Driving in the Fog: Latency Measurement, Modeling, and Optimization of LTE-based Fog Computing for Smart Vehicles

Yong Xiao[*], Marwan Krunz[†], Haris Volos[‡], and Takashi Bando[‡]

*School of Electronic Information and Communications, Huazhong University of Science & Technology, Wuhan, China

†Department of ECE, University of Arizona, Tucson, AZ

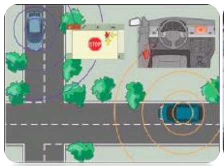
‡ Silicon Valley Innovation Center, DENSO International America, Inc.

IEEE SECON'19, Boston, MA

Smart Vehicular Services

Road safety and Efficiency

Basic Safety Services

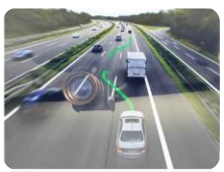

Collision avoidance


Traffic lights guidance


Road conditions

Autonomous Driving


Platooning


Cooperative driving


High-definition map

Telematics


Remote Driving


Remote Vehicle Health Monitoring

Navigation


Parking

Infotainment

YouTube

Video Streaming


Music


Mobile Office

CNN

News

Source: Huawei

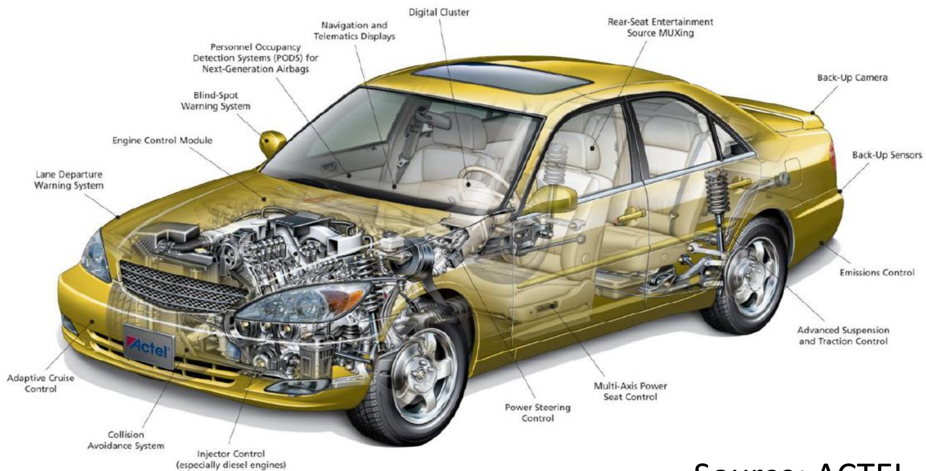
Features of Future Smart Cars

- ✓ Always connected
- ✓ Computing capabilities
- ✓ Environment awareness
- ✓ Storage space

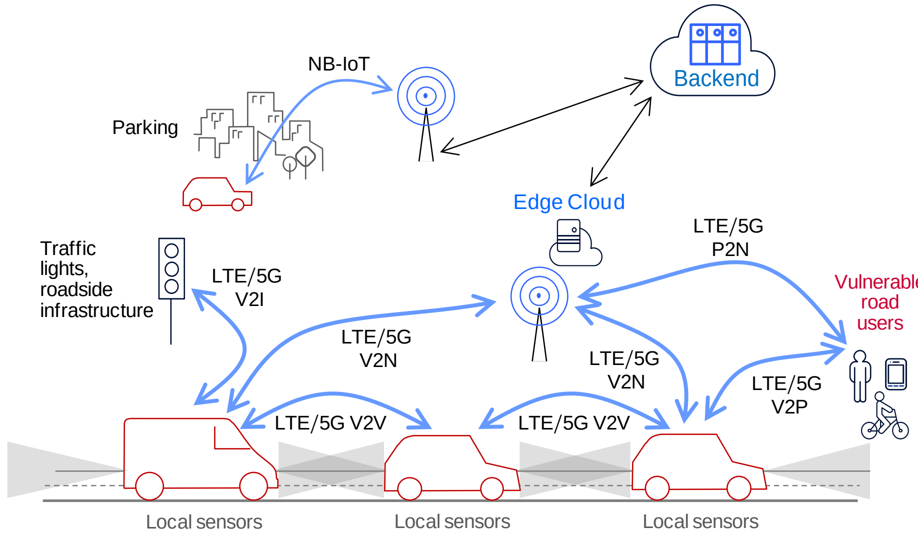
THE UNIVERSITY OF ARIZONA.
TUCSON ARIZONA

Connected Vehicles

In-Vehicle Processing vs. Connected Vehicle



Source: ACTEL



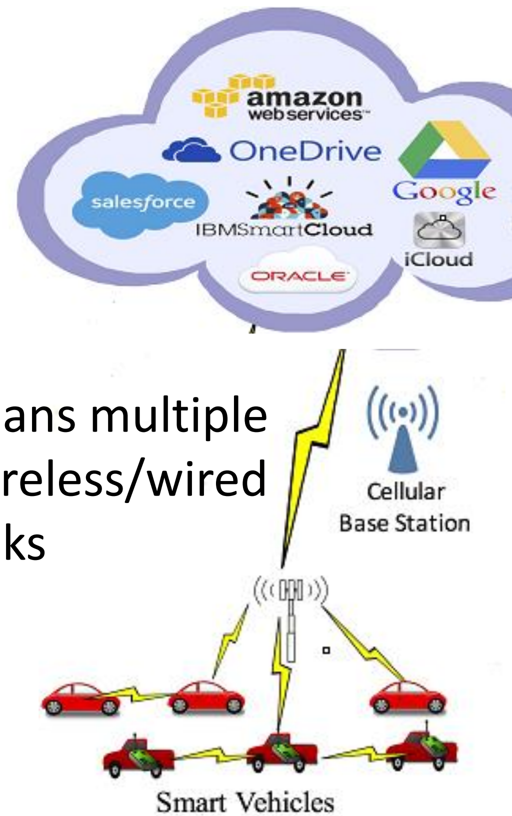
Source: 5GAA

- Limited computing/storage capability
- Blind spot
- Limited/no traffic updates
- Quick decision

- Relaxed computing/storage limits
- No blind spot
- Instantaneous traffic updates
- Requires ultra-low latency comm.

Cloud-based Connected Smart Vehicles

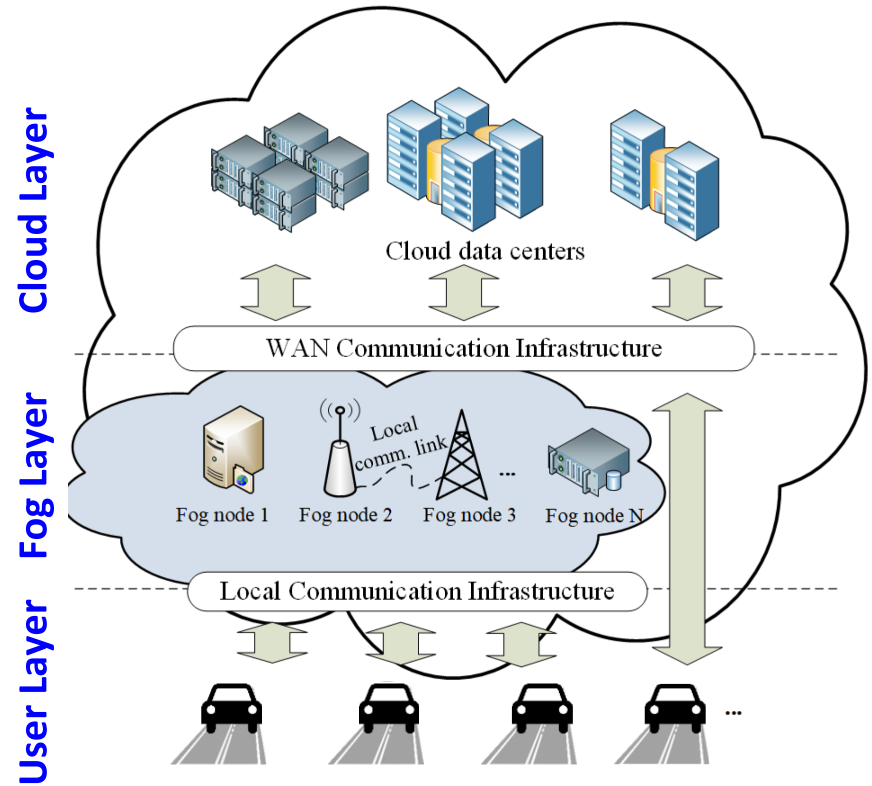
Traditional Cloud Computing



Spans multiple wireless/wired links

- 🙂 High computational performance
- 🙁 Unpredictable latency and connection reliability

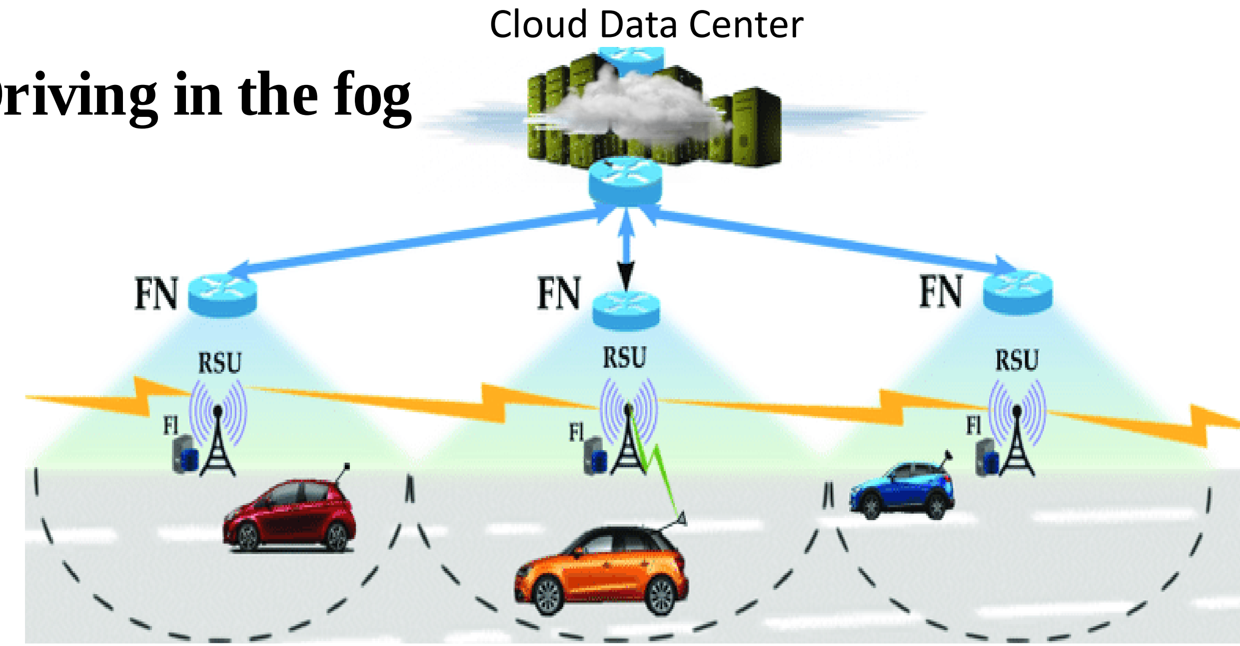
Hierarchical Fog/Cloud Computing



- 🙂 Cloud data center has been supplemented by fog nodes
- 🙂 Low latency
- 🙂 High reliability

Latency Challenge

Driving in the fog



Research Problem:
Develop approaches to reduce last-mile latency in existing LTE networks

- 3GPP recommends ~ 10 msec RTT for UEs across LTE networks (optimal conditions)
- Recent reports and our measurements suggest that this latency is far too challenging to achieve in existing LTE networks

Key Contributions

- ✓ **AdaptiveFog**: Vehicle-to-fog framework for multi-MNO LTE networks
- ✓ Novel distance metric (**weighed K-R distance**) to quantify latency performance of different MNO networks
- ✓ Measurement-driven modeling of V-to-fog and V-to-cloud latencies
- ✓ Optimal policy for dynamic selection of LTE provider & fog/cloud server

Outline

AdaptiveFog Framework

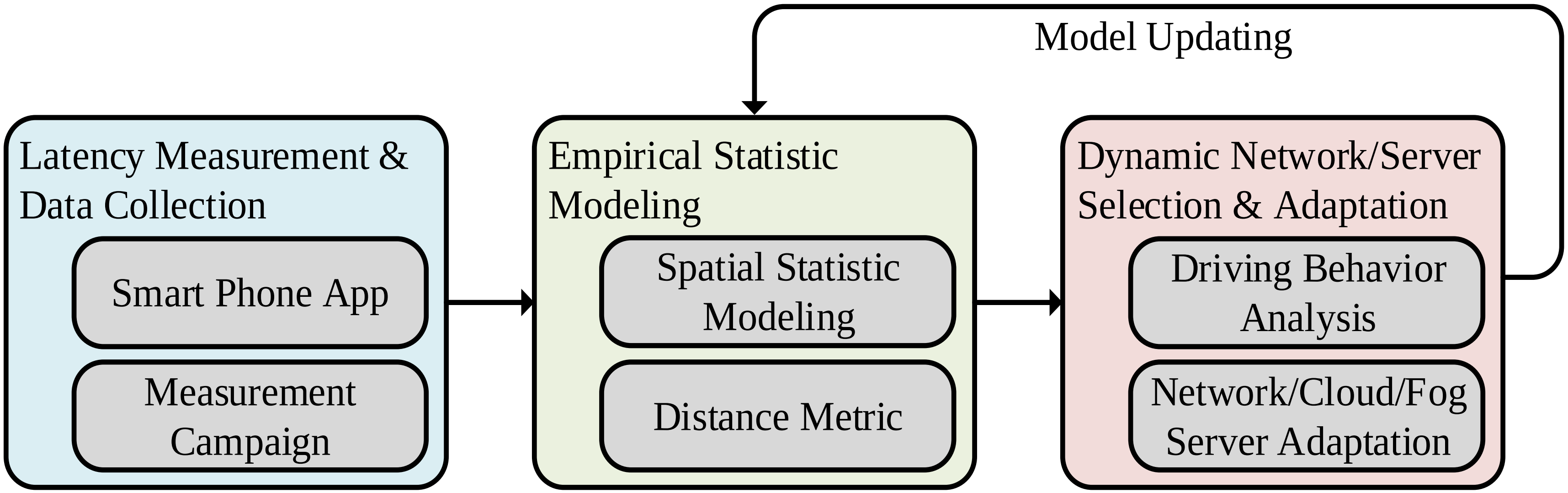
Latency Measurements & Modeling

Dynamic Network/Server Selection & Adaptation

Conclusions

AdaptiveFog Framework

AdaptiveFog is a novel framework for the UE to dynamically switch between available MNO networks and cloud/fog servers on the move

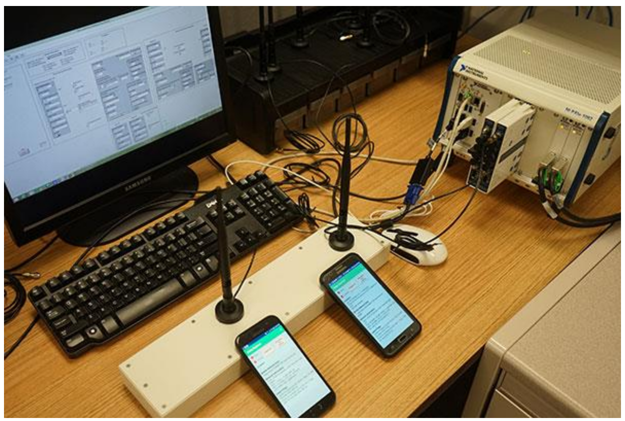


Latency Measurements

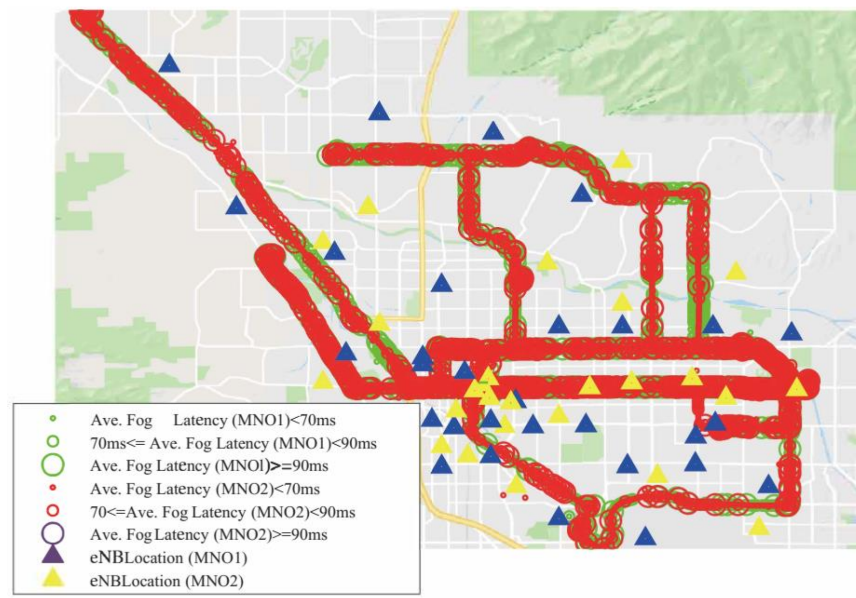
Extensive measurement campaign in two cities (San Francisco & Tucson)

Tens of traces of fog & cloud latencies collected over several months

Fixed-location as well as “in-vehicle” measurements using a custom app



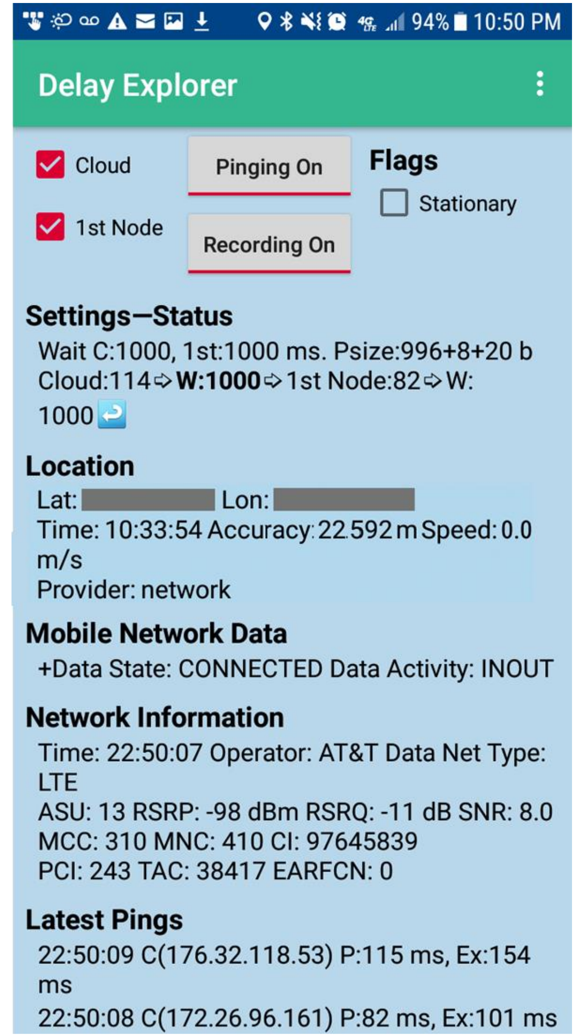
Example routes (Tucson, AZ)



Smartphone App

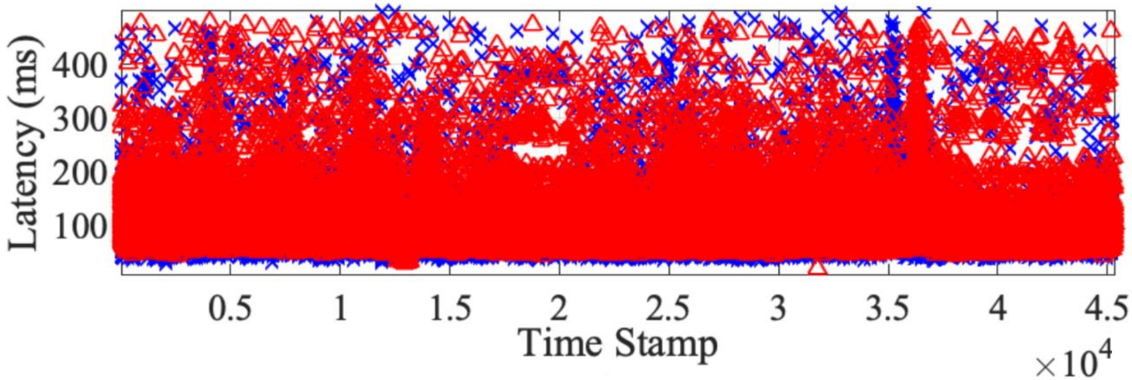
Delay Explorer

- Periodically Ping IP address of
  - 1st node in LTE network
  - Amazon cloud server (West coast)
- Record RTT of two MNOs networks (Sprint and AT&T)
- Record other info (location, time stamp, GPS coordinate, etc.)



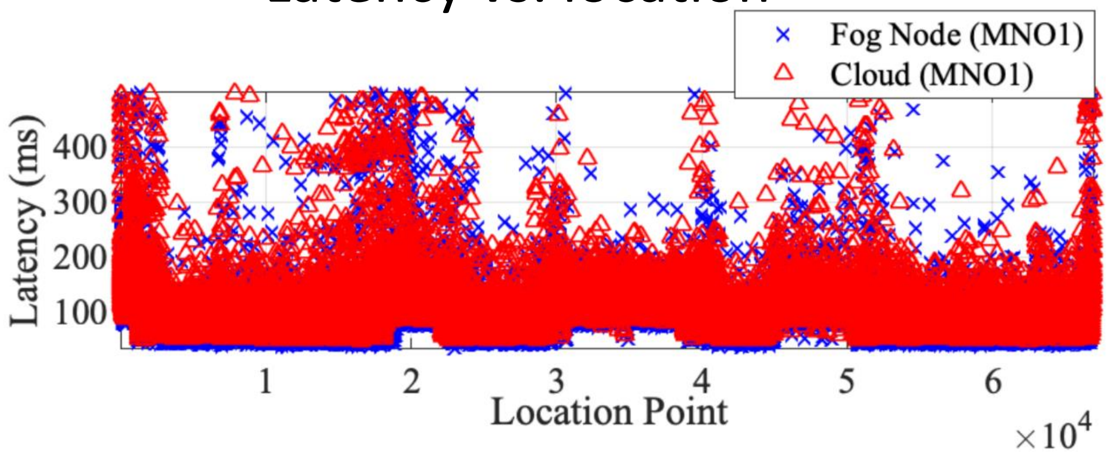
Preliminary Observations

Latency vs. time stamp



(second)

Latency vs. location



(Location Point)

- No correlation between RTT and time stamp
- Noticeably different latency patterns at different locations

Latency Statistics

Cloud vs. Fog

Traces			L1 Fixed	L2 Fixed	All Fixed	R1(Drive) (6.1m/s)	R2(Drive) (15.7m/s)	All Drive
MNO1	Fog Latency (ms)	Mean	62	72	70	83	96	88
		STD	18	16	18	28	29	34
		Median	55	71	68	77	91	85
		Conf. 90%	85	86	85	115	121	120
	Cloud Latency (ms)	Mean	74	87	85	94	108	96
		STD	15	15	21	26	29	33
		Median	71	88	86	92	108	94
		Conf. 90%	88	100	104	124	129	128
MNO2	Fog Latency (ms)	Mean	72	64	72	85	80	83
		STD	14	17	15	52	46	51
		Median	71	93	71	69	67	66
		Conf. 90%	84	87	86	132	112	131
	Cloud Latency (ms)	Mean	87	74	88	119	125	124
		STD	13	13	17	50	47	54
		Median	88	71	90	108	117	109
		Conf. 90%	99	87	102	166	133	100

Fixed loc vs. driving

Key Observations:

- MNOs vary significantly in some locations.
- When averaging over all traces, two MNOs exhibit similar behavior
- Difference between cloud and fog is around 10 ms in average

Distance Metric

Weighted Confidence

- Confidence level of service type i

$$F_i = \Pr(x \le r_i)$$

Max tolerable latency for service *i*

- *Proportionally weighted confidence level*

$$\hat{F} = \sum_{i \in \mathcal{M}} w_i F_i$$

Weight of service *i* (e.g., probability of service i arrival)

Set of all supported services

Weighted Kantorovich-Rubinstein (K-R) Distance

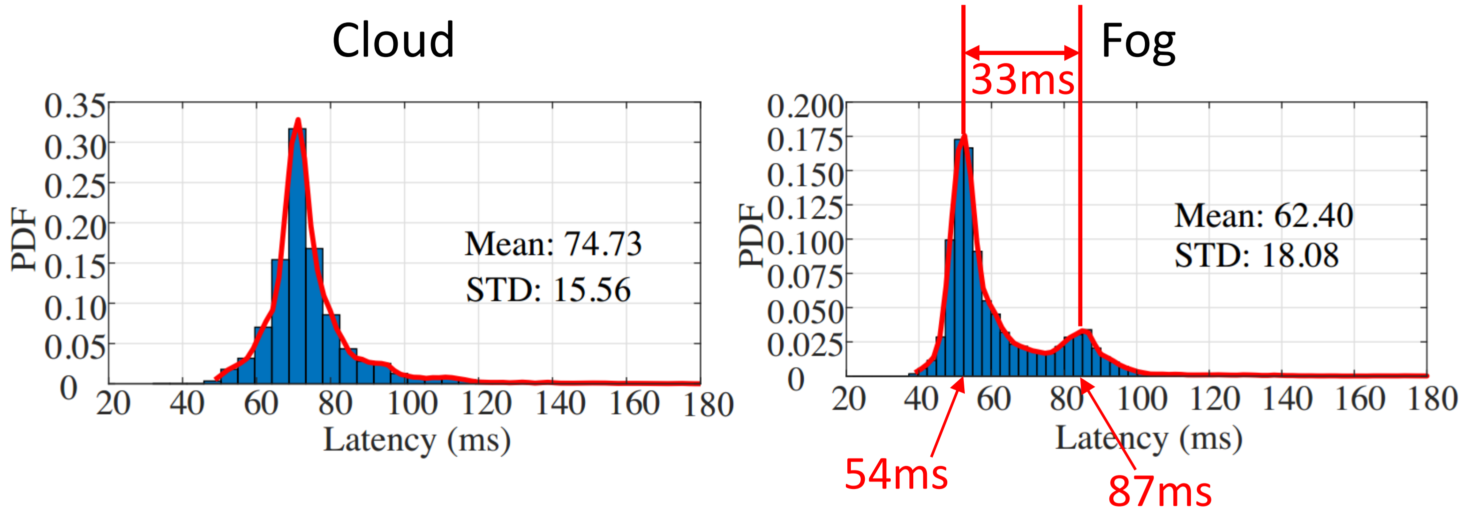
- Performance difference between two MNOs/servers (e.g., cloud and fog)

$$K(F,G) = \sum_{i \in \mathcal{M}} w_i \left[F_i - G_i\right]$$

Performance of the same service *i* offered by two MNOs/servers

Empirical Modeling of Latency

Fixed Location



PDF of fog latency can be fitted by a bimodal Gamma distribution

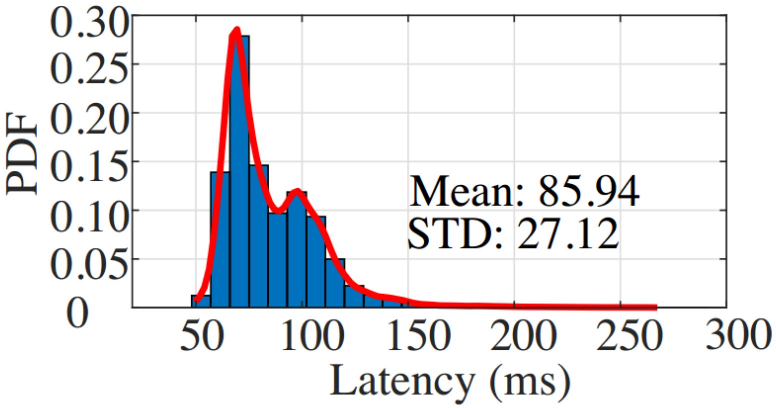
Difference (~ 33ms) between two peaks is caused by

- SR retransmission periodicity (~ 20 to 40 msec)
- HARQ retransmission delay (~ 1 to 8 msec)

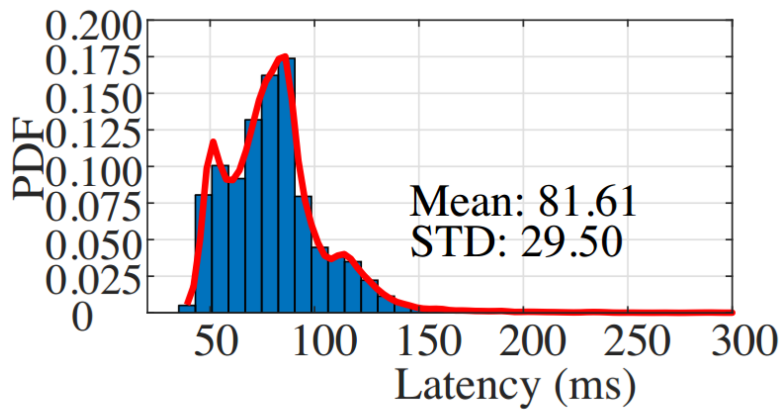
Empirical Modeling of Latency

While Driving

Cloud



Fog

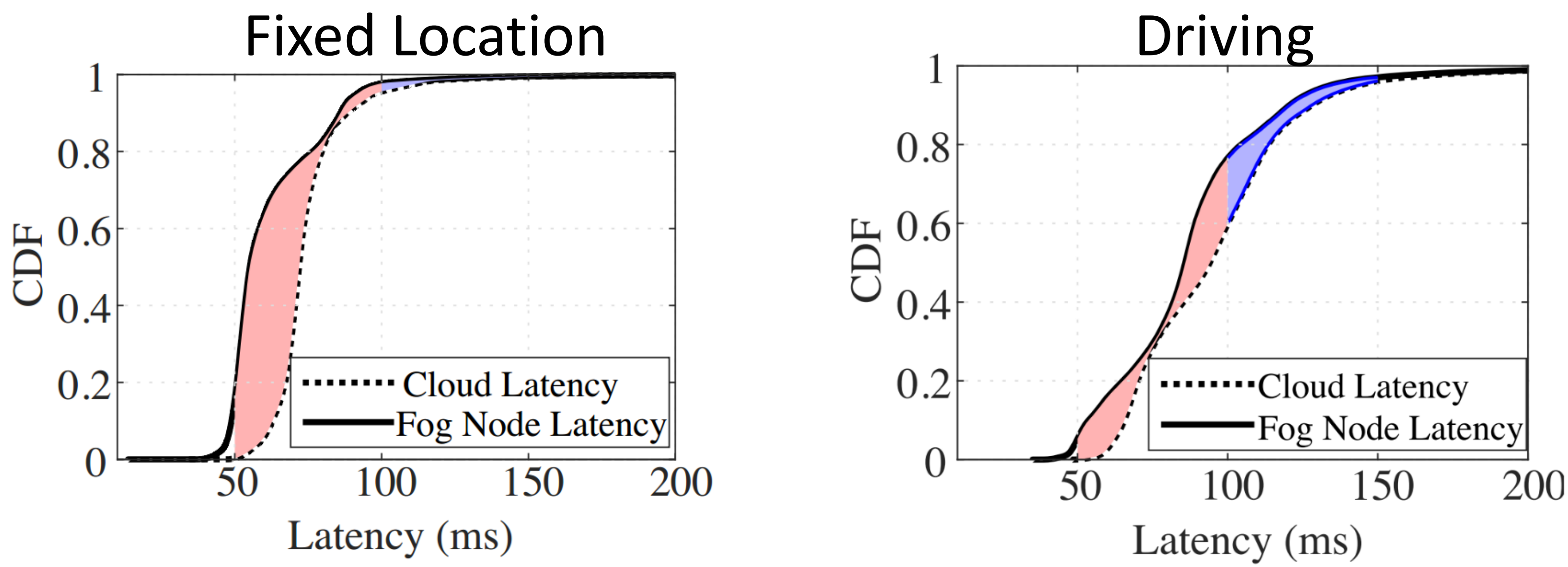


Compared to fixed location latency:

- Mobility contributes to around 10-20ms latency increase
- Variance increases significantly

Empirical Modeling of Latency

K-R Distance between Fog and Cloud



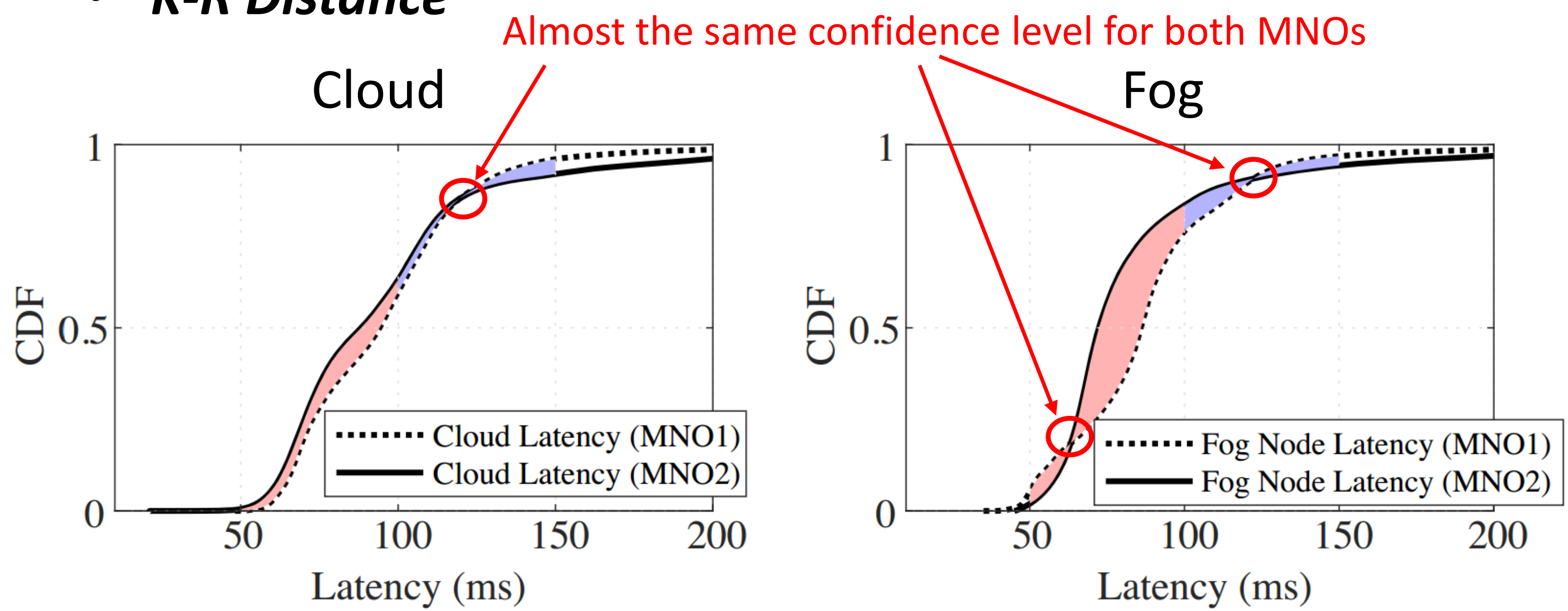
- Fixed-location
  - ➢ min K-R distance is at 85 ms (=0.23%)
  - ➢ max K-R distance is at 63 ms (=58.6%)
- Driving:
  - ➢ min K-R distance is at 74ms (=0.55%)
  - ➢ max K-R distance is at 57ms (=18%)

Negligible for most applications

Compared to fixed loc, K-R dis in driving is much smaller

Empirical Modeling: Different MNOs

- ***K-R Distance***



- Cloud:
  - max K-R distance at 88ms (=25.79%)
  - MNO 2>MNO 1 (<131ms); MNO 2<MNO 1 (>131ms);
- Fog:
  - MNO1>MNO2 (<64ms and >125ms); MNO1<MNO2 (btw. 64ms and 125ms)

Optimal Network/Server Selection & Adaptation

State
Driving speed, location and LTE network of UE

Action
UE decides whether or not to switch to another LTE network

State Transition Function
Probability of state transitioning

Formulate network adaptation and fog/cloud server selection as a Markov decision process (MDP)

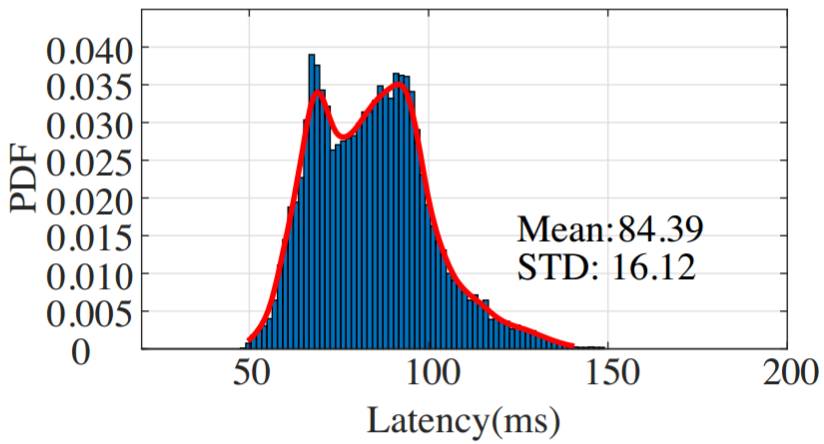
Utility Function
Maximize confidence level for UE

Selection & Adaptation

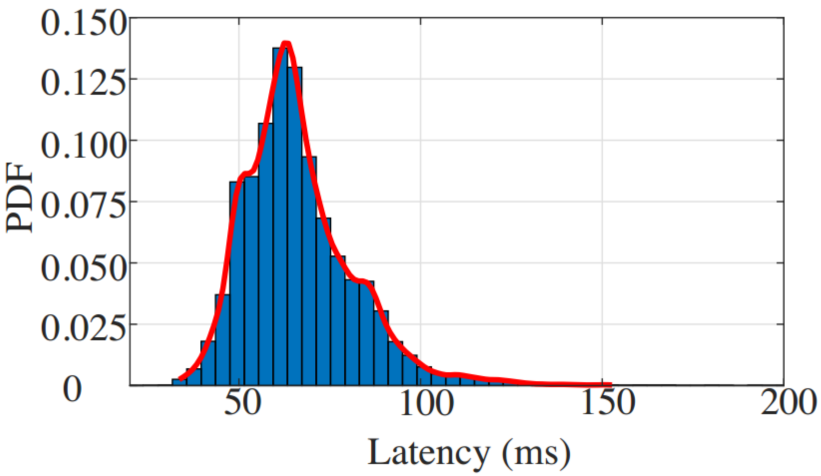
Optimal Network/Server Selection & Adaptation

- ***Empirical PDFs***

Cloud



Fog



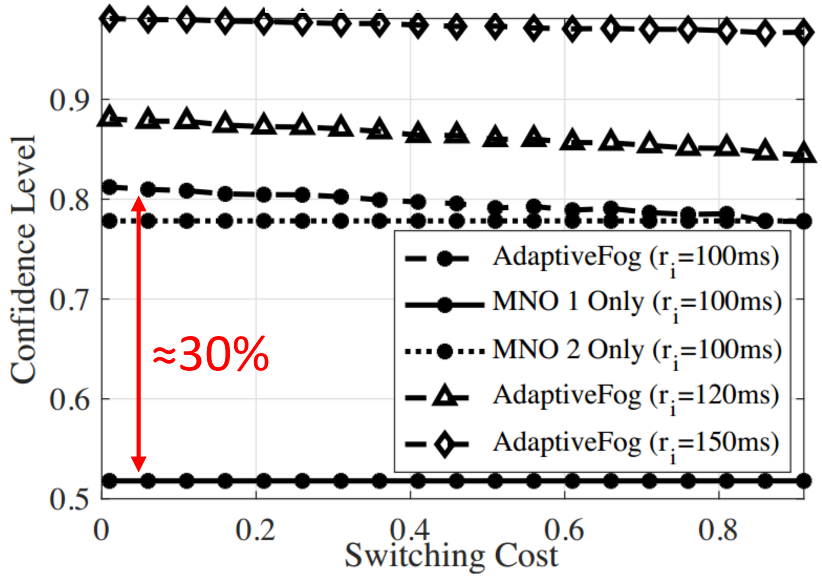
Compared to the single MNO case, AdaptiveFog

- reduces RTT in around 15ms (fog) and 9ms (cloud)
- reduces STD by half

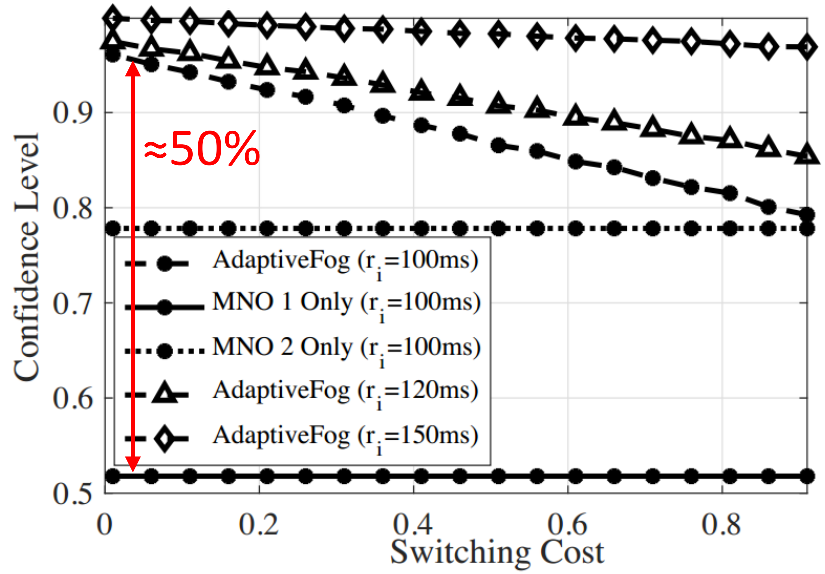
Optimal Network/Server Selection & Adaptation

- ***Confidence level***

Cloud



Fog



AdaptiveFog

- ➢ achieves almost 30% improvement in confidence level for cloud
- ➢ achieves almost 50% improvement in confidence level for fog

Summary

- ✓ AdaptiveFog is the first framework supporting the vision of 5GAA for supporting multi-operator connection in smart vehicle
- ✓ Compared to average/instantaneous latency value, confidence level is a more realistic metric to quantify service performance
- ✓ AdaptiveFog achieves 30% and 50% improvement in confidence level for fog and cloud
- ✓ Future work: Extending AdaptiveFog into more generally scenarios (e.g., with processing latencies offered by different fog service providers)

For more information, please contact:

Yong Xiao (yongxiao@hust.edu.cn)