Inferring Internet Worm Temporal Characteristics

Qian Wang\textsuperscript{1}  Zesheng Chen\textsuperscript{1}  
Kia Makki\textsuperscript{1}  Niki Pissinou\textsuperscript{1}  Chao Chen\textsuperscript{2}

\textsuperscript{1}Department of Electrical and Computer Engineering  
Florida International University

\textsuperscript{2}Department of Engineering  
Indiana University - Purdue University Fort Wayne

IEEE GLOBECOM 2008, 12/03/2008
Outline

1. Introduction
   - What is the problem?
   - What are we going to do?

2. Estimating the Host Infection Time
   - Estimating the Host Infection Time
   - Comparison of Estimators
   - Simulation Results

3. Estimating the Worm Infection Sequence
   - An Illustrated Scenario
   - Simulation Results

4. Summary and Future Works
1. **Introduction**
   - What is the problem?
   - What are we going to do?

2. **Estimating the Host Infection Time**
   - Estimating the Host Infection Time
   - Comparison of Estimators
   - Simulation Results

3. **Estimating the Worm Infection Sequence**
   - An Illustrated Scenario
   - Simulation Results

4. **Summary and Future Works**
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?
Introduction

Estimating the Host Infection Time

Estimating the Worm Infection Sequence

Summary and Future Works

What is the problem?

What are we going to do?

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

1. Introduction
   - What is the problem?
   - What are we going to do?

2. Estimating the Host Infection Time
   - Estimating the Host Infection Time
   - Comparison of Estimators
   - Simulation Results

3. Estimating the Worm Infection Sequence
   - An Illustrated Scenario
   - Simulation Results

4. Summary and Future Works
Internet Worm Tomography

Solution

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

1. Introduction
   - What is the problem?
   - What are we going to do?

2. Estimating the Host Infection Time
   - Estimating the Host Infection Time
   - Comparison of Estimators
   - Simulation Results

3. Estimating the Worm Infection Sequence
   - An Illustrated Scenario
   - Simulation Results

4. Summary and Future Works
Host Infection Time

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

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

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

- $\Pr(\delta = k) = p \cdot (1 - p)^{k-1}, \quad k = 1, 2, 3, \ldots$
How to estimate $t_0$?

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

$$\Pr (\text{hit event}) = 1 - \left( 1 - \frac{\omega}{\Omega} \right)^s = p.$$
How to estimate $t_0$?

**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 $t_0$?

The problem is reduced to estimating $\mu$

$$\hat{t}_0 = t_1 - \hat{\mu}.$$
How to estimate $t_0$?

Infected host $\rightarrow$ Darknet $\rightarrow$ Monitor

$\delta_0 \rightarrow \delta_1 \rightarrow \delta_i \rightarrow \delta_{i+1} \rightarrow \delta_{n-1} \rightarrow \delta_n \rightarrow$ Observed hit times

$E(\delta) = \mu.$

The problem is reduced to estimating $\mu$

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

Pr(δ) is maximized when δ = 1.

Naïve Estimator (NE) of μ

\[ \hat{\mu}_{\text{NE}} = 1. \]
**Naïve Estimator**

The Naïve Estimator (NE) of $\mu$ is given by:

$$\hat{\mu}_{\text{NE}} = 1.$$
Method of Moments Estimator

Equating sample mean with unobservable real mean.

\[
\hat{\mu}_{\text{MME}} = \bar{\delta} = \frac{1}{n-1} \sum_{i=1}^{n-1} \delta_i = \frac{t_n - t_1}{n-1}.
\]
Method of Moments Estimator

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

- Equating sample mean with unobservable real mean.

An illustration of Darknet observations.
Maximum Likelihood Estimator

- Finding the value of parameter $\mu$ 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

Finding the value of parameter $\mu$ 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

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

An illustration of Darknet observations.
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

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

\[
\begin{align*}
\hat{\alpha} &= \bar{t} - \hat{\beta} \cdot \bar{i} \\
\hat{\beta} &= \frac{i \cdot \bar{t} - \bar{i} \cdot \bar{t}}{i^2 - (\bar{i})^2}.
\end{align*}
\]

**Linear Regression Estimator (LRE) of \( \mu \)**

\[
\hat{\mu}_{\text{LRE}} = \hat{\beta} = \hat{t}_1 - \hat{t}_0.
\]
Linear Regression Estimator

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

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

- We then have

\[
\begin{align*}
\hat{\alpha} &= \bar{t} - \hat{\beta} \cdot \bar{i} \\
\hat{\beta} &= \frac{\bar{i} \cdot t - \bar{i} \cdot \bar{t}}{\bar{i}^2 - (\bar{i})^2}.
\end{align*}
\]

**Linear Regression Estimator (LRE) of \( \mu \)**

\[ \hat{\mu}_{\text{LRE}} = \hat{\beta} = \hat{t}_1 - \hat{t}_0. \]
Outline

1. Introduction
   - What is the problem?
   - What are we going to do?

2. Estimating the Host Infection Time
   - Estimating the Host Infection Time
   - Comparison of Estimators
   - Simulation Results

3. Estimating the Worm Infection Sequence
   - An Illustrated Scenario
   - Simulation Results

4. Summary and Future Works
Comparison of Estimators

Compare the performance of the estimators

\[
\begin{align*}
\text{Bias}(\hat{\mu}) &= E(\hat{\mu}) - \mu \\
\text{Var}(\hat{\mu}) &= E[(\hat{\mu} - E(\hat{\mu}))^2] \\
\text{MSE}(\hat{\mu}) &= E[(\hat{\mu} - \mu)^2] = \text{Bias}^2(\hat{\mu}) + \text{Var}(\hat{\mu}).
\end{align*}
\]

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

<table>
<thead>
<tr>
<th>(\hat{\mu})</th>
<th>Bias((\hat{\mu}))</th>
<th>Var((\hat{\mu}))</th>
<th>MSE((\hat{\mu}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{\mu}_{\text{NE}} = 1)</td>
<td>(1 - \frac{1}{p})</td>
<td>0</td>
<td>(\frac{(1-p)^2}{p^2})</td>
</tr>
<tr>
<td>(\hat{\mu}_{\text{MME}} = \frac{t_n - t_1}{n-1})</td>
<td>(0)</td>
<td>(\frac{1-p}{p^2(n-1)})</td>
<td>(\frac{1-p}{p^2(n-1)})</td>
</tr>
<tr>
<td>(\hat{\mu}_{\text{LRE}} = \frac{i \cdot \bar{t} - \bar{t} \cdot \bar{t}}{i^2 - (\bar{t})^2})</td>
<td>(0)</td>
<td>(\frac{6(n^2+1)(1-p)}{5n(n^2-1)p^2})</td>
<td>(\frac{6(n^2+1)(1-p)}{5n(n^2-1)p^2})</td>
</tr>
</tbody>
</table>
Comparison of Estimators

- Compare the performance of the estimators

\[
\begin{align*}
\text{Bias}(\hat{\mu}) &= E(\hat{\mu}) - \mu \\
\text{Var}(\hat{\mu}) &= E[(\hat{\mu} - E(\hat{\mu}))^2] \\
\text{MSE}(\hat{\mu}) &= E[(\hat{\mu} - \mu)^2] = \text{Bias}^2(\hat{\mu}) + \text{Var}(\hat{\mu}).
\end{align*}
\]

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

<table>
<thead>
<tr>
<th>$\hat{\mu}$</th>
<th>Bias($\hat{\mu}$)</th>
<th>Var($\hat{\mu}$)</th>
<th>MSE($\hat{\mu}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\mu}_{NE} = 1$</td>
<td>$1 - \frac{1}{p}$</td>
<td>0</td>
<td>$\frac{(1-p)^2}{p^2}$</td>
</tr>
<tr>
<td>$\hat{\mu}_{MME} = \frac{t_n - t_1}{n-1}$</td>
<td>0</td>
<td>$\frac{1-p}{p^2(n-1)}$</td>
<td>$\frac{1-p}{p^2(n-1)}$</td>
</tr>
<tr>
<td>$\hat{\mu}_{LRE} = \frac{i \cdot t - \bar{t} \cdot \bar{t}}{i^2 - (\bar{t})^2}$</td>
<td>0</td>
<td>$\frac{6(n^2+1)(1-p)}{5n(n^2-1)p^2}$</td>
<td>$\frac{6(n^2+1)(1-p)}{5n(n^2-1)p^2}$</td>
</tr>
</tbody>
</table>
Comparison of Estimators

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

<table>
<thead>
<tr>
<th>$\hat{t}_0$</th>
<th>Bias($\hat{t}_0$)</th>
<th>Var($\hat{t}_0$)</th>
<th>MSE($\hat{t}_0$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{t}<em>{0\text{NE}} = t_1 - \hat{\mu}</em>{\text{NE}}$</td>
<td>$\frac{1-p}{p}$</td>
<td>$\frac{1-p}{p^2}$</td>
<td>$\frac{(1-p)(2-p)}{p^2}$</td>
</tr>
<tr>
<td>$\hat{t}<em>{0\text{MME}} = t_1 - \hat{\mu}</em>{\text{MME}}$</td>
<td>0</td>
<td>$\frac{1-p}{p^2} \cdot \frac{n}{n-1}$</td>
<td>$\frac{1-p}{p^2} \cdot \frac{n}{n-1}$</td>
</tr>
<tr>
<td>$\hat{t}<em>{0\text{LRE}} = t_1 - \hat{\mu}</em>{\text{LRE}}$</td>
<td>0</td>
<td>$\frac{1-p}{p^2} \cdot \frac{5n^3+6n^2-5n+6}{5n(n^2-1)}$</td>
<td>$\frac{1-p}{p^2} \cdot \frac{5n^3+6n^2-5n+6}{5n(n^2-1)}$</td>
</tr>
</tbody>
</table>

Theorem

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

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

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

<table>
<thead>
<tr>
<th>$\hat{t}_0$</th>
<th>Bias($\hat{t}_0$)</th>
<th>Var($\hat{t}_0$)</th>
<th>MSE($\hat{t}_0$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{t}<em>{0\text{NE}} = t_1 - \hat{\mu}</em>{\text{NE}}$</td>
<td>$\frac{1-p}{p}$</td>
<td>$\frac{1-p}{p^2}$</td>
<td>$\frac{(1-p)(2-p)}{p^2}$</td>
</tr>
<tr>
<td>$\hat{t}<em>{0\text{MME}} = t_1 - \hat{\mu}</em>{\text{MME}}$</td>
<td>0</td>
<td>$\frac{1-p}{p^2} \cdot \frac{n}{n-1}$</td>
<td>$\frac{1-p}{p^2} \cdot \frac{n}{n-1}$</td>
</tr>
<tr>
<td>$\hat{t}<em>{0\text{LRE}} = t_1 - \hat{\mu}</em>{\text{LRE}}$</td>
<td>0</td>
<td>$\frac{1-p}{p^2} \cdot \frac{5n^3+6n^2-5n+6}{5n(n^2-1)}$</td>
<td>$\frac{1-p}{p^2} \cdot \frac{5n^3+6n^2-5n+6}{5n(n^2-1)}$</td>
</tr>
</tbody>
</table>

Theorem

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

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

1. Introduction
   - What is the problem?
   - What are we going to do?

2. Estimating the Host Infection Time
   - Estimating the Host Infection Time
   - Comparison of Estimators
   - Simulation Results

3. Estimating the Worm Infection Sequence
   - An Illustrated Scenario
   - Simulation Results

4. Summary and Future Works
Simulation Results

(a) Changing $\Omega$.  
(b) Changing $s$.  
(c) Changing $T$.

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

1. Introduction
   - What is the problem?
   - What are we going to do?

2. Estimating the Host Infection Time
   - Estimating the Host Infection Time
   - Comparison of Estimators
   - Simulation Results

3. Estimating the Worm Infection Sequence
   - An Illustrated Scenario
   - Simulation Results

4. Summary and Future Works
How it works?

A scenario of the worm infection sequence.

- $Pr_{NE}(error) = Pr(t_{1A} - 1 > t_{1B} - 1)$.
- $Pr(error) = Pr(t_{1A} - \frac{1}{\rho_A} > t_{1B} - \frac{1}{\rho_B})$.

Probability of Error Detection

$E[Pr_{NE}(error)] > E[Pr(error)]$. 
How it works?

A scenario of the worm infection sequence.

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

Probability of Error Detection

\[ \mathbb{E}[\Pr_{\text{NE}}(\text{error})] > \mathbb{E}[\Pr(\text{error})]. \]
Outline

1. Introduction
   - What is the problem?
   - What are we going to do?

2. Estimating the Host Infection Time
   - Estimating the Host Infection Time
   - Comparison of Estimators
   - Simulation Results

3. Estimating the Worm Infection Sequence
   - An Illustrated Scenario
   - Simulation Results

4. Summary and Future Works
Simulation Results

**Table:** A sample run of simulations.

<table>
<thead>
<tr>
<th>$S_i$</th>
<th>$\hat{S}_{iNE}$</th>
<th>$\hat{S}_{iMME}$</th>
<th>$\hat{S}_{iLRE}$</th>
<th>$t_0$</th>
<th>$\hat{t}_{0NE}$</th>
<th>$\hat{t}_{0MME}$</th>
<th>$\hat{t}_{0LRE}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>114</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>85</td>
<td>98</td>
<td>74</td>
<td>73</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>105</td>
<td>165</td>
<td>116</td>
<td>116</td>
</tr>
<tr>
<td>520</td>
<td>498</td>
<td>533</td>
<td>534</td>
<td>593</td>
<td>622</td>
<td>589</td>
<td>589</td>
</tr>
<tr>
<td>521</td>
<td>433</td>
<td>488</td>
<td>477</td>
<td>594</td>
<td>611</td>
<td>581</td>
<td>580</td>
</tr>
</tbody>
</table>

**Sequence Distance**

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

**Table:** A sample run of simulations.

<table>
<thead>
<tr>
<th>$S_i$</th>
<th>$\hat{S}_{i\text{NE}}$</th>
<th>$\hat{S}_{i\text{MME}}$</th>
<th>$\hat{S}_{i\text{LRE}}$</th>
<th>$t_0$</th>
<th>$\hat{t}_{0\text{NE}}$</th>
<th>$\hat{t}_{0\text{MME}}$</th>
<th>$\hat{t}_{0\text{LRE}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>114</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>85</td>
<td>98</td>
<td>74</td>
<td>73</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>105</td>
<td>165</td>
<td>116</td>
<td>116</td>
</tr>
<tr>
<td>520</td>
<td>498</td>
<td>533</td>
<td>534</td>
<td>593</td>
<td>622</td>
<td>589</td>
<td>589</td>
</tr>
<tr>
<td>521</td>
<td>433</td>
<td>488</td>
<td>477</td>
<td>594</td>
<td>611</td>
<td>581</td>
<td>580</td>
</tr>
</tbody>
</table>

Sequence Distance

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

(a) Changing $\Omega$

(b) Changing $\sigma$

(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

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

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

**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

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