Capacity Estimation and Auditability of Network Covert Channels

Balaji R. Venkatraman

Hewlett-Packard Company
Corporate Network Services
Palo Alto, CA - 94043
balaji@corp.hp.com

R. E. Newman-Wolfe

Computer & Information Sciences Dept.
University Of Florida,
Gainesville, Florida - 32611.
nemo@chameleon.cis.ufl.edu

Abstract

Classical covert channel analysis has focused on channels available on a single computer: timing channels and storage channels. In this paper, we characterize network covert channels. Potential network covert channels are exploited by modulating transmission characteristics. We distinguish between spatial covert channels, caused due to variation in the relative volume of communication between nodes in the network, and temporal covert channels caused due to variation in transmission characteristics over time, extending the work of Girling[5]. A model for obtaining spatially neutral transmission schedule is given in [16, 17]. Temporally neutral transmissions are characterized and scheduling policies to generate temporally neutral transmission schedules are given in [23]. In this paper, we estimate the covert channel capacity using an adaptive scheduling policy, modeling the system as a mode secure system[1]. Based on our measurements on University of Florida campus-wide backbone network (UFNET), we discuss the auditability of network covert channels and suggest some handling policies to reduce the capacity of these covert channels to TCSEC acceptable levels.

1 Introduction

Covert channels are communication channels that allow a process to transfer information in a manner that violates the system's security policy[2]. Traditionally, covert channels in computer systems are classified into two categories, storage and timing covert channels. Covert storage channels involve the direct or indirect writing of storage by one process (the sender of covert message) and the direct or indirect reading of the storage location by another process (the recipient of the covert message). Covert timing channels are exploited when the sender process modulates the use of its own resources in a manner that affects the real time response of the receiver process[2]. In both cases, for a covert channel to exist, the sender and the receiver process must share some common computational resource, usage of which can be modulated in a predetermined manner. Network covert channels, on the other hand, are exploited by the manipulation of communication resources or transmission characteristics. Processes that use the covert network channel do not manipulate the resources generally local to the machine on which the computation occurs.

A related problem is traffic analysis or the disclosure of information by inference caused by the passive monitoring of network traffic by an eavesdropper. By observing the volume of communication (number of packets) either between specific source-destination pairs or the overall volume of communication between nodes in the network, the eavesdropper can gain some insight into the behavior of the network users. Network covert channels exist due to the active participation of the sender (and perhaps the receiver processes) to modulate transmission characteristics. Different transmission characteristics can be modulated to exploit covert channels of varying capacity. We distinguish between two types of network covert channels: *spatial covert channels* are exploited by modulating the volume of communication between nodes and *temporal covert channels* are exploited by modulating the order, frequency, or duration of transmission. These are in addition to the modulation of the extrinsic characteristics of packets like the packet size, the application generating the packet, etc. The potential capacity of several simple network covert channels is several orders of magnitude more than the capacity of computational covert channels and therefore deserves more attention.

In the following section, we discuss relevant work in the area of traffic analysis and covert channel analysis. In section 2.1, we discuss briefly a model for spatial neutrality, followed by a discussion of the criteria for temporal neutrality in section 2.2 and the scheduling policies used to generate temporally neutral transmission schedules. In section 3, we discuss techniques to reduce the capacity of covert channels and in section 3.1, we discuss techniques to estimate covert channel capacity based on Millen's[12] application of Shannon's information theory[20] and in section 3.1.1, we discuss Browne's mode secure systems[1] use a mode secure system to estimate network covert channels under our model. In section 4 we discuss the auditability of noiseless covert channels in section 4.1

followed by a discussion of various handling policies to reduce the covert channel capacity in section 5. In section 6, we give our conclusions and suggestions for future work.

2 Related Work

The primary goal of a communication subsystem or network is to provide remote access to users and facilitate resource sharing, which increases both its vulnerability to attacks by wiretappers and intruders as well as the potential for covert channels. The issues involved in the security of communication systems differ from those of computer systems because network security depends on factors such as the network architecture, the topology, connectivity, the communication protocols used, the security of each individual node connected to the network, and the security of each link interconnecting the nodes. Each of the above must be both individually secure and securely composable for the communication subsystem to be deemed secure.

Potential security violations in networks include unauthorized release or modification of information, and unauthorized denial of use of resources[26]. While active attacks cause information modification or denial of resources, in a passive attack, the intruder simply releases the contents of a message or mounts a traffic analysis attack to infer user behavior or exploit certain covert channels.

One of the goals of communication security is the prevention of traffic analysis, by which an intruder may deduce important information from the mere presence of message traffic in a network [26, 19]. This information may yield clues to the activity or intentions of unsuspecting network users, or may provide a covert channel for communication between an intruder and an accomplice within the system. Traffic analysis countermeasures must mask the amount and nature of traffic between origin-destination pairs within the network. The precision with which an intruder can analyze these traffic patterns determines the amount of information that he can infer about the network user[29].

Most previous work has used the ISO's (International Standards Organization) open system interconnection (OSI) network architecture for describing threats and countermeasures [8, 26, 27, 19, 28]; we will do the same. The standard approach to preventing unauthorized release of information in a network is encryption. A major issue is the OSI level(s) at which encryption is performed; we refer the interested reader to Padlipsky et al.[18].

The two basic approaches to communication security are *link-oriented* security measures, which provide security by protecting message traffic independently on each communication link, and *end-to-end* security measures, which provide protection for each message from its source to destination node[26]. Link-oriented security demands that all intermediate network nodes participate in the network's security, and is vulnerable to compromise of a single node. End-to-end measures reveal at least destination addresses, but allow a subset of nodes to implement network security without requiring the entire network to do so. With the increasing use of public networks, this is an important consideration.

2.1 High Level Prevention of Traffic Analysis: Spatial Neutrality

To achieve protection beyond that offered by encryption at the transport layer at a cost less than that exacted by sending dummy packets alone, we propose an approach in which the transport layer entities may send a message to a destination other than the true destination of the message[16]. Transport entities must agree to forward these rerouted messages to their true destination when they are received and initially decoded. In this way, we manipulate the initial traffic matrix so that each host sends every other host in the network the same volume of messages, that is, it presents the intruder with a *neutral traffic matrix*. Regardless of the original traffic pattern, the intruder observing the manipulated traffic pattern will see only even communication levels. Thus, the intruder cannot derive useful information regarding the original traffic patterns. We call this the *spatial neutrality criterion*.

The forms of traffic manipulation we consider in this model are padding and a restricted form of rerouting. Padding consists of introducing dummy packets, which, since they are encrypted, are indistinguishable from the actual communication[5]. One way to achieve spatial neutrality is to pad the traffic matrix so that the volume of each internode communication is the same. Since padding alone is expensive, we use rerouting to smooth the traffic matrix first and then pad to get a neutral traffic matrix. Rerouting sends a packet first to an intermediate node, which upon processing is directed to forward the packet to its true destination. The nodes participating in this security model agree to honor this restricted form of rerouting so that in the resulting neutral traffic matrix, the original traffic is delivered to its intended destination by the second hop. Encryption may be used to maintain secrecy so that an observer cannot distinguish between a rerouted packet from a non-rerouted one and to maintain end-to-end confidentiality of the packet contents. The cost of achieving spatially neutral traffic matrix is discussed in [17] and [25].

The basic model guarantees spatial neutrality by eliminating the variation in the relative volume of traffic between each pair of nodes. However, we are also concerned with the temporal variation in traffic and the possible introduction of covert channels due to the variation in the transmission characteristics over time. Note that although the total volume of traffic communicated between any pair of nodes in the network is

the same to satisfy the spatial neutrality criterion, the source could transmit the packets in a burst or could spread out its transmissions over a period of time. The model imposes no restriction on the transmission schedule and therefore a knowledgeable user might be able to communicate with his accomplice by timing the transmissions, thus introducing covert channels.

2.2 Temporal Neutrality of Transmission Schedules

To address this concern, in addition to requiring that the traffic matrix be spatially neutral, we require the transmission schedule be temporally neutral to eliminate potential covert channels. In [23], we propose two transmission scheduling policies that will satisfy our primary goal of prevention of traffic analysis and the prevention of covert channels due to temporal variation in packet transmission schedule. The static scheduling policy generates spatially and temporally neutral transmission schedules but is unresponsive to changes in system load. The adaptive scheduling policy can adapt to long term fluctuations in system load at the expense of allowing certain covert channels.

We can describe a transmission schedule by the following five characteristics: the volume (V), frequency (F), order (O), nature (N), and length of communication (L). We denote these characteristics by the tuple $< V, F, O, N, L >$. Depending on the information that can be encoded in any of these transmission characteristics, a covert channel may exist.

Definition: *A temporally neutral transmission schedule is one in which none of the members of the tuple* $< V, F, O, N, L >$ *can be used to encode any information and communicate surreptitiously via a covert channel.*

2.3 Scheduling Policies for Temporal Neutrality

The transmission schedule that satisfies each of the restrictions in $< V, F, O, N, L >$ is temporally neutral. In other words, the intruder cannot gain any useful information regarding the traffic matrix, source and/or destination user identity, etc., just by observing the flow of packets on the network. In effect, a user in collusion with an accomplice should not be able to use the volume, frequency, order, nature of communication, and the duration (length) of transmission in the network to exchange information surreptitiously.

For analysis purposes, we will use a slotted time system in which a "slot" is the basic time unit during which a given node may send or receive at most one packet on a slotted medium. We assume that at most one node can transmit per slot, so $n(n-1)$ slots are needed for all pairs to communicate. This model may be modified for point-to-point networks. A period is a set of successive slots during which one phase of the transmission schedule is carried out. In our model, a period consists of $n(n-1)$ active slots and m idle slots. In the static scheduling policy, m is a constant; in the adaptive scheduling policy, m may vary over time. A set of successive periods during which the number of idle slots, m, does not vary is a cycle.

2.3.1 The Adaptive Scheduling Policy

In this section, we discuss the adaptive scheduling policy in terms of the $< V, F, O, N, L >$ criteria. For a more thorough exposition, refer [23, 24].

For our discussion on temporal neutrality, we assume spatial neutrality. The volume of communication between each node has to be the same to satisfy the neutrality criterion. The frequency and order of transmission is predetermined and is fixed for the entire duration of a cycle. This satisfies the V, F, and O criteria of $< V, F, O, N, L >$. If we build the communication protocol on a NTCB and fix extrinsic packet characteristics like the packet size and encryption algorithm, we satisfy the N restriction of $< V, F, O, N, L >$. Since there are exactly $n(n-1)$ active slots in the period, one for each pair of nodes in the system and m idle slots, the length of the period is fixed at $n(n-1)+m$ for the duration of the cycle. This is the simple static scheduling policy and is temporally neutral[23].

The only restriction that we cannot satisfy in the adaptive scheduling policy is the L restriction of the tuple $< V, F, O, N, L >$. This is because the nodes may change the number of idle slots, changing L, the duration (length) of transmission. Though this could potentially introduce a low bandwidth, noisy covert channel, we feel that this tradeoff may be acceptable in order to have an adaptive scheduling policy.

When a node or a group of nodes see a need to change the number of idle slots to accommodate additional traffic, they initiate the negotiation process with the NTCB. As shown in figure 1, after a sustained increase in the load, the nodes negotiate and decide to decrease the number of idle slots per period. The number of active slots in the period remain the same, but the total period length decreases (in this case by one slot). Note that the length of the period L and therefore the transmission characteristic has changed. It is this possibility that prevents us from guaranteeing the L restriction in $< V, F, O, N, L >$ and leaves open the possibility of a covert channel.

It should be noted that no single node can affect the active and idle slot times significantly without reaching a consensus with other nodes in the network. Therefore the potential of a single node to change the transmission schedule is very limited. For example, an user may try to change the load on a particular node in an attempt to change the transmission characteristics, which could be observed by the accomplice on the network, thus creating a covert channel. However, the negotiation protocol considers the cumulative ef-

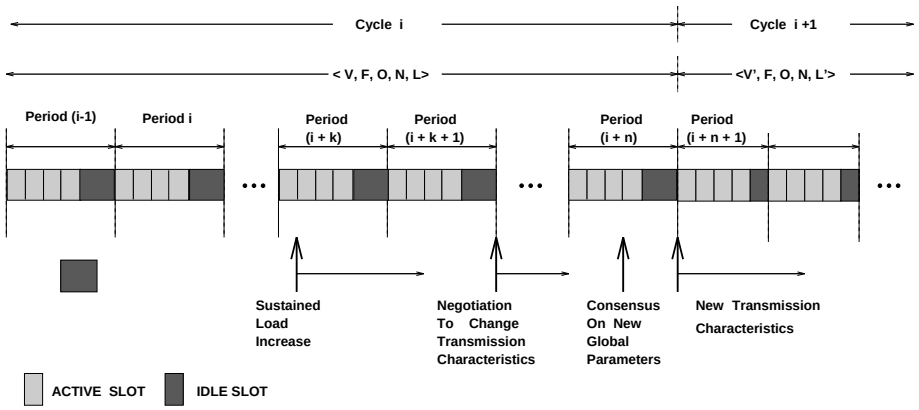

Figure 1: The Adaptive Scheduling Policy

fects of several individual node's (user's) actions, so the effects of any single node on the transmission characteristics is relatively minor.

2.4 Covert Channels

The National Computer Security Center's recently published *Guide to Understanding Covert Channel Analysis of Trusted Systems*[15] is a comprehensive summary of techniques used in covert channel analysis and related areas. This document discusses storage and timing channels that arise due to the sharing of "computational resource" among subjects at different security levels.

We refer the interested reader to Huskamp[7] for a detailed discussion on covert timing channels, and to Tsai[22] for a discussion on covert storage channels and to NCSC[15] for a summary of techniques to identify covert channels and estimate their capacity, and a discussion of the various covert channel handling mechanisms. The covert channel guideline[15] also presents TCSEC requirements for covert channel analysis and includes additional recommendations corresponding to B2-A1 evaluation classes.

In Girling[5], various covert channels in LANs are identified and their behavior discussed in the context of the ISO OSI framework. Covert channels due to address field encoding, length of data block and time between successive transmissions were discussed and solutions proposed. Experiments conducted to estimate the covert channel capacity concluded that high capacity covert channels could exist in high bandwidth networks. The author concludes that physical security may be more effective against covert channels than the use of encryption or complex mechanisms to reduce their capacity.

We partially disagree with the Girling's conclusions; state–of–the–art computer and communication systems require more protection than that offered by physical security. Further, for wide-area networks, such physical security may not be feasible. With rapid advances in the understanding of such complex systems and the easy availability of sophisticated technology, we believe that the scope and nature of attacks mounted on a secure system has undergone a drastic change for the worse. Security loopholes that were too difficult or intractable to exploit until a few years ago are now easy prey to computer scientists and hackers alike. Effective countermeasures are required to ensure operation security in real-time without an undue penalty on system performance. Our model for prevention of traffic analysis is a step in that direction.

In spite of the directions in Voydock[27], Girling[5], TCSEC[2], TNI[14] and the covert channel guideline[15] there exists no accepted definition of network covert channels. Therefore, we propose the following definition of network covert channels.

Definition: *Covert channels in communication subsystems that implement a valid interpretation of a consistent security policy, are based on the observation of the extrinsic characteristics of the communication without necessarily having access to the information contained within messages (due to encryption) or the necessity to modulate internal states or variables.*

Using the above definition, we can readily identify the covert channels discussed in Girling[5] and in Venkatraman[23].

3 Covert Channel Handling

The following techniques, discussed in [15, 21, 22, 24], can be used to reduce covert channel capacity.

1. **Elimination of Covert Channels:**
   Covert channels can be eliminated either by forbidding any resource sharing between subjects at different levels in a multi-level secure system or by eliminating features and mechanisms in the operation of the system that could potentially cause covert channels.

This approach eliminates covert channels and guarantees a system free of covert channels, but, in general, is very restrictive and expensive to enforce. Any restriction on resource sharing will most likely be at the expense of system performance and resource utilization. Also redesigning user interface to be less user friendly and cryptic is not a very appealing solution.

2. **Capacity Limitation:**
   The capacity of known covert channels can be reduced by the introduction of noise or delays into channel operation. The objective of such methods is to reduce the maximum channel capacity to a level where the cost of further reduction or elimination of these covert channels is prohibitive.

   Examples of capacity limitation techniques include Fuzzy Time[6], Pump[9], etc. Fuzzy Time reduces the capacity of covert timing channels by making all clocks in the system noisy. Fuzzy time was successfully implemented in the VAX security kernel to reduce the capacity of the covert timing channel without severe performance degradation. In the reliable Pump, a buffer is inserted between Low and High to respond with ACKs at probabilistic times, thereby introducing noise into any potential timing channel.

   However, in general, problems with these approaches lie in being able to quantify noise correctly and the difficulty in being able to introduce delays and subsequently tune the appropriate TCB primitives and other scheduling operations. While the introduction of delay and noise tends to degrade system performance and necessitates restating the Quality of Service (QOS) guarantees, the penalty imposed on system performance may be justified given the objective of containing covert channel capacity.

3. **Auditing the Use of Covert Channel:**
   In this approach, the existence of covert channels is known and may be exploited by the users of a secure system. However, auditing the usage of such channels acts as a deterrence to potential users of the channel. The difficulty in this approach is in being able to unambiguously distinguish between innocuous user activity and actual covert channel usage. A related problem is to be able to detect every such usage of the covert channel.

   Audit trails can be used to determine the use of covert channels and its capacity. However, determining which specific events need to be monitored and recorded by audit mechanisms to ensure that all covert channel usage is detected is a nontrivial task.

We adopt mode security (discussed in section 3.1.1), a different approach to capacity reduction in conjunction with several audit policies (discussed in section 4) to contain the capacity of the network covert channel.

3.1 Covert Channel Capacity Estimation Methods

In this section, we discuss formal and informal techniques to estimate the covert channel capacity. A capacity estimation method based on Shannon's information theory is presented in Millen[12]. In this method, the assumptions are that the covert channels are noiseless; that other than the sender and receiver, there are no unconfined processes in the system during channel operation; and the sender-receiver synchronization takes a negligible amount of time [12, 15]. With these assumptions, one can model most covert channels as finite state machines and compute the maximum attainable capacity.

The capacity, C, of a noiseless discrete channel with symbols of different length is given by Shannon[20] and modified by Moskowitz[13] as

$$C = \lim_{t \to \infty} sup \, \frac{\log_2 \, N(t)}{t}$$

where N(t) is the number of possible symbol sequences of time t.

In the following subsection, we model our system as a mode secure system and estimate the capacity of network covert channel.

3.1.1 Capacity Analysis Using Mode Based Security

Mode security, proposed by Browne[1], seeks to contain the capacity of covert channel by partitioning the resources dynamically "for a limited period of time," based on factors such as resource request, utilization, etc. During the time interval when the resource partitioning remains fixed, the machine is said to be in a "mode". Each mode of the machine behaves as a lattice separable system which implies that there is no interaction between different security levels in the system. Therefore when the machine is operating in a mode, there can be no covert channels. Based on resource requirements, the system periodically re-partitions the resources among security levels. This re-partitioning of resources may lead to covert channels[1].

In mode secure systems, the capacity of a covert channel depends on two factors: first, the frequency with which the system changes its mode and second, the number of unique modes to which the system can transition during a mode change. We can reduce the covert channel bandwidth by either limiting the frequency of mode change or limiting the number of legal transitions at each mode or both.

In our model, a cycle(see figure 1) is defined as the time interval between changes in transmission characteristics and can be set either to a fixed amount

of time or a fixed number of periods. If both the time and the number of periods are variable, in his attempts to maximize the channel capacity, the sender will try to force the minimum of the two, thereby maximizing the frequency of mode change.

If the cycle consists of a fixed number of periods, then we can estimate the channel capacity using the information theory based technique, the difference being that the symbols take longer time to be transmitted. If t_i was the time required to transmit the symbol assuming one period per cycle, we will now require Pt_i time to transmit the same symbol, where P is the number of periods per cycle. This implies that the channel capacity will be reduced by a factor of P. Note that the symbol transmission time is independent of the previous symbol transmitted, simplifying the information theoretic analysis.

Let us consider the case when the cycle length is fixed and does not depend on the number of periods. Let T_c be the length of a cycle and let $T_c \gg T_p$, where T_p is the length of a period. $M-1$ is the maximum number of idle slots allowable in any period and $m \in [0 \cdots M-1]$. Time is normalized to slot times. Let t_m denote the time required to transmit a symbol m. Then $t_m \approx T_c$, which implies $t_m \approx t_{m'}$, for all symbols m, m'. The channel capacity C is simply

$$C = \lim_{t \to \infty} sup \frac{\log_2 M^n}{nT_c} = \frac{\log_2(M)}{T_c}.$$

For a given number of idle slots, as the number of active slots increase, the channel capacity decreases. This is to be expected as the fraction of transmission capacity that is potentially available for covert transmission is reduced. Note that for a given number of active slots $d = n(n-1)$ as the maximum number of idle slots in a period is increased, the channel capacity tends to approach the limiting channel capacity.

4 Auditable Covert Channels

An effective method of handling known covert channels is to deter its potential users. The existence of the covert channel is known to users, who may attempt to exploit the channel but the deterrence mechanism discourages such channel use. Covert channel auditing is the main deterrence mechanism and is effective only when the covert channel use can be detected unambiguously, i.e., discovery of covert channel use must be certain. Covert channel auditing must not be circumventable, and false detection of covert channel use must be avoided[21].

Our measurements on ECSNET (Engineering Consulting Services Network, a subnet on UFNET) showed that on a typical day, 19,888,635 packets were exchanged over a 15 hour period, the mean packet size being 291 bytes. Since there are 25 nodes in the network, the mean number of packets transmitted per node per minute is 884 packets. We also observed the maximum load on the network to be 35,050 packets per minute or 1,402 packets per node per minute and the minimum load to be 2,680 packets per minute or 107 packets per node per minute. We will use these traffic measurements to determine the covert channel capacity and propose audit mechanisms.

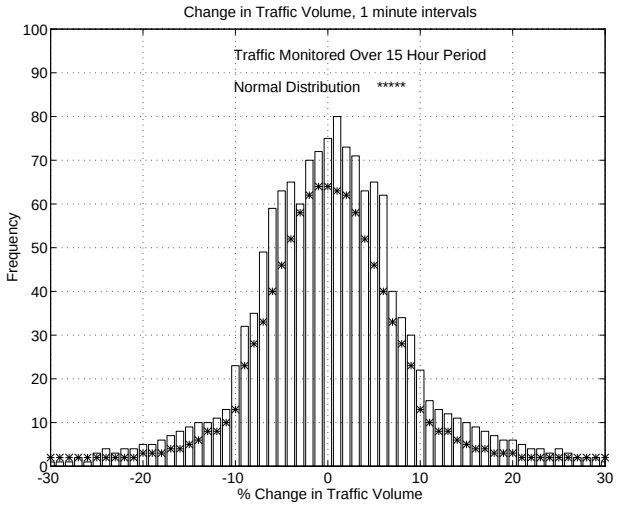

Figure 2: Percentage Change In Volume: ECSNET Traffic

Figure 2 shows the distribution of the percentage change in traffic volume over one minute intervals. We accept approximately 95% of the variations in traffic characteristics to be "normal" and exclude them from scrutiny. Two standard deviations for the distribution shown in figure 2 is a 23.95 percentage change in traffic volume. Therefore we set the threshold for auditability, θ, to be 24%, i.e., any variation in the traffic volume that is at least 24% of the current volume is audited.

The justification for setting the audit threshold at 24% is that the "normal" sustained burst volume in the network was observed to be approximately 15% for TCP traffic and 22% for UDP traffic under average load conditions (see section 4.1.1 and refer [23]). If the audit threshold is set lower that 24%, then the channel becomes extremely noisy.
Another reason for choosing the audit threshold is that based on the above measurements, a 5% audit factor will generate 45 events over a 15 hour period for the network. With 25 nodes in the network and assuming 100 bytes per audit record, the audit log can grow as large as 100k bytes just for ECSNET subnetwork. If measurements are done at a smaller granularity, or if the system load varies considerably over the measurement duration, then the audit logs can grow much larger. In such cases, even a 5% audit factor might be too high from a practical standpoint.

For our model, we must estimate traffic in terms of slot times based on our traffic observations and per-

formance evaluation of the spatial neutrality model. Under the given environment and load conditions, our approach has an overhead of a factor of four to achieve spatial and temporal neutrality[17, 25]. Hence unaudited channel capacities will be based on loads four times the observed load. ECSNET is a 10 Mbps LAN; with average packet size being 291 bytes, we derive the slot transmission time as 0.23 msec. With sufficient guard band and allowance for other overhead included, the effective slot time is 0.5 msec. Since the number of nodes in the network, n, is 25, our model prescribes a minimum of $n(n-1) = 600$ active slots in a period. At peak load conditions, the period consists entirely of active slots with a period length of 0.3 seconds or 200 periods per minute. At this rate, the maximum sustainable packet transmission rate is 4800 packets per node per minute.

In section 4.1, we estimate the capacity of auditable covert channels assuming that the channel is noiseless and suggest simple but effective handling techniques. Note that this is an informal estimate of channel capacity in which we assume fixed cycle length. A formal analysis of channel capacity will require an information theory based approach. Such an analysis is very involved even for simple cases and we will not pursue it in this paper and refer the interested reader to [24].

4.1 Auditability of Noiseless Covert Channels

In this section, we discuss the auditing of covert channels for the average load case. We assume that the channel is noiseless, i.e., any change to the traffic load is caused only by the sender and that the sender has the maximum effect over changing the traffic volume over successive periods. Other nodes in the network contribute a constant volume of traffic. To derive the maximum possible covert channel capacity, the sender is assumed to operate at its maximum transmission capacity.

We also assume that there are 25 nodes in the network, the slot time is 0.5 msec and the cycle time, T_c, is one minute. In order to compute the channel capacity for a given traffic load, we determine the maximum and minimum number of idle slots in a period for a -24% and a $+24\%$ change in the traffic volume respectively. For each audited channel, we also derive the channel capacity with and without handling. The covert channel handling methods reduce channel capacities by reducing the number of states to which the system can transition. This reduction in the number of states is achieved by increasing the granularity of change in the number of idle slots in a period. We discuss cases where the granularity is fixed at ± 10, ± 50, ± 100 and when the granularity is a fraction of the total number of slots in the period. This gives us the range of the number of idle slots within a period and knowing the handling policy, we can determine the number of states that are possible. Once we know the number of states, we can easily determine the covert channel capacity.

4.1.1 Auditing Noiseless Covert Channels for Average Load

The number of actual packets transmitted on average is 884 packets per node per minute; using our approach the number of packets exchanged is $884 \times 4 = 3536$ packets per node per minute. Since each node exchanges one packet with every other node in the network each period, we have $3536/24 = 147.3$ periods per minute or 0.407 seconds per period. Assuming the slot time to be 0.0005 seconds per slot, we have $0.407/0.0005 = 814$ slots per period. Since the number of active slots per period is $n(n-1) = 600$, the number of idle slots per period is 214 for this load level. We observed sustained burst lengths of approx. 770 packets for UDP traffic and 530 packets TCP traffic, giving 22% and 15% respectively of the average load on the network.

Since we accept any variation less than $\pm 24\%$ as normal, we now estimate the capacity for a variation of $\pm 24\%$. A 24% change in traffic volume yields 3536 ± 849 packets per node per minute. Following the computations shown above, $3536+849 = 4385$ packets per node per minute. This yields $4385/24 = 182.71$ periods per minute or 0.328 seconds per period for $0.3284/0.0005 = 657$ slots per period with 57 idle slots. Similarly, $3536 - 849 = 2687$ packets per node per minute. This yields $2687/24 = 111.958$ periods per minute or 0.5359 seconds per period for $0.5359/0.0005 = 1072$ slots per period with 472 idle slots.

Therefore the range of idle slots is $472 - 57 + 1 = 416$ states. This can encode $\log_2 416 = 9$ bits. Therefore the capacity of this channel is $9/60 = 0.15$ bps.

4.1.2 Handling Policy

1. If we assume that the granularity of change in the number of idle slots, $\delta = \pm 100$ from the previous period, then there are 5 states, needing 3 bits to encode the states. This yields a channel capacity of $3/60 = 0.05$ bps.

2. If we assume that the granularity of change in the number of idle slots, $\delta = \pm 50$ from the previous period, then there are 9 states, needing 4 bits to encode the states. This yields a channel capacity of $4/60 = 0.06$ bps.

3. If we assume that the granularity of change in the number of idle slots, $\delta = \pm 10$ from the previous period, then there are 42 states, needing 6 bits to encode the states. This yields a channel capacity of $6/60 = 0.1$ bps.

Table 1: Channel Capacities for a Noiseless Channel

	Minimum Load	Average Load	Maximum Load
Basic Load			
load, pkt/node/min	428	3536	3648
# slots	6728	814	790
# idle slots	6128	214	190
Unaudited Channels			
-24% load	325	2687	2772
$+24\%$ load	531	4385	4524
# idle slots for -24% load	8260	472	440
# idle slots for $+24\%$ load	4822	57	40
Range of idle slots	3439	416	401
Channel Capacity, bps	0.20	0.15	0.15
Channel Handling			
idle ± 100, Cap. bps	0.1	0.05	0.05
idle ± 50, Cap. bps	0.12	0.06	0.07
idle ± 10, Cap. bps	0.15	0.1	0.1
$\delta = 0.1 \times i$ *slots*	672	81	79
# states in $\pm 24\%$ load	7	7	7
$\delta = 0.1 \times \#$ *slots*, Cap. bps	0.05	0.05	0.05

4. If we assume that the granularity of change in the number of idle slots, δ, is a fraction, α, of the basic number of slots in the period, then we have $\delta = \alpha \times$ *slots*.

   For example, if $\alpha = 0.1$, then $\delta = 0.1 \times 814 = 81$. To determine the total number of states, we compute the number of states for $\pm 24\%$ variation in traffic. When the variation in traffic is $+24\%$, the number of idle slots is 57. Therefore the number of states in the positive direction is $\lceil \frac{214-57}{81} \rceil = 2$. Similarly, when the variation in traffic is -24%, the number of idle slots is 472. Therefore the number of states in the negative direction is $\lceil \frac{472-214}{81} \rceil = 4$. Therefore the total number of states is 2+4+1=7. This can be encoded by $\log_2 7 = 3$ bits, giving a channel capacity of 0.05 bps.

4.2 Discussion of Noiseless Channel Capacity

Table 1 shows the covert channel capacity for a noiseless channel with and without handling for minimum, average and maximum load cases. We see that the period length is inversely proportional to the network load. As the load increases, the number of idle slots per period reduces thereby reducing period length. A shorter period yields a greater fraction of active slots and therefore greater utilization of transmission capacity. Since the covert channel capacity depends on the maximum number of idle slots per period, it is clear that with increasing load, i.e., with increasing effective utilization, the covert channel capacity reduces. A detailed discussion of the relation between the load and channel capacity is given in the next section. The handling policies reduce the number of states to which the system can transition by varying the granularity of change in number of idle slots. With increasing granularity, the number of possible states decreases, reducing covert channel capacity. We refer the interested reader to Venkatraman[24] for a similar analysis of noisy channels.

5 Effect of Handling Policies on Channel Capacity

Figure 3 shows the channel capacities for different auditability thresholds, θ. The top three curves represent the channel capacities for a noiseless channel without handling for minimum load, average and effective maximum load, and the actual maximum load as indicated. The channel capacities for the average and the effective maximum load are almost the same because the load in both cases is almost the same (see table 1). From the figure, we can see that as the auditability threshold increases, the variability in the system load that is accepted as "normal" increases leading to higher covert channel capacities. Channel capacity also depends on the total load on the system. In a lightly loaded system, each period contains a large number of idle slots and therefore the potential

covert channel has high capacity. On the other hand, under maximum load conditions, there are very few idle slots, if any, in each period. Therefore the channel capacity is lower. The $2\sigma, 3\sigma$ and 4σ values of covert channel capacity are marked; as expected, by accepting a larger variation in the traffic change as normal, we allow a potential covert channel of larger capacity.

The lower three curves represent the channel capacities for a noiseless channel with handling for minimum, average and maximum loads. Using the proportional handling policy where the granularity of change in the number of idle slots, δ, is a fraction, α, of the basic number of slots in the period, then we have $\delta = \alpha \times slots$. In this case, we choose $\alpha = 0.1$. We see that this handling technique reduces covert channel capacity by more than 50 percent compared to the corresponding channel with no handling. Note also that the channel capacity after handling in the average and minimum load conditions is almost the same. This is due to the quantization effects caused by the granularity of change in the number of idle slots.

Figure 4 shows the effect of different handling policies at different load conditions on the covert channel capacity for a noiseless channel. From the figure, we can see that as the load increases, the covert channel capacity reduces. We also note that as the granularity of change in the number of idle slots increases, the channel capacity decreases. The solid curve represents the case when no handling mechanism is used. We see that proportional handling is more effective than other simpler handling policies. The covert channel capacity when $\alpha = 0.2$ is constant at 0.03 bps and does not depend on the system load.

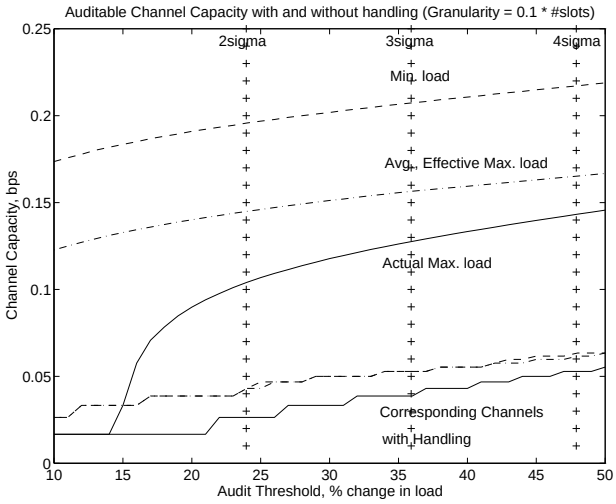

Figure 3: Noiseless Covert Channel for Different θ

5.1 Factors Affecting Channel Capacity and Auditability

The factors affecting the capacity and auditability of covert channels are the system load, the cycle length, the granularity of change in the number of idle slots and the auditability threshold. We will briefly discuss the effects of each of these factors below.

1. System load

As the system load increases, the number of packets exchanged increases. This has the effect of reducing the number of idle slots per period. The number of active slots per period is $n(n-1)$ to guarantee spatial neutrality. A lower bound on the number of slots per period and consequently on the number of periods per minute or the period length is imposed when the period consists of only active slots and zero idle slots. In this case, the capacity of the system is completely assigned and any excess load is backlogged for future transmission.

As seen in Figure 4, as the system load increases, the covert channel capacity decreases. The effect of handling policies on covert channel capacity is as discussed in the previous section.

2. Cycle length, T_c

The channel capacity depends on the number of states to which the system can transition, which is the range of idle slots that a period can contain. This range is computed by finding the number of idle slots per period for the maximum and minimum variation in load. The covert channel capacity is given by $\frac{(\log_2\ states)}{T_c}$. Therefore increasing the cycle length by some factor will reduce the channel capacity by the same factor.

While increasing cycle length reduces the channel capacity, it gives the sender more time to effect the changes to system load, i.e., the sender can smooth out the variation over a longer period. This implies that the interval over which the variability of traffic was considered to be "normal" will be extended, making it easier to audit the channel.

3. Granularity of change in number of idle slots, δ

Given a range of idle slots that a period can contain, the number of states depends on the granularity of change, δ. The coarser the granularity, smaller is the number of possible states, thereby reducing the channel capacity.

As discussed earlier, figure 5 shows the effect of granularity on covert channel capacity. The solid curve shows the channel capacity under maximum load conditions and the dashed curve shows the capacity under minimum and average load conditions. Due to the handling policies, the channel capacities at minimum and average load conditions are almost the same.

Coarser granularity means that the sender should be able to cause a large enough change in the system load

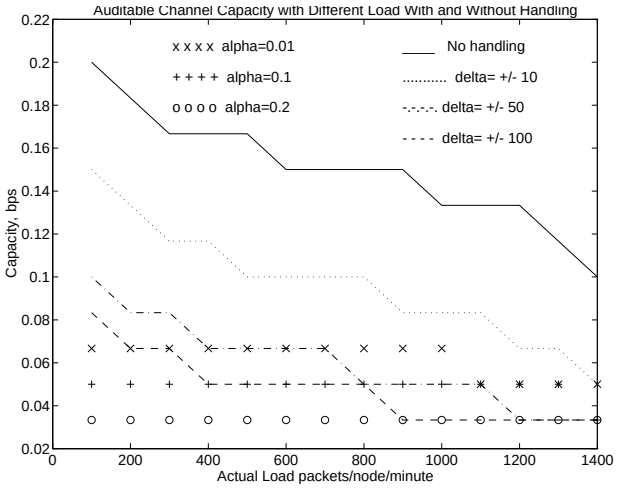

Figure 4: Noiseless Covert Channel Capacities for Different Load

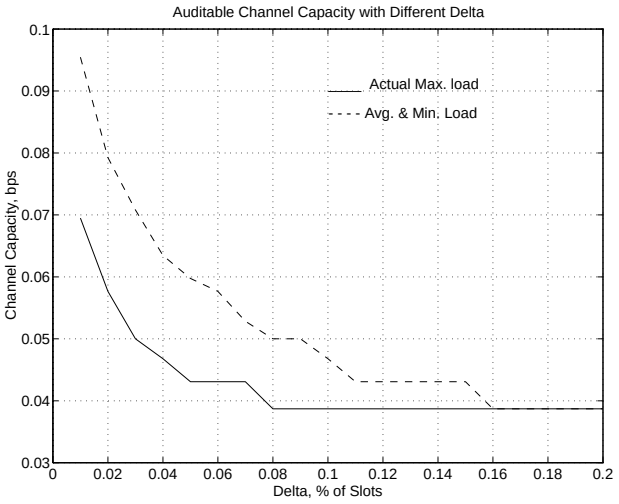

Figure 5: Noiseless Covert Channel Capacities for Different Delta

for the number of idle slots to change. This improves the auditability of the channel. Finer granularity allows the sender to change the load by a very small fraction of the current load and still manage to transmit a symbol. However, coarser granularity reduces the system's responsiveness. The system will remain in the current state unless the system load has significantly changed. This might lead to poor quality of service and under-utilization of system resources.

4. Auditability threshold, θ

Auditability threshold, θ, determines the variation in system load that will be accepted as "normal"; any variation above this threshold is a candidate for audit analysis. This threshold is determined by studying the traffic characteristics and is set to a value such that most of the variations that occur during the course of normal system operation is excluded from scrutiny.

The threshold should be set so that no potential covert channel usage is undetected and false detection of covert channel use is avoided. A larger than optimal threshold will allow potential usage of the covert channel to go undetected. A smaller than optimal threshold will trigger too many spurious audit events using up expensive resources and reducing the confidence in the audit mechanism by auditing innocuous events. Figure 3 shows the effect of different auditability thresholds on covert channel capacity.

5.2 Limiting Channel Capacity to Desired Levels

Having discussed the effects of various parameters on covert channel capacity, in this section we give a procedure to determine the values of the parameters for a desired channel capacity. We note that while parameters such as the number of nodes in the network, system load and slot time are fixed by the underlying network, the network administrator can decide parameters such as the auditability threshold, θ, and the granularity of change in the number of idle slots, δ.

The auditability threshold primarily depends on the variability of the system load and the degree of "completeness" desired of the auditability analysis. The granularity of change in the number of idle slots depends on the covert channel capacity that we are willing to accept in lieu of system responsiveness and the handling policy.

For a given number of nodes in the network, n, the slot time, T_s, the number of idle slots in a period for any particular load is given by
Number of idle slots = Total number of slots − Number of Active slots,

$$idle = \frac{(n-1)}{T_s \times load} - n(n-1),$$

where load is in packets per node.

Since the covert channel capacity depends on the number of states, we have to find the range of the number of idle slots in a period. Since the auditability threshold is θ, the variation in load that we need to consider is $load \pm \theta$.

$$\begin{aligned} \text{Range of idle slots} &= (\text{idle slots at } load - \theta \times load) \\ &\quad - (\text{idle slots at } load + \theta \times load) \end{aligned}$$

$$\begin{aligned} range = &[\tfrac{(n-1)}{T_s \times (load - \theta \times load)} - n(n-1)] \\ &- [\tfrac{(n-1)}{T_s \times (load + \theta \times load)} - n(n-1)] + 1 \end{aligned}$$

where the number of active slots, $n(n-1)$, is deducted.

Simple algebraic manipulation gives

$range = \frac{a}{load}(\frac{1}{(1-\theta)} - \frac{1}{(1+\theta)}) + 1$

where $a = \frac{(n-1)}{T_s}$.

Therefore we have

$range = \frac{a}{load}\frac{2\theta}{(1-\theta^2)} + 1$

From this relation, we know the maximum covert channel capacity without handling. In handling the covert channels, we attempt to reduce it potential capacity by reducing the range of states possible. This is done by varying the granularity of change in the number of idle slots either by a constant amount or by making the change proportional to the system load.

Once we know the range of the idle slots in a period, we can derive the covert channel capacity, C, as

$C = \frac{\log_2(\frac{range}{\delta})}{T_c}$ bps

$\log_2(range) - \log_2 \delta = T_c\, C$, or

$\log_2 \delta = \log_2(range) - T_c\, C$

where δ is the granularity of change in the number of idle slots.

If we are given C, the desired covert channel capacity, we can find the granularity of the change in the number of idle slots as

$\log_2 \delta \geq \log_2(range) - T_c\, C$, or

$\delta \geq 2^{\log_2(range) - T_c\, C}$, or

$\delta \geq \frac{range}{2^{T_c\, C}}$

If the handling policy is fixed, then the the granularity, δ, is as computed. If we use proportional handling policy, then $\delta = \alpha \times slots$, where $slots$ is the total number of slots in a period. Using the above relation, we can determine the granularity of the change in the number of idle slots such that the covert channel remains constant regardless of the changes to the system load.

5.3 Summary of Auditability Analysis

As discussed above, network covert channels can be effectively contained using relatively simple audit mechanisms. Our analysis of the traffic on ECSNET showed that the normal variation in the traffic is $\pm$24% of the current load on the system. Therefore we set the audit threshold at 24%.

Covert channel capacity is estimated for average load condition with and without handling policies. It is observed that the handling policies reduces the channel capacity to TCSEC acceptable levels. We have also estimated the channel capacity for various auditability thresholds, θ, granularity of change in the number of idle slots, δ, and the covert channel capacity for various load conditions with and without any handling mechanism for noiseless covert channels.

While it is difficult to audit noisy channels, audit mechanisms that employ knowledge of recent network behavior by maintaining a history of resource utilization for each node and the network as a whole can better audit noisy covert channels. Any change in a particular node's traffic is compared with previous variations and if the current variation is unusual, then the node is audited.

An active intruder can continuously vary system load either to actually exploit the covert channel to communicate with an accomplice or to fabricate an history of constantly varying traffic characteristics. Such a behavior will foil the above handling policy. Therefore in addition to monitoring individual nodes for variations in traffic characteristics, monitoring the variability of variations in traffic characteristics of a node can give additional insights into the behavior of the nodes and detect potential covert channel usage.

6 Conclusion

In this paper, we define network covert channels and give examples of covert channels due to spatial and temporal variation in transmission characteristics. Modeling the system as a mode based security system and using adaptive scheduling policy, we estimate the capacity of network covert channels. Using traffic characteristics from our measurements on UFNET, we audit certain covert channels and use various handling policies to reduce the capacity of network covert channels. We see that using the proportional handling policy, irrespective of network load, we can reduce the capacity of covert channels to TCSEC acceptable levels.

We wish to point out that while awareness of time of day variation in traffic volume might help a network trojan horse to better exploit a network covert channel, we ignored such variations in this study and used the average traffic volume during the observation period to characterize network traffic. It should be noted that a trojan horse hoping to exploit the time of day variation will need continuous access to the network over a period of time in order to monitor and collect network utilization statistics. This activity is expensive in terms of the time (e.g., strategic value of the information will diminish over time), processing and storage resources required. Also, covert communication on such a channel needs additional synchronization, such as with the time of the day clock.

Our efforts are currently directed towards estimating the capacity of spatial covert channels. We are also

interested in evaluating the effectiveness of random routing policy which yield traffic matrices that are almost neutral.

Acknowledgements

We wish to thank the anonymous reviewers for suggesting, among other things, that we justify the selection of the audit threshold and the fact that we do not consider the time of the day variation in our model. We would also like to thank Mr. Sam Madani of Hewlett-Packard Company for sharing his insights into the LAN traffic characteristics.

References

[1] R. Browne, "Mode Security: An Infrastructure for Covert Channel Suppression," *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, Oakland, California, May 16-18, 1994. pp. 39-55.

[2] Department of Defense. Department of Defense Trusted Computer System Evaluation Criteria. Report DOD 5200.28-STD, Department of Defense, Washington DC, December 1985.

[3] A. L. Donaldson, J. McHugh, and K. A. Nyberg, "Covert Channels in Trusted LANs," *Proceedings of the 11th National Computer Security Conference*, Washington DC. October 17-20, 1988, pp. 226-232.

[4] K. W. Eggers and P. W. Mallet, "Characterizing Network Covert Storage Channels," *Aerospace Computer and Communication Security Conference*, Washington DC. 1988, pp. 275-279.

[5] G. C. Girling, "Covert Channels in LAN's," *IEEE Transactions on Software Engineering*, Vol SE-13, No. 2, February, 1987, pp. 292-296.

[6] Hu Wei-Ming, "Reducing Timing Channels With Fuzzy Time", *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, Oakland, California, May 20-22, pp. 8-20, 1991. *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, Oakland, California, April 27-29, 1987, pp. 60-66.

[7] J. C. Huskamp, "Covert Communication Channels in Timesharing Systems," Technical Report UCB-CS-78-02, Ph.D. Thesis, University of California, Berkeley, California, 1978.

[8] International Standards Organization. Final Text of DIS 7498-2, Information Processing Systems–OSI Reference Model–Part 2: Security Architecture, ISO/IEC, International Standards Organization, July 1988.

[9] M. H. Kang and I. S. Moskowitz, "A Pump for Rapid, Reliable, Secure Communication," *Proceedings of the 1st ACM Conference on Computer and Communications Security*, pp. 119-129, 1993.

[10] K. E. Kirkpatrick, "Standards for Network Security," Proceedings of the 11th National Computer Security Conference, Washington DC, 17-20 October, 1988, pp. 201-211.

[11] J. K. Millen, "Covert Channel Capacity," *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, Oakland, California, April 27-29, 1987, pp. 60-66.

[12] J. K. Millen, "Finite State Noiseless Covert Channels," *Proceedings of Computer Security Foundations Workshop II*, Franconia, New Hampshire, 1989. pp. 81-86.

[13] I. S. Moskowitz and A. R. Miller, "Simple Timing Channels", *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, Oakland, California, May 16-18, 1994, pp. 56-64.

[14] National Computer Security Center. Trusted Network Interpretation Environments Guideline, Report NCSC-TG-005 Version-1, National Computer Security Center, Ft. George G. Meade, Maryland, August 1987.

[15] National Computer Security Center. A Guide To Understanding Covert Channel Analysis of Trusted Systems, Report NCSC-TG-030, Version 1, National Computer Security Center, Ft. George G. Meade, Maryland, November 1993.

[16] R. E. Newman-Wolfe and B. R. Venkatraman, "High Level Prevention of Traffic Analysis," *Seventh Annual Computer Security and Applications Conference*, San Antonio, Texas, December 2-6, 1991, pp. 102-109.

[17] R. E. Newman-Wolfe and B. R. Venkatraman, "Performance Analysis of a Method for High Level Prevention of Traffic Analysis," *Eighth Annual Computer Security and Applications Conference*, San Antonio, Texas, November 30 - December 4, 1992, pp. 123-130

[18] M. A. Padlipsky, D. W. Snow, and P. A. Karger, "Limitations of End-to-End Encryption in Secure Computer Networks," Electronic Systems Division Technical Report, ESDTR 78-158 (AD A059221), The MITRE Corporation, Bedford, Massachusetts, May 1978.

[19] L. S. Rutledge and L. J. Hoffman, "A Survey of Issues in Computer Network Security," *Computers and Security*, Vol. 5, 1986, pp. 296-308.

[20] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana, Illinois, 1949.

[21] S. Shieh and V. D. Gligor, "Auditing the Use of Covert Storage Channels in Secure Systems," *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, Oakland, California, 1990, pp. 285-295.

[22] C. Tsai, "Covert Channel Analysis in Secure Computer Systems," PhD Dissertation, University of Maryland, College Park, 1987.

[23] B. R. Venkatraman and R. E. Newman-Wolfe, "Transmission Schedules To Prevent Traffic Analysis" *Ninth Annual Computer Security and Applications Conference*, Orlando, Florida, December 6-10, 1993, pp. 108-115.

[24] B. R. Venkatraman, "Prevention of Traffic Analysis and Associated Covert Channels", Ph.D. Dissertation, University of Florida, Gainesville, 1994.

[25] 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", *Tenth Annual Computer Security and Applications Conference*, Orlando, Florida, December 5-9, 1994.

[26] V. L. Voydock and S. T. Kent, "Security Mechanisms in High-Level Network Protocols," *ACM Computing Surveys*, Vol. 15, No. 2, June 1983, pp. 135-171.

[27] V. L. Voydock and S. T. Kent, "Security in High-Level Network Protocols," *IEEE Communications Magazine*, July 1985, pp. 12-24.

[28] R. Ward, "OSI Network Security and the NTCB," *Lecture Notes in Computer Science*, No. 396, Springer-Verlag, 1989, New York, pp. 67-74.

[29] M. Wolf, "Covert Channels in LAN Protocols," *Lecture Notes in Computer Science*, 396, Springer-Verlag, New York, 1989, pp. 67-74.