Exploring and Improving BitTorrent Topologies

Christian Decker
BitTorrent

- Filesharing protocol
- Peers form Ad-Hoc networks (swarm)
- Trackers to join the swarm
- Trading pieces between peers
BitTorrent: handshake

- Peers exchange handshakes before trading
  - Protocol identifier
  - Protocol extensions
  - peer_id
  - Torrent info_hash
Exploring
Exploring Swarm Topologies

Random topology:

- Tracker return random peers
- Peers chose random neighbors
Exploring Swarm Topologies

Random topology:
  • Tracker return random peers
  • Peers chose random neighbors

Studies to explore BitTorrent topologies:
  • Experimental setup
  • Traffic log
  • Instrumented clients
Exploring Swarm Topologies

Random topology:
- Tracker return random peers
- Peers chose random neighbors

Studies to explore BitTorrent topologies:
- Experimental setup
- Traffic log
- Instrumented clients
- *Live swarms*
Scanning method

A

B

S

"Hello I’m A"
Challenges when moving to real swarms

- Scanning takes time
Challenges when moving to real swarms

- Scanning takes time
- Invisible part of a swarm
Challenges when moving to real swarms

- Scanning takes time
- Invisible part of a swarm
- Error detection

CDF of connection uptimes

CDF of connections

Uptime [s]

Uptime

CDF of connections

0 500 1000 1500 2000 2500 3000 3500

0.0 0.2 0.4 0.6 0.8 1.0

CDF of connection uptimes

Uptime [s]
Evaluation: topology sample size

58% peers cannot be scanned
Evaluation: topology sample size

58% peers cannot be scanned

But we can scan either endpoint of a connection:
Evaluation: coverage

How fast can we scan for all possible connections?

![Graph showing coverage vs. swarm size for different concurrency levels.](image-url)
Evaluation: coverage

How fast can we scan for all possible connections?

![Graph showing coverage across different scan times and concurrency levels.](image-url)
Trading with random peers, that may be halfway around the globe. Closer peers may be available.

\[ \sigma(a, b) = \begin{cases} 
1 & \text{a and } b \text{ are connected} \\
0 & \text{otherwise} 
\end{cases} \]
Locality

Trading with random peers, that may be halfway around the globe. Closer peers may be available.

\[ \sigma(a, b) = \begin{cases} 
1 & a \text{ and } b \text{ are connected} \\
0 & \text{otherwise} 
\end{cases} \]

\[ L = \frac{\sum_{a,b} \sigma(a, b) \cdot d(a, b)}{\sum_{a,b} \sigma(a, b)} \cdot \mathbb{E}[D]^{-1} \]
Trading with random peers, that may be halfway around the globe. Closer peers may be available.

\[ \sigma(a, b) = \begin{cases} 
1 & \text{a and } b \text{ are connected} \\
0 & \text{otherwise}
\end{cases} \]

\[ \mathcal{L} = \frac{\sum_{a,b} \sigma(a, b) \cdot d(a, b)}{\sum_{a,b} \sigma(a, b)} \cdot \mathbb{E}[D]^{-1} \approx 1 \]
Locality: evaluation

- $\sigma(a, b)$ provided from scanning method
- $d(a, b)$ provided by MaxMind GeoIP Database

$L = 1.062 > 1$
Locality: evaluation

- $\sigma(a, b)$ provided from scanning method
- $d(a, b)$ provided by MaxMind GeoIP Database
- 33 recently uploaded torrents
- International torrents
- 50-500 peers
Locality: evaluation

- $\sigma(a, b)$ provided from scanning method
- $d(a, b)$ provided by MaxMind GeoIP Database
- 33 recently uploaded torrents
- International torrents
- 50-500 peers

BitTorrent is not locality aware!

$$L = 1.062 > 1$$
Improving
Suggesting

Peer Exchange (PEX):
- Reduce tracker load
- Increase trading partners
Peer Exchange (PEX):

- Reduce tracker load
- Increase trading partners
- *Suggest nearby peers*
Suggesting peers with PEX

1. Identify new peers
2. Find nearby peers
3. Connect to new peers
4. Send suggestions as PEX message
Suggesting peers with PEX

- No special access
- No shaping or blocking
- Widely supported
- ISPs get information for free
Suggesting peers with PEX

+ No special access
+ No shaping or blocking
+ Widely supported
+ ISPs get information for free
  - Limited influence
How are we doing?

- $\mathcal{L}: 1.062 \rightarrow 0.994$.
- 6.3% improvement

Connection length comparison

Influenced Swarms
Untouched Swarms
Conclusion

A

B

? "Hello I'm A"

S

Influenced Swarms

Untouched Swarms

Connection length comparison

Connection length [km]

Probability

0

2000 4000 6000 8000 10000 12000 14000 16000 18000 20000

0.00 0.05 0.10 0.15 0.20 0.25 0.30 ●

●
Conclusion

A

B

"Hello I'm A"

S

?  

B

"You'll like A"

S

Connection length comparison

<table>
<thead>
<tr>
<th>Connection length [km]</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>2000</td>
<td>0.05</td>
</tr>
<tr>
<td>4000</td>
<td>0.10</td>
</tr>
<tr>
<td>6000</td>
<td>0.15</td>
</tr>
<tr>
<td>8000</td>
<td>0.20</td>
</tr>
<tr>
<td>10000</td>
<td>0.25</td>
</tr>
<tr>
<td>12000</td>
<td>0.30</td>
</tr>
</tbody>
</table>

●

Influenced Swarms

Untouched Swarms
Thank you, questions?