On Estimating Tight-Link Bandwidth Characteristics over Multi-Hop Path

Seong-Ryong Kang

Joint work with Xiliang Liu, Amit Bhati, and Dmitri Loguinov

Internet Research Lab
Department of Computer Science
Texas A&M University, College Station, TX 77843

July 6, 2006
Agenda

• Introduction
  ─ Motivations and goals

• Background
  ─ Definition of bandwidth
  ─ End-to-end Internet-path and single-hop model

• Envelope
  ─ Envelope packet-trains
  ─ Phase-based individual link measurement

• Performance of Envelope
  ─ Estimation accuracy and asymptotic behavior

• Wrap-up
Motivation

• Bandwidth estimation is an important area of Internet research
  — Helps to understand network path characteristics
  — Potentially can help various Internet applications

• Majority of available bandwidth estimation processes do not provably converge to the correct values
Motivation 2

• Furthermore, none of the existing techniques can correctly measure the tight-link capacity over a multi-hop path
  — Tight link is the link with the minimum available bandwidth of a path
Goals

• Develop a provably accurate estimation technique
• Measure both capacity and available bandwidth of the tight link
• Work for multi-hop paths under arbitrary cross-traffic
**Definition of Bandwidth**

- **Bottleneck bandwidth**
  - The capacity of the slowest link of an end-to-end path
  - The slowest link is often called *narrow* link

- **Available bandwidth**
  - The smallest average unused bandwidth along the end-to-end path
**Definition Bandwidth 2**

- Available bandwidth of the path: \( A = 12 \)

```
S → R_1 (50) → R_2 (20) → R_3 (40) → R_4 → D
```

- Tight link, capacity = 50
- Narrow link bottleneck capacity = 20
End-to-End Internet Path

The sender injects probe packets with inter-packet spacing \( x \)

Due to expansion/compression in pre-tight links, inter-packet spacing \( x_t \) is different from the initial spacing \( x \)

\( x_t \) is altered at router \( R_t \) to be \( y_t \) by random noise \( w_t \)

The receiver samples inter-arrival dispersion \( y \)
**Single-Hop Model**

- Assumes that cross-traffic in non-tight links does not change inter-packet spacings of the probe packets
  - That is, $x = x_t$ and $y = y_t$

- Derives the mean output dispersion $E[y]$ under arbitrary cross-traffic
  - Other single-hop models (Dovrolis *et al.* INFOCOM 2001 and Melander *et al.* GLOBECOM 2000) rely on a constant-rate fluid model of cross-traffic
Single-Hop Model 2

- Extracts capacity and available bandwidth of the tight link from $E[y]$ with high accuracy
- For details, see the paper (Kang et al. INCP 2004)
Envelopes

- Recursively extends the single-hop results to multi-hop paths
  - Sample statistics of each hop independently

- Measures both capacity and available bandwidth of the tight link under arbitrary cross-traffic
  - Can also measure non-tight link in certain path and cross-traffic conditions

- Provides asymptotic accuracy
  - Estimates converge to the true value after a sufficiently long measurement process
• Recursive extension:
  — Treats inter-packet spacing $x_k$ of probe traffic arriving at router $R_k$ as the inter-departure delay $y_{k-1}$ of the previous router $R_{k-1}$

\[ y_{k-1} = x_k \]

\[ y_k = x_{k+1} \]
• Necessary conditions for measuring link $R_k$
  — Spacing between two probe packets must be small when arriving at router $R_k$
  — However, the departure spacing from the router must be large to preserve its mean along the path suffix

• How do we satisfy these conditions?
  — By using Envelope packet-trains and TTL-limited dropping of probe packets at select routers
  — This does not require special “cooperation” from the routers
Envelope 4

• Envelope packet train

- An envelope packet-train includes $N$ probe packets $P_1$, $..., P_N$ surrounded by two Envelope packets $E_1$ and $E_2$
- Delays between two probe packets are small
- Delays between two Envelope packets become large by selecting a large $N$
To obtain the departure spacing from router $R_k$, all probe packets $P_1, \ldots, P_N$ are dropped at router $R_{k+1}$ using TTL limiting.

The probe packets sample queueing dynamics at the desired router $R_k$.

The surviving envelope packets carry spacing $z_k$ that is $N+1$ times larger than the departure spacing $y_k$ from the router $R_k$. 
Envelope 6

• For $M$ links in the end-to-end path, Envelope takes $M-1$ measurement phases
  — Each phase focuses on a particular router $R_k$
  — Obtains the mean spacing $E[z_k]$ exiting from $R_k$

• Using $E[z_{k-1}]$ measured in the previous phase, Envelope estimates the available bandwidth $A_k$ and capacity $C_k$ of the router $R_k$

• For details of bandwidth extraction process, see the paper
Simulation Topology

- $S_1$ to $R_1$
  - 100 Mb/s
  - 5 ms
- $R_1$ to $R_2$
  - 100 Mb/s
  - 5 ms
- $R_2$ to $S_2$
  - 100 Mb/s
  - 5 ms
- $S_2$ to $D_2$
  - 100 Mb/s
  - 5 ms
- $D_2$ to $S_3$
  - 100 Mb/s
  - 5 ms
- $S_3$ to $R_3$
  - 100 Mb/s
  - 5 ms
- $R_3$ to $C_2$
  - 100 Mb/s
  - 5 ms
- $C_2$ to $R_4$
  - 100 Mb/s
  - 5 ms
- $R_4$ to $S_4$
  - 100 Mb/s
  - 5 ms
- $S_4$ to $D_4$
  - 100 Mb/s
  - 5 ms
- $D_4$ to $PR$
  - 100 Mb/s
  - 5 ms
- $PR$ to $C_4$
  - 100 Mb/s
  - 5 ms
- $C_4$ to $R_5$
  - 100 Mb/s
  - 5 ms
- $R_5$ to $R_1$
  - 100 Mb/s
  - 5 ms
## Simulation Setup

<table>
<thead>
<tr>
<th></th>
<th>$C_1$</th>
<th>$A_1$</th>
<th>$C_2$</th>
<th>$A_2$</th>
<th>$C_3$</th>
<th>$A_3$</th>
<th>$C_4$</th>
<th>$A_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case-I</td>
<td>5</td>
<td>1</td>
<td>100</td>
<td>50</td>
<td>100</td>
<td>40</td>
<td>1.5</td>
<td>0.3</td>
</tr>
<tr>
<td>Case-II</td>
<td>2</td>
<td>0.4</td>
<td>1.5</td>
<td>0.25</td>
<td>0.8</td>
<td>0.4</td>
<td>1.5</td>
<td>0.35</td>
</tr>
<tr>
<td>Case-III</td>
<td>1.5</td>
<td>0.3</td>
<td>100</td>
<td>50</td>
<td>100</td>
<td>40</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Case-IV</td>
<td>20</td>
<td>4</td>
<td>15</td>
<td>2.5</td>
<td>8</td>
<td>4</td>
<td>15</td>
<td>3.5</td>
</tr>
<tr>
<td>Case-V</td>
<td>2</td>
<td>0.4</td>
<td>0.8</td>
<td>0.4</td>
<td>1.5</td>
<td>0.25</td>
<td>2</td>
<td>0.4</td>
</tr>
</tbody>
</table>

- Darkly shaded values in each row represent the tight-link capacity and available bandwidth of the path.
- We also lightly shade the narrow-link capacity in cases when it is different from the tight link.
Estimation Accuracy

- Relative error metrics:

\[ e_{C_i} = \frac{|C_i - C'_i|}{C_i}, \quad e_{A_i} = \frac{|A_i - A'_i|}{A_i} \]

- \( e_{C_i} \) and \( e_{A_i} \) are relative estimation errors of \( C_i \) and \( A_i \), respectively
- \( C_i \) is the true capacity of link \( i \) and \( C'_i \) is its estimate
- \( A_i \) is the true available bandwidth of link \( i \) and \( A'_i \) is its estimate
### Estimation Accuracy 2

- **CBR cross-traffic**

<table>
<thead>
<tr>
<th></th>
<th>Case-I</th>
<th>Case-II</th>
<th>Case-III</th>
<th>Case-IV</th>
<th>Case-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_{C_1}$</td>
<td>0.94%</td>
<td>2.39%</td>
<td>0.17%</td>
<td>0.15%</td>
<td>10.76%</td>
</tr>
<tr>
<td>$e_{A_1}$</td>
<td>7.75%</td>
<td>1.57%</td>
<td>3.74%</td>
<td>6.99%</td>
<td>4.20%</td>
</tr>
<tr>
<td>$e_{C_2}$</td>
<td>—</td>
<td>0.35%</td>
<td>—</td>
<td>2.36%</td>
<td>2.47%</td>
</tr>
<tr>
<td>$e_{A_2}$</td>
<td>—</td>
<td>2.09%</td>
<td>—</td>
<td>5.62%</td>
<td>8.71%</td>
</tr>
<tr>
<td>$e_{C_3}$</td>
<td>—</td>
<td>3.76%</td>
<td>—</td>
<td>0.65%</td>
<td>4.13%</td>
</tr>
<tr>
<td>$e_{A_3}$</td>
<td>—</td>
<td>7.07%</td>
<td>—</td>
<td>2.04%</td>
<td>5.71%</td>
</tr>
<tr>
<td>$e_{C_4}$</td>
<td>1.56%</td>
<td>0.60%</td>
<td>—</td>
<td>12.11%</td>
<td>21.19%</td>
</tr>
<tr>
<td>$e_{A_4}$</td>
<td>2.38%</td>
<td>3.05%</td>
<td>—</td>
<td>9.86%</td>
<td>17.59%</td>
</tr>
</tbody>
</table>
### Estimation Accuracy 3

- TCP cross-traffic

<table>
<thead>
<tr>
<th></th>
<th>Case-I</th>
<th>Case-II</th>
<th>Case-III</th>
<th>Case-IV</th>
<th>Case-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_{C_1}$</td>
<td>0.21%</td>
<td>0.40%</td>
<td>0.46%</td>
<td>8.55%</td>
<td>4.22%</td>
</tr>
<tr>
<td>$e_{A_1}$</td>
<td>12.07%</td>
<td>0.27%</td>
<td>0.98%</td>
<td>21.23%</td>
<td>16.60%</td>
</tr>
<tr>
<td>$e_{C_2}$</td>
<td>—</td>
<td>3.62%</td>
<td>—</td>
<td>0.26%</td>
<td>5.90%</td>
</tr>
<tr>
<td>$e_{A_2}$</td>
<td>—</td>
<td>4.22%</td>
<td>—</td>
<td>3.29%</td>
<td>10.06%</td>
</tr>
<tr>
<td>$e_{C_3}$</td>
<td>—</td>
<td>10.79%</td>
<td>—</td>
<td>9.41%</td>
<td>10.06%</td>
</tr>
<tr>
<td>$e_{A_3}$</td>
<td>—</td>
<td>15.44%</td>
<td>—</td>
<td>23.30%</td>
<td>5.82%</td>
</tr>
<tr>
<td>$e_{C_4}$</td>
<td>0.24%</td>
<td>10.04%</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>$e_{A_4}$</td>
<td>3.30%</td>
<td>11.53%</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>
Asymptotic Behavior

- CBR cross-traffic (case I)

(a) $e_C$ for case I

(b) $e_A$ for case I
Asymptotic Behavior 2

- TCP cross-traffic (case I)

(a) $e_C$ for case I

(b) $e_A$ for case I
**Performance Comparison**

- Available bandwidth under TCP cross-traffic

<table>
<thead>
<tr>
<th>Case</th>
<th>Relative estimation error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Envelope</td>
</tr>
<tr>
<td>Case-I</td>
<td>3.30%</td>
</tr>
<tr>
<td>Case-II</td>
<td>4.22%</td>
</tr>
<tr>
<td>Case-III</td>
<td>0.98%</td>
</tr>
<tr>
<td>Case-IV</td>
<td>3.29%</td>
</tr>
<tr>
<td>Case-V</td>
<td>5.82%</td>
</tr>
</tbody>
</table>
## Performance Comparison 2

- Bottleneck bandwidth under TCP cross-traffic

<table>
<thead>
<tr>
<th></th>
<th>Relative estimation error</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Envelope</td>
<td>CapProbe</td>
<td>Pathrate</td>
<td></td>
</tr>
<tr>
<td>Case-I</td>
<td>0.24%</td>
<td>40.95%</td>
<td>40.93%</td>
<td></td>
</tr>
<tr>
<td>Case-II</td>
<td>10.79%</td>
<td>39.12%</td>
<td>32.50%</td>
<td></td>
</tr>
<tr>
<td>Case-III</td>
<td>0.46%</td>
<td>35.78%</td>
<td>48.10%</td>
<td></td>
</tr>
<tr>
<td>Case-IV</td>
<td>9.41%</td>
<td>50.60%</td>
<td>20.62%</td>
<td></td>
</tr>
<tr>
<td>Case-V</td>
<td>5.90%</td>
<td>51.62%</td>
<td>45.62%</td>
<td></td>
</tr>
</tbody>
</table>
Wrap-up

- Envelope measures both bandwidth metrics of the tight-link under arbitrary cross-traffic
- Its estimates are asymptotically accurate
- It is based on recursive extension of the single-hop results to multi-hop paths
- Future work
  - Implementation and deployment of Envelope
  - Further reduction of probe traffic required