Steganography Assisted Tor

Mike Freyberger

March 2015

1 Introduction

Tor, the onion router, is an anonymity system that allows for privacy and decreased censorship during web browsing. However, there are a few attacks that can still mitigate the increase in privacy that Tor provides. One of the most popular attacks is a timing analysis attack. This attack can take place when there are two attackers, one eavesdropping the victim client and one eavesdropping the victim server. Both of the eavesdroppers work together to determine correlations between the packet traces in order to determine if the client is communicating with the server. Interesting research has been done to evaluate the effectiveness of application level defenses. This paper is interested in determining if there are possible innovations within the physical layer that can protect Tor against the timing analysis attack. Most specifically, this paper will investigate how the Tor system can be assisted by the use of Steganography.

2 Steganography

2.1 What is Steganography?

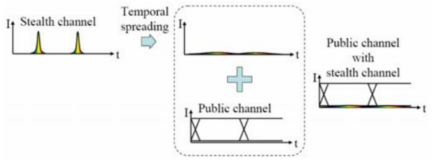


Figure 1: Temporal spreading allows stealth messages to be concealed within public messages.

Steganography is concealing a message within another message. Therefore, within optical communications, stenography refers sending a stealth message that is concealed within the public message. This is achieved by using a high dispersive fiber to flatten out the signal, referred to as temporal spreading, and then the message becomes a low amplitude signal that can be carried secretly in the noise of the public message. Given the amount of amplifiers in an optical communication system, there is a lot of amplified spontaneous emission noise. As long as the spread signal falls below the amplitude of the noise, it can be carried securely in the public channel. This temporal spreading is achieved by group velocity dispersion (GVD). Figure 1 demonstrates how the public channel does not look any different once the stealth messages have been added to it. Given that a high dispersive fiber is used to transmit the stealth, the inverse of that dispersion must be in place at the receiver end in order for the stealth signal to be discovered. The fiber that is used at the receiver end is called a dispersion compensating fiber. The dispersion parameter that is associated with the group velocity dispersion can be positive or negative.

$$\sum_{i=1}^{n} L_i D_i = 0 \tag{1}$$

Equation (1) must be held in order for the stealth message to be recovered on the receiver end. L_i is the length of one of the fibers in the system, and D_i is the dispersion parameter associated with that fiber. Figure 2 demonstrates the use of a highly dispersive fiber at the input, and a dispersion compensation fiber at the output.

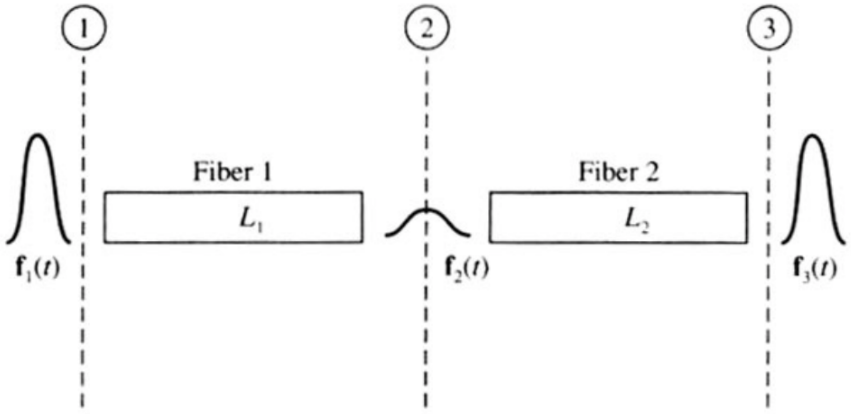


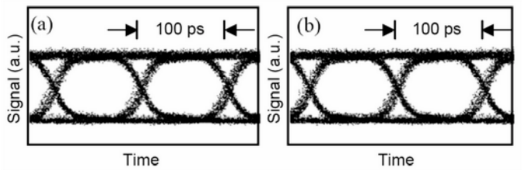
Figure 2: Demonstrating the use of a highly dispersive fiber in Fiber 1, and a dispersion compensating fiber in Fiber 2, in order to recover the original signal.

2.2 Time Domain Analysis

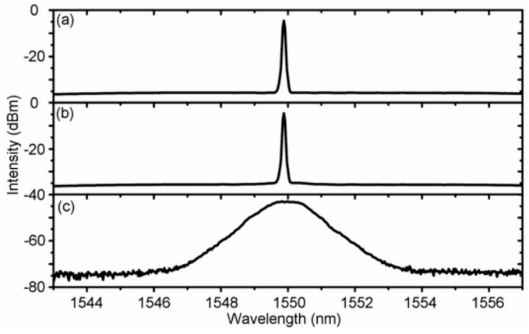
The most important part of Steganography is that the stealth signal is concealed in both the time and frequency domain. To demonstrate the stealth signal is concealed in the time domain look at Figure 3 (left). This Figure shows that the signal does not look any different in the time domain when the stealth signal is present.

2.3 Frequency Domain Analysis

Concealing the signal in the frequency domain is also integral to making the stealth message completely concealed from an eavesdropper. There are two main approaches that effectively conceal the signal in the frequency domain. The first approach simply makes sure the stealth signal has the same frequency components as the public signal. In this case, the frequency spectra of the public signal with and without the stealth signal would look the exact same. The other approach is to spread the frequency spectra, and



Eye diagram of the signal. (a) Without stealth signal. (b) With steal signal.



Frequency spectra of public signal. (a) The spectra without the stealth signal. (b) The spectra with the stealth signal. (c) The spectra of the stealth signal alone. Notice it's intensity is much smaller than the public channel intensity.

Figure 3: (left) Stealth signal concealed in the time domain. (right) Stealth signal concealed in the frequency domain.

decrease the amplitude so that spectra of the stealth signal goes below the noise level. These techniques can be combined effectively in order to completely conceal the stealth signal within the public signals frequency spectra. Figure 3 (right) demonstrates that the public channel's frequency spectra looks the same with and without the stealth signal.

3 Tor

Tor, the onion router, provides anonymity online. Tor protects the clients privacy by providing a browser that can browse the Internet without being monitored.

3.1 The Problem Tor Solves

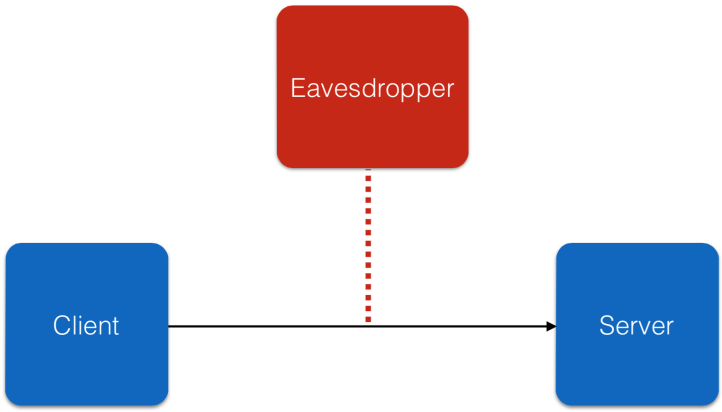


Figure 4: Simple connection directly between the client and server. The eavesdropper can compromise this link, and determine when the client is communicating with the server.

Without the use of Tor, Internet connections can be monitored directly. Most connections are in the form presented in Figure 4. In this case, the client is directly communicating with the server. Even though it is probable that the actual contents of the messages are encrypted, the presence of a communication between the client and server would be known to the the attacker. For instance, if the attacker wants to know if I (the victim client) am accessing content from cnn.com (the victim server), the eavesdropper will be able to identify this communication due the fact packets are present on the link. In order to mitigate this simple attack, Tor sets up a circuit of three relays for all Internet connections. A possible circuit of connections is demonstrated in Figure 5. Before accessing any web page, a client using the Tor browser determines a set of three Tor node to communicate through, and this connection path completely separates the client from the server. If the victim client is using Tor, the eavesdropper will have no means of definitely determining there is communication between the victim client and victim server.

3.2 How Tor Works

In order to better understand how Steganography can ultimately integrate with Tor, it is important to understand the fundamentals empowering Tor. The fundamentals of Tor are symmetric and asymmetric encryption, and the idea of an onions layers.
1) The first step is selecting the three Tor nodes to communicate through. There are roughly 7,000 Tor nodes in the Tor system according to Tor metrics. The more nodes in the system increases the possible configurations a client can select, and this ultimately increases the privacy of the system.
2) Now the path is selected, the initiating clients establishes a symmetric key with all of the relays in the circuit. In order to establish the symmetric keys, the client and the relays use asymmetric cryptography.

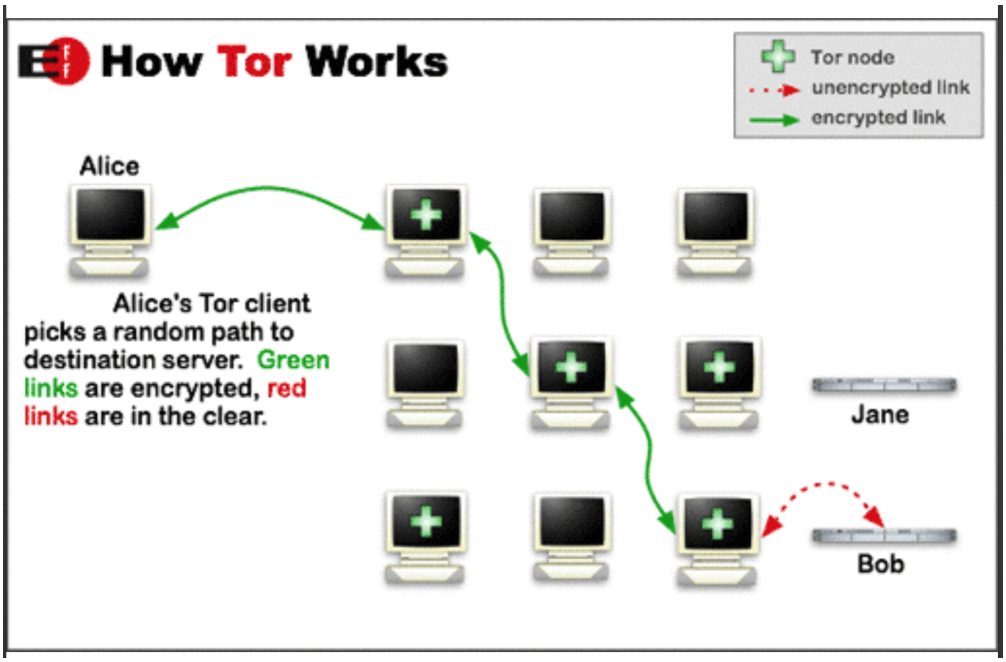


Figure 5: The Tor system selects a set of three Tor node relays to communicate through in order to increase privacy throughout the Internet.

The client can encrypt the symmetric keys with the Tor Nodes public keys, and then Tor Nodes can each decrypt using their respective private keys. This establishes three symmetric keys in the system, one for each node.
3) The client encrypts the web request or any Internet message with all three symmetric keys. Imagine the message as the core of the onion, and each symmetric key is one layer of the onion. A good diagram of this is provided in Figure 6.
4) Each node will peel off one layer of the onion by decrypting the message with their own symmetric key. Once Tor Node 3 decrypts with his symmetric key, the original message (the core of the onion), will be all that's left.

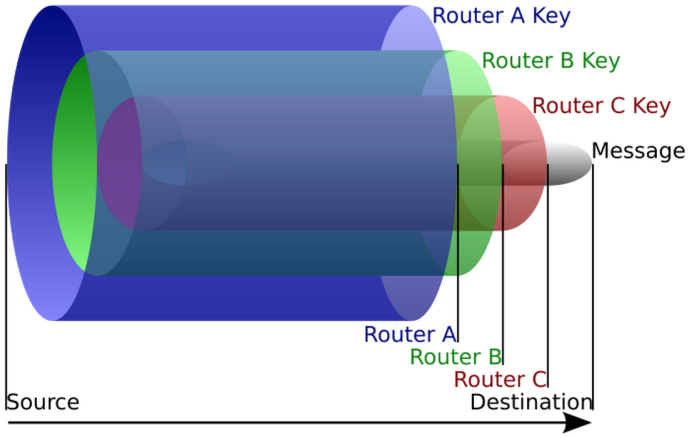


Figure 6: The onion encryption scheme. The message is wrapped in three symmetric keys. Each Node (Router) peels off one layer until the actual message is delivered to the destination.

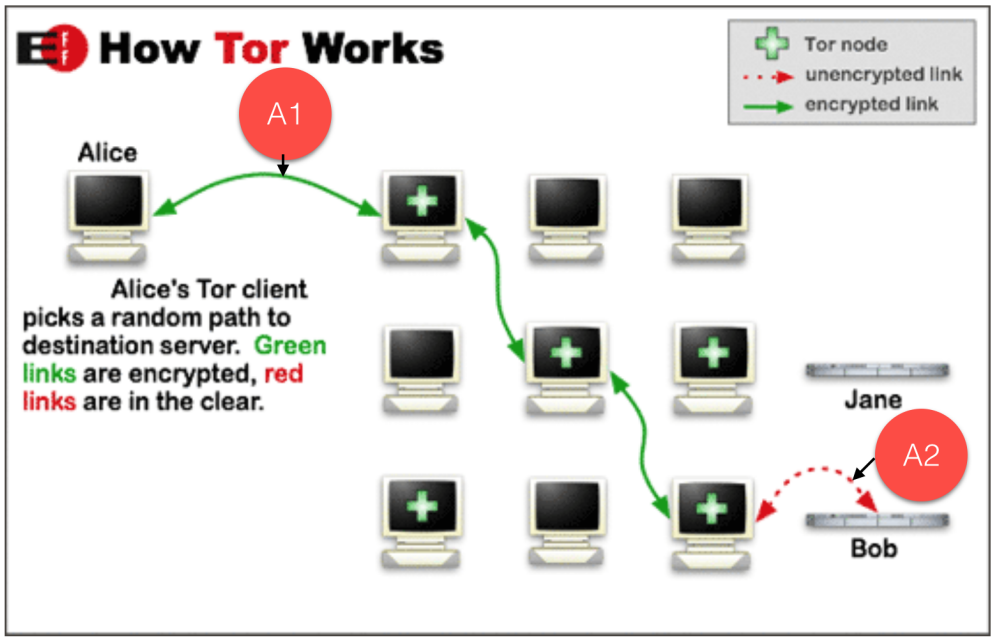


Figure 7: Timing Analysis attack with one attacker eavesdropping the inbound link, and another attacker eavesdropping the outbound link.

5) The exiting relay delivers the unencrypted message to the server. Ideally the application layer will provide some layer of security, such as SSL or TLS, so that the exiting relay cannot investigate the packet for any personal data such as user name, password, address, etc..
6) The server then responds to the message, and the same path is used for the return path. During the return, each relay encrypts the message with his symmetric key. Therefore, at the end when Tor Node 1 delivers content to the source, the content will be fully wrapped with all three symmetric keys.
7) Finally, the originator will remove all three layers of encryption by decrypting with all three symmetric keys that the originator created in the first place, and the originator will be left with the response from the server.

This design effectively prevents the passive eavesdropper from determining the client is communicating with the server. However, there are some sophisticated attacks that mitigate the Tor system design, and still can determine the victim client is communicating with the victim server.

3.3 Timing Analysis on Tor

The Tor system is vulnerable to a Timing Analysis attack, which requires an attack connected to the inbound and outbound link. This attack is pictured in Figure 7. In this attack, the two attackers work together in order to determine how probable it is that the packets the inbound attacker witnesses are the same packets the outgoing attacker witnesses. For instance, say the inbound attacker sees a communication trace that looks like Figure 8 top and the outbound attacker sees a communication trace that looks like Figure 8 bottom. These two attackers can conclude with high probability the victim client and server are in communication because of the correlation between the two communication traces. It is possible for network congestion, dropped packets, slow intermediate links, or any uncertain in the network to cause the two streams to look rather different. However, for many networks the inbound and outbound attacker can work together in order to break past the privacy barrier the Tor system tries to provide.

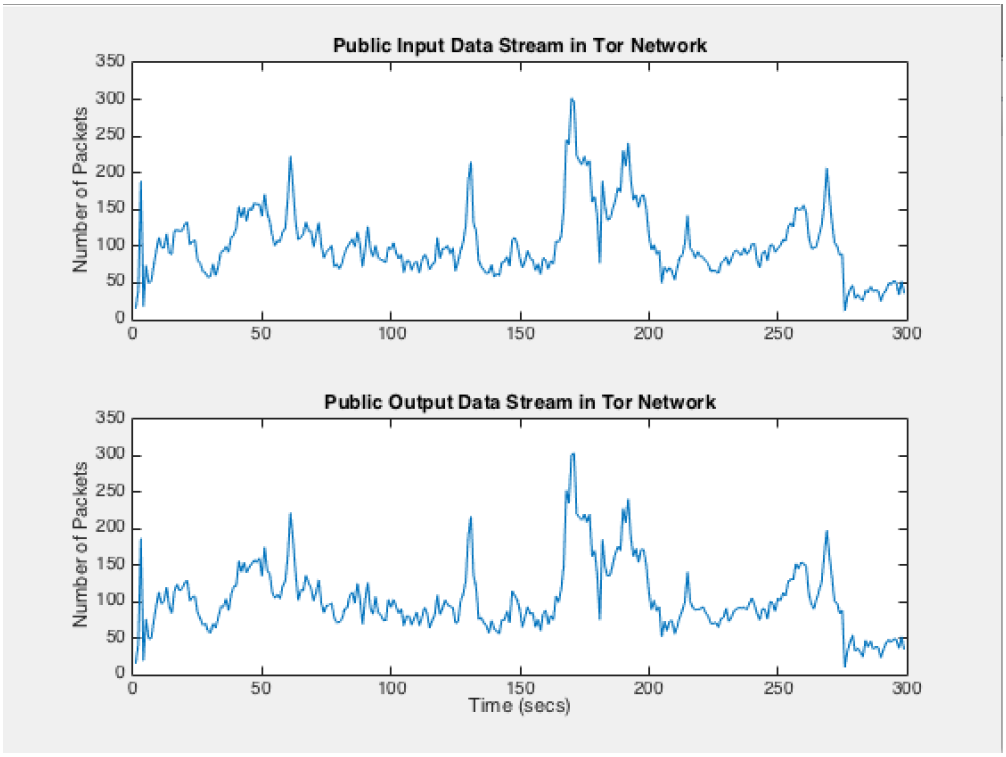


Figure 8: The input and output communication stream in the Tor network typically look identical.

4 Steganography Assisted Tor

In order to prevent Timing Analysis Attacks on Tor, Steganography can be used between the client and the first node relay, which is demonstrated in Figure 9. With the stealth channel at this place in the communication system, schemes can be developed that use a combination of the public and stealth channel so tat the attackers see communication streams that do not have a high correlation.

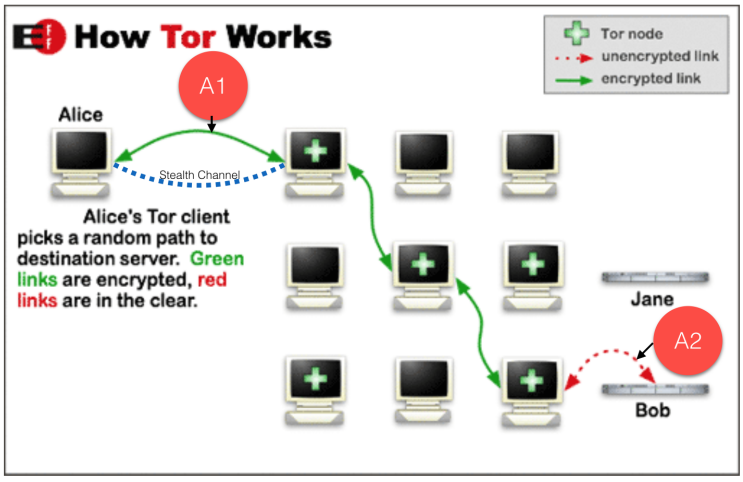


Figure 9: In order to prevent timing analysis attacks a stealth channel is added between the client and the first Tor relay.

5 Simulation Model

5.1 Timing Analysis

Our attacker can count the number of packets on a network link that occur within a fixed time window. This fixed time window is much less than the time of the connection. For instance, in the our data set, the fixed time window is 1 second, and connection time is 300 seconds; therefore, there are 300 distinct time segments each with a measured packet count. Each attacker, the one connected to the inbound link and the one connected to the outbound link, has the same fixed time window, and can measure the packet counts. Using the packet counts for each time segment, the attackers can determine the correlation using (2) below. [1] The attacker selects also selects a *threshold* value, and if the correlation is about that value, than the attacker considers the two sequences of communication to be on the same communication path.

5.2 Correlation Function

The following equation is used to calculate the correlation between the packet counts measured on the inbound and outbound link:

$$r(d) = \frac{\sum_i ((x_i - \mu)(x'_{i+d} - \mu'))}{\sqrt{\sum_i (x_i - \mu)^2} \sqrt{\sum_i (x'_{i+d} - \mu')^2}} \tag{2}$$

x_i and x'_i are the number of packets in the ith timing segment on the inbound link and outbound link respectively. μ and μ' are the average number of packets in a timing segment, averaged across the entire time of connection, for each sequence. d is the delay, which is 0 for our simulations.

5.3 False Positive & False Negative

False positive rate and false negative rate are important metrics in determining the effectiveness of schemes that use Steganography assisted Tor. False positives refer to case when the attackers believe the two sequences are on the same communication link, when they are actually not. For instance, if victim A is going to cnn.com and nytimes.com is the victim server, the packet counts measured by the inbound and outbound attacker should not appear to be correlated. However, if the correlation value calculated by the attackers on the two sequences yields a correlation value above the threshold, the attackers will incorrectly conclude that victim A is communicating with nytimes.com. On the other, a false negative occurs when the attackers conclude that the two sequences are not on the same communication link, when they actually are. For instance, if victim A is going to nytimes.com and nytimes.com is the victim server, the packet counts measured by the inbound and outbound attacker should appear to be correlated. However, if the correlation value calcuated by the attackers yields a correlation value below the threshold, the attackers will incorrectly conclude that victim A is not communication with nytimes.com.

5.4 Dataset

The dataset used for the simulations came from 50 different instances of uploading a file for 300 seconds in the Tor environment. Each instance of uploading the file had a slightly modified setting, so no two streams are exactly alike. Using all of these different inbound and outbound packet streams, we are able to determine the attackers false positive and false negative rate for different threshold values.

5.5 Defense Metric

The defense metric used to measure the effectiveness of a steganography scheme will be the area under the curve when the false positive rate is plotted against 1 - false negative rate. Each rate is calculated for all threshold values from 0 to 1. This curve will be nearly linear for very effective schemes. The ideal curve will be a linear line connecting the origin and (1,1). Therefore, the ideal area under curve is .5. Non-ideal

[1] The attacker is limited to this information, and cannot use the start and end time of the connection to determine correlation.

schemes, will spike up to (0,1) very rapidly. This type of curve indicates even though false positive rates may change with threshold, the false negative rate is relatively constant.

5.6 Tradeoff

In order to achieve an ideal area under curve, the steganography scheme must use the stealth channel for 100% of the packets. However, since the stealth channel has less bandwidth, this type of scheme will also have less bandwidth. The greater percentage of packets that are sent over the stealth channel will lead to a smalled bandwidth on channel between the client and the first Tor node.

6 Simulation Results

6.1 No Steganography

The first simulation uses zero stenography to demonstrate the effectiveness of timing analysis attacks. In this attack for almost every possible threshold value the false negative rate is 0%, therefore the correlation between two parties are actually communicating is very high. The average correlation between two parties without any steganography is .99. The effectiveness plot is provided in Figure 10.

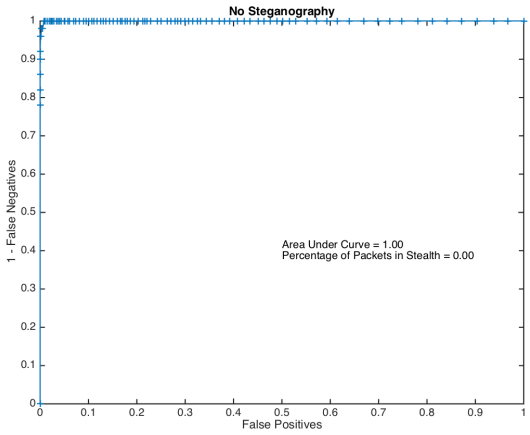


Figure 10: No Steganography. All packets are sent in the public channel.

6.2 Full Steganography

On the other extreme, we can use the stealth channel to send all of the packets between the client and first Tor node. In this case, there will be no correlation because the attacker eavesdropping on the inbound link will see no packets on the public channel. This scheme will completely remove the threat of timing analysis attacks. However, this scheme has the largest effect on bandwidth, as the public channel is not used at all, and the bandwidth is entirely defined by the bandwidth on the stealth channel. The effectiveness plot is provided in Figure 11.

6.3 Deterministic Partial Steganography

Rather than completely using the public channel or the stealth channel for the whole communication, this scheme sends a particular percentage of packets on the stealth channel throughout the communication. This scheme will hurt the bandwidth less than the full steganography approach, however, this scheme will not be as effective as the full steganography approach. In fact, even though this scheme uses the stealth channel for some percentage of the packets in the, this scheme is ineffective. Even if this scheme sends 90% of the packets in the stealth channel, this scheme will not benefit the client's privacy because timing analysis attacks can

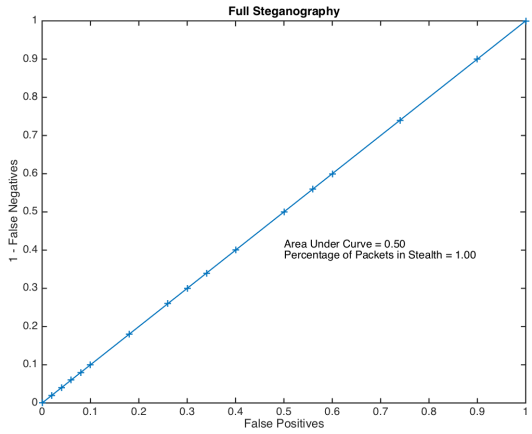


Figure 11: Full Steganography. All packets are sent in the stealth channel.

still be used. In the case 90% of packets are sent in the stealth channel, the attackers just begin to expect a different number of packets in each timing window, but the difference will always be 90%. For instance if the inbound attacker sees 1 packet in the first second, the outbound attacker will see 10 packets. Then, during the next second the inbound attacker might see 3 packets, and the outbound attacker will see 30 packets. Therefore, there will still be a high correlation between two parties that actually communicating. When 90% of packets are deterministically sent in the stealth channel, the average correlation between two parties is .98. The effectiveness curve is provided in Figure 12.

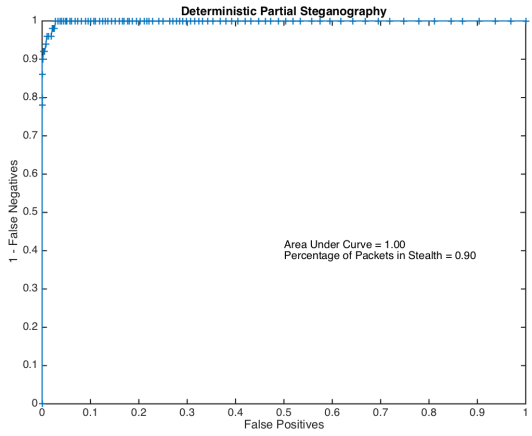


Figure 12: Deterministic Partial Steganography. 90% of packets are sent in the stealth channel.

6.4 Random Partial Steganography

This scheme uses the stealth channel for a percentage of the packets as well. However, this channel selects a random percentage at the beginning of every time segment, and sends that percentage on the stealth channel. A random percentage from 0 to 1 is selected. This scheme is more effective than the deterministic scheme, and averages out to send roughly 50% of packets on the stealth channel. This scheme would be difficult to implement because there will be a lot of switching between the public and stealth channel. Furthermore, some random percentages would be impossible to achieve. For instance, if the random percentage is set to be 40%, and only 3 packets come in during the frame, it's impossible to send 40% of the packets on the stealth channel. The effectiveness plot is provided in Figure 13.

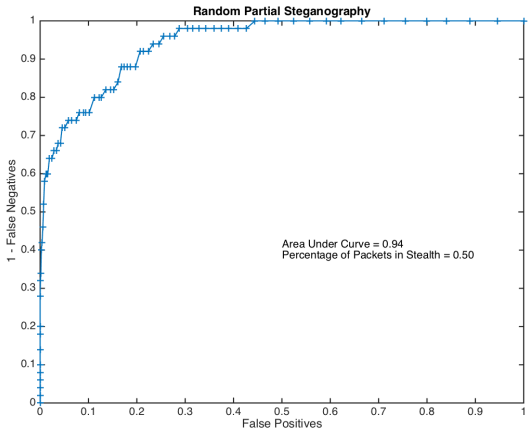


Figure 13: Random Partial Steganography. 50% of packets are sent in the stealth channel.

6.5 Switch Steganography

This scheme either uses complete stealth or complete public channel for each time segment. Before each time segment, the switch is set randomly to either the public or stealth channel. Essentially a coin is flipped before each time interval: if heads, use the stealth channel, otherwise use the public channel. Once again this scheme averages out to send 50% of packets on the stealth channel. Furthermore, this scheme would be significantly easier to implement because the only one channel needs to be used each time segment, so there is less switching between channels. The effectiveness plot is provided in Figure 14.

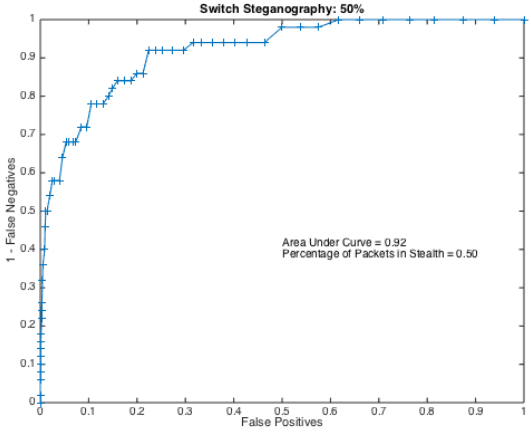


Figure 14: Switch Steganography 50%. Either steganography is completely used, or not used at all for each time segment. This is decided randomly with a coin flip. 50% of packets are sent in the stealth channel.

This scheme can be modified to a use a ten sided die in the beginning rather than a coin. In this case, only if the die rolls a 1 will the public channel be used. This scheme will lead to roughly 90% of packets using the stealth channel. This scheme is naturally more effective than the previous switch steganography scheme where only 50% packets use the stealth channel. This modification leads to an effectiveness value of .79.

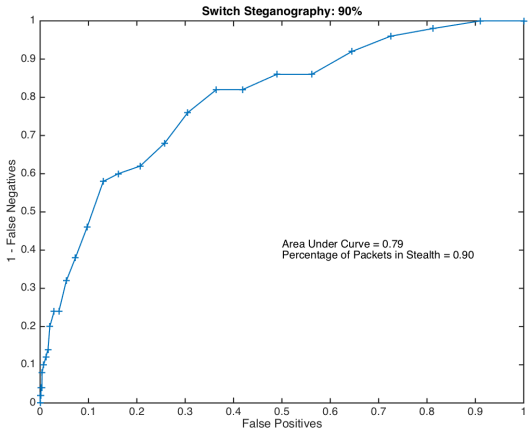


Figure 15: Switch Steganography 90%. Either steganography is completely used, or not used at all for each time segment. This is decided by selecting a random value from 0-9, and only using the public channel if the value is a 1. 90% of packets are sent in the stealth channel.

6.6 Selective Partial Steganography

This scheme keeps a running calculation of the mean number of packets in a time segment. If the number of packets in the next time segment is greater than the current mean value, than 50% of the packets in that time segment are sent over the stealth channel. This scheme introduces a fixed delay of one time segment, which is required to count to the number of packets in each time segment. This scheme is the most effective as the percentage of packets in the stealth channel is only 35% and is the most effective in terms of area under the effectiveness curve. The effectiveness plot if provided in Figure 16.

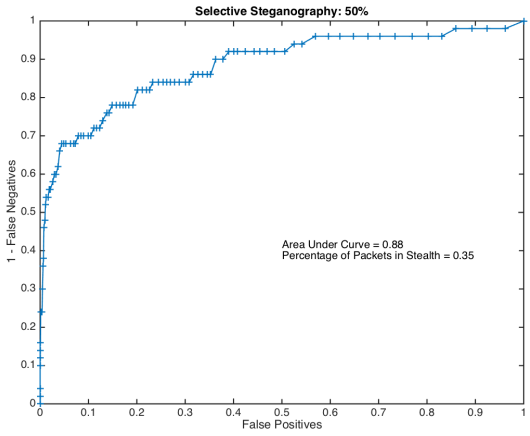


Figure 16: Selective Partial Steganography 50%. Either steganography is 50% used, or not used at all for each time segment, depending on whether or not the packet count in the time segment is greater than the running average. 35% of packets are sent in the stealth channel.

This scheme can use the same logic, but when the packet count in the time segment is greater than the running average, 80% of packets are sent in the stealth channel. This naturally leads a greater percentage of packets sent in the stealth channel, but also makes the effectiveness nearly linear. This scheme leads to the effectiveness curve found in Figure 17.

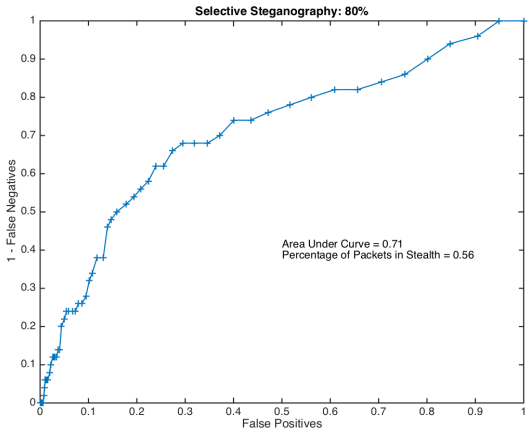


Figure 17: Selective Partial Steganography 80%. Either steganography is 80% used, or not used at all for each time segment, depending on whether or not the packet count in the time segment is greater than the running average. 56% of packets are sent in the stealth channel.

6.7 Tradeoff

The final tradeoff curve is provided in Figure 18. This graph show Full and No Steganography at the far extremes. The most ideal point would be at the (0,.5), because this would indicate we used the high bandwidth public channel for all communication, and the effectiveness curve was perfectly linear. However, this point is impossible to achieve, because if we do not use the stealth channel at all, the timing analysis attack will be very effective. Ideal points will get close to the (0,.5) point because this means the amount stealth channel used is low, so bandwidth is not negatively impacted, but the effectiveness is still great. From the curve below, the most effective scheme is the selective partial steganography, as this scheme decreases the effectiveness of timing analysis without having a terrible impact on bandwidth.

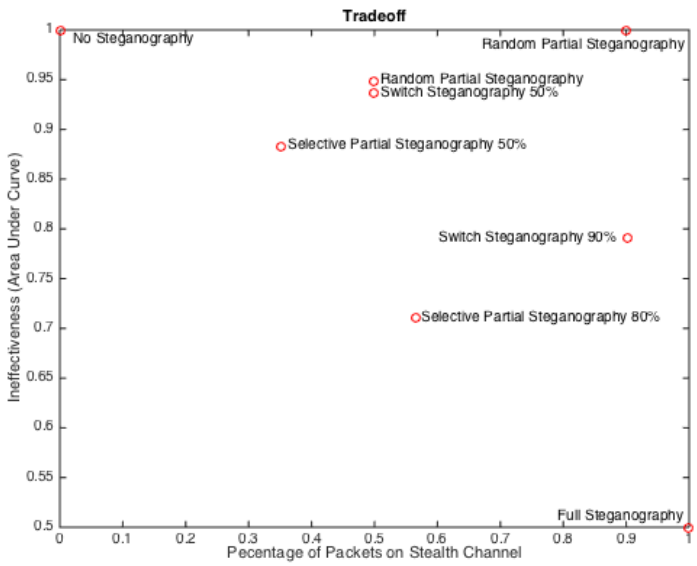


Figure 18: Tradeoff curve.

References

[1] *Optical steganography based on amplified spontaneous emission noise*. Ben Wu,* Zhenxing Wang, Yue Tian, Mable P. Fok, Bhavin J. Shastri, Daniel R. Kanoff, and Paul R. Prucnal.

[2] *Studying Timing Analysis on the Internet with SubRosa*. Hatim Daginawala and Matthew Wright

[3] *Timing analysis in low-latency mix networks: attacks and defenses*. Vitaly Shmatikov and Ming-Hsiu Wang

[4] *Timing Attacks in Low-Latency Mix Systems (Extended Abstract)*. Brian N. Levine, Michael K. Reiter, Chenxi Wang, and Matthew Wright