NRL Release Number 05-1226-0842

Multiple Access Covert Channels*

Ira S. Moskowitz
Center for High Assurance Computer Systems
Naval Research Laboratory
Washington, DC 20375
moskowitz@itd.nrl.navy.mil

Richard E. Newman
CISE Department
University of Florida
Gainesville, FL 32611-6120
nemo@cise.ufl.edu

Abstract

In this paper we consider the situation of multiple malicious transmitters attempting to covertly communicate with a single receiver. We show how the situation of non-collaborating transmitters can be modeled by multiple access channels. The simpler situation of collaborating transmitters is used as a bounding result. We also discuss the surprising results of Gaarder and Wolf that feedback can increase capacity, unlike the situation for standard covert channel analysis. This is of importance when dealing with the network scenario.
KEYWORDS: covert channel, multiple access channel, feedback, anonymity, Mix

1 Introduction

Classically, covert channel analysis has concerned itself with the situation of one transmitter and one receiver. The only exception that we can find for this in the literature is that of the Network Pump™ [8]. However, in [8], even though the situation is brought to light, it is not analyzed. The situation of multiple transmitters attempting to communicate covertly with one receiver also comes up when dealing with anonymity systems [16]. Recent work [13, 14, 15] discusses how *quasi-anonymous channels* arise in anonymity systems as covert channels that exist due to the lack of perfect anonymity. The quasi-anonymous channels considered though only deal with a single transmitter and a single receiver. In this paper we consider multiple covert channel transmitters.

For the sake of simplicity in this paper we assume that all channels are discrete and memoryless (with stationary distributions). The mathematical foundations for this paper, *multiple access channels*, were first hinted at in [19], and then put on firm ground in [1]. The definitive explanation can be found in [4].

1.1 Anonymity Example

In [13, 14] the situation of senders communicating with their recipients[1] from one private enclave to another is considered. Each enclave is protected by a Mix-firewall. The Mix-firewall hides the sender/recipient pairing. As in [13] we assume that every time unit t (tick) a sender either sends or does not send a single message from Enclave_1 to Enclave_2.

Eve is tapping the line between the enclaves. Eve can count the number of messages per t that go from Enclave_1 to Enclave_2, and Eve also knows how many possible senders there are in Enclave_1. We assume that there is a malicious sender Alice in Enclave_1 who wishes to communicate covertly with Eve. By Alice sending, or not sending a message, each t, Alice affects the message count of Eve. This covert channel is the quasi-anonymous channel in this anonymity system (see Fig. 1). Alice is the transmitter and Eve is the receiver in the quasi-anonymous channel.

The other senders in Enclave_1 act in a clueless manner (hence their names as $\text{Clueless}_i, i = 1, \ldots N$), that is, they act independently of Alice and they act independently of each other in an identical manner as i.i.d. Bernoulli random variables where p is the probability that they send a message from Enclave_1 to

*Research supported by the Office of Naval Research.

[1] We use the terms transmitters and receivers when discussing Shannon communication channels. We use the terms senders and recipients when discussing other type communication. This is done to avoid confusion between the receiver in a covert channel and the recipient in an anonymity network.

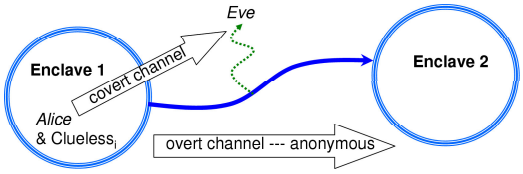

Figure 1: One Transmitter

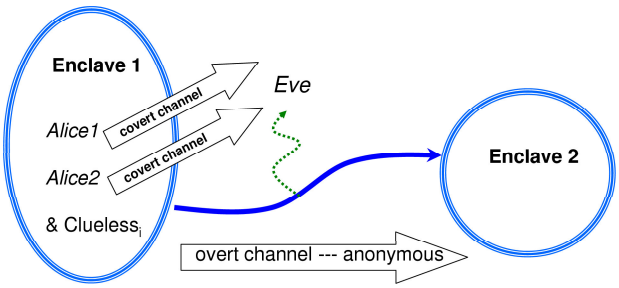

Figure 2: Two Transmitters — Anonymity Example

Enclave$_2$. The obvious fact that the capacity decreases to zero as N increases was illustrated in [13, 14], and rigorously proved in [11]. It is worth noting that the rigorous proof involved rather sophisticated results concerning the asymptotic behavior of the differences of divergent series. This does not bode well for more complex covert channel models of anonymity systems.

Example 1 — the anonymity example: Now we assume that instead of one malicious Alice there are in fact *two* malicious Alices. Furthermore, we assume that the Alices do not collaborate with each other. This may come about because the Alices may be in sub-enclaves within Enclave$_1$, or that any communication from Alice$_1$ to Alice$_2$ would arouse suspicion. Each Alice$_i$ still wishes to covertly communicate with Eve using the quasi-anonymous channel between each Alice$_i$ and Eve. The difficulty is that since they are not collaborating, they act as noise with respect to each other and may lessen the communication. Let us make these thoughts more precise.

There exists a quasi-anonymous channel from Alice$_1$ to Eve, and another quasi-anonymous channel from Alice$_2$ to Eve. Each quasi-anonymous channel is a covert channel because the Mix-firewall ideally should stop any such communication between Alice$_i$ and Eve. This above system of two covert channels is *Example 1, the anonymity example* (see Fig. 2). Note that the anonymity example involves storage channels.

1.2 NRL Network Pump^{TM}

In [8] the Network Pump^{TM} (see Fig. 3) was discussed as a solution to a secure, reliable, pragmatic, and robust method of sending messages up from several "Lows" to several "Highs." When a Low sends to a High, message acknowledgments, or ACKs, are required for reliability. Unfortunately ACKs can be used to send information from High to Low, which is against our wishes (Low can "talk" to High, but High should not be able to "talk" to Low in order to prevent High information leakage). Even if the ACKs are stripped down, the timing of the ACKs forms the basis of a covert timing channel from a High to a Low. The Network Pump^{TM} moderates the timing of the ACKs to moderate (but not eliminate entirely) the covert channel threat, while at the same time not degrading system performance in an intolerable manner. The interested reader is directed to the literature for more details on the Pump idea. Keep in mind that the covert channels that pertain to the Pump are *timing channels*. The thrust of this paper is on the easier to analyze storage channels.

In the Network Pump^{TM} each Low, L_i may send to any High, H_j. With respect to covert channels, in [8] it was assumed that the Highs were not collaborating and the covert channel analysis looked at each covert channel from H_j to L_i separately. *A fortiori* it was implicit that there there was no pre-arranged agreement between the H_js. This is important because there was no attempt of multiple H_js to communicate to a single L_i

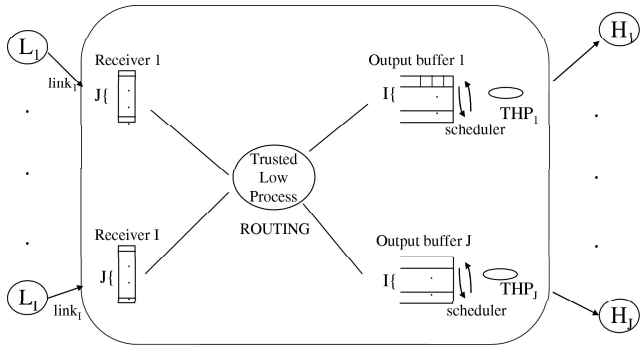

Figure 3: Network Pump^{TM} internals

Now, we wish to consider what happens when this is not the case. This is *Example 2, the Pump example.* In this paper we wish to consider H_1 and H_2 each attempting to communicate covertly with a specific Low. Thus, we have simplified matters by assuming that there is only one Low and there are two Highs.

This forms the basis of of the Pump example. We see that the Pump example and the anonymity example have a similar mathematical basis. This is the gist of what we wish to explore in this paper. Note that the Pump example is based upon timing channels. Therefore, we will not be able to use the capacity results that exist in the literature for the Pump example. Therefore, a full analysis of the Pump example is put off until we, or others, develop a theory of multiple access communication channels where the symbols take different amounts of time. However, we have included a discussion of the Pump for motivation for the simpler cases discussed here, and because of the importance of the Pump itself.

2 Multiple Access Channels

We will use a heuristic definition of a multiple access channel (MuAC) in this paper. For more details we point the interested reader to [5, Eq. 4], [10, Sec. III.A], [4, Sec. 14.3], or [7, Sec. II]. Note that, except when considering the Pump example, all symbols take the same time to transmit from input to output. Hence, time is not a consideration; thus the units of rate, mutual information, and capacity are in bits per symbol (this will be assumed and not written out each time). Therefore, with respect to covert channels, we are dealing with (covert) storage channels [9], not (covert) timing channels [12]. Of course the Pump example is a timing channel. As noted we include the discussion of the Pump example as motivation for studying multiple access channels and for showing the need for more theory in this area.

We emphasize that the two Highs are not collaborating once transmission begins. However they may have knowledge of each other's probabilistic behavior, which does not change over time. In fact they may agree *a priori* upon a protocol and coding strategy before beginning their transmissions. This is necessary so that the transmissions can assist each other in the passage of covert information, rather than hindering it. This fact does not seem to be well thrashed out in the information theory literature. However, once transmission begins the two transmitters share no further information[2], in fact they do not even know what each other is transmitting. That they are aware of each other's existence and have a static plan for shared transmissions we refer to as the *existence assumption.*

2.1 Review of Shannon channel

Recall that in a discrete and memoryless communication channel à la Shannon [17] we have one input modeled by the transmitter random variable X, taking on values x_i, and one output modeled by the receiver random variable Y, taking on the values y_j. The probability transition channel matrix determines the noise relationships in the channel. The (i, j) entry of the probability transition channel matrix

[2]This assumption is relaxed in the feedback section.

is the transition (conditional) probability $P(Y = y_j|X = x_i) = p(y_j|x_i)$. From the distributions of X, Y, and the conditional distribution of the $p(y_j|x_i)$ we can determine[3] the mutual information[4] to $I(X;Y) = I(Y;X)$ as

$$I(X;Y) = H(X) - H(X|Y),$$

where

$$H(X) = -\sum_i p(x_i) \log p(x_i)$$

is the entropy of X and

$$H(X|Y) = -\sum_{j,i} p(y_j)p(x_i|y_j) \log p(x_i|y_j)$$

is the conditional entropy. The capacity C is the maximum rate at which we can send information across the channel from the transmitter to the receiver with asymptotically small probability of error. Rates less than or equal to C are considered achievable, rates higher than C will probabilistically have non-trivial error. Shannon has shown that

$$C = \max_X I(X,Y)$$

(Note in the maximization process the possible non-trivial values of X are fixed at x_i, but the probabilities $p(x_i)$ vary.)

We intentionally redundantly state this as: we may transmit reliably (with asymptotically zero probabilistic error), through proper coding, at rates R that satisfy

$$0 \leq R \leq C\ .$$

Such rates are said to be achievable. The reason for this restatement will become clear in the next subsection. The interval $[0, C]$ is taken[5] as the *capacity region*; some call it the achievable rate region [2].

2.2 Multiple Access Channel Model

A *multiple access channel* (MuAC) has multiple inputs modeled by X_i, corresponding to multiple transmitters, and a single output (the receiver) modeled by Y. For the sake of simplicity let us assume throughout this paper that there are only two inputs, X_1 (taking on discrete values x_{1_i}) and X_2 (taking on discret values x_{2_j}), which transmit to Y (taking on discrete values y_k). In a MuAC the two inputs are not collaborating with each other. However, this is not to say that the two inputs do not have knowledge of each other's existence or overall probabilistic behavior. They do have knowledge of each others probabilistic behavior and they agree on a coding strategy/protocol before starting their transmission. This is the existence assumption that we mentioned earlier. Recall though, once they start their transmissions they act independently and with no further knowledge of each other. This point is often overlooked in the network information literature (e.g. [4]). We refer to this shared knowledge and protocol prior to transmission as their *a priori* knowledge.

Hence X_1 and X_2 each transmit independently and separately to Y, but with their *a priori* knowledge [10]. So there are two discrete memoryless channels: CH_1 which is the channel from X_1 to Y, and CH_2 which is the channel from X_2 to Y. Each channel CH_i has capacity C_i. Each X_i may transmit, with the proper coding, at some rate $R_i \leq C_i$ (these are the achievable rates).

The interesting question is what happens to transmission rates when both channels are in use together? Do they help each other, hurt each other, or have no effect upon each other? To answer this we must generalize the idea of the probability transition channel matrix to include a third dimension. Therefore, when dealing with MuACs we consider the transition probability $p(y_k|x_{1_i}, x_{2_j})$. The transition probabilities determine the noise in the channel.

[3] All logarithms are base two.

[4] We use the semi-colon ";" to represent the mutual information between two random variables, and use the comma "," represent a joint distribution between two random variables. Note that sometimes the comma is notationally used to represent mutual information between two random variables, which would cause confusion with the joint distribution here.

[5] For the standard Shannon channel with one input and one output this terminology is usually not employed. It is usually reversed for the multiple input situation discussed below. However, we feel that it makes sense to include the "one dimensional" situation as a special case.

The model of the MuAC assumes that one has knowledge of the distribution of the X_i and the transition probabilities. Let R_i be the rate for a code for CH_i. One may send information across both channels using a separate code for each channel. Each channel has its own rate. However, we may consider the two codes and the two rates as a 2-tuple and analyze the average joint error across both channels. If the error is asymptotically negligible (as for the Shannon channel) then the *rate pair* (R_1, R_2) *is* said to be *achievable*[6] for the MuAC [4]. Following [4], we define the *capacity region* for the MuAC as the closure of the set of achievable rate pairs.

In section 2.1 we defined the mutual information between two discrete random variables $I(X;Y)$. First though we need to expand the definition of entropy and conditional entropy for discrete random variables $A_1, \ldots, A_n, B_1 \ldots B_m$, following [3] as:

$$H(A_1, \ldots, A_n) = -\sum_{a_1,\ldots,a_n} p(a_1, \ldots, a_n) \log p(a_1, \ldots, a_n)$$

$$H(B_1, \ldots, B_m | A_1, \ldots A_n) = -\sum_{a_1,\ldots,a_n,b_1,\ldots,b_m} p(a_1, \ldots, a_n, b_1, \ldots, b_m) \log p(b_1, \ldots, b_m | a_1, \ldots, a_n) \ .$$

We next generalize the definition of mutual information for discrete random variables A, B, C (see [4, Sec. 2.5])

$$I(A; B|C) = H(A|C) - H(A|B, C)$$

and

$$I(A, B; C) = H(A, B) - H(A, B|C) \ .$$

Given a set of points Γ, the smallest convex set that contains those points is called the *convex hull* of Γ. This term is well-known in the field of computational geometry. With all the above we are now ready for the main mathematical underpinnings of this paper. In [4, Th.14.3.1] it is shown that that

Theorem 1 *The capacity region for a MuAC is the convex hull of the set of rate pairs* (R_1, R_2) *that satisfy:*

$$0 \leq R_1 \leq I(X_1; Y|X_2), \textit{ and} \tag{1}$$

$$0 \leq R_2 \leq I(X_2; Y|X_1), \textit{ and} \tag{2}$$

$$0 \leq R_1 + R_2 \leq I(X_1, X_2; Y). \tag{3}$$

We see that our capacity region is now (unlike for the Shannon channel) something geometrically of interest. If we attempt to (maximally) transmit at capacity across each channel we will most likely run into trouble, and introduce error, because of the third condition above: $0 \leq R_1 + R_2 \leq I(X_1, X_2; Y)$. This third condition is where the "action is." It describes how the two channels interfere with each other in the quest for a large achievable rate. The reason one uses the convex hull determined by Eqs. (1), (2), and (3) is that a timesharing process is used to send across each channel. The details of course are in the proof [4].

Definition 1 *A covert channel that is modeled by a MuAC is said to be a multiple access covert channel (MuACC).*

By our previous discussion such covert channels must be storage channels. (Of course we need a theory for dealing with multiple timing type channels, as in the Pump example.) We feel that it is important to introduce and to study MuACCs. The area of covert channel analysis has not touched on MuACs before. We will show that MuACCs introduce another dimension to the field of high-assurance computing which must be taken into account when analyzing the security of systems.

[6] When dealing with the Shannon channel, the capacity forms an upper bound for rates whose maximum probability of error approaches zero. It can be shown that using average probability of error suffices [3, Lemma 3.5.3]. However, for MuACs the error of the codes that give us the rate pairs is only considered to be average error. It seems to be unknown if the capacity region forms bounds for codes with rate pairs whose maximum error goes to zero.

2.3 Anonymity Example revisited

Consider the anonymity example where there are two Alices and no Clueless senders. So, we have a covert channel from Alice_1 to Eve and a covert channel from Alice_2 to Eve, by assumption the Alice_i do not collaborate with each other. We see that this is a MuACC since both Alice_1 and Alice_2 are attempting to covertly communicate with Eve. What is the capacity region for this MuACC?

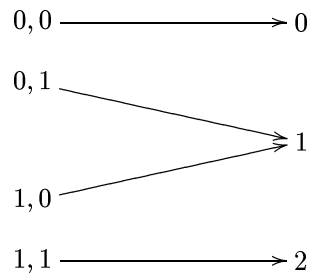

Figure 4: Channel transition diagram

First, let us consider each channel separately. Assume that there is only Alice_1. Alice_1 either sends or does not send a message from Enclave_1 to a recipient in Enclave_2. Eve can only count messages going from Enclave_1 to Enclave_2. Therefore, Eve either receives a 0 or a 1. The capacity of this (not so) covert channel is 1. Now, what happens when we also have an Alice_2 (there are still no clueless users)? The actions of Alice_2 function as noise for CH_1, the covert channel between Alice_1 and Eve. As shown in [13], the capacity across CH_1 varies from 1 (no noise) down to 1/2 (maximum noise) when there is one other transmitter acting in as a Bernoulli random variable with parameter p. The situation of maximum noise corresponds to $p = 1/2$, and capacity is 1, when $p = 0$ or $p = 1$.

We now continue with our study of the MuACC. We represent the possible inputs to the MuACC as a 2-tuple. That is (a,b) means that Alice_1 inputs a, while Alice_2 inputs b. If the dual input is (0,0), Eve receives a 0. Eve then knows that both Alices input a 0 and there is no noise. The same holds if the dual input is (1,1), Eve receives the message count of 2, and knows that both Alices input a 2. The noise comes in when the input is either (0,1) or (1,0). In this situation Eve receives a 1, and only knows that one Alice input a 1, and another Alice input a 0, but Eve does not know which Alice did what. However, we see that if Alice_1 is content to always transmit a 0 (achieving a throughput rate of 0 on Channel_1) then Alice_2 can transmit at any rate up to 1 on Channel_2, and visa versa. These facts correspond to the left and bottom boundaries, respectively, in Fig. 5. These can also be taken as the boundary values in Eqs. (1) and (2).

The more interesting question is what happens when both Alices are acting in a non-trivial manner? (Note that our analysis follows directly from [4, Ex. 14.3.3].) Assume that Alice_1 is maximally transmitting across Channel_1 to Eve, that is Channel_1 has a capacity of 1. In this situation Alice_1 is sending 0s and 1s with equal probabilities of 1/2. In this situation Channel_2 is a binary erasure channel with an erasure factor of 1/2. Hence the capacity of Channel_2 is 1/2. Similarly, when Channel_2 transmits at rate 1, Channel_1 has a maximum rate of 1/2. These combined rates correspond to the points (1/2,1) and (1,1/2) in Fig. 5. They also represent the extrema of Eq. (3). Thm. 1 states that the capacity region is the convex hull of the set of rate pairs satisfying Eqs. (1), (2), and (3). Thus, we see that by "connecting the points" (0,0), (1,0), (1,1/2), (1/2,1), and (0,1) we have the capacity region as shown in Fig. 5.

It is certainly if interest that we can achieve a maximum joint combined rate of 3/2. Of course this is under our assumption that they are not collaborating while transmitting. The next subsection shows that if the Alices do collaborate while transmitting they can, not surprisingly, do better than a combined rate of 3/2. However, for this simple example at least, the two Alices do not do much better.

2.4 Collaborating MuACC

Of course, keep in mind that the channels are not collaborating and their transmissions are independent of each other. However, if Alice_1 and Alice_2 conspire prior to their communications with Eve, they could possibly split a large file between then and thus transmit at a rate of 3/2. Therefore, we see that the

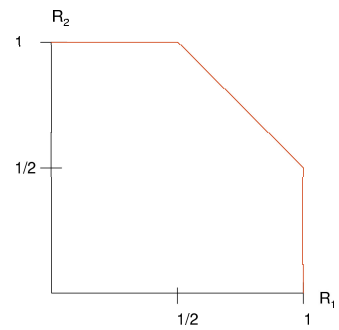

Figure 5: 2 Alices only: Capacity Region

lesson learned is that in a network covert channel scenario one must look at more than the individual covert channel capacities. The true throughput for covert communication is at a sub-additive level of the individual capacities.

If the two Alices are collaborating synchronously (acting as a single transmitter) they can achieve a capacity of $\log 3 \approx 1.58$. In this case we have three output symbols for Eve: 0,1,2 (and the situation is modeled as a standard covert channel.) By collaboration Alice_1 and Alice_2 can send these symbols noiselessly to Eve. Note that $\log 3$ is only slightly larger than 3/2. Of course, analysis between the collaborating and non-collaborating cases must be studied for more complex situation before any conclusions can be drawn. However, $\log 3$ is the maximum rate at which Eve can receive information. There are three output symbols, so one cannot do better than $\log 3$. Therefore, we see without transmitting together the best the covert transmitters can hope for is 1.5 bits per symbol, and by acting as one transmitter this can be raised to $\log 3$ bits per symbol. However, we consider the case of the two Alices acting as one transmitter to be too extreme. We still stick with the assumption that the Alices, even though they may agree on coding strategies prior to transmission, do not collaborate once transmission has started. Therefore, the maximum combined throughput is 1.5 bits per symbol. Or is it? We will return to this issue in the section on Feedback.

3 Clueless$_i$ as noise

So far in all of our concrete work we have concentrated on the very simple example of two Alices and no Clueless users. Unfortunately, the three equations comprising Thm. 1 are quite difficult to work with. Certainly adding Clueless users increases the noise and hence lessens the combined rates. We have also constrained ourselves to only two active covert transmitters (Alice_1 and Alice_2) in our examples. We can certainly have many such Alices. The purpose of this paper was to introduce the concept of multiple access communication channels to the covert channel community. This paper is far from a complete exposition. It is meant to whet the appetite of the reader for the areas of covert channel analysis that have not been considered before. The results with no clueless transmitters can also stand on their own as bounding cases. We conclude this paper with a very interesting and surprising result of Gaarder and Wolf, that for multiple access channels, feedback can *increase* the combined rates. This has serious implication for the covert channel analysis that was done for the network Pump [8].

4 Feedback

By our results above we know that the maximal combined rate pair sums to 3/2. This can be achieved for example by the rate pairs (1,1/2), (1/2,1), (3/4,3/4), etc.. The rate pair of (3/4,3/4) comes about because the capacity region is the convex hull of the rate pairs satisfying Eqs. (1),(2), and (3). In [6], a rate pair of (.76,.76) is constructed. Of course, this is not the same scenario that we presented above. In the above it was tacitly assumed that there was no feedback from Eve to the Alices. For a single input discrete memoryless channel this need not be explicitly stated since feedback does not increase

capacity [18]. At first this result seems counterintuitive, but the genius of Shannon's coding theorem takes all cases into account. What is surprising is that this does not hold for MuACs. Gaarder and Wolf [6] demonstrated this fact interestingly enough for a channel just like the one we have been analyzing.

We will now show that if Eve is allowed to send feedback to Alice_1 and Alice_2 that a rate pair of (.76,.76) can be achieved. Thus we have a rate pair with a combined rate of $1.52 > 1.5$. Gaarder and Wolf [6] use a simple technique with a clever proof to show that (.76,.76) is achievable.

Each Alice_i knows what was received. Thus, if the Alice_i know that Eve received a 0, or a 2, they know that Eve received the symbols without noise, and all is fine. However, if the feedback to the Alice_i is that Eve received the symbol 1, they know that there is noise and Eve does not know if the channel input was (0,1) or (1,0). However, the Alice_i use this to their advantage. The Alice_i agree to just attempt to send the input for Alice_1, the coding/decoding strategy agreed upon on both ends is that after the symbol 1 is received the Alice_i will retransmit the symbol of Alice_1, the symbol for Alice_2 will then be the mod 2 complement of the Alice_1 symbol. The Alice_i actually have 3 symbols to play with, not just two, since they can now noiselessly send (0,0), (1,1), and, without loss of generality (0,1). So they have an input range of $\log 3$ bits. This is only after the noisy symbol of 1 is received by Eve. N is chosen so that $.76N$ is an integer K. To achieve a rate pair of (.76,.76) both Alice_1 and Alice_2 must transmit 2^K messages in N uses of the channel. This is accomplished by each Alice_i sending K uncoded bits (of course if there was not any noise when Eve received a 2 we would be done. Let Q be the number of transmissions for which Eve received a 1. We are left with $N - K = .24N$ uses of the channel to try to "get the noise out". As discussed above the Alice_i can actually send 3 symbols in each of these type uses. So, as long as $2^Q \leq 3^{N-K}$ the noise can be taken out, and we would be able to send $2^{.76N}$ distinct and noiseless messages in N uses of the MuACC.

This now boils down to showing that the probability $p_e = P(2^Q > 3^{(N-K)})$ can be made as small as desired. We may rewrite p_e as $p_e = P(Q > .24N \log 3)$. Recall that Q is the number of transmissions where Eve receives a 1. Q may be modeled as a binomial random variable with parameters $N, 1/2$, this is since there are N trials and each outcome (0,0), (0,1), (1,0), and (1,1) is equally likely (since there is no bias in whether the Alice_i send (0,1) or (1,0)). Therefore half of the trials result in the output 1 to Eve, hence the 1/2 parameter. Therefore, Q has mean $\mu = K/2$ and variance $\sigma^2 = K/4$. Thus, with $K = .76N$, $\mu = .38N$ and $\sigma^2 = .19N$. Since
$p_e = P(Q - \overline{Q} > .24N \log 3 - \overline{Q}) = P(Q - \overline{Q} > .38039N - .38N)$
$= P(Q - \overline{Q} > .00039N)$, and $P(Q - \overline{Q} > .00039N) < P(|Q - \overline{Q}| > .00039N)$ we have by Chebyshev's inequality that $p_e \leq \frac{\sigma^2}{(.00039N)^2} = \frac{.19}{(.00039)^2 N}$. Thus we see that as N grows p_e approaches zero, so the error can be made as small as possible with a rate pair of (.76,.76). Gaarder and Wolf never claimed their method was optimal, in fact if we attempt the same procedure with at rate pair of (.77,.77) we have non-trivial asymptotic error. What is so important about Gaarder and Wolf's example is that it is above (.75,.75). The actual bounds are unknown. However, we do know that the combined rate pair cannot be greater than $\log 3 = 1.5850$, since there are only three output symbols (0,1,2) received by Eve. Therefore the true capacity region, if we allow feedback is greater than what is given by Thm. 1.

5 Conclusion

This has been a brief introduction to the area of MuACCs in covert channels. In it, we consider only the noise introduced by multiple transmitters (i.e., there are no clueless senders). Clueless senders act as noise to the Alices, but we still must consider some sub-additive measure of the individual capacities.

Here, we have only considered the simple case of two conspiring Alices; there could be more covert channel senders. Future work will study the effects of more transmitters, as well as the effects of clueless senders on the capacity. It will compare these to the effects of clueless senders on a single transmitter with multiple symbols.

The simplified Mix under consideration is a timed Mix, so the channel is a storage channel. In the case of threshold Mixes, the output is always the same (a constant number of messages each time it fires, sent to the other Mix-firewall), but the time between firing varies. Hence, it is a timing channel. We know of no theoretical or other type results for dealing with multiple access type channels where the time values are the information carrying symbols. This is an open area of research that should be investigated.

We also note that the best coding and transmission strategy that Alice_1 can use when Alice_2 is also transmitting may be different from the best coding and transmission strategy she can use when Alice_2 is not transmitting, even at the same channel rate for Alice_1. Since we assume that neither Alice knows whether or when the other Alice is transmitting, their coding method and transmission strategy must accommodate these contingencies. It is easy to require that both Alices continuously exercise the channel, sending dummy messages that are discarded by Eve when they have nothing to send, but this seems wasteful. In fact, since the absence of transmissions by the other Alice should reduce noise in the channel, it should become more reliable when the other Alice stops sending for some time. However, this has not been shown here, and begs for further investigation.

The purpose of this paper is to point out how the theoretical tools of network information theory allow us to examine covert channel in networks in a new light. We can no longer simply study the covert channels in isolation to get a complete gauge of the potential amount of information leakage. We must see how multiple channels can act in unison to leak information.

6 Acknowledgments

We thank the reviewers for their helpful comments.

References

[1] R. Ahlswede. Multi-way communication channels. In *Proceedings 2nd International Symposium on Information Theory*, pages 23–52, Tsahkadsor, Armenian SSR, 1971. Budapest, Hungary: Hungarian Acad. Sci., 1973.

[2] Abdelaziz Amraoui, Sanket Dusad, and Rüdiger Urbanke. Achieving general points in the 2-user gaussian MAC without time-sharing or rate-splitting by means of iterative coding. In *Proceedings 2002 IEEE International Symposium on Information Theory*, page 334, Lausanne, Switzerland, 2002.

[3] Robert B. Ash. *Information Theory.* Dover, 1965.

[4] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory.* Wiley, 1991.

[5] N. Thomas Gaarder and Jack K. Wolf. The capacity region of a multiple-access discrete memoryless channel can increase with feedback. *IEEE Transactions on Information Theory*, 21(1):100–102, 1975.

[6] N. Thomas Gaarder and Jack K. Wolf. The capacity of a multiple-access discrete memoryless channel can increase with feedback. *IEEE Transactions on Information Theory*, 21(1):100–102, 1995.

[7] Te Sun Han. An information-spectrum approach to capacity theorems for the general multiple-access channel. *IEEE Tran. Information Theory*, 44(7):2773–2795, 1998.

[8] Myong H. Kang, Ira S. Moskowitz, and Daniel C. Lee. A network Pump. *IEEE Transactions on Software Engineering*, 22(5):329–328, 1998.

[9] Butler W. Lampson. A note on the confinement problem. *Communications of the ACM*, 16(10):613–615, 1973.

[10] Edward C. Van Der Meulen. A survey of multi-way channels in information theory. *IEEE Transactions on Information Theory*, 23(1):1–37, 1977.

[11] Allen R. Miller and Ira S. Moskowitz. Difference of sums containing products of binomial coefficients and their logarithms. *SIAM Review — to appear*, 2004.

[12] Ira S. Moskowitz and Myong H. Kang. Covert channels — here to stay? In *Proc. COMPASS'94*, pages 235–243, Gaithersburg, MD, June 27- July 1 1994. IEEE Press.

[13] Ira S. Moskowitz, Richard E. Newman, Daniel P. Crepeau, and Allen R. Miller. Covert channels and anonymizing networks. In *ACM WPES*, pages 79–88, Washington, October 2003.

[14] Ira S. Moskowitz, Richard E. Newman, and Paul F. Syverson. Quasi-anonymous channels. In *IASTED CNIS*, pages 126–131, New York, December 2003.

[15] Richard E. Newman, Vipan R. Nalla, and Ira S. Moskowitz. Anonymity and covert channels in simple timed mix-firewalls. In *Workshop on Privacy Enhancing Technologies*, page TBD, Toronto, May 2004. Springer-Verlag, LNCS TBD.

[16] Andreas Pfitzmann and Michael Waidner. Networks without user observability – design options. In *Proc. of EUROCRYPT 1985*. Springer-Verlag, LNCS 219, 1985.

[17] Claude E. Shannon. The mathematical theory of communication. *Bell Systems Technical Journal*, 30:50–64, 1948.

[18] Claude E. Shannon. The zero error capacity of a noisy channel. *IRE Trans. on Information Theory*, Vol. IT-2:S8–S19, September 1956.

[19] Claude E. Shannon. Two-way communication channels. In *Proceedings Fourth Berkeley Symposium Probability and Statistics*, pages 611–644, Berkeley, CA, 1985.