Research Methods
in computer science
Fall 2021

Lecture 16

Omprakash Gnawali
October 18, 2021
Agenda

Conference Updates
System Comparison Experiments
Level of Details
HW8
CS Experiments Today

Artifact Comparison Experiments

Run the new artifact
Run best-known prior work
Compare

Simulations + “Real” experiments
Wireless Experiments Today

Protocol Comparison Experiments
Run the new protocol
Run best-known prior work
Compare

Simulations + Testbed experiments
Serial Experiments

Run one protocol at a time
Compare the results

Difficult to distinguish the contribution of these variables
Environment
Protocol mechanisms
Concurrent Experiments

Run multiple protocols concurrently
Compare the results

Advantages
Consistent environment for both the protocols

Concerns
Contention of different types
Results from Serial CTP vs LQI Experiment on Tutornet

- Delivery
- Cost
- Path Length
- Churn/node-hr

Comparison between CTP and LQI for different metrics.
Results from Concurrent CTP vs LQI Experiment on Tutornet

<table>
<thead>
<tr>
<th>Measure</th>
<th>LQI</th>
<th>CTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delivery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Path Length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Churn/node-hr</td>
<td></td>
<td>8</td>
</tr>
</tbody>
</table>

The graph shows a comparison of LQI and CTP across different measures such as delivery, cost, and path length. The churn/node-hr measure for CTP is significantly higher than for LQI.
Putting Concurrent Methodology to Use: Expts. with External Interference

Engineered Scenario

Both protocols *struggle* in the same environment.
Putting Concurrent Methodology to Use: Experiments in a Dynamic Network

CTP and LQI react differently to dynamics.
Uncontrolled environment does not imply we cannot do fair comparisons
Data

System 1

System 2
Level of Details
At What Level of Detail?

Descriptions
  System and algorithm
  Experiments
  Datasets
  Results
We use all available nodes in every experiment. In some testbeds, this means the set of nodes across experiments is almost but not completely identical, due to backchannel connectivity issues. However, we do not prune problem nodes. In the case of Motelab, this approach greatly affects the computed average performance, as some nodes are barely connected to the rest of the network.
5.1 Methodology

We conducted our experiments on a tiered network testbed with several Stargate nodes and 40 TelosB motes. All nodes are located above the false ceiling across multiple rooms and hallways on a floor of a large office building. The wireless environment above the false ceiling is harsh, with some links experiencing above 30% packet loss rates. All nodes run the Tenet stack modified to support AEM. In most experiments, we use a single Tenet master node. We configured the mote radios to transmit at -8.906 dBm, which results in a tree with 4-hop depth.
Experimental Methodology and Metrics  We now compare the performance of Tenet-PEG and mote-PEG. Our experiments are conducted on the testbed shown in Figure 7. This testbed consists of 56 Tmotes and 6 Stargates deployed above the false ceiling of a single floor of a large office building. The Stargate and mote radios are assigned non-interfering channels. This testbed represents a realistic setting for examining network performance as well as for evaluating PEGs. The false ceiling is heavily obstructed, so the wireless communication that we see is representative of harsh environments. The environment is also visually obstructed, and thus resembles say, a building after a disaster, in which a pursuit-evasion sensor network might aid the robotic search for survivors.
Results from the same Testbed
Network Metric

Converting these subjective descriptions to a more quantitative description
END and CTP Performance

[Graphs showing END vs. Delivery and Goodput vs. [pkts/sec] for Class A and Class B, with data points for CTP and Arbutus marked in different colors.]
The SFC dataset includes 4.4 million labeled faces from 4,030 people each with 800 to 1200 faces, where the most recent 5% of face images of each identity are left out for testing. This is done according to the images’ time-stamp in order to simulate continuous identification through aging. The large number of images per person provides a unique opportunity for learning the invariance needed for the core problem of face recognition...

“See the supplementary material for more details about SFC.”
Supplementary Material:
DeepFace: Closing the Gap to Human-Level Performance in Face Verification

Yaniv Taigman       Ming Yang       Marc’Aurelio Ranzato       Lior Wolf
Facebook AI Research     Menlo Park, CA, USA
Tel Aviv University     Tel Aviv, Israel
{yaniv, mingyang, ranzato}@fb.com     wolf@cs.tau.ac.il
“We evaluate the throughput and delay benefits of CQIC using the Google Nexus device to download content from a Google server via a popular cellular network provider. Reflecting a common CDN scenario, this server is located near the network of the mobile carrier such that the cellular channel is the bottleneck link...”

[Lu 2015]
HW8

Describe the results from your research.

Please include Data+graphs+insights.

About one page in length.