On the Effectiveness of Low Latency Anonymous Network in the Presence of Timing Attack

Jing Jin
Department of Computer Science
George Mason University
Fairfax, VA 22030, USA
jjin3@gmu.edu

Xinyuan Wang
Department of Computer Science
George Mason University
Fairfax, VA 22030, USA
xwangc@gmu.edu

Abstract

In this paper, we introduce a novel metric that can quantitatively measure the practical effectiveness (i.e. anonymity) of all anonymous networks in the presence of timing attack. Our metric is based on a novel measurement of the distortion of the packet timing between the incoming and the outgoing flows to and from the anonymous network and it uses wavelet based analysis to measure the variability of the distortion. To the best of our knowledge, our approach is the first practical method that can quantitatively measure the packet timing distortion between flows that may have gone through such transformations as flow mixing/spliting/merging, adding chaff, packet dropping.

To validate our anonymity metric, we have conducted real-time timing attacks on various deployed anonymous networks such as Tor, anonymizer.com and have used the timing attack results as the ground truth for validating our anonymity metric. We have found strong correlation between our anonymity metric and the timing attack results. Our metric measurements and timing attack results show that the circuit rotation in Tor network could significantly increase its resistance to timing attack at the cost of more timing disturbances to the normal users. In addition, we have found that adding constant rate chaff (i.e. cover traffic) has diminishing effect in anonymizing packet flows.

Keywords: Dependability benchmarking, Measurement techniques, Networking and networked systems, Reliability, availability and safety, Security.

1 Introduction

Privacy and anonymity are a major concern for Internet users. To provide anonymous, real-time communication (e.g., Internet browsing) for Internet users, many low-latency, anonymous networks (e.g. Anonymizer.com [3], Crowds [25], Onion Routing [24], Tor [10], Hordes [28], Web Mixes [6], Tarzan [13]) have been proposed to disguise the identity and correspondence between the communicating parties.

However, the timing constraint imposed by the requirement of low-latency makes low-latency networks susceptible to the timing-attack [12, 30, 20, 21, 15, 29, 32, 33, 23], which essentially exploits the timing correlation between the original flow and the anonymized flow to correlate them. Since no practical low-latency anonymous network could completely eliminate the timing correlation between the original flow and the anonymized flow, all practical low-latency anonymous networks are subject to timing attacks. Therefore, it is important to understand the negative impact of timing attack on low-latency anonymous networks.

To evaluate the resilience of various low-latency anonymous networks against the timing attack, we need a metric to quantitatively measure the effectiveness of various anonymous networks in the presence of timing attack. Such a generic metric not only lets us to compare different deployed anonymous networks, but also enables us to analyze new anonymity techniques in order to design better anonymous network in the presence of timing attack.

In this paper, we propose a novel metric that can quantitatively measure the practical effectiveness (i.e. anonymity) of all anonymous networks in the presence of timing attack. Recognizing that the objective of anonymous network is to disguise the anonymized flow as much as possible so that it is hard to be correlated with the corresponding original flow in the packet timing domain, we build our anonymity metric upon a novel measurement of the packet timing distortion between the incoming and the outgoing flows to and from the anonymous network. To the best of our knowledge, our approach is the first practical method that can quantitatively measure the packet timing distortion

between flows that may have gone through such transformations as flow mixing/spliting/merging, adding chaff, packet dropping. We use wavelet-based multiresolution analysis (MRA) to capture the variability of the timing distortion at all scales, and quantify the effectiveness of low-latency anonymous network in the presence of timing attack as the wavelet-based energy.

To validate our anonymity metric, we have conducted real-time timing attack on various deployed anonymous networks such as Tor, anonymizer.com, findnot, stegnos and have used the timing attack results as the ground truth for validating our anonymity metric. We have found strong correlation between our anonymity metric and the timing attack results. Our analytical and empirical results show that the circuit rotation in Tor network could significantly increase its resistance to timing attack at the cost of more timing disturbances to the normal users. In addition, we have found that adding constant rate chaff (i.e. cover traffic) has diminishing effect in anonymizing packet flows.

The rest of the paper is organized as follows. In section 2, we briefly overview related works in anonymous network and timing attack. In section 3, we build our wavelet-based energy metric upon a new packet timing distortion measurement and describe several properties of the new metric. In section 4, we empirically validate our new anonymity metric with real time experiments on Tor, anonymizer.com, Steganos, and findnot.com. We conclude in section 5.

2 Related works

Since Chaum [7] first introduced the mix network for anonymous email, a number of low-latency anonymous networks [3, 1, 2, 24, 25, 28, 6, 17, 13, 27, 10] have been proposed, developed and deployed. Notably, Onion Routing [24] and its second generation, Tor [10], use a sequence of proxies and public key encryption to protect the transport of TCP flows. Crowds [25] uses randomly selected proxies to make it hard to track the sender and receiver. However, none of these methods were designed to provide the unlinkability of sender and receiver. Both NetCamo [17] and Tarzan [13] use cover traffic to anonymize the real-time traffic. Hordes [28] uses multicasting to provide sender anonymity. P5 [27] uses broadcast to provide sender-, receiver-, and sender-receiver anonymity assuming the adversary is passive. Among the deployed low-latency anonymous networks, anonymizer.com [3] is the most popular commercial anonymous communication service in the USA and Tor [10] is the most popular open source low-latency anonymous network.

To exploit the timing constraint of the low-latency anonymous networks, a number of timing attacks [12, 30, 20, 21, 31, 15, 29, 32, 33, 23] have been proposed and identified. Specifically, Wang et al. [32] have developed an active flow watermarking scheme that has successfully "penetrated" anonymizer.com [3]. Yu et al. [33] have developed similar flow watermarking scheme based DSSS (direct-sequence spread spectrum) technique. Murdoch and Zieliński [22] conducted passive traffic analysis on sample data collected from Internet exchanges.

Berthold at el [6] defined the degree of anonymity as the log of the number of users in the system. Both Diaz [9] and Serjantov and Danezis [26] proposed the information-theoretic metric to measure the anonymity. Danezis [8] further applied the metric [26] to continuous-time mixes, where inter-arrival time of the messages is Poisson distributed. Zhu et al [34] investigated the relationship between the anonymity degree and information leakage from an anonymous network. Hopper et al [19] studied the information leak due to the knowledge of network latency (RTT) and to what extent such information leak could be used to compromise anonymity.

To the best of our knowledge, none of the existing anonymity metrics has considered active timing attack in their models. Therefore, no existing anonymity metric can measure the effectiveness of low-latency anonymous network under timing attack.

3 Wavelet-Based Energy Metric of Anonymity

In this section, we present an energy-based metric to quantitatively measure the effectiveness of anonymous network in the presence of the timing attack. We first describe the model of anonymous communication in the presence of timing attack and then discuss how to measure the packet timing distortion between two flows that may have different number of packets. We define the effectiveness of anonymous network as the variability of the packet timing distortion it introduces and we use wavelet based energy plot to measure the variability at multiple resolutions. We demonstrate several important properties of the newly proposed wavelet based metric of the anonymous network.

3.1 Low-Latency Anonymous Network Model and Packet Timing Distortion

Figure 1 illustrates the low-latency anonymous network model we use in this work. We view the anonymous network as a black box, and we assume there is no attacker inside the black box. We only consider the incoming and outgoing flows. Specifically, we assume all the incoming and outgoing flows are encrypted and there exists no observable correlation between the content of incoming flow and outgoing flow. The incoming flow X enters the low-latency anonymous network and goes through various transforms such as repacketization (i.e. combining several pack-

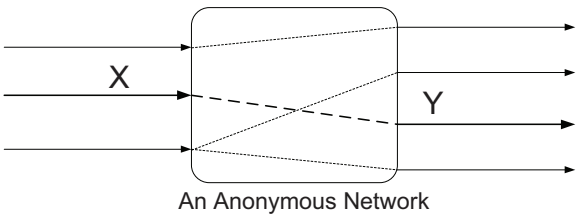

Figure 1. Low Latency Anonymous Network Model.

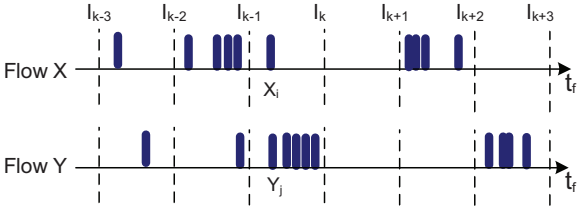

Figure 2. Packet Timing Distortion Due to Flow Transformation.

ets into one packet, splitting packet into multiple packets), flow spliting/mixing, adding chaff, packet dropping. Y is the corresponding outgoing flow of flow X. Note that one incoming flow may have multiple corresponding outgoing flows when there is flow splitting, and one outgoing flow may have multiple corresponding incoming flows when there is flow mixing. Furthermore, the anonymous network may introduce bounded timing perturbation on each packet of outgoing flows.

Despite significant transformations done by low-latency anonymous networks, previous work [32] has shown that, by transparently watermarking the flow at the packet timing domain, it is possible for the attacker to uniquely identify the transformed outgoing flows and correlate them with their corresponding incoming flows. Such timing attacks essentially exploit the fact that low-latency anonymous networks do not eliminate all the mutual information between the incoming flows and the outgoing flows in the timing domain. The less the amount of mutual information is left in the timing domain, the more resilient the anonymous network is against timing attack. Apparently the amount of mutual information in the timing domain is adversely affected by the packet timing distortion by the anonymous network.

Given a latency bound, different anonymity techniques (e.g., flow mixing, packet dropping) may generate different packet timing distortions, which in turn have different adverse impact on the mutual information in the timing domain. Therefore, it is necessary to quantitatively measure the packet timing distortion in order to study the effectiveness of low-latency anonymous networks.

3.2 Measuring Packet Timing Distortion

While there are many works on measuring the network delay jitter, they are not suitable for measuring the packet timing distortion between the original incoming flow and the anonymized outgoing flow. This is because there is no guaranteed one-one correspondence between packets of the incoming and the outgoing flows. In other words, some (e.g., dropped) packet in the incoming flow may have no corresponding packet in the outgoing flow, and some (e.g., bogus or chaff) packet in the outgoing flow may have no corresponding packet in the incoming flow. In addition, the packet timing of the incoming and outgoing flows may not be precisely synchronized. All these make it difficult to quantitatively measure the packet timing distortion between two flows.

To address these issues in measuring the packet timing distortion, we develop an interval based approach which divides a flow duration into time intervals of equal length and build the packet timing distortion metric upon the aggregated time difference between two flows at the granularity of the time interval. Such an approach allows us to quantitatively measure the packet timing distortion between two flows even if they don't have same number of packets.

Let X be the incoming flow to a low latency anonymous network, and flow Y be the corresponding outgoing flow. Due to the low latency nature, flow X and flow Y have about the same duration. We use T_f to represent the duration of both flows X and Y. Assume flow X has $n > 0$ packets $p_0^x, \ldots, p_{n-1}^x$ and flow Y has $m > 0$ packets $p_0^y, \ldots, p_{m-1}^y$. We use $t(p_i^x)$ and $t(p_i^y)$ to represent the timestamp of the i-th packet of flow X and Y respectively. Here m and n may be different.

We divide the flow duration T_f into $\lceil \frac{T_f}{T} \rceil$ time intervals of equal length $T > 0$, and use $S(i)$ to represent the start point of interval i. Apparently, packet p_i falls into interval $\lfloor \frac{t(p_i)-t(p_0)}{T} \rfloor$. We use $n(f,i)$ to denote the number of packets in interval i of flow f, and $\bar{t}(f,i)$ to denote the mean of the timestamp of packets in interval i of flow f. When $n(f,i) = 0$, we define $\bar{t}(f,i) = 0$.

For interval i ($i > 0$) of flow f, we define *fpne*(f,i) to be the index of the first, previous, non-empty interval that is before interval i. For the first interval of flow f, we define *fpne*$(f,0) = 0$.

Let

$$x(f,i) = [\bar{t}(f,i) - \bar{t}(f,\textit{fpne}(f,i))] \times n(f,i) \qquad (1)$$

We define the aggregated time difference of interval i between flow f_1 and flow f_2 to be

$$d(f_1,f_2,i) = [x(f_1,i) - x(f_2,i)] \times S(i+1) \qquad (2)$$

Note $d(f_1,f_2,i)$ could be positive, negative or zero. For example, in Figure 2, flow X has more packets than flow

Y in the $(k-2)$-th interval $[S(k-2), S(k-1))$ and *fpne*$(X, k-2)$ = *fpne*$(Y, k-2)) = k-3$. As a result, $d(X, Y, k-2) > 0$. On the other hand, $d(X, Y, k-1) < 0$ since flow X has less packets than flow Y in the $(k-1)$-th interval $[S(k-1), S(k))$ and *fpne*$(X, k-1)$ = *fpne*$(Y, k-1)) = k-2$. $d(X, Y, k) = 0$ since both flow X and flow Y have no packet in the k-th interval $[S(k), S(k+1))$.

We further define the overall packet timing distortion between flow f_1 and flow f_2 to be vector

$$D(f_1, f_2) = \langle d(f_1, f_2, 0), \ldots, d(f_1, f_2, \lceil \frac{T_f}{T} \rceil - 1) \rangle \quad (3)$$

3.3 Wavelet-Based Energy Plot

In this subsection, we analyze the variability of packet timing distortion between two flows via wavelet-based Multi Resolution Analysis (MRA). Specifically, we use the wavelet-based statistical estimator developed by Abry and Veitch [5, 11]. The wavelet-based MRA takes a sequence of data as input and transforms the sequence of data into a number of wavelet coefficients at different resolutions and a low-resolution approximation. The output of discrete wavelet transform (DWT) gives the detail coefficients (from the high-pass filter) and the approximation coefficients (from the low-pass filter). The wavelet energy plot shows the variance of the wavelet detail coefficients at all time scales.

Given flow X, flow Y and interval size $T_0 > 0$, we can obtain a packet timing distortion vector $\langle d(X, Y, 0), \ldots, d(X, Y, \lceil \frac{T_f}{T_0} \rceil - 1) \rangle$ from equation (3) and feed this vector to the wavelet-based MRA as input. Based on the input vector, the wavelet-based MRA generate a series of vectors of different scales j:

$$D(X, Y, j) = \langle d_{j,0}, \ldots, d_{j,n_j-1} \rangle \quad (4)$$

where $n_j = \lceil \frac{T_f}{T_j} \rceil$, $T_j = 2^j T_0$ $(j = 0, 1, ...)$ and $d_{j,k} = d_{j-1,2k} + d_{j-1,2k+1}$ for $j > 0$.

Let $C_{D(X,Y,j)}(p)$ be the pth $(p = 0...N_j - 1)$ wavelet detail coefficient at scale j for jth vector $D(X, Y, j)$, where $N_j = 2^{-j} n_j$ is the number of wavelet detail coefficients at scale j. The energy at time scale j is defined as the variance of the coefficients. When $E(C_{D(X,Y,j)}(p)) = 0$, the energy at scale j is

$$e_j = \frac{\sum_{p=0}^{N_j-1} [C_{D(X,Y,j)}(p)]^2}{N_j} \quad (5)$$

Here the wavelet-based MRA assumes that $D(X, Y, j)$ is covariance stationary in that 1) for a given j, the mean of $D(X, Y, j)$ is constant; and 2) the covariance between any $d_{j,k}$ and $d_{j,k'}$ only depends on $|k - k'|$.

From a linear algebra's perspective, a wavelet detail coefficient of the wavelet transform can be thought as an inner product of a high pass filter $\mathbf{g}$ (i.e. a vector of length l) and a vector $\langle d_{j,2p}, ..., d_{j,2p+l-1} \rangle$.

We first consider Haar (Daubechies 2, $l = 2$) wavelet of scale j, whose high pass filter $\mathbf{g} = \langle g_0, g_1 \rangle = \langle \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \rangle$. We have

$$\begin{aligned} C_{D(X,Y,j)}(p) &= \frac{1}{\sqrt{2^{j-1}}} \mathbf{g} \cdot \hat{\mathbf{d}}_{\mathbf{j-1}} \\ &= \frac{1}{\sqrt{2^{j-1}}} (g_0 d_{j-1,2p} + g_1 d_{j-1,2p+1}) \\ &= \frac{1}{\sqrt{2^j}} (d_{j-1,2p} - d_{j-1,2p+1}) \end{aligned} \quad (6)$$

Therefore, a Haar wavelet coefficient essentially reflects the difference between an even-numbered element and an odd-numbered element of the $(j-1)$th scale vector.

Let $\Delta d_{j-1,p} = d_{j-1,2p} - d_{j-1,2p+1}$, the energy e_j in equation (5) at scale j for the Haar wavelet becomes

$$e_j = 2^{-j} \frac{\sum_{p=0}^{N_j-1} \Delta d_{j-1,p}^2}{N_j} \quad (7)$$

Since $E(\Delta d_{j-1,p}) = 0$, the energy e_j at scale j can be thought as the variance of the data variation $\Delta d_{j-1,p}$.

Similarly, Daubechies 6 wavelet transform [11] uses a highpass filter $\mathbf{g} = \langle g_0, g_1, g_2, g_3, g_4, g_5 \rangle$. The pth D6 wavelet coefficient at scale j is

$$\begin{aligned} C_{D(X,Y,j)}(p) &= \frac{1}{\sqrt{2^{j-1}}} \mathbf{g} \cdot \hat{\mathbf{d}}_{\mathbf{j-1}} \\ &= \frac{1}{\sqrt{2^{j-1}}} (\sum_{q=0}^{5} g_q d_{j-1,2p+q}) \end{aligned} \quad (8)$$

where $\hat{\mathbf{d}}_{\mathbf{j-1}} = \langle d_{j-1,2p}, d_{j-1,2p+1}, ..., d_{j-1,2p+5} \rangle^T$. Since $\sum_{k=0}^{5} g_k = 0$[1] and $E(d_{j-1,k}) = E(d_{j-1,k'})$ for all $k \neq k'$, $E(C_{D(X,Y,j)}(p)) = 0$. The energy at scale j for the D6 wavelet becomes

$$e_j = 2^{-j+1} \frac{\sum_{p=0}^{N_j-1} (\sum_{q=0}^{5} g_q d_{j-1,2p+q})^2}{N_j} \quad (9)$$

The wavelet-based energy plot shows the logarithm of energy $log_2(e_j)$ at all time scales, which reflects the variability of the input sequence of data at different scales. The more variable the input data, the higher the energy will be. For example, constant rate data series should have the lowest energy since it has the least variability. Figure 3 shows the energy plots (the logarithm of energy) of fixed length (8192) data series of various distributions. The energy of constant data series is very close to zero ($< 2^{-28}$). In addition, the energy of linear increasing data series is also close

[1] g_0=0.035226, g_1=0.085441, g_2=-0.135011, g_3=-0.459878, g_4=0.806892, g_5=-0.332671

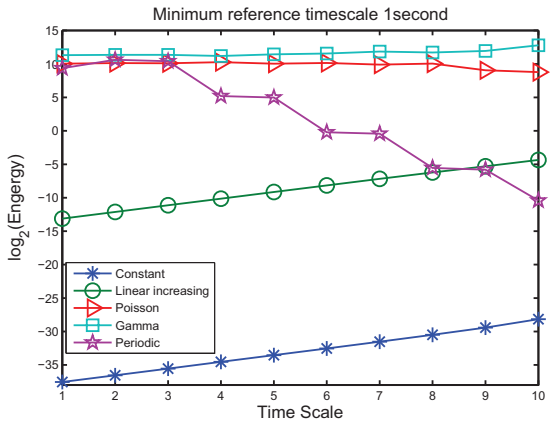

Figure 3. Energy Plots of Data Series of Various Distributions

to zero. Both Poisson and Gamma distributed data series have fixed energy, and periodic (with repeating pattern) distributed data series has decreasing energy level approaching to zero with increasing time scale.

3.4 Properties of the Energy-Based Metric

Given packet flows X, Y and interval size T_0, we can get the packet timing distortion between X and Y $D(X,Y)$ from equations (1) (2) (3). We can further get the energy of the packet timing distortion between X and Y at scale j from equations (4) (5). We use $e_j(D(X,Y))$ to denote the energy of the packet timing distortion between flows X and Y at scale j. The energy-based metric $e_j(D(X,Y))$ has the following properties

1. **Zero energy for no distortion** $e_j(D(X,X)) = 0$ for all j.

   It is easy to see that $D(X,X,j) = \langle 0,\ldots,0\rangle$ for all j. Therefore, $C_{D(X,X,j)}(p) = 0$ for all j and p.

2. **Commutativity or Symmetry** $e_j(D(X,Y)) = e_j(D(Y,X))$ for all j. From equations (2) and (3), $D(X,Y) = -D(Y,X)$. Therefore, $C_{D(X,Y,j)}(p) = -C_{D(Y,X,j)}(p)$ from equations (6) and (8). From equation (5) we have $e_j(D(X,Y)) = e_j(D(Y,X))$.

3. **Zero energy change by adding a constant to the distortion** $e_j(D(X,Y)) = e_j(D(X,Y) + \hat{c})$ where $\hat{c} = \langle c,\ldots c\rangle$ be any vector of any constant c of the same number of elements as that of $D(X,Y)$. Adding a constant vector $\hat{c}$ to the distortion vector $D(X,Y)$ is equivalent to adding some constant vector $\hat{c'} = \langle c',\ldots,c'\rangle$ to $\hat{\mathbf{d}}_{\mathbf{j-1}}$ for all j. Since $\sum_{k=0}^{l-1} g_k = 0$ for Daubechies l wavelet, we have

$$\begin{aligned} & \mathbf{g}\cdot(\hat{\mathbf{d}}_{\mathbf{j-1}} + \hat{c'}) \\ = \; & \mathbf{g}\cdot\hat{\mathbf{d}}_{\mathbf{j-1}} + \mathbf{g}\cdot\hat{c'} \\ = \; & \mathbf{g}\cdot\hat{\mathbf{d}}_{\mathbf{j-1}} \end{aligned}$$

   Therefore, the energy plot captures only the variability of the packet timing distortion and it ignores any constant changes on each element of the packet timing distortion.

4. **Constant energy plot change by multiplying the distortion by a non-zero constant** Suppose we multiply each element of $D(X,Y)$ with $a \neq 0$, then $e_j(aD(X,Y)) = a^2 e_j(D(X,Y))$ or $\log(e_j(aD(X,Y))) = 2\log(a) + \log(e_j(D(X,Y)))$. In other words, multiplying the distortion by a non-zero constant will move the energy plot up or down by a constant. This means that changing the unit of the packet timing distortion will not affect the shape of the energy plot nor will it affect the relative distance between the energy plots of different distortions of different pairs of flows at any scale.

Note while property 3) states adding constant to the distortion vector will not change the energy of the distortion, it does not mean adding constant rate packets to one flow will not change the energy of the packet timing distortion. In fact, mixing or adding constant rate packet flow to flow Y may change the energy of the packet timing distortion between X and Y in that $e_j(D(X,Y)) \neq e_j(D(X,Y+\hat{c}))$ where $\hat{c}$ is a constant rate packet flow. However, as we will show empirically in section 4, adding constant rate flow to a flow will have diminishing impact on the energy of the packet timing distortion. Specifically, $e_j(D(X,Y+\hat{c_1})) \approx e_j(D(X,Y+\hat{c_1}+\hat{c_2}))$ where $\hat{c_1}$ and $\hat{c_2}$ are constant rate packet flows.

4 Evaluation

In this section, we empirically evaluate our energy-based metric via both real time and offline experiments. The goals of this evaluation are two fold. First, we want to validate our energy-based metric and see if it really measures the practical effectiveness of low-latency anonymous network in the presence of timing attack. Second, we want to gain further insight from applying the energy-based metric to various anonymity techniques.

To validate the new metric, we conduct timing attack on four leading low-latency anonymous networks Anonymizer.com, Tor, Findnot.com and Steganos and use the timing attack results as the ground truth about the effectiveness in the presence of timing attack.

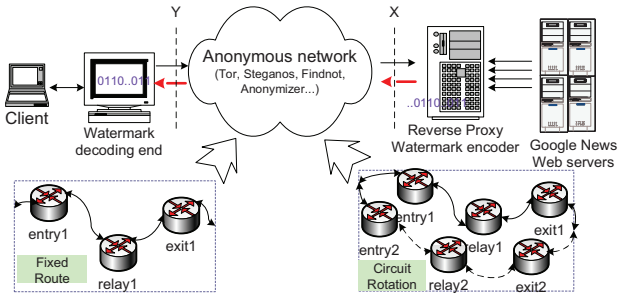

Figure 4. Timing Attack against Anonymous Web Traffic

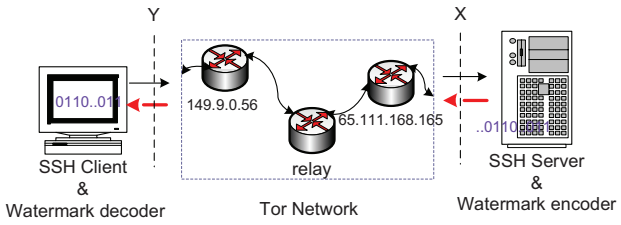

Figure 5. Timing Attack against SSH traffic Anonymized by Tor

We further evaluate, both analytically and empirically, the impacts of Tor circuit rotation and adding constant rate chaff on the effectiveness of the low-latency anonymous networks in the presence of timing attack.

4.1 Experimental Setup

Figure 4 shows the experimental setup for the timing attack on web traffic anonymized by various anonymous networks. Specifically, we have setup an Apache web server as the reverse proxy for `news.google.com` on a Dell Dimension 3000 PC running Redhat Enterprise 4 Linux of kernel 2.6.10. We have also configured the Apache web server to use non-persistent HTTP protocol so that every request and response would use a separate TCP connection. To watermark the web traffic from the Apache web server to the client, we have installed the watermark encoder at the host where the Apache reverse proxy ran. To decode the watermarked flow out of the anonymous network, we have setup a Dell Precision 390 running Redhat Enterprise 4 as a NAT router, which would route the client's web request traffic to the entry point of the chosen anonymous network and forward the web response traffic from the anonymous network to the original client.

At the client, we have manually generated the Web Traffic by clicking some URL on news.google.com once every 8 to 10 seconds. At the NAT router, We have collected the anonymized web response traffic and decoded the watermark from the the network traces collected there.

Figure 5 shows the configuration for the timing attack

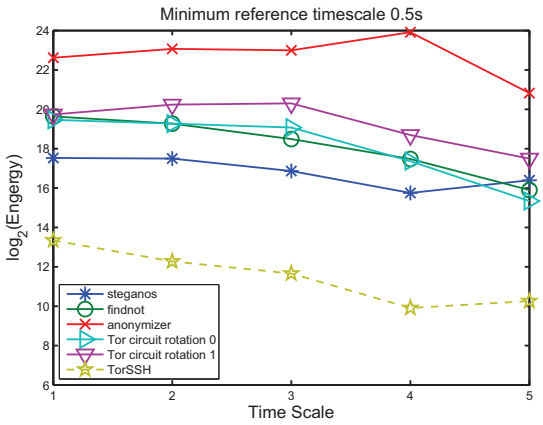

Figure 6. Energy Plots of the Timing Distortion by Various Anonymous Networks

on SSH traffic anonymized by Tor. We have setup an SSH server and the watermark encoder at one Linux machine. From the Tor network status web site [4], we have chosen to use 'croeso' with IP address 149.9.0.56 as the fixed entry node of Tor and 'illuminata' with IP address 65.111.168.165 as the fixed exit node of Tor. We have generated SSH traffic at the client by randomly typing echoing commands with average 3 characters per second, and we have decoded the watermark at the SSH client machine.

4.2 Empirical Validation of the Energy-Based Metric

We have conducted the timing attack [32] on four representative low-latency networks: Anonymizer.com [3], Tor [10], Findnot.com [1] and Steganos [2], and have used the timing attack results as the ground truth for validating the energy-based metric. The idea is to watermark the incoming flow to the anonymous network and try to recover the watermark from the corresponding outgoing (i.e. anonymized) flow. The better we can recover the watermark from the anonymized flow, the weaker the anonymous network is.

We have routed web traffic from Google news through all the above 4 anonymous networks (shown in Figure 4). We have also routed SSH traffic through Tor (shown in Figure 5). We have watermarked [32] the return web traffic and the return SSH traffic toward the web browser and the SSH client respectively with 32 bit watermark of redundancy 5 using 500ms interval size and 350ms maximum timing adjustment. This requires the to-be-watermarked flow to be at least 160 seconds long. We have collected both the incoming flow (i.e. the original flow) denoted as X and the corresponding outgoing flow (i.e. the anonymized flow) denoted as Y to and from all the anonymous networks, all of which are at least 180 seconds long.

Table 1. Average of the Logarithm of Energy E_{avg} and Watermark Decoding Success Rate of Timing Attacks of Redundancy 5 (180 Seconds) on Various Anonymous Networks

Anonymities	E_{avg}	Decoding	Packet rate(pkt/s)
Tor SSH	11.49	96.87%	4.23
Steganos	16.81	96.43%	12.77
Findnot	18.16	91.00%	13.53
Tor CR_0	18.10	87.03%	13.13
Tor CR_1	19.29	72.15%	11.57
Anonymizer	22.68	59.37%	8.01

We have set the smallest scale $T_0 = 0.5$ second, and we have measured the energy $e_j(D(X,Y))$ of the timing distortion between X and Y for all pair of traffic at scales 1 to 5. Figure 6 shows the energy plots of the timing distortion by Anonymizer.com, Tor, Findnot.com and Steganos. For web traffic, Anonymizer.com has the highest energy at all scales, Steganos has the lowest energy, Findnot has about the same energy level as Tor when it has no circuit rotation. We have found that one circuit rotation in Tor would increase the energy at all scales. This suggests that Tor circuit rotation increases the variability of the timing distortion between the original flow and the anonymized flow by Tor.

We have decoded the outgoing (i.e. anonymized) watermarked flows from various anonymous networks, Table 1 summarizes the watermarking decoding results. Anonymizer.com has the lowest watermark decoding success rate 59.37% for web traffic, which means it has the strongest resistance to timing attack among all the anonymous networks we have tested. Tor SSH has the highest watermark decoding success rate 96.87%, which means it is very weak in front of timing attack. Column E_{avg} in Table 1 shows the average of the logarithm of energy in all scales in Figure 6, and it is clearly negatively correlated with the watermarking decoding success rate. This confirms the intuition that the higher the energy of the timing distortion by a anonymous network, the stronger resistance it has against the timing attack. This indicates that our energy-based metric can indeed measure the practical effectiveness of low-latency anonymous network in the presence of timing attack.

4.3 Impact of Tor Circuit Rotation

As shown in the left bottom part of Figure 4, a Tor circuit consists of 3 nodes: the entry, the relay and the exit nodes. Normally, the Tor circuit is fixed for web traffic. However, Tor network could have circuit rotation due to changes of network condition (e.g., overloading of one Tor node). As

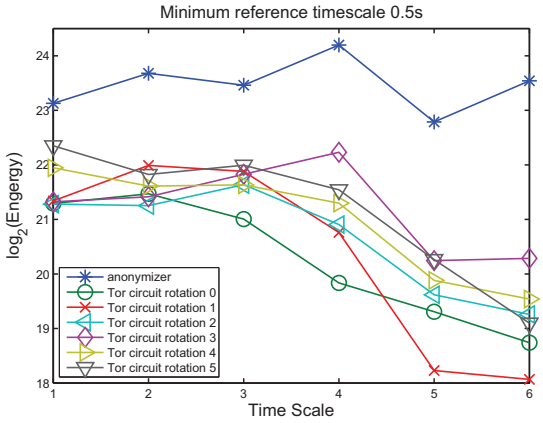

Figure 7. Energy Plots of the Timing Distortion by Anonymizer.com and Tor with 0 to 5 Circuit Rotations

shown in Figure 6, circuit rotation in Tor would increase the energy of the timing distortion of Tor at all scales. It is worthwhile to study the impact of circuit rotation of Tor on the effectiveness of Tor in the presence of timing attack.

We have conducted real-time experiments on Tor and have manually triggered 1 to 5 circuit rotations by changing the relay node in different runs. Specifically, we have randomly select the multiple Tor nodes from the lists shown in the real-time Tor network status monitor [4] as the new relay node for the circuit rotation. The following Tor debug information shows one instance of 5 Tor circuit rotations happened in a duration of 420 seconds (16:39:08-16:46:08)

16:39:08.072 [info] connection_ap_handshake_send_begin():
Address/port sent, ap socket 11, n_circ_ id 19098 # Initial Circuit
... ...
16:43:58.104 [info] connection_ap_handshake_send_begin():
Address/port sent, ap socket 15, n_circ_ id 6172
16:45:28.117 [info] connection_ap_handshake_send_begin():
Address/port sent, ap socket 16, n_circ_ id 6173 # Final Circuit

We have watermarked the incoming flows to Tor and Anonymizer.com with 32 bit watermark of redundancy 10, 350ms maximum timing adjustment and 500ms interval size. We have collected both the incoming flow (i.e. the original flow) denoted as X and the corresponding outgoing flow (i.e. the anonymized flow) denoted as Y to and from Tor in our experiments with 0 to 5 circuit rotations, all of which are at least 350 seconds long to accommodate up to 5 Tor circuit rotations.

Table 2 shows the number of packets of both the incoming and outgoing flows to and from Tor. It shows Tor will normally decrease the number of packets when it anonymizes a packet flow.

Using minimum time scale $T_0 = 0.5$ second, we have measured the energy $e_j(D(X,Y))$ of the timing distortions

Table 2. Number of Packets of Flows in Tor Circuit Rotation Experiments

Circuit Rotations	Incoming flow X	Outgoing flow Y
Tor CR_0	8350	7926
Tor CR_1	7919	7891
Tor CR_2	6583	6472
Tor CR_3	6052	6055
Tor CR_4	5904	5328
Tor CR_5	6105	5469

Table 3. Average of the Logarithm of Energy E_{avg} and Watermark Decoding Success Rate of Timing Attacks of Redundancy 10 (180 Seconds) on Tor with 0-5 Circuit Rotations

Anonymities	E_{avg}	Decoding	Packet rate(pkt/s)
Tor CR_0	20.17	91.25%	24.01
Tor CR_1	20.37	85.93%	21.25
Tor CR_2	20.66	85.1%	18.66
Tor CR_3	21.22	82.14%	16.85
Tor CR_4	20.98	75.0%	15.16
Tor CR_5	21.17	72.91%	15.66
Anonymizer	23.47	69.7%	8.83

of Tor with 0-5 circuit rotations. Figure 7 shows the energy plots for the timing distortion by anonymizer.com and Tor with different circuit rotations. It clearly shows that circuit rotation in Tor will increase the energy at all scales. However, Anonymizer still have the highest energy compared with all Tor circuit rotations.

We have decoded the outgoing flows from Anonymizer.com and Tor with 0 to 5 circuit rotations with redundancy 10. Table 3 shows the watermark decoding success rate and the average of the logarithm of energy at all scales. The average logarithm of energy and the watermark success rate in Table 3 are generally negatively correlated.

When there is no circuit rotation, Tor is very weak in front the timing attack [32]. For example, time attack of redundancy 10 [3] can achieve 91.25% watermark bit decoding success rate on the web traffic anonymized by Tor without circuit rotation. However, each additional Tor circuit rotation will decrease the watermark decoding success rate – making Tor more resistent to timing attack. 5 Tor circuit rotations can decrease the watermark bit decoding success rate of the same timing attack to 72.91%. At the same time, Tor circuit rotation tends to increase the energy of the timing distortion by Tor. We have noticed that Tor circuit rotation will introduce long idle period in the outgoing

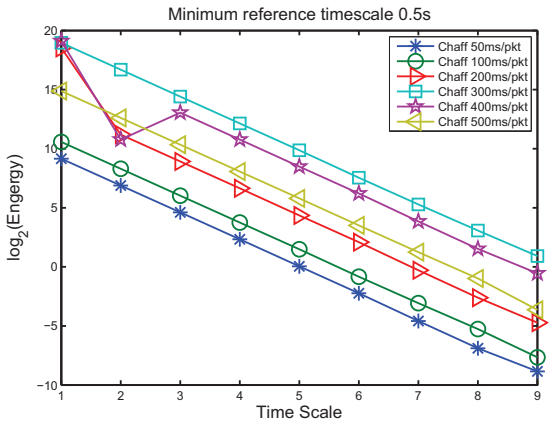

Figure 8. Energy Plots of the Timing Distortions by Adding Constant Rates Chaff to a Synthetic Constant Rate Flow

flow, which makes the anonymized flow more bursty thus have higher energy. However, Anonymizer.com still has the lowest watermark decoding success rate than Tor with 5 circuit rotations. This timing attack result is consistent with the energy plot shown in Figure 7.

In summary, Tor circuit rotation increases Tor's resistance to timing attack at the cost of more timing disturbances to Tor user. The wavelet-based energy plot can accurately model the impact of Tor circuit rotation.

4.4 Impact of Adding Constant Rate Chaff

Adding cover traffic or chaff has long been believed an effective way to anonymize network flows, and a number of proposed anonymous systems [18, 16, 13] are based on adding cover traffic. However, previous research [14] suggested that adding constant rate chaff is not optimal in achieving anonymity. Recent work by Wang et al. [32] has proved both analytically and empirically that adding cover traffic has fundamental limitations in anonymizing sufficiently long flows. Specifically, adding constant rate chaff has been shown to have neglectable impact on the timing attack developed by Wang et al. [32].

We have evaluated the impact of adding constant rate chaff on our wavelet-based energy metric using both synthetic flow and real flows.

We have generated a 3000 seconds long synthetic flow X of constant rate of one packet every 300ms, then we have add flow X with chaff of different constant rates: one chaff packet every 100, 200, 300, 400 and 500ms to the chaffed flow Y. Using the same minimum scale $T_0 = 0.5$ second, we have measured the energy of the timing distortions between synthetic flow X and the chaffed flows Y of different

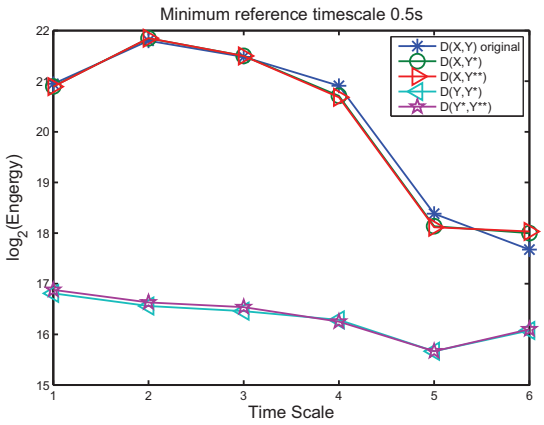

Figure 9. Energy Plots of the Timing Distortions by Adding Constant Rate Chaff to Real Flows

chaff rates. Figure 8 shows the corresponding energy plots of the timing distortions caused by adding different rates of chaff. It clearly shows that adding constant rate chaff to constant rate flow only introduces timing distortion of approaching to zero energy.

To understand the impact of adding constant rate chaff to non-constant rate flow, we have used real flows collected from experiments on Tor. Specifically, X is a 350-second long incoming flow to Tor and Y is the corresponding 350-second long outgoing flow from Tor. Here X is not watermarked and we have not manually triggered any Tor circuit rotation. We have added one chaff packet every 50ms to Y to get the chaffed flow Y^*. As a result, flow Y^* has 14212 packets while the original Y has 7213 packets. In other words, the chaff added is about 100% of the original packets. We have further added one chaff packet every 99ms to Y^* to get Y^{**}.

Figure 9 shows the measured logarithm of energy of the timing distortions between X and Y, X and Y^*, X and Y^{**}, Y and Y^*, Y and Y^{**}. Despite significant amount of chaffed added, $e_j(D(X,Y)) \approx e_j(D(X,Y^*)) \approx e_j(D(X,Y^{**}))$ and $e_j(D(Y,Y^*)) \approx e_j(D(Y,Y^{**}))$ at all scales j. These analytical results are consistent with the findings of previous works [32, 33] that adding constant rate chaff does not increase the resistance to the timing attack. Therefore, our anonymity metric can model the diminishing effect of adding constant rate chaff to anonymous network.

5 Conclusion

All practical low-latency anonymous networks are susceptible to timing attack due the timing constraint imposed by the low-latency requirement. A quantitative metric is needed in order to understand and analyze the negative impact of timing attack on low-latency anonymous networks.

The key contribution of this paper is that we have developed a wavelet-based energy metric that can quantitatively measure the practical effectiveness of all low-latency anonymous networks in the presence of timing attack. Our metric is based on a novel measurement of the timing distortion of the packet timing between two flows that may have different number of packets, and it uses the wavelet-based energy of the timing distortion to represent the practical effectiveness of the low-latency anonymous network in the presence of timing attack.

We have validated our wavelet-based energy metric with real-time experiments on leading low-latency anonymous networks Anonymizer.com, Tor, Findnot.com and Steganos, and we have used the timing attack results as the ground truth. We have found strong correlation between the timing attack results and the metric measurements. We have further found that the circuit rotation in Tor network could substantially increase Tor's resistance to the timing attack at the cost of more timing disturbances to the normal users. Our wavelet-based energy metric has also confirmed that adding constant rate chaff (i.e. cover traffic) has diminishing effect in anonymizing network flows.

In our future work, we plan to use the newly developed wavelet-based energy metric to systematically analyze various anonymity techniques such as batch, flow transformation and seek more effective anonymity techniques against the timing attack.

Acknowledgments

The authors would like to thank the anonymous reviewers for their insightful comments, Daryl Veitch for making the wavelet-based MRA code available for generating energy plots. This work was partially supported by NSF Grants CNS-0524286 and CT-0627493.

References

[1] Findnot. http://findnot.com.
[2] Steganos. http://www.steganos.com.
[3] The Anonymizer. http://anonymizer.com.
[4] Tor Network Status. http://torstatus.kgprog.com.
[5] P. Abry and D. Weitch. Wavelet Analysis of Long-Range Dependent Traffic. *IEEE Transactions on Information Theory*, 44(1):2–15, January 1998.
[6] O. Berthold, H. Federrath, and S. Köpsell. Web MIXes: A System for Anonymous and Unobservable Internet Access. In *Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*, pages 115–129, July 2000.
[7] D. Chaum. Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms. *Communications of the ACM*, 24(2):84–88, 1981.

[8] G. Danezis. The Traffic Analysis of Continuous-Time Mixes. In *Proceedings of the 4th International Workshop of Privacy Enhancing Technologies workshop (PET 2004)*, pages 35–50, Toronto, Canada, May 2004.

[9] C. Diaz, S. Seys, J. Claessens, and B. Preneel. Towards Measuring Anonymity. In *Proceedings of the 2nd International Workshop of Privacy Enhancing Technologies Workshop (PET 2002)*, pages 54–68, San Francisco, CA, USA, 2002. Springer-Verlag.

[10] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The Second-Generation Onion Routing. In *Proceedings of the 13th USENIX Security Symposium*, pages 303–320, San Diego, CA, USA, August 2004. USENIX.

[11] D.Weitch. Code for the Estimation of Scaling Exponents. *http://www.cubinlab.ee.mu.oz.au/ darry*.

[12] E. Felton and M. Schneider. Timing Attacks on Web Privacy. In *Proceedings of the 7th ACM Conference on Computer and Communications Security (CCS 2000)*, Athens, Greece, November 2000.

[13] M. J. Freedman and R. Morris. Tarzan: A Peer-to-Peer Anonymizing Network Layer. In *Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS 2002)*, pages 193–206, Washington, DC, November 2002.

[14] X. Fu, B. Graham, R. Bettati, and W. Zhao. On Countermeasures to Traffic Analysis Attacks. In *Proceedings of the 2003 IEEE Workshop on Information Assurance*, pages 188–195, West Point, NY, USA, June 2003.

[15] M. Gogolewski, M. Klonowski, and M. Kutylowski. Local View Attack on Anonymous Communication. In *Proceedings of the 10th European Symposium On Research In Computer Security (ESORICS 2005)*, September 2005.

[16] Y. Guan, X. Fu, D. Xuan, P. U. Shenoy, R. Bettati, and WeiZhao. Netcamo: Camouflaging Network Traffic for QoS-Guaranteed Mission Critical Applications. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans, Special Issue on Information Assurance*, 31(4):253–265, 2001.

[17] Y. Guan, X. Fu, D. Xuan, P. U. Shenoy, R. Bettati, and W. Zhao. Netcamo: Camouflaging network traffic for qos-guaranteed. *IEEE Transactions on Systems, Man, and Cybernetics*, 34(4):253–265, July 2001.

[18] Y. Guan, C. Li, D. Xuan, R. Bettati, and W. Zhao. Preventing Traffic Analysis for Real-Time Communication Networks. In *Proceedings of Military Communications Conference (MILCOM 1999)*, pages 744–750, November 1999.

[19] N. Hopper, E. Y. Vasserman, and E. Chan-Tin. How Much Anonymity does Nework Latency Leak. In *Proceedings of the 14th ACM Conference on Computer and Communications Security (CCS 2007)*, pages 82–91, Alexandra, Virginia, October 2007.

[20] B. N. Levine, M. K. Reiter, C. Wang, and M. K. Wright. Timing Attacks in Low-Latency Mix-Based Systems. In A. Juels, editor, *Proceedings of Financial Cryptography (FC '04)*, pages 251–265. Springer-Verlag, LNCS 3110, February 2004.

[21] S. J. Murdoch and G. Danezis. Low-cost Traffic Analysis of Tor. In *Proceedings of the 2005 IEEE Symposium on Security and Privacy (S&P 2005)*, pages 183–195, May 2005.

[22] S. J. Murdoch and P. Zielinski. Sampled Traffic Analysis by Internet-Exchange-Level Adversaries. In *Proceedings of the 7th International Symposium of Privacy Enhancing Technologies workshop (PET 2007)*, pages 167–183, Ottawa, Canada, June 2007.

[23] Y. J. Pyun, Y. H. Park, X. Wang, D. S. Reeves, and P. Ning. Tracing Traffic through Intermediate Hosts that Repacketize Flows. In *Proceedings of the 26th Annual IEEE Conference on Computer Communications (Infocom 2007)*, Anchorage, Alaska, USA, May 2007. IEEE.

[24] M. G. Reed, P. F. Syverson, and D. M. Goldschlag. Anonymous Connections and Onion Routing. *IEEE JSAC Copyright and Privacy Protection*, 16(4):482–494, 1998.

[25] M. Reiter and A. Rubin. Crowds: Anonymity for web transactions. *ACM TISSEC*, 1(1):66–92, 1998.

[26] A. Serjantov and G. Danezis. Towards an Information Theoretic Metric for Anonymity. In *Proceedings of the 2nd International Workshop of Privacy Enhancing Technologies workshop (PET 2002)*, pages 41–53, San Francisco, CA, USA, 2002. Springer-Verlag.

[27] R. Sherwood, B. Bhattacharjee, and A. Srinivasan. P5: A protocol for Scalable Anonymous Communication. In *Proceedings of 2002 IEEE Symposium on Security and Privacy (S&P 2002)*, Oakland, California, USA, May 2002.

[28] C. Shields and B. N. Levine. A Protocol for Anonymous Communication over the Internet. In *Proceedings of the 7th ACM Conference on Computer and Communications Security (CCS 2000)*, pages 33–42, Athens, Greece, November 2000.

[29] V. Shmatikov and M.-H. Wang. Timing Analysis in Low-Latency Mix Networks: Attacks and Defenses. In *Proceedings of the 11th European Symposium On Research In Computer Security (ESORICS 2006)*, Hamburg, Germany, September 2006.

[30] Q. Sun, D. R. Simon, Y. Wang, W. Russell, V. N. Padmanabhan, and L. Qiu. Statistical Identification of Encrypted Web Browsing Traffic. In *Proceedings of the 2002 IEEE Symposium on Security and Privacy (S&P 2002)*, Berkeley, California, USA, May 2002.

[31] X. Wang, S. Chen, and S. Jajodia. Tracking Anonymous Peer-to-Peer VoIP Calls on the Internet. In *Proceedings of the 12th ACM Conference on Computer and Communications Security (CCS 2005)*, pages 81–91, Alexandra, VA, November 2005. ACM.

[32] X. Wang, S. Chen, and S. Jajodia. Network Flow Watermarking Attack on Low-Latency Anonymous Communication Systems. In *Proceedings of the 2007 IEEE Symposium on Security & Privacy (S&P 2007)*, pages 116–130, Oakland, CA, May 2007.

[33] W. Yu, X. Fu, S. Graham, D. Xuan, and W. Zhao. DSSS-Based Flow Marking Technique for Invisible Traceback. In *Proceedings of 2007 IEEE Symposium on Security and Privacy (S&P 2007)*, Oakland, California, USA, May 2007.

[34] Y. Zhu and R. Bettati. Anonymity vs. Information Leakage in Anonymity Systems. In *Proceedings of the 25th IEEE International Conference on Distributed Computing Systems (ICDCS 2005)*, pages 514–524, Washington, DC, USA, June 2005.