Inferring Internet Worm Temporal Characteristics

Qian Wang[1] Zesheng Chen[1]
Kia Makki[1] Niki Pissinou[1] Chao Chen[2]

[1]Department of Electrical and Computer Engineering
Florida International University

[2]Department of Engineering
Indiana University - Purdue University Fort Wayne

IEEE GLOBECOM 2008, 12/03/2008

Outline

Outline

Internet Worm Temporal Characteristics

Host Infection Time

When exactly does a specific host get infected?

Worm Infection Sequence

What is the host infection order of worm propagation?

Internet Worm Temporal Characteristics

Host Infection Time

When exactly does a specific host get infected?

Worm Infection Sequence

What is the host infection order of worm propagation?

Internet Worm Temporal Characteristics

Host Infection Time

When exactly does a specific host get infected?

Worm Infection Sequence

What is the host infection order of worm propagation?

Why is it important?

Host Infection Time

- Forensic analysis of an infected host.
- Reconstruction of the worm infection sequence.

Worm Infection Sequence

- Understand worm propagation characteristics.
- Identify patient zero or initially infected hosts.

Why is it important?

Host Infection Time

- Forensic analysis of an infected host.
- Reconstruction of the worm infection sequence.

Worm Infection Sequence

- Understand worm propagation characteristics.
- Identify patient zero or initially infected hosts.

Why is it important?

Host Infection Time

- Forensic analysis of an infected host.
- Reconstruction of the worm infection sequence.

Worm Infection Sequence

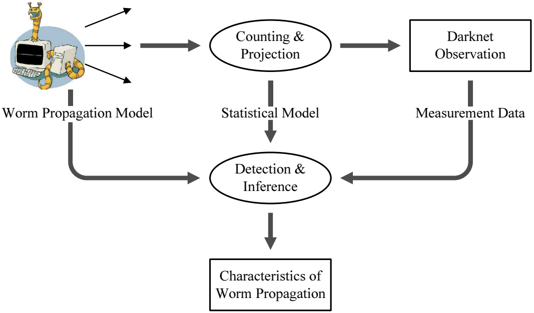
- Understand worm propagation characteristics.
- Identify patient zero or initially infected hosts.

Outline

Internet Worm Tomography

Solution

Inferring the characteristics of Internet worms from the observations of Darknet that are the routable but unused IP address space.



Internet worm tomography.

Why Darknet?

Source Detection and Defenses

Detect infected hosts in the local networks.

Middle Detection and Defenses

Reveal the appearance of worms by analyzing the traffic going through routers.

Destination Detection and Defenses

- Monitor malicious or unintended traffic arriving at Darknet.
- Offer unique advantages in observing large-scale network explosive events.

Why Darknet?

Source Detection and Defenses

Detect infected hosts in the local networks.

Middle Detection and Defenses

Reveal the appearance of worms by analyzing the traffic going through routers.

Destination Detection and Defenses

- Monitor malicious or unintended traffic arriving at Darknet.
- Offer unique advantages in observing large-scale network explosive events.

Why Darknet?

Source Detection and Defenses

Detect infected hosts in the local networks.

Middle Detection and Defenses

Reveal the appearance of worms by analyzing the traffic going through routers.

Destination Detection and Defenses

- Monitor malicious or unintended traffic arriving at Darknet.
- Offer unique advantages in observing large-scale network explosive events.

What are we going to do?

- 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.
- Rajab *et al.* use the same data and study the "infection and detection times" to infer the worm infection sequence.

Our approach

Employ **statistical estimation** techniques to Internet worm tomography.

What are we going to do?

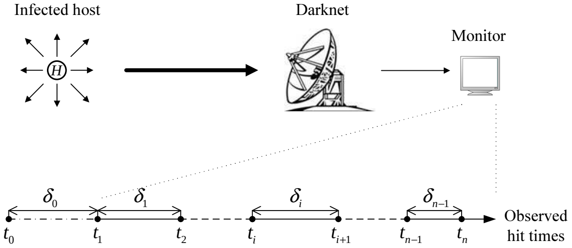
- Kumar *et al.* use network telescope data and analyze the pseudo-random number generator to reconstruct the "who infected whom" infection tree of the Witty worm.
- Rajab *et al.* use the same data and study the "infection and detection times" to infer the worm infection sequence.

Our approach

Employ **statistical estimation** techniques to Internet worm tomography.

Outline

Host Infection Time

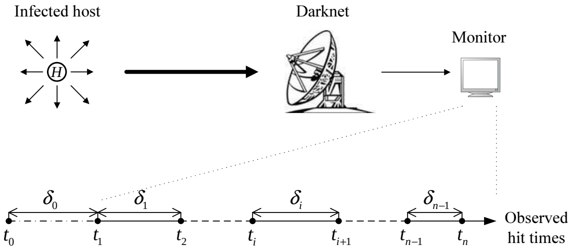


An illustration of Darknet observations.

Host Infection Time

Given the Darknet observations $t_1, t_2, \cdots, t_n$, what is the best estimate of t_0?

Host Infection Time

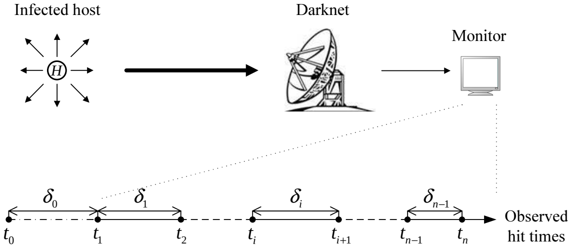


An illustration of Darknet observations.

Host Infection Time

Given the Darknet observations $t_1, t_2, \cdots, t_n$, what is the best estimate of t_0?

How to estimate t_0?



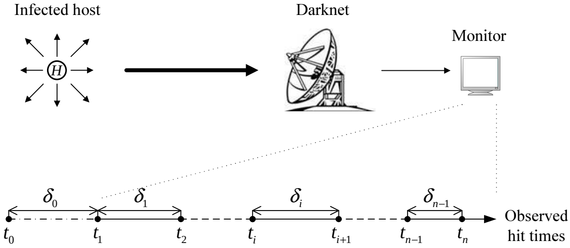
An illustration of Darknet observations.

- Hit event: Darknet observing at least one scan from the same infected host in a time unit.

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

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

How to estimate t_0?



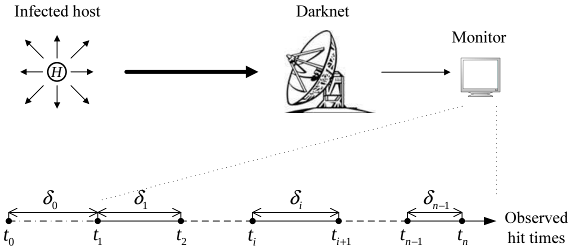
An illustration of Darknet observations.

- Hit event: Darknet observing at least one scan from the same infected host in a time unit.

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

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

How to estimate t_0?



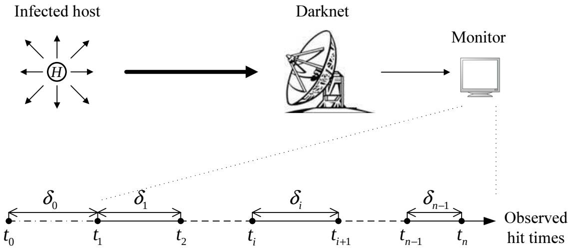
An illustration of Darknet observations.

- Hit event: Darknet observing at least one scan from the same infected host in a time unit.

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

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

How to estimate t_0?



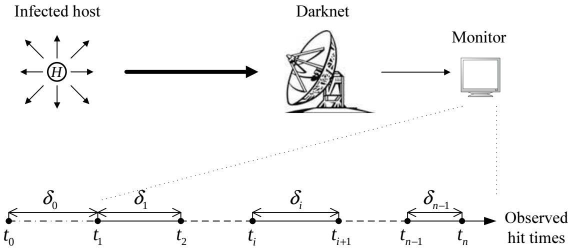
An illustration of Darknet observations.

$$\mathrm{E}(\delta) = \mu.$$

The problem is reduced to estimating μ

$$\hat{t}_0 = t_1 - \hat{\mu}.$$

How to estimate t_0?



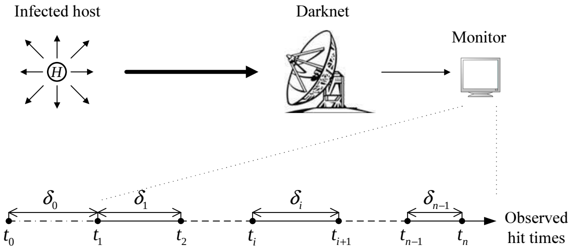
An illustration of Darknet observations.

$$\mathrm{E}(\delta) = \mu.$$

The problem is reduced to estimating μ

$$\hat{t}_0 = t_1 - \hat{\mu}.$$

Naïve Estimator



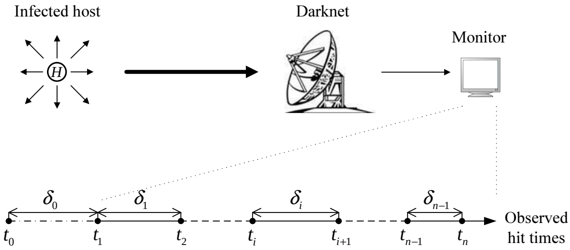
An illustration of Darknet observations.

- $\Pr(\delta)$ is maximized when $\delta = 1$.

Naïve Estimator (NE) of μ

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

Naïve Estimator



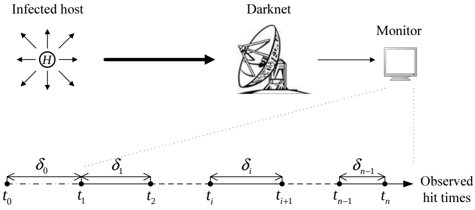
An illustration of Darknet observations.

- $\Pr(\delta)$ is maximized when $\delta = 1$.

Naïve Estimator (NE) of μ

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

Method of Moments Estimator



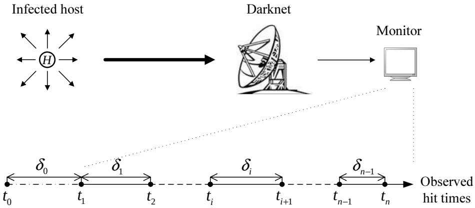
An illustration of Darknet observations.

- Equating sample mean with unobservable real mean.

Method of Moments Estimator (MME) of μ

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

Method of Moments Estimator



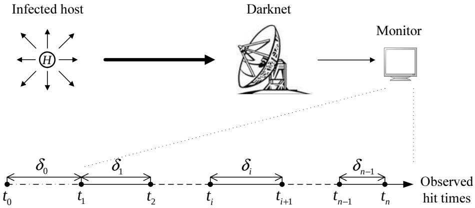
An illustration of Darknet observations.

- Equating sample mean with unobservable real mean.

Method of Moments Estimator (MME) of μ

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

Maximum Likelihood Estimator

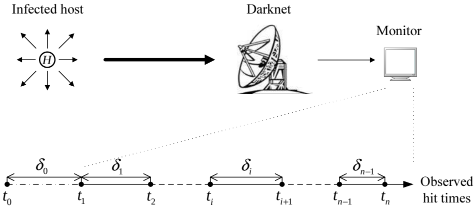


An illustration of Darknet observations.

- Finding the value of parameter μ which makes the likelihood function a maximum.
- Likelihood function
  - Probability for the occurrence of observed Darknet samples.

$$L(\mu) = \prod_{i=1}^{n-1} \Pr(\delta_i; \mu).$$

Maximum Likelihood Estimator

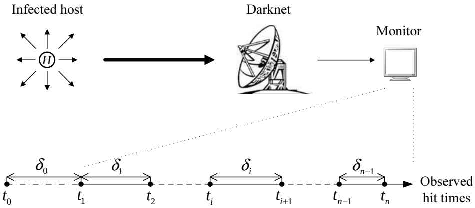


An illustration of Darknet observations.

- Finding the value of parameter μ which makes the likelihood function a maximum.
- Likelihood function
  - Probability for the occurrence of observed Darknet samples.

$$\mathrm{L}(\mu) = \prod_{i=1}^{n-1} \Pr(\delta_i; \mu).$$

Maximum Likelihood Estimator

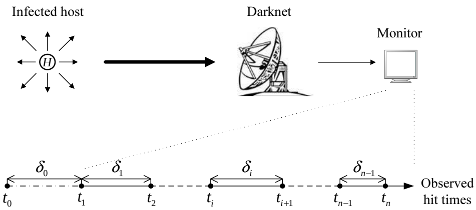


An illustration of Darknet observations.

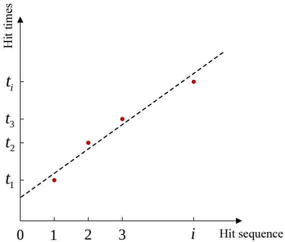
Maximum Likelihood Estimator (MLE) of μ

$$\hat{\mu}_{\mathrm{MLE}} = \arg\max_{\mu} \mathrm{L}(\mu) = \frac{1}{n-1}\sum_{i=1}^{n-1}\delta_i = \frac{t_n - t_1}{n-1}.$$

Linear Regression Estimator



An illustration of Darknet observations.

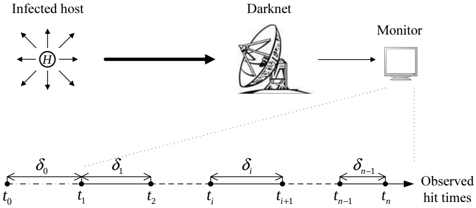


Linear regression model.

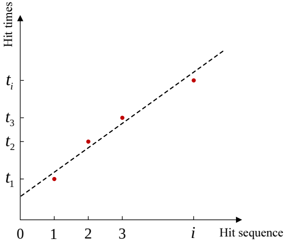
- Assuming scanning rate of an individual infected host is time-invariant.
- The relationship between t_i and i can be described by a linear regression model

$$t_i = \alpha + \beta \cdot i + \varepsilon_i.$$

Linear Regression Estimator



An illustration of Darknet observations.



Linear regression model.

- Assuming scanning rate of an individual infected host is time-invariant.
- The relationship between t_i and i can be described by a linear regression model

$$t_i = \alpha + \beta \cdot i + \varepsilon_i.$$

Linear Regression Estimator

- Choose the coefficients that minimize the residual sum of squares (RSS)

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

- We then have

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

Linear Regression Estimator (LRE) of μ

$$\hat{\mu}_{\text{LRE}} = \hat{\beta} = \hat{t}_1 - \hat{t}_0.$$

Linear Regression Estimator

- Choose the coefficients that minimize the residual sum of squares (RSS)

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

- We then have

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

Linear Regression Estimator (LRE) of μ

$$\hat{\mu}_{\text{LRE}} = \hat{\beta} = \hat{t}_1 - \hat{t}_0.$$

Outline

Comparison of Estimators

- Compare the performance of the estimators

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

Table: Comparison of estimator properties ($\hat{\mu}$).

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

Comparison of Estimators

- Compare the performance of the estimators

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

Table: 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	$\frac{(1-p)^2}{p^2}$
$\hat{\mu}_{\text{MME}} = \frac{t_n - t_1}{n-1}$	0	$\frac{1-p}{p^2(n-1)}$	$\frac{1-p}{p^2(n-1)}$
$\hat{\mu}_{\text{LRE}} = \frac{\overline{i \cdot t} - \overline{i} \cdot \overline{t}}{\overline{i^2} - (\overline{i})^2}$	0	$\frac{6(n^2+1)(1-p)}{5n(n^2-1)p^2}$	$\frac{6(n^2+1)(1-p)}{5n(n^2-1)p^2}$

Comparison of Estimators

Table: Comparison of estimator properties ($\hat{t}_0$).

$\hat{t}_0$	Bias($\hat{t}_0$)	Var($\hat{t}_0$)	MSE($\hat{t}_0$)
$\hat{t}_{0\,\mathrm{NE}} = t_1 - \hat{\mu}_{\mathrm{NE}}$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$	$\frac{(1-p)(2-p)}{p^2}$
$\hat{t}_{0\,\mathrm{MME}} = t_1 - \hat{\mu}_{\mathrm{MME}}$	0	$\frac{1-p}{p^2} \cdot \frac{n}{n-1}$	$\frac{1-p}{p^2} \cdot \frac{n}{n-1}$
$\hat{t}_{0\,\mathrm{LRE}} = t_1 - \hat{\mu}_{\mathrm{LRE}}$	0	$\frac{1-p}{p^2} \cdot \frac{5n^3+6n^2-5n+6}{5n(n^2-1)}$	$\frac{1-p}{p^2} \cdot \frac{5n^3+6n^2-5n+6}{5n(n^2-1)}$

Theorem

When the Darknet observes a sufficient number of hits (i.e., $n \gg 1$) and $p \ll 1$,

$$MSE(\hat{t}_{0\,MME}) = MSE(\hat{t}_{0\,MLE}) \approx MSE(\hat{t}_{0\,LRE}) \approx \frac{1}{2} MSE(\hat{t}_{0\,NE}).$$

Comparison of Estimators

Table: Comparison of estimator properties ($\hat{t}_0$).

$\hat{t}_0$	Bias($\hat{t}_0$)	Var($\hat{t}_0$)	MSE($\hat{t}_0$)
$\hat{t}_{0\,\mathrm{NE}} = t_1 - \hat{\mu}_{\mathrm{NE}}$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$	$\frac{(1-p)(2-p)}{p^2}$
$\hat{t}_{0\,\mathrm{MME}} = t_1 - \hat{\mu}_{\mathrm{MME}}$	0	$\frac{1-p}{p^2} \cdot \frac{n}{n-1}$	$\frac{1-p}{p^2} \cdot \frac{n}{n-1}$
$\hat{t}_{0\,\mathrm{LRE}} = t_1 - \hat{\mu}_{\mathrm{LRE}}$	0	$\frac{1-p}{p^2} \cdot \frac{5n^3+6n^2-5n+6}{5n(n^2-1)}$	$\frac{1-p}{p^2} \cdot \frac{5n^3+6n^2-5n+6}{5n(n^2-1)}$

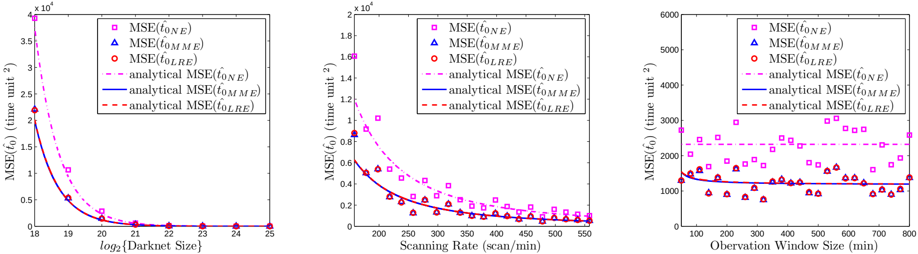
Theorem

When the Darknet observes a sufficient number of hits (i.e., $n \gg 1$) *and* $p \ll 1$,

$$MSE(\hat{t}_{0\,MME}) = MSE(\hat{t}_{0\,MLE}) \approx MSE(\hat{t}_{0\,LRE}) \approx \frac{1}{2} MSE(\hat{t}_{0\,NE}).$$

Outline

Simulation Results

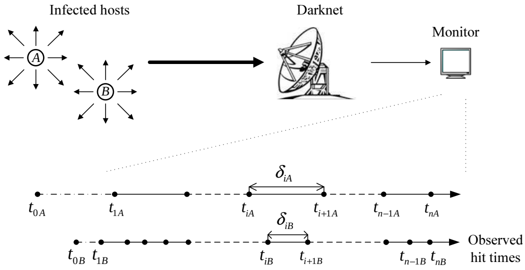


(a) Changing Ω. (b) Changing s. (c) Changing T.

Figure: Comparison of MSE($\hat{t}_0$).

Outline

How it works?



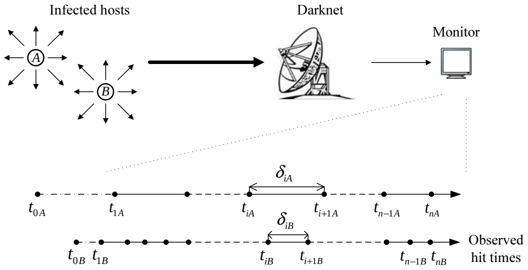
A scenario of the worm infection sequence.

- $\Pr_{NE}(\text{error}) = \Pr(t_{1A} - 1 > t_{1B} - 1)$.
- $\Pr(\text{error}) = \Pr(t_{1A} - \frac{1}{p_A} > t_{1B} - \frac{1}{p_B})$.

Probability of Error Detection

$$E\left[\Pr_{NE}(\text{error})\right] > E\left[\Pr(\text{error})\right].$$

How it works?



A scenario of the worm infection sequence.

- $\Pr_{\mathrm{NE}}(\text{error}) = \Pr(t_{1\mathrm{A}} - 1 > t_{1\mathrm{B}} - 1)$.
- $\Pr(\text{error}) = \Pr(t_{1\mathrm{A}} - \frac{1}{p_{\mathrm{A}}} > t_{1\mathrm{B}} - \frac{1}{p_{\mathrm{B}}})$.

Probability of Error Detection

$$\mathrm{E}\left[\Pr_{\mathrm{NE}}(\text{error})\right] > \mathrm{E}\left[\Pr(\text{error})\right].$$

Outline

Simulation Results

Table: A sample run of simulations.

S_i	$\hat{S}_{i\,\mathrm{NE}}$	$\hat{S}_{i\,\mathrm{MME}}$	$\hat{S}_{i\,\mathrm{LRE}}$	t_0	$\hat{t}_{0\,\mathrm{NE}}$	$\hat{t}_{0\,\mathrm{MME}}$	$\hat{t}_{0\,\mathrm{LRE}}$
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
:	:	:	:	:	:	:	:

- Sequence Distance

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

Simulation Results

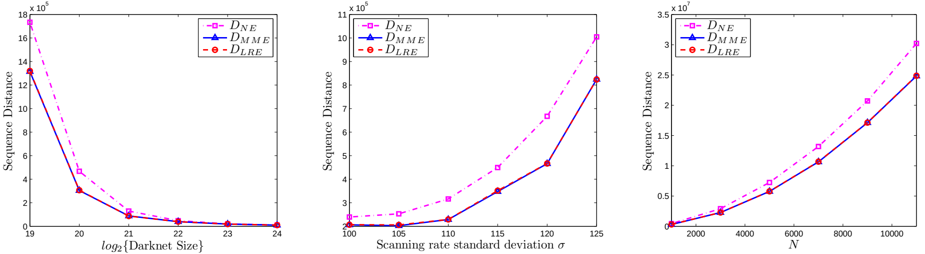
Table: A sample run of simulations.

S_i	$\hat{S}_{i\,\mathrm{NE}}$	$\hat{S}_{i\,\mathrm{MME}}$	$\hat{S}_{i\,\mathrm{LRE}}$	t_0	$\hat{t}_{0\,\mathrm{NE}}$	$\hat{t}_{0\,\mathrm{MME}}$	$\hat{t}_{0\,\mathrm{LRE}}$
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
:	:	:	:	:	:	:	:

- Sequence Distance

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

Simulation Results



(a) Changing Ω (b) Changing σ (c) Changing N

Figure: Comparison of sequence distance.

Summary

Host Infection Time

- Propose method of moments, maximum likelihood, and linear regression statistical estimators.
- Show analytically and empirically

$$\text{MSE}(\hat{t}_{0_{\text{MME}}}) = \text{MSE}(\hat{t}_{0_{\text{MLE}}}) \approx \text{MSE}(\hat{t}_{0_{\text{LRE}}}) \approx \frac{1}{2}\text{MSE}(\hat{t}_{0_{\text{NE}}}).$$

Worm Infection Sequence

- Extend our proposed estimators to infer the worm infection sequence.
- Demonstrate our method performs much better than the naïve estimator.

Summary

Host Infection Time

- Propose method of moments, maximum likelihood, and linear regression statistical estimators.
- Show analytically and empirically

$$\text{MSE}(\hat{t}_{0\text{MME}}) = \text{MSE}(\hat{t}_{0\text{MLE}}) \approx \text{MSE}(\hat{t}_{0\text{LRE}}) \approx \frac{1}{2}\text{MSE}(\hat{t}_{0\text{NE}}).$$

Worm Infection Sequence

- Extend our proposed estimators to infer the worm infection sequence.
- Demonstrate our method performs much better than the naïve estimator.

Summary

Host Infection Time

- Propose method of moments, maximum likelihood, and linear regression statistical estimators.
- Show analytically and empirically

$$\mathrm{MSE}(\hat{t}_{0_{\mathrm{MME}}}) = \mathrm{MSE}(\hat{t}_{0_{\mathrm{MLE}}}) \approx \mathrm{MSE}(\hat{t}_{0_{\mathrm{LRE}}}) \approx \frac{1}{2}\mathrm{MSE}(\hat{t}_{0_{\mathrm{NE}}}).$$

Worm Infection Sequence

- Extend our proposed estimators to infer the worm infection sequence.
- Demonstrate our method performs much better than the naïve estimator.

Summary

Host Infection Time

- Propose method of moments, maximum likelihood, and linear regression statistical estimators.
- Show analytically and empirically

$$\mathrm{MSE}(\hat{t}_{0_{\mathrm{MME}}}) = \mathrm{MSE}(\hat{t}_{0_{\mathrm{MLE}}}) \approx \mathrm{MSE}(\hat{t}_{0_{\mathrm{LRE}}}) \approx \frac{1}{2}\mathrm{MSE}(\hat{t}_{0_{\mathrm{NE}}}).$$

Worm Infection Sequence

- Extend our proposed estimators to infer the worm infection sequence.
- Demonstrate our method performs much better than the naïve estimator.

Summary

Host Infection Time

- Propose method of moments, maximum likelihood, and linear regression statistical estimators.
- Show analytically and empirically

$$\mathrm{MSE}(\hat{t}_{0\mathrm{MME}}) = \mathrm{MSE}(\hat{t}_{0\mathrm{MLE}}) \approx \mathrm{MSE}(\hat{t}_{0\mathrm{LRE}}) \approx \frac{1}{2}\mathrm{MSE}(\hat{t}_{0\mathrm{NE}}).$$

Worm Infection Sequence

- Extend our proposed estimators to infer the worm infection sequence.
- Demonstrate our method performs much better than the naïve estimator.

Future Works

Future Works

- What if packets can be lost?
- What if scanning rate of an infected host can vary?
- What about other scanning methods?

Questions