Capturing User Friendship in WLAN Traces

Wei-jen Hsu and Ahmed Helmy

Department of Electrical Engineering
University of Southern California
{weijenhs, helmy}@usc.edu
http://nile.usc.edu/MobiLib/

INFOCOM Poster and Demo Session 2006
Motivation

• Intuitively, sub-groups of people on university campuses show closer relationship.
  – How is such “relationship” reflected in the WLAN traces?
  – How to define metrics for such behavior?
  – What are the inferences of these sub-groups of friends on network structures?

• We use WLAN from 4 different sources to understand grouping behavior realistically.
Wireless LAN traces used

<table>
<thead>
<tr>
<th>Trace source</th>
<th>Unique users</th>
<th>Unique APs</th>
<th>Unique buildings</th>
<th>Trace duration</th>
<th>User type</th>
<th>Environment</th>
<th>Trace collection method</th>
<th>Analyzed part in this paper</th>
<th>Users in analyzed part</th>
<th>Labels used in graphs</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIT[1]</td>
<td>1,366</td>
<td>173</td>
<td>3</td>
<td>Jul. 20 '02 to Aug. 17 '02</td>
<td>Generic</td>
<td>3 Engineer buildings</td>
<td>Polling</td>
<td>Whole trace</td>
<td>1,366</td>
<td>MIT-cons MIT-rel</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mar. 2004</td>
<td>5,416</td>
<td>Dart-04 Dart-rel Dart-cons</td>
</tr>
<tr>
<td>UCSD[2]</td>
<td>275</td>
<td>518</td>
<td>N/A</td>
<td>Sep. 22 '02 to Dec. 8 '02</td>
<td>PDA only</td>
<td>Whole campus</td>
<td>Polling</td>
<td>Sep. 22 '02 to Oct. 21 '02</td>
<td>275</td>
<td>UCSD</td>
</tr>
<tr>
<td>USC</td>
<td>4,548</td>
<td>79</td>
<td>73</td>
<td>Dec 03-Now (trap) Apr 20 05-Now (detail)</td>
<td>Generic</td>
<td>Whole campus</td>
<td>Event-based</td>
<td>Apr. 20, '05 to May. 19 '05</td>
<td>4,528</td>
<td>USC</td>
</tr>
</tbody>
</table>

- Traces from environments with various settings.
- In each trace we have AP association history of individual nodes.
- Objective: Capture “close relationship” between nodes.

INFOCOM Poster and Demo Session 2006
Friendship Index

- Intuitively, if two nodes appear together “more often”, it indicates closer relationship between them.

- Define friendship index based on two metrics:
  - Encounter time between the node pair
    \[ Frd_t (A, B) = \frac{E_t (A, B)}{Online\_time(A)} \]
  - Encounter count between the node pair
    \[ Frd_c (A, B) = \frac{E_c (A, B)}{Session(A)} \]

\( Et(A,B) \): Total encounter duration between node A, B.
\( Ec(A,B) \): Encounter count between A,B
\( Online\_time(A) \): Total online duration for node A
\( Session(A) \): Total number of sessions for node A
Friendship Index

- Friendship is very skewed: Few pairs of nodes have high friendship index (Exponential dist.)
Encounter-Relationship Graph

- Put a link to connect the node pairs if they ever encounter with each other… What does the graph look like?

**Regular Graph**
- High path length
- High clustering

**Random Graph**
- Low path length
- Low clustering

**Small World graph**

High clustering as regular graph
Low path length as random graph
Encounter-Relationship Graph

- It is a connected graph
  Disconnected ratio drops to below 10% for trace duration longer than 1 day.

- It is a SmallWorld graph
  Normalized clustering coef. (CC) close to regular graph and average Path Length (PL) close to random graph

Normalized PL/CC: 1 corresponds to regular graphs, 0 corresponds to random graphs
ER Graph with Friends

• Sort the encountered nodes according to friendship indexes for each node, and include only part of them in the ER graph. How does it change the graph property?
  – Higher tendency of clustering if only top-friends are included.
  – Higher average path length and disconnected ratio if only top-friends are included.
  – The above observations are consistent regardless of which definition of friendship index is chosen.
ER Graph with Friends

- Using friendship index based on *encounter time*

**Clustering Coefficient**

**Avg. Path Length**

- Using high-ranked friends only in the ER graph leads to graph properties closer to regular graphs. Using low-ranked friends leads to graph properties closer to random graphs.

INFOCOM Poster and Demo Session 2006
ER Graph with Friends

- Using friendship index based on encounter count

- Using high-ranked friends only in the ER graph leads to graph properties closer to regular graphs. Using low-ranked friends leads to graph properties closer to random graphs.

INFOCOM Poster and Demo Session 2006
Conclusion

- Friendship between nodes in WLAN traces defined based on encounter time or encounter count.
- Friendship is skewed: Few node pairs with high friendship index, many with low friendship index.
- Encounters link nodes into connected SmallWorld graphs.
- Including nodes with high friendship indexes make the encounter-relationship graphs shift toward regular graphs.
Implication

• While people tend to trust others with close relationship, random links may be the key to maintain a connected network.

References


