CableMon: Improving the Reliability of Cable Broadband Networks via Proactive Network Maintenance

Jiyao Hu*, *Zhenyu Zhou**, Xiaowei Yang,
Jacob Malone, and Jonathan W Williams

*Jiyao Hu and Zhenyu Zhou, placed in the alphabetic order, are the lead student authors and contributed equally to this work.

Broadband Networks are Important

E-Commerce

Remote Learning

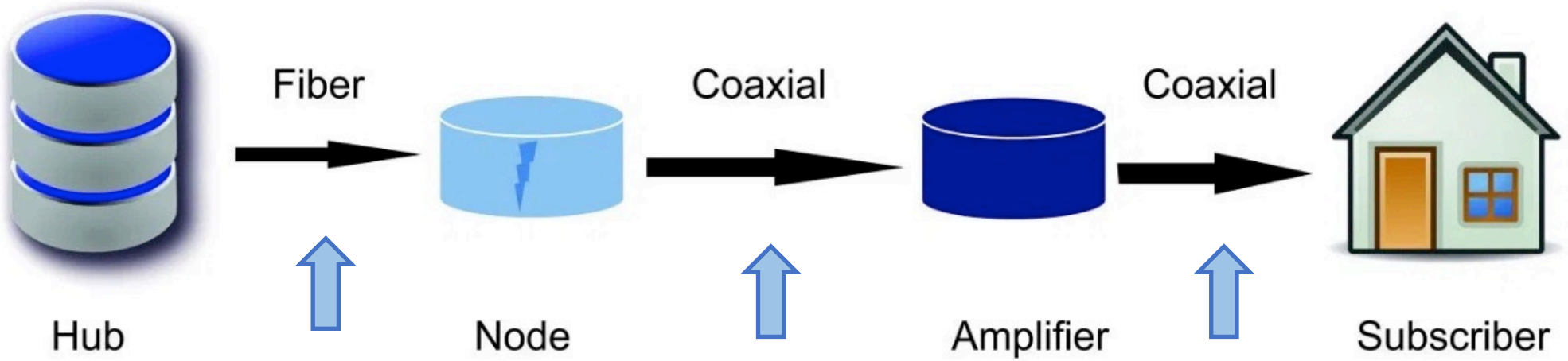
Work from Home

Entertainment

Cable Broadband: One of Few Choices in U.S.

- Availability to U.S. homes
  - VDSL: 43%
  - FTTP: 29%
  - Cable: 93%

Cable Networks are Unreliable



- Coaxial cables are vulnerable to radio frequency (RF) interference
- Aging can lead to problems
  - Cable shielding erosion
  - Loose connectors
  - Broken amplifiers
  - …
- Reliablity of broadband is at most 99% (<< FCC‘s requirement for PSTN, 99.99%)

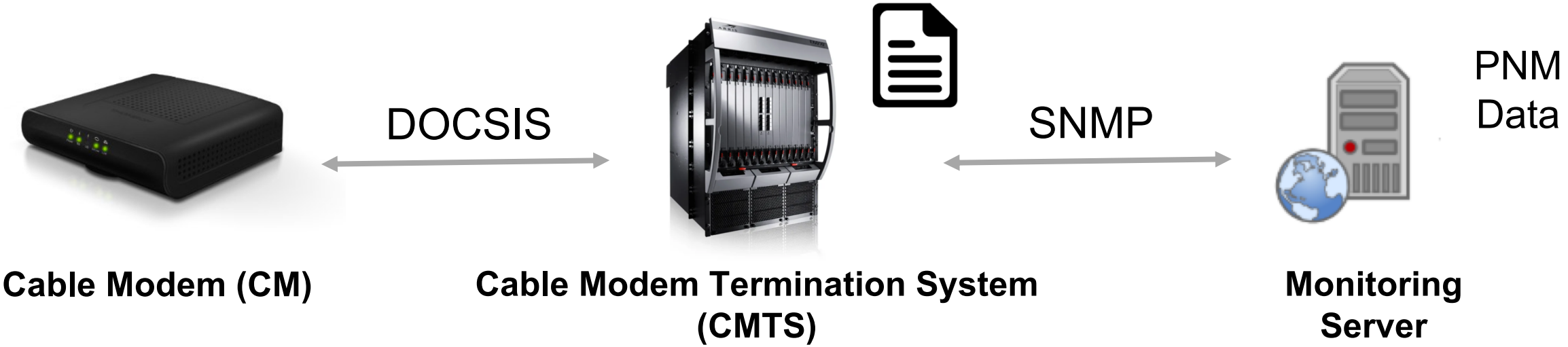
Goal

Improving the reliability of cable broadband networks

Roadmap

- Limitations of Existing Work
- CableMon Design
  - High-level Idea
  - Challenges
  - Solutions
  - ISP Deployment
- Evaluation
- Conclusion

Proactive Network Maintenance (PNM)



- Cable industry developed the PNM framework
- CMTS communicates with CM periodically to obtain PNM data
- A monitoring server collects all CMs' PNM data
  - E.g., Signal to Noise Ratio (SNR)

Existing Work

- Manually-set thresholds
  - A metric below/above a threshold: Repair the network
  - Hard to determine a proper threshold manually
  - Current recommendations: High false positives
    - In one of our studies, over 25% modems need repair following the PNM Best Practice document

Roadmap

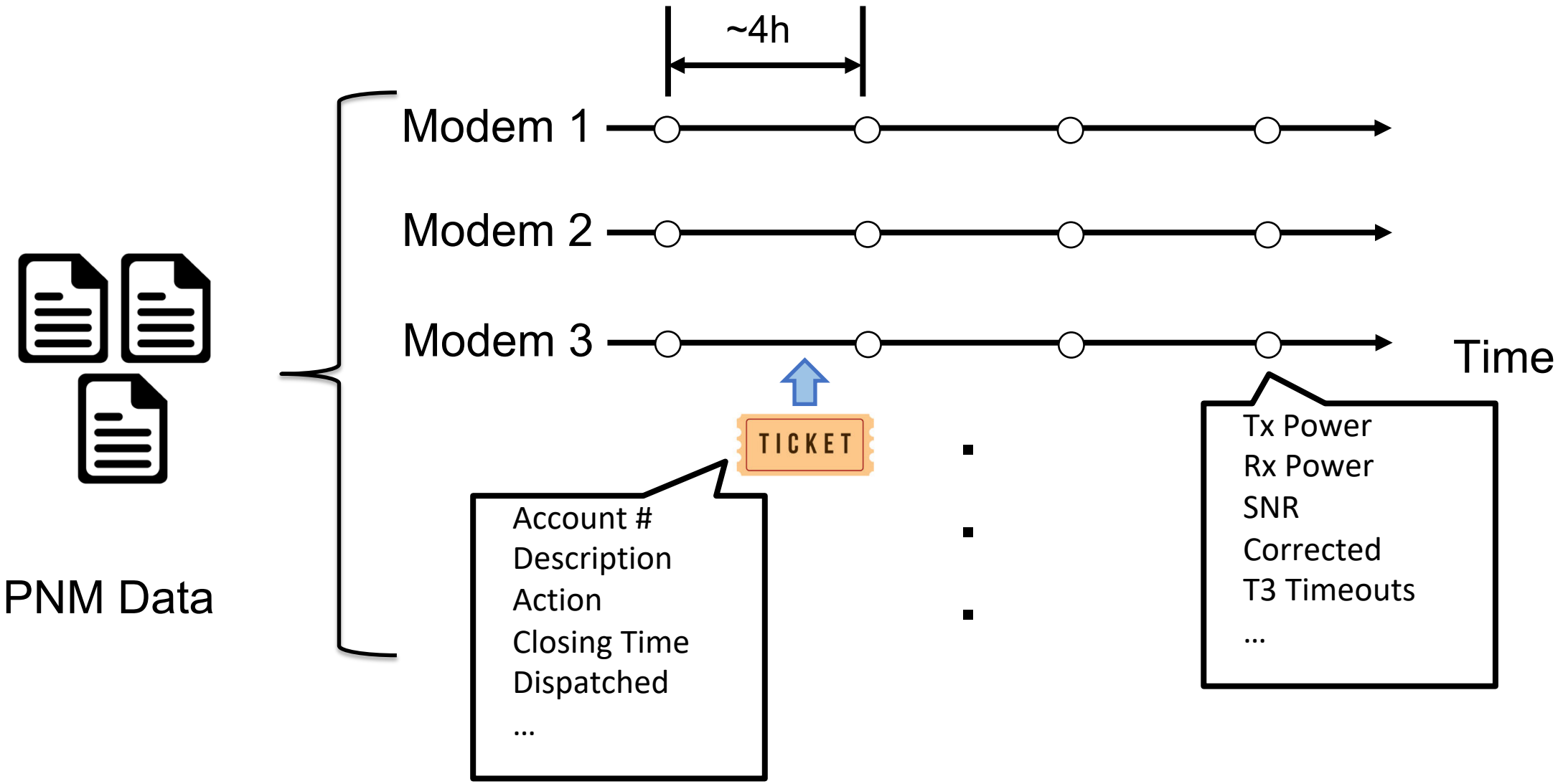
- Limitations of Existing Work
- CableMon Design
  - High-level Idea
  - Challenges
  - Solutions
  - ISP Deployment
- Evaluation
- Conclusion

High-level Idea

- Use customer trouble tickets as hints to learn what patterns of PNM data indicate network faults

- Design goals
  - Save operational costs of ISPs by forecasting troubles before customer calls
  - No manual labeling
  - No extensive parameter tuning
  - Efficient

Datasets: PNM Data and Customer Trouble Tickets

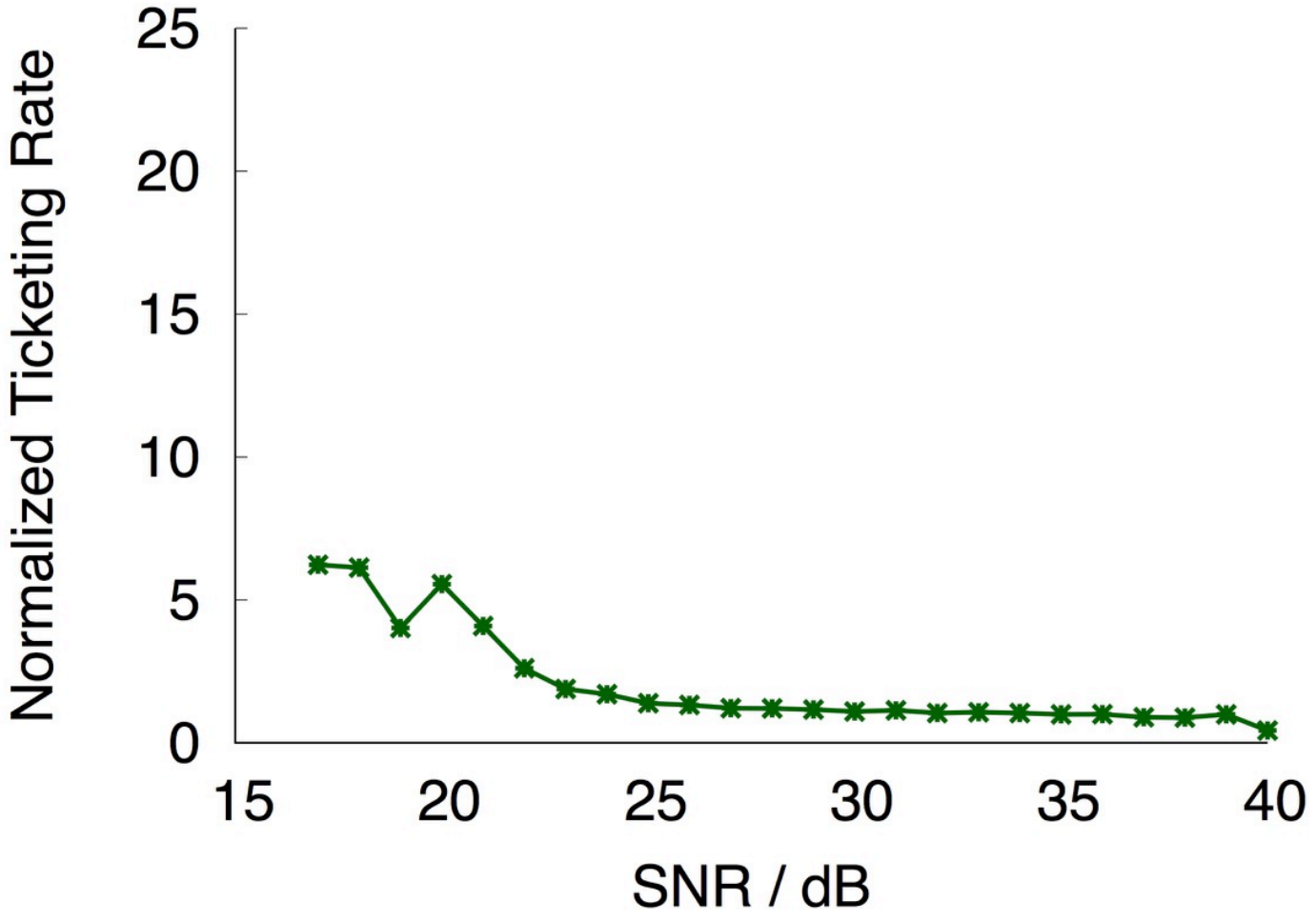


Tickets Correlate with Network Faults

Ticketing Rate:

$$\frac{\text{\# of tickets}}{\text{time}}$$

Tickets: Hints to detect network faults

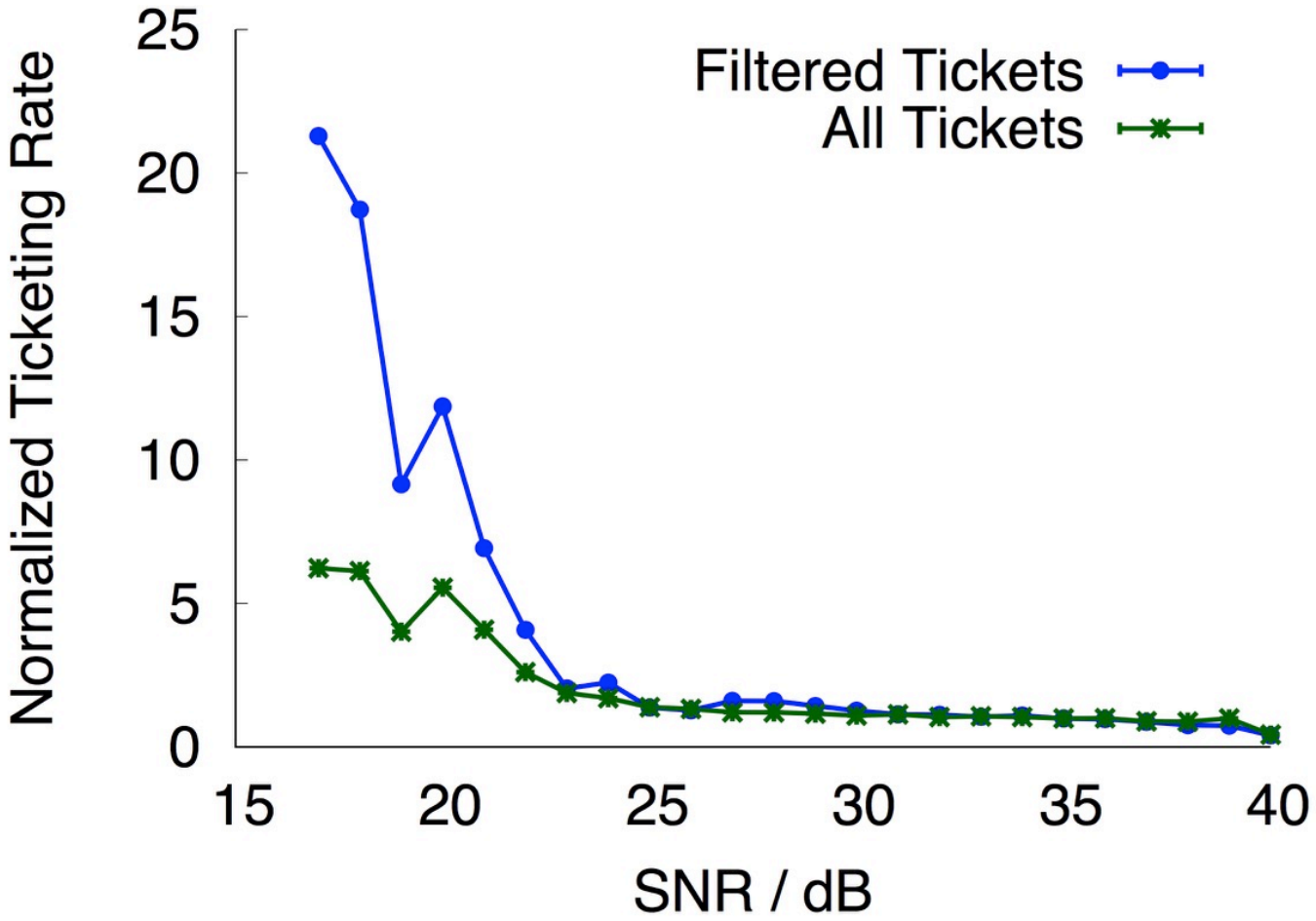


Challenges

- Tickets are noisy
  - Customers may call for network-irrelevant issues
  - Customers may not call when network faults occur

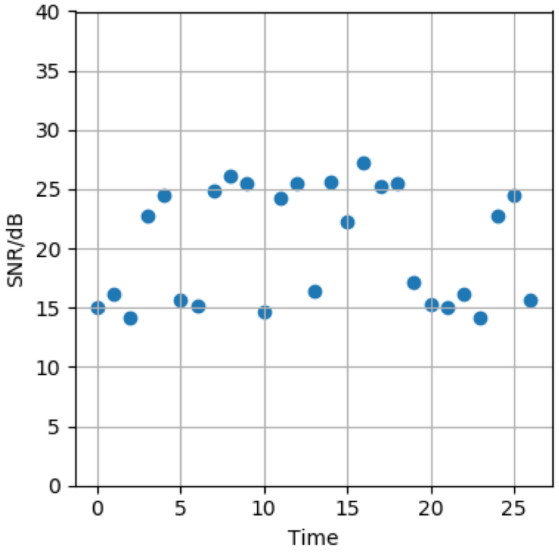
- PNM data
  - Instantaneous channel conditions, not sufficient for fault detection
  - Including environmental noise, not an accurate description of channel conditions

Filter Non-network Related Tickets

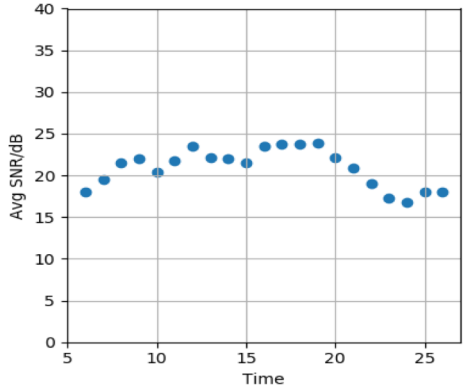


- Tickets are filtered according to the Description, Action, Dispatched, etc.
- Network tickets correlate better with PNM data values

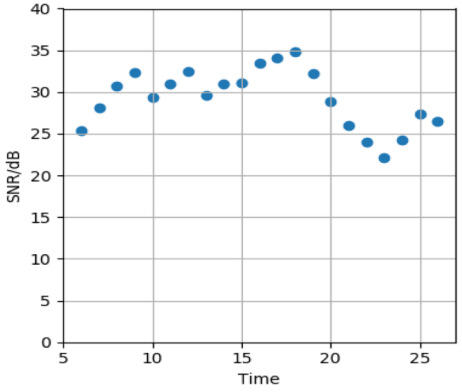
Feature Generation



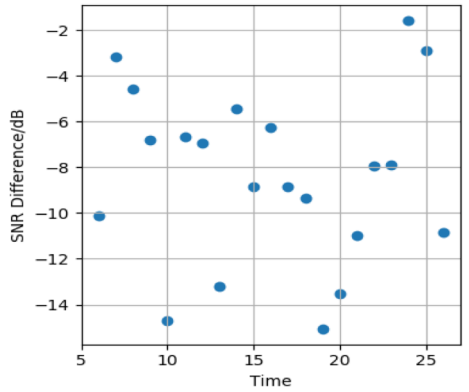
Statistical Models



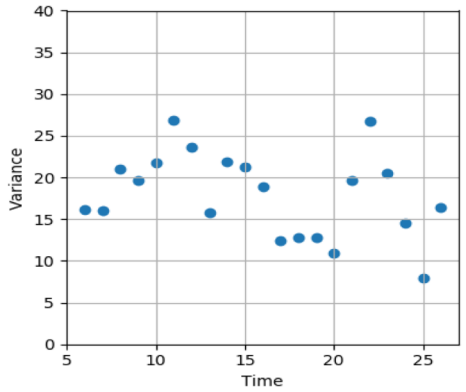
Moving Average



Weighted MA

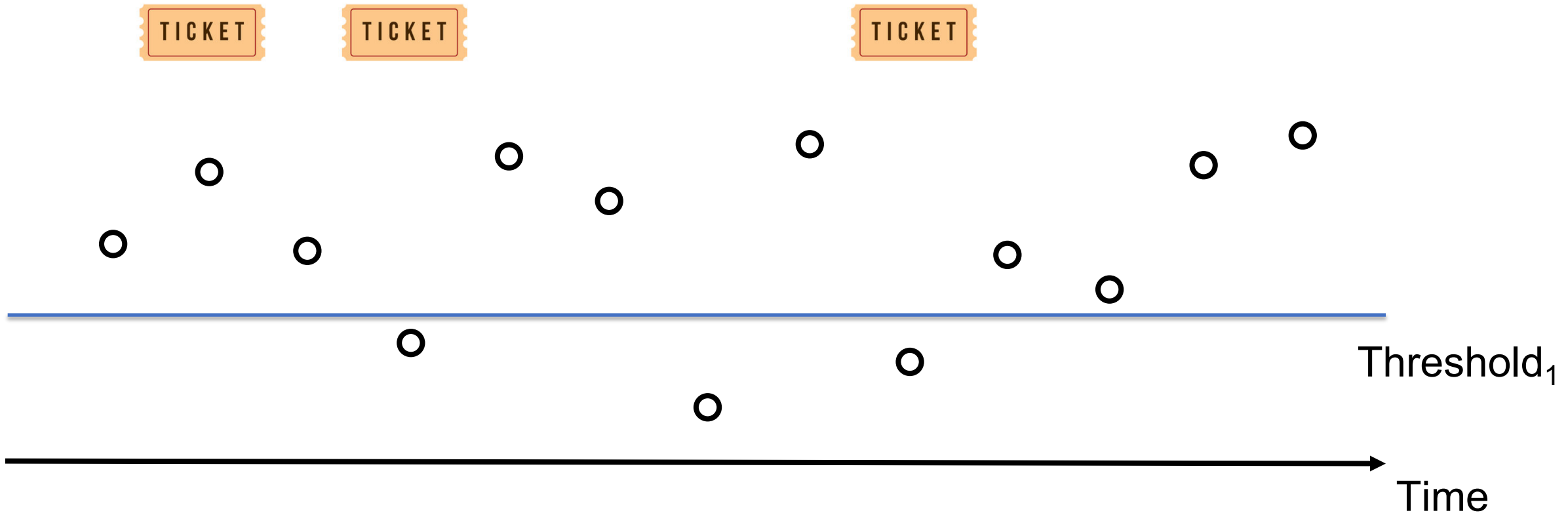


WMA Difference

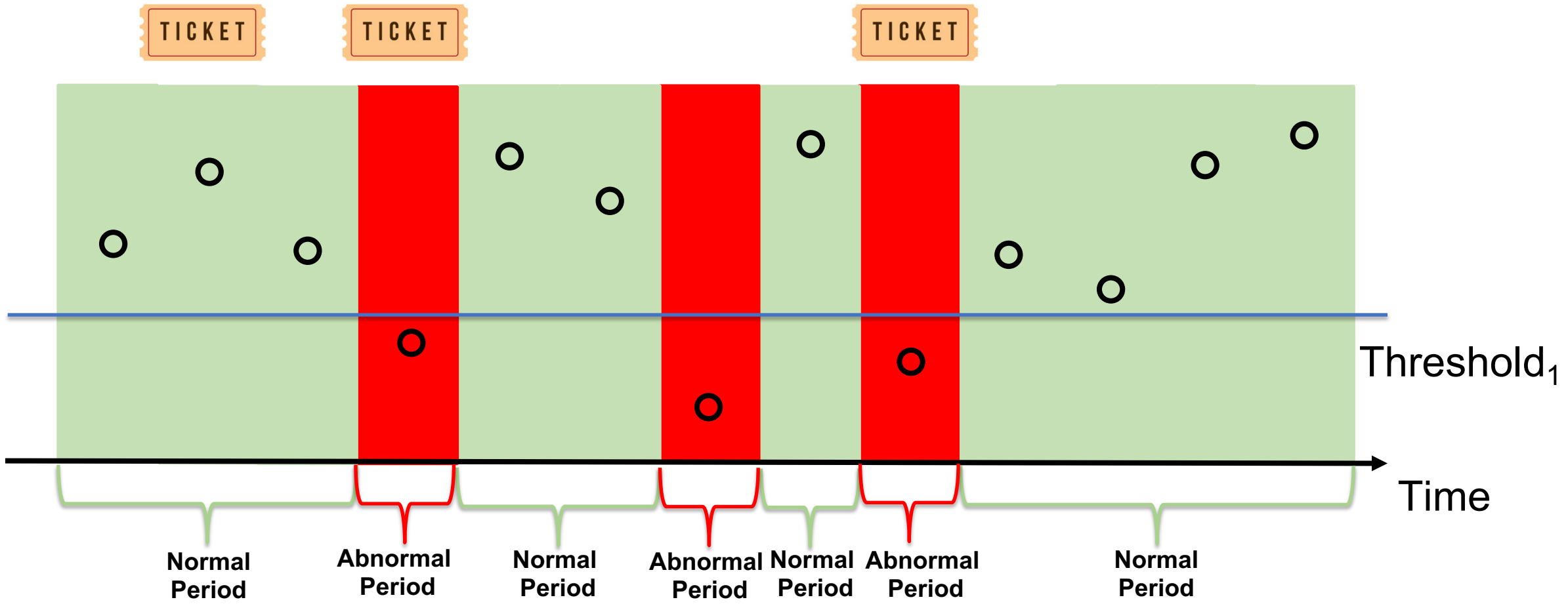


Variance

Determine the Threshold



Determine the Threshold

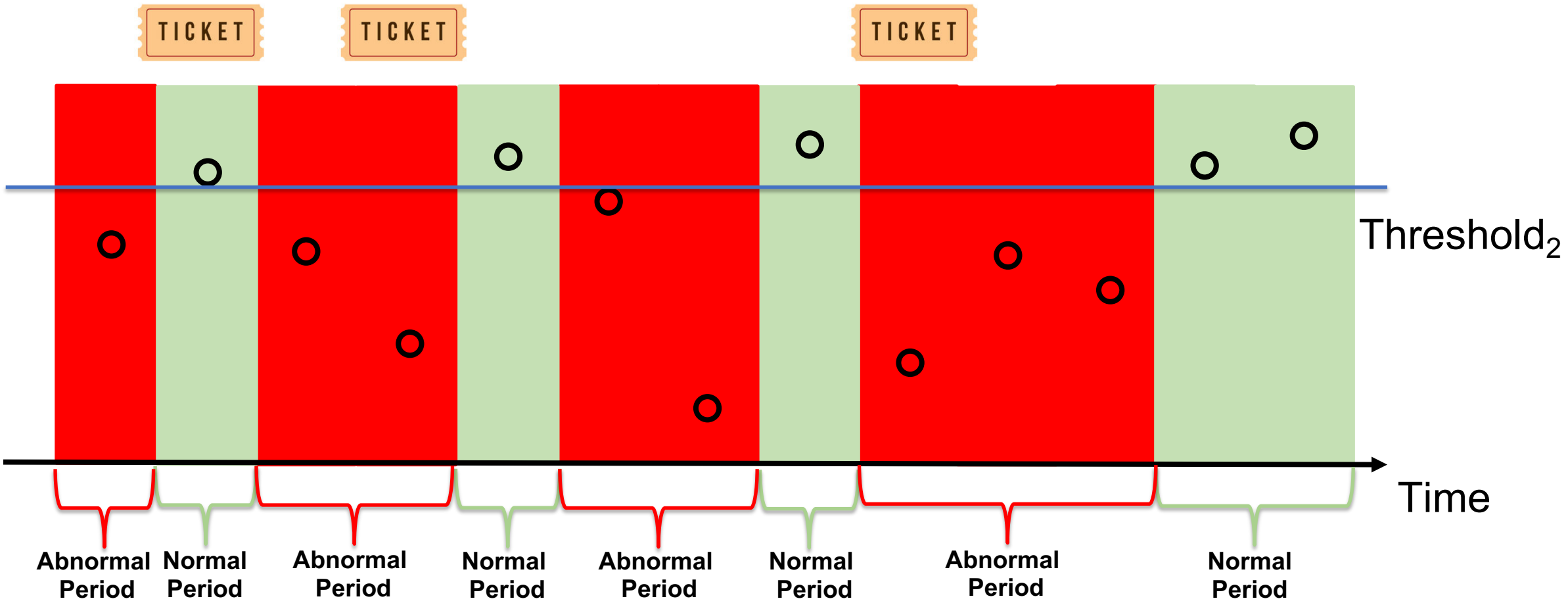


Normal:
- 1 ticket
- 10 data points (~40 hours)
- Ticketing Rate: 1/40 hours

Abnormal:
- 2 tickets
- 3 data points (~12 hours)
- Ticketing Rate: 2/12 hours

$$\text{Ticketing Rate Ratio} = \frac{\text{Abnormal Ticketing Rate}}{\text{Normal Ticketing Rate}} = 6.67$$

Determine the Threshold

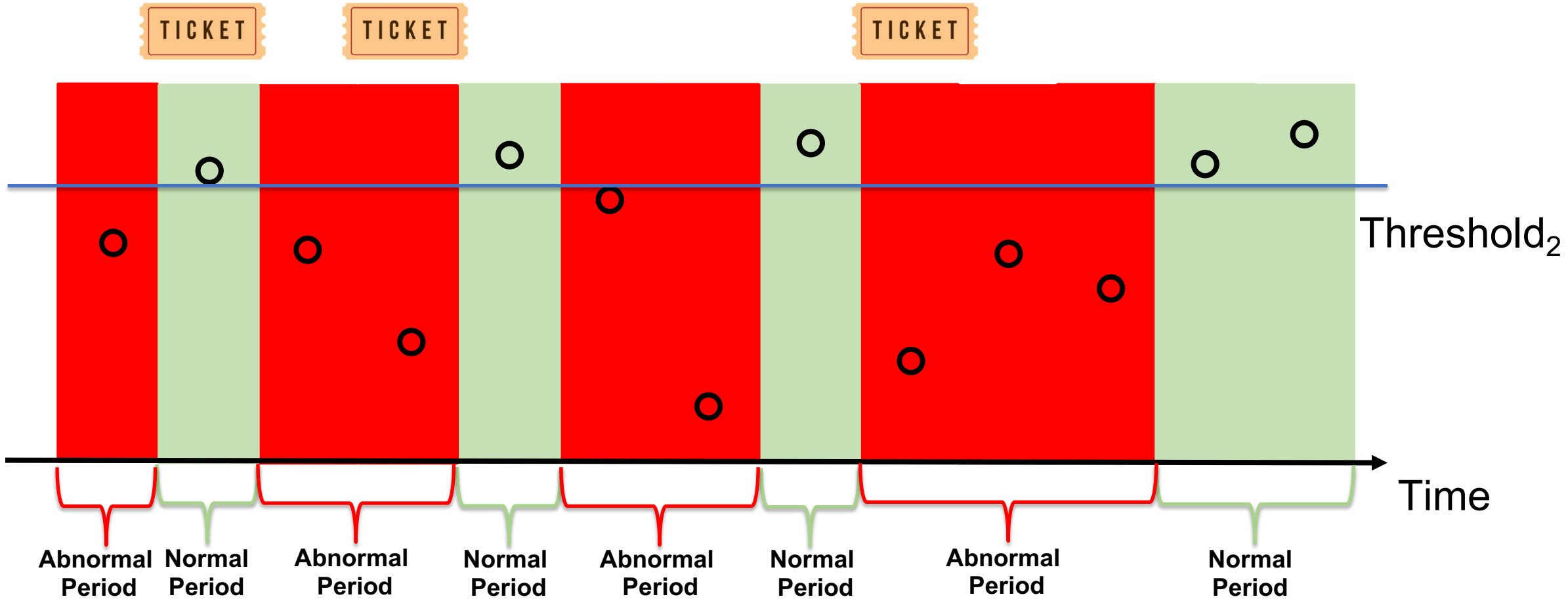


Normal:
- 1 ticket
- 5 data points (~20 hours)
- Ticketing Rate: 1/20 hours

Abnormal:
- 2 tickets
- 8 data points (~32 hours)
- Ticketing Rate: 2/32 hours

$$\text{Ticketing Rate Ratio} = \frac{\text{Abnormal ticketing rate}}{\text{Normal ticketing rate}} = 1.25$$

Determine the Threshold



Threshold =
Argmax(Ticketing Rate Ratio)

$$\text{Ticketing Rate Ratio} = \frac{\text{Abnormal ticketing rate}}{\text{Normal ticketing rate}}$$

CableMon Properties

- Ticketing Rate Ratio (TRR)

  1. is monotonously decreasing w.r.t. both false positives (FPs) and false negatives (FNs) under the assumption ticketing noise is uniformly distributed

  2. is maximized iff. both FPs and FNs are 0

Formal proof can be found in the paper

Selecting Top Features

- We select the features that achieve the highest ticketing rate ratio
- For the features generated from the same metric, we will select at most two

Top Five Features

Feature	Ticketing Rate Ratio
snr-var	14.49
uncorrected-var	7.66
rxpower-wma-diff	5.31
t3timeouts-wma-diff	4.93
t4timeouts-var	4.18

Combining Features

- If one feature is abnormal, the data point is labelled as abnormal
  - Different features may detect different types of network faults
  - Use the threshold(s) that maximize(s) the ticketing rate ratio for each feature

How ISPs Use CableMon

- Proactive detection
  - Make a dispatch decision only after a fault persists
  - Convert pointwise detection to abnormal event detection using a sliding window algorithm
    - See paper for more details

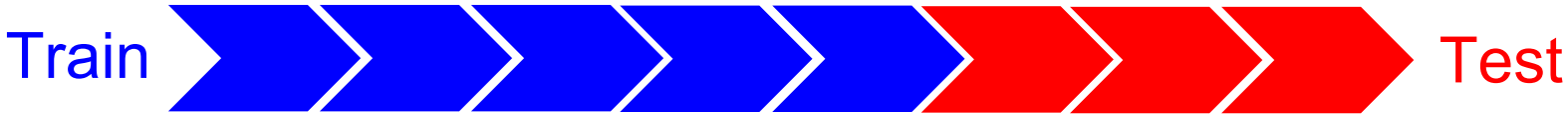
- Diagnosis
  - Determine whether the customer reported problem is network relevant

Roadmap

- Limitations of Existing Work
- CableMon Design
  - High-level Idea
  - Challenges
  - Solutions
  - ISP Deployment
- Evaluation
- Conclusion

Evaluation

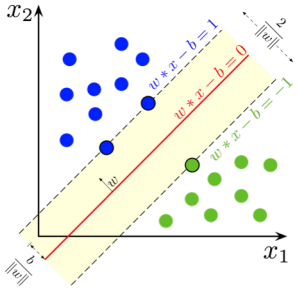
- Dataset



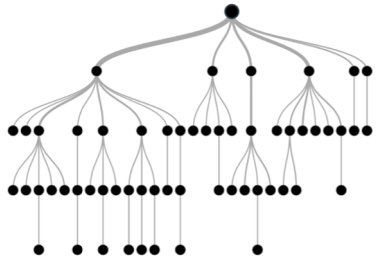
Jan June Aug, 2019

- Other Approaches

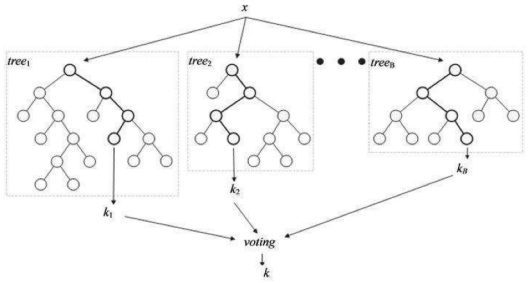
ISP Tools



SVM



Decision Tree



Random Forest

Evaluation

- Methodology
  - Ideal: Real world deployment (working in progress)
  - Experiments: Emulate how ISPs use CableMon
    - Run the sliding window algorithm
    - Count the number of tickets arrived during an abnormal event

- Metrics
  - Ticket prediction accuracy
  - Ticket coverage
  - Normalized ticketing rate

Results

Many PNM irrelevant tickets

	Ticket Prediction Accuracy	Ticket Coverage	Normalized Ticketing Rate
CableMon	81.92%	22.99%	3.55
SVM	75.64%	12.54%	2.02
Random Forest	73.14%	14.21%	2.24
Decision Tree	68.93%	15.53%	2.52
Comcast's Tool	23.48%	2.21%	1.18
AnonISP's Tool	10.04%	25.13%	0.98

Results

Many PNM irrelevant tickets

	Ticket Prediction Accuracy	Ticket Coverage	Normalized Ticketing Rate
CableMon	81.92%	22.99%	3.55
SVM	75.64%	12.54%	2.02
Random Forest	73.14%	14.21%	2.24
Decision Tree	68.93%	15.53%	2.52
Comcast's Tool	23.48%	2.21%	1.18
AnonISP's Tool	10.04%	25.13%	0.98

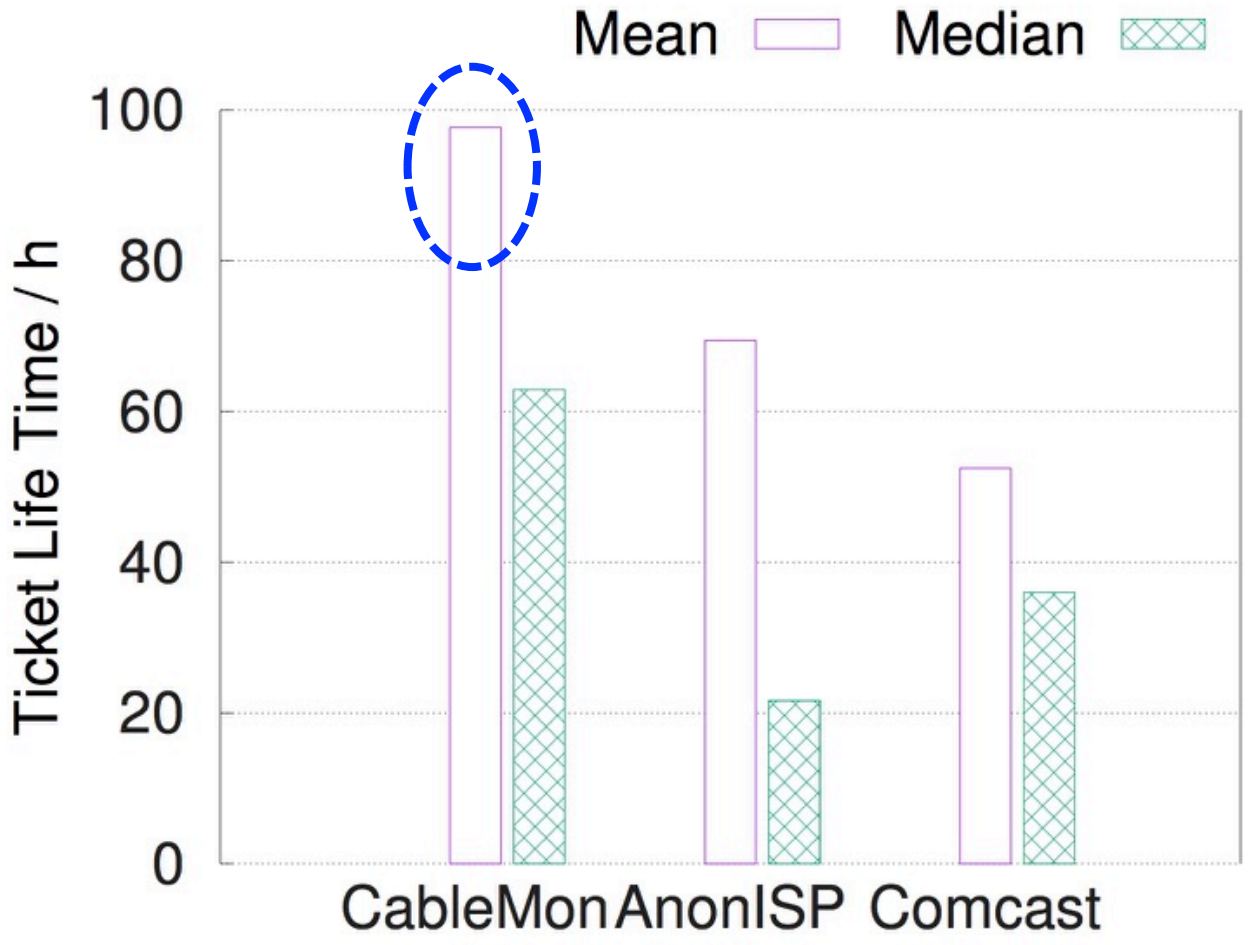
Results

Many PNM irrelevant tickets

	Ticket Prediction Accuracy	Ticket Coverage	Normalized Ticketing Rate
CableMon	81.92%	22.99%	3.55
SVM	75.64%	12.54%	2.02
Random Forest	73.14%	14.21%	2.24
Decision Tree	68.93%	15.53%	2.52
Comcast's Tool	23.48%	2.21%	1.18
AnonISP's Tool	10.04%	25.13%	0.98

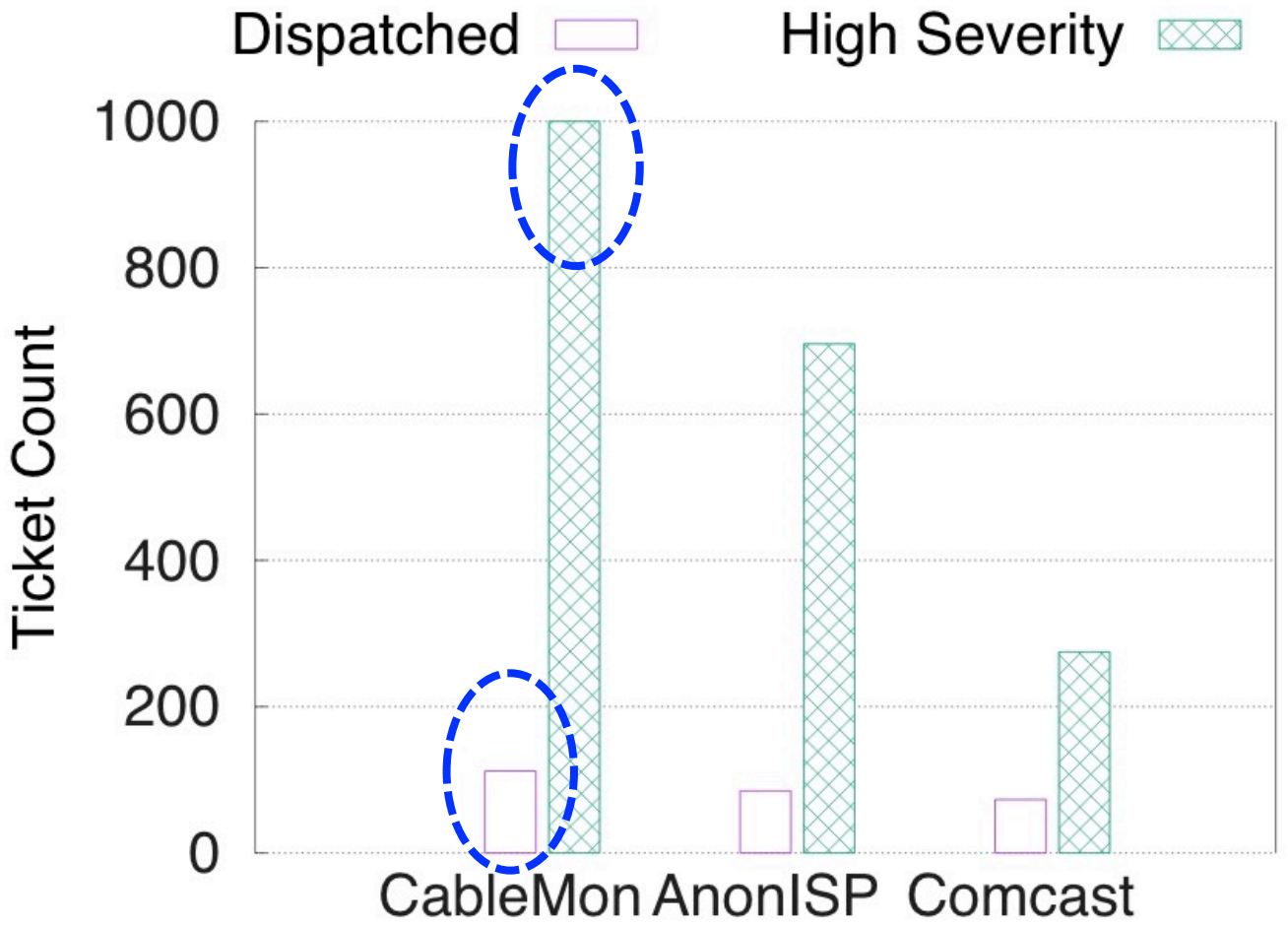
Ticket Life Time

- Tickets detected by CableMon
  - last for longer time
  - indicate harder issues



Type of Tickets

- Tickets detected by CableMon
  - are more severe
  - need more dispatches



More results can be found in the paper

Conclusion

- CableMon to detect network faults
- Use tickets as hints: No manual labeling
- Overcome the noise from both PNM data and customer trouble tickets
- Achieve high ticket prediction accuracy, and moderate ticket coverage

Thanks for your attention!

Questions?