Inferring Internet Worm Temporal Characteristics

Qian Wang[1], Zesheng Chen[1], Kia Makki[1], Niki Pissinou[1], and Chao Chen[2]
[1]Department of Electrical & Computer Engineering
Florida International University
Miami, FL 33174
E-mails: {qian.wang, zchen, makkik, pissinou}@fiu.edu

[2]Department of Engineering
Indiana University - Purdue University Fort Wayne
Fort Wayne, IN 46805
E-mail: chen@engr.ipfw.edu

***Abstract*— Internet worm attacks pose a significant threat to network security. In this work, we coin the term *Internet worm tomography* as inferring the characteristics of Internet worms from the observations of Darknet or network telescopes that are routable but unused IP addresses. Under the framework of Internet worm tomography, we attempt to infer worm temporal behaviors such as the host infection time and the worm infection sequence, and thus pinpoint patient zero. Specifically, we introduce statistical estimation techniques and propose method of moments, maximum likelihood, and linear regression estimators. We show analytically and empirically that our proposed estimators can better infer worm temporal characteristics than a naive estimator that has been used in the previous work.**

I. Introduction

Since Code Red and Nimda worms were released in 2001, epidemic-style attacks have caused enormous damages. Internet worms can spread so rapidly that existing defense systems cannot respond until they have infected most vulnerable hosts. For example, the Slammer worm infected more than 90% of vulnerable machines within 10 minutes on January 25th, 2003 [15]. Therefore, worm attacks present a significant threat to the Internet.

To counteract these notorious epidemic-style attacks, many detection and defense strategies have been studied in recent years. According to where the detectors are located, these strategies can generally be classified into three categories: *source detection and defense*, locating infected hosts in the local networks [17], [11]; *middle detection and defense*, revealing the appearance of worms by analyzing the traffic going through routers [19], [8], [13]; and *destination detection and defense*, monitoring unwanted traffic arriving at *Darknet or network telescopes*, a globally routable address space where no active services or servers reside [1], [2], [3], [4], [5]. There are two types of Darknet: *active Darknet* that responds to malicious scans to elicit the payloads of the attacks [3], [4], and *passive Darknet* that observes unwanted traffic passively [2], [5].

In this work, we focus on the destination detection and defense. Specifically, we study the problem of inferring the characteristics of Internet worms from Darknet observations. We refer to such a problem as *Internet worm tomography*, as illustrated in Fig. 1. Most worms use scan-based methods to find targets and have to guess the IP addresses of destinations. Thus, Darknet can observe partial scans from infected hosts. Combined with the worm propagation model and the statistical model, the Darknet observations can be used to detect the worm appearance [18], [6] and infer the worm characteristics (*e.g.*, infection rate [21], number of infected hosts [14], and worm infection sequence [12], [16]).

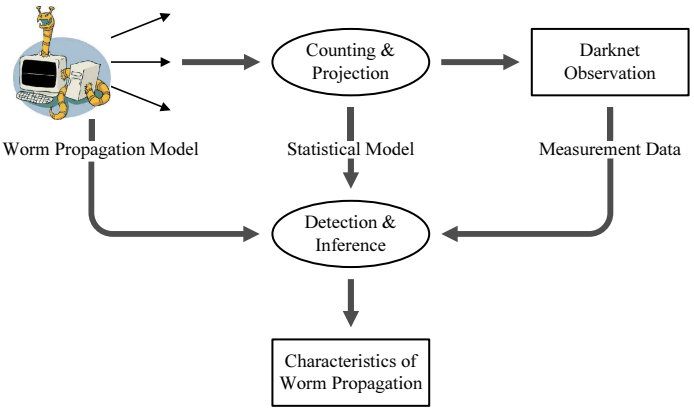


Fig. 1. Internet worm tomography.

Internet worm tomography is named after *network tomography*, and bears similarities to it. Network tomography infers the characteristics of the internal network (*e.g.*, link loss rate, link delay, and topology) through the observations from end systems [7], [10]. Network tomography can be formulated as a linear inverse problem. Internet worm tomography, however, cannot be translated into the linear inverse problem due to the specific properties of worm propagation, and thus presents new challenges.

Several works have applied the Internet worm tomography to infer worm temporal behaviors. For example, Moore *et al.* apply network telescope observations to compute the likelihood of the start times of infection events being in particular ranges around the observed times [14]. Kumar *et al.* use network telescope data and analyze the pseudo-random number generator to reconstruct the "who infected whom" infection tree of the Witty worm [12]. Rajab *et al.* use the same data and study the "infection and detection times" to infer the worm infection sequence [16]. The previous works, however, have not applied statistical estimation techniques to Internet worm tomography to further improve the inference accuracy. For example, a simple, naive estimator is used in [16] to infer the host infection time and the worm infection sequence.

The goal of this paper is to infer the Internet worm temporal characteristics accurately under the framework of Internet worm tomography by applying the statistical estimation techniques. Specifically, we attempt to answer the following questions:

- *Host infection time:* When exactly does a specific host get infected? We propose method of moments, maximum likelihood, and linear regression statistical estimators to infer the infection time. We show analytically and empirically that the mean squared error of our proposed estimators can be almost half of that of the naive estimator.
- *Worm infection sequence:* What is the order that hosts are infected by worm propagation? The key idea behind reconstructing the temporal infection sequence is to use our proposed estimators to infer the infection time of each infected host accurately. Our simulation results show that our algorithms can pinpoint patient zero and perform much better than the algorithm proposed in [16].

The remainder of this paper is organized as follows. Section II introduces estimators for inferring the host infection time. Section III presents our algorithms in estimating the worm infection sequence. Section IV gives simulation results. Finally, Section V concludes the paper.

II. Estimating the Host Infection Time

We use the Darknet observations to estimate when a host gets infected. As shown by Fig. 2, a host is infected at time t_0. Darknet can observe some scans from this host and record scan arrival times or hit times $t_1, t_2, \cdots, t_n$, where n is the number of hit events for an infected host. Here, we use *hit* to denote the event that a worm scan hits the Darknet. To estimate the host infection time t_0, we assume that there is no packet loss in the Internet. We also assume that an infected host uses the actual source IP address and does not apply IP spoofing, which is the case for TCP worms. Moreover, we assume that the infected host uses a constant scanning rate. We set s as the scanning rate or the number of scans sent by an infected host per time unit. In this paper, we focus on random scanning that selects targets randomly. Our estimation techniques, however, can be potentially extended to other scanning methods such as localized scanning [9].

The problem of estimating the host infection time can then be stated as follows: Given the Darknet observations $t_1, t_2, \cdots, t_n$, what is the best estimate of t_0? To study this problem, we consider a discrete-time system. Since a worm scans the IPv4 address space with Ω addresses (*i.e.,* $\Omega = 2^{32}$) and Darknet monitors ω addresses, the probability for a worm scan to hit the Darknet is ω/Ω. Thus, the probability of a hit event in the discrete-time system or Darknet observing at least one scan from the same infected host in a time unit is

$$\Pr(\text{hit event}) = 1 - \left(1 - \frac{\omega}{\Omega}\right)^s = p. \tag{1}$$

Denote δ_0 as the time interval between when a host gets infected and when Darknet observes the first scan from this host, *i.e.,* $\delta_0 = t_1 - t_0$. Denote δ_i as the time interval between i-th hit and $(i+1)$-th hit on Darknet, *i.e.,* $\delta_i = t_{i+1} - t_i$, $i \geq 1$. Thus, $\delta_0, \delta_1, \cdots, \delta_{n-1}$ are independent and identical distributed (i.i.d.) and follow a geometric distribution with parameter p, *i.e.,*

$$\Pr(\delta = k) = p \cdot (1-p)^{k-1}, k = 1, 2, 3, \cdots \tag{2}$$

$$\mathrm{E}(\delta) = \frac{1}{p} = \mu, \qquad \mathrm{Var}(\delta) = \frac{1-p}{p^2}. \tag{3}$$

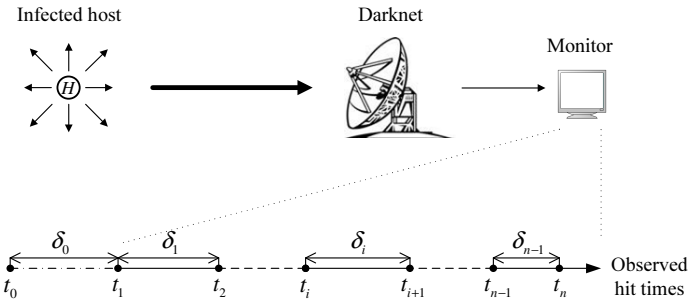


Fig. 2. An illustration of Darknet observations.

TABLE I
Notations used in this paper.

Notations	Definition
Ω	Size of the scanning space ($\Omega = 2^{32}$)
ω	Size of the Darknet
s	Scanning rate (scans/time unit)
p	Probability that at least one scan from the same infected host hits the Darknet in a time unit
t_0	Host infection time
$\hat{t_0}$	Estimated infection time
t_i	Discrete time tick when the infected host hits the Darknet for the i-th time ($i \geq 1$)
δ_i	Time interval between two consecutive hits of the Darknet ($\delta_i = t_{i+1} - t_i$, $i \geq 1$)
μ	Mean of δ
$\hat{\mu}$	Estimation of μ

Denote μ as the mean value of δ and $\hat{\mu}$ as the estimate of μ. We then estimate t_0 by subtracting $\hat{\mu}$ from t_1, *i.e.,*

$$\hat{t_0} = t_1 - \hat{\mu}. \tag{4}$$

Therefore, our problem is reduced to estimate μ. Table I summarizes the notations used in this paper.

A. Naive Estimator

Since δ follows the geometric distribution as described by (2), $\Pr(\delta)$ is maximized when $\delta = 1$. Then, a *naive estimator* (NE) of μ is

$$\hat{\mu}_{\text{NE}} = 1. \tag{5}$$

Thus, the NE of t_0 is

$$\hat{t_{0_{\text{NE}}}} = t_1 - \hat{\mu}_{\text{NE}} = t_1 - 1. \tag{6}$$

Note that $\hat{t_{0_{\text{NE}}}}$ depends only on t_1, ignoring $t_2, t_3, \cdots, t_n$. This estimator has been used in [16] to infer the host infection time and the worm infection sequence. In this paper, however, we consider more advanced estimation methods.

B. Method of Moments Estimator

Since $\mathrm{E}(\delta) = \mu$, we then design a *method of moments estimator* (MME), *i.e.,*

$$\hat{\mu}_{\text{MME}} = \overline{\delta} = \frac{1}{n-1}\sum_{i=1}^{n-1}\delta_i = \frac{t_n - t_1}{n-1}. \tag{7}$$

Thus, the MME of t_0 is

$$\hat{t_{0_{\text{MME}}}} = t_1 - \hat{\mu}_{\text{MME}} = t_1 - \frac{t_n - t_1}{n-1}. \tag{8}$$

Note that $\hat{t_{0_{\text{MME}}}}$ is not only related to t_1, but also n and t_n.

C. Maximum Likelihood Estimator

Rewrite the probability mass function of δ with respect to μ

$$\Pr(\delta;\mu) = \frac{1}{\mu}\left(1-\frac{1}{\mu}\right)^{\delta-1}, \delta = 1,2,3,\cdots. \tag{9}$$

Since $\delta_1, \delta_2, \cdots, \delta_{n-1}$ are i.i.d., the likelihood function is given by the following product

$$\begin{aligned} \mathrm{L}(\mu) &= \prod_{i=1}^{n-1} \Pr(\delta_i;\mu) && (10)\\ &= \left(\frac{1}{\mu}\right)^{n-1}\left(1-\frac{1}{\mu}\right)^{(\sum_{i=1}^{n-1}\delta_i)-(n-1)}. && (11)\end{aligned}$$

We then design a *maximum likelihood estimator* (MLE), *i.e.,*

$$\hat{\mu}_{\text{MLE}} = \arg\max_{\mu} \mathrm{L}(\mu). \tag{12}$$

Instead of maximizing $\mathrm{L}(\mu)$, we maximize $ln\mathrm{L}(\mu)$. That is,

$$\frac{d}{d\mu} ln\mathrm{L}(\mu) = 0 \Longrightarrow \hat{\mu}_{\text{MLE}} = \frac{1}{n-1}\sum_{i=1}^{n-1}\delta_i = \frac{t_n - t_1}{n-1}, \tag{13}$$

which has the same expression as the MME. Thus,

$$\hat{t_{0_{\text{MLE}}}} = t_1 - \hat{\mu}_{\text{MLE}} = t_1 - \frac{t_n - t_1}{n-1}. \tag{14}$$

D. Linear Regression Estimator

Under the assumption that the scanning rate of an individual infected host keeps constant over time, the relationship between t_i and i can be described by a linear regression model as illustrated in Fig. 3, *i.e.,*

$$t_i = \alpha + \beta \cdot i + \varepsilon_i, \tag{15}$$

where α and β are coefficients, and ε_i is the error term. To fit the observation data, we apply the least squares method. That is, we choose the coefficients that minimize the residual sum of squares (RSS)

$$\text{RSS} = \sum_{i=1}^{n}[t_i - (\alpha + \beta \cdot i)]^2. \tag{16}$$

The minimum RSS occurs when the partial derivatives with respect to the coefficients are zero

$$\begin{cases} \dfrac{\partial \text{RSS}}{\partial \alpha} = -2\sum_{i=1}^{n}(t_i - \alpha - \beta\cdot i) = 0 \\ \dfrac{\partial \text{RSS}}{\partial \beta} = -2\sum_{i=1}^{n} i\cdot(t_i - \alpha - \beta\cdot i) = 0, \end{cases} \tag{17}$$

which leads to

$$\begin{cases} \hat{\alpha} = \bar{t} - \hat{\beta}\cdot\bar{i} \\ \hat{\beta} = \dfrac{\overline{i\cdot t} - \bar{i}\cdot\bar{t}}{\overline{i^2} - (\bar{i})^2}, \end{cases} \tag{18}$$

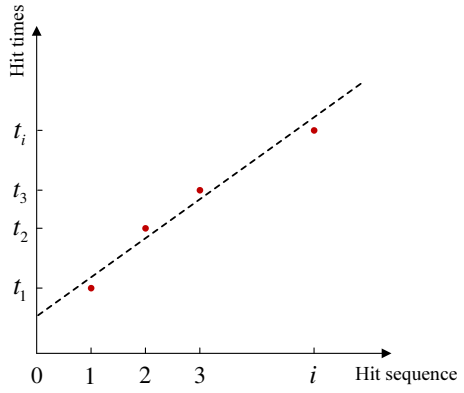


Fig. 3. Linear regression model.

TABLE II
COMPARISON OF ESTIMATOR PROPERTIES ($\hat{\mu}$).

$\hat{\mu}$	Bias($\hat{\mu}$)	Var($\hat{\mu}$)	MSE($\hat{\mu}$)
$\hat{\mu}_{\text{NE}} = 1$	$1-\frac{1}{p}$	0	$\frac{(1-p)^2}{p^2}$
$\hat{\mu}_{\text{MME}} = \hat{\mu}_{\text{MLE}} = \frac{t_n - t_1}{n-1}$	0	$\frac{1-p}{p^2(n-1)}$	$\frac{1-p}{p^2(n-1)}$
$\hat{\mu}_{\text{LRE}} = \frac{\overline{i\cdot t} - \bar{i}\cdot\bar{t}}{\overline{i^2}-(\bar{i})^2}$	0	$\frac{6(n^2+1)(1-p)}{5n(n^2-1)p^2}$	$\frac{6(n^2+1)(1-p)}{5n(n^2-1)p^2}$

where the bar symbols denote the average values

$$\begin{cases} \bar{i} = \dfrac{1}{n}\sum_{i=1}^{n} i, \quad \overline{i^2} = \dfrac{1}{n}\sum_{i=1}^{n} i^2 \\ \bar{t} = \dfrac{1}{n}\sum_{i=1}^{n} t_i, \quad \overline{i\cdot t} = \dfrac{1}{n}\sum_{i=1}^{n} i\cdot t_i. \end{cases} \tag{19}$$

We then design a *linear regression estimator* (LRE), *i.e.,*

$$\hat{\mu}_{\text{LRE}} = \hat{\beta} = \hat{t_1} - \hat{t_0}. \tag{20}$$

Thus, the LRE of t_0 is

$$\hat{t_{0_{\text{LRE}}}} = t_1 - \hat{\mu}_{\text{LRE}} = t_1 - \frac{\overline{i\cdot t} - \bar{i}\cdot\bar{t}}{\overline{i^2} - (\bar{i})^2}. \tag{21}$$

There is another way to estimate t_0, which uses the point of interception shown in Fig. 3 as the estimation of t_0, *i.e.,*

$$\hat{t_{0_{\text{LRE}}}}' = \hat{\alpha} = \bar{t} - \hat{\mu}_{\text{LRE}}\cdot\bar{i}. \tag{22}$$

However, we find that the mean squared error of $\hat{t_{0_{\text{LRE}}}}'$ increases when n increases, which makes this estimator undesirable.

E. Comparison of Estimators

To compare the performance of these estimators, we compute the bias, the variance, and the mean squared error (MSE). For example, for estimating μ,

$$\begin{cases} \text{Bias}(\hat{\mu}) &= \mathrm{E}(\hat{\mu}) - \mu \\ \text{Var}(\hat{\mu}) &= \mathrm{E}[(\hat{\mu} - \mathrm{E}(\hat{\mu}))^2] \\ \text{MSE}(\hat{\mu}) &= \mathrm{E}[(\hat{\mu}-\mu)^2] = \text{Bias}^2(\hat{\mu}) + \text{Var}(\hat{\mu}). \end{cases} \tag{23}$$

Table II summarizes the results of NE, MME (or MLE), and LRE for estimating μ[1]. It is noted that MME and LRE are

[1]Due to page limitation, we only give final results and omit the derivations.

TABLE III

COMPARISON OF ESTIMATOR PROPERTIES ($\hat{t_0}$).

$\hat{t_0}$	Bias($\hat{t_0}$)	Var($\hat{t_0}$)	MSE($\hat{t_0}$)	
$\hat{t_{0\text{NE}}} = t_1 - \hat{\mu}_{\text{NE}}$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$	$\frac{(1-p)(2-p)}{p^2}$	$(\approx \frac{2(1-p)}{p^2}, p \ll 1)$
$\hat{t_{0\text{MME}}} = \hat{t_{0\text{MLE}}} = t_1 - \hat{\mu}_{\text{MME}}$	0	$\frac{1-p}{p^2} \cdot \frac{n}{n-1}$	$\frac{1-p}{p^2} \cdot \frac{n}{n-1}$	$(\approx \frac{1-p}{p^2}, n \gg 1)$
$\hat{t_{0\text{LRE}}} = t_1 - \hat{\mu}_{\text{LRE}}$	0	$\frac{1-p}{p^2} \cdot \frac{5n^3+6n^2-5n+6}{5n(n^2-1)}$	$\frac{1-p}{p^2} \cdot \frac{5n^3+6n^2-5n+6}{5n(n^2-1)}$	$(\approx \frac{1-p}{p^2}, n \gg 1)$

unbiased, while NE is biased. Moreover, MME and LRE have smaller MSE than NE if $n > 2$ and $p < 0.5$. Specifically, when $n \to \infty$, MSE($\hat{\mu}_{\text{MME}}$) $\to 0$ and MSE($\hat{\mu}_{\text{LRE}}$) $\to 0$. It is observed that MME is slightly better than LRE in terms of MSE.

Similarly, we compute the bias, the variance, and the MSE of the estimators for estimating t_0 in Table III. We also observe that MME (or MLE) and LRE are unbiased, whereas NE is biased. Moreover, MSE($\hat{t_{0\text{MME}}}$) and MSE($\hat{t_{0\text{LRE}}}$) are smaller than MSE($\hat{t_{0\text{NE}}}$), whereas MSE($\hat{t_{0\text{MME}}}$) is the smallest one when $n > 3$ and $p < 0.5$. Specifically, the size of Darknet, ω, is much smaller than Ω ($= 2^{32}$), which leads to $p \ll 1$. Thus, when the Darknet observes a sufficient number of hits (*i.e.,* $n \gg 1$), we have MSE($\hat{t_{0\text{MME}}}$) $\approx$ MSE($\hat{t_{0\text{LRE}}}$) $\approx \frac{1}{2}$MSE($\hat{t_{0\text{NE}}}$). That is, the MSE of our proposed estimators is almost half of that of the naive estimator.

III. ESTIMATING THE WORM INFECTION SEQUENCE

We extend our proposed estimators for inferring the worm infection sequence. Our approach is that we first estimate the infection time of each infected host. Then, we infer the infection sequence based on these infection times. That is, if $\hat{t_{0\text{A}}} < \hat{t_{0\text{B}}}$, we infer that host A is infected before host B. It is noted that the algorithm used in [16] to infer the worm infection sequence can be regarded as using this approach with the naive estimator for estimating the host infection time.

The naive estimator, however, can potentially fail to infer the worm infection sequence in some cases. Fig. 4 shows an example, where hosts A and B get infected at $t_{0\text{A}}$ and $t_{0\text{B}}$ respectively, and $t_{0\text{A}} < t_{0\text{B}}$. Moreover, these two infected hosts have such scanning rates $s_{\text{A}} < s_{\text{B}}$ that Darknet observes that $t_{1\text{A}} > t_{1\text{B}}$. If the naive estimator is used, $\hat{t_{0\text{A}}} > \hat{t_{0\text{B}}}$, and thus host A is incorrectly inferred to be infected after host B. Intuitively, if our proposed estimators are applied, it is possible to obtain $\hat{t_{0\text{A}}} < \hat{t_{0\text{B}}}$ and thus recover the infection sequence. Therefore, we expect that our algorithms can provide more accuracy in estimating the worm infection sequence from Darknet observations.

IV. SIMULATION RESULTS

A. Estimating the Host Infection Time

To evaluate the performance of our proposed estimators, we simulate the behavior of a host infected by the Code Red worm. The host is infected at time tick 0 and uses a constant scanning rate. The time unit is set to 20 seconds. The Darknet records hit times for the time interval of an observation window size. Each point in Fig.s 5-7 is averaged over 100 independent runs.

We then consider the effects of the Darknet size, the scanning rate, and the observation window size on the performance of the estimators. Fig. 5 compares the performance of NE, MME, and

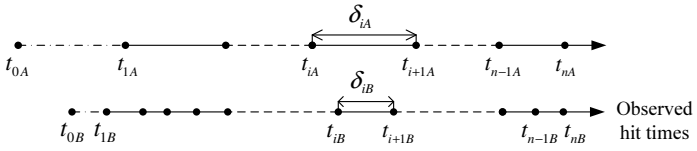


Fig. 4. A scenario of the worm infection sequence.

LRE with different Darknet sizes from 2^{18} to 2^{25}, a scanning rate of 358 scans/min, and an observation window size of 800 mins. The three sub-figures show the mean of estimators for μ, the mean of estimators for t_0, and the MSE of estimators for t_0. Fig. 6 compares the three estimators with different scanning rates from 158 scans/min to 558 scans/min, a Darknet size of 2^{20}, and an observation window size of 800 mins. Similarly, Fig. 7 is with different observation window sizes from 50 mins to 800 mins, a scanning rate of 358 scans/min, and a Darknet size of 2^{20}. It is observed that for all cases, our proposed estimators have better performance (*i.e.,* unbiasedness and smaller MSE) than the naive estimator in estimating the host infection time. Specifically, the simulation results verify that the MSE of our estimators is almost half of that of the naive estimator, when the observation window size is sufficiently large (*e.g.,* $>$ 200 mins).

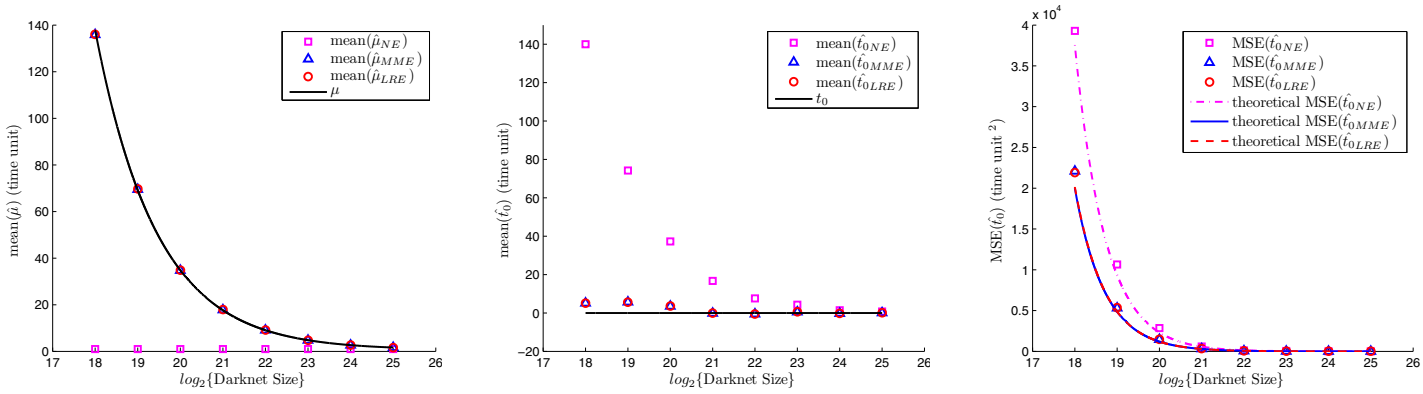
B. Estimating the Worm Infection Sequence

To evaluate the performance of our algorithms in estimating the worm infection sequence, we simulate the propagation of the Code Red worm. The simulator is extended from the code provided by [20]. The Code Red worm has a vulnerable population of 360,000. Different infected hosts may have different scanning rates. Thus, we assign a scanning rate (scans/min) from a normal distribution $N(358, \sigma^2)$ to a newly infected host. Moreover, we start our simulation at time tick 0 and from one infected host, *i.e.,* patient zero. The time unit is set to 20 seconds. Each point in Fig. 8 is averaged over 20 independent runs. Table IV gives the results of a sample run with a Darknet size of 2^{20}, an observation window size of 1,600 mins, and σ = 110. In the table, S_i is the actual infection sequence, whereas $\hat{S}_i$ is the estimated sequence. In this example, we find that MME and LRE can pinpoint the patient zero successfully, while NE fails.

To compare the performance of estimators quantitatively, we consider a simple l_1 *sequence distance*, *i.e.,*

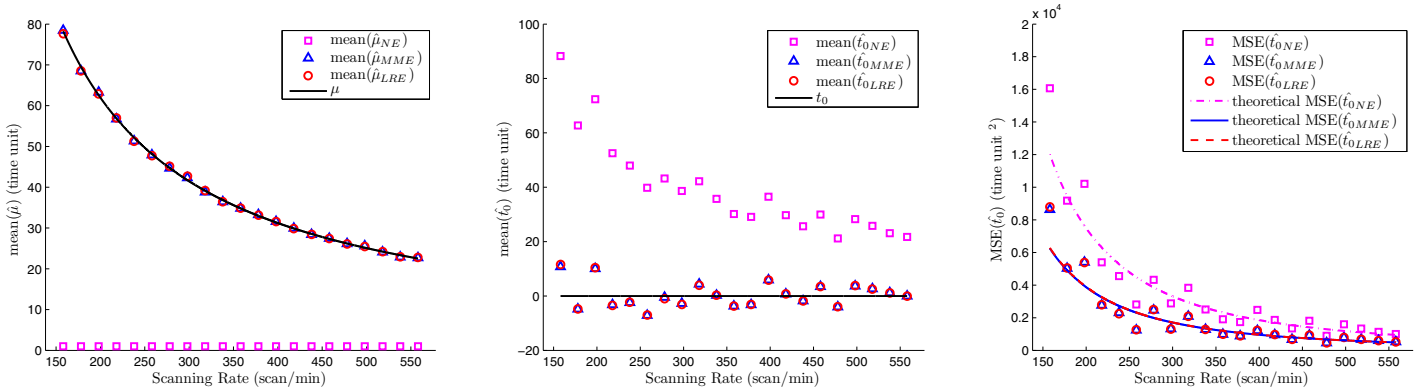
$$D = \sum_{i=1}^{N} \left| S_i - \hat{S}_i \right|, \tag{24}$$

where $S_i = i$ and N is the length of the infection sequence considered. Note that the smaller the sequence distance is, the



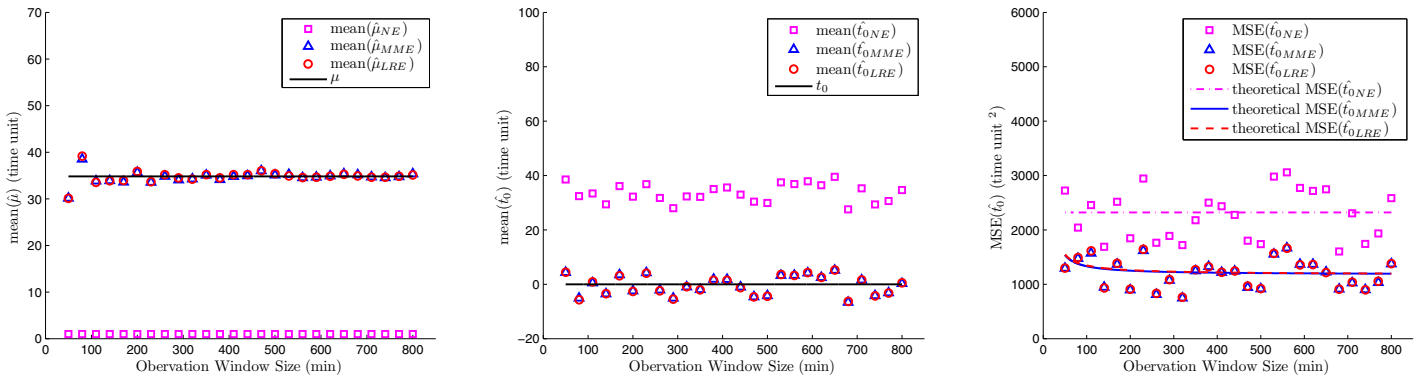
(a) Comparison of $\hat{\mu}$. (b) Comparison of $\hat{t_0}$. (c) Comparison of MSE($\hat{t_0}$).

Fig. 5. Simulation results of changing the Darknet size (all cases are for scanning rate: 358 scans/min, observation window size: 800 mins).



(a) Comparison of $\hat{\mu}$. (b) Comparison of $\hat{t_0}$. (c) Comparison of MSE($\hat{t_0}$).

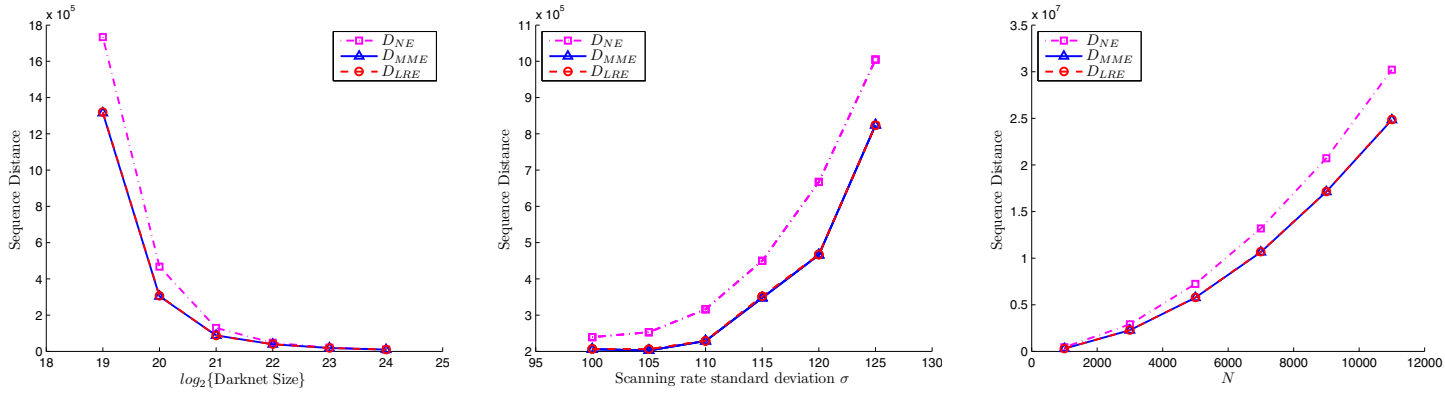
Fig. 6. Simulation results of changing the scanning rate (all cases are for Darknet size: 2^{20} IP addresses, observation window size: 800 mins).



(a) Comparison of $\hat{\mu}$. (b) Comparison of $\hat{t_0}$. (c) Comparison of MSE($\hat{t_0}$).

Fig. 7. Simulation results of changing the observation window size (all cases are for scanning rate: 358 scans/min, Darknet size: 2^{20} IP addresses).

better the estimator performance is. Fig. 8(a) shows the sequence distances of NE, MME, and LRE with varying Darknet sizes from 2^{19} to 2^{24}, an observation window size of 1,600 mins, N = 1,000, and σ = 115. It is observed that when the Darknet size increases, the performance of all estimators improves dramatically. Moreover, the performance of MME and LRE is always better than that of NE. For example, when the Darknet size equals 2^{19}, MME and LRE improve the inference accuracy by 24%, compared with NE. Fig. 8(b) demonstrates the sequence distances of these three estimators by changing the standard deviation of the scanning rate (*i.e.,* σ) from 100 to 125. In the figure, the Darknet size is 2^{20}, the observation window size is 1,600 mins, and N = 1,000. It is noted that when σ increases, the performance of all estimators deteriorates. The



(a) (b) (c)

Fig. 8. Simulation results of the sequence distance. (a) Changing the Darknet size (N = 1,000, observation window size: 1,600 mins, scanning rate: $N(358, 115^2)$). (b) Changing the scanning rate standard deviation (N = 1,000, observation window size: 1,600 mins, Darknet size: 2^{20} IP addresses). (c) Changing the length of the infection sequence considered (observation window size: 1,600 mins, Darknet size: 2^{20} IP addresses, scanning rate: $N(358, 115^2)$).

TABLE IV
A SAMPLE RUN OF SIMULATIONS.

S_i	$\hat{S}_{i\text{NE}}$	$\hat{S}_{i\text{MME}}$	$\hat{S}_{i\text{LRE}}$	t_0	$\hat{t_0}_{\text{NE}}$	$\hat{t_0}_{\text{MME}}$	$\hat{t_0}_{\text{LRE}}$
1	2	1	1	0	114	20	20
2	1	2	2	85	98	74	73
3	3	3	3	105	165	116	116
:	:	:	:	:	:	:	:
520	498	533	534	593	622	589	589
521	433	488	477	594	611	581	580
:	:	:	:	:	:	:	:

performance of MME and LRE, however, is always better than that of NE. For example, when σ = 120, MME and LRE reduce the sequence distance by 30%, compared with NE. In Fig. 8(c), we increase the length of the infection sequence considered, N, from 1,000 to 11,000. Here the Darknet size is 2^{20}, the observation window size is 1,600 mins, and σ = 115. It is intuitive that the sequence distances of all estimators become larger as N increases. However, MME and LRE are always better than NE. For example, when N = 7,000, MME and LRE beat NE by 20%. Therefore, our proposed estimators perform much better than the naive estimator.

V. CONCLUSIONS

In this paper, we have attempted to understand the temporal characteristics of Internet worms, under the framework of Internet worm tomography, through both analysis and simulation. Specifically, we have proposed method of moments, maximum likelihood, and linear regression estimators to infer the host infection time and reconstruct the worm infection sequence. We have shown analytically and empirically that the mean squared error of our proposed estimators can be almost half of that of the naive estimator in estimating the host infection time. As a result, our estimation techniques have been demonstrated to perform much better than the algorithm used in [16] in estimating the worm infection sequence.

ACKNOWLEDGMENT

This work was supported in part by DHS grant 2008-ST-062-000012.

REFERENCES

[1] Darknet. [Online]. Available: http://www.cymru.com/Darknet/.
[2] Network Telescope. [Online]. Available: http://www.caida.org/research/security/telescope/.
[3] Honeypots: Tracking Hackers. [Online]. Available: http://www.tracking-hackers.com/.
[4] Internet Motion Sensor. [Online]. Available: http://ims.eecs.umich.edu/.
[5] Internet Sink. [Online]. Available: http://wail.cs.wisc.edu/anomaly.html.
[6] T. Bu, A. Chen, S. Wiel, and T. Woo, "Design and Evaluation of a Fast and Robust Worm Detection Algorithm," in *Proc. IEEE INFOCOM*, Apr. 2006.
[7] R. Caceres, N.G. Duffield, J. Horowitz, and D. Towsley, "Multicast-based Inference of Network-internal Loss Characteristics," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2462–2480, Nov. 1999.
[8] X. Chen and J. Heidemann, "Detecting Early Worm Propagation through Packet Matching," *Technical Report ISI-TR-2004-585*, Feb. 2004.
[9] Z. Chen, C. Chen, and C. Ji, "Understanding Localized-Scanning Worms," in *Proc. IEEE IPCCC*, Apr. 2007.
[10] M. Coates, A. Hero, R. Nowak, and B. Yu, "Internet Tomography," *IEEE Signal Processing Magazine*, pp. 47–65, May 2002.
[11] J. Jung, V. Paxson, A. Berger, and H. Balakrishnan, "Fast Portscan Detection Using Sequential Hypothesis Testing," in *Proc. IEEE Symposium on Security and Privacy*, May 2004.
[12] A. Kumar, V. Paxson, and N. Weaver, "Exploiting Underlying Structure for Detailed Reconstruction of an Internet-scale Event," in *Proc. Internet Measurement Conference*, 2005.
[13] A. Lakhina, M. Crovella, and C. Diot, "Mining Anomalies Using Traffic Feature Distributions," in *Proc. ACM SIGCOMM*, Aug. 2005.
[14] D. Moore, C. Shannon, G. M. Voelker, and S. Savage, "Network Telescopes: Technical Report," *Technical Report*, Jul. 2004.
[15] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver, "Inside the Slammer Worm," *IEEE Security and Privacy*, vol. 1, no. 4, pp. 33–39, Jul. 2003.
[16] M. A. Rajab, F. Monrose, and A. Terzis, "Worm Evolution Tracking via Timing Analysis," in *Proc. Workshop on Rapid Malcode (WORM)*, Nov. 2005.
[17] N. Weaver, S. Staniford, and V. Paxson, "Very Fast Containment of Scanning Worms," in *Proc. 13th Usenix Security Conference*, Aug. 2004.
[18] J. Wu, S. Vangala, L. Gao, and K. Kwiat, "An Effective Architecture and Algorithm for Detecting Worms with Various Scan Techniques," *NDSS*, 2004.
[19] Y. Xie, V. Sekar, D. A. Maltz, M. K. Reiter, and H. Zhang, "Worm Origin Identification Using Random Walks," in *Proc. IEEE Symposium on Security and Privacy*, May 2005.
[20] C. C. Zou, Internet Worm Propagation Simulator. [Online]. Available: http://www.cs.ucf.edu/~czou/research/wormSimulation/simulator-codered-100run.cpp.
[21] C. C. Zou, W. Gong, D. Towsley, and L. Gao, "The Monitoring and Early Detection of Internet Worms," *IEEE/ACM Transactions on Networking*, vol. 13, no. 5, pp. 967–974, Oct. 2005.