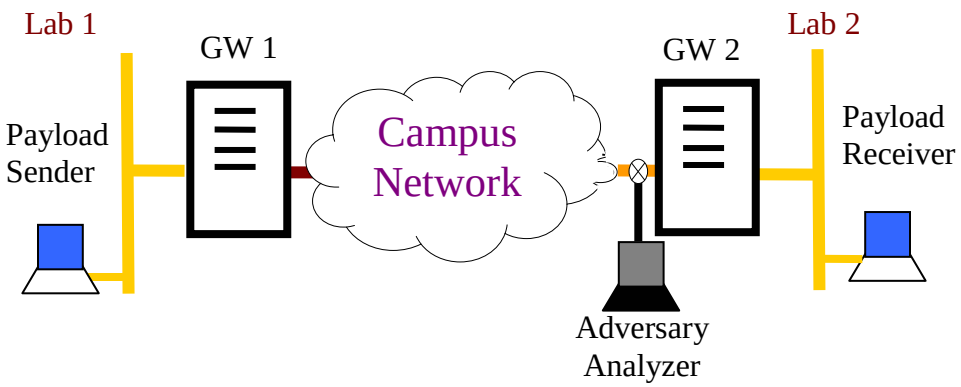
In each case, we collect data continuously for a complete day (24 hours). The data for the case of Texas A&M campus network was collected on March 24, 2003 while the data for the wide are network case was collected on March 26, 2003. Figure 9 (a) and Figure 9 (b) display the detection rate throughout the observation period when sample size is *1000*. We make the following observations:

1. When the padded traffic traverses only the Texas A&M campus network, the detection rates of sample entropy and sample variance are high for almost all the time we collected data. This means that over a medium-size enterprise network like the Texas A&M campus, the crossover traffic has limited influence on the

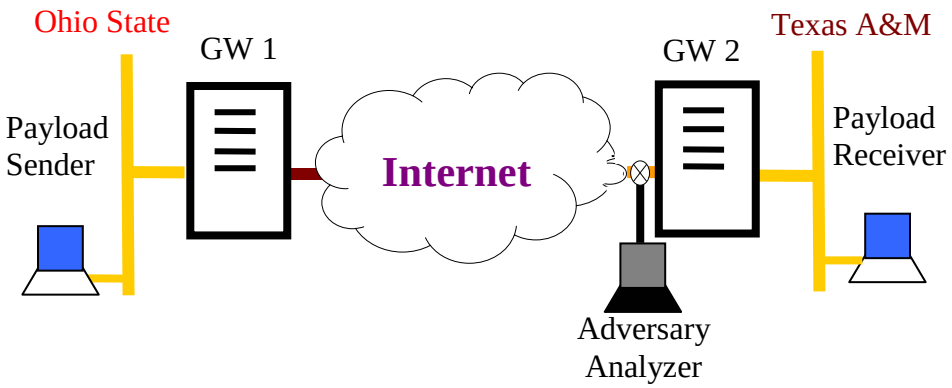
padded traffic's PIAT. Consequently, we would not recommend CIT padding for use in such an environment.

2. When the padded traffic traverses more network elements, such as the span of the Internet between Ohio State University and Texas A&M University, the detection rates are low. This is because the padded traffic experiences congestion at a large number of routers and switches, and its PIAT is seriously distorted with a relatively large $\sigma^2_{net}$.
3. In the case of wide area networks, sample entropy and sample variance can experience over *65*% detection rates during periods of relatively low network activity (such as at 2:00AM). This means that CIT padding may still be sufficiently safe even if the adversary is very remote.

Experimental results in Figure 9 match our theoretical analysis in Section 3.3.2. This further validates the correctness of our theoretical framework for analyzing the passive link load analysis attack and their countermeasures.




(a) Experiment Setup within Texas A&M University




(b) Network Setup between Ohio and Texas

Figure 8. Experiment Setup over Campus and the Internet for Passive Attacks

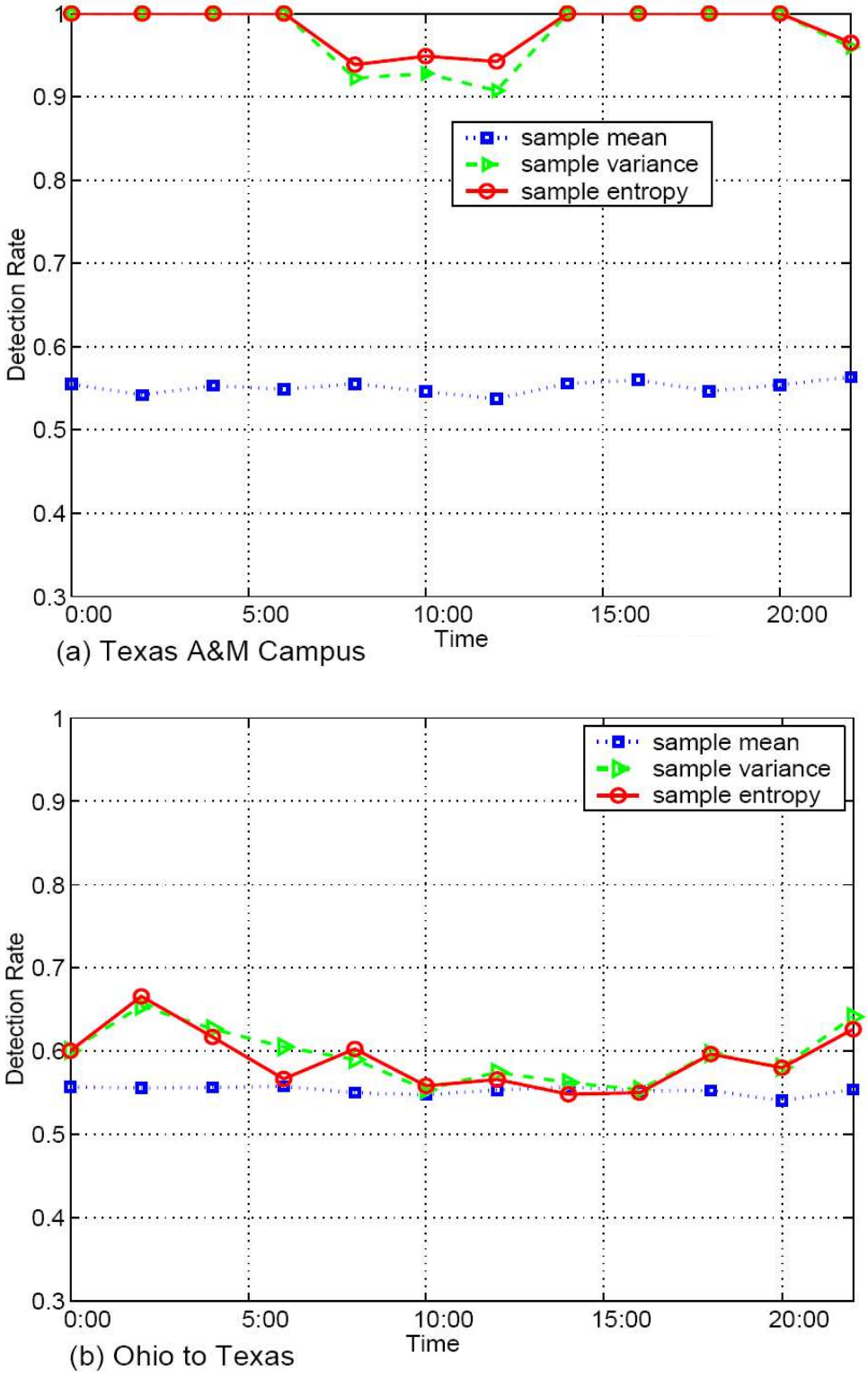



Figure 9. Empirical Detection Rate for Passive Attacks over Campus and Internet

## 3.5. Theorem Proof

In this subsection, we prove theorems introduced in Section 3.3.2.

### 3.5.1. Proof of Theorem 3.1

Theorem 3.1: Using sample mean as the classification feature gives rise to an estimated detection rate

$$v_{\bar{X}} \approx 1 - \frac{1}{\sqrt{2\left(1/\sqrt{r} + \sqrt{r}\right)}} \tag{3.31}$$

where $r$ is defined in (3.15).

*Proof:* The distribution of sample mean for a normal distribution $N(\mu, \sigma^2)$ is still a normal one, $N(\mu, \sigma^2/n)$. Thus, sample mean $\bar{X}_l$ for the case of the payload traffic rate being low has a normal distribution

$$f_l(x) = N\left(\mu, \sigma_1^2\right) = N\left(\mu, \frac{\sigma_l^2}{n}\right) \tag{3.32}$$

Similarly, sample mean $\bar{X}_h$ for the high payload traffic rate has a normal distribution

$$f_h(x) = N\left(\mu, \sigma_2^2\right) = N\left(\mu, \frac{\sigma_h^2}{n}\right) \tag{3.33}$$

Since $\bar{X}_l$ and $\bar{X}_h$ are normally distributed, we can use the Bhattacharyya bound [27] to estimate the error rate as follows:

$$\varepsilon_{\bar{X}} \le \sqrt{P(\omega_l)P(\omega_h)}\int \sqrt{f(x \mid \omega_l) f(x \mid \omega_h)}dx \tag{3.34}$$

Substituting (3.32), (3.33), and $P(\omega_l)=P(\omega_h)=0.5$ into (3.34) and carrying out the integration, we have

$$\varepsilon_{\bar{X}} \le \frac{1}{2}\exp(-k) \tag{3.35}$$

where

$$k = \frac{1}{2}\ln\frac{\frac{\sigma_l^2+\sigma_h^2}{2}}{\sqrt{\sigma_l^2\sigma_h^2}} \tag{3.36}$$

Substituting (3.15) into (3.36) and some rearranging, we have

$$k = \frac{1}{2}\ln\frac{\left(1/\sqrt{r}+\sqrt{r}\right)}{2} \tag{3.37}$$

Substituting (3.37) into (3.35), the error rate is given by

$$\varepsilon_{\bar{X}} \le \frac{1}{\sqrt{2\left(1/\sqrt{r}+\sqrt{r}\right)}} \tag{3.38}$$

The detection rate $v_{\bar{X}}$ then satisfies the following:

$$v \quad = \quad 1-\varepsilon \tag{3.39}$$

$$\geq \quad 1-\frac{1}{\sqrt{2\left(1/\sqrt{r}+\sqrt{r}\right)}} \tag{3.40}$$

Thus, we can use the lower bound of (3.40) as the estimation of detection rate by sample mean. The theorem is proven.

In the following, we prove that detection rate *r* is an increasing function of *r* when the adversary use sample mean as the feature statistic. To prove that $v_{\bar{X}}$ increases with *r*, we need to prove that the term $t(r)=\sqrt{2\left(1/\sqrt{r}+\sqrt{r}\right)}$ in (3.17) increases with *r*.

$$dt/dr=\frac{1}{2\sqrt{r}}\left(1-\frac{1}{\sqrt{r}}\right) \tag{3.41}$$

Since $r \geq 1$, $dt/dr \geq 0$. Thus $t(r)$ increases with *r*, and the decision rate $v_{\bar{X}}$ increases with *r*.

### 3.5.2. Proof of Theorem 3.2

Theorem 3.2**:** Using sample variance as the classification feature gives rise to an estimated detection rate

$$v_Y \approx \max\left(1-\frac{C_Y}{n-1},0.5\right) \tag{3.42}$$

where *C* is calculated as follows:

$$C_Y = \frac{1}{2\left(1-\frac{1}{r-1}\log r\right)^2} + \frac{1}{2\left(\frac{r}{r-1}\log r - 1\right)^2} \tag{3.43}$$

and $r$ is defined in (3.15).

*Proof:* Denote $\chi_{n-1}^2$ as a random variable with a *chi square* distribution $f_{\chi_{n-1}^2}(x)$ with freedom *n-1*, which is defined as follows,

$$f_{\chi_{n-1}^2}(x) = \frac{x^{\frac{n-1}{2}} \exp\left(-\frac{x}{2}\right)}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)} \tag{3.44}$$

where $x > 0$. Denote $Y$ as the random variable of sample variance. Then $(n-1)Y/\sigma^2$ has a chi square distribution with freedom *n-1* [31]. That is

$$\chi_{n-1}^2 = \frac{n-1}{\sigma^2} Y \tag{3.45}$$

From (3.45), we get

$$Y = \frac{\sigma^2}{n-1}\chi_{n-1}^2 \tag{3.46}$$

From chi square's properties, we have sample variance's mean $\overline{Y}$ as

$$\overline{Y} = \sigma^2 \tag{3.47}$$

and its variance *var*($Y$) as

$$\mathrm{var}(Y) = \frac{2\sigma^4}{n-1} \tag{3.48}$$

To get sample variance's PDF at sample size *n*, we first compute its distribution function

$$P(Y < y) = P\left(\frac{\sigma^2}{n-1}\chi^2_{n-1} < y\right) \tag{3.49}$$

$$= P\left(\chi^2_{n-1} < \frac{n-1}{\sigma^2}y\right) \tag{3.50}$$

Differentiating the two sides of (3.50), we have the density function

$$f_Y(y) = f_{\chi^2_{n-1}}\left(\frac{n-1}{\sigma^2}y\right)\frac{n-1}{\sigma_l^2} \tag{3.51}$$

We denote $Y_l$ as the random variable of sample variance of padded traffic's PIAT at the low rate payload traffic. Substituting (3.44) into (3.51), we then derive $Y_l$'s density function $f_{Y_l}(y)$

$$f_{Y_l}(y) = \frac{\left(\frac{n-1}{\sigma_l^2}y\right)^{\frac{n-1}{2}-1}\exp\left(-\frac{n-1}{2\sigma_l^2}y\right)}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}\frac{n-1}{\sigma_l^2} \tag{3.52}$$

Similarly, $Y_h$ is the random variable of sample variance of padded traffic's PIAT at high rate payload traffic, and its density function $f_{Y_h}(y)$ is

$$f_{Y_h}(y)=\frac{\left(\frac{n-1}{\sigma_h^2}y\right)^{\frac{n-1}{2}-1}\exp\left(-\frac{n-1}{2\sigma_h^2}y\right)}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}\frac{n-1}{\sigma_h^2} \tag{3.53}$$

To get the detection rate, we calculate the cross point $y_c$ of $f_{Y_l}(y)$ and $f_{Y_h}(y)$

$$f_{Y_l}(y_c)= f_{Y_h}(y_c) \tag{3.54}$$

After lengthy arithmetic operations, we have

$$y_c = \log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2\sigma_l^2}{\sigma_h^2-\sigma_l^2} \tag{3.55}$$

Now we use Chebynov inequality for the estimation of error rate if the adversary uses sample variance as the feature statistic. The distance from the mean of $Y_l$ to the cross point $y_c$ is denoted as $D_l$

$$D_l = y_c - \bar{Y}_l \tag{3.56}$$

Substituting (3.47) and (3.55) into (3.56), we have

$$D_l = \left(\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2}{\sigma_h^2-\sigma_l^2}-1\right)\sigma_l^2 \tag{3.57}$$

Denoting $c_l$ as the ratio of $D_l$ to the standard deviation of $Y_l$

$$c_l = \frac{D_l}{\sqrt{\mathrm{var}(Y_l)}} \tag{3.58}$$

Substituting (3.48) and (3.57) into (3.58), we have

$$c_l = \frac{\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2}{\sigma_h^2 - \sigma_l^2} - 1}{\sqrt{\frac{2}{n-1}}} \tag{3.59}$$

Similarly, denoting $D_h$ as the distance from the mean of $Y_h$ to the cross point $y_c$,

$$D_h = \sigma_h^2 - y_c = \sigma_h^2\left(1 - \log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_l^2}{\sigma_l^2 - \sigma_l^2}\right) \tag{3.60}$$

Then $c_h$, the ratio of $D_h$ and the standard deviation of $Y_h$, can be calculated as follows

$$c_h = \frac{1 - \log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_l^2}{\sigma_h^2 - \sigma_l^2}}{\sqrt{\frac{2}{n-1}}} \tag{3.61}$$

When sample size $n$ is big (>40), we can assume that a chi square PDF is symmetrical. Thus from Chebynov inequality, we can get the error rate $e_Y$

$$e_Y \le \frac{\frac{1}{2c_l^2} + \frac{1}{2c_h^2}}{2} \tag{3.62}$$

Thus, the detection rate $v_Y$ can be calculated as

$$v_Y = 1 - e_Y \tag{3.63}$$

$$\geq 1 - \frac{1}{4c_l^2} - \frac{1}{4c_h^2} \tag{3.64}$$

Substituting (3.15), (3.59) and (3.61) into (3.64), we have

$$v_Y \geq 1 - \frac{C_Y}{n-1} \tag{3.65}$$

where

$$C_Y = \frac{1}{2\left(1 - \frac{1}{r-1}\log r\right)^2} + \frac{1}{2\left(\frac{r}{r-1}\log r - 1\right)^2} \tag{3.66}$$

Thus, we use the lower bound of $v_Y$ as the estimation of the detection rate. Since that $v_Y$ must be greater than 50%, we can get (3.19). The theorem is proved.

In the following, we prove that $v_Y$ is an increasing function of $r$. That is, we need to prove that $C_Y$ in (3.20) is a decreasing function of $r$. For terms in (3.20), we have the following denotations

$$C_{Y1}(r) = \frac{\log r}{r-1} \tag{3.67}$$

$$C_{Y2}(r) = \frac{r}{r-1} \tag{3.68}$$

If $C_{Y1}$ is decreasing function and $C_{Y2}$ is an increasing function, then $C_Y$ is a decreasing function.

We first prove that $C_{Y1}$ is a decreasing function. Let $r=e^x$, we have

$$C_{Y1}(r) = c_{y1}(x) = \frac{x}{e^x - 1} \tag{3.69}$$

$$\frac{dC_{Y1}}{dr} = \frac{dc_{y1}}{dr} \tag{3.70}$$

$$= \frac{dc_{y1}}{dx}\frac{dx}{dr} \tag{3.71}$$

$$= \frac{e^x - 1 - xe^x}{r\left(e^x - 1\right)^2} \tag{3.72}$$

Since $r>1$, $x>0$, the denominator of (3.72) is greater than $0$. We have the Taylor expansion of the numerator of (3.72) as follows

$$e^x - 1 - xe^x = \sum_{n=1}^{\infty}\left(\frac{1}{n!} - \frac{1}{(n-1)!}\right)x^n \tag{3.73}$$

Since

$$\frac{1}{n!} - \frac{1}{(n-1)!} < 0 \tag{3.74}$$

So

$$e^x - 1 - xe^x < 0 \tag{3.75}$$

Thus

$$\frac{dC_{Y1}}{dr} = \frac{dc_{y1}}{dr} < 0 \tag{3.76}$$

$C_{Y1}$ is a decreasing function in terms of $r$.

Now we prove $C_{Y2}$ is an increasing function of $r$. Let $r = e^x$, we have,

$$C_{Y2}(r) = \frac{r \log r}{r-1} = c_{y2}(x) = \frac{xe^x}{e^x - 1} \tag{3.77}$$

and

$$\frac{dC_{Y2}}{dr} = \frac{dc_{y2}}{dr} = \frac{e^{2x} - e^x - xe^x}{r\left(e^x - 1\right)^2} \tag{3.78}$$

Since r>1, x>0, the denominator of (3.78) is greater than *0*. The Taylor expansion of the numerator of (3.78) is as follows

$$e^{2x} - e^x - xe^x = \sum_{n=1}^{\infty}\left(\frac{2^n - 1 - n}{n!}\right)x^n \tag{3.79}$$

Since

$$\forall n > 0, \frac{2^n - 1 - n}{n!} \geq 0 \tag{3.80}$$

we have

$$e^{2x} - e^{x} - xe^{x} \geq 0 \tag{3.81}$$

So

$$\frac{dC_{Y2}}{dr} = \frac{dc_{y2}}{dr} > 0 \tag{3.82}$$

and $C_{Y2}(r)$ is an increasing function of $r$.

Thus, we have proved that is an increasing function in terms of $r$.

3.5.3. Proof of Theorem 3.3

Theorem 3.3: Using sample entropy as the classification feature gives rise to an estimated detection rate

$$v_{\tilde{H}} \approx \max\left(1 - \frac{C_H}{n}, 0.5\right) \tag{3.83}$$

where $C_H$ is calculated as follows:

$$C_{\tilde{H}} = \frac{1}{2\left(\log\left(\frac{r}{r-1}\log r\right)\right)^2} + \frac{1}{2\left(\log\left(\frac{r-1}{\log r}\right)\right)^2} \tag{3.84}$$

and $r$ is defined in (3.15).

*Proof:* A normal distribution's differential entropy can be calculated as

$$H = \frac{\log 2\pi\sigma^2 + 1}{2} \tag{3.85}$$

Here we use sample variance *Y* defined in (3.18) to estimate sample entropy $\tilde{H}$

$$\tilde{H} = \frac{\log 2\pi Y + 1}{2} \tag{3.86}$$

To get sample entropy's PDF, we first derive its distribution,

$$P\left(\tilde{H} < h\right) = P\left(\frac{\log 2\pi Y + 1}{2} < h\right) = P\left(Y < \frac{e^{2h-1}}{2\pi}\right) \tag{3.87}$$

Differentiating two sides of (3.87), we get sample entropy's PDF

$$f_{\tilde{H}}(h) = f_Y\left(\frac{e^{2h-1}}{2\pi}\right) = P\left(Y < \frac{e^{2h-1}}{2\pi}\right)\left(\frac{e^{2h-1}}{2\pi}\right)' \tag{3.88}$$

Denote $H_l$ as the sample entropy of padded traffic's PIAT at the low-rate payload traffic, and $f_{\tilde{H}_l}(h)$ as $H_l$'s PDF. Denote $H_h$ as the sample entropy of padded traffic's PIAT at the high-rate payload traffic, and $f_{\tilde{H}_h}(h)$ as $H_h$'s PDF. To get the detection rate, we need to calculate the cross point $h_c$ of $f_{\tilde{H}_l}(h)$ and $f_{\tilde{H}_h}(h)$

$$f_{\tilde{H}_l}(h_c) = f_{\tilde{H}_h}(h_c)' \tag{3.89}$$

By lengthy arithmetic operations

$$h_c = \frac{\log\left[2\pi \log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2\sigma_l^2}{\sigma_h^2-\sigma_l^2}\right]+1}{2} \tag{3.90}$$

For the ease of estimation, we approximate sample entropy's mean as follows

$$E\left(\tilde{H}\right) = \frac{\log 2\pi\sigma^2 + 1}{2} \tag{3.91}$$

That is, we approximate the entropy estimator in (3.86) as an unbiased one.

Now we use Chebynov inequality for the error rate estimation. The distance from the mean of $H_l$ to the cross point $h_c$ is denoted as $D_l$

$$D_l = \frac{\log\left[2\pi \log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2\sigma_l^2}{\sigma_h^2-\sigma_l^2}\right]}{2} - \frac{\log 2\pi\sigma_l^2 + 1}{2} \tag{3.92}$$

$$= \frac{\log\left[\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2}{\sigma_h^2-\sigma_l^2}\right]}{2} \tag{3.93}$$

Using Taylor expansion over (3.86) and by appropriate approximation, we get

$$\mathrm{var}\left(\tilde{H}\right) \approx \frac{1}{2n} \tag{3.94}$$

(3.94) is the same result as for the histogram-based entropy estimator in [32]. Denoting $c_l$ as the ratio of $D_l$ to the standard deviation of $h_l$

$$c_l = \frac{\log\left[\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2}{\sigma_h^2-\sigma_l^2}\right]}{2\sqrt{\frac{1}{2n}}} \tag{3.95}$$

Similarly, denoting $D_h$ as the distance from the mean of $H_h$ to the cross point $h_c$ and $c_h$ as the ratio of $D_h$ to the standard deviation of $h_h$, we have

$$D_l = \frac{\log 2\pi\sigma_h^2 + 1}{2} - \frac{\log\left[2\pi\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2\sigma_l^2}{\sigma_h^2-\sigma_l^2}\right]+1}{2} \tag{3.96}$$

$$= \frac{\log\left[\frac{\sigma_h^2-\sigma_l^2}{\sigma_l^2\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)}\right]}{2} \tag{3.97}$$

$$c_h = \frac{\log\left[\frac{\sigma_h^2-\sigma_l^2}{\sigma_l^2\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)}\right]}{2\sqrt{\frac{1}{2n}}} \tag{3.98}$$

So error rate is calculated as

$$e_{\tilde{H}} \le \frac{1}{4c_l^2} + \frac{1}{4c_h^2} = \frac{C_{\tilde{H}}}{n} \tag{3.99}$$

where

$$C_{\tilde{H}} = \frac{1}{2\left[\log\left(\left(\log\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2}{\sigma_h^2-\sigma_l^2}\right)\right]^2} + \frac{1}{2\left[\log\left(\frac{\sigma_h^2-\sigma_l^2}{\sigma_l^2\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)}\right)\right]^2} \tag{3.100}$$

Substitute (3.15) into (3.100), we have

$$C_{\tilde{H}} = \frac{1}{2\left[\log\left(\frac{r}{r-1}(\log r)\right)\right]^2} + \frac{1}{2\left[\log\left(\frac{r-1}{\log r}\right)\right]^2} \tag{3.101}$$

Since

$$v_{\tilde{H}} \geq 1 - e_{\tilde{H}} \tag{3.102}$$

Substituting (3.102) into (3.99), we have

$$v_{\tilde{H}} \geq 1 - \frac{C_{\tilde{H}}}{n} \tag{3.103}$$

In this section, we use the lower bound of $v_{\tilde{H}}$ as the estimation of the detection rate by sample entropy. Consider that detection rate must be greater than 50% and we can get (3.21) in Theorem 3.3. The theorem is proved.

In the following we prove that $v_{\tilde{H}}$ is an increasing function in terms of *r*. That is, we need to prove that $C_{\tilde{H}}$ is a decreasing function of *r*. Denote

$$C_{H1}(r) = \frac{r-1}{\log r} \tag{3.104}$$

$$C_{H2}(r) = \frac{r}{r-1} \tag{3.105}$$

If $C_{H1}(r)$ in (3.104) and $C_{H2}(r)$ in (3.105) are increasing functions of $r$, $C_{\tilde{H}}$ is a decreasing function of $r$.

We can see that $C_{H1}(r)=1/ C_{Y1}(r)$, where $C_{Y1}(r)$ is defined in (3.67). Since $C_{Y1}(r)$ is a decreasing function of $r$, $C_{H1}(r)$ is an increasing function $r$. $C_{H2}(r)=C_{Y2}(r)$, where $C_{Y2}(r)$ is defined in (3.68). We have proved that $C_{Y2}(r)$ is an increasing function, so is $C_{H2}(r)$.

Thus $v_{\tilde{H}}$ is of $r$, where $r>1$.

3.5.4. Proof of Theorem 3.4

Theorem 3.4: The optimal bin size $\Delta h$ for the histogram-based entropy estimator can be calculated as follows:

$$\Delta x = \frac{6\sigma}{I} \tag{3.106}$$

where $\sigma$ is the standard deviation of the underlying distribution and the number of bin $I$ can be calculated as follows:

$$(I-1)I^2 = 3n \tag{3.107}$$

*Proof:* We now develop theory calculating the optimal bin size for histogram based entropy estimation. In [32], for a histogram based entropy estimation,

$$E\left(\tilde{H}\right) \approx H - \frac{I-1}{2n} + \frac{1}{24}\left(\frac{\Delta x}{\sigma}\right)^2 \tag{3.108}$$

where $\tilde{H}$ is the estimated entropy, $H$ is original entropy of random variable $X$, $I$ is the number of bins, $n$ is the sample size, $\Delta x$ is the bin size and $\sigma$ is $X$'s standard deviation.

If random variable $X$ is normally distributed,

$$\mathrm{var}\left(\tilde{H}\right) \approx \frac{1}{2}n^{-1} \tag{3.109}$$

Moreover, for a normally distributed random variable, since the probability of $x>3\sigma$ is very small and ignorable, we have the following relationship between bin size $\Delta x$, the number of bins $I$ and random variable $X$'s standard deviation $\sigma$,

$$6\sigma = I\Delta x \tag{3.110}$$

Thus

$$\frac{\Delta x}{\sigma} = \frac{I}{6} \tag{3.111}$$

We can calculate the mean square error (MSE) of the entropy estimation as follows,

$$E\left[\left(H - \tilde{H}\right)^2\right] = VAR\left(\tilde{H}\right) + \left[BIAS\left(\tilde{H}\right)\right]^2 \tag{3.112}$$

where

$$\left[BIAS\left(\tilde{H}\right)\right] = E\left(\tilde{H}\right) - H \tag{3.113}$$

Substitute (3.108) and (3.109) into (3.112),

$$E\left[\left(H-\tilde{H}\right)^2\right]=\frac{1}{2}n^{-1}+\left(\frac{1}{24}\left(\frac{\Delta x}{\sigma}\right)^2-\frac{I-1}{n}\right)^2 \tag{3.114}$$

To derive a minimum mean square error, the second term in (3.114) should be minimized,

$$\frac{1}{24}\left(\frac{\Delta x}{\sigma}\right)^2=\frac{I-1}{n} \tag{3.115}$$

Substituting (3.111) into (3.115), we have

$$\frac{1}{24}\left(\frac{6}{I}\right)^2=\frac{I-1}{2n} \tag{3.116}$$

Thus

$$(I-1)I^2=3n \tag{3.117}$$

$$\Delta x=\frac{6\sigma}{I} \tag{3.118}$$

## 3.6. Summary

While researchers have proposed link padding as an effective way to prevent traffic analysis, there was no systematic method to analyze the security of a system under the traffic analysis attacks before our study results were released. We provide an effective analysis model for the evaluation of different padding strategies aimed at camouflaging

the payload traffic rates under traffic analysis attacks. We define as our security metric detection rate, which is the probability that the rate of payload traffic is recognized. We believe that our analysis methods can be widely used to analyze other security systems for different objectives under traffic analysis attacks.

By statistical analysis of different feature statistics (sample mean, sample variance, and sample entropy) of the padded traffic PIATs, we found that sample variance and sample entropy can exploit the correlation between payload traffic rate and PIATs of padded traffic, when the padded traffic is dumped and explored next to the sender gateway or at a remote site across one or more congested routers. The reason for the failure of CIT padding is that payload traffic causes small disturbances to the timer's interval, which is used to control packet sending. Moreover, the higher the user traffic rate, the larger the disturbance of the PIAT of the padded traffic.

After a careful analysis, we propose VIT link padding as an alternative to the most common CIT link padding. Both theoretical analysis and empirical results validate the effectiveness of VIT padding strategy. The importance of the VIT padding technique is validated by extensive experiments showing that CIT link padding may be compromised even at a remote site behind noisy routers.

In this portion of the dissertation research, we discuss the simple case where two classes of traffic rates should be distinguished. Our technique can be easily extended to multiple traffic rates by performing more offline training. It is also straight-forward to use our technique to track the user payload traffic rate variation by measuring sample entropy of the PIAT of the corresponding padded traffic.

# 4. ACTIVE LINK-LOAD ANALYSIS ATTACKS AND COUNTERMEASURES

In this section, we address issues related to the active link-load analysis attack and their countermeasures. In particular, we study how an adversary may discover link load by actively pinging victim networks and analyzing statistics of ping round trip times, (i.e., the active link load analysis attack) even if link load is protected by the VIT padding analyzed in Section 3. We will also develop and evaluate countermeasures against this kind of attack.

## 4.1. Models

This section first presents the network model and then discusses link padding mechanisms used as countermeasures for traffic analysis attacks. Finally, we define the model of adversary who uses statistical pattern recognition strategies for active *ping* probing attacks on these security systems, which employ link padding mechanisms.

### 4.1.1. Network Model

We use a similar network model as shown in Figure 1. But we assume that the payload traffic has been protected by VIT link padding implemented on gateways $GW_A$ and $GW_B$, which allow cross traffic to pass through.

### 4.1.2. Adversary Model

We first present our assumptions on the capabilities of the adversary, i.e. threat model.

1. The adversary cannot obtain information from packet contents, which are perfectly encrypted. All (payload or dummy) packets have a constant size and dummy packets cannot be distinguished from payload packets.
2. The adversary is external, that is, she is not a participant of either Subnet A or B and does not compromise sender and receiver gateways. The adversary can only obtain access to the two subnets in seemingly legal ways, such as pinging the two gateways.
3. The adversary has complete knowledge about the gateway machines and the countermeasure algorithms used for preventing traffic analysis. Thus, the adversary can simulate the entire system, including the gateway machines, to obtain a priori knowledge about traffic behavior. In many studies on information security, it is a convention that we make worst case assumptions like this. But, we will show in this section, even without the capability of simulating the system, the adversary can also track the traffic rate changing pattern by a method introduced in this section.

Now we discuss the adversary strategy in terms of an active link load analysis attack. Recall that the motivation of link padding is to ensure traffic confidentiality, i.e., to prevent the adversary from performing traffic analysis and inferring critical characteristics of the payload traffic exchanged over unprotected networks. We limit the adversary's interest to *payload traffic rate*, that is, the rate at which payload traffic is exchanged between protected subnets. Specifically, we assume that there is a set of discrete payload traffic rates $\{\omega_1, \ldots, \omega_m\}$. At a given time, the rate of payload traffic

from the sender will be one of those rates. Consequently, the objective of the adversary is to identify at which of the *m* rates the payload is transmitted. But, we will also demonstrate how the adversary may take the approach outlined in this section to track the continuous changing pattern of the payload traffic.

We consider that, based on these assumptions, the adversary may deploy a sophisticated ping probing attack aimed at determining the payload traffic rate. In the attack, the adversary pings the sender gateway $GW_A$, analyzes the statistics of round trip times (RTT) of these ping packets and tries to derive Subnet A's payload traffic rate (even if $GW_A$ uses VIT padding). We use this ping attack as a model to analyze a much larger class of active probing attacks.

The adversary can analyze his sample of ping RTT data based on Bayes decision theory [27]. The entire attack strategy consists of two phases: Offline training phase and online recognition phase. We will describe them below.

The offline training phase can be decomposed into the following steps:

1. Collecting training data: The adversary reconstructs the entire link padding system and collects timing information of ping packets at different payload traffic rates.
2. Preprocessing training data: From the timing data, the adversary derives RTT information at different payload traffic rates.
3. Selecting feature from preprocessed training data: The adversary selects a statistic of the RTT sample of size *n*. This statistic is called *a feature* and will be used for traffic rate classification. Possible features we study in this section are

*sample mean*, *sample variance*, and *sample entropy*. The adversary then derives the *Probability Density Functions* (PDF) of the selected statistical feature. As histograms are usually too coarse for the distribution estimation, we assume that the adversary uses the *Gaussian kernel estimator of PDF* [1], which is effective in our problem domain.

4. Selecting decision rule: Based on the PDFs of statistical features for different payload traffic rates, Bayes decision rules are derived. Recall that there are *m* possible payload traffic rates. The Bayes decision rule can be stated as follows:

*The sample represented by feature s corresponds to payload rate* $\omega_i$ *if*

$$\forall j \in [1, m], P(\omega_i | s) \geq P(\omega_j | s) \tag{4.1}$$

*That is,*

$$f(s|\omega_i)P(\omega_i) \geq f(s|\omega_j)P(\omega_j) \tag{4.2}$$

*where* $f(s|\omega_i)$ *is PDF of feature s conditioned on payload traffic rate* $\omega_i$, $P(\omega_i)$ *is the a priori probability that the payload traffic is transmitted at rate* $\omega_i$, *and* $P(\omega_i|s)$ *is the post priori probability that the payload traffic is sent at rate* $\omega_i$ *when the collected sample has the measured feature equal to s*

Once the adversary completes her training phase, she can perform the classification at run time. We assume that the adversary has the means to ping the gateways $GW_A$ and $GW_B$. In particular, when she intends to determine the current payload rate, the adversary collects a sample of ping RTTs. She calculates the value of the statistical feature from

the collected sample and then uses the Bayes decision rules derived in the offline training phase to determine the payload traffic rate.

## 4.2. Overview of Countermeasures to Link-Load Analysis Attacks

In Section 3.2, we have introduced CIT padding and VIT padding methods to protect the link load from passive link load analysis attacks. Recall the implementation of these methods. In CIT Padding, a periodic timer is used on $GW_A$ to control the sending of padded traffic. In VIT padding, a non-periodic timer is used to $GW_A$ to control the sending of padded traffic.

As we have shown in Section 3.4, VIT padding can effectively counter passive link-load analysis attacks while the commonly used CIT padding fails under passive link-load analysis attacks.

In this section of the dissertation, we will show that the active link load analysis attack can defeat VIT padding. To counter this kind of active attack, we have to perturb both the payload traffic and cross traffic through $GW_A$.

## 4.3. Performance Metric and Analysis

### 4.3.1. Detection Rate as Performance Metric

Given models described in the previous section, we would like to evaluate the security of the system in Figure 1 in terms of detection rate. Recall that d*etection rate* is defined as the probability that the adversary can correctly identify the payload traffic rate. In this section, we derive the closed-form formulae for detection rates when the adversary uses

sample mean, sample variance, or sample entropy, as the statistical feature, respectively. Our formulae will be approximate due to the complexity of the problem. Nevertheless, they do correctly reflect the impact of various system parameters, including the type of padded traffic, sample size, and statistical feature used. These relationships are useful for understanding the nature of the attack and designing effective countermeasures. In the next section, we will see that experimental data well matches the detection rate predicted by our approximation formulae.

Let $\{X_1, \ldots, X_n\}$ be a sample of ping RTT with sample size $n$. *Sample mean* $\overline{X}$, *sample variance* $Y$, and *sample entropy* $\tilde{H}$ are defined below:

Sample Mean**:**

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{4.3}$$

Sample Variance**:**

$$Y = \frac{\sum_{i=1}^{n} (X_i - m)^2}{n-1} \tag{4.4}$$

Sample Entropy**:**

$$\tilde{H} \approx \sum_i \frac{k_i}{n} \log \frac{k_i}{n} + \log \Delta x \tag{4.5}$$

Note that in (4.5) we use the histogram-based entropy estimation developed in [32], where $k_i$ is the number of sample points in the $i^{th}$ bin, and $\Delta x$ is the bin size of the histogram. In Theorem 4.1, we derive a method to calculate the optimal bin size for the estimation of entropy. Refer to Section 4.5.1 for the proof of Theorem 4.1.

Theorem 4.1: The optimal bin size $\Delta h$ for the histogram-based entropy estimator can be calculated as follows:

$$\Delta x = \frac{6\sigma}{I} \tag{4.6}$$

where $\sigma$ is the standard deviation of the underlying distribution and the number of bin $I$ can be calculated as follows:

$$(I-1)I^2 = 3n \tag{4.7}$$

Below, we derive close-form formulae for simple cases in which the user payload traffic has two statuses: low rate $\omega_l$ and high rate $\omega_h$.

### 4.3.2. Derivation of Detection Rate

We consider two kinds of systems: stable systems where the payload traffic rate remains stable and periodic systems where the payload traffic rate changes periodically.

#### 4.3.2.1. Detection Rate for Stable Systems

The ping RTT can be represented as a random variable. Let $RTT_{low}$ and $RTT_{high}$ be random variables of RTT when the user payload traffic rate is low and high respectively. Denote their means as $\mu_l$ and $\mu_h$ and variances as $\delta_l^2$ and $\delta_h^2$ respectively. Also we define $r$ as the ratio between $\delta_l^2$ and $\delta_h^2$.

$$r = \frac{\delta_h^2}{\delta_l^2} \tag{4.8}$$

The following theorems provide closed-form formulae for estimation of detection rate when sample mean, sample variance, and sample entropy are used as feature statistics respectively.

Theorem 4.2: Using sample mean as the classification feature gives rise to an estimated detection rate

$$v_{\bar{x}} \approx 1-\left(\exp\left(-\frac{1}{4}\frac{(\mu_h-\mu_l)^2}{\sigma_h^2+\sigma_l^2}\right)\right)^n \frac{1}{\sqrt{2\left(1/\sqrt{r}+\sqrt{r}\right)}} \tag{4.9}$$

Theorem 4.3: Using sample variance as the classification feature gives rise to an estimated detection rate

$$v_Y \approx \max\left(1-\frac{C_Y}{n-1},0.5\right) \tag{4.10}$$

where $C_Y$ is calculated as follows:

$$C_Y = \frac{1}{2\left(1-\frac{1}{r-1}\log r\right)^2}+\frac{1}{2\left(\frac{r}{r-1}\log r-1\right)^2} \tag{4.11}$$

Theorem 4.4: Using sample variance as the classification feature gives rise to an estimated detection rate

$$v_{\tilde{H}} \approx \max\left(1-\frac{C_H}{n},0.5\right) \tag{4.12}$$

where $C_H$ is calculated as follows:

$$C_{\tilde{H}} = \frac{1}{2\left(\log\left(\frac{r}{r-1}\log r\right)\right)^2} + \frac{1}{2\left(\log\left(\frac{r-1}{\log r}\right)\right)^2} \tag{4.13}$$

Refer to Section 4.5.2 for the proof of Theorem 4.2, Section 4.5.3 for the proof of Theorem 4.3 and Section 4.5.4 for the proof of Theorem 4.4, respectively.

We have a number of observations from the above theorems:

1. For sample mean, detection rate increases exponentially with sample size n. Thus, if there is a small difference between $\mu_h$ and $\mu_l$, detection rate will increase dramatically with sample size. Furthermore, detection rate decreases when variance $\delta_l^2$ and $\delta_h^2$ increase.
2. For sample variance, the detection rate is an increasing function in terms of sample size n. When $n \to \infty$, the detection rate is *100*%. This means that if the payload traffic lasts for a sufficient time at one rate, then the adversary can obtain a sample of sufficiently large size, and he may detect the payload traffic rate by sample variance of ping RTT. Furthermore, the detection rate is an increasing function of *r* defined in (4.8), where $r \geq 1$. That is, the smaller r, the closer the two variances under different payload traffic rates and, intuitively, the lower the corresponding detection rate. When $r = 1$, the detection rate is *50*%. That is, the probing attack using sample variance will fail.

3. For sample entropy, the detection rate is also an increasing function in terms of sample size n. Also, the detection rate is also an increasing function of r in (4.8), where $r \geq 1$. When $r = 1$, the detection rate reaches $50\%$.

4.3.2.2. Detection Rate for Periodic Systems

In this kind of system, the payload traffic rate changes periodically with time. Thus, a sample of ping RTTs may span payload traffic of different rates. A sample of ping RTTs can be partitioned into a few segments. Each segment corresponds to an interval during which the payload traffic rate is either low or high. Assume that a sample has $L$ possible partitions: $\{Partition_i: 1 \leq i \leq L\}$ in terms of segment length. For example, if we have a sample of one ping RTT, we have two possible partitions: $Partition_1$ = {the one ping RTT is collected when the payload traffic rate is low} and $Partition_2$ = {the one ping RTT is collected when the payload traffic rate is high}.

We denote a correct detection of the payload rate by an adversary as the one in which the adversary discovers the payload traffic rate while he collects the first RTT of the sample. It can be seen that the first RTT may be collected when the user payload traffic rate is either low or high.

To derive the detection rate for this kind of system, we can derive the occurrence probability $Pr(Partition_i)$ of $Partition_i$ and the average recognition error rate conditioned on this partition case, $Pr(error|Partition_i)$. Then we have the general form of detection rate formula $v_d$ for this kind of system as follows:

$$v_d = 1 - \sum \Pr(error \mid Partition_i) \Pr(Partition_i) \quad (4.14)$$

For the case of two payload traffic rates, we assume that the payload traffic at each rate lasts for half of the payload rate changing period, $M$ is the number of ping RTTs in half of this period[4] and $n$ is the sample size, we have the following theorem.

Theorem 4.5: When sample size $n < M$, a closed form of detection rate is as follows:

$$v_d = 1 - \left( \frac{M - n + 1}{M} \varepsilon + \frac{n - 1}{2M} \right) \quad (4.15)$$

where error rate $\varepsilon = 1\text{-}v$, and $v$ can be calculated in (4.9), (4.10) or (4.12) when the adversary uses different features respectively.

Refer to Section 4.5.5 for the proof. From Theorem 4.5, we have the following observations:

1. When the ping packet rate is fixed, the payload rate changing period is larger, $M$ is larger and thus $v_d$ is larger. This is intuitive.
2. $v_d$ has a complicated relationship with sample size $n$ because of ε's relation with $n$. From our experiments and later analysis, we can see that given $M$, detection rate $v_d$ has a maximum value at some $n$.

## 4.4. Evaluations

In this section, we evaluate the security of a system under an active traffic analysis attack. We will also demonstrate how well the theoretical analysis of detection rate from the previous section approximates results from experiments designed to reflect real-life

[4] Ping packets are sent out at a constant rate.

situations. In the series of experiments that we conducted, we assume that the adversary uses a high-performance network analyzer, such as Agilent's J6841A, to dump ping packets. In terms of experimental environments, we consider the following three cases: laboratory (i.e., local area network - LAN), campus networks (i.e., metropolitan area network - MAN), and Internet, (i.e., wide area network - WAN).

$GW_A$ and $GW_B$ in Figure 10 run TimeSys Linux/Real-Time. To counter passive traffic analysis attacks, VIT padding is used. The timer interval satisfies a normal distribution $N(10ms, 3ms^2)$, which is a reasonable setting for resisting passive traffic analysis attacks. Thus, the average rate of padded traffic between the two security gateways is *100* packets per second (pps). The payload has two average rate states: *10* pps and 40pps. We assume both rates occur in equal probability. Note that for such a system, the detection rate for the adversary is lower-bounded at 50%. For all the experiments, the adversary uses an appropriate rate of ping packets with size of 512 bytes.

### 4.4.1. Experiments in a Laboratory Environment

The experiment setup is shown in Figure 10. The advantage of experimenting in a laboratory environment is that we can control the cross traffic over the network. The disadvantage is that the generated cross traffic may not reflect the characteristics of a real network.

The two gateways are connected by a Marconi ESR-5000 enterprise switch router. Subnet C is connected to the router as the cross traffic (noise) generator while the cross traffic receiver is located in Subnet D. The cross traffic shares the outgoing link of the

router, creating a case where the cross traffic has an influence on the padded traffic. The adversary pings the sender gateway behind the Marconi router.

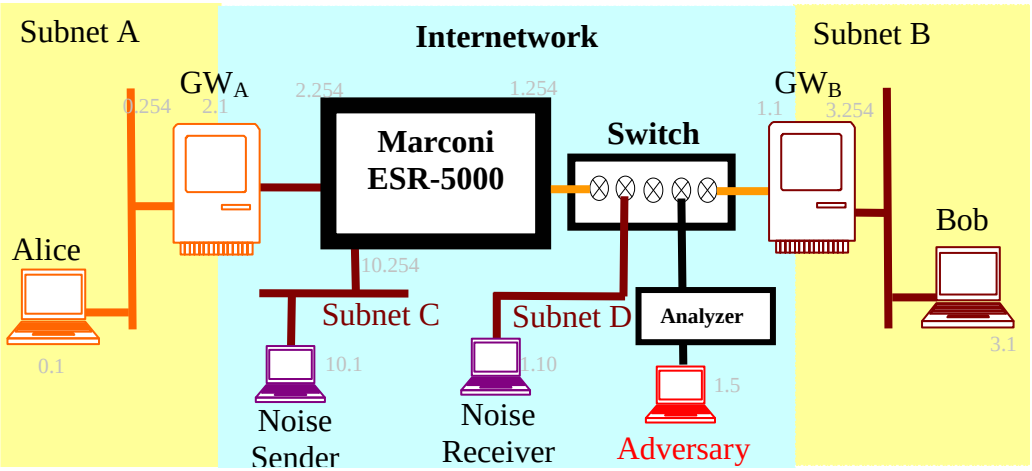



Figure 10. Experiment Setup in Laboratory for Active Attacks

4.4.1.1. Results of Stable Systems

Figure 11 (a) and Figure 11 (b) show the detection rate by different features for cases of with and without cross traffic. We have the following observations:

1. As the sample size increases, as shown in Figure 11 (a), detection rates for sample mean, sample variance, and sample entropy increase and approach 100%. This shows that when payload traffic lasts for a sufficiently long time at a given rate, an adversary can use these three features to determine the payload traffic rate with 100% accuracy, even if VIT padding has been used. This means that

security systems using padding fail under active traffic analysis attacks. Furthermore, the trend of theoretical detection rate curves coincides well with the trend of empirical curves for the three features.

2. From Figure 11 (a) and Figure 11 (b), sample entropy is a fairly robust feature in detecting the payload traffic rate. This is because sample entropy defined in (5) is not sensitive to outliers, which influence the performance of sample mean and sample variance, especially when there is cross traffic.
3. In Figure 11 (b), as the link utilization increases, the detection rates of the three features decrease. This is because the cross traffic between Subnet C and Subnet D interferes with ping traffic. In theory, compared to the ping RTT variances $\sigma_l^2$ and $\sigma_h^2$ in the no cross traffic case, both these variances are increased in the cross traffic case, by a quantity relative to the cross traffic. This will cause a decrease in r. As Theorem 4.2 predicts, the detection rate by all three features drops.

We have seen that systems with VIT padding fail under active traffic analysis attacks. The reason of this failure lies in the subtle interaction between the traffic padding system and the probing traffic. While $GW_A$'s network subsystem processes payload packets from Subnet A, the processing of ping packets is delayed. A higher rate of payload traffic causes more delay on ping packets. This means that sample mean, sample variance, and sample entropy of the RTT of the probing packets at a given sample size $n$ are changed, and there is some correlation between the user payload traffic rate and sample mean, sample variance, and sample entropy of the RTT of the probing packets. The adversary can exploit this correlation to discover the user payload traffic rate.

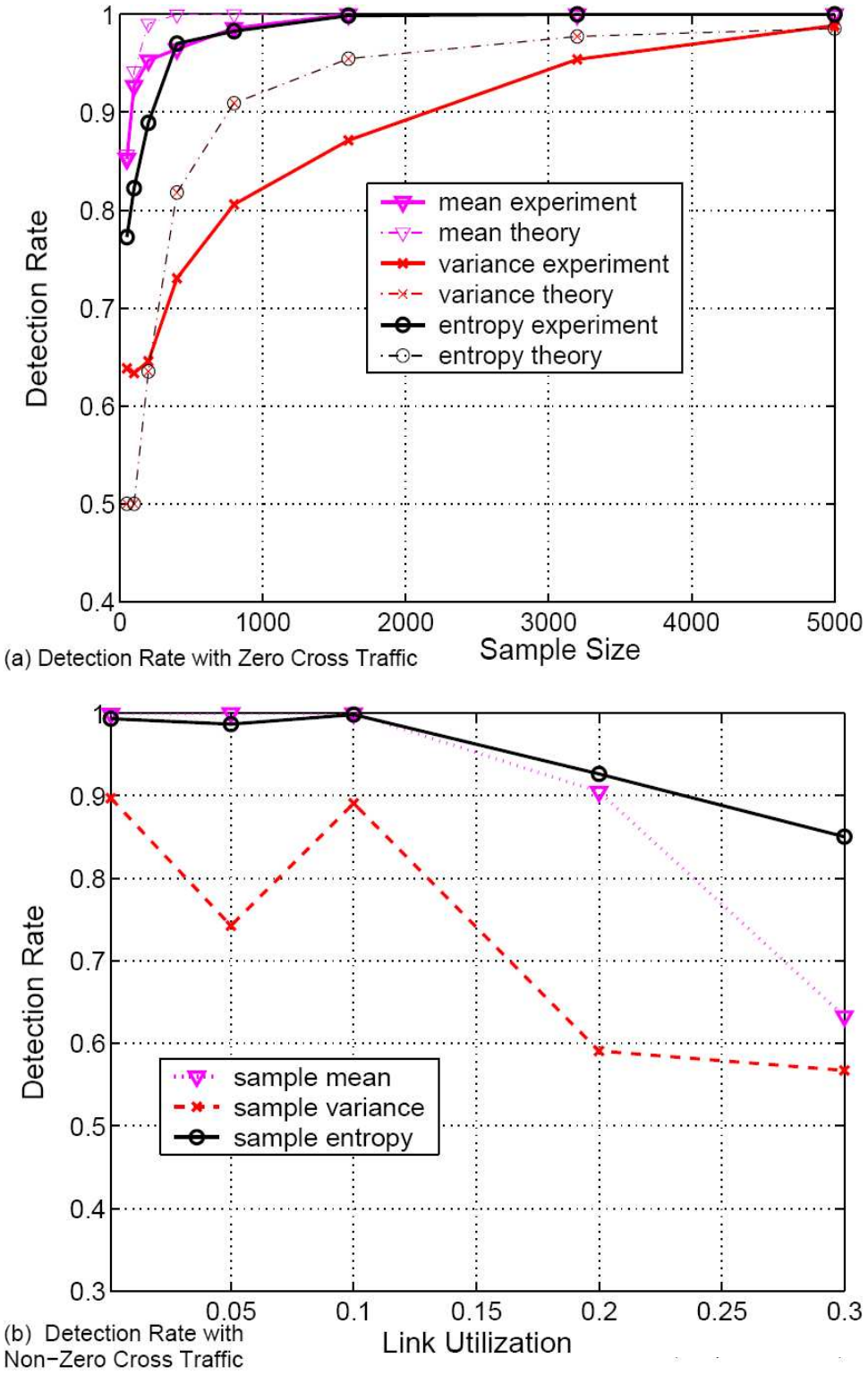
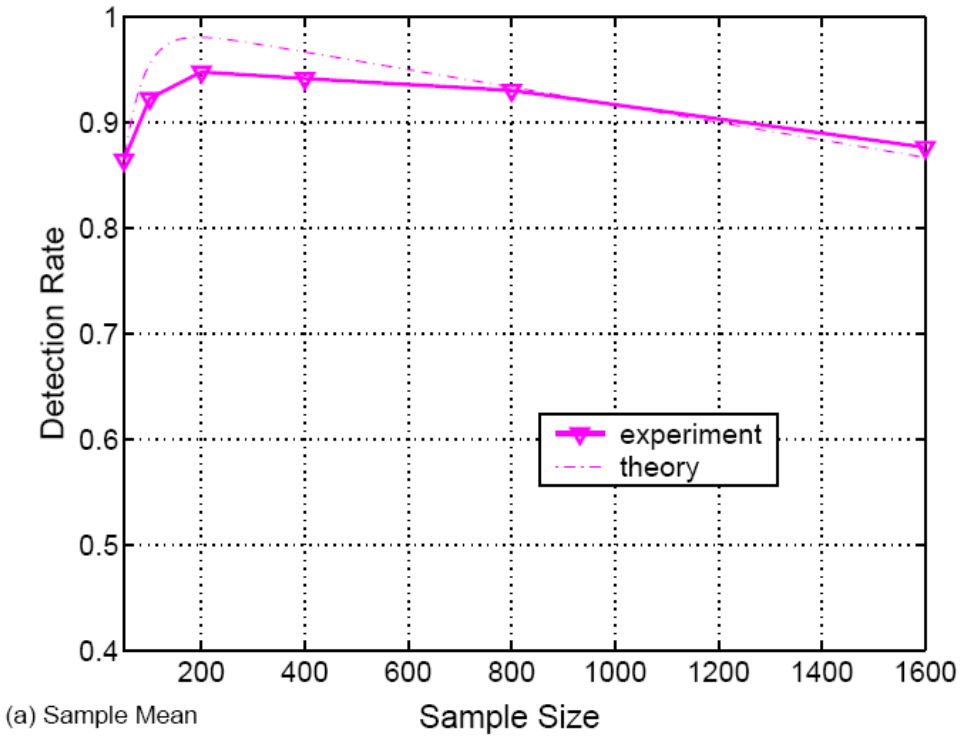



Figure 11. Detection Rate for Stable Payload Traffic

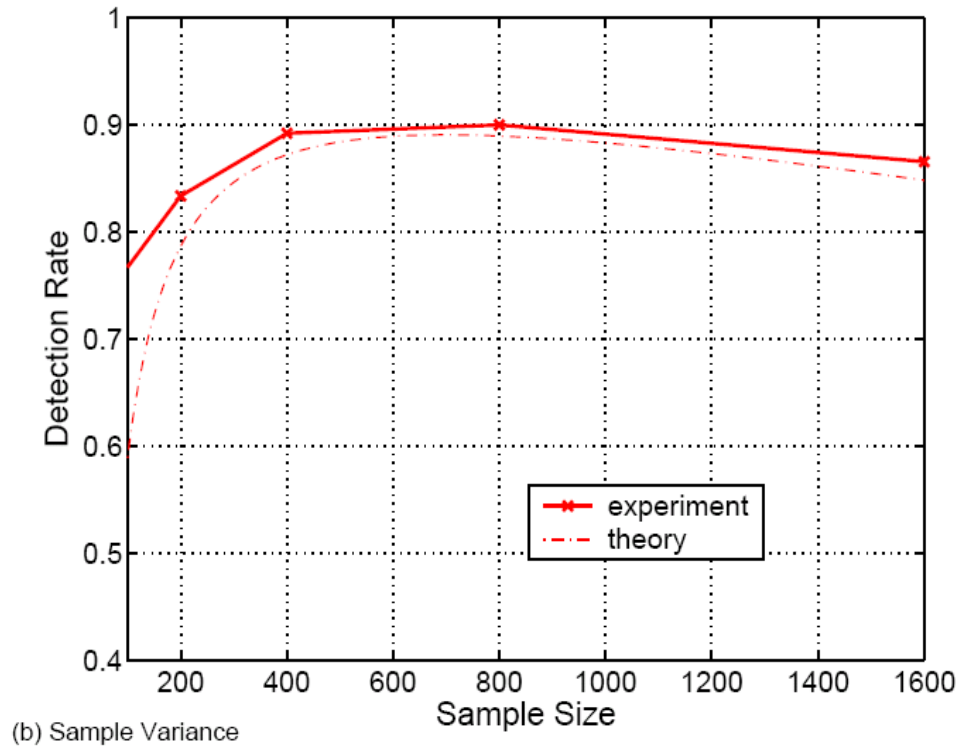
#### 4.4.1.2. Results of Periodical Systems

Figure 12 (a), (b) and (c) give detection rates for periodical systems. The traffic rate changes periodically between 10pps and 40pps. The period in our experiments is 1 minute. Figure 13 illustrates how the adversary can track continuously changing payload traffic rate by probing attacks. We have the following observations.

1. In all figures, the theoretical curves match the empirical curves generally well. This validates various approximate assumptions made when we derive Theorem 4.5.
2. As Theorem 4.5 predicts, there exists a maximum detection rate as sample size changes. That is, in practice, when the ping probing attack is deployed, the adversary has to choose an appropriate sample size to maximize its detection rate.
3. In Figure 13, sample entropy (sample size = 2,000) is used to track the changing pattern of the user payload traffic rate while the payload traffic rate has three statuses: 0 pps, 10 pps, and 40 pps. Each rate lasts for 5 minutes in one period. It is clear that the adversary can use sample entropy to reconstruct the payload traffic rate pattern very well. This further confirms the effectiveness of probing attacks.




(a) Sample Mean




(b) Sample Variance

Figure 12. Detection Rate for Periodical Systems

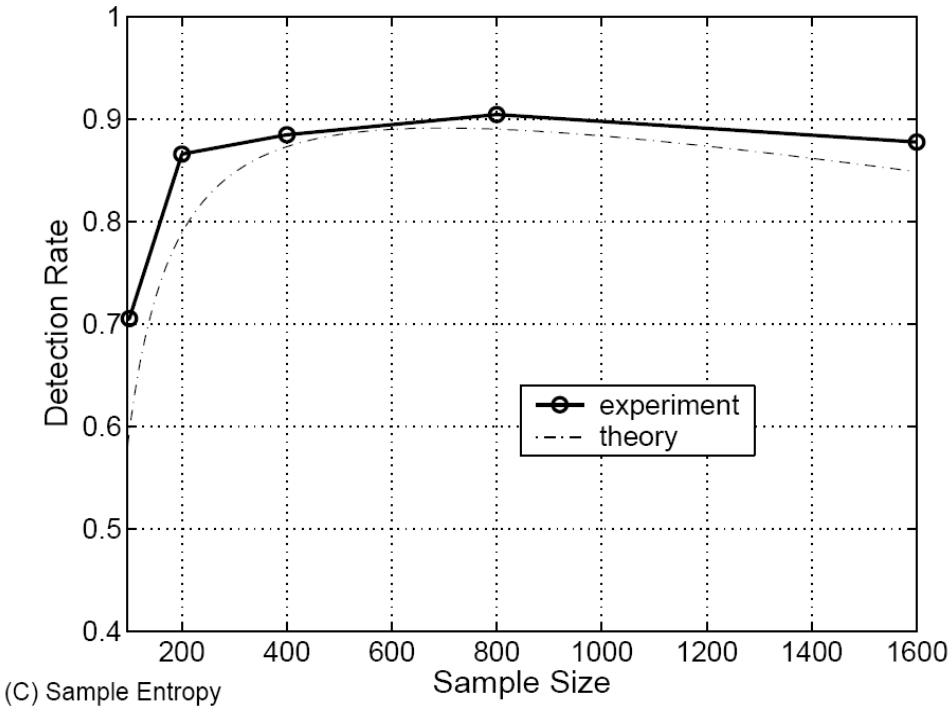



Figure 12. Continued

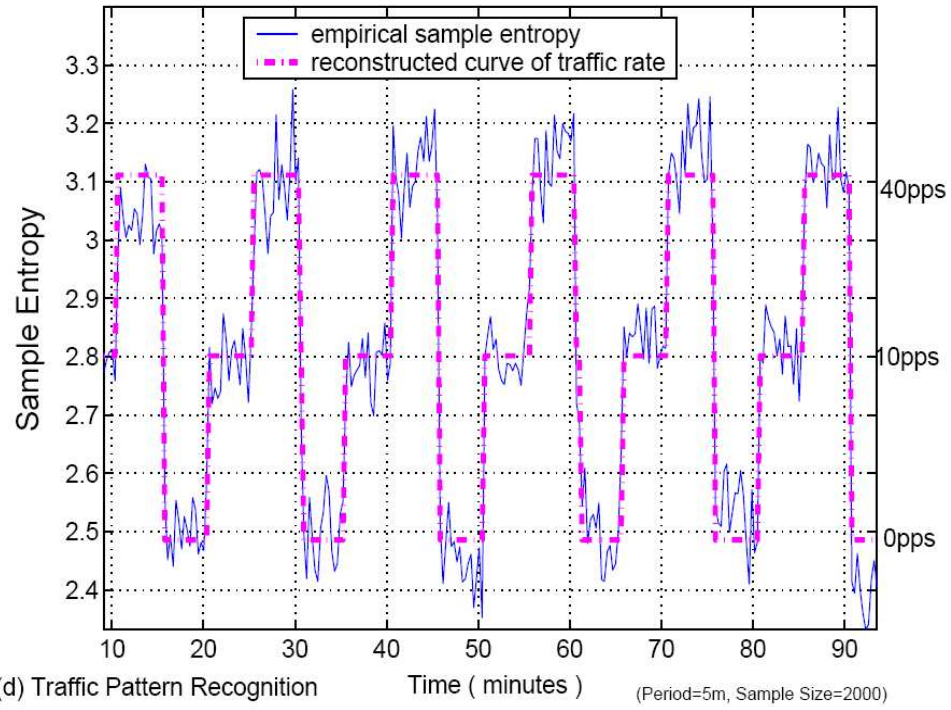



Figure 13. Tracking of the Changing Pattern of the User Payload Traffic Rate

4.4.1.3. Security Improvement

From Theorem 4.2, Theorem 4.3, and Theorem 4.4, we know that when $r$ decreases, and/or $\sigma_l^2$ and $\sigma_h^2$ increase, the detection rate decreases. To reduce $r$ and increase $\sigma_l^2$ and $\sigma_h^2$, we may intentionally introduce a random delay to ping packets. This is similar to the effect of adding noise to the RTT of ping packets by cross traffic. We assume that the random delay satisfies a normal distribution $N(\mu_T, \sigma_T^2)$. It can be perceived that an appropriate selection of $\mu_T$ and $\sigma_T$ will dramatically reduce the detection rate. To validate this approach, we again use the configuration in Figure 10 as the experimental network setup. There is no cross traffic. Figure 14 gives the detection rate by different statistics when ping packets are delayed by a random interval, which satisfies a normal distribution $N(10ms, 3ms^2)$. We observe that the detection rate achievable by the adversary at different feature statistics approaches 50% (the minimum detection rate for two class recognition) at a large sample size.

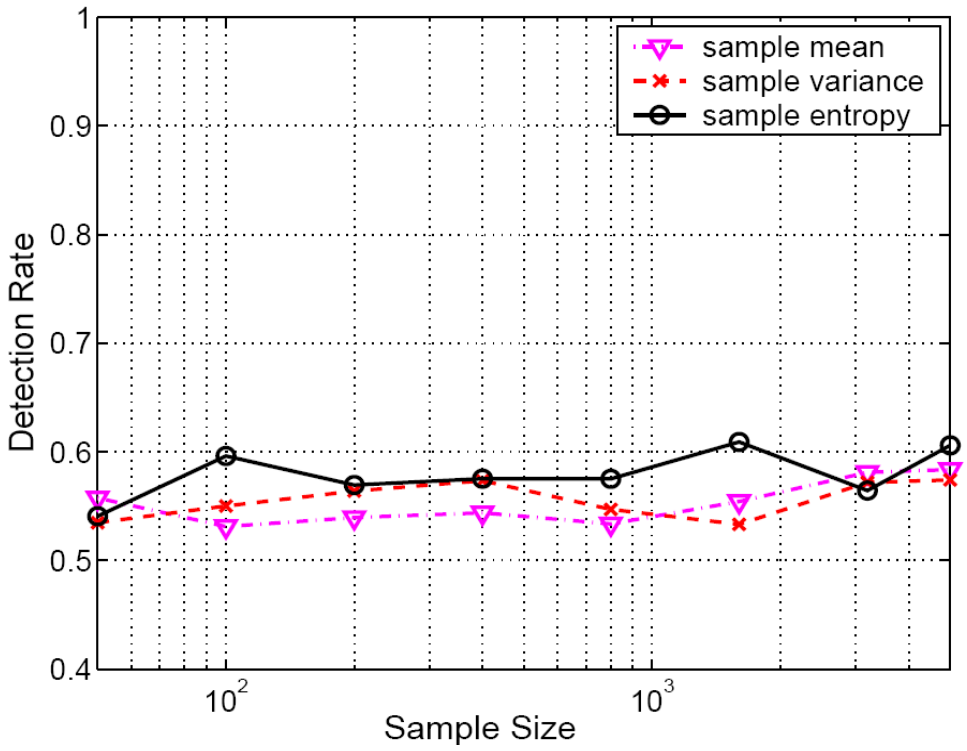



Figure 14. Detection Rate by RTT of Delayed Ping Packets with Zero Cross Traffic

In addition to this random delay method, other means may also be used to counter the active traffic analysis attacks. To defeat the ping-based probing attack, one can disable the ping service on security gateways, but the disadvantage of this method is that ping often is a useful service for debugging a network, e.g., to check if $GW_A$ is alive. Sometimes, one cannot sacrifice functionality for the sake of security.

Another method for countering active ping probing attacks is that one should avoid sending payload traffic at a constant rate for long periods of time. For example, in a peer-to-peer anonymous file sharing system, the file should be split into small pieces

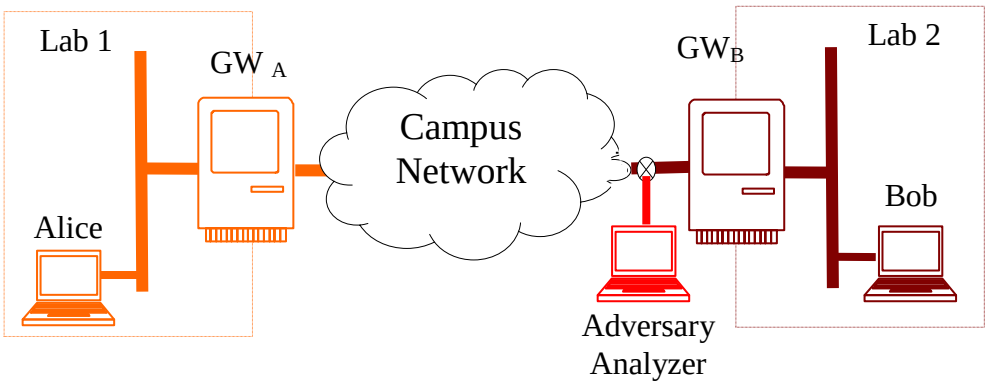
before uploading and downloading. Consequently, the detection rate will decrease since the adversary cannot get a sufficiently large sample for a constant rate of payload traffic.

### 4.4.2. Experiments over Campus and Wide Area Networks

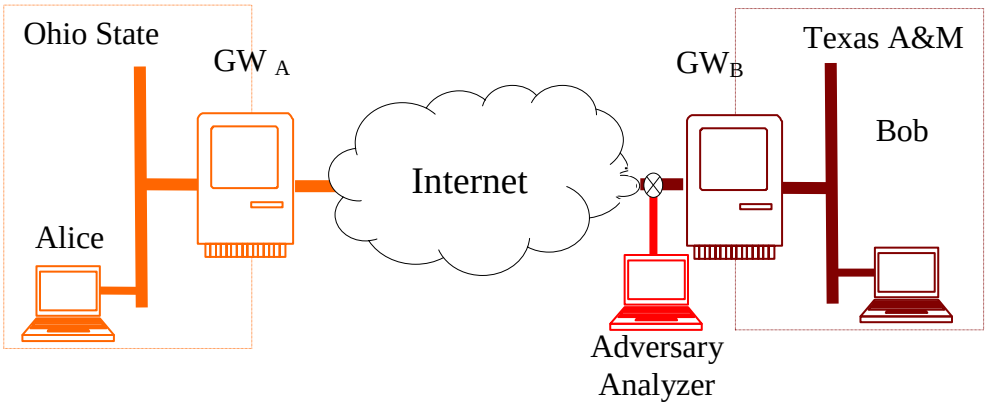
In this subsection, we examine the detection rate when the ping traffic of the adversary traverses a campus network and the internet, respectively.

Figure 15 shows the setup for the experiments discussed in this subsection. In both cases, the observation point of the adversary is located right in front of the receiver gateway and thus maximally far from the sender. Figure 15 (a) is a setup for experiments over our local campus network. That is, the ping traffic goes through our local campus network before it reaches the sender's gateway. Figure 15 (b) is a setup for experiments over the Internet between a remote campus network and our local campus network. Here, the sender workstation and the sender gateway are located at the remote campus network. The ping traffic goes through the Internet and arrives at the remote campus network. We note that in this case, the path from the sender's workstation to the receiver's workstation spans 15 or more routers.

In each case, we collect data continuously for 24 hours. The data for the case of our local campus network was collected on July 16, 2003, while the data for the wide area network case was collected on July 14, 2003.




(a) Experiment Setup within Texas A&M University



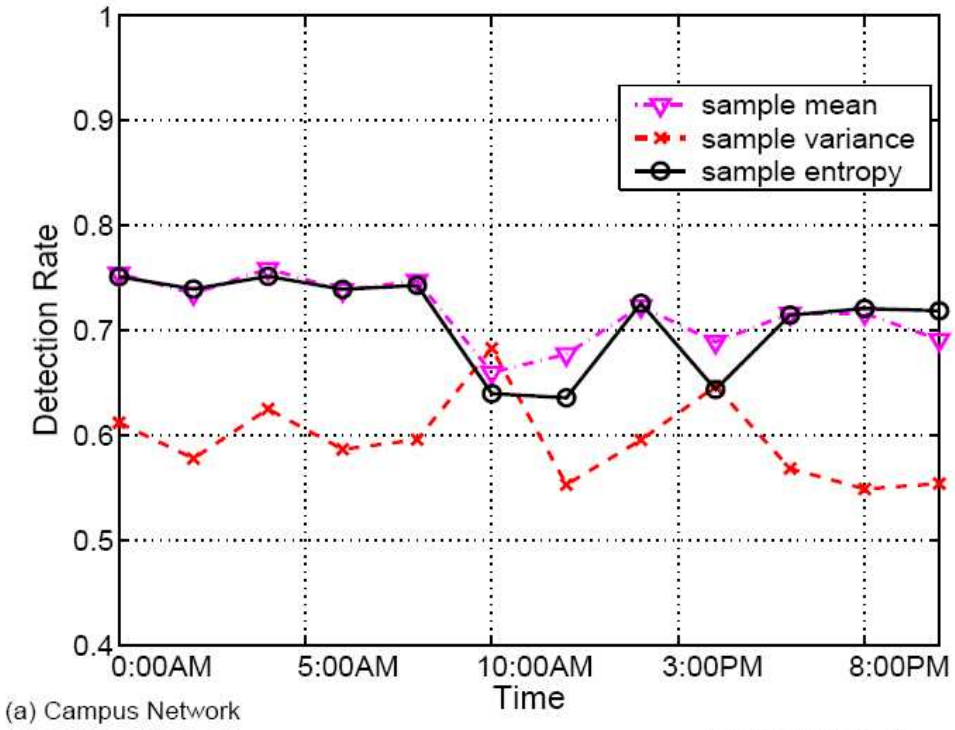

(b) Network Setup between Ohio and Texas

Figure 15. Experiment Setup over Campus and Internet for Active Attacks

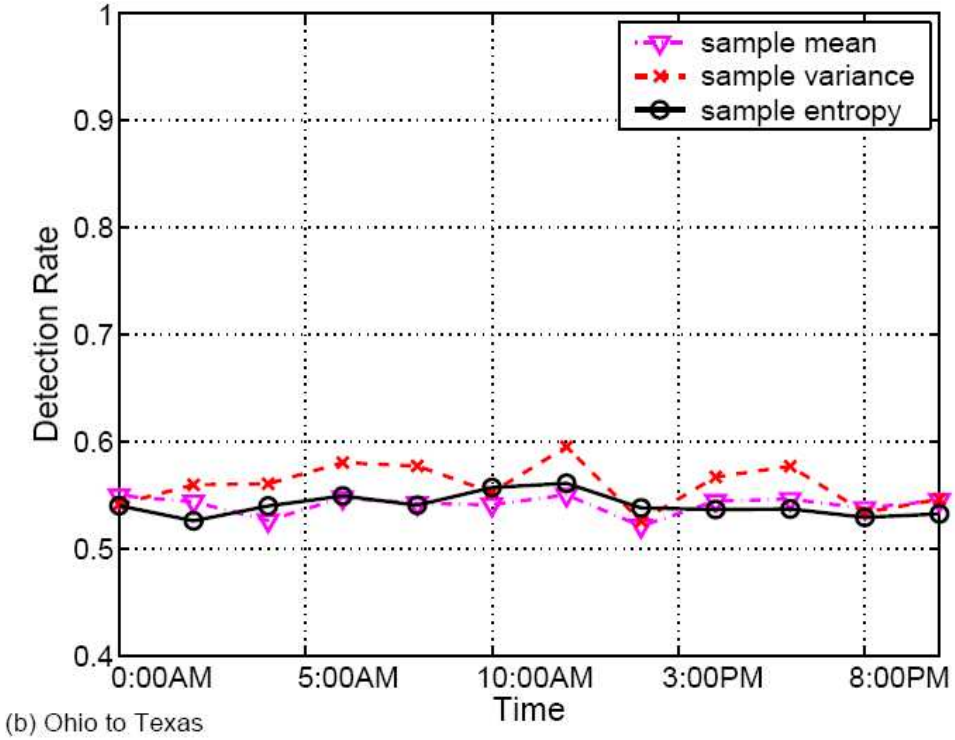
Figure 16 (a) and (b) show the detection rate throughout the observation period. We have the following observations:

1. When ping traffic traverses only our local campus network, the detection rates of sample entropy and sample mean can approach about *75*%. This means that over a medium-sized enterprise network like our campus network, the cross traffic does have an influence on the ping traffic, but systems using VIT padding scheme alone still cannot resist ping probing attacks effectively.
2. When the padded traffic traverses more network elements, such as the Internet between the remote campus network and our local campus network, the detection rates are much lower. This is because ping traffic has a low scheduling priority at a large number of routers and switches, and the RTT of ping packets is seriously distorted.

Experimental results in Figure 16 match our theoretical analysis in Section 4.3.2. This further validates the correctness of our theoretical framework for analyzing the active link load analysis attack and their countermeasures.




(a) Campus Network




(b) Ohio to Texas

Figure 16. Empirical Detection Rate for Active Attacks over Campus and Internet

## 4.5. Theorem Proof

In this section, we prove theorems introduced in Section 4.3.2.

### 4.5.1. Proof of Theorem 4.1

Theorem 4.1: The optimal bin size $\Delta h$ for the histogram-based entropy estimator can be calculated as follows:

$$\Delta x = \frac{6\sigma}{I} \tag{4.16}$$

where $\sigma$ is the standard deviation of the underlying distribution and the number of bin $I$ can be calculated as follows:

$$(I-1)I^2 = 3n \tag{4.17}$$

*Proof:* We now develop theory calculating the optimal bin size for histogram based entropy estimation. In [32], for a histogram based entropy estimation,

$$E(\tilde{H}) \approx H - \frac{I-1}{2n} + \frac{1}{24}\left(\frac{\Delta x}{\sigma}\right)^2 \tag{4.18}$$

where $\tilde{H}$ is the estimated entropy, $H$ is original entropy of random variable $X$, $I$ is the number of bins, $n$ is the sample size, $\Delta x$ is the bin size and $\sigma$ is $X$'s standard deviation.

If random variable $X$ is normally distributed,

$$\mathrm{var}(\tilde{H}) \approx \frac{1}{2}n^{-1} \tag{4.19}$$

Moreover, for a normally distributed random variable, since the probability of $x>3\sigma$ is very small and ignorable, we have the following relationship between bin size $\Delta x$, the number of bins $I$ and random variable $X$'s standard deviation $\sigma$,

$$6\sigma = I\Delta x \tag{4.20}$$

Thus

$$\frac{\Delta x}{\sigma} = \frac{I}{6} \tag{4.21}$$

We can calculate the mean square error (MSE) of the entropy estimation as follows,

$$E\left[\left(H - \tilde{H}\right)^2\right] = VAR\left(\tilde{H}\right) + \left[BIAS\left(\tilde{H}\right)\right]^2 \tag{4.22}$$

where

$$\left[BIAS\left(\tilde{H}\right)\right] = E\left(\tilde{H}\right) - H \tag{4.23}$$

Substitute (4.18) and (4.19) into (4.22),

$$E\left[\left(H - \tilde{H}\right)^2\right] = \frac{1}{2}n^{-1} + \left(\frac{1}{24}\left(\frac{\Delta x}{\sigma}\right)^2 - \frac{I-1}{n}\right)^2 \tag{4.24}$$

To derive a minimum mean square error, the second term in (4.24) should be minimized,

$$\frac{1}{24}\left(\frac{\Delta x}{\sigma}\right)^2 = \frac{I-1}{n} \tag{4.25}$$

Substituting (4.21) into (4.25), we have

$$\frac{1}{24}\left(\frac{6}{I}\right)^2 = \frac{I-1}{2n} \tag{4.26}$$

Thus

$$(I-1)I^2 = 3n \tag{4.27}$$

$$\Delta x = \frac{6\sigma}{I} \tag{4.28}$$

Below, we derive close-form formulae for simple cases in which the user payload traffic has two statuses: low rate $\omega_l$ and high rate $\omega_h$.

4.5.2. Proof of Theorem 4.2

Theorem 4.2: Using sample mean as the classification feature gives rise to an estimated detection rate

$$v_{\bar{x}} \approx 1 - \left(\exp\left(-\frac{1}{4}\frac{(\mu_h - \mu_l)^2}{\sigma_h^2 + \sigma_l^2}\right)\right)^n \frac{1}{\sqrt{2\left(1/\sqrt{r} + \sqrt{r}\right)}} \tag{4.29}$$

*Proof:* The distribution of sample mean of a normal distribution $N(\mu, \sigma^2)$ is still a normal one, $N(\mu, \sigma^2/n)$. Thus, sample mean $\overline{X}_l$ for the case when the payload traffic rate is low has a normal distribution

$$f_l(x|\omega_l) = N(\mu_l, \sigma_1^2) = N\left(\mu_l, \frac{\sigma_l^2}{n}\right) \tag{4.30}$$

Similarly, sample mean $\overline{X}_h$ for the high payload traffic rate has a normal distribution

$$f_h(x|\omega_h) = N(\mu_h, \sigma_2^2) = N\left(\mu_h, \frac{\sigma_h^2}{n}\right) \tag{4.31}$$

Since $\overline{X}_l$ and $\overline{X}_h$ are normally distributed, we can use the Bhattacharyya bound [27] to estimate the error rate as follows:

$$\varepsilon_{\overline{X}} \le \sqrt{P(\omega_l)P(\omega_h)}\int \sqrt{f(x \mid \omega_l) f(x \mid \omega_h)}dx \tag{4.32}$$

Substituting (4.30), (4.31), and $P(\omega_l)=P(\omega_h)=0.5$ into (4.32) and carrying out the integration, we have

$$\varepsilon_{\overline{X}} \le \frac{1}{2}\exp(-k) \tag{4.33}$$

where

$$k = \frac{1}{4}\frac{(\mu_h - \mu_l)^2}{\sigma_h^2 + \sigma_l^2} + \frac{1}{2}\ln\frac{\frac{\sigma_l^2 + \sigma_h^2}{2}}{\sqrt{\sigma_l^2 \sigma_h^2}} \tag{4.34}$$

After substituting (4.34) into (4.33) and some rearranging, we have

$$k = \frac{1}{4}\frac{(\mu_h - \mu_l)^2}{\sigma_h^2 + \sigma_l^2} + \frac{1}{2}\ln\frac{\left(1/\sqrt{r} + \sqrt{r}\right)}{2} \tag{4.35}$$

Substituting (4.35) into (4.33), the error rate is given by

$$\varepsilon_{\bar{X}} \le e^{-\frac{1}{4}\frac{(\mu_h - \mu_l)^2}{\sigma_h^2 + \sigma_l^2}} \frac{1}{\sqrt{2\left(1/\sqrt{r} + \sqrt{r}\right)}} \tag{4.36}$$

The detection rate $v_{\bar{X}}$ then satisfies the following:

$$v = 1 - \varepsilon \tag{4.37}$$

$$\ge 1 - e^{-\frac{1}{4}\frac{(\mu_h - \mu_l)^2}{\sigma_h^2 + \sigma_l^2}} \frac{1}{\sqrt{2\left(1/\sqrt{r} + \sqrt{r}\right)}} \tag{4.38}$$

Thus, we can use the lower bound of (4.38) as the estimation of detection rate by sample mean. The theorem is proven.

4.5.3. Proof of Theorem 4.3

Theorem 4.3: Using sample variance as the classification feature gives rise to an estimated detection rate

$$v_Y \approx \max\left(1 - \frac{C_Y}{n-1}, 0.5\right) \tag{4.39}$$

where *C* is calculated as follows:

$$C_Y = \frac{1}{2\left(1-\frac{1}{r-1}\log r\right)^2} + \frac{1}{2\left(\frac{r}{r-1}\log r - 1\right)^2} \tag{4.40}$$

and $r$ is defined in (4.8).

*Proof:* Denote $\chi_{n-1}^2$ as a random variable with a *chi square* distribution $f_{\chi_{n-1}^2}(x)$ with freedom *n-1*, which is defined as follows,

$$f_{\chi_{n-1}^2}(x) = \frac{x^{\frac{n-1}{2}} \exp\left(-\frac{x}{2}\right)}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)} \tag{4.41}$$

where $x > 0$. Denote $Y$ as the random variable of sample variance. Then $(n-1)Y/\sigma^2$ has a chi square distribution with freedom *n-1* [29]. That is

$$\chi_{n-1}^2 = \frac{n-1}{\sigma^2} Y \tag{4.42}$$

From (4.42), we get

$$Y = \frac{\sigma^2}{n-1}\chi_{n-1}^2 \tag{4.43}$$

From chi square's properties, we have sample variance's mean $\overline{Y}$ as

$$\overline{Y} = \sigma^2 \tag{4.44}$$

and its variance *var*(*Y*) as

$$\mathrm{var}(Y) = \frac{2\sigma^4}{n-1} \tag{4.45}$$

To get sample variance's PDF at sample size *n*, we first compute its distribution function

$$P(Y < y) = P\left(\frac{\sigma^2}{n-1}\chi^2_{n-1} < y\right) \tag{4.46}$$

$$= P\left(\chi^2_{n-1} < \frac{n-1}{\sigma^2}y\right) \tag{4.47}$$

Differentiating the two sides of (4.47), we have the density function

$$f_Y(y) = f_{\chi^2_{n-1}}\left(\frac{n-1}{\sigma^2}y\right)\frac{n-1}{\sigma_l^2} \tag{4.48}$$

We denote $Y_l$ as the random variable of sample variance of padded traffic's PIAT at the low rate payload traffic. Substituting (4.41) into (4.48), we then derive $Y_l$'s density function $f_{Y_l}(y)$

$$f_{Y_l}(y) = \frac{\left(\frac{n-1}{\sigma_l^2}y\right)^{\frac{n-1}{2}-1}\exp\left(-\frac{n-1}{2\sigma_l^2}y\right)}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}\frac{n-1}{\sigma_l^2} \tag{4.49}$$

Similarly, $Y_h$ is the random variable of sample variance of padded traffic's PIAT at high rate payload traffic, and its density function $f_{Y_h}(y)$ is

$$f_{Y_h}(y) = \frac{\left(\frac{n-1}{\sigma_h^2} y\right)^{\frac{n-1}{2}-1} \exp\left(-\frac{n-1}{2\sigma_h^2} y\right)}{2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right)} \frac{n-1}{\sigma_h^2} \tag{4.50}$$

To get the detection rate, we calculate the cross point $y_c$ of $f_{Y_l}(y)$ and $f_{Y_h}(y)$

$$f_{Y_l}(y_c) = f_{Y_h}(y_c) \tag{4.51}$$

After lengthy arithmetic operations, we have

$$y_c = \log\left(\frac{\sigma_h^2}{\sigma_l^2}\right) \frac{\sigma_h^2 \sigma_l^2}{\sigma_h^2 - \sigma_l^2} \tag{4.52}$$

Now we use Chebynov inequality for the estimation of error rate if the adversary uses sample variance as the feature statistic. The distance from the mean of $Y_l$ to the cross point $y_c$ is denoted as $D_l$

$$D_l = y_c - \overline{Y}_l \tag{4.53}$$

Substituting (4.44) and (4.52) into (4.53), we have

$$D_l = \left(\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right) \frac{\sigma_h^2}{\sigma_h^2 - \sigma_l^2} - 1\right) \sigma_l^2 \tag{4.54}$$

Denoting $c_l$ as the ratio of $D_l$ to the standard deviation of $Y_l$

$$c_l = \frac{D_l}{\sqrt{\mathrm{var}(Y_l)}} \tag{4.55}$$

Substituting (4.45) and (4.54) into (4.55), we have

$$c_l = \frac{\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2}{\sigma_h^2 - \sigma_l^2} - 1}{\sqrt{\frac{2}{n-1}}} \tag{4.56}$$

Similarly, denoting $D_h$ as the distance from the mean of $Y_h$ to the cross point $y_c$,

$$D_h = \sigma_h^2 - y_c = \sigma_h^2\left(1 - \log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_l^2}{\sigma_l^2 - \sigma_l^2}\right) \tag{4.57}$$

Then $c_h$, the ratio of $D_h$ and the standard deviation of $Y_h$, can be calculated as follows

$$c_h = \frac{1 - \log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_l^2}{\sigma_h^2 - \sigma_l^2}}{\sqrt{\frac{2}{n-1}}} \tag{4.58}$$

When sample size $n$ is big (>40), we can assume that a chi square PDF is symmetrical. Thus from Chebynov inequality, we can get the error rate $e_Y$

$$e_Y \le \frac{\frac{1}{2c_l^2} + \frac{1}{2c_h^2}}{2} \tag{4.59}$$

Thus, the detection rate $v_Y$ can be calculated as

$$v_Y = 1 - e_Y \tag{4.60}$$

$$\geq 1 - \frac{1}{4c_l^2} - \frac{1}{4c_h^2} \tag{4.61}$$

Substituting (4.8), (4.56) and (4.58) into (4.61), we have

$$v_Y \geq 1 - \frac{C_Y}{n-1} \tag{4.62}$$

where

$$C_Y = \frac{1}{2\left(1 - \frac{1}{r-1}\log r\right)^2} + \frac{1}{2\left(\frac{r}{r-1}\log r - 1\right)^2} \tag{4.63}$$

Thus, we use the lower bound of $v_Y$ as the estimation of the detection rate. Since that $v_Y$ must be greater than 50%, we can get (4.39). The theorem is proved.

In the following, we prove that $v_Y$ is an increasing function of $r$. That is, we need to prove that $C_Y$ in (4.63) is a decreasing function of $r$. For terms in (4.63), we have the following denotations

$$C_{Y1}(r) = \frac{\log r}{r-1} \tag{4.64}$$

$$C_{Y2}(r) = \frac{r}{r-1} \tag{4.65}$$

If $C_{Y1}$ is decreasing function and $C_{Y2}$ is an increasing function, then $C_Y$ is a decreasing function.

We first prove that $C_{Y1}$ is a decreasing function. Let $r=e^x$, we have

$$C_{Y1}(r) = c_{y1}(x) = \frac{x}{e^x - 1} \tag{4.66}$$

$$\frac{dC_{Y1}}{dr} = \frac{dc_{y1}}{dr} \tag{4.67}$$

$$= \frac{dc_{y1}}{dx}\frac{dx}{dr} \tag{4.68}$$

$$= \frac{e^x - 1 - xe^x}{r(e^x - 1)^2} \tag{4.69}$$

Since $r>1$, $x>0$, the denominator of (4.69) is greater than $0$. We have the Taylor expansion of the numerator of (4.69) as follows

$$e^x - 1 - xe^x = \sum_{n=1}^{\infty}\left(\frac{1}{n!} - \frac{1}{(n-1)!}\right)x^n \tag{4.70}$$

Since

$$\frac{1}{n!} - \frac{1}{(n-1)!} < 0 \tag{4.71}$$

So

$$e^x - 1 - xe^x < 0 \tag{4.72}$$

Thus

$$\frac{dC_{Y1}}{dr} = \frac{dc_{y1}}{dr} < 0 \tag{4.73}$$

$C_{Y1}$ is a decreasing function in terms of $r$.

Now we prove $C_{Y2}$ is an increasing function of $r$. Let $r = e^x$, we have,

$$C_{Y2}(r) = \frac{r \log r}{r-1} = c_{y2}(x) = \frac{xe^x}{e^x - 1} \tag{4.74}$$

and

$$\frac{dC_{Y2}}{dr} = \frac{dc_{y2}}{dr} = \frac{e^{2x} - e^x - xe^x}{r\left(e^x - 1\right)^2} \tag{4.75}$$

Since r>1, x>0, the denominator of (4.75) is greater than *0*. The Taylor expansion of the numerator of (4.75) is as follows

$$e^{2x} - e^x - xe^x = \sum_{n=1}^{\infty} \left( \frac{2^n - 1 - n}{n!} \right) x^n \tag{4.76}$$

Since

$$\forall n > 0, \frac{2^n - 1 - n}{n!} \geq 0 \tag{4.77}$$

we have

$$e^{2x} - e^{x} - xe^{x} \geq 0 \tag{4.78}$$

So

$$\frac{dC_{Y2}}{dr} = \frac{dc_{y2}}{dr} > 0 \tag{4.79}$$

and $C_{Y2}(r)$ is an increasing function of $r$.

Thus, we have proved that is an increasing function in terms of $r$.

4.5.4. Proof of Theorem 4.4

Theorem 4.4: Using sample entropy as the classification feature gives rise to an estimated detection rate

$$v_{\tilde{H}} \approx \max\left(1 - \frac{C_H}{n}, 0.5\right) \tag{4.80}$$

where $C_H$ is calculated as follows:

$$C_{\tilde{H}} = \frac{1}{2\left(\log\left(\frac{r}{r-1}\log r\right)\right)^2} + \frac{1}{2\left(\log\left(\frac{r-1}{\log r}\right)\right)^2} \tag{4.81}$$

and $r$ is defined in (4.8).

*Proof:* A normal distribution's differential entropy can be calculated as

$$H = \frac{\log 2\pi\sigma^2 + 1}{2} \tag{4.82}$$

Here we use sample variance $Y$ defined in (4.4) to estimate sample entropy $\tilde{H}$

$$\tilde{H} = \frac{\log 2\pi Y + 1}{2} \tag{4.83}$$

To get sample entropy's PDF, we first derive its distribution,

$$P\left(\tilde{H} < h\right) = P\left(\frac{\log 2\pi Y + 1}{2} < h\right) = P\left(Y < \frac{e^{2h-1}}{2\pi}\right) \tag{4.84}$$

Differentiating two sides of (4.84), we get sample entropy's PDF

$$f_{\tilde{H}}(h) = f_Y\left(\frac{e^{2h-1}}{2\pi}\right) = P\left(Y < \frac{e^{2h-1}}{2\pi}\right)\left(\frac{e^{2h-1}}{2\pi}\right)' \tag{4.85}$$

Denote $H_l$ as the sample entropy of padded traffic's PIAT at the low-rate payload traffic, and $f_{\tilde{H}_l}(h)$ as $H_l$'s PDF. Denote $H_h$ as the sample entropy of padded traffic's PIAT at the high-rate payload traffic, and $f_{\tilde{H}_h}(h)$ as $H_h$'s PDF. To get the detection rate, we need to calculate the cross point $h_c$ of $f_{\tilde{H}_l}(h)$ and $f_{\tilde{H}_h}(h)$

$$f_{\tilde{H}_l}(h_c) = f_{\tilde{H}_h}(h_c)' \tag{4.86}$$

By lengthy arithmetic operations

$$h_c = \frac{\log\left[2\pi \log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2\sigma_l^2}{\sigma_h^2-\sigma_l^2}\right]+1}{2} \tag{4.87}$$

For the ease of estimation, we approximate sample entropy's mean as follows

$$E\left(\tilde{H}\right) = \frac{\log 2\pi\sigma^2 + 1}{2} \tag{4.88}$$

That is, we approximate the entropy estimator in (4.83) as an unbiased one.

Now we use Chebynov inequality for the error rate estimation. The distance from the mean of $H_l$ to the cross point $h_c$ is denoted as $D_l$

$$D_l = \frac{\log\left[2\pi \log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2\sigma_l^2}{\sigma_h^2-\sigma_l^2}\right]}{2} - \frac{\log 2\pi\sigma_l^2 + 1}{2} \tag{4.89}$$

$$= \frac{\log\left[\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2}{\sigma_h^2-\sigma_l^2}\right]}{2} \tag{4.90}$$

Using Taylor expansion over (4.90) and by appropriate approximation, we get

$$\mathrm{var}\left(\tilde{H}\right) \approx \frac{1}{2n} \tag{4.91}$$

(4.91) is the same result as for the histogram-based entropy estimator in [32]. Denoting $c_l$ as the ratio of $D_l$ to the standard deviation of $h_l$

$$c_l = \frac{\log\left[\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2}{\sigma_h^2-\sigma_l^2}\right]}{2\sqrt{\frac{1}{2n}}} \tag{4.92}$$

Similarly, denoting $D_h$ as the distance from the mean of $H_h$ to the cross point $h_c$ and $c_h$ as the ratio of $D_h$ to the standard deviation of $h_h$, we have

$$D_l = \frac{\log 2\pi\sigma_h^2 + 1}{2} - \frac{\log\left[2\pi\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2\sigma_l^2}{\sigma_h^2-\sigma_l^2}\right]+1}{2} \tag{4.93}$$

$$= \frac{\log\left[\frac{\sigma_h^2-\sigma_l^2}{\sigma_l^2\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)}\right]}{2} \tag{4.94}$$

$$c_h = \frac{\log\left[\frac{\sigma_h^2-\sigma_l^2}{\sigma_l^2\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)}\right]}{2\sqrt{\frac{1}{2n}}} \tag{4.95}$$

So error rate is calculated as

$$e_{\tilde{H}} \le \frac{1}{4c_l^2} + \frac{1}{4c_h^2} = \frac{C_{\tilde{H}}}{n} \tag{4.96}$$

where

$$C_{\tilde{H}} = \frac{1}{2\left[\log\left(\left(\log\frac{\sigma_h^2}{\sigma_l^2}\right)\frac{\sigma_h^2}{\sigma_h^2-\sigma_l^2}\right)\right]^2} + \frac{1}{2\left[\log\left(\frac{\sigma_h^2-\sigma_l^2}{\sigma_l^2\log\left(\frac{\sigma_h^2}{\sigma_l^2}\right)}\right)\right]^2} \tag{4.97}$$

Substitute (4.8) into (4.97), we have

$$C_{\tilde{H}} = \frac{1}{2\left[\log\left(\frac{r}{r-1}(\log r)\right)\right]^2} + \frac{1}{2\left[\log\left(\frac{r-1}{\log r}\right)\right]^2} \tag{4.98}$$

Since

$$v_{\tilde{H}} \geq 1 - e_{\tilde{H}} \tag{4.99}$$

Substituting (4.98) into (4.96), we have

$$v_{\tilde{H}} \geq 1 - \frac{C_{\tilde{H}}}{n} \tag{4.100}$$

In this section, we use the lower bound of $v_{\tilde{H}}$ as the estimation of the detection rate by sample entropy. Consider that detection rate must be greater than 50% and we can get (4.80) in Theorem 4.3. The theorem is proved.

In the following we prove that $v_{\tilde{H}}$ is an increasing function in terms of *r*. That is, we need to prove that $C_{\tilde{H}}$ is a decreasing function of *r*. Denote

$$C_{H1}(r) = \frac{r-1}{\log r} \tag{4.101}$$

$$C_{H2}(r) = \frac{r}{r-1} \tag{4.102}$$

If $C_{H1}(r)$ in (4.101) and $C_{H2}(r)$ in (4.102) are increasing functions of $r$, $C_{\tilde{H}}$ is a decreasing function of $r$.

We can see that $C_{H1}(r)=1/ C_{Y1}(r)$, where $C_{Y1}(r)$ is defined in (4.64). Since $C_{Y1}(r)$ is a decreasing function of $r$, $C_{H1}(r)$ is an increasing function $r$. $C_{H2}(r)=C_{Y2}(r)$, where $C_{Y2}(r)$ is defined in (4.65). We have proved that $C_{Y2}(r)$ is an increasing function, so is $C_{H2}(r)$.

Thus $v_{\tilde{H}}$ is of $r$, where $r>1$.

4.5.5. Proof of Theorem 4.5

Theorem 4.5: When sample size $n < M$, a closed form of detection rate is as follows:

$$v_d = 1 - \left( \frac{M-n+1}{M} \varepsilon + \frac{n-1}{2M} \right) \tag{4.103}$$

where error rate $\varepsilon = 1-v$, and $v$ can be calculated in (4.9), (4.10) or (4.12) when the adversary uses different features respectively.

*Proof:* We assume that the user payload traffic rate is either low (denoted as $\omega_l$) or high (denoted as $\omega_h$) and the rate varies periodically with each rate lasting for half of the period. Recall how the adversary does the rate recognition. After getting a sample, she calculates its feature measurement $s$, whose PDF is $p(s)$. Offline she obtains the decision

boundary $d$. For two class ($\omega_l$, $\omega_h$) classification problem, when $s<d$, the sample is classified as from $\omega_l$, otherwise from $\omega_h$.

First let's check the elements of a sample of size $n$: there are $m$ elements, $X_1, \ldots, X_m$, from $\omega_l$, whose RTTs satisfy $N(\mu_l, \sigma_l^2)$, and the other elements, $Y_1, \ldots, Y_{n-m}$, are from $\omega_h$, whose RTTs satisfy $N(\mu_h, \sigma_h^2)$. Thus, the sample is $\{X_1, \ldots, X_m, Y_1, \ldots, Y_{n-m}\}$.

For a sample consisting of packets from both $\omega_l$ and $\omega_h$, the adversary should classify the sample as belonging to the rate whose packets show up first in the sample. Let's calculate the error rate under this definition of a correct classification.

Define $\omega_{l,k}$ as the event that elements from $\omega_l$ show up first, and $\omega_{h,k}$ as the event that $k$ elements from $\omega_h$ show up first, then

$$\Pr(error) = \sum_{k=1}^{n} \Pr(error|\omega_{l,k}) \Pr(\omega_{l,k}) + \sum_{k=1}^{n} \Pr(error|\omega_{h,k}) \Pr(\omega_{h,k}) \tag{4.104}$$

where

$$\Pr(error|\omega_{l,k}) = \Pr(s_{\omega_{l,k}} > d) \tag{4.105}$$

$$\Pr(error|\omega_{h,k}) = \Pr(s_{\omega_{h,k}} > d) \tag{4.106}$$

Under the assumption of that $n < M$, we now check different cases of $\omega_{l,k}$ and $\omega_{h,k}$ for a sample of size $n$ ($<M$).

We have the following cases in which the payload traffic rate is classified as $\omega_l$: ($M$-$n+1$) number of $\omega_{l,n}$, $\{\omega_{l,n-1}, \omega_{h,1}\}$, $\{\omega_{l,n-2}, \omega_{h,2}\}$, …, $\{\omega_{l,1}, \omega_{h,n-1}\}$. We have the following

cases in which the payload traffic rate is classified as $\omega_h$: ($M$-$n$+1) number of $\omega_{h,n}$, $\{\omega_{h,n-1}, \omega_{l,1}\}$, $\{\omega_{h,n-2}, \omega_{l,2}\}$, …, $\{\omega_{h,1}, \omega_{l,n-1}\}$. Thus

$$\Pr(\omega_{l,k}) = \begin{cases} \dfrac{M-n+1}{2M}, & k = n \\ \dfrac{1}{2\mathrm{M}}, & 1 \le k \le n-1 \end{cases} \tag{4.107}$$

and

$$\Pr(\omega_{h,k}) = \begin{cases} \dfrac{M-n+1}{2M}, & k = n \\ \dfrac{1}{2\mathrm{M}}, & 1 \le k \le n-1 \end{cases} \tag{4.108}$$

When we don't consider the order of elements in a sample,

$$\omega_{l,k} = \omega_{h,n-k}, \ when \ k \neq n \tag{4.109}$$

Thus

$$\Pr(s_{l,k}) = \Pr(s_{h,n-k}), \ when \ k \neq n \tag{4.110}$$

Noticing this fact, we have

$$\Pr(s_{\omega_{l,k}} > d)\frac{1}{2M} + \Pr(s_{h,n-k} < d)\frac{1}{2M} = \frac{1}{2M} \tag{4.111}$$

Substitute (4.107) and (4.108) into (4.104) and reorder the elements,

$$\Pr(error) = \sum \Pr(error|\omega_{l,k})\Pr(\omega_{l,k}) + \sum \Pr(error|\omega_{h,k})\Pr(\omega_{h,k}) \quad (4.112)$$

$$\begin{aligned} = & \frac{M-n+1}{2M}\left(\Pr(error|\omega_{l,n}) + \Pr(error|\omega_{h,n})\right) \\ & + \frac{1}{2M}\sum_{k=1}^{n-1}\left(\Pr(error|\omega_{l,k}) + \Pr(error|\omega_{h,n-k})\right) \end{aligned} \quad (4.113)$$

Since we assume that the high-rate traffic and low-rate traffic have the same probability to happen, it's easy to see that

$$\varepsilon = \frac{1}{2}\left(\Pr(error|\omega_{l,n}) + \Pr(error|\omega_{h,n})\right) \quad (4.114)$$

where ε is the error rate for the static case of the two-traffic-rate classification. Obviously,

$$\Pr(error|\omega_{l,k}) + \Pr(error|\omega_{h,n-k}) = \Pr(s_{\omega_{l,k}} > d) + \Pr(s_{\omega_{h,n-k}} < d) \quad (4.115)$$

$$= \mathbf{1} \quad (4.116)$$

Substitute (4.114) and (4.116) into (4.113),

$$\Pr(error) = \frac{M-n+1}{M}\varepsilon + \frac{n-1}{2M} \quad (4.117)$$

Since success rate *v = 1 - Pr*(*error*), we have (4.15) instantly.

## 4.6. Summary

In this section, we have evaluated the security of systems under active traffic analysis attacks. To demonstrate the threat from such attacks, we assume that the adversary uses ping probing to derive user payload traffic rates. We found that by measuring statistics of the round trip time of ping packets injected into security gateways, the adversary can achieve its objective to determine and track the payload traffic rates. This is true even if a powerful link padding scheme, such as VIT padding, has been used.

Of the possible statistics, sample entropy is an effective and robust feature statistic to explore the correlation between user payload traffic rate and the round trip time of probing ping packets. The reason for the success of the exploit is that payload traffic causes small disturbances to the RTT of ping packets. Moreover, the higher the user traffic rate, the larger this disturbance and hence a larger entropy.

Under the framework of statistical pattern recognition, we have formally modeled systems with different statistical features. Our empirical results match well with our theoretical analysis. Our framework can be easily extended to analyze other statistical analysis attacks. We have also conducted extensive experiments in various situations including a laboratory, campus networks, and the Internet. We have found that for campus networks, the ping probing attack can still achieve a high detection rate. This extensive empirical data consistently demonstrate the usefulness of our formal model and correctness of detection rate predicted by the closed-form formulae.

To improve the system security, we have proposed a random delay method to counter the active probing attack. Our experiments and theories have shown the effectiveness of this scheme.

# 5. CONNECTIVITY ANALYSIS ATTACKS AND COUNTERMEASURES

In this section, we study connectivity analysis attacks and their countermeasures. In particular, we focus connectivity analysis attacks against communication privacy in a wireless anonymous communication system. In this kind of attack, an adversary embeds a recognizable pattern of marks into wireless traffic flows by electromagnetic interference. We propose a new countermeasure based on digital filtering technology.

Concerns about privacy have gained more attention with the rapid growth and public acceptance of the Internet as a means of communication and information dissemination. Communication privacy has become necessary and legitimate in many scenarios, such as anonymous web browsing, E-Voting, and E-Commerce. In each of these scenarios, encryption alone cannot hide the communication relationship between users [11][21].

Since Chaum [12] pioneered the basic idea of the anonymous communication system referred to as mixes for hiding the communication relationship between users, researchers have developed various anonymity systems for different applications. Guan, Fu, Xuan, Shenoy, Bettati, and W. Zhao [24][35] study the problem of preventing passive traffic analysis attacks in a mission critical system. In this section, we study how to hide communication relation from active traffic analysis attacks in a general (best-effort) communication system.

Although a significant amount of effort has been made in wired networks, not enough attention has been paid to hide communication relationship in wireless environments. In this section, we consider a broad range of wireless networks, ranging

from networks with all links being wireless to hybrid wired and wireless networks. The wireless links can be either 802.11 (or its extensions) or Bluetooth. A wireless network may use existing mix techniques to provide anonymity for flow-based applications such as anonymous web browsing. We study three mix batching approaches, which are feasible for a flow-based wireless mix network and find that they are all susceptible to a new flow-level attack, which we call the flow marking attack.

The remainder of this section is organized as follows: We first introduce the wireless mix network model and adversary threat model. We then discuss flow marking attack and related issues and use experiments to evaluate its threat against communication privacy. Finally we develop a digital filter-based countermeasure to flow marking attacks and empirically prove its feasibility.

## 5.1. Models

In this section, we first present the model of mix network, and then describe the wireless mix network model used in this section. Finally, we introduce the threat model.

### 5.1.1. Mix Network

A traditional mix is a relay server for anonymous email communication [12]. It has a public key which senders use to encrypt messages. A mix operates as follows:

1. The sender attaches the receiver address to the message and encrypts the entire package by using the mix's public key;
2. The mix collects a batch of messages (from different senders), and decrypts them to obtain the receiver addresses;

3. Finally the mix sends decrypted messages out in a rearranged order to their corresponding receivers. Batching and reordering are necessary techniques for a mix to reduce or eliminate the correlation between input messages and output messages. This kind of correlation may help an adversary to identify flow connectivity.

A mix network consists of multiple mix servers and can provide enhanced anonymity. In a mix network, senders route their messages through a series of mixes. Therefore, even if an adversary compromises one mix and discovers the correlation between its input and output messages, other mixes along the path can still provide the necessary anonymity [17][36][37]. Figure 17 illustrates the route selection for one message. A sender can choose different routes for each message or use one route for all her messages.

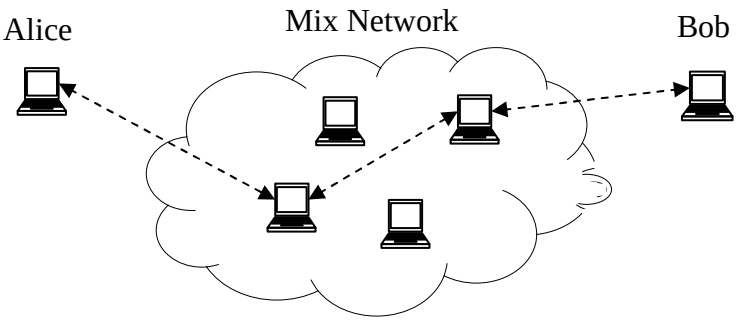



Figure 17. Mix Network

Message-based mix networks have been extended to flow-based networks for applications such as anonymous FTP, Web browsing, video and audio transmission, and many other low-latency applications. In the context of an IP network, the relay servers in Figure 17 form an overlay network and forward packets instead of messages.

Researchers have paid attention to attacks exploring *packet-level correlation* in anonymous communication systems. However, this is not sufficient and sometimes misleading, since most of today's communications are flow based, with the majority using TCP. On the Internet, TCP flows constitute 60%~90% of the Internet traffic and UDP flows constitute 10%~40% [38][39], while all other protocols combined produce less than 5% traffic. On the Sprint IP backbone, new applications such as distributed file sharing and streaming media using TCP and UDP flows constitute 60% of the traffic on some links, while 30% is web traffic [40]. Traffic flows consist of rich features that can be explored to compromise anonymity systems.

Major differences between flow-based systems and message-based systems are as follows:

1. Flow-based systems usually do not use dummy packets to pad the traffic in order to counter traffic analysis attacks. This is because dummy packets consume additional bandwidth and reduce efficiency [27].
2. Flow-based systems usually adopt static routing, i.e., one path per flow, in order to avoid the difficulty and overhead caused by using multiple routes for TCP connections and to prevent intersection attacks [37]. This practice coincides with

the design of several existing systems, including Crowds [41], Tor [42], and many others.

3. Batching and reordering increase the (worst case) delay and are less preferred methods in flow based systems [23]. However, they may be necessary to counter packet-level timing correlation attacks.

In this section, we will investigate the security of flow-based systems with several different configurations. In [43], a complete list of batching strategies for a message-based mix has been provided to counter message-level timing attacks. In our opinion, not all of them are appropriate for flow-based systems. For example, in a threshold mix, a mix can transmit the batch of packets only if the number of packets it collects has gone beyond a pre-defined threshold. This may cause serious problems for traffic of TCP flows, for instance, if the first (SYN) packet cannot be exchanged between a sender and receiver, the TCP flow cannot start, and hence, the entire mix network may not be stable. We select three batching strategies which appear to be feasible for a flow-based mix network. We summarize them in Table 1.

Table 1. Batching Strategies

| **Strategy Index** | **Name** | **Adjustable Parameters** | **Algorithm** |
|---|---|---|---|
| $S_0$ | Simple Proxy | *None* | No batching or reordering |
| $S_1$ | Timed Mix | $<t>$ | If timer of period $t$ fires, send all the packets queued in the last interval |
| $S_2$ | Stop-and-go Mix (Continuous Mix) | $<\mu, \sigma^2>$ | Each packet is assigned a delay (deadline) satisfying a distribution with mean $\mu$ and variance $\sigma^2$. A packet is sent out when its deadline is reached |

5.1.2. Wireless Networks

In this subsection, we introduce the wireless network model used in this section. We note that there are two popular radio frequency (RF) technologies: IEEE 802.11 [44] (and its extensions such as 802.11a/b/g) and Bluetooth [45].

The IEEE 802.11 standards are widely adopted for wireless LAN (WLAN). Two types of WLAN are supported: one is the infrastructure mode and the other ad-hoc mode. In the infrastructure mode, a station acts as the access point (AP) centrally controlling the WLAN, and other mobile units communicate with the AP. A WLAN in the infrastructure mode is denoted as the basic service set (BSS). In the ad-hoc mode, an AP

does not exist. All mobile units (MUs) communicate within their transmission range. Ad-hoc routing protocols, such as DSDV [46], DSR [47], AODV [48], and many others, have been developed to extend the range and flexibility of ad-hoc networks. A WLAN in the ad hoc mode is also denoted as an Independent Basic Service Set (IBSS). An Extended Service Set (ESS) consists of multiple BSS/IBSS interconnected by access points and a distribution system, such as ethernet.

Bluetooth is a low cost, low-power, short range radio technology, originally designed as a cable replacement to connect devices such as mobile phone handsets, headsets, and portable computers. In Bluetooth, a group of two to eight Bluetooth units forms a *piconet*, sharing the same wireless channels (hopping sequence). In a piconet, any, but only one, unit can act as the *master* of the piconet, and the others are *slaves*. The master implements centralized control, and only communication between the master and slaves is allowed. The communication between two slaves is relayed via the master. Piconets can be interconnected and form a *scatternet*. Routing algorithms are proposed in [49] and many others for efficient communication between Bluetooth units (BU) in a scatternet. We focus on the standard of Bluetooth 1.1 because of its popularity.

### 5.1.3. Wireless Mix Network

Most anonymity communication systems are built as overlay networks. Thus, wireless units (MUs or BUs) can use mixing strategies discussed above and form a wireless mix network. This section assumes an ESS-like network with combined wireless (Bluetooth or 802.11) and wired links, in which any host can act as a mix. For example, in Figure

17, Alice (sender) and Bob (receiver) can be mobile units, and they may communicate with each other through a wireless or wired mix network.

### 5.1.4. Adversary Model

In the following, we summarize the adversarial assumptions considered in this section:

1. The content of wireless communication between legal participants is protected by underlying encryption algorithms and immune to any attack.
2. The adversary is an external one and therefore is not a legal participant of the wireless network.
3. The adversary can passively eavesdrop on the communication session. We will show that eavesdropping wireless links can be easily realized in Section 5.3.
4. The adversary can actively interfere with wireless networks by injecting interference traffic. We assume that the adversary uses a reasonably good directional antenna, allowing it to interfere with a selected victim with minimum disturbance to other wireless units [50][51][52].

## 5.2. Flow Marking Attack

In this subsection, we introduce the flow marking attack and related issues.

### 5.2.1. Overview and Problem Definition

Figure 18 illustrates the basic idea of a flow marking attack. Alice is communicating with Bob through a wireless mix network. The goal of the adversary is to find if Alice is communicating with Bob. A component of the adversary, *interferer*, first embeds a series of *marks* into Alice's traffic by interfering with her link. Another component of the adversary, *sniffer*, then eavesdrops on Bob's inbound traffic. The interferer and the sniffer communicate with each other and/or report their actions and findings to the adversary headquarter. If the sniffer discovers a pattern of marks in Bob's traffic that is similar to that embedded by the interferer, the adversary can be sure that Alice is indeed communicating with Bob.

Thus, the general problem of the flow marking attack can be defined as follows: given a series of marks embedded into a flow, how can an adversary recognize them at a location somewhere along this path of the same flow?

Flow marking is a general technique and can be used in both wired and wireless networks. In wired networks, an adversary may explore TCP's characteristics and use efficient denial of service approaches [53] to introduce marks. In wireless networks, an interferer can use electromagnetic interference to embed marks into traffic. This is the focus of this section.

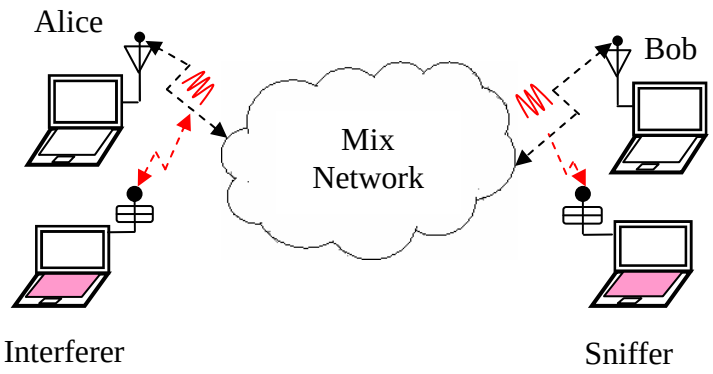



Figure 18. Flow Marking Attack Scenario

### 5.2.2. Issues of Flow Marking Attack

From the viewpoints of both adversaries and defenders, there are four critical issues related to the problem of flow marking attacks:

1. How can an adversary introduce marks into traffic flows and intercept traffic?
2. How can an adversary effectively recognize if marks exist?
3. How effective and efficient can the flow marking attack be in reality?
4. How can we counter flow marking attack to minimize its effectiveness?

We intend to address these issues in the following sections.

## 5.3. Mark Embedding and Traffic Interception

In this section, we discuss two key issues related to embedding marks into wireless traffic and intercepting wireless traffic.

### 5.3.1. Overview of Radio Frequency Communication

The physical layer of IEEE 802.11 and Bluetooth is where interference may happen. IEEE 802.11 (and its extensions) has two different physical layers: frequency hopping (FHSS) layer and direct sequence (DSSS) layer[5]. Bluetooth uses FHSS. Both IEEE 802.11 and Bluetooth use license-free ISM (industrial, scientific, and medical) radio frequency (RF) band from 2.4GHz to 2.5GHz. This band is divided into many channels.

In this section, we assume that an adversary uses a (laptop) computer equipped with an 802.11b (DSSS) PCMCIA card to apply the interference and hence to embed marks. Below we will focus on how the interference and interception can be realized. Related RF specifications are based on the regulation of American Federal Communications Commission. Please refer to [43] and [45] for RF regulations in other regions.

### 5.3.2. Interfering with and Intercepting Wireless Communication

It's straightforward to interfere with and intercept 802.11 DSSS communication. There are 11 channels available, Channels 1 to 11. Hosts in the same channel can interfere with and intercept one another. Furthermore, only Channels 1, 6 and 11 are free of interference with each other, but adjacent channels may interfere with each other.

In FHSS, the transceiver must be synchronized. With both 802.11 and Bluetooth, the ISM band is divided into 79×1MHz channels. The synchronized transmitter and receiver communicate on a series of channels, denoted as *hopping pattern* or *hopping sequence* and only remain on one channel for a predefined amount of time, denoted as *dwell time*.

[5] Today, most of 802.11 products use DSSS because of its high throughput.

An 802.11 DSSS device can interfere with an 802.11 FHSS device since 802.11 FHSS hopping sequence visits the DSSS channel and its adjacent channels regularly, hence potentially causing interference with each other. Intercepting the 802.11 FHSS traffic is not difficult since 802.11 FHSS has only 78 possible hopping sequences divided into 3 sets, and the adversary can know the whole hopping sequence by observing a small fragment of communication using an appropriate spectrum analyzer [54]. Then the adversary can adjust her own 802.11 FHSS device to synchronize with the victim 802.11 device and intercept the traffic. Of course, a full ISM band analyzer can easily intercept 802.11 FHSS traffic.

In general, an 802.11 DSSS device can cause more interference to Bluetooth traffic than to 802.11 FHSS traffic since a Bluetooth device visits a fixed DSSS channel more frequently. Bluetooth's hopping sequence has a dwell time of 625 microseconds, which corresponds to 1600 hops/s. The Bluetooth specification also requires that the hopping sequence distribute the hop frequencies equally over the 79 MHz during a short time interval. An 802.11 FHSS device's hopping rate is often in the order of tens of hops per second.

It is still possible to intercept Bluetooth communication, although Bluetooth's hopping sequence has a very long period length and does not show repetitive patterns over a short time interval. The method has a few defects. First, a piconet uses clear-text frequency hopping sequence (FHS) packets to exchange hopping sequence information between the master and slaves. An adversary can intercept FHS packets, synchronize with the master, and then eavesdrop on the communication. Second, the adversary may

have sophisticated Bluetooth listening devices to sniff the communication [55]. Again, a full ISM band analyzer can easily intercept Bluetooth traffic.

## 5.4. Mark Recognition by Feature Frequency

In this section, we address two issues of the flow marking attack: how to choose an effective pattern of marks and how to recognize marks.

### 5.4.1. Effective and Efficient Marks

An effective pattern of marks for flow marking attacks must demonstrate uniqueness. That is, the adversary can be certain of recognizing the same series of marks at one location as the one it introduces at another location. Because of the inherent nature of the Internet traffic, an arbitrary pattern of marks may not be effective and efficient for flow marking attacks.

In this section, we demonstrate that a periodic pattern of marks can be effective and efficient. That is, an adversary may use *on-off traffic* with a period of $T_I$, denoted as *interference period*, to interfere with the victim traffic. During an *on period*, the interfering device transmits traffic at a rate as high as possible. This will reduce the available bandwidth for the victim traffic or disrupt packets of the victim traffic. During an *off period*, the interfering device becomes silent and the victim traffic gains the lost bandwidth. In this way, the adversary forces the victim traffic to adapt to the pattern of the interfering traffic and the victim traffic develops a replicate pattern. The adversary can choose a relatively unique interference period (compared with the background noise

traffic) to achieve a series of unique and strong marks within the network. We use an on period (approximately) equal to the off period, with each lasting for $T_I/2$.

Depending on where interference is deployed in the path of the flow, the flow marking attack can have different effects on different types of flows. For TCP flows, the interference location can be flexible. The adversary can apply interference at any point along a TCP flow's path (i.e., at the sender, an intermediate mix, an intermediate hop, or the receiver). Since TCP uses a loop-control mechanism [56], a TCP flow will demonstrate the similar periodicity along its path from the sender to the receiver. For UDP traffic, an adversary may have to deploy the attack as close to the sender as possible. We will focus on the effect of flow marking attacks on TCP flows because of their dominant status on the Internet.

### 5.4.2. Flow Marking Attack Framework

Now we summarize the framework of flow marking attacks based on pattern recognition [27] in Figure 2. Although we have introduced a similar framework in Section 1, we prefer to give a relatively detailed introduction in the current context for the ease of understanding the topic in this section.

Recall that in a flow marking attack, an adversary tries to discover if Alice is communicating with Bob by checking if the intentionally embedded pattern of marks exist in both Alice's outbound traffic and Bob's inbound traffic. The adversary has to decide what the pattern is and how to evaluate its existence.

Generally speaking, the goal of the pattern recognition process is to use classifiers to *classify* an unknown pattern as belonging to one of several existing pattern *classes* with

the help of a feature (or a vector of features). A classifier is trained from training data. In a flow marking attack, there are only two classes of events:

$\omega_0$: Alice does not communicate with Bob

$\omega_1$: Alice communicates with Bob (5.1)

Following the common practice, a pattern recognition system for flow marking attacks consists of two phases: (a) offline training phase and (b) online mark recognition phase.

Figure 2 (a) is the flowchart of offline training phase. In the following, we will discuss each step of this phase for flow marking attacks illustrated in Figure 18.

1. Collecting training data: The adversary emulates the entire system. The interferer interferes with Alice's wireless link and dumps Alice's interfered traffic or records the adversary's own interference traffic. In general, an adversary may not achieve a perfect periodic interference and needs either her own or Alice's traffic to derive the actual interference period. The sniffer intercepts Bob's inbound traffic.
2. Preprocessing training data: The collected data sample will be divided into segments, each of which contains packets within an interval, $T_s$, denoted as *sampling interval*. Thus, the number of packets in each segment forms a time series. The number of segments in the sample is denoted as *sample size*. This time series of packet counts is denoted as follows:

$$X(T_I, T_S) = \{x_1, \mathrm{K}, x_N\} \quad (5.2)$$

where $T_I$ is the interference period, $N$ is the sample size and $x_i$ the number of packets in the $i^{th}$ segment. We denote *sample length* as the lasting time of the traffic sample and it is equal to $NT_s$.

3. Selecting feature from preprocessed training data: This is the key step for flow marking attacks. An appropriate feature extracted from $X(T_I, T_s)$ should represent the pattern of marks.

   In flow marking attacks, because the adversary artificially introduces periodicity into the victim traffic, when Fourier transform is applied to $X(T_I, T_s)$, strong amplitudes will be observed around the frequency of $1/T_I$, denoted as *feature frequency*.

4. Selecting decision rule: If the sniffer can observe the feature frequency in Bob's traffic, she can be sure that Alice is communicating with Bob. Here, we have an implicit assumption: without interference, the amplitude at the feature frequency is not significant. The adversary collects training traffic without applying the interference and derives this *a priori* knowledge of statistics of the amplitude during the offline training phase.

In this section, we assume the adversary uses the Bayes decision rule for flow marking attacks. To use the Bayes decision rule, the adversary collects training traffic without applying the interference and derives the *a priori* probability density function (PDF) of the amplitude at the supposed feature frequency. She collects data by emulating the flow marking attack and obtains the PDF of the amplitude at the feature

frequency. From these two statistics, the adversary generates the following Bayes decision rule:

*The amplitude at the feature frequency implies $\omega_1$ if*

$$p(\omega_1 \mid a) \geq p(\omega_0 \mid a) \tag{5.3}$$

*That is,*

$$p(a \mid \omega_1)\Pr(\omega_1) \geq p(a \mid \omega_0)\Pr(\omega_0) \tag{5.4}$$

where $a$ is the measured amplitude of the feature frequency, $Pr(\omega_i)$ ($i=0,1$) is the *a priori* probability that Alice is communicating with Bob or not (set as 50% in this section), $p(\omega_i|a)$ is the *a posteriori* probability that Alice is communicating with Bob when the collected sample has the amplitude at the feature frequency, $p(a|\omega_0)$ is the PDF of the amplitude of the feature frequency conditioned when Alice and Bob are not communicating with each other and $p(a|\omega_1)$ is the PDF of the amplitude of the feature frequency conditioned when Alice and Bob are communicating with each other.

From (5.4), the decision boundary $d$ can be derived if we solve the following equation:

$$p(a \mid \omega_1)\Pr(\omega_1) = p(a \mid \omega_0)\Pr(\omega_0) \tag{5.5}$$

Thus, the rule is,

*Alice is communicating with Bob if $a>d$.*

Refer to Figure 2 (b) for procedures in the online recognition phase. The procedure is similar to the offline recognition phase. The difference is that here, the network is

realistic. During the online recognition phase, the adversary interferes with Alice's wireless traffic and measures the feature frequency of Bob's inbound traffic. Then the adversary makes a decision by the Bayes decision rule based on the measured amplitude at the feature frequency.

### 5.4.3. Detection Rate as Evaluation Criterion

Detection rate is defined as the probability that an adversary correctly recognizes the fact that Alice is communicating with Bob. To derive the detection rate for the Bayes decision system, the adversary has to estimate *a posteriori* probability distribution of the feature frequency power amplitude in power spectrum for classes $\omega_0$ and $\omega_1$. We assume that the adversary uses a Gaussian kernel function based method to estimate density functions [28].

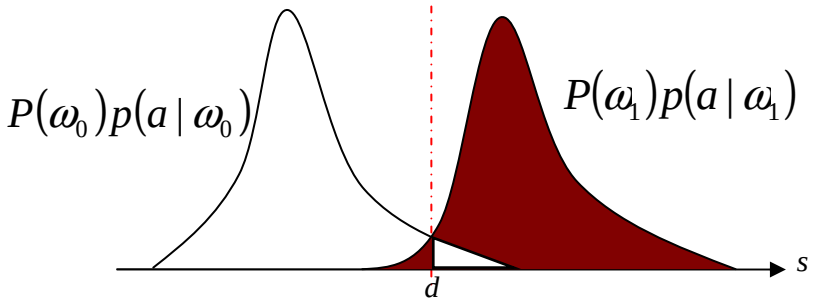



Figure 19. Bayes Decision Rule for Flow Marking Attack

As showed in Figure 19, once $p(a|\omega_0)$ and $p(a|\omega_1)$ are derived, detection rate can be calculated in (5.6).

$$v = P(\omega_0)\int_{-\infty}^{d} f(s|\omega_0)ds + P(\omega_1)\int_{d}^{\infty} f(s|\omega_1)ds \quad (5.6)$$

5.4.4. Selection of Interference Interval and Sampling Interval

In flow marking attacks discussed above, there are two parameters: interference period $T_I$ and sampling interval $T_s$. These parameters are critical to the effectiveness and efficiency of a flow marking attack.

We claim that the sampling interval should be smaller than half of the interference period. That is,

$$T_s < T_I / 2 \quad (5.7)$$

This claim can be justified as follows. When we count packets in a sampling interval and derive the packet count time series in Step 3, as shown in Figure 2 (a) and Figure 2 (b), this process is similar to a *zero-order hold* [57] sampling process. We know the feature frequency is $1/T_I$, which has to be preserved for the best effectiveness of flow marking attack. Nyquist's sampling theorem [57] suggests that to preserve this feature frequency, the sampling rate should be at least twice of the feature frequency. That is,

$$1/T_s < 2/T_I \quad (5.8)$$

Thus, (5.7) is justified.

The selection of interference period $T_I$ is not arbitrary either. As discussed above, the interference traffic during the on period of the interference period has to decrease the victim traffic rate, and the off period has to be long enough so that the victim traffic can gain the lost bandwidth. Clearly, for 802.11 DSSS, 802.11 FHSS and Bluetooth, there are different requirements for choosing $T_I$ because of their different physical and protocol characteristics.

The interference period cannot be too long since in practice, a flow may only last for a short time. For example, the duration of a FTP session is determined by the corresponding file size.

Interference period is also related to the requirement of sample length for the effectiveness of flow marking attack. By the definition of a feature frequency, an adversary must sample for at least one complete cycle of interference. Otherwise, she could not recognize the feature frequency [57]. Thus, the sample length of $NT_s$ should be greater than the interference period, i.e.,

$$T_I \leq NT_s \tag{5.9}$$

## 5.5. Evaluation of Flow Marking Attack

In this section, we empirically show the failure of a wireless mix network under a flow marking attack (FMA) in a laboratory environment and discuss related issues.

### 5.5.1. Experiment Environment

Figure 20 illustrates the experiment setup in the laboratory. It is a typical one-mix anonymous communication network with wireless links, i.e., an ESS-like wireless network. Alice uses FTP to download a file from Bob through a mix. To simplify our discussion, we assume that only Alice's link is wireless, and she communicates with other parts of the network through a machine performing access-point-like functions. We also install NISTNet [58] on this access-point-like computer to simulate delay and other network dynamics when necessary. One computer acts as a noise generator to inject noise traffic. In this way, we can evaluate the impact of noise on the performance of flow marking attacks.

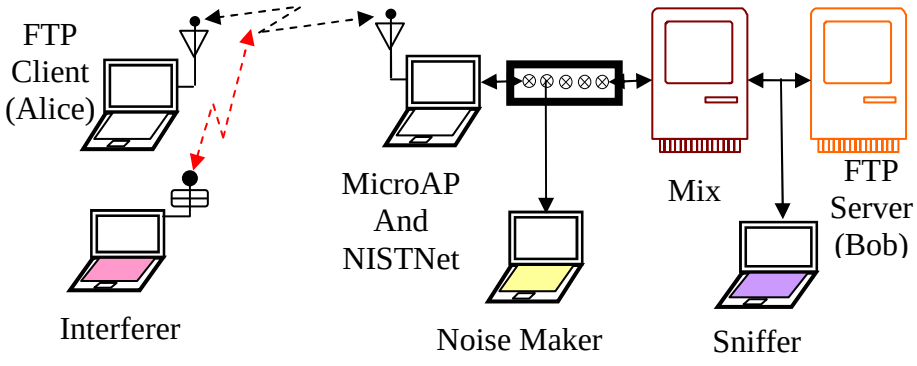



Figure 20. Experiment Setup

Mixing strategies are implemented on the TimeSys/Real Time Linux operating system for its timer accuracy [29]. We integrate the mix control module performing batching and reordering functions into Linux's firewall sub-system *Netfilter* [59], and firewall rules are used to specify what traffic should be protected.

We use WaveLAN silver PC card as 802.11 DSSS devices, Spectrum24 LA 3021 PC card as 802.11 FHSS devices, and Belkin Bluetooth PC card as Bluetooth devices. Wireless traffic and wired traffic is dumped by tcpdump [60]. Wireless channels can be changed by iwconfig [61].

In our experiments, a timed mix timer has a period of 100ms. The stop-and-go mix assigns an exponentially distributed delay to packets with average delay of *25*ms. This delay cannot be too long, otherwise it may cause a large number of packet reordering and hence seriously disrupt normal behavior of TCP.

### 5.5.2. Failure of Mix Networks under FMA

Figure 21 shows the power spectrum by 64-point FFT for a stop-and-go mix network with an 802.11 DSSS wireless link. We can see that the feature frequency, 2Hz ($1/T_I$), has a strong amplitude compared to the case without flow marking attacks, in which every frequency component has roughly equal amplitudes.

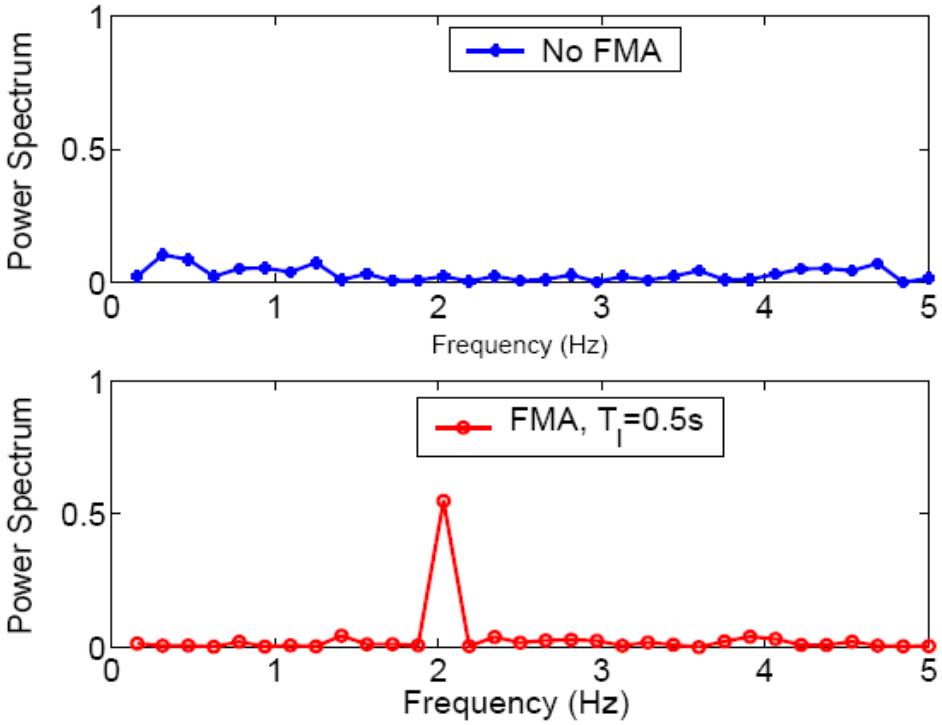
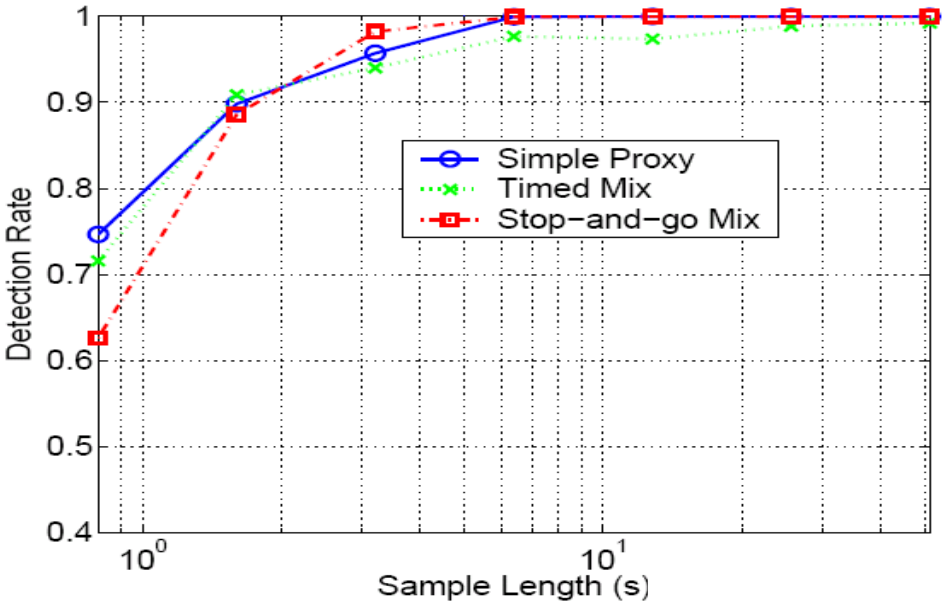
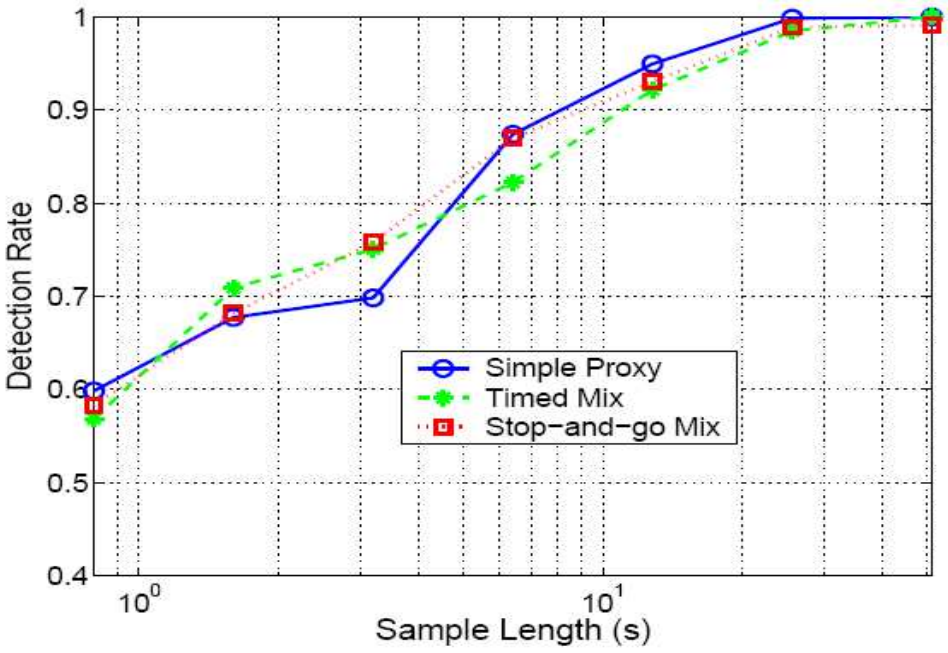



Figure 21. Power Spectrum of 802.11 DSSS Traffic for Stop-and-go Mix

Figure 22 shows the relationship between detection rate and sample length for all the three mixing techniques in Table 1 and three types of wireless links. In all the experiments, interference period $T_I = 0.5$s and sampling interval $T_s = 0.1$s. 802.11 DSSS and 802.11 FHSS links have a bandwidth capacity of *2*Mbps while the Bluetooth link has a bandwidth capacity of *1*Mbps.

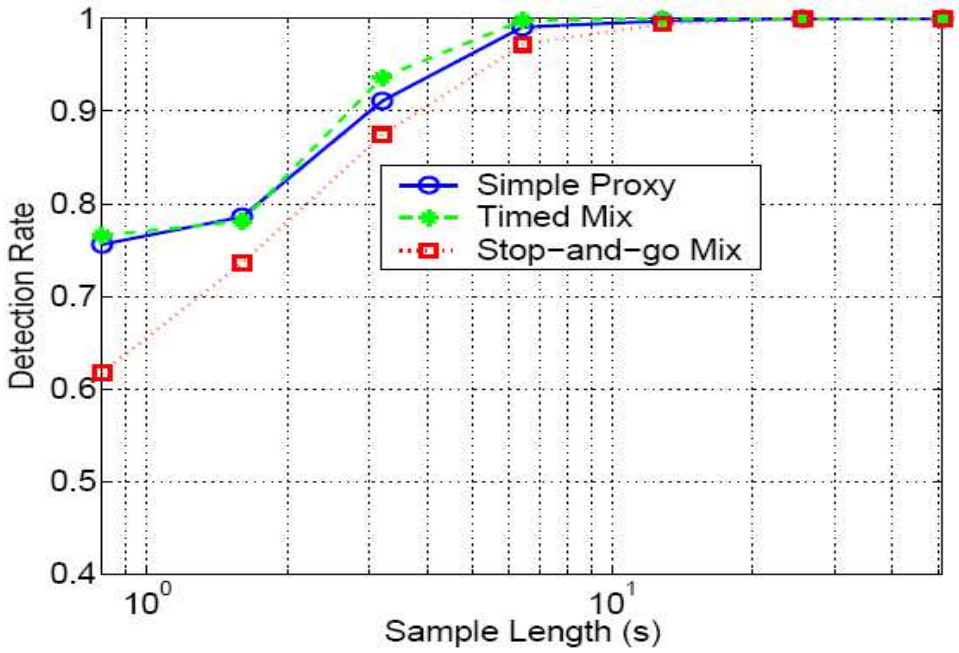



(a) 802.11 DSSS




(b) 802.11 FHSS

Figure 22. Detection Rate by Flow Marking Attack




(c) Bluetooth

Figure 22. Continued

We have the following observations from Figure 22:

1. A wireless anonymous communication system may completely fail under flow marking attacks. As sample length increases, a flow marking attack can achieve a detection rate of *100*% in all cases as shown in Figure 22.
2. An adversary only needs a few seconds of sampling to achieve a detection rate of *100*%. This shows that flow marking attacks can be effective and efficient for online piracy tracing even if an anonymous file exchange service is used on the Internet since most of the file downloading takes longer than a few seconds.

### 5.5.3. Detection Rate vs. Different Wireless Links

Figure 23 compares detection rate for the three different wireless links. Stop-and-go mixes are used in experiments. As we analyze above, since an adversary can use the same 802.11 DSSS channel to interfere with the victim 802.11 DSSS wireless link, she achieves the highest detection rate in this case. Because of a higher hopping rate, a Bluetooth FHSS link is more susceptible to the 802.11 DSSS interference than an 802.11 FHSS link, where the Spectrum24 PC card has a hopping rate of *10* hops/s. The adversary achieves higher detection rate in the case of interfering with a Bluetooth link.

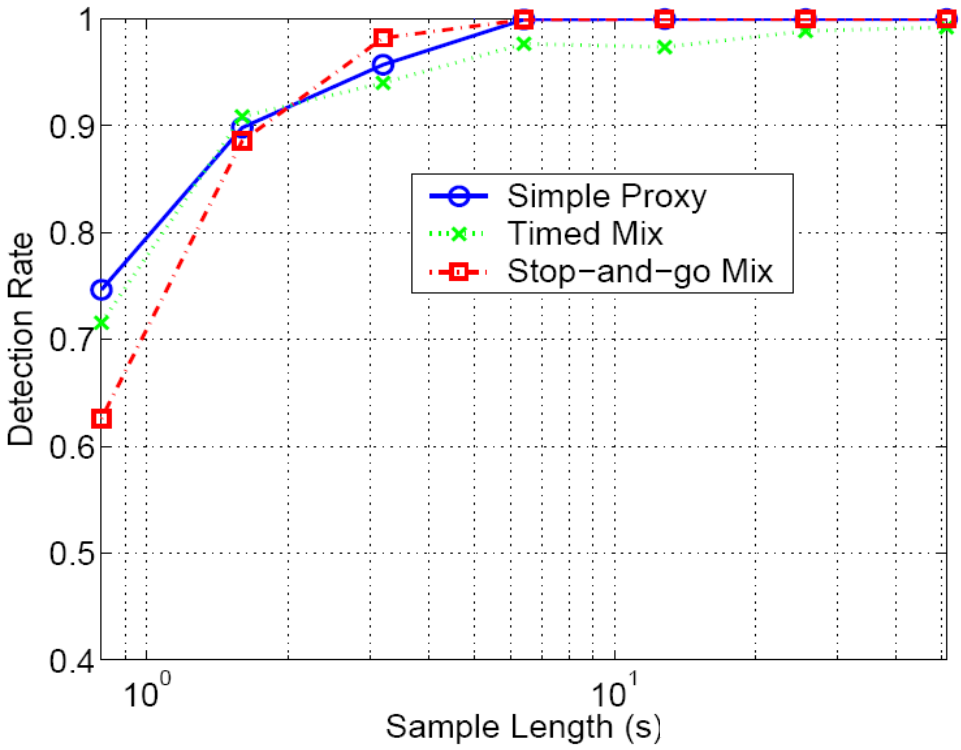



Figure 23. Detection Rate for Different Types of Wireless Links

In the following, we concentrate on properties of flow marking attacks of 802.11 DSSS wireless links. For other cases, we have similar results.

### 5.5.4. Sample Length to Achieve Detection Rate of 95%

Figure 24 shows the minimum amount of time an adversary takes to achieve a detection rate of 95% for each interference period. From Figure 24, we can see that at the interference period of 0.5s, it takes the adversary about 1.6 seconds to achieve a detection rate of 95%. This indicates that there is an *optimal* interference period by which the sample length is minimized. That is, flow marking attacks can be very effective and efficient when the adversary operates in the optimal mode.

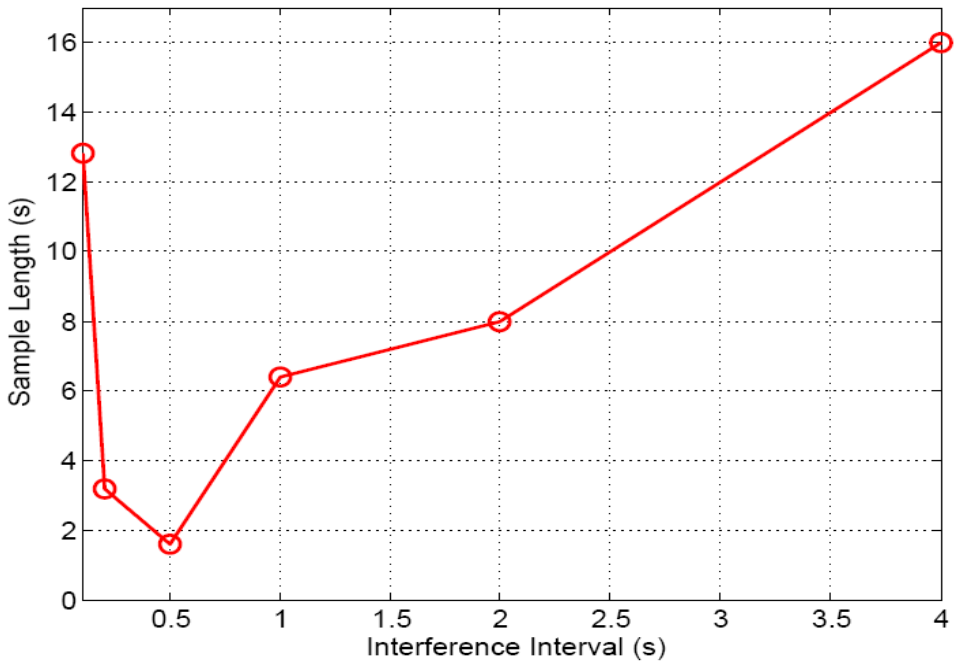



Figure 24. Sample Length Required to Achieve Detection Rate of 95%

We note that the curve is concave up. The reason is as follows: if the interference period is too small, a TCP flow does not have enough time to reduce the rate during interference and to increase the rate during the silent time of the flow marking attack. Thus, the introduced pattern would have a weak appearance in the TCP flow. It may take more time to effectively detect the pattern of marks. On the other hand, if the interference period is very long, from (5.9), we know the flow marking attack needs at lease one interference period of sample to be effective. Thus, the longer the interference period, the larger the sample length. Clearly, the large sample length is caused by the unnecessarily large interference period.

### 5.5.5. Impact of Noise Traffic

Figure 25 shows noise's impact on the effectiveness of a flow marking attack. We use *r* to represent the ratio of the number of noise traffic packets to the number of TCP payload packets. The noise traffic is generated with an inter-arrival time satisfying a Pareto distribution with the shape parameter of 1.5 [62].

We have the following observations:

1. Noise traffic has a clear impact on the performance of a flow marking attack. We can see that as r increases, detection rate decreases. The reason is that noise traffic introduces randomness into the aggregated traffic, and the power spectrum at the feature frequency would have more randomly distributed energy with more noise traffic. This decreases detection rate.

2. Noise traffic's impact on the flow marking attack is limited. We can see that an adversary may still achieve a detection rate of 100% even if r ≈ 5, which corresponds to a 60% utilization rate for Bob's 10Mbps link.

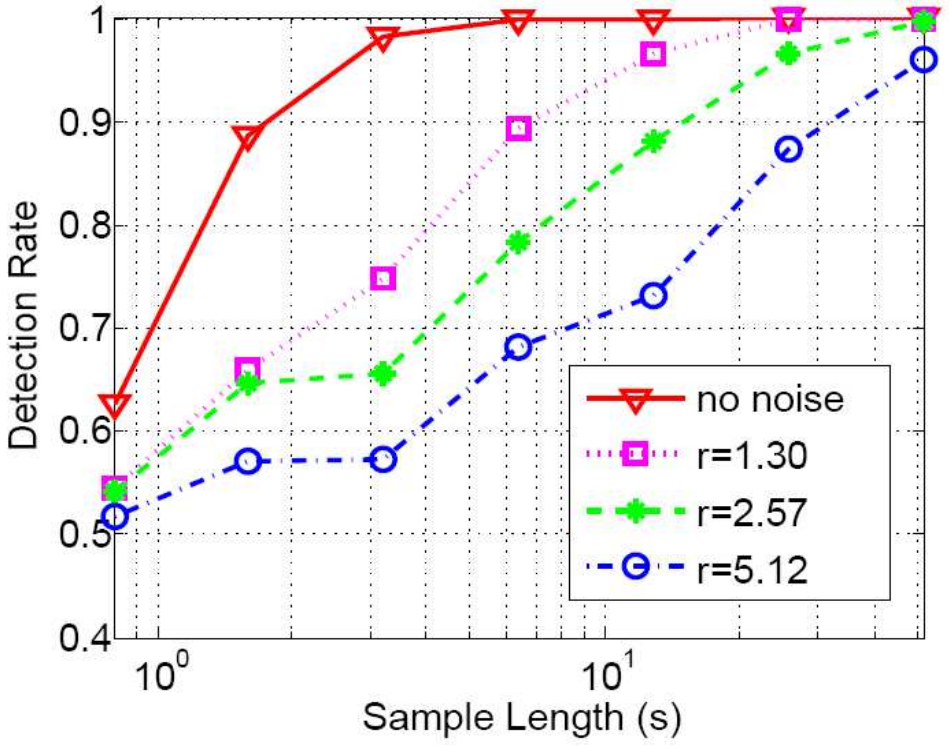



Figure 25. Detection Rate vs. Noise Traffic

## 5.6. Countermeasures by Filtering

### 5.6.1. Overview

In this section, we develop possible countermeasures to flow marking attacks. Our idea is based on signal processing theory. That is, we use digital filters to filter out possible feature frequencies introduced by adversaries.

The filter-based countermeasure works as follows:

1. We deploy filters at locations where traffic shaping and filtering is needed.
2. The filter utilizes a periodic timer of period $T_f$ to sample the traffic rate. It buffers packets arriving in its current timer interval, say the $n^{th}$ interval, and counts the number of packets, $x(n)$, in this interval[6].
3. Then, we can calculate the required number, $y(n)$, of packets we should transmit in order to filter out feature frequencies by using the following formula:

$$y(n) = \sum_{k=0}^{M} a(k)x(n-k) - \sum_{l=1}^{M} b(l)y(n-l) \qquad (5.10)$$

where $M$ is the filter order, $x(n - k)$ and $y(n – k)$ are the number of input packets and output packets of the filter, respectively, during the past $k^{th}$ interval, and $a(k)$ and $b(k)$ are filter coefficients, which are discussed in 5.6.2. Please refer to [63] for general knowledge of the design of a recursive (IIR) filter specified in (5.10).

4. The filter then acts depending on different values of $x(n)$ and $y(n)$: If $x(n) \geq y(n)$, the filter sends out $y(n)$ payload (user) packets when the timer fires and holds the

[6] In fact, $x(n)$ is the sum of incoming packets in the current interval and packets left over from the previous interval. Refer to Step 4.

remaining $x(n)$-y($n$) payload packets, which will be counted into the next round of incoming packets, i.e., $x(n+1) = x(n+1) + x(n) - y(n)$. If $x(n) < y(n)$, the filter generates $y(n) - x(n)$ dummy packets and sends them out with $x(n)$ payload packets;

### 5.6.2. Selection of Filter Coefficients

Filter coefficients have to be carefully chosen for the best performance in countering flow marking attacks. To achieve this objective, we first determine the possible feature frequency band ($F_l$, $F_u$), which should be filtered out. In reality, the interference frequency of an adversary is bounded due to the following reasons: It takes time for the victim traffic to respond to the interference and to reduce its rate. Time is also needed for the victim traffic to gain the bandwidth when the interference stops. This gives feature frequency an upper bound, $F_u$. Moreover, a traffic flow only lasts for a limited interval, for example, the duration of a FTP session is determined by the file size. This gives feature frequency a lower bound, $F_l$.

Then we set a sufficiently large filter order and use the *yulewalk* function from Matlab to derive the filter coefficients $a(k)$ ($k=0, \ldots, M$) and $b(l)$ ($l=1, \ldots, M$). The filter is of a band-stop type, as we just filter out the band of possible feature frequencies. The benefit is that details of traffic are kept and the number of dummy packets can be reduced.

### 5.6.3. Evaluation of Filter-based Countermeasure

Figure 26 gives the detection rate when we put a filter of $T_f = 0$ on MicroAP in Figure 20, where a stop-and-go mix is used. The interference period is 0.5s and the filter has an order of *20*.

We can see that detection rate approaches *50%*, which is the minimum value in a two-class pattern recognition. So traffic filtering can be used as an effective countermeasure for flow marking attacks in combination with mixes in a wireless mix network.

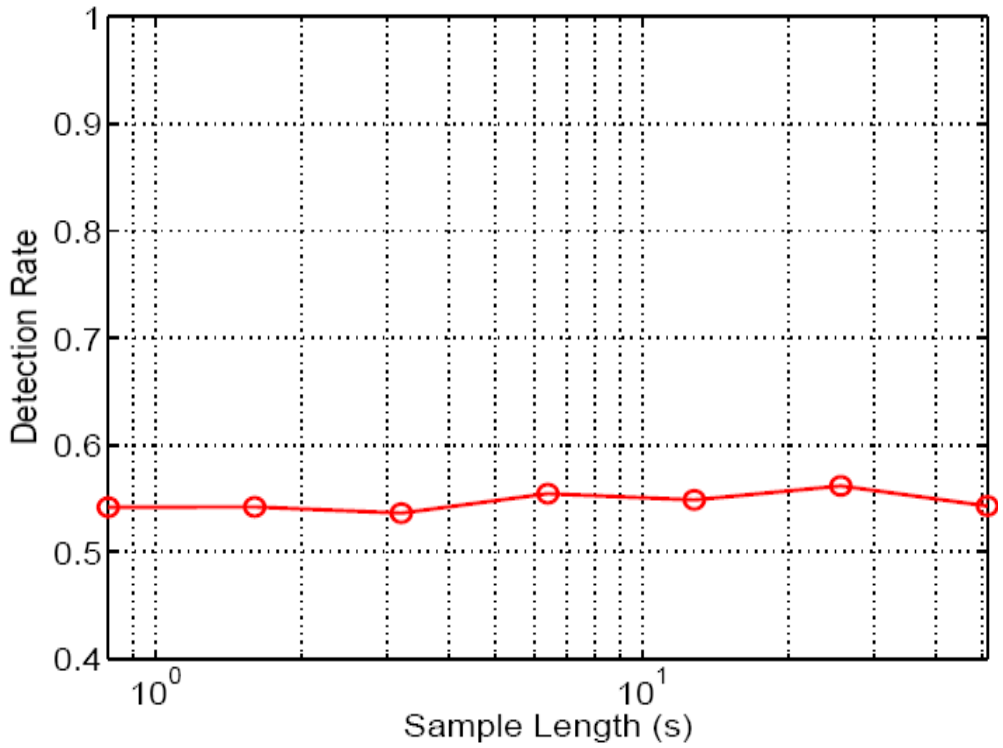



Figure 26. Detection Rate with Filter-based Countermeasure

## 5.7. Summary

This section studies the performance of an anonymous wireless communication system under flow marking attacks. Detection rate is defined as the probability that the adversary finds the communication relationship of “Alice” and “Bob”: whether or not they are communicating with each other. We show that it takes only a few seconds for an adversary to achieve a detection rate of 100%. This is, in a wireless environment, flow marking attacks can be very effective and efficient even if traditional mix technologies are used.

To counter flow marking attacks, we introduce a countermeasure that uses digital filters to filter out the suspect band of feature frequencies. Our filter is an IIR recursive one. We empirically demonstrate the success of this digital filter based countermeasure. With a filter deployed in a wireless mix network, the detection rate can be maintained near the minimum value of *50*%.

# 6. CONCLUSIONS

In this dissertation, we study traffic analysis attacks and their countermeasures. Traffic analysis attacks are aimed at deriving critical information by analyzing the statistics of traffic flows. This kind of attack challenges the design of traditional systems where encryption is typically used as the main method for protecting security and privacy. However, it is obvious that encryption alone cannot protect many important characteristics of network traffic that may be mission critical and require protection. We focus on studying the threat from link-load analysis attacks and connectivity analysis attacks and relevant countermeasures to them.

Link padding is to counter traffic analysis attacks. We introduce statistical pattern recognition as a formal framework for analyzing the security of communication systems under traffic analysis attacks. We find that the commonly used CIT padding still can reveal the payload traffic rate if an adversary measures statistics such as sample entropy and sample variance of packet inter-arrival times. We discover that VIT padding can be very effective against these kinds of passive link load analysis attacks.

Using the framework of statistical pattern recognition, we also study the security of VIT padding under active link-load analysis attacks. To demonstrate the threat from such attacks, we consider ping probing attacks aimed at deriving user payload traffic rates. We found that by measuring statistics of the round trip time of ping packets injected into security gateways, the adversary can break the system, track the user payload traffic

changing pattern, and discover exactly the payload traffic rate that security gateways attempt to protect. This is true even if a powerful method such as VIT padding is used.

In addition, we study connectivity analysis attacks, which are aimed at discovering the flow connectivity between a pair of users. We find that a class of active connectivity analysis attacks, namely flow marking attacks, can be both effective and efficient in wireless networks. We propose filter-based approaches to counter these kinds of attacks. Our experimental data shows that our approach can help to protect flow connectivity information when the system is under connectivity analysis attacks.

In this dissertation, not only do we have experimental results about the above attacks and countermeasures in a laboratory environment, but we also theoretically analyze these attacks and countermeasures and perform extensive experiments on campus networks and the Internet to further validate our findings. We believe that our methodology will provide a solid foundation for studying the entire spectrum of traffic analysis attacks and their countermeasures.

# REFERENCES


[1] B. Caswell and M. Roesch, *Snort*, 2004, available from http://www.snort.org/.

[2] C. E. Shannon, "Communication Theory of Secrecy Systems," *Bell Sys. Tech. J.*, vol. 28, pp. 656–715, 1949.

[3] P. Baran, "On Distributed Communications: IX Security, Secrecy, and Tamper-free Considerations," Memo RM-3765-PR, Santa Monica, California: Rand Corp., August 1964.

[4] V. Voydoc and S. Kent, "Security Mechanisms in High-level Network Protocols," *ACM Computing Surveys*, vol. 15, no. 2, pp. 135–171, June 1983.

[5] R. E. Newman-Wolfe and B. R. Venkatraman, "High Level Prevention of Traffic Analysis," *Proceedings of the Seventh Annual Computer Security Applications Conference*, pp. 102–09, San Antonio, Texas, December 1991.

[6] R. E. Newman-Wolfe and B. R. Venkatraman, "Performance Analysis of a Method for High Level Prevention of Traffic Analysis," *Proceedings of the Eighth Annual Computer Security Applications Conference*, pp. 123 –130, San Antonio, Texas, November 1992.

[7] B. R. Venkatraman and R. E. Newman-Wolfe, "Performance Analysis of a Method for High Level Prevention of Traffic Analysis Using Measurements from a Campus Network," *Proceedings of the Tenth Annual Computer Security Applications Conference*, pp. 288 –297, Orlando, Florida, December 1994.

[8] B. Timmerman, "A Security Model for Dynamic Adaptive Traffic Masking," *Proceedings of New Security Paradigms Workshop*, pp. 107-116, Great Langdale, Cumbria, UK, September 1997.

[9] H. T. Kung, C-M Cheng, K. S. Tan, and S. Bradner, "*Design and Analysis of an IP-layer Anonymizing Infrastructure*," *Proceedings of the Third DARPA Information Survivability Conference and Exposition (DISCEX)*, pp. 62-77, Washington D.C., April 2003.

[10] Y. Guan, X. Fu, D. Xuan, P. U. Shenoy, R. Bettati, and W. Zhao, "Netcamo: Camouflaging Network Traffic for QoS-guaranteed Critical Applications," *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans, Special Issue on Information Assurance*, vol. 31, no. 4, pp. 253–265, July 2001.

[11] D. X. Song, D.Wagner, and X. Tian, "Timing Analysis of Keystrokes and Timing Attacks on SSH," *Proceedings of the 10th USENIX Security Symposium*, pp. 337-352, Washington, D.C., August 2001.

[12] D. Chaum, "Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms," *Communications of the ACM*, vol. 4, no. 2, pp 84-88, February 1981.

[13] P. F. Syverson, D. M. Goldschlag, and M. G. Reed, "Anonymous Connections and Onion Routing," *Proceedings of IEEE Symposium on Security and Privacy*, pp. 44–54, Oakland, California, May 1997.

[14] J. Raymond, "Traffic Analysis: Protocols, Attacks, Design Issues and Open Problems," *Proceedings of International Workshop on Designing Privacy Enhancing Technologies: Design Issues in Anonymity and Unobservability*, pp.10-29, Berkeley, California, January 2001.

[15] A. Back, U. Möller, and A. Stiglic, "Traffic Analysis Attacks and Trade-offs in Anonymity Providing Systems," *Proceedings of the 4th International Workshop on Information Hiding*, pp.245-257, Pittsburgh, Pennsylvania, April 2001.

[16] A. Back, I. Goldberg, and A. Shostack, *Freedom Systems 2.1 Security Issues and Analysis*, May 2001, available from http://www.freehaven.net/anonbib/cache/ freedom21-security.pdf.

[17] G. Danezis, R. Dingledine, and N. Mathewson, "Mixminion: Design of a Type III Anonymous Remailer Protocol," *Proceedings of IEEE Symposium on Security and Privacy*, pp. 2-15, Berkeley, California, May 2003.

[18] M. J. Freedman and R. Morris, "Tarzan: A Peer-to-peer Anonymizing Network Layer," *Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS)*, Washington, D.C., pp. 193-206, November 2002.

[19] G. Danezis, "The Traffic Analysis of Continuous-Time Mixes," *Proceedings of Privacy Enhancing Technologies*, pp. 35-50, Toronto, Canada, May 2004.

[20] J. Kong and X. Hong, "ANODR: Anonymous on Demand Routing with Untraceable Routes for Mobile Ad-hoc Networks," *Proceedings of the 4th ACM International Symposium on Mobile Ad Hoc Networking & Computing*, pp.291-302, Annapolis, Maryland, June 2003.

[21] Q. Sun, D. R. Simon, Y. Wang, W. Russell, V. N. Padmanabhan, and L. Qiu, "Statistical Identification of Encrypted Web Browsing Traffic," *Proceedings of IEEE Symposium on Security and Privacy*, pp. 19-30, Berkeley, California, 2002.

[22] A. Hintz, "Fingerprinting Websites Using Traffic Analysis," May 2002, available from http://guh.nu/projects/ta/safeweb/safeweb.html.

[23] A. Serjantov and P. Sewell, "Passive Attack Analysis for Connection-based Anonymity Systems," *Proceedings of the 8th European Symposium on Research in Computer Security*, pp. 116-131, Gjøvik, Norway, 2003.

[24] Y. Guan, X. Fu, R. Bettati, and W. Zhao, "An Optimal Strategy for Anonymous Communication Protocols," *Proceedings of the 24th IEEE International Conference on Distributed Computing Systems (ICDCS)*, pp. 257-270, Vienna, Austria, July 2002.

[25] Y. Zhu, X. Fu, B. Graham, R. Bettati, and W. Zhao, "On Flow Correlation Attacks and Countermeasures in Mix Networks," *Proceedings of Privacy Enhancing Technologies*, pp. 207-225, Toronto, Canada, May 2004.

[26] B. N. Levine, M. K. Reiter, C. Wang, and M. Wright, "Timing Attacks in Low-Latency Mix-based Systems," *Proceedings of Financial Cryptography (FC)*, pp. 251-265, Key West, Florida, February 2004.

[27] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed., New York: John Wiley & Sons, 2001.

[28] B. Silverman, *Density Estimation for Statistics and Data Analysis*, Monographs on Statistics and Applied Probability. London: Chapman & Hall, 1986.

[29] TimeSys, "Timesys Linux Docs," 2003, available from http://www.timesys.com/index.cfm?hdr=homeheader.cfm&bdy=homebdylibrary.cfm.

[30] S. Ghosh and R. Rajkumar, "Resource Management of the OS Network Subsystem," *Proceedings of the Fifth IEEE International Symposium on Object-Oriented Real-Time Distributed Computing*, pp. 271–279, Newport Beach, California, April 2002.

[31] G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed., Pacific Grove, California: Duxbury/Thomson Learning, 2002.

[32] R. Moddemeijer, "On Estimation of Entropy and Mutual Information of Continuous Distributions," *Signal Processing*, vol. 16, no. 3, pp.233–246, 1989.

[33] Agilent Technologies, "Agilent j6841a Network Analyzer Software," March 2002, available from http://onenetworks.comms.agilent.com/NetworkAnalyzer/J6841A.asp.

[34] Marconi Corporation, "ESR-5000 and ESR-6000 Enterprise Switch Routers," 2003, available from http://www.marconi.com/html/products/esr50006000.htm.

[35] Y. Guan, "A Study on Countermeasures Against Traffic Analysis Attacks," Ph.D. Dissertation, Texas A&M University, 2002

[36] M.Wright, M. Adler, B. N. Levine, and C. Shields, "An Analysis of the Degradation of Anonymous Protocols," *Proceedings of the Network and Distributed Security Symposium (NDSS)*, pp. pp. 20-31, San Diego, California, February 2002.

[37] M. Wright, M. Adler, B. N. Levine, and C. Shields, "Defending Anonymous Communication against Passive Logging Attacks," *Proceedings of the IEEE Symposium on Security and Privacy*, pp. 28-43, Berkeley, California, May 2003.

[38] M. Fomenkov, K. Keys, D. Moore, and K. Claffy, "Longitudinal Study of Internet Traffic in 1998-2003," *Proceedings of the Winter International Symposium on Information and Communication Technologies*, pp. 23-28, Cancun, Mexico, January 2004.

[39] K. Thompson, G.Miller, and R.Wilder, Wide-area Internet Traffic Patterns and Characteristics. *IEEE Network Magazine*, vol. 11, no. 6, pp. 10-23, November/ December 1997.

[40] C. Fraleigh, S. Moon, C. Diot, B. Lyles, and F. Tobagi, "Packet-level Traffic Measurements from a Tier-1 IP Backbone," Sprint ATL Technical Report, TR01-ATL-110101, Burlingame, California, November 2001.

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

[42] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The Second-generation Onion Router," *Proceedings of the 13th USENIX Security Symposium*, pp. 303-320, San Diego, California, August 2004.

[43] A. Serjantov, R. Dingledine, and P. Syverson, "From a Trickle to a Flood: Active Attacks on Several Mix Types," *Proceedings of Information Hiding Workshop*, pp. 36-52, Noordwijkerhout, Netherlands, October 2003.

[44] IEEE-SA Standards Board, *Part 11: Wireless LAN Media Access Control (MAC) and Physical Control Specifications (802.11)*. Piscataway, New Jersey: IEEE, Inc., 1999.

[45] Bluetooth Special Interest Group (SIG), *Specification of the Bluetooth System, Version 1.1*, 2001, available from https://www.bluetooth.org/spec/.

[46] C. E. Perkins and P. Bhagwat, "Highly Dynamic Destination-sequenced Distance-Vector Routing (DSDV) for Mobile Computers," *Proceedings of Special Interest Group on Data Communications (SIGCOMM)*, pp. 234-244, London, UK, September 1994.

[47] D. B. Johnson and D. A. Maltz, "Dynamic Source Routing in Ad-hoc Wireless Networks," *Mobile Computing*, vol. 353, pp. 153-181, 1996.

[48] C. Perkins and E. Royer, "Ad-hoc On-demand Distance Vector Routing," In *RFC 3561*, July 2003, available from http://www.faqs.org/rfcs/rfc3561.html.

[49] R. Kapoor and M. Gerla, "A Zone Routing Protocol for Bluetooth Scatternets," *Proceedings of IEEE Wireless Communications and Networking Conference*, vol. 3, pp. 1459-1464, New Orleans, Louisiana, March 2003.

[50] K. Krizman, T. Biedka, and T. S. Rappaport, "Position Location: Fundamentals, Implementation Strategies, and Sources of Error," *Proceedings of IEEE Vehicular Technology Conference*, pp. 919-923, Phoenix, Arizona, May 1997.

[51] S. J. Lee, W. Su, and M. Gerla, "Wireless Ad-hoc Routing with Mobility Prediction," *Mobile Networks and Applications*, vol. 6, no. 4, pp. 351–360, 2000.

[52] G. Liu and G. J. Maguire, "A Class of Mobile Motion Prediction Algorithms for Wireless Mobile Computing and Communication," *Mobile Networks and Applications - Special Issue: Routing in Mobile Communications Networks*, vol. 1, no. 2, pp. 113–121, October 1996.

[53] A. Kuzmanovic and E. W. Knightly, "Low-rate TCP Targeted Denial of Service Attacks," *Proceedings of ACM Special Interest Group on Data Communications* (*SIGCOMM*), pp. 75-86, Oakland, California, August 2003.

[54] R. Dixon, *Spread Spectrum Systems*, 2nd Edition. New York: John Wiley & Sons, Inc., 1984.

[55] A. Laurie, "Serious Flaws in Bluetooth Security Lead to Disclosure of Personal Data," October 2003, available from http://www.thebunker.net/release-bluestumbler.htm.

[56] J. Padhye, V. Firoiu, D. Towsley, and J. Krusoe, "Modeling TCP Throughput: A Simple Model and Its Empirical Validation," *Proceedings of ACM Special Interest Group on Data Communications (SIGCOMM)*, pp. 303-314, Vancouver, B.C., Canada, August 1998.

[57] A. V. Oppenheim, A. S.Willsky, and S. H. Nawab, *Signals and Systems*. 2nd ed., Upper Saddle River, New Jersey: Prentice-Hall, 1997.

[58] M. Carson and D. Santay, “Nistnet - a Linux-based Network Emulation Tool,” *Computer Communication Review*, vol. 33, no. 3, pp. 111–126, July 2003.

[59] Netfilter. 2003, available from http://netfilter.samba.org/.

[60] tcpdump, 2004, available from http://www.tcpdump.org/.

[61] J. Tourrilhes, “Wireless Tools for Linux,” 2004, available from http://www.hpl.hp .com /personal/Jean_Tourrilhes/Linux/Tools.html.

[62] W.Willinger, M. S. Taqqu, R. Sherman, and D. V.Wilson, “Self-similarity through High-variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level,” *IEEE/ACM Transactions on Networking*, vol. 5, no. 1, pp. 71–86, 1997.

[63] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1975.

# VITA


| | |
|---|---|
| Name: | Xinwen Fu |
| Address: | Dakota State University<br>College of Business and Information Systems<br>East Hall<br>820 North Washington Avenue<br>Madison, SD 57042 |
| Email Address: | xinwenfu@gmail.com |


Education:

Ph.D. in Computer Engineering, December 2005, Texas A&M University, U.S.A.

M.S. in Electrical Engineering, December 1998, Graduate School of University of Science and Technology of China, China

B.S. in Electrical Engineering, July 1995, Xi'an Jiaotong University, China

Research Interests: Network Security and Privacy, Information Assurance, Computer Networking, Distributed Systems, Mobile and Wireless Networks

Major Honors and Awards:

- 1st place, Oral Competition in the Student Research Week, Texas A&M University, 2005
- Graduate Student Research Excellence Award, Department of Computer Science, Texas A&M University, 2004
- 2nd Place Graduate Winner, ACM International Student Research Contest, 2002
- Elite Scholarship, Chinese Academy of Sciences, 1996
- Annual Student Scholarship, Xi'an Jiaotong University, 1991 - 1995