REPRINT

Simple Timing Channels

Ira S. Moskowitz and *Allen R. Miller*

FROM:
Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy, Oakland, CA, May 16-18 1994, pages 56-64, IEEE Press.

CONTACT:
Ira S. Moskowitz, Information Technology Division, Mail Code 5543, Naval Research Laboratory, Washington, DC 20375.

E-MAIL:
moskowit@itd.nrl.navy.mil

COMMENTS:
As of June 23, 1994 this reprint has corrected a typographical error that occurred on page 59, column 1, line 6 of the proof of Corollary 1.1 in the actual published paper.
We have changed the bottom index of the second sum from a 0 to a 1 .

Simple Timing Channels

Ira S. Moskowitz

Information Technology Division
Naval Research Laboratory
Washington, DC 20375

Allen R. Miller

Department of Mathematics
George Washington University
Washington, DC 20052

Abstract

We discuss the different ways of defining channel capacity for certain types of illicit communication channels. We also correct some errors from the literature, offer new proofs of some historical results, and give bounds for channel capacity. Special function techniques are employed to express the results in closed form. We conclude with examples.

1 Introduction

Even the most securely designed computer systems may inadvertently contain covert (communication) channels between specific users/processes of different security levels. Such covert channels can thwart efforts to prevent higher level information from being accessible to a lower level. Specifically, as in [26], we consider a multi-user computer system, where there are two specific user/processes designated High and Low. We assume that Bell-LaPadula type security procedures [2] have been set up so that Low may not read High's files and High may not write to Low's files. However, it may be possible for High to pass information to Low over a covert channel that unintentionally exists in the system.

In this paper we are interested in a specific type of covert channel, a *timing channel*. A timing channel exists if it is possible for High to interfere with the system response time to an input by Low. Therefore, a timing channel is a communication channel where the output alphabet is constructed from different time values (see [30]). Timing channels with noise and/or memory have been studied by the security community, for example [22]. However, the thrust of this paper is the analysis of timing channels that are discrete, memoryless, and noiseless. We will call such a timing channel a *simple timing channel* (STC).

From a security viewpoint, the capacity of a covert channel is the standard metric with which to measure its potential damage. In fact, the value of the capacity leads to different levels of secure system certification [7]. However, STC's may crop up on their own, e.g. in the disk arm channel [8], or as in the recent paper by Mathur and Keefe [20]. Further, STC's can be used as capacity bounds for more complicated types of timing channels [13]; i.e. STC's may give a worst case scenario. STC's therefore warrant special attention.

Implicit in the study of timing channels is the assumption that Low always receives the same response; it is the *time* at which Low receives the response that forms the output alphabet. If Low receives different responses, all taking the same amount of time, then we are in the situation of a "storage channel" [17]. If Low receives different responses at different times, then the resulting covert channel is termed a *mixed channel*. We will examine mixed channels in future work.

EXAMPLE 1: Say that Low wishes to play Chess. Chess can have multiple users, the effect being that response time increases from 1ms to 2ms when there is more than one user. By High playing or not playing Chess while Low is playing, High can send a 2 symbol alphabet to Low. If there are other users besides High, and Low cannot distinguish them from High, then this transmission is noisy. In fact, if we look at capacity in units of bits/transmission, we simply have the Z-channel [10, 4] (a two symbol channel where one of the symbols is transmitted perfectly). However, the capacity in terms of bits/ms is more complicated [28].

As mentioned, an important measure of the potential damage of a STC is the (channel) capacity. In our studies of STC capacity, we noticed some inconsistencies in the definition of capacity [27]. We discuss this and also offer a novel and simple proof of one of the major theorems concerning the capacity of STC's (and certain communication channels in general), thus providing a firm theoretical foundation on which to base our covert channel analysis.

We give bounds for the capacities of STC's when exact closed form solutions are intractable or unnecessary. Often, for security, an upper bound on capacity will suffice. We have given an example of this in previous work [13]. However, when one institutes system modifications to lessen the capacity of STC's, performance tends to suffer [12, 13]. Therefore, the tighter we can make the capacity bounds, the better. We make a detailed study of both upper and lower bounds by examining the roots of trinomials. Finally, we apply our work to STC's from the database world.

Note 1 *The notation* log *will always mean the base 2 logarithm and* ln *will mean the natural logarithm. We assume that there is a way of measuring time, and the unit of time is a tick and all measurements are integral multiples of one tick.*

2 Asymptotic Definition of Capacity for a STC

In a STC, High (transmitter) has an input alphabet consisting of the symbols $\{s_1, ..., s_k\}$. Low (receiver) has the output alphabet $\{t_1, ..., t_k\}$, where each distinct integer t_j is the amount of time, in units of ticks, for the symbol s_j to be transmitted over the channel and $t_j < t_{j'}$ if $j < j'$. We say that the above STC has an alphabet of size k and use the notation $T(t_1, \ldots, t_k)$ to denote the above STC when we wish to be specific.

Only one response is being sent to Low, but High is able to vary the time t_j it takes for that response to arrive at Low. Since Low can distinguish between the different t_j values, this is equivalent to a discrete noiseless channel with k different symbols, each taking distinct times to be transmitted over the channel [27].

A transmission through the STC can be viewed as a sequence whose terms are s_j's. Since the STC is memoryless, the choice of symbols being sent is unconstrained; thus, all sequences are allowed. The length of the sequence is defined as the sum of the t_j's corresponding to the s_j's comprising the sequence. Thus, the length of the sequence is equal to the total transmission time of the sequence.

Definition 1 *With respect to a given STC, let S be the set of all sequences whose terms are from the set of symbols. Let S_n be the subset of S consisting of sequences whose length is n, $n \in \mathbb{Z}^+$, and let $|S_n|$ denote the cardinality of S_n. Also, if $s \in S$, we let $|s|$ denote the length of s.*

Since information is passed over the STC by sending different sequences of symbols, we see that the ratio $|S_n|/n$, as n gets large, gives a measure of the amount of information being sent [27]. This leads us to the following definition.

Definition 2 *(Krause) The capacity (C) of a STC is given by*

$$C = \limsup_{n\to\infty} \frac{\log |S_n|}{n} \, . \qquad (1)$$

The units of capacity are bits per tick. To specifically identify the capacity of $T(t_1, \ldots, t_k)$ we will use the notation $C_{T(t_1,\ldots,t_k)}$.

Shannon's original paper used the ordinary limit instead of the limit superior. The ordinary limit does not exist for many channels of interest. Our definition is a restatement of Krause's [15] definition of capacity (see also [29]). The following example shows the problem of using the ordinary limit.

EXAMPLE 2: Say that we only have two symbols s_1 and s_2, and that $t_1 = a$ and $t_2 = b$. Take any $s \in S_n$, where s consists of $c(i)$ terms of s_i. Hence $n = c(1)a + c(2)b$. We see that the greatest common divisor of a and b must also divide n. Therefore, a necessary (but not necessarily sufficient) condition for $S_n \neq 0$ is that n be a multiple of the greatest common divisor of a and b. For instance if $a = 2$ and $b = 4$, $|S_{2n+1}| = 0$ (even time values can never give a sequence of odd length). Therefore, the limit of $(\log |S_n|)/n$ is not always defined. We also see that in general $|S_n|$ cannot be asymptotic to λ^n, where λ is the positive root of an associated characteristic polynomial, as Shannon states [27].

Note 2 *Our/Krause's definition of capacity, Equation (1), is well-defined because if there are m symbols then $|S_n| \leq m^n$, and $(\log |S_n|)/n \leq \log m$. Hence, C is well-defined and bounded from above by $\log m$.*

Others have gotten around the problem with the ordinary limit versus the limit superior by slightly redefining $|S_n|$ so that it is non-decreasing [9, 5]. In fact, in [6], where the problem with the ordinary limit is also noted, Csiszár goes into a detailed analysis of different measures of $|S_n|$ leading to equivalent definitions of capacity.

Since log is an increasing function and $(\log |S_n|)/n = \log \sqrt[n]{|S_n|}$, we can also express the capacity as

$$C = \log \limsup_{n\to\infty} \sqrt[n]{|S_n|} \, . \qquad (2)$$

We extend the definition of $|S_n|$ to all integers by letting $|S_{-|n|}| = 0$, if $n \neq 0$, and defining $|S_0| = 1$. This extension makes sense because the empty sequence is the only sequence of length zero, and there are no sequences of negative length. Therefore, the $|S_n|$ satisfy the following recurrence relation

$$|S_n| = \sum_j |S_{n-t_j}| + \delta_{0n} \qquad (3)$$

where δ_{0n} is the Kronecker delta which is needed to make both sides of the equation equal to one when $n = 0$.

Now let us apply the z-transform [24] to both sides of Equation (3) and we arrive at the formal equations

$$\begin{aligned} \sum_{n=0}^{\infty} |S_n| z^n &= \sum_j \sum_{n=0}^{\infty} |S_{n-t_j}| z^n + \sum_{n=0}^{\infty} \delta_{0n} z^n \\ &= \sum_j z^{t_j} \sum_{n=0}^{\infty} |S_{n-t_j}| z^{n-t_j} + 1 \end{aligned}$$

$$= \sum_{t_j} z^{t_j} \sum_{n=0}^{\infty} |S_n| z^n + 1 .$$

These give us the formal equation

$$\sum_{n=0}^{\infty} |S_n| z^n = \frac{1}{1 - \sum_j z^{t_j}} ,$$

where $\frac{1}{1-\sum_j z^{t_j}}$ is referred to as the generating function [11] of the power series $\sum_{n=0}^{\infty} |S_n| z^n$. The above series manipulations and formal equations are valid in the disk about the origin where $\sum_{n=0}^{\infty} |S_n| z^n$ is analytic. Recall the root test for convergence of a power series:

Root Test — The power series $\sum_{n=0}^{\infty} a_n z^n$ converges absolutely for

$$|z| < \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|a_n|}}$$

and diverges if the inequality is reversed.

The number $1/\limsup_{n\to\infty} \sqrt[n]{|a_n|}$ is called the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n z^n$. We will show that the radius of convergence of $\sum_{n=0}^{\infty} |S_n| z^n$, denoted by R, is non-zero and, is in fact, equal to the (unique) real positive root of $1 - \sum_j z^{t_j}$.

Lemma 1 $R > 0$

PROOF: By note 2 we know that $1/m \leq 1/\limsup_{n\to\infty} \sqrt[n]{|S_n|}$; therefore, the power series converges when $|z| < 1/m$, so $R > 0$. □

Hence, there is a neighborhood about the origin in which $\sum_{n=0}^{\infty} |S_n| z^n$ is analytic. Since $\frac{1}{1-\sum_j z^{t_j}}$ is a rational function which is non-zero at the origin, we know that it too is analytic about the origin. By the uniqueness of power series representation for an analytic function [19, Thm. 3.2.5], the MacLaurin series of $\frac{1}{1-\sum_j z^{t_j}}$ must be $\sum_{n=0}^{\infty} |S_n| z^n$. Since the poles of $\frac{1}{1-\sum_j z^{t_j}}$ are precisely the roots of $1 - \sum_j z^{t_j}$, we see that the root(s) of smallest magnitude determine the largest disk about the origin in which $\frac{1}{1-\sum_j z^{t_j}}$ is analytic. Since, for a function of a complex variable, analyticity is the same as convergence of the power series, we see that the above smallest magnitude is exactly R.

Lemma 2 *The polynomial $1 - \sum_j z^{t_j}$ has one positive root r and any other root must have magnitude at least equal to $|r|$.*

Proof: Let the magnitude of the complex number z be denoted by $\zeta, \zeta \geq 0$

$$\begin{aligned} |1 - \sum_j z^{t_j}| &\geq 1 - |\sum_j z^{t_j}| \\ &\geq 1 - \sum_j |z|^{t_j} \\ &= 1 - \sum_j \zeta^{t_j} \end{aligned}$$

We will show that $h(\zeta) \equiv 1 - \sum_j \zeta^{t_j}$ has a unique positive root. Note that $h'(\zeta) < 0$ so $h(\zeta)$ is a decreasing function. This, along with the fact that $h(0) > 0$ and $\lim_{\zeta\to\infty} h(\zeta) < 0$, tells us that $h(\zeta)$ has a unique root r in $(0, \infty)$. In fact, since $\sum_j 1 \geq 2$ we see that $h(1) < 0$ so $r \in (0, 1)$. A root of $h(\zeta)$ is obviously a root of $1 - \sum_j z^{t_j}$. If ρ is any root of $1 - \sum_j z^{t_j}$ we have that $0 \geq 1 - \sum_j |\rho|^{t_j}$, with equality only for $|\rho| = r$ (since $h(\zeta)$ has a unique positive root); hence, $r \leq |\rho|$ since $h(\zeta)$ is a decreasing function. □

So $R = r$, but $R = 1/\limsup_{n\to\infty} \sqrt[n]{|S_n|}$. This tells us that

$$\limsup_{n\to\infty} \sqrt[n]{|S_n|} = r^{-1}$$

and $C = \log r^{-1}$, $1 < r^{-1}$. By noting that the inverse of the positive root of $1 - \sum z^{t_j}$ is the same as the positive root of $1 - \sum x^{-t_j}$, $x \in \mathbb{R}$, we arrive at the following theorem of Shannon. However, Shannon's sketched proof [27] is incomplete because it relies on a fact from the asymptotic behavior of finite-difference equations which, as we discussed earlier, does not apply.

Theorem 1 *The capacity of the STC $T(t_1, \ldots, t_k)$ is*

$$C = \log \omega$$

where $\omega > 1$ is the unique positive root of $1 - \sum_j x^{-t_j}$ (we may specifically identify ω as $\omega_{T(t_1,\ldots,t_k)}$).

Krause [15] was the first to give a rigorous proof of the above theorem. He obtained the result by using Dirichlet series instead of power series. However, the Dirichlet series approach is much more complicated than our proof. Kuich [16, Thms. 1,5] has done work similar to ours, but in relation to the entropy of context-free languages. Although Theorem 1 has appeared quite often in the literature, our proof is simpler and more direct than the previous proofs.

We also see that the problem of channel capacity is actually an algebraic problem. Due to the importance of the equation

$$1 - \sum_{j=1}^{k} x^{-t_j} = 0$$

we refer to it as the *characteristic equation* of the STC and will denote the *characteristic polynomial* $1 - \sum_j x^{-t_j}$ as $\chi(x)$. The characteristic equation may also be written as

$$x^{t_k} - (x^{t_k - t_1} + \cdots + 1) = 0 \, .$$

Corollary 1.1 The bounds $0 < C < 1$ are best possible.

Proof: First we will show that $C \in (0,1)$ and then that these bounds are tight. Since $\omega > 1$, it is trivial that $C > 0$. We wish to solve $\chi(x) = 0$, already knowing that the solution is in $(1, \infty)$. Since there are at least two output symbols $\chi(1) = 1 - \sum_j 1^{-t_j} < 0$. Since $\sum_j 2^{-t_j} < \sum_{i=1}^{\infty} 2^{-i} = 1$, we see that $\chi(2) > 0$. Therefore, ω must be in $(1,2)$ hence $C < 1$.

We can always make C as close to 0 as we wish by just choosing larger and larger values of t_j. For example, if the channel has two symbols and $t_1 = q$ and $t_2 = 2q$, the characteristic equation is $x^{2q} - x^q - 1 = 0$ which has positive root $(\frac{1+\sqrt{5}}{2})^{1/q}$. Therefore, $\omega \to 1$ and hence $C \to 0$, as $q \to \infty$.

To show that 1 is the *least* upper bound of C is a little trickier. Assume that the STC alphabet has n symbols and that $t_i = i$. Then the characteristic equation is $x^n - (x^{n-1} + x^{n-2} + \cdots + 1) = 0$. The solution in the interval $(1,2)$ is the same as the solution in the interval $(1,2)$ of $x^n - \frac{x^n - 1}{x-1} = 0$. Therefore, ω must obey the equation $\omega = 2 - \frac{1}{\omega^n}$, (also see [3])and because ω is bounded away from 1 we see that $\omega \to 2$ as $n \to \infty$ and hence $C \to 1$. $\square$

Note 3 The above can be extended to mixed channels, or, in general, to finite state discrete noiseless channels (see [22] for security applications), with the caveat being that the characteristic polynomial can have coefficients other than ± 1. Hence, the bounds on C will also change.

3 Average Mutual Information

Consider a discrete memoryless channel where X represents the input random variable with distribution $P(X = s_i) = p_i$, and Y represents the output random variable. Let $H(X)$ denote the entropy of X and $I(X,Y)$ the mutual information (in units of bits per transmission). The mutual information in units of bits per tick for a discrete memoryless channel is

$$I_t = \frac{I(X,Y)}{E(T)} \tag{4}$$

where $E(T)$ is the mean time for a symbol to be transmitted over the channel, see [25, 26, 28]. Of course, for a STC this reduces to

$$I_t = \frac{H(X)}{E(T)} \, .$$

(Since the channel is memoryless, the distribution on X is stationary; this corresponds to the unconstrained symbol condition mentioned in the previous section.) A rigorous study of Equation (4) for the memoryless channel in general has been given by Verdú [28]. In fact, he proves generalizations of the fact that the maximum value of I_t *is* the channel capacity. However, for a finite state discrete noiseless channel Shannon states and proves that the maximum value of I_t is the channel capacity in [27, appendix 4], see also [18]. Krause [15] gave a beautiful proof of the following theorem solely by relying on the inequality $\log x \le x - 1$.

Theorem 2 *(Shannon) For a STC,* $\max I_t = \log \omega$, *where* ω *is the positive root of the characteristic polynomial* $1 - \sum_j x^{-t_j}, x \ge 0$. *Furthermore, the distribution on* X *that achieves the maximum value is given by* $p_j = \omega^{-t_j}$.

Therefore, $\max I_t = C$ for the STC. Thus, we see two very different (the asymptotic and mutual information approaches), but equivalent, ways of defining capacity. Fully understanding the theory behind the definitions can assist in making the proper approximations necessary to construct models of covert channels [13]. For example, by using the mutual information definition of capacity, we can quickly see that $C \le \frac{\log n}{E(T)}$. A slight difference in the approximation of covert channel capacity can rapidly increase to a large error as the speed of a computer system increases. Therefore, we must be careful in the mathematical models of covert channels that we propose and analyze.

The mutual information definition of capacity is the proper way to look at channels with noise, see Equation (4). The asymptotic approach does not generalize. However, it is Shannon's [27] coding theorems that give the power to the definitions of capacity as the upper bound on errorless communication rates.

EXAMPLE 3: In [25, 26] we looked at a noisy timing channel where the first symbol arrived between 1 and 2 ticks and the second symbol arrived at 2 ticks. The noiseless version of this channel has the symbols arriving at 1 and 2 ticks, respectively, [22]. The characteristic polynomial of this channel is $1 - (x^{-2} + x^{-1})$. The positive root of $\chi(x)$ is $\frac{1+\sqrt{5}}{2}$, hence $C = \log \frac{1+\sqrt{5}}{2}$. In fact, by using Theorem 2, we see that the value of p that maximizes the mutual information, p_c, is given by $p_c = \left(\frac{1+\sqrt{5}}{2}\right)^{-1} = \frac{-1+\sqrt{5}}{2}$.

4 Bounds

A piece of software that is unintentionally inserted into a computer system and which is capable of exploiting a security flaw (such as a covert channel) is called a *Trojan horse*. We are concerned with the damage, in terms of capacity, that a Trojan horse might cause. Of course, we are not discussing the nature of the information being passed, only the rate at which it is being passed. Different exploitations of a security flaw can lead to different covert channels and hence to different capacities. We wish to study the different exploitations possible with a specific type of flaw. The flaw of concern is one that allows High to modulate Low response time.

For example, a simple exploitation allows High to modulate the response time by only one value. To be specific, let a be the smallest amount of time that it takes for Low to receive a response to a particular input. Therefore, if High does nothing, Low will receive its response after a time duration of a ticks. Let the smallest amount of time that High can add to this response time be d. Obviously, the Trojan horse wishes for d to be as small as possible to increase the capacity of the timing channel.

A simple exploitation by the Trojan horse will have High affect or not affect Low's response time. This may be the only exploitation available to the Trojan horse. This is the STC $T(a, a+d)$ with a 2 symbol input and output alphabet.

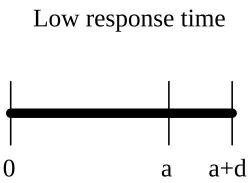

Figure 1: $T(a, a+d)$ Simple Exploitation

The characteristic equation of $T(a, a+d)$ is

$$x^{a+d} - x^d - 1 = 0 \ .$$

A more complex exploitation arises if High is able to delay the response to Low in multiples of d. Then the capacity of $T(a, a+d)$ is not a true measurement of the possible damage, in terms of capacity, that the Trojan horse can cause. For example, assume that High can delay the response to Low by $d, 2d, \ldots, (n-1)d$. Hence, this channel has an alphabet of n symbols and is just $T(a, a+d, \ldots, a+(n-1)d)$

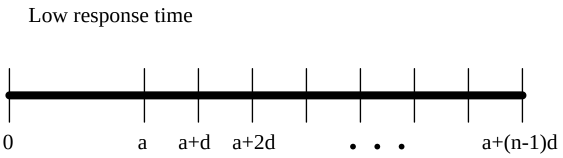

Figure 2: $T(a, a+d, \ldots, a+(n-1)d)$ Complex Exploitation

The characteristic equation of $T(a, a+d, \ldots, a+(n-1)d)$ is

$$1 - (x^{-a} + x^{-(a+d)} + \ldots + x^{-a-(n-1)d}) = 0 \ . \quad (5)$$

It is obvious that $C_{T(a,a+d,\ldots,a+(n-1)d)} \geq C_{T(a,a+d)}$, since any code for transmitting over $T(a, a+d)$ is also a code for $T(a, a+d, \ldots, a+(n-1)d)$. By studying the roots of the characteristic polynomial we see that $C_{T(a,a+d,\ldots,a+(n-1)d)}$ is in fact *strictly* greater than $T(a, a+d)$. Therefore, $C_{T(a,a+d)}$ is a lower bound for $C_{T(a,a+d,\ldots,a+(n-1)d)}$. Equation (5) may be written as (since $x > 1$, is the region of interest)

$$\begin{aligned} 1 - x^{-a}\left(\frac{1-x^{-nd}}{1-x^{-d}}\right) &= 0\ , \quad \text{or} \\ 1 - \frac{x^{-a}}{1-x^{-d}} + \left(\frac{x^{-a}}{1-x^{-d}}\right)x^{-nd} &= 0\ . \end{aligned}$$

As we increase n, the positive root of $\chi(x)$ also increases. This follows because $1 - \sum_{i=0}^{n-1} x^{-(a+id)} > 1 - \sum_{i=0}^{n'-1} x^{-(a+id)}$, if $n' > n$. Therefore, we know that when a and d are fixed and $n \geq 2$, that ω is bounded away from 1. By this we mean that there exists an $\epsilon > 0$ such that $\omega \in [1+\epsilon, 2)$. We let M denote the maximum value of $\frac{x^{-a}}{1-x^{-d}}$ on the closed interval $[1+\epsilon, 2]$. Since

$$1 - \frac{x^{-a}}{1-x^{-d}} < \chi(x) < 1 - \frac{x^{-a}}{1-x^{-d}} + Mx^{-dn}$$

and the functions are increasing for $x \in (1,2)$, we see that ω approaches, from the left, the root of $1 - \frac{x^{-a}}{1-x^{-d}}$, as $n \longrightarrow \infty$.

Therefore, we can (tightly) bound the capacity for this exploitation by investigating the positive solution of

$$1 - \frac{x^{-a}}{1-x^{-d}} = 0 \quad (6)$$

for $x \in (1, 2]$. This can be interpreted as the characteristic equation of a STC with infinitely many symbols, each one taking time $t_i = a + (i-1)d$, because this is what the limiting behavior of the Equation (5) is. We can rewrite Equation (6) as

$$1 - (x^{-a} + x^{-d}) = 0 \ .$$

Note 4 *For $a \neq d$, $1 - (x^{-a} + x^{-d})$ is $\chi(x)$ for the STC $T(a, d)$.*

This result is very useful because it says that:

Theorem 3 *The capacity for a STC where the smallest time is a and the time can be moderated by multiples of d is, for $a \neq d$, bound from above by the capacity of a channel with a 2 symbol alphabet where $t_1 = d$ and $t_2 = a$, and, for $a = d$, bound from above by $\log \sqrt[a]{2} = a^{-1}$.*

Even though this rate of information transfer may involve symbols whose duration is less than a, there is nothing mysterious going on because we are simply dealing with a huge number of symbols. For example, if we were to send 64 distinct symbols across a channel, each symbol taking 2 ticks, the rate of information transfer would be 3 bits per tick, even though each symbol takes 2 ticks to pass over the channel.

Thus, from Theorem 3 and our previous discussions we have

Corollary 3.1

$$\begin{array}{rcll} C_{T(a,a+d)} \leq & C_{T(a,a+d,\ldots,a+(n-1)d)} & \leq C_{T(a,d)} & , a \neq d \\ a^{-1}\log\dfrac{1+\sqrt{5}}{2} \leq & C_{T(a,2a,\ldots,na)} & \leq a^{-1} & , a = d\ . \end{array}$$

5 Trinomial Equations

Summarizing the main results from the last section, we see that we may bound the capacity of the STC $T(a, a+d, \ldots, a+nd)$ by $\log \omega_{T(a,d)}$ from above and $\log \omega_{T(a,a+d)}$ from below for $a \neq d$. Therefore we wish to obtain a closed form solution for the positive root $\omega_{T(a,d)}$ of $1-(x^{-a}+x^{-d})$ and the positive root $\omega_{T(a,a+d)}$ of $1-(x^{-a}+x^{-(a+d)})$.

We have three cases to consider.

Case 1: $a > d$
This arises when the Trojan horse can affect the response time by an amount of time less than the original response time. An example of this would be if a response to Low involves scanning an entire disk, and High is able to add small amounts of delay to the response.

Case 2: $a < d$
This arises by modulating the response time by values that are much bigger than the original response. This does not seem to be as likely an exploitation as Case 1.

Case 3: $a = d$
Here, the response time is locked into fixed multiples of a. Corollary 3.1 tells us the bounds. This case can itself be used as a worst case scenario by letting a be the quickest possible response time. Therefore, we are left with cases 1 and 2 to analyze.

We wish to simplify the (upper bound) polynomials by expressing, as before, the characteristic equation $1-(x^{-a}+x^{-d}) = 0$ using positive exponents. In case 1 this becomes $x^a - x^{a-d} - 1 = 0$ and in case 2 this becomes $x^d - x^{d-a} - 1 = 0$. Therefore, we will study trinomial equations of the form

$$x^N - x^{N-Q} - 1 = 0, \quad N > Q > 0 \tag{7}$$

with N and Q being either a or d, depending on whether we are in the upper bound part of cases 1 or 2, and N, Q are $a+d, a$ when we are looking at the lower bound for any of the cases.

6 Roots of the Trinomial

Although many authors have investigated the solutions of algebraic trinomial equation (see the extensive bibliography in Belardinelli [1]), the positive root of Equation (7) may be expressed elegantly by employing Mellin's result [21] and Wright's Psi function. Thus, for real k the positive root of the trinomial equation

$$y^N + ky^{N-Q} - 1 = 0\ , \qquad N > Q > 0 \tag{8}$$

is given by

$$y = {}_1\Psi_1^* \left[\begin{array}{ll} (\frac{1}{N}, \frac{N-Q}{N}) & ; \\ (1+\frac{1}{N}, \frac{-Q}{N}) & ; \end{array} -k \right] \tag{9}$$

where Q and N are real numbers such that

$$|k| < (Q/N)^{-Q/N}(1-Q/N)^{Q/N-1} \leq 2\ . \tag{10}$$

Wright's Psi function in Equation (9) is defined by the series representation

$${}_1\Psi_1^* \left[\begin{array}{l} (\alpha, A); \\ (\beta, B); \end{array} z \right] = \frac{\Gamma(\beta)}{\Gamma(\alpha)} \sum_{n=0}^{\infty} \frac{\Gamma(\alpha + An)}{\Gamma(\beta + Bn)} \frac{z^n}{n!}\ .$$

Miller and Moskowitz have shown in [23] that the Wright function ${}_1\Psi_1^*[z]$ may be expressed in various ways as a finite sum of generalized Gaussian hypergeometric functions when A and B are rational numbers. Hence, setting $k = -1$ in Equation (8) and verifying that the inequality (10) holds, the positive root ω of Equation (7) for integers $N > Q \geq 1$, may be written in the three ways show in figure 3.

The above closed form solutions enable us to tightly bound capacity. The closed forms are useful from both a numerical and theoretical standpoint. Since they are in closed form, a numerical answer can be easily calculated. One of their uses is in seeing how sensitive/robust the bounds are to slight perturbations in channel exploitation. This is especially important in light of the exponential growth of processor speeds. In other words, a slight difference in capacity on this year's machine might be a severe problem with next year's faster machine.

Also, the closed form solutions let us see the exact functional relation between the capacity and various channels parameters. This is apparent by examing the closed forms presented in this section. One could glean some information through numerical methods, but never as much as through closed form expressions.

$$\omega = 1 + \frac{1}{N}\sum_{r=1}^{N} \frac{\Gamma(\frac{1}{N}+\frac{N-Q}{N}r)}{\Gamma(1+\frac{1}{N}-\frac{Q}{N}r)r!}$$

$$\cdot {}_{N+1}F_N\left[\begin{array}{l} 1,\mu(r),\ldots,\mu(r)+\frac{N-Q-1}{N-Q},\nu(r),\ldots,\nu(r)+\frac{Q-1}{Q} \\ \frac{r+1}{N},\ldots,\frac{r+N}{N} \end{array} \begin{array}{l} ; \\ ; \end{array} \xi\right]$$

where

$$\mu(r) = \frac{r}{N}+\frac{1}{N(N-Q)}\ , \quad \nu(r) = \frac{r}{N} - \frac{1}{NQ}$$

and

$$\xi = (-Q)^Q(N-Q)^{N-Q}(-x/N)^N\ ;$$

$$\omega^{-1} = 1 + \frac{1}{N}\sum_{r=1}^{N} \frac{\Gamma(\frac{1}{N}+\frac{Q}{N}r)}{\Gamma(1+\frac{1}{N}-\frac{N-Q}{N}r)}\ \frac{(-1)^r}{r!}$$

$$\cdot {}_{N+1}F_N\left[\begin{array}{l} 1,\mu(r),\ldots,\mu(r)+\frac{Q-1}{Q},\nu(r),\ldots,\nu(r)+\frac{N-Q-1}{N-Q} \\ \frac{r+1}{N},\ldots,\frac{r+N}{N} \end{array} \begin{array}{l} ; \\ ; \end{array} \xi\right]$$

where

$$\mu(r) = \frac{r}{N}+\frac{1}{NQ}\ , \quad \nu(r) = \frac{r}{N} - \frac{1}{N(N-Q)}$$

and

$$\xi = (-Q)^Q(N-Q)^{N-Q}(x/N)^N\ ;$$

$$\omega^{N-Q} = \frac{\delta}{N} + (-1)^{N-Q}\left(\frac{N-Q}{N}\right)\sum_{r=1}^{N-1}\frac{(1+Q-N+\frac{N-Q}{N}r)_{N-Q-1}(\frac{Q-N}{N}r)_{r-1}}{(\frac{Q-N}{N}r)_{N-Q}}\ \frac{(-1)^{r-1}}{(r-1)!}$$

$$\cdot {}_NF_{N-1}\left[\begin{array}{l} 1,\frac{r}{N}-\frac{1}{Q},\frac{r}{N},\ldots,\frac{r}{N}+\frac{Q-2}{Q},\frac{r}{N}+\frac{1}{N-Q},\ldots,\frac{r}{N}+\frac{N-Q-1}{N-Q} \\ \frac{r+1}{N},\ldots,\frac{r+N-1}{N} \end{array} \begin{array}{l} ; \\ ; \end{array} \xi\right]$$

where

$$\delta = \begin{cases} 0 & Q>1 \\ 1 & Q=1 \end{cases}\ , \qquad \xi = Q^Q(Q-N)^{N-Q}(-1/N)^N\ .$$

Figure 3: Different Expressions for the Root of Equation (7)

7 Database STC's

A database is updated when a user successfully enters new information into the database. Covert channels may arise in a database if High can influence Low's ability to enter new information and/or the time at which Low gets a "receipt" of its update. In [13], High is able to influence the time at which Low receives an acknowledgement of its update by removing or not removing messages from an intermediary communications buffer. This results in a noisy timing channel. Kang and Moskowitz showed how STC's could be used as a worst case analysis for these more complicated timing channels. The study of these bounds is documented in [13].

A more traditional and fundamental approach to database security involves the study of update transactions. Recently, Mathur and Keefe [20] discussed specific covert channels that arise in a database scheduler implementing certain concurrency control and recovery protocols. These covert channels, which are in fact STC's, arise from the desire to ensure atomicity. (Atomicity means that a transaction either happens or does not happen, it does not partially happen.)

Mathur and Keefe identify and analyze in detail three specific types of STC's arising from High manipulating subtransactions: the delayed sibling subtransaction STC; the delayed reader subtransaction STC; and the compensation STC. We will examine only the delayed reader subtransaction STC here. Mathur and Keefe argue that their channels suffice for a worst case analysis. In their statement immediately preceding section 4.1 of [20] they assert:

> *However, if we use exactly two symbols in the channel and ensure that the durations of these symbols is the smallest possible, i.e., there exists no covert channel scenario which requires fewer operations to transmit a symbol, then indeed we have achieved the maximum bandwidth possible in the channel. It turns out that this is possible to establish for each of the above scheduling schemes.*

Presumably they are limiting themselves to channels with two symbols. If, in fact, High can send more than two symbols the capacity increases. (Keefe and Mathur have noted this problem in later work [14]) Our Corollary 3.1 shows the limits of this increase. Let us examine the Mathur and Keefe delayed reader subtransaction STC in detail.

The delayed reader subtransaction STC:
In this series of subtransactions Low writes a data item, commits the update, reads the update, and records the the time it waited. All four of these actions take a standard amount of time t_{op}. However, it is possible for High to delay the Low commit by one unit of t_{op}. Thus, we have a STC with $t_1 = 4t_{op}$ and $t_2 = 5t_{op}$. Of course t_{op} is a certain fixed number of ticks. We assume that one tick is one ms for the sake of comparison with Mathur and Keefe. Therefore, t_{op} is γ ms, where γ depends on the subtransaction speed of the database. The characteristic equation for this STC is $1-(x^{-4t_{op}}+x^{-5t_{op}}) = 0$. Let $\xi = x^{t_{op}}$, then $C = \frac{1}{t_{op}} \log \eta$ bits/ms, where η is the positive root of $\xi^5 - \xi - 1$. This gives us

$$C = \frac{1}{\gamma}\, 223.18 \text{ bits/second}$$

by using both Mathematica and MathCad. Note that Mathur and Keefe express their values in tabular rather than functional form and they get 217.59 bits/second instead of our 223.18 bits/second when $\gamma = 1$. We attribute this to round-off errors.

In our terminology $a = 4$ and $d = 1$, so we can upper bound the capacity by $\frac{1}{\gamma} C_{T(4t_{op},t_{op})} = \frac{1}{\gamma}\, 464.96$ bits/second. Therefore, we see that using just a two symbol input alphabet is far from a worst case scenario. If High is able to send symbols taking 4,5,6,...,n units of t_{op}, with n large, the above upper bound must be considered as the true worst case exploitation by the Trojan horse.

8 Summary

In this paper we have analyzed the twofold importance of STC's, both in and of themselves and as bounds for more complicated types of timing channels. We have shown how STC's themselves can be bound by two symbol STC's, the capacity of which we found by studying the roots of trinomials. We offer closed form solutions for these trinomials. We have presented different ways of defining capacity and have cleared up Shannon's original definition of capacity. Further, we have presented a new proof of a fundamental result from information theory. We conclude this paper by applying our analysis and bounding results to a previously studied STC.

References

[1] Giuseppe Belardinelli. *Fonctions Hypergéométriques de Plusieurs Variables et Résolution Analytique des Équations Algébriques Générales.* Gauthier-Villars, Paris, 1960. Fascicule 145.

[2] D.E. Bell and L.J. La Padula. *Secure Computer System: Unified Exposition and Multics Interpretation, MTR-2997.* MITRE Corp., Bedford, MA, March 1976.

[3] Leon Brillouin. *Science and Information Theory.* Academic Press, New York, NY, 2nd edition, 1962.

[4] Oliver L. Costich and Ira S. Moskowitz. Analysis of a storage channel in the two-phase commit protocol. In *Proc. of The Computer Security Foundations Workshop 4*, pages 201–208, Franconia, NH, June 1991.

[5] Imre Csiszár. Simple proofs of some theorems on noiseless channels. *Information and Control*, 14:285–298, 1969.

[6] Imre Csiszár. Noiseless channels. *Problemy Peredachi Informatsii*, 6:3–15, October-December 1970. (translation available).

[7] Department of Defense, National Computer Security Center. *Trusted Computer System Evaluation Criteria 5200.28-STD*, December 1985.

[8] B.D. Gold, R.R. Linde, R.J. Peeler, M. Schaefer, J.F. Scheid, and P.D. Ward. A security retrofit of vm/370. In *AFIPS Conference Proceedings, 1979 National Computer Conference*, volume 48, pages 335–344, Montvale, NJ, 1979.

[9] Stanford Goldman. *Information Theory.* Prentice-Hall, NY, 1953.

[10] Solomon W. Golomb. The limiting behavior of the Z-channel. *IEEE Transactions on Information Theory*, page 372, May 1980.

[11] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics.* Addison-Wesley, Reading, MA, 1989.

[12] Wei-Ming Hu. Reducing timing channels with fuzzy time. In *Proc. of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy*, pages 8–20, Oakland, CA, 1991.

[13] Myong H. Kang and Ira S. Moskowitz. A pump for rapid, reliable, secure communication. In *Proc. of the 1st ACM Conference on Computer and Communications Security*, pages 119–129, Fairfax, VA, November 1993.

[14] T.F. Keefe and A.G. Mathur. The concurrency control and recovery problem for multilevel update transactions in multilevel secure database systems. 1993 Preprint.

[15] Ralph M. Krause. Channels which transmit letters of unequal duration. *Information and Control*, (5):13–24, 1962.

[16] Werner Kuich. On the entropy of context-free languages. *Information and Control*, 16:173–200, 1970.

[17] Butler W. Lampson. A note on the confinement problem. *Communications of the ACM*, 16(10), October 1973.

[18] Richard S. Marcus. Discrete Noiseless Coding. Master's thesis, Massachusetts Institute of Technology, 1957. Department of Electrical Engineering.

[19] Jerrold E. Marsden and Michael J. Hoffman. *Basic Complex Analysis.* W.H. Freeman, New York, 2nd edition, 1987.

[20] Amit G. Mathur and Thomas F. Keefe. The concurrency control and recovery problem for multilevel update transactions in MLS systems. In *Proc. of the Workshop on Computer Security Foundations VI*, pages 10–23, Franconia, NH, June 1993.

[21] Hj. Mellin. Zur theorie der trinomischengleichungen. *Ann. Ac. Sc. Fenn.*, 7(7), 1915.

[22] Jonathan K. Millen. Finite-state noiseless covert channels. In *Proc. of The Computer Security Foundations Workshop II*, pages 81–86, Franconia, NH, June 1989.

[23] Allen R. Miller and Ira S. Moskowitz. Reduction of a class of Fox-Wright Psi functions for certain rational parameters. 1993 Preprint.

[24] Michael K. Molloy. *Fundamentals of Performance Modeling.* Macmillan, New York, 1989.

[25] Ira S. Moskowitz. Variable noise effects upon a simple timing channel. In *Proc. of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy*, pages 362–372, Oakland, CA, May 1991.

[26] Ira S. Moskowitz and Allen R. Miller. The channel capacity of a certain noisy timing channel. *IEEE Transactions on Information Theory*, 38(4):1339–1344, July 1992.

[27] Claude E. Shannon and Warren Weaver. *The Mathematical Theory of Communication.* University of Illinois Press, Urbana, IL, 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).

[28] Sergio Verdú. On channel capacity per unit cost. *IEEE Transactions on Information Theory*, 36(5):1019–1030, September 1990.

[29] J. Todd Wittbold. Controlled signalling systems and covert channels. In *Proc. of The Computer Security Foundations Workshop II*, pages 87–104, Franconia, NH, June 1989.

[30] John C. Wray. An analysis of covert timing channels. In *Proc. of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy*, pages 2–7, Oakland, CA, May 1991.