Proceedings of the IEEE Symposium on Security and Privacy, Pages 2–11, Oakland, California, May 6–8, 1996

An Analysis of the Timed Z-channel

Ira S. Moskowitz Steven J. Greenwald Myong H. Kang
Information Technology Division, Mail Code 5540
Center for High Assurance Computer Systems
Naval Research Laboratory
Washington, DC 20375 *

Abstract

Our timed Z-channel (a general case of the Z-channel) appears as the basis for a large class of covert channels. Golomb analyzed the Z-channel, a memoryless channel with two input symbols and two output symbols, where one of the input symbols is transmitted with noise while the other is transmitted without noise, and the output symbol transmission times are equal. We introduce the timed Z-channel, where the output symbol transmission times are different. Specifically, we show how the timed Z-channel applies to two examples of covert timing channel scenarios: a CPU scheduler, and a token ring network. We then give a detailed analysis of our timed Z-channel. We report a new result expressing the capacity of the timed Z-channel as the log of the root of a trinomial equation. This changes the capacity calculation from an optimization problem into a simpler algebraic problem and illustrates the relationship between the noise and time factors. Further, it generalizes Shannon's work on noiseless channels for this special case. We also report a new result bounding the timed Z-channel's capacity from below. Finally, we show how an interesting observation that Golomb reported for the Z-channel also holds for the timed Z-channel.

1. Introduction

Covert timing channels arise from resource sharing in MLS systems. High can pass information to Low, by either interfering with, or refraining from interfering with, the timing of Low's activities. In most of these systems this interference is noisy. The simplest model for such interference, where the output alphabet consists of time values, is what we call the *timed Z-channel*. Knowledge of the characteristics of the timed Z-channel should allow the system designer to engineer countermeasures to this danger.

We discuss in detail two scenarios where the timed Z-channel may occur as a serious threat. Our first scenario is a generalization of the well-known CPU scheduling channel [17, 25], as discussed in a mathematical sense by Huskamp [9, section 4]. It is very important to understand noisy versions of this scenario because many researchers are currently investigating countermeasures to this scenario and its variants (e.g., [8, 7, 30, 10]). Note that McCullough's [20] "half-bit channels" may be analyzed as timed Z-channels.

Our second scenario is quite different, dealing with a theoretical MLS computer network organized as a token ring topology. We show how a timed Z-channel can be exploited as a covert channel in a specific configuration of this network. Considering the current popularity of ring topologies for networks (e.g., FDDI, FDDI-II, etc.), we feel that understanding the behavior of any potential threat to MLS implementations of this type of network is desirable. We demonstrate that a covert timed Z-channel threat exists under certain circumstances. Given the existence of such a threat, we feel that the designers of MLS token ring networks should be aware of the issues and mathematical tools needed to recognize this network threat.

Some of the important questions being investigated that relate to this paper follow.

- Is capacity large enough to be of concern?
- Can understanding the mathematical interplay between the noise and time variables in this covert channel be used to lessen capacity?
- How would the intentional introduction of noise affect both capacity and system performance?

Because of the above, we feel that an analysis of the capacity of the timed Z-channel is of great importance.

*{moskowitz,greenwald,mkang}@itd.nrl.navy.mil

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1996		**00-00-1996 to 00-00-1996**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
An Analysis of the Timed Z-channel	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Research Laboratory,Center for High Assurance Computer Systems,4555 Overlook Avenue, SW,Washington,DC,20375**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**10**	

The known capacity results on the Z-channel do not extend to the timed Z-channel. Theorem 1 and Corollary 1 (the main mathematical results of this paper) show how capacity can be easily expressed as the log of a zero of a trinomial. This lets us transform a complicated optimization problem into an algebraic problem. In turn, this gives us a simple method for calculating capacity and seeing the interplay between the noise and timing factors. Knowledge of the interplay between these various terms can lead us to a better understanding of how to lessen capacity without degrading performance. This has been seen in papers such as [3, 4], where noise is introduced in the system to lessen capacity. The best defense to covert channel threats is a thorough understanding and analysis of covert channel behavior. Knowledge of similar system behavior was of great significance during the design and implementation of the NRL Pump [12, 13, 14, 23].

We now present our two scenarios in detail, give some background on the Z-channel, and then give a detailed exposition and analysis of our timed Z-channel.

2. Covert Channel Scenarios

There are two scenarios presented here. The first is very well-known in the field of covert channel analysis. The second is a previously undiscovered covert timing channel existing in an interesting three level environment (to date, most discussions of covert channel attacks deal with only two levels).

2.1. Scenario 1 — The CPU scheduler type channel

The following type of configuration is common in computer systems. Assume that there are two levels L1 and L2, where L2 > L1.

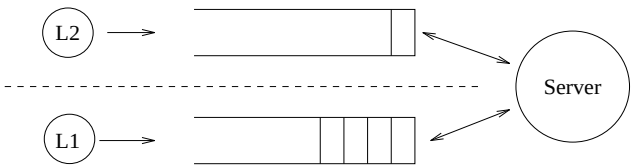

Figure 1. A Queuing Diagram

There are two queues (i.e., Q2 and Q1), where L2 processes put their jobs/messages into Q2 and L1 processes put their jobs/messages into Q1. A server, which is a shared resource, provides service in a round-robin fashion. A server may be a CPU which processes jobs from two different levels. This is illustrated in figure 1.

In such a scenario there exists a well-known covert timing channel from L2 to L1 [17, 25, 9]. We also note that studies of this type of covert channel have been useful in other areas, such as the analysis of the generalized version of the NRL Pump [15]. For simplicity, let us assume that each job takes time δ. Assume that an L1 process (e.g., BL1) submits jobs and observes the time to complete each job. If an L2 process (e.g., BL2) does not submit any job, each job in Q1 takes time δ to complete. If BL2 submits a job, BL1 observes time 2δ to complete one of its jobs (i.e., time δ for the BL2 job and an additional time δ for the BL1 job). Therefore, we have the noiseless timing channel illustrated in figure 2.

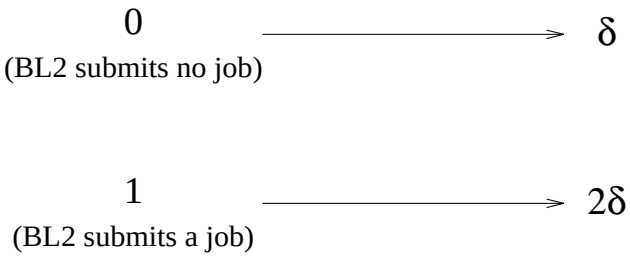

Figure 2. A Noiseless Timing Channel

The capacity, in bits per time unit, of this noiseless channel is known to be $\delta^{-1} \log \frac{1+\sqrt{5}}{2}$, [26, 21]. BL2 may decide to communicate with BL1 even in the presence of noise. The noise can be introduced by another L2 process (e.g., GL2) that does not have any intention of communicating with BL1. If GL2 submits a job and BL2 does not submit a job (i.e., BL2 attempts to send a binary 0), BL1 still observes time 2δ which will be interpreted as a binary 1 from BL2. Therefore, we have a covert timed Z-channel where $p + q = 1$ as illustrated in figure 3. This type of channel, where capacity (bits per time unit) is not known, is the focus of this paper.

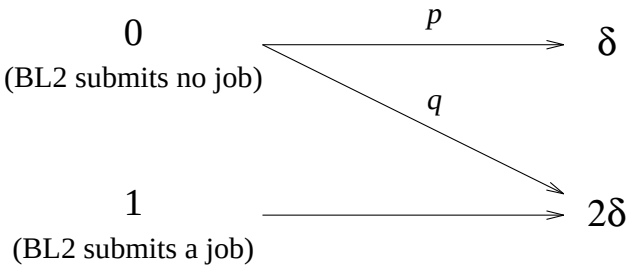

Figure 3. A Timed Z-channel

2.2. Scenario 2 — A Token Ring Timing Channel

A *token ring* is a type of computer network organized as a set of stations arranged in a ring topology,

either physically, through serial connections of transmission media such as twisted pairs (e.g., the IEEE 802.5 Token Ring Local Area Network standard [27]) or fiber optic links (e.g., the Fiber Distributed Data Interface [5, 1, 11]), or logically (e.g., the IEEE 802.4 Token Bus Local Area Network standard [29]). Information is sequentially transferred from station to station, circulating around the ring. Each station on the ring has a unique address, and each information packet transmitted on the ring contains (among many other items) a destination address, a source address, and a special bit used to detect when a packet has been received. If a station wishing to transmit data has access to the medium, it transfers an information packet onto the ring, where the packet circulates unidirectionally from one station to the next. Eventually, the destination station copies the information as it passes by on the ring, and modifies the special bit to serve as an ACK to the source station. If all is working well, the packet eventually is received by the originating station and is removed from the ring.

A station gets the right to transmit a packet on the ring when it detects a special message called the *token* passing on the ring medium. The token is a symbol of authority passed between stations. It is used as a method for enforcing mutual exclusion among stations contending for the ring transmission medium (only one station is allowed to control the ring at any given time). A station wishing to transmit a packet captures the token as it passes by, and holds it until it is finished with its transmission (the time the token is held can vary, or can be isochronous in the case of FDDI-II [28]). Generally, there is a maximum period of time that a station may hold the token (and therefore control the network). Once a token holding station is finished with its transmissions, it releases the token to its downstream neighbor. General token ring management issues, such as initialization, error recovery, etc., are not dealt with in this paper, but are available in the references noted above. The reader who is not familiar with token ring networks may take it for granted that mechanisms exists for error recovery, ring initialization, adding and subtracting stations, and so forth.

A token holding station wishing to transmit information formats a packet containing (among other things) its source and the destination address, and transmits the packet to its downstream neighbor station. That station, via a hardware mechanism in the network interface, examines the destination address in the packet header, and if the packet is not destined for that station, it passes the packet on to the next station in the ring. If a station finds that a packet is destined for itself, it *receives* the packet, modifies the special bit in the packet signifying that it received the packet, and passes the packet on to its downstream neighbor station. Eventually, the packet will travel to the originating station (i.e., the station holding the token). The originating station then examines the special bit to verify that the packet was received. It then passes (releases) the token on to the next station in the ring.

Of interest to us is an MLS token ring network. This is a token ring network where each station has a level. Communication between stations must obey the Bell-LaPadula (BLP) principles [2]. Such a token ring would allow only reception of packets by a station at a level that *dominates* the level of the transmitting station. Of course, the token is passed independently of any BLP considerations.

Such an MLS token ring network (loosely based on the IEEE 802.4 and 802.5 standards) would require hardware not needed for a non-MLS token ring network. A modification to the hardware would be needed so that all packets would have a label attached, indicating their *Mandatory Access Control* (MAC) level (e.g., Low, Medium, High). In addition, we would want the hardware to enforce our MAC BLP-like policies, so that a packet sent from a station at a higher level could not be received by a station at a lower level (i.e., in such a case, the packet would be sent unread to the downstream station). This could be easily accomplished by modifying (if necessary) the ring physical layer interface hardware of each station so that it fails to recognize any packets that have been transmitted by a station at a level higher than itself.

A covert timing channel may exist in such an MLS token ring network when there are exactly three stations, S_l, S_m, and S_h, and where each station has the respective levels of Low, Medium, and High (see figure 4; PHY is our trusted physical layer interface). The actual physical location of the stations on the ring is unimportant.

Let us examine our hypothetical MLS Token Ring network with three stations, each with a unique MAC level. According to the BLP policy, there are only three allowable transmissions: S_l to S_m, S_l to S_h, and S_m to S_h. Suppose S_m wishes to send information to S_l in violation of the ring's MLS policy. Further suppose that S_m and S_l are acting together to exploit the covert timing channel that follows. There are only two basic time intervals that are of interest, from the viewpoint of S_m and S_l: $3t$ and $6t$.

If no station wishes to transmit a message, then the token circulates completely around the ring in time $3t$, where t is the time to transmit a message from one station to another (we have assumed for the sake of

simplicity in this explanation that each link transmission takes the same amount of time). Explicitly, if the token is initially held by S_l, it takes time t to transmit the token to S_l's downstream neighbor station, it takes time t for S_l's downstream neighbor station to transmit the token to S_l's upstream neighbor station, and it takes time t to transmit the token from S_l's upstream neighbor station back to S_l. Therefore, the total time for this event is $3t$.

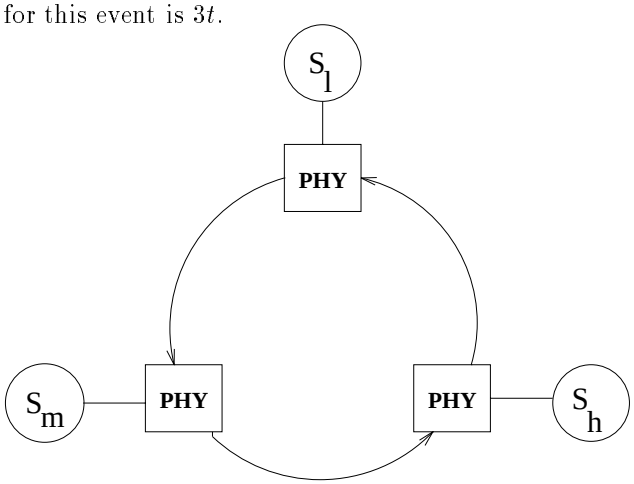

Figure 4: An MLS Token Ring with 3 Levels and 3 Stations

If any station (say S_l's downstream neighbor) wishes to transmit a message, then the time between S_l initially passing the token, and then next receiving it is time $6t$. In other words, if the token is initially held by S_l, it takes time t to transmit the token to S_l's downstream neighbor station (suppose this station happens to be S_m as shown in figure 4). S_m then captures the token, and sends a message (necessarily destined for S_h in this case) to its downstream neighbor station, S_h, taking time t. S_h then receives the message, modifies the special bit indicating that it received the message, and transmits the message to its downstream station (S_l), taking time t. S_l's ring interface hardware notices that the message is not destined for S_l, so it passes the message back to S_m, again taking time t. S_m is now finished with its transmission, so it passes the token to its downstream neighbor (S_h) once more taking time t. S_h has nothing to send to anyone (due to the BLP policy) so it transmits the token back to S_l in time t. Therefore, the total time for this event is $6t$.

So we see that there are only two possible output symbol times that exist ($3t$ and $6t$). We believe the two output symbol times can be exploited by a covert timing channel which we now describe.

Suppose that S_m wishes to send information to S_l in violation of the BLP policy, and that S_m and S_l are cooperating (i.e., they are using the following protocol). S_l remains completely passive, noting only when it has the token and how much time has elapsed since it last held the token (i.e., when S_l transmits the token, it starts a timer used to note how much time has elapsed until it next receives the token). These times will be either $3t$ (i.e., no transmissions have occurred), or $6t$ (a transmission has occurred from S_m to S_h — the only BLP allowable transmission that S_l does not initiate). If S_m wishes to send S_l a binary zero, then when it holds the token it transmits nothing (i.e., it just passes the token on to its downstream neighbor) and when S_l next holds the token it will notice that only time $3t$ has passed, and will interpret this event as a binary zero. If S_m wishes to send S_l a binary one, then when it holds the token, it sends a message to S_h (e.g., a simple "ping" would suffice) and when S_l next holds the token it will notice that time $6t$ has passed and will interpret this event as a binary one.

Noise (we discount thermal noise because it is a very unlikely event in most modern ring networks, generally on the order of 10^{-9} for FDDI) is present in this system as legitimate messages. This is because legitimate messages must be sent at some point, or there would be no reason for the network to exist. Of the 3 allowable communications, two of them are originated by S_l which obviously is aware of its own transmissions and can treat these events as a temporary suspension of covert channel operations (S_m would also be aware of this through either receiving a message from S_l, or by inference). Only one valid transmission, from S_m to S_h could introduce noise into this system. However, the noise introduced takes exactly time $6t$, the same time as the second symbol. Therefore, we are dealing with a timed Z-channel.

The situation described is quite possible. In any MLS Token Ring network that has three or more stations, there will, at some point in the life of the network, exist a three station configuration (this is due to the method for generating the network and adding new stations — the interested reader is referred to [29] and [27]). It is certainly possible that the three station configuration will not contain three different levels. However, unless this can be ruled out, a timed Z-channel threat must be taken seriously. In addition, since FDDI and FDDI-II are organized as token rings, and given the current popularity of them, we feel it likely that such a scenario may occur. This is especially so in the isochronous FDDI-II, which lends itself to real-time applications. Note that in 1977 Karger [16, Ch. 11] mentioned that covert channels could arise by modulating inter-packet transmission times in generic distributed systems.

2.3. Discussion about the Scenarios

Note that we have restricted ourselves to the simplified case of only two output symbols in the presentation of the noisy timing channel scenarios. It has been shown for noiseless timing channels [24] that introducing extra symbols can greatly increase the capacity. The full problem requires further research, since we did not find it as tractable for noisy timing channels.

Also, note that we are using time values t_1 and $2t_1$, when we could have used t_1 and $t_1 + \epsilon$ where ϵ need not equal t_1. The capacity analysis we do in this paper is the more general case where the two time values are not necessarily multiples of one another.

Before we can perform an analysis of channels with more than two output symbols, we must first understand the simpler case of the timed Z-channel. We hope for a more complete analysis in the future.

3. Mathematical Background: Golomb's Z-channel

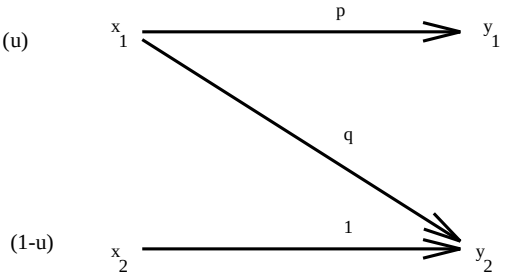

Figure 5. The Z-channel

In [6] Golomb succinctly analyzed the mutual information and channel capacity of the Z-channel (see figure 5). The Z-channel[1] is a discrete memoryless channel with two input symbols. One symbol, x_2, is transmitted without noise and received as y_2, while the other input symbol x_1 is transmitted with noise and received as either y_1 or y_2. We let u represent the probability $P(x_1)$ that x_1 is the input. Therefore, $P(x_2) = 1 - u$. The amount of noise (constant with each transmission) is given by the following channel matrix ($q = 1 - p$), which describes the conditional probabilities.

$$\begin{pmatrix} P(y_1 \mid x_1) & P(y_2 \mid x_1) \\ P(y_1 \mid x_2) & P(y_2 \mid x_2) \end{pmatrix} = \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix}$$

[1] We use a channel where the second symbol is transmitted noiselessly. The standard descriptions of the Z-channel have the first symbol transmitted noiselessly. There is no difference other than pictorially. However, our reflected picture helps with the physical intuition of the timed Z-channel introduced later.

Let us now calculate the mutual information I (in units of bits per symbol). I is the difference in entropies $H(Y) - H(Y \mid X)$. The output entropy $H(Y) = H(y_1, y_2)$, where we are using the shorthand notation[2] of $H(a, b) = -\{a \log a + b \log b\}$. We condition by

$$P(y_j) = P(y_j \mid x_1)P(x_1) + P(y_j \mid x_2)P(x_2)$$

to facilitate our calculations. Thus, $P(y_1) = up$ and $P(y_2) = 1 - up$. So, $H(Y) = H(up, 1 - up)$. The conditional entropy is

$$H(Y \mid X) =$$

$$P(x_1)H(Y \mid x_1) + P(x_2)H(Y \mid x_2) =$$

$$P(x_1)H(p, q) + P(x_2)H(0, 1) = uH(p, q) \ .$$

Thus we see that [6, Eq. 1]

$$I = H(up, 1 - up) - uH(p, q) \ .$$

In brief, a timed Z-channel is identical to Golomb's Z-channel except that the output symbol y_2 has a greater transmission time than y_1.

Implicit in our discussions of Golomb's Z-channel is that the symbols take the same amount of time to be transmitted. The time that the symbols take to transmit is not an issue. Therefore, units are in bits per symbol when we are dealing with the Z-channel. To avoid confusion, I (I_t) will be mutual information in bits per symbol (tick, the time unit), and C (C_t) will be capacity in bits per symbol (tick). In general, for discrete memoryless channels, if all symbols take the same time τ to be transmitted, we have $I_t = \tau^{-1} I$ and $C_t = \tau^{-1} C$. However, for the timed Z-channel, transmission times of the two symbols are different, so C_t is not a multiple of C. In the next section we analyze the timed Z-channel.

4. The Timed Z-channel

We see from our above discussions that we may abstract our covert timing channel scenarios to a memoryless channel with two input symbols x_1 and x_2. Input symbol x_2 is transmitted without noise and is interpreted as the output symbol y_2, taking time t_2. Input symbol x_1 either arrives with probability p taking time t_1 and is interpreted as the output symbol y_1, or it arrives with probability $q = 1 - p$ taking time t_2 and is interpreted as the output symbol y_2. Note the output symbols are distinguished by the differing time values, whereas in the Z-channel it is not the time that distinguishes them, but the fact that $y_1 \neq y_2$. We will use

[2] All logarithms are base 2.

the notation "$Z(\epsilon)$-channel", where ϵ is the difference in time $t_2 - t_1 \geq 0$, for the timed Z-channel (see figure 6). Of course, the channel matrix of the $Z(\epsilon)$-channel is identical to that of the Z-channel. Therefore, I for the $Z(\epsilon)$-channel is the same as I for the Z-channel. Note that the Z-channel is just a $Z(0)$-channel.

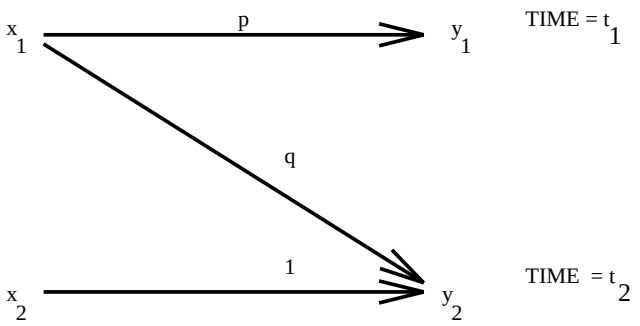

Figure 6. The $Z(\epsilon)$-channel

Let T represent the Bernoulli random variable that describes the time that an output symbol arrives. The probability $P(T = t_j)$, or written more simply as $P(t_j)$, is the probability $P(y_j)$. Hence, the expected value (mean) of T is $E(T) = t_1(up) + t_2(1 - up)$. Since $\epsilon = t_2 - t_1$ we have

$$E(T) = t_1 + \epsilon(1 - up)\,.$$

We see that if the channel is noiseless that $E(T)|_{q=0} = t_1 + \epsilon(1 - u)$. If the channel is totally noisy (useless) we see that $E(T)|_{q=1} = t_1 + \epsilon$.

We now wish to analyze the mutual information in bits per tick, $I_t = \frac{I}{E(T)}$. From our above equations we see that

$$I_t = \frac{H(up, 1 - up) - uH(p, q)}{t_1 + \epsilon(1 - up)}\,.$$

We wish to calculate the capacity in units of bits per tick, C_t, for the $Z(\epsilon)$-channel. Verdú [31] has shown that $C_t = \max_u I_t$, and that this is the proper measure of maximal asymptotically error-free information flow. We can find C_t by setting $\frac{dI_t}{du} = 0$ and solving for u. We denote the value of u that maximizes I_t by u_c.

$$\frac{dI_t}{du} =$$

$$\frac{[t_1 + \epsilon(1-up)][p\log(\frac{1-up}{up}) - H(p,q)] + \epsilon p[H(up,1-up) - uH(p,q)]}{[t_1 + \epsilon(1-up)]^2}\,.$$

Setting the derivative to zero gives us

$$0 = t_1 p \log\left(\frac{1-up}{up}\right) - t_1 H(p,q) + \epsilon p \log\left(\frac{1-up}{up}\right)$$
$$-\epsilon H(p,q) - \epsilon u p^2 \log\left(\frac{1-up}{up}\right)$$
$$+\epsilon u p H(p,q) - \epsilon u p^2 \log(up) - \epsilon p(1-up)\log(1-up) - \epsilon u p H(p,q)$$

which simplifies to

$$\left(\frac{t_1+\epsilon}{p}\right) H(p,q) = t_1 \log(\frac{1-up}{up}) + \epsilon \log\left(\frac{1-up}{up}\right)$$
$$-\epsilon up \log(\frac{1-up}{up}) - \epsilon up \log(up) - \epsilon \log(1-up) + \epsilon up \log(1-up)$$

which further reduces to

$$\left(\frac{t_1+\epsilon}{p}\right) H(p,q) = t_1 \log(\frac{1-up}{up}) - \epsilon \log(up)\,.$$

This simplifies to

$$\begin{aligned}
\log(1-up)^{t_1}(up)^{-(t_1+\epsilon)} &= \log\left(p^p q^q\right)^{-\left(\frac{t_1+\epsilon}{p}\right)} \\
(1-up)^{t_1} &= \left(p^p q^q\right)^{-\left(\frac{t_1+\epsilon}{p}\right)} (up)^{(t_1+\epsilon)} \\
1 - up &= \left(p^p q^q\right)^{-\left(\frac{t_1+\epsilon}{t_1 p}\right)} (up)^{\frac{t_1+\epsilon}{t_1}}\,.
\end{aligned}$$

Letting $\kappa = \left(pq^{q/p}\right)^{1/t_1}$ and $\gamma^{-t_1} = up$ we have the trinomial equation

$$1 - [(\kappa\gamma)^{-(t_1+\epsilon)} + \gamma^{-t_1}] = 0 \qquad (1)$$

which we refer to as the characteristic equation of the $Z(\epsilon)$-channel. Eq. (1) gives us the following useful identities:

$$(\kappa\gamma)^{-(t_1+\epsilon)} + \gamma^{-t_1} = 1$$
$$1 - \gamma^{-t_1} = (\kappa\gamma)^{-(t_1+\epsilon)}\,.$$

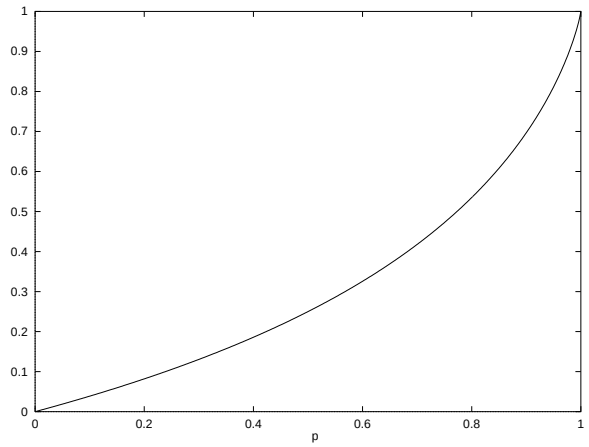

Figure 7. $pq^{q/p}$

Since the term $pq^{q/p}$ is so important we will include a plot of it (see figure 7). Keep in mind that $pq^{q/p}$ is defined by its limiting values of 0 and 1 at $p = 0$ and $p = 1$, respectively.

Of course, the variable γ in Eq. (1) is functionally dependent upon p, but Eq. (1) is very appealing because if $p = 1$, Eq. (1) reduces to

$$1 - [\gamma^{-(t_1+\epsilon)} + \gamma^{-t_1}] = 0 \qquad (2)$$

and $u_c = {r_1}^{-t_1}$, where r_1 is the unique positive root of Eq. (2). This is of interest because Shannon has shown that $C_t = \log r_1$, [26, 24, 22] (for this noiseless channel).

Now let us attempt to find a general closed form for the capacity of the $Z(\epsilon)$-channel. Using the identity $u_c p = {r_p}^{-t_1}$, where r_p is the positive root of Eq. (1) we see that

$$H(u_c p, 1-u_c p) = H\left({r_p}^{-t}, (\kappa r_p)^{-(t_1+\epsilon)}\right)$$

$$= -{r_p}^{-t_1}\log {r_p}^{-t_1} - (\kappa r_p)^{-(t_1+\epsilon)}\log(\kappa r_p)^{-(t_1+\epsilon)}$$

$$= t_1 {r_p}^{-t_1}\log r_p + (t_1+\epsilon)(\kappa r_p)^{-(t_1+\epsilon)}\left[\log r_p + \log\kappa\right]$$

$$= \left[t_1\left({r_p}^{-t_1} + (\kappa r_p)^{-(t_1+\epsilon)}\right) + \epsilon(\kappa r_p)^{-(t_1+\epsilon)}\right]\log r_p$$

$$+(t_1+\epsilon)(\kappa r_p)^{-(t_1+\epsilon)}\log\kappa$$

$$= \left[t_1 + \epsilon(\kappa r_p)^{-(t_1+\epsilon)}\right]\log r_p + (t_1+\epsilon)(\kappa r_p)^{-(t_1+\epsilon)}\log\kappa \ .$$

Since $H(p,q) = -\log(p^p q^q)$ we see:

$$\begin{aligned} u_c H(p,q) &= -u_c\log(p^p q^q) \\ &= -u_c p\log(pq^{q/p}) \\ &= -u_c p\log\kappa^{t_1} \\ &= -t_1 u_c p\log\kappa \\ &= -t_1 {r_p}^{-t_1}\log\kappa \ . \end{aligned}$$

This gives us $H(u_c p, 1-u_c p) - u_c H(p,q) =$

$$\left[t_1 + \epsilon(\kappa r_p)^{-(t_1+\epsilon)}\right]\log r_p$$

$$+(t_1+\epsilon)(\kappa r_p)^{-(t_1+\epsilon)}\log\kappa + t_1 {r_p}^{-t_1}\log\kappa$$

$$= \left[t_1 + \epsilon(\kappa r_p)^{-(t_1+\epsilon)}\right]\log r_p$$

$$+\left[t_1\left({r_p}^{-t_1} + (\kappa r_p)^{-(t_1+\epsilon)}\right) + \epsilon(\kappa r_p)^{-(t_1+\epsilon)}\right]\log\kappa$$

$$= \left[t_1 + \epsilon(\kappa r_p)^{-(t_1+\epsilon)}\right](\log r_p + \log\kappa)$$

$$= \left[t_1 + \epsilon(\kappa r_p)^{-(t_1+\epsilon)}\right]\log\kappa r_p \ .$$

The mean time to receive a symbol, with respect to the maximizing value of u is

$$E(T)|_{u=u_c} = t_1 + \epsilon(1-u_c p) = t_1 + \epsilon(\kappa r_p)^{-(t_1+\epsilon)} \ .$$

Since $C_t = I_t(u_c) = \dfrac{H(u_c p, 1-u_c p) - u_c H(p,q)}{E(T)|_{u=u_c}}$ we see that

Theorem 1 *$C_t = \log\kappa r_p$, where r_p is the positive root of $1 - [(\kappa\gamma)^{-(t_1+\epsilon)} + \gamma^{-t_1}] = 0$.*

By changing variables in Eq. (1) by letting $w = \kappa\gamma$, we see that κr_p is the positive root of $1 - [w^{-(t_1+\epsilon)} + (pq^{q/p})w^{-t_1}] = 0$. Therefore we have the following Corollary to Theorem 1.

Corollary 1 *$C_t = \log x_p$, where x_p is the positive root of $1 - [w^{-(t_1+\epsilon)} + (pq^{q/p})w^{-t_1}] = 0$.*

Theorem 1 and Corollary 1 collapse to Shannon's result [26] for $p = 1$. Corollary 1 mimics the form of Shannon's result expressing C_t as the logarithm of a zero of a polynomial. However, our polynomial does not have coefficients all 1 as Shannon's. Note that for the $Z(0)$-channel our result gives us $C_t|_{\epsilon=0} = \frac{1}{t_1}\log(1+pq^{q/p})$, which can be easily obtained from Golomb's paper. This is an intriguing generalization of Shannon's result on capacity.

We also note that our results turn the capacity calculation from an optimization problem into a much simpler algebraic problem. The algorithm necessary to calculate the capacity is nothing more than a simple root finder.

Let us see what else we can glean from the above results. We will use the notation $C_t(p)$ for the capacity of the $Z(\epsilon)$-channel with noise $q = 1-p$ (for the noiseless case the capacity is then $C_t(1)$). We are interested in the behavior of $C_t(p)$ as p varies from 1 to 0 (i.e., from noiseless to useless) for t_1 and ϵ fixed. By Theorem 1 we have $C_t(p) = \log\kappa r_p = \log r_p + \log\kappa$. Since $\kappa = \left(pq^{q/p}\right)^{1/t_1}$ we see that we have $C_t(p) = \log r_p + \frac{\log(pq^{q/p})}{t_1}$. We are adding a negative (we are not including the trivial comparison where $p = 1$) term to $\log r_p$. We will use this expression to bound $C_t(p)$ from below. Consider the equation $1 - [(\kappa\gamma)^{-(t_1+\epsilon)} + \gamma^{-t_1}] = 0$. This is equivalent to the equation $\gamma^{t_1+\epsilon} = \kappa^{-(t_1+\epsilon)} + \gamma^{\epsilon}$.

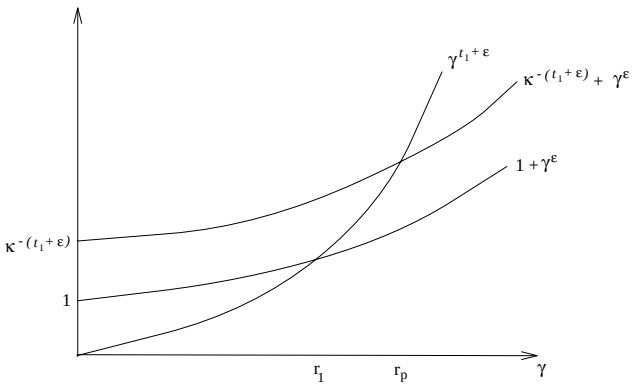

Figure 8. Comparison of the roots

Therefore, r_p is the value of γ where the plot of $\gamma^{t_1+\epsilon}$ intersects the plot of $\kappa^{-(t_1+\epsilon)} + \gamma^{\epsilon}$, and r_1 is the value of γ where the plot of $\gamma^{t_1+\epsilon}$ intersects the plot of $1+\gamma^{\epsilon}$ (see figure 8). Since $\kappa^{-(t_1+\epsilon)} > 1$ we see that $r_p > r_1$. Hence,

$$\begin{aligned} \log r_1 &< \log r_p \\ \log r_1 + \log\kappa &< \log r_p + log\kappa \\ C_t(1) + \log\kappa &< C_t(p) \ . \end{aligned}$$

Therefore the capacity $C_t(p)$ is never less than $C_t(1) + \log\kappa$. Since $C_t(p) < C_t(1)$ we have:

Theorem 2 *For the $Z(\epsilon)$-channel the capacity $C_t(p)$ is always bounded as*

$$\max(0,\ C_t(1) + \log\kappa) \leq C_t(p) \leq C_t(1)\ .$$

We may interpret this as a feasibility region where $C_t(p)$ must lie (see figure 9).

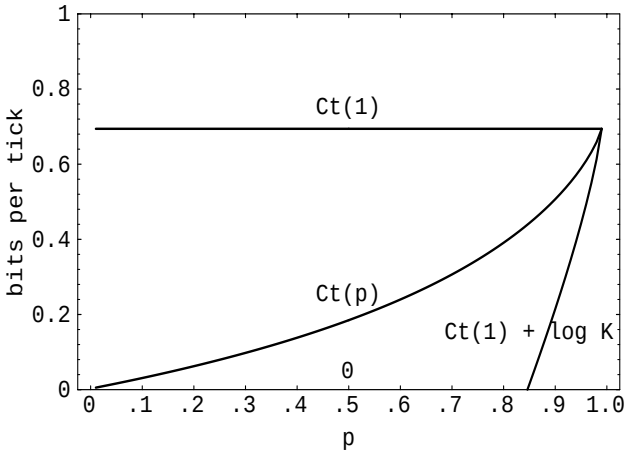

Figure 9. Illustration of Theorem 2 for $t_1 = \epsilon = 1$

Note that if we used Corollary 1 instead of Theorem 1 we would only obtain the known and obvious bound that $C_t(p) \leq C_t(1)$.

5. Comments on the very noisy $Z(\epsilon)$-channel

Majani [18] has done a systematic study of "very noisy" channels in his dissertation. We are only concerned with examining the $Z(\epsilon)$-channel as $p \to 0^+$ (which is a very noisy channel). Golomb looked at the Z-channel under the same behavior and noted some interesting behavior. We will show that the $Z(\epsilon)$-channel behaves similarly.

Golomb [6] showed that

Theorem 3 (Golomb) *For the Z-channel,*

$$\lim_{p\to 0^+} u_c = 1/e.$$

This result is remarkable because it implies that even though x_1 is received more and more often as y_2, we still must send x_1 a large amount of the time to achieve capacity. Further, Golomb showed that u_c varies only from 1/2 to $1/e$, as p varies from 1 to 0. This is of interest in light of a recent result of Majani and Rumsey (they are concerned with units of bits per symbol, not time) [18, 19]:

Theorem 4 (Majani & Rumsey)
For a binary-input discrete memoryless channel with $C > 0$, C is achieved when the the probability of the first symbol being input is in the interval $(1/e, 1-1/e)$.

Of course then the probability for the second symbol is also in the same interval. Therefore, Golomb's example when $p \to 0^+$ represents the limiting case of Majani and Rumsey's result, and shows that Theorem 4 is the best possible bound for u_c . Majani and Rumsey have noted that this does not hold, in general, for more than two input symbols. Note that the result does not hold for timing channels.

Counter-Example: Take the noiseless ($q = 0$) $Z(\epsilon)$-channel with $t_1 = 1$ and $\epsilon = 29$, then $u_c \approx .919 > 1 - 1/e$.
Therefore, Majani and Rumsey's result will not hold for timing channels.

Let us see if Golomb's result generalizes to the timed Z-channel by studying the $Z(\epsilon)$-channel when it is very noisy, i.e., $p \to 0^+$. As we remarked above, Golomb showed for the Z-channel that $u_c \to 1/e$. Therefore, we must study the root of Eq. (1) when $p = 0$. This is a little tricky because both κ and γ are functions of p. Recalling that $\gamma^{-t_1} = up$, so $\gamma^{-\epsilon} = (up)^{\epsilon/t_1}$ we may express Eq. (1) as:

$$up[(pq^{q/p})^{-1} q^{-(\epsilon q)/(t_1 p)} p^{-\epsilon/t_1} u^{\epsilon/t_1} p^{\epsilon/t_1} + 1] = 1\ .$$

Which, by letting $\delta = q^{q/p}$ simplifies to

$$p\delta u + (1/\delta)^{\epsilon/t_1} u^{1+\epsilon/t_1} = \delta\ . \qquad (3)$$

By using L'Hôpital's rule we see that $\lim_{p\to 0^+} \ln\delta = -1$, hence $\delta \to e^{-1}$. Hence as $p \to 0^+$, Eq. (3) collapses to $e^{\epsilon/t_1} u^{1+\epsilon/t_1} = e^{-1}$, so we have:

Theorem 5 *For the $Z(\epsilon)$-channel* $\lim_{p\to 0^+} u_c = e^{-1}$.

We find this result to be of great mathematical interest. It is worth noting that even though Majani and Rumsey's result on u_c values does not generalize to the timed Z-channel, Golomb's result on the boundary behavior of u_c does. Theorem 5 might be of use in the design of a code to exploit a very noisy timed Z-channel.

We wonder what the correct generalization of Majani and Rumsey's results are for the timed Z-channel. Results such as these are quite useful in estimating capacity, when a closed form might be difficult to derive.

6. Conclusion

We examined two scenarios where the timed Z-channel can appear. One of them is a well-known family of covert channels dating back to work on the CPU scheduling channel. The other scenario is interesting because it has three levels, is in a network environment, and is previously unknown. We view these scenarios as a serious threat. We discussed generalizations of the timed Z-channel and mentioned how complicated the mathematics would be to solve for their capacity. In addition to the fact that the timed Z-channel is present in real scenarios, it is also, in general, a good basis for an analytical study of noisy timing channels. We gave background on Golomb's classic work on the Z-channel. We then defined the timed Z-channel formally, and presented new results on its capacity. These results showed us the relationship between the noise and timing factors. The results also gave us a theorem bounding the capacity in terms of other well-known results. We also discovered that a very interesting mathematical artifact, the limiting behavior of the critical probability for the Z-channel, generalizes to the timed Z-channel.

We hope that our results will be of use to the designers of MLS systems. We feel that describing and thoroughly analyzing this threat is the first step in its management. We also feel that the serious threat of our timed Z-channel may not be restricted to just our example scenarios. In future work we plan on investigating other scenarios, as well as systems having output alphabets with more than two timed symbols.

7. Acknowledgments

The authors appreciate discussion and correspondence with Solomon Golomb, Jason Martin, John McLean, Bruce Montrose and the referees.

References

[1] G.S. Alijani and R.L. Morrisson. An evaluation of IEEE 802 protocols and FDDI in real-time distributed systems. In *Proceedings of the 15th Conference on Local Computer Networks*, pages 334–342, Los Alamitos, California, 30 September–3 October 1990.

[2] D.E. Bell and L.J. LaPadula. Secure computer system: Unified exposition and multics interpretation. Technical Report MTR-2997, The MITRE Corporation, Bedford, Massachusetts, March 1976. Available from the National Technical Information Service as report number: AD A023 588.

[3] O.L. Costich and I.S. Moskowitz. Analysis of a storage channel in the two phase commit protocol. In *Proceedings of the Computer Security Foundations Workshop IV*, pages 201–208, Franconia, New Hampshire, June 1991.

[4] R. David, S. Song, and R. Mukkamala. Supporting security requirements in multilevel real-time databases. In *Proceedings of the IEEE Symposium on Security and Privacy*, pages 199–210, Oakland, California, May 1995.

[5] D. Dodds. FDDI - a new standard for high speed fiber optic data networks. In *WESCANEX 88: Digital Communications Conference Proceedings*, pages 63–66, New York, New York, May 1988.

[6] S.W. Golomb. The limiting behavior of the Z-Channel. In *IEEE Transactions on Information Theory*, vol. 26, no. 3, page 372, May 1980.

[7] J.W. Gray III. On introducing noise into the Bus-Contention Channel. In *Proceedings of the IEEE Symposium on Research in Security and Privacy*, pages 90-98, Oakland, California, May 1993.

[8] W. Hu. Reducing timing channels with fuzzy time. In *Proceedings of the IEEE Symposium on Research in Security and Privacy*, pages 8–20, Oakland, California, May 1991.

[9] J.C. Huskamp. Covert Communication Channels in Timesharing Systems. PhD thesis, UC Berkeley, May 1978. Available as UCB-CS-78-02 and also as Electronics Research Laboratory Memo No. ERL-M78/37, UC Berkeley.

[10] J.V. Janeri, D.B. Darby, and D.D. Schnackenberg. Building Higher Resolution Clocks in Covert Timing Channels. In *Proceedings of The Computer Security Foundations Workshop VIII*, pages 85–95, Kenmare, County Kerry, Ireland, June 1995.

[11] R. Jain. FDDI: current issues and future plans. *IEEE Communications Magazine*, 31(9):98–105, September 1993.

[12] M.H. Kang and I.S. Moskowitz. A Pump for rapid, reliable, secure communication. In *Proceedings of the ACM Conference on Computer & Communication Security '93*, pages 119–129, Fairfax, Virginia, 1993.

[13] M.H. Kang and I.S. Moskowitz. A data Pump for communication. Naval Research Laboratory Memo report 5540-95-7771, 1995.

[14] M.H. Kang, I.S. Moskowitz, and D.C. Lee. A network version of the Pump. In *Proceedings of the IEEE Symposium in Security and Privacy*, pages 144–154, Oakland, California, May 1995.

[15] M.H. Kang and I.S. Moskowitz. A generalized version of the Pump. Preprint 1995.

[16] P.A. Karger A non-discretionary access control for decentralized computing systems. MIT/LCS/TR-179 MIT Lab for Computer Science, May 1977

[17] S.B. Lipner. A comment on the confinement problem. In *Proceedings of the 5th Symposium on Operating System Principles*, pages 192–196, University of Texas, Austin, November 1975.

[18] E.E. Majani. A model for the study of very noisy channels, and applications. PhD thesis, California Institute of Technology, Pasadena, California, 1988.

[19] E.E. Majani and H. Rumsey. Two results on binary-input discrete memoryless channels. In *Proceedings of the IEEE International Symposium on Information Theory*, page 104, Budapest, Hungary, June 1991.

[20] D. McCullogh. Covert channels and degrees of insecurity. In *Proceedings of the Computer Security Foundations Workshop I*, pages 1–33, Franconia, New Hampshire, June 1988.

[21] J.K. Millen. Finite-state noiseless covert channels. In *Proceedings of the Computer Security Foundations Workshop II*, pages 81–86, Franconia, New Hampshire, June 1989.

[22] A.R. Miller and I.S. Moskowitz. Reduction of a class of Fox-Wright Psi functions for certain rational parameters. In *Computers & Mathematics with Applications*, vol. 30, no. 11, November 1995.

[23] B.E. Montrose and M.H. Kang. An Implementation of the Pump: Event Driven Pump. Naval Research Laboratory Memo. Report 5540-95-7782, 1995.

[24] I.S. Moskowitz and A.R. Miller. Simple timing channels. In *Proceedings of the IEEE Symposium in Security and Privacy*, pages 56–64, Oakland, California, May 1994.

[25] M. Schaefer, B. Gold, R. Linde., and J. Scheid. Program confinement in KVM/370. In *Proceedings of the Annual Conference, ACM*, pages 404–410, Seattle, Washington, October 1977.

[26] C. Shannon and W. Weaver. The mathematical theory of communication. University of Illinois Press, 1949. Also appeared as a series of papers by Shannon in the Bell System Technical Journal, July 1948, October 1948 (A Mathematical Theory of Communication), January 1949 (Communication in the Presence of Noise).

[27] Technical Committee on Computer Communications of the IEEE Computer Society. *IEEE Standards for Local Area Networks: Token Ring Access Method and Physical Layer Specifications*. The Institute of Electrical and Electronics Engineers, Inc., New York, New York, September 1989.

[28] M. Teener and R. Gvozdanovic. FDDI-II operation and architectures. In *Proceedings of the 14th Conference on Local Computer Networks*, pages 49–61, Minneapolis, Minnesota, October 1989.

[29] Token-Passing Bus Access Method Working Group. *Information Processing System, Local Area Networks, Part 4: Token-passing bus access method and physical layer specifications*. The Institute of Electrical and Electronics Engineers, Inc., New York, New York, 1990.

[30] J.T. Trostle. Modeling a fuzzy time system. In *Proceedings of the IEEE Symposium on Research in Security and Privacy*, pages 82–89, Oakland, California, May 1993.

[31] S. Verdú. On channel capacity per unit cost. In *IEEE Transactions on Information Theory*, vol. 36, no. 5 pages 1019–1030, Sept. 1990.