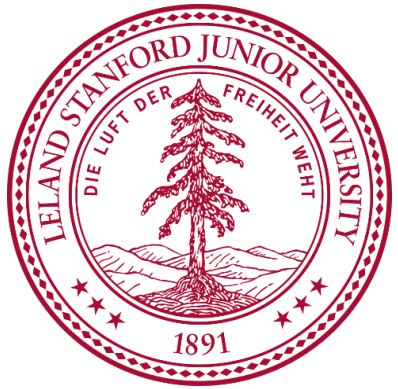
Broadcast LTE Data Reveals Application Type

Arjun Balasingam, Manu Bansal, Rakesh Misra, Rahul Tandra, Aaron Schulman, Sachin Katti

Mobicom 2017, Snowbird, UT



UHANA

UC San Diego

Growth in Mobile Connectivity

- Internet-connected mobile devices are ubiquitous
- Mobile traffic expected to increase 7x by 2021
- More data routed through cloud

Need for Secure Wireless Systems

- Cloud-based mobile apps are integral to our lives
- Potential to leave trail of personal data
- We need *secure* mobile systems!

Our Contributions

We show that it is possible to *infer the type of application being hosted by any radio session* from only its radio resource allocation patterns.

1. A phone can infer *its own* application type from radio-layer data.
2. *Anyone* can identify all applications being served by cell tower.

Prior Work

Analysis of PHY-layer Data

- Characterize congestion [LTEye]
- Video streaming [piStream]

Application Classification

- Other layers in the networking stack

Our Work

- First application classifier from broadcast radio-layer data
- Can be done *without* any privileged information

Data Collection and Experiments

Testbed

QXDM logs data at phone's modem

Application Classifier

- Parse QXDM logs
- Derive Metrics
- Apply classifier

LTE (Long Term Evolution)

- Standard for high-speed wireless communication today
- Transmitted data is sent over *uplink* or *downlink*
- Downlink channels with user data
  - PDSCH (encrypted)
  - PDCCH (broadcast)
- LTE layer = radio layer = PHY layer

In this work, we focus on data broadcast over PDCCH.

Our Application Classifier

- We classify a broad set of popular mobile applications:
  - File download
  - Web browsing
  - Video streaming
  - Video conferencing

The PROMINENCE Metric

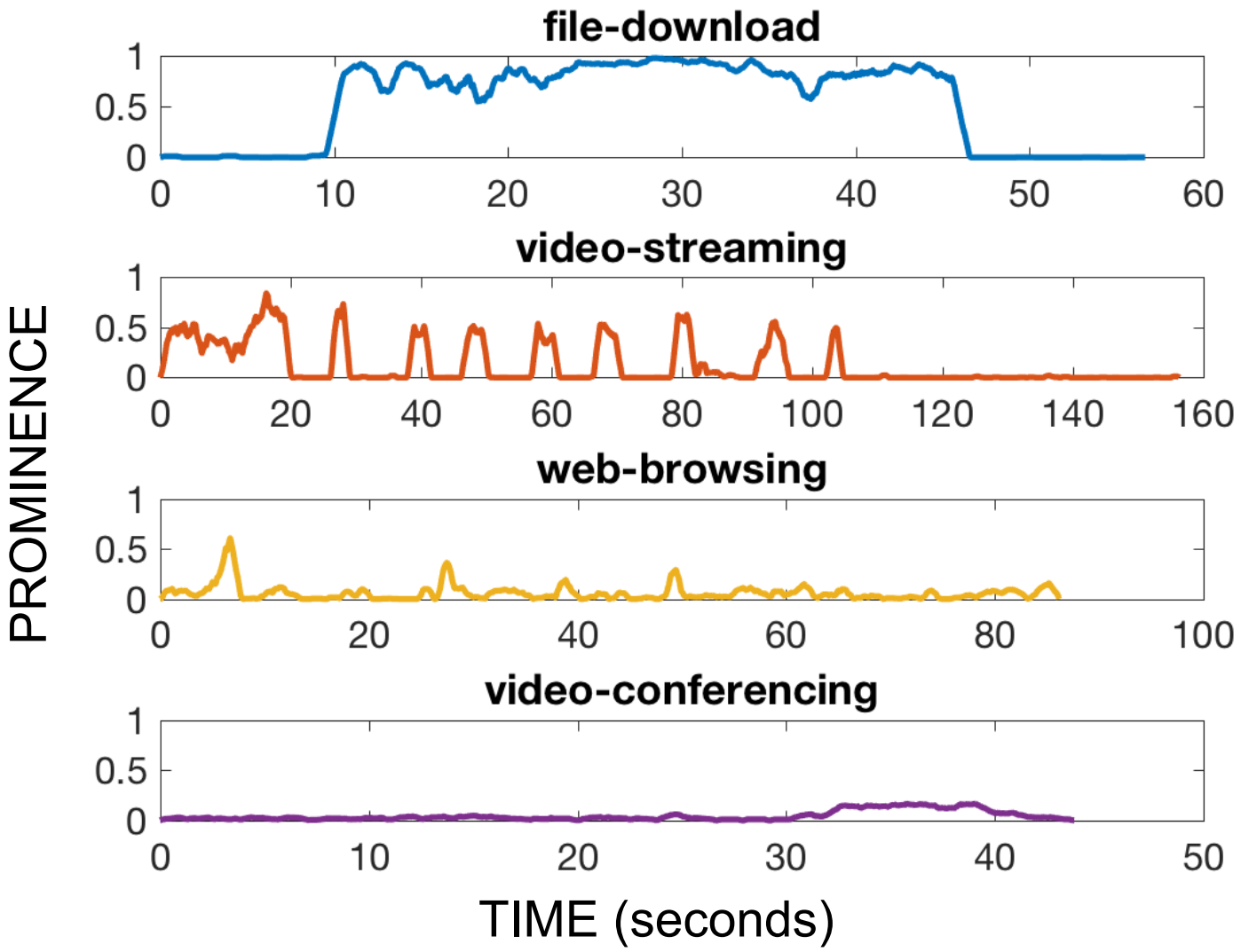
Derived Metrics

- SCHEDTIME
  - # of ms where session was scheduled resource blocks (RBs)
- SESSDUR
  - session duration (in ms)

$$\text{PROMINENCE} = \frac{\text{SCHEDTIME}}{\text{SESSDUR}}$$

- Simple to compute
- Captures traffic arrival patterns
- Abstracts out other session-specific factors

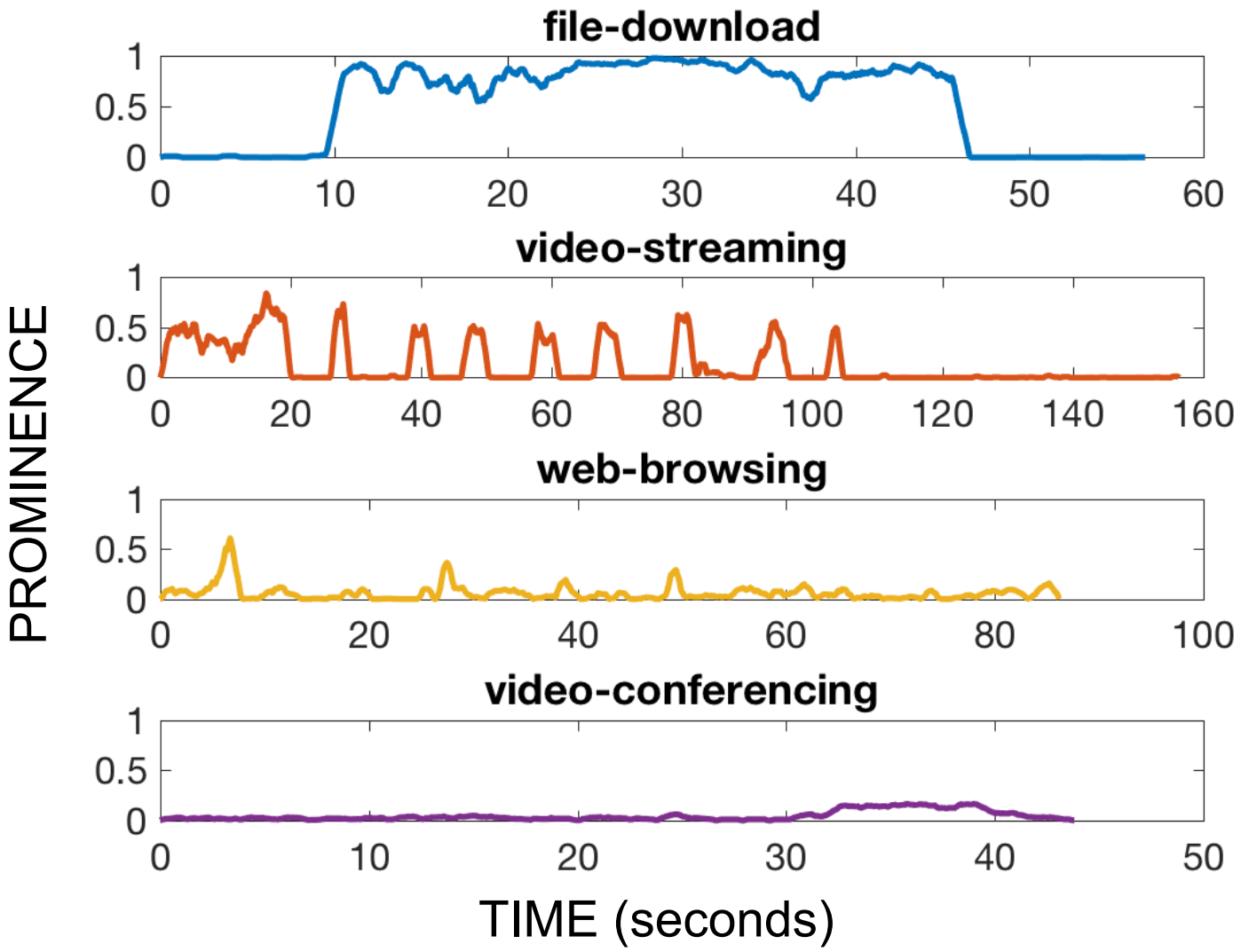
PROMINENCE Signatures



- Time series of PROMINENCE computed in a moving window of 1 second
- Different classes of applications have distinct PROMINENCE signatures

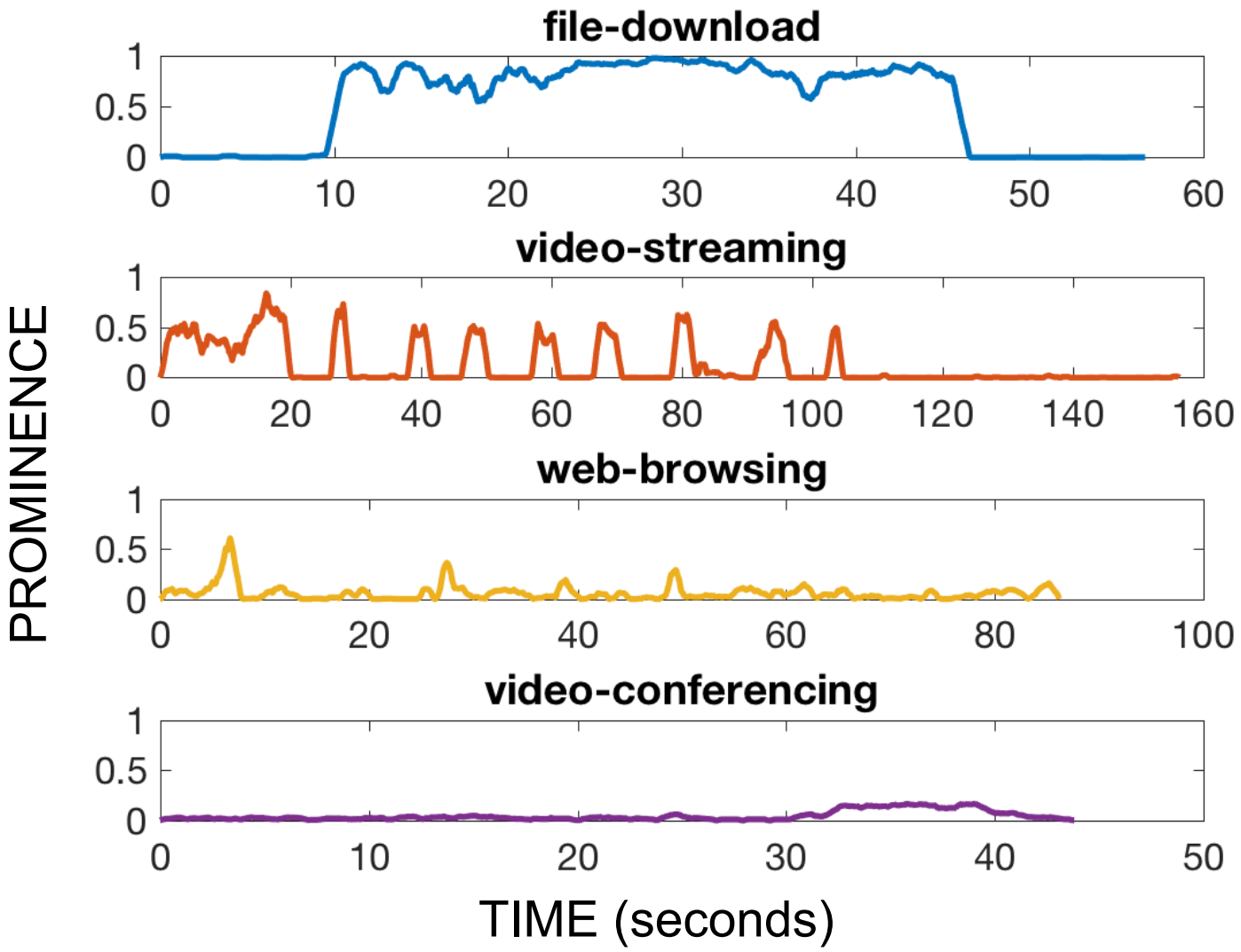
We leverage this insight to design our classifier.

PROMINENCE Signatures



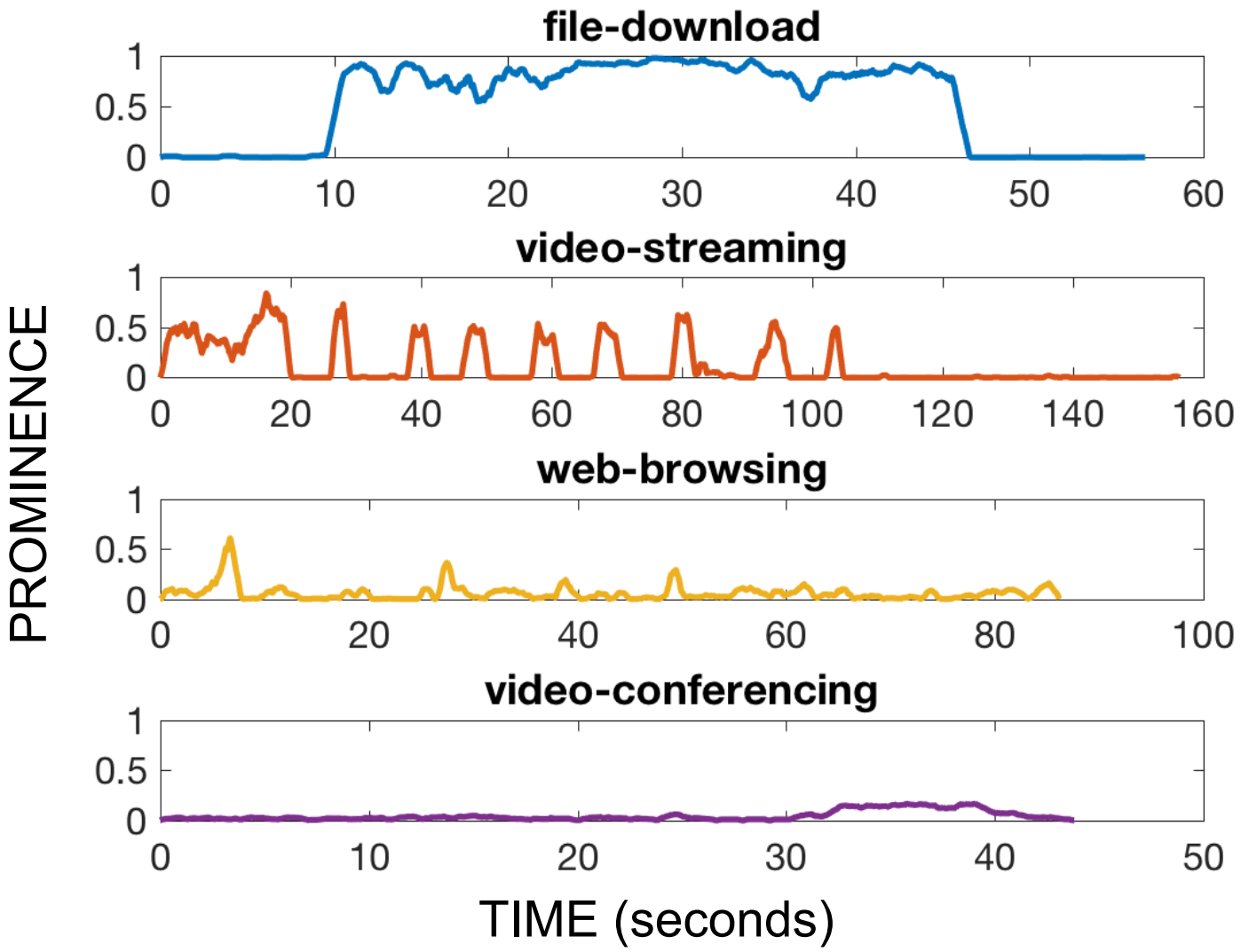
- File download
  - Full buffer
  - Highly prominent
  - Scheduled on 80% of RBs/sec

PROMINENCE Signatures



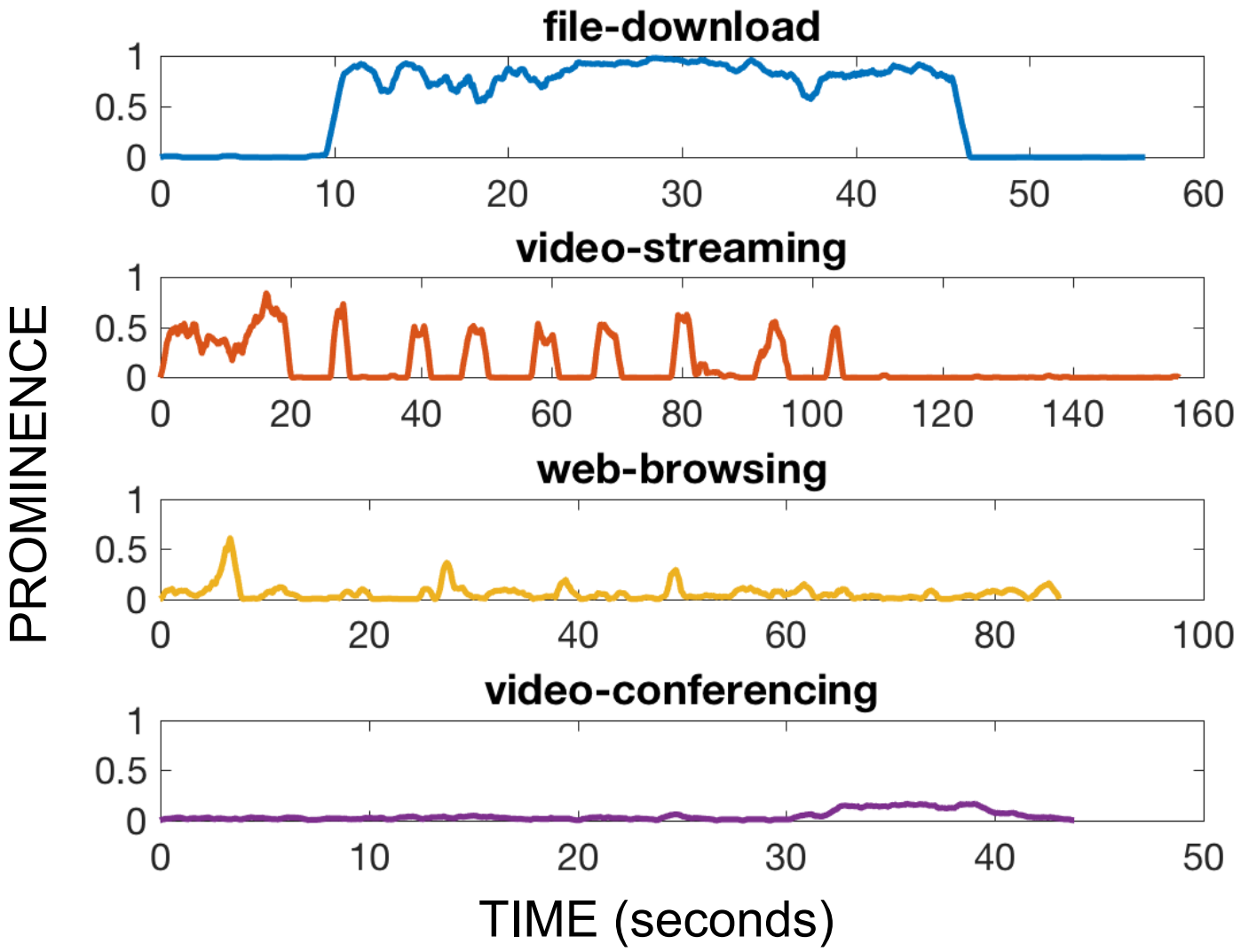
- Video streaming
  - Periodicity of segment download followed by idle period

PROMINENCE Signatures



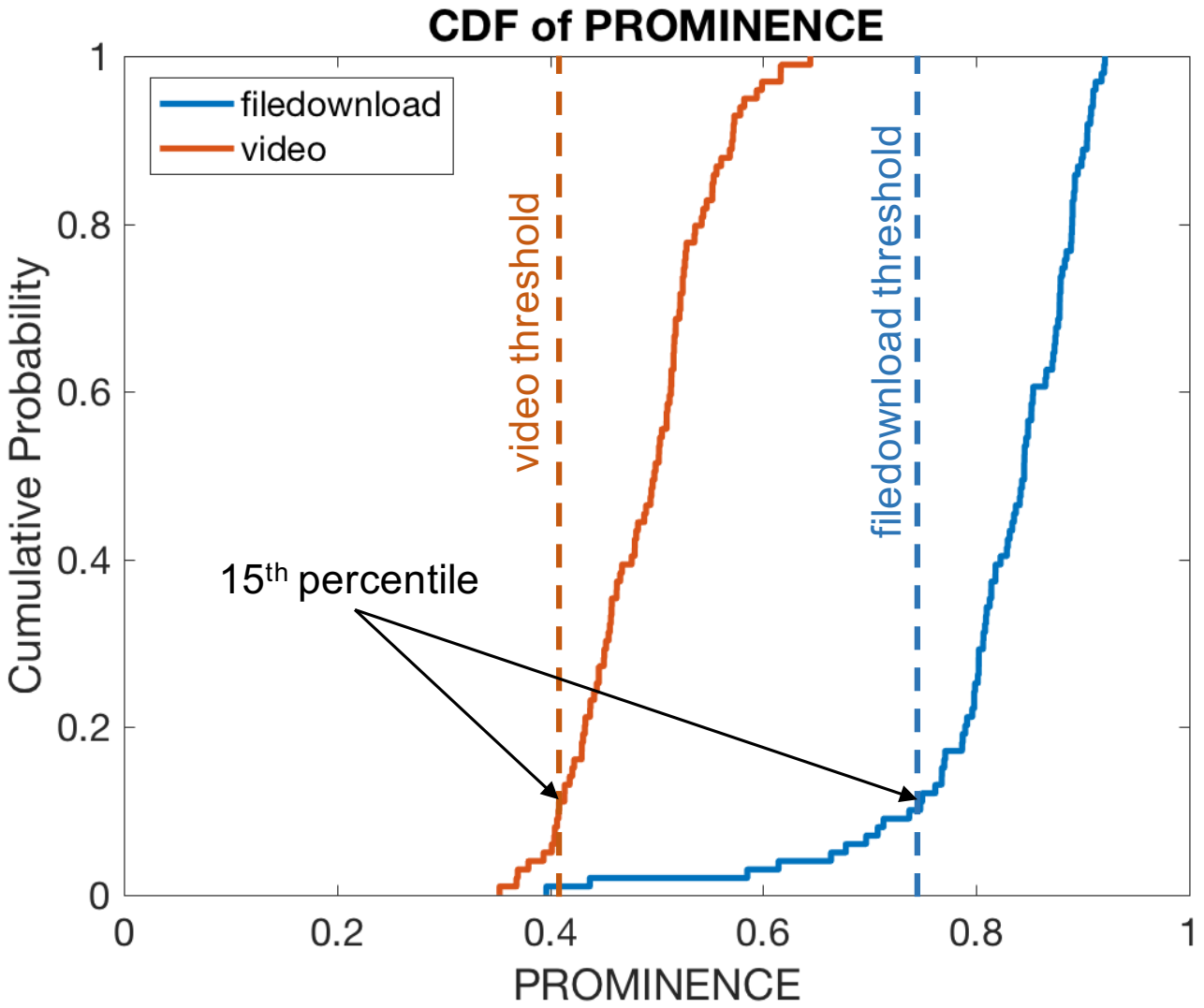
- Web browsing
  - Several brief file downloads

PROMINENCE Signatures



- Video conferencing
  - Low PROMINENCE
  - Data transmitted as needed

PROMINENCE as a Feature



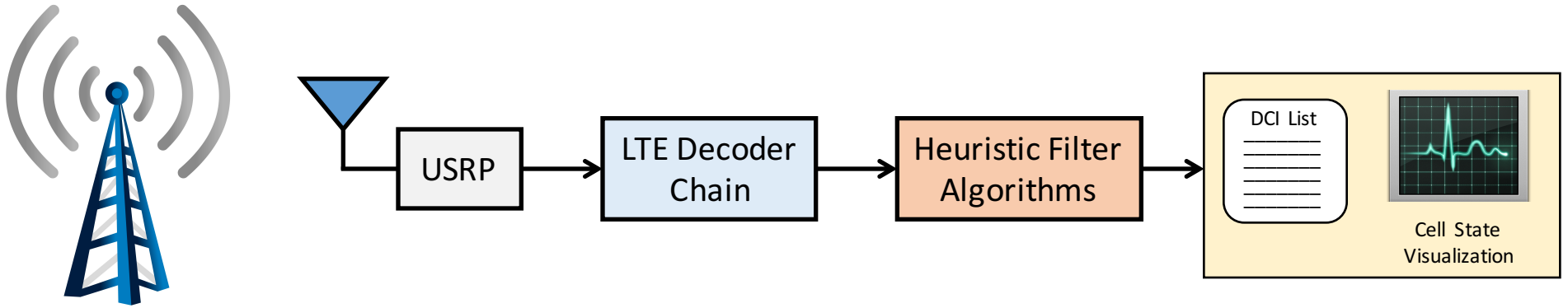
- CDF of PROMINENCE over 100 runs each of file download and video
- PROMINENCE is highly repeatable
  - Spread < 25%
  - Non-overlapping distributions

Our classifier is based on a simple thresholding of PROMINENCE.

Application Classifier

- We have shown so far…
  - A phone can determine its own application from radio-layer data decoded at its modem
- Next, we will show…
  - Anyone can infer all applications being served by a cell tower from the same type of data

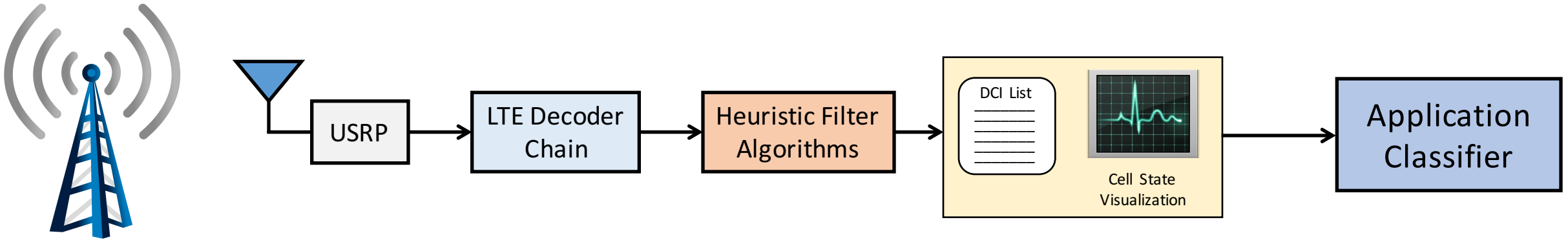
eNBsniffer: A Cell-Wide View of Resource Allocation



- *Passive* LTE PHY-layer sniffer
- Includes
  - Off-the-shelf USRP
  - MATLAB LTE decoder
  - Heuristic filter algorithms
- Validated over field tests: < 5% false negative error (for favorable RF conditions)

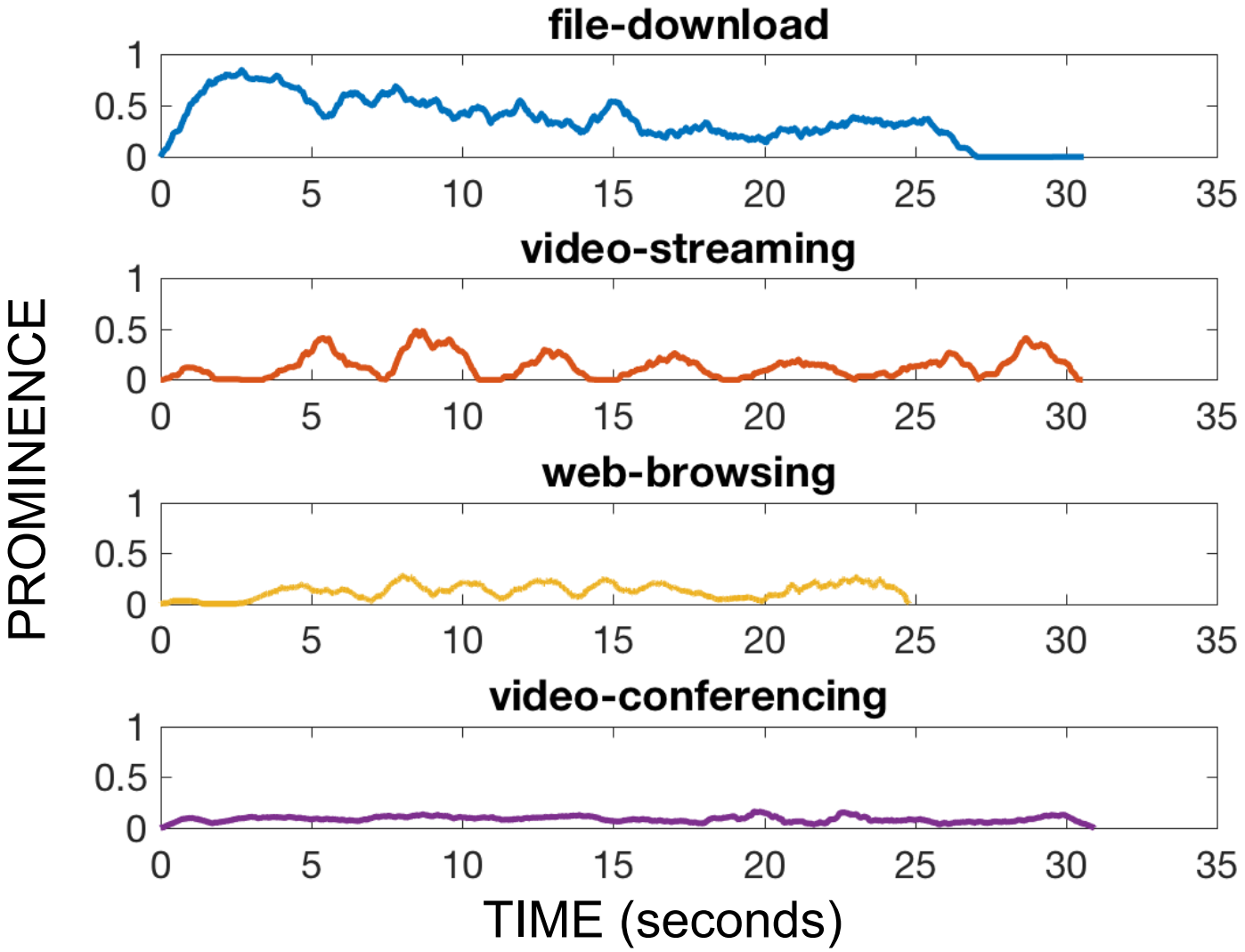
eNBsniffer decodes DCIs for *all* users served by cell tower.

eNBsniffer + Application Classifier



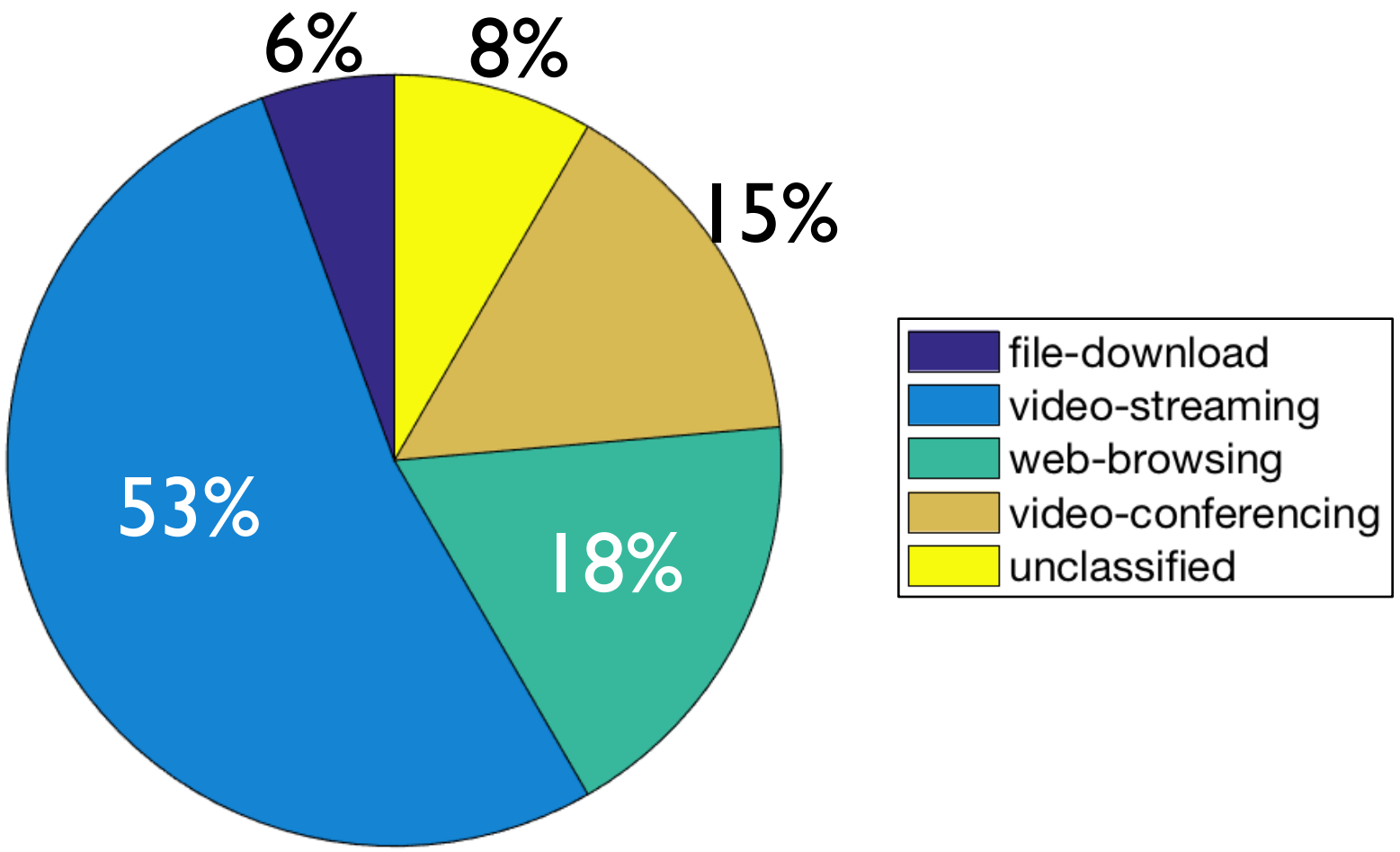
- Apply classifier on data from eNBsniffer
- Tag each user connected to the cell tower with the type of application being run on phone

Analysis of Congested eNB in Downtown Palo Alto



PROMINENCE signatures from eNBsniffer data

Analysis of Congested eNB in Downtown Palo Alto



Breakdown of applications served by eNB during lunch hours

Future Directions: Heuristic Refinement

- Expand classifier to broader class of apps
- Validate generality on different video clients
- Verify robustness to different schedulers
- Identify hidden patterns in PROMINENCE signatures (with ML)

Privacy Implications

- This work raises several privacy concerns
  - e.g. Hackers could isolate desired applications to attack
- Encourages an open discussion about security of LTE protocols
  - Mask features that can exploited to infer application type
  - “Pad” traffic to make signature ambiguous

Thank you!