DISTRIBUTED RESEARCH ON EMERGING APPLICATIONS & MACHINES
Department of Computational & Data Sciences
Indian Institute of Science, Bangalore

DREAM:Lab

Case Studies of IoT Applications

Yogesh Simmhan (Lecture)
Suman Raj, Shreya Mishra (Tutorials)

भारतीय विज्ञान संस्थान INDIAN INSTITUTE OF SCIENCE

Computational and Data Sciences

ACM India Summer School on IoT & Embedded Systems, IIIT-B
12 Jul, 2021

Era of IoT: Sensing and Communication

- Affordable Sensing
  - Present within infrastructure
  - Deployed as part of appliances, commercial building, Manufacturing Industries
    - Surveillance cameras, Electric & Water meters, Pollution sensors
    - Thermostat, parking indicators, equipment monitoring
  - Smart phones as a human sensing platform
    - Allows crowd sourced sensing

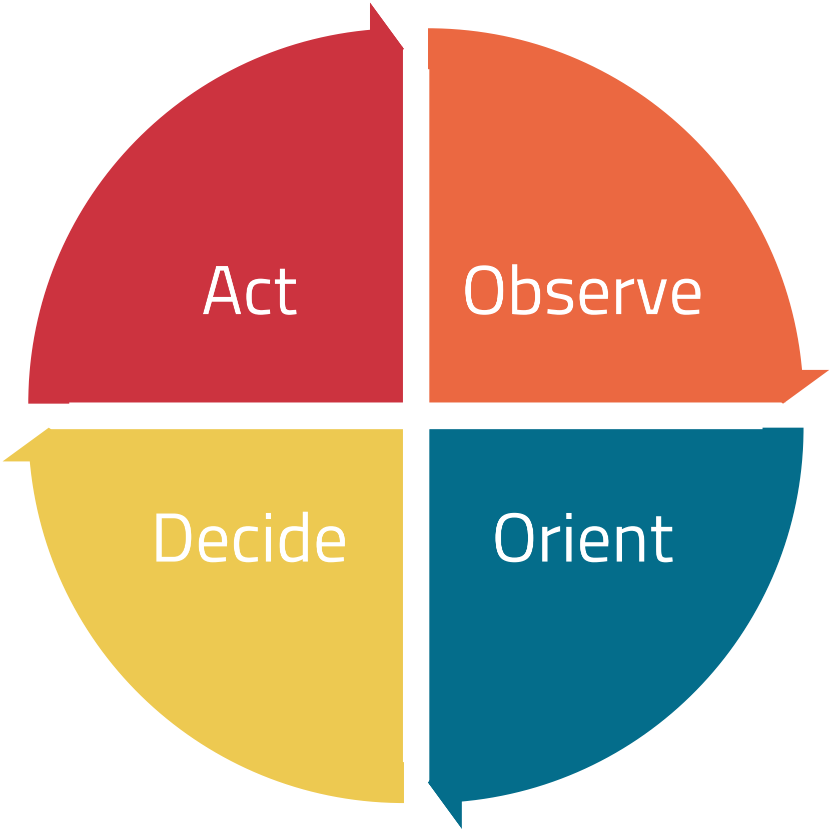
- Pervasive Communication
  - 2G/3G/4G/5G
  - Low power, ad hoc networks, NBIoT
  - Move data between sensors/actuators and analytics

Era of IoT: Push for *smarter* services

- What can we do with this sensing/actuation/communication fabric?

- Faster, reliable, affordable citizen services
- eGovernance for transparency in services
- Information-driven decisions
  - Visibility into the infrastructure & human system
  - Automation to take rapid action
  - Monitoring impact of actions

OODA Loop



An intrinsic property of IoT Applications

IoT as a Big Data Challenge

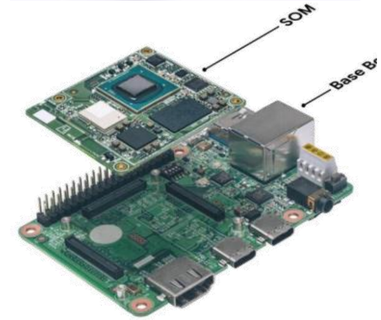
- Information-driven decision making requires rapid access to data
- Data from city-scale infrastructure can be large
- Data from sensors can arrive rapidly
- Scalable Big Data Platforms & Analytics are needed for decision making

Cloud Computing: An IoT Enabler

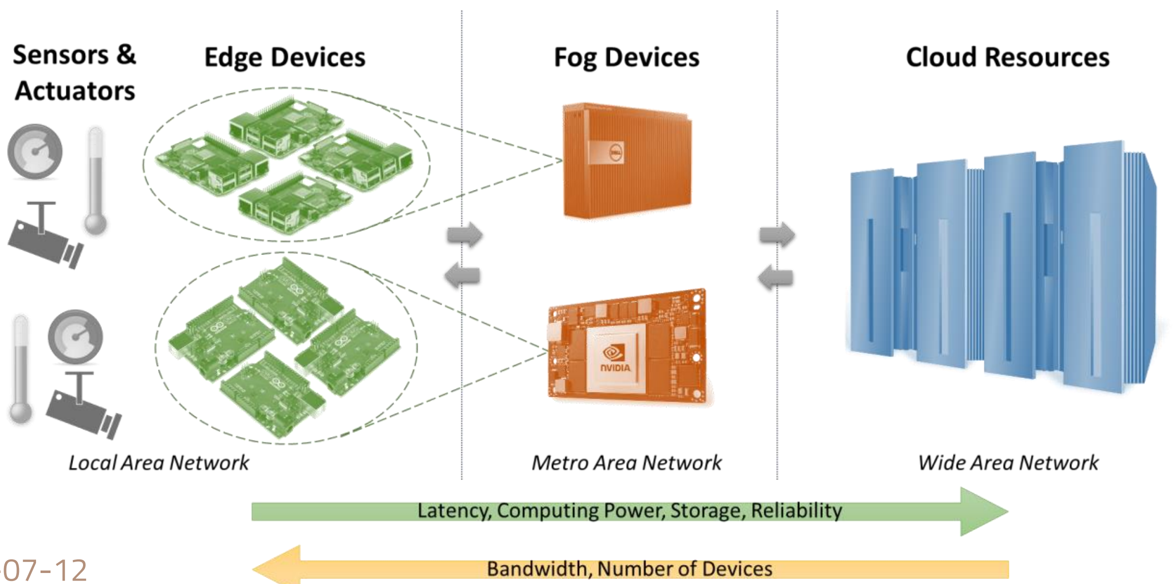
- Large Data Centers
- 1000's of racks, 100k servers
- Virtualized infrastructure
  - On-demand: Rent VMs by the minute
  - 100 machines for 10mins costs same as 1 machine for 1000 minutes
- Makes use of economies of scale. Much cheaper than on-premises servers.
  - Reduce operations costs (energy, personnel)
  - Ensures capital costs (servers) are fully used
- *Have you used the Cloud?*
- Traditionally: Move IoT data to the Cloud, process/store it

Edge & Fog Computing

- Captive edge devices abound
  - Micro-controllers on sensors and motes, “RPi-class” gateway devices, “Edge TPU”
  - Almost free
- Enhanced by fog/micro data centers
  - Accelerated workstations (e.g., TX1), small clusters
  - Serves a community or city.
  - Like a content distribution network (CDN)
  - These can complement cloud computing



The Edge TPU development kit—SOM (above) and base board (below)



IoT Applications

- What problems are solved?
- What technologies are used?
- How do we integrate different components?
- How do we validate the solution?
- What are the research challenges?
- What are the gaps and limitations?

Additional Reading

- Misra, P.; Simmhan, Y. & Warrior, J. , Towards a Practical Architecture for Internet of Things: An India-centric View, IEEE Internet of Things Newsletter, 2015
- Simmhan, Y. & Perera, Big Data Analytics Platforms for Real-Time Applications in IoT, Big Data Analytics: Methods and Applications, 2016
- Varshney, P. & Simmhan, Y., Characterizing Application Scheduling on Edge, Fog and Cloud Computing Resources, Software: Practice and Experience, 2020, Vol. 50 (5) , pp. 558-595

IoT Case Studies

- **Smart Cities**
  - Demand Response Optimization in Smart Power Grids
  - Equitable and safe smart water distribution
  - Low-cost urban air quality monitoring
- **Smart Health**
  - COVID-19 contact tracing
- **Smart Mobility**
  - Drones as autonomous platforms

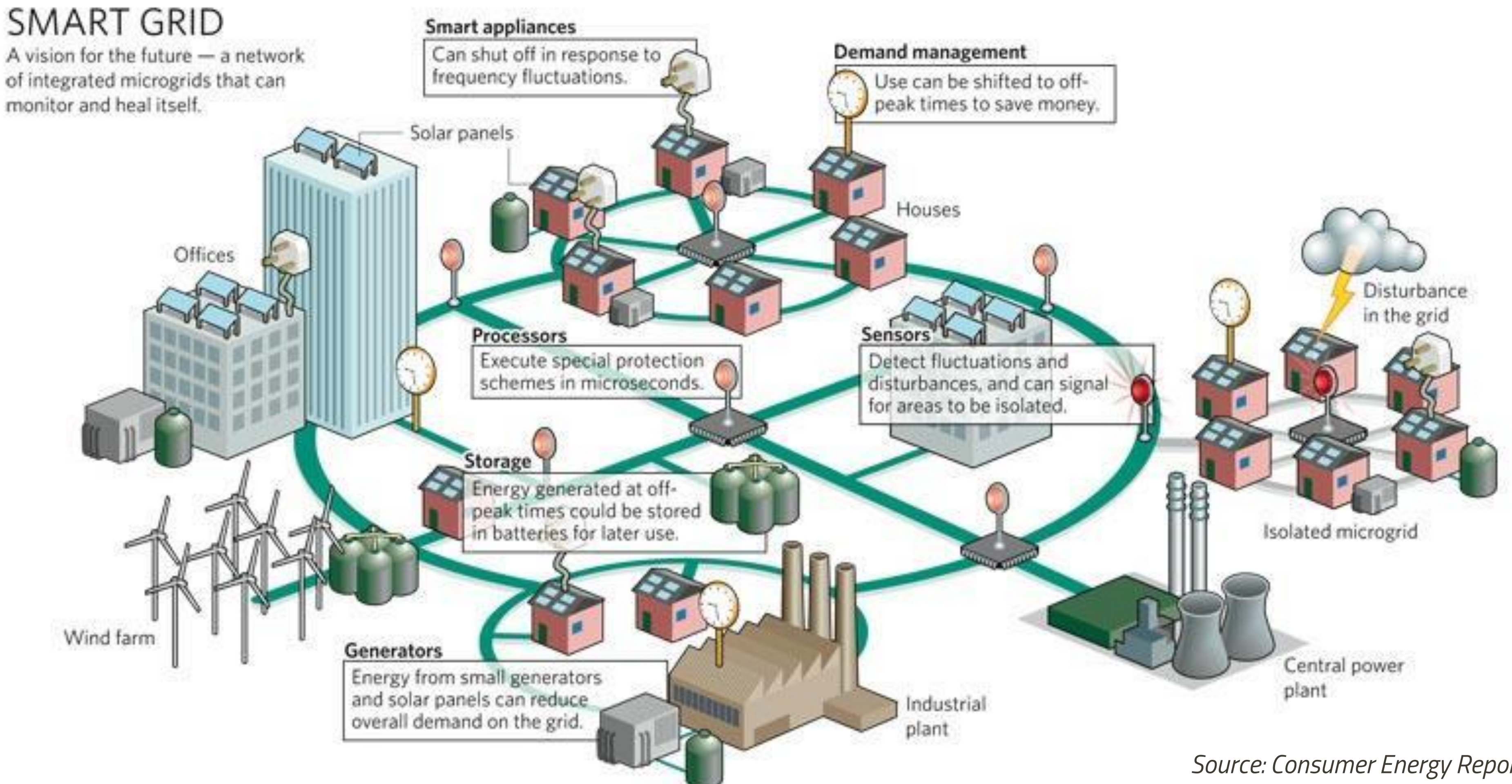
Smart Cities

Demand-response Optimizing for Smart Power Grids

Acknowledgements: V. Prasanna, USC

Smart Grid Cyber-Physical System

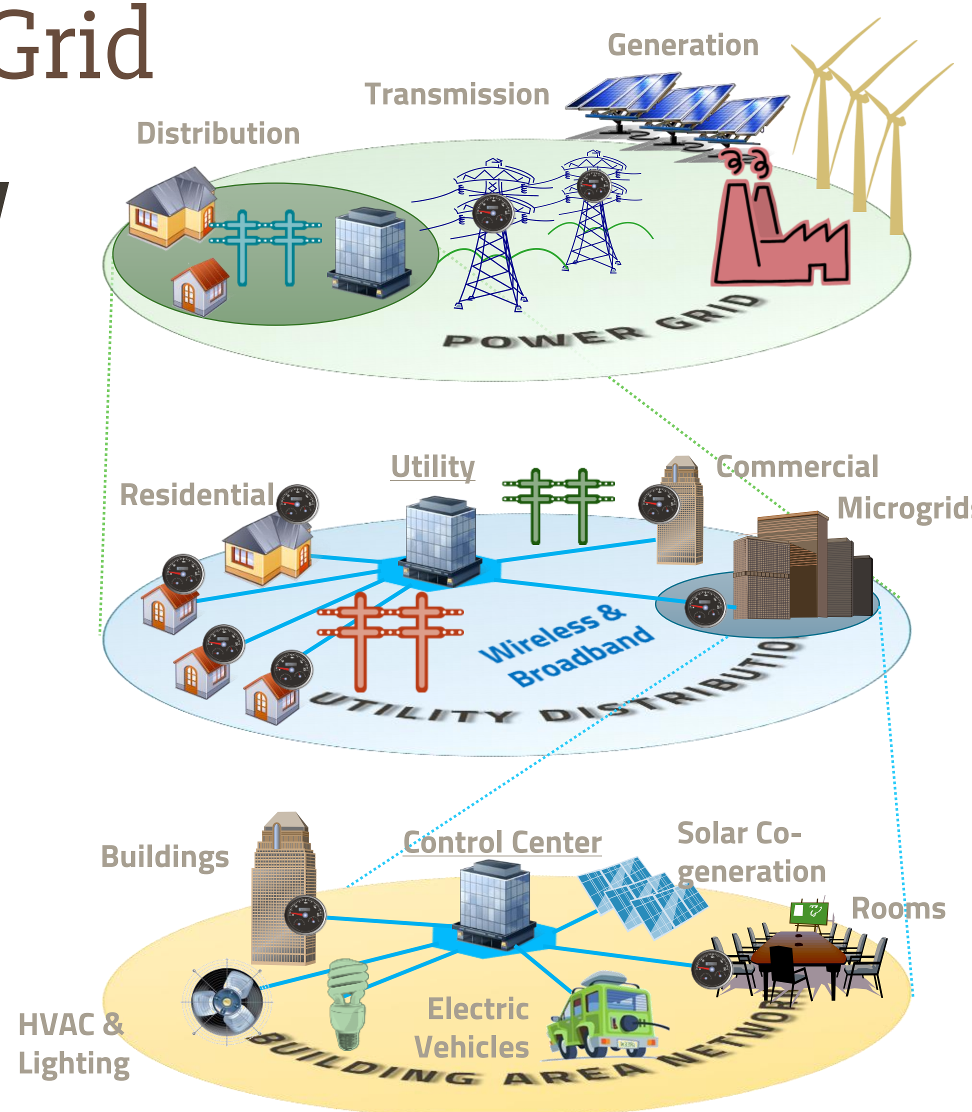
Networked sensors & actuators – *Building Sensors, Smart Meters & PMUs* – offer unprecedented "realtime" visibility & control



Source: Consumer Energy Report

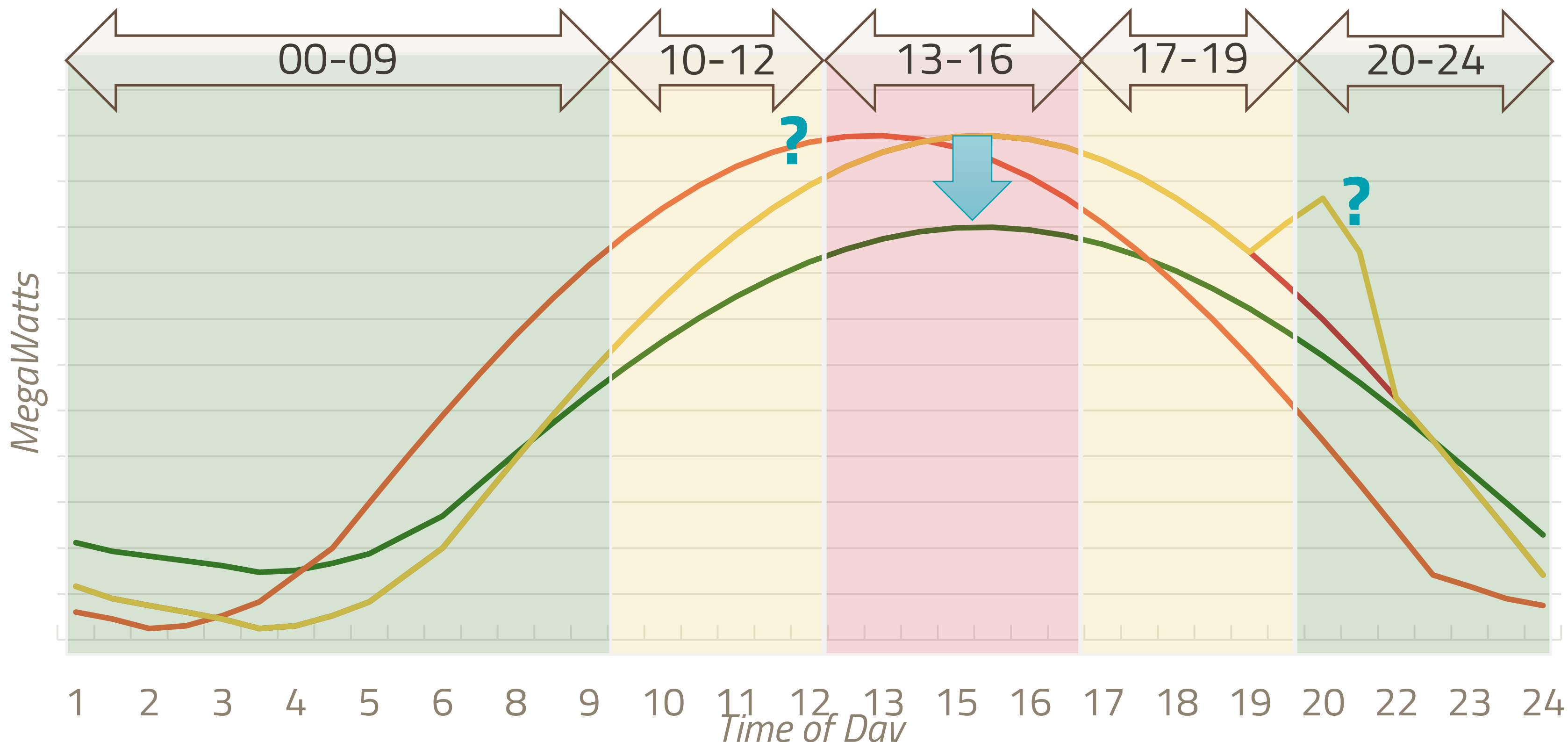
Los Angeles Power Grid

- **Largest Municipal Utility in the US**
  - Dept of Water & Power
  - *3.7M Residents, 1.4M Cust.*
  - Regulated by City
- **Vertically Integrated**
  - Owns *Generation, Transmission & Distribution*
  - Facilities in *CA,AZ,UT,OR,NV*
- **Power Use**
  - Annual: 25,000 GWh
  - 7200MW peak load
  - *~1% of US Load*



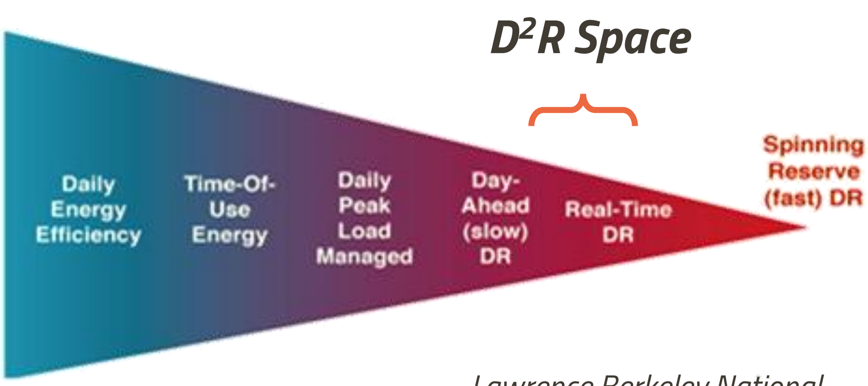
Demand Response Optimization

- Adjustment of power demand thru' *Shedding, Shifting and Shaping*, by the consumer at specific periods … **<u>or in response to utility signals</u>**



Dynamic Demand Response (D^2R)

- Supply & Demand are dynamic
  - Help decide when, by how much & how to reduce demand … *in near realtime!*
- *Dynamic* decision variables
  - *Start time (t)*
  - *Duration (Δt)*
  - *Depth (kWh)*
  - *Selection of customers*
  - *Selection of strategy*



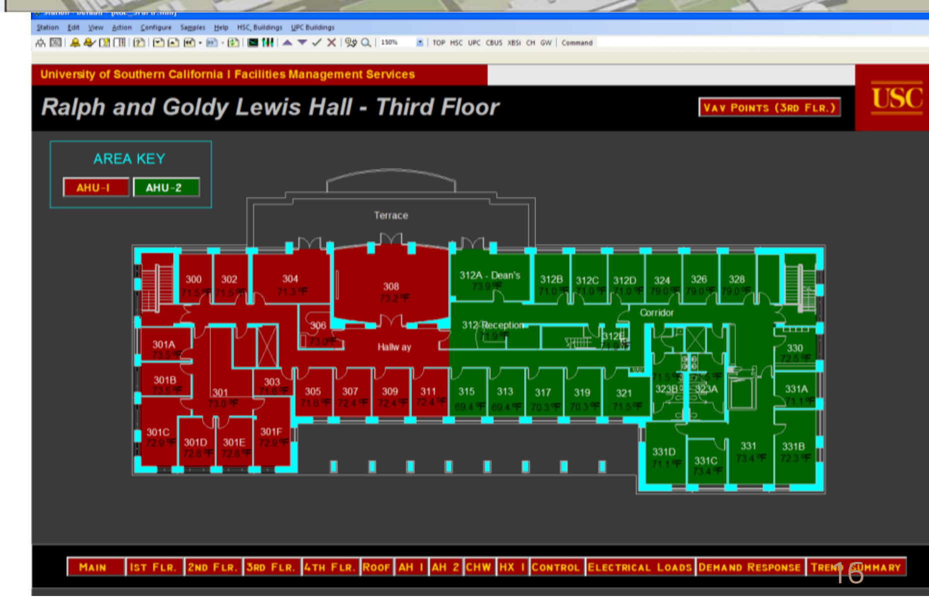
Lawrence Berkeley National Lab

➲ *Scalable Software & Analytics Stack to help put the "Smart" in Smart Grids*

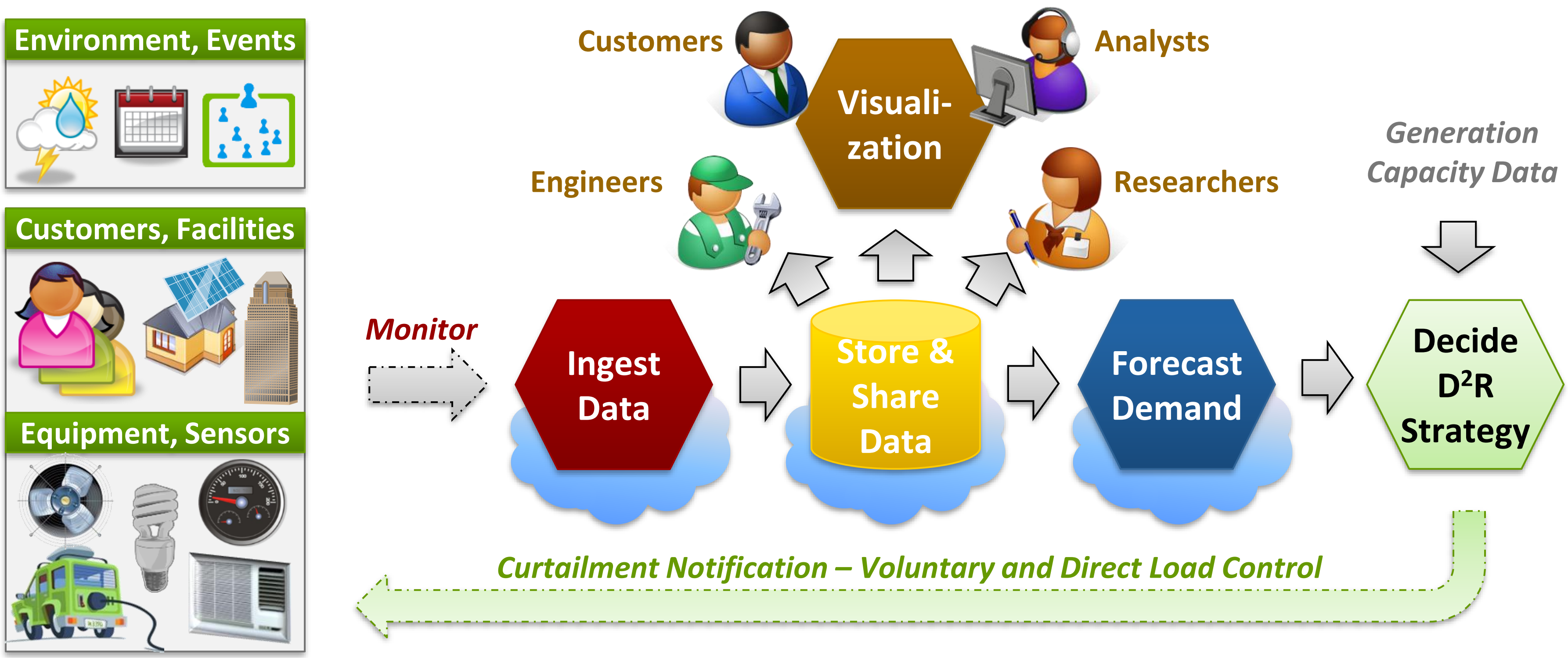
USC Campus Microgrid

- 170 buildings (UPC & HSC), 50K people
  - *Office, Acad & Lab, Restaurants, Dorms*
- Instrumented sensors on LAN
  - *Status, load, flows*
  - *O(1000) sensors*
    - » *O(100) active use*
- Direct load control
  - *HVAC, Lighting, Lifts*
- No Co-Generation

USC FMS



D²R Data & Control Flow

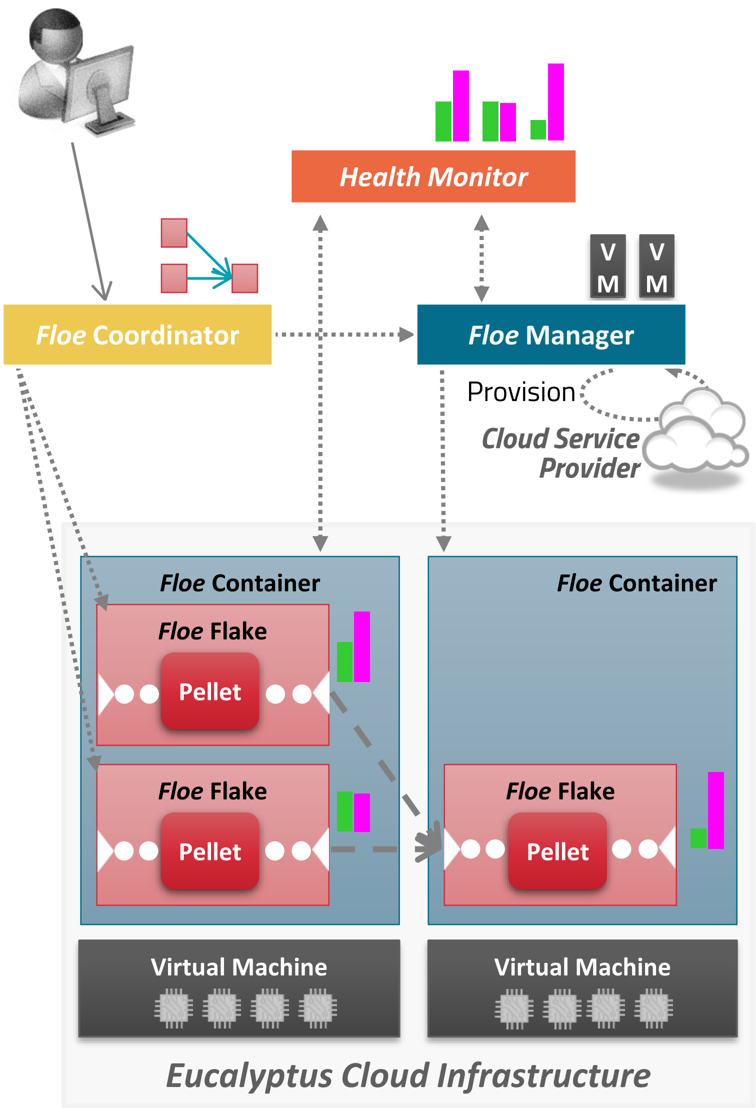


Feeds & Speeds

- Utility has O(M) Sensors, O(min) sampling
- Campus has O(K) Sensors, O(min) sampling
- Sensor data used for modelling & prediction
  - Need low latency data acquisition & analysis
  - Conserve Bandwidth & Computation
  - Be intelligent!
    - Throttle rates
    - Elastic scaling of compute resource
    - Use sampling of customers
  - Be adaptable
    - Easy specification, Modular extensions *(data sources, analytics)*

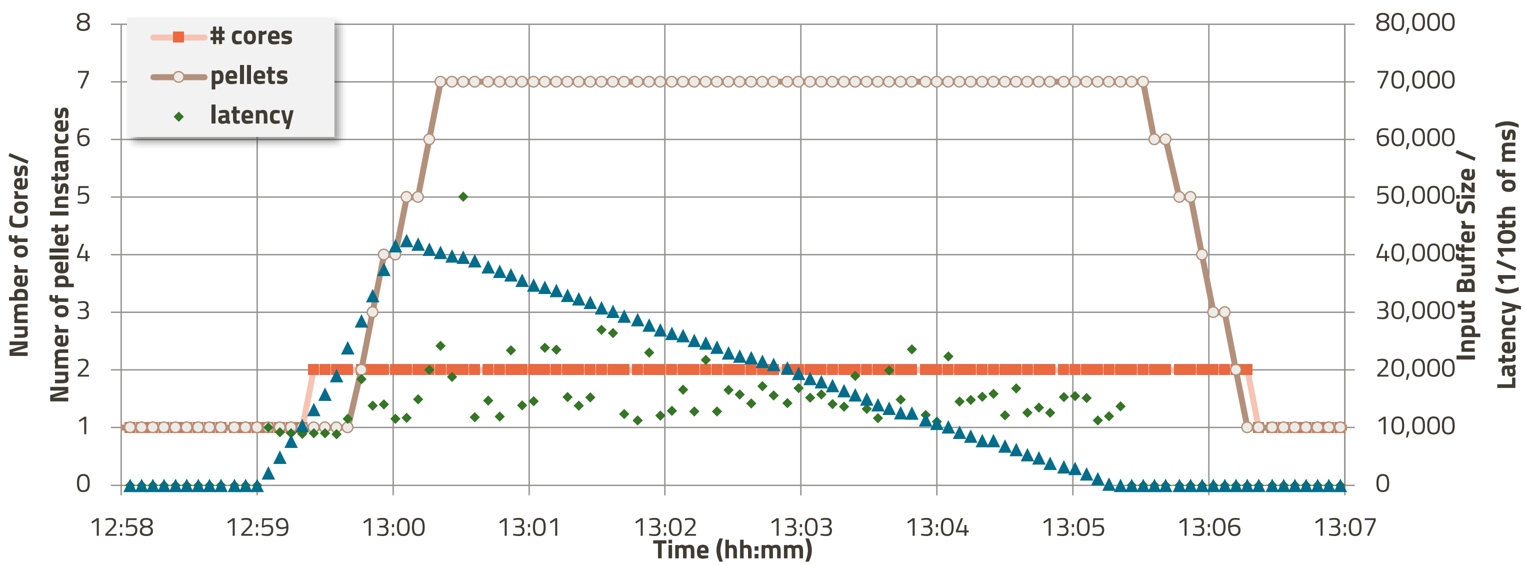
Floe: Continuous Dataflow Engine for Clouds

- Flexible dataflow engine
  - Stream & file processing
  - M:N, multiports, push/pull
- Distributed coordination
  - Component model
- Dynamic deployment & recomposition
  - Cloud-friendly
- Adaptability & adaptivity
  - Optimized scalability for dynamic applications

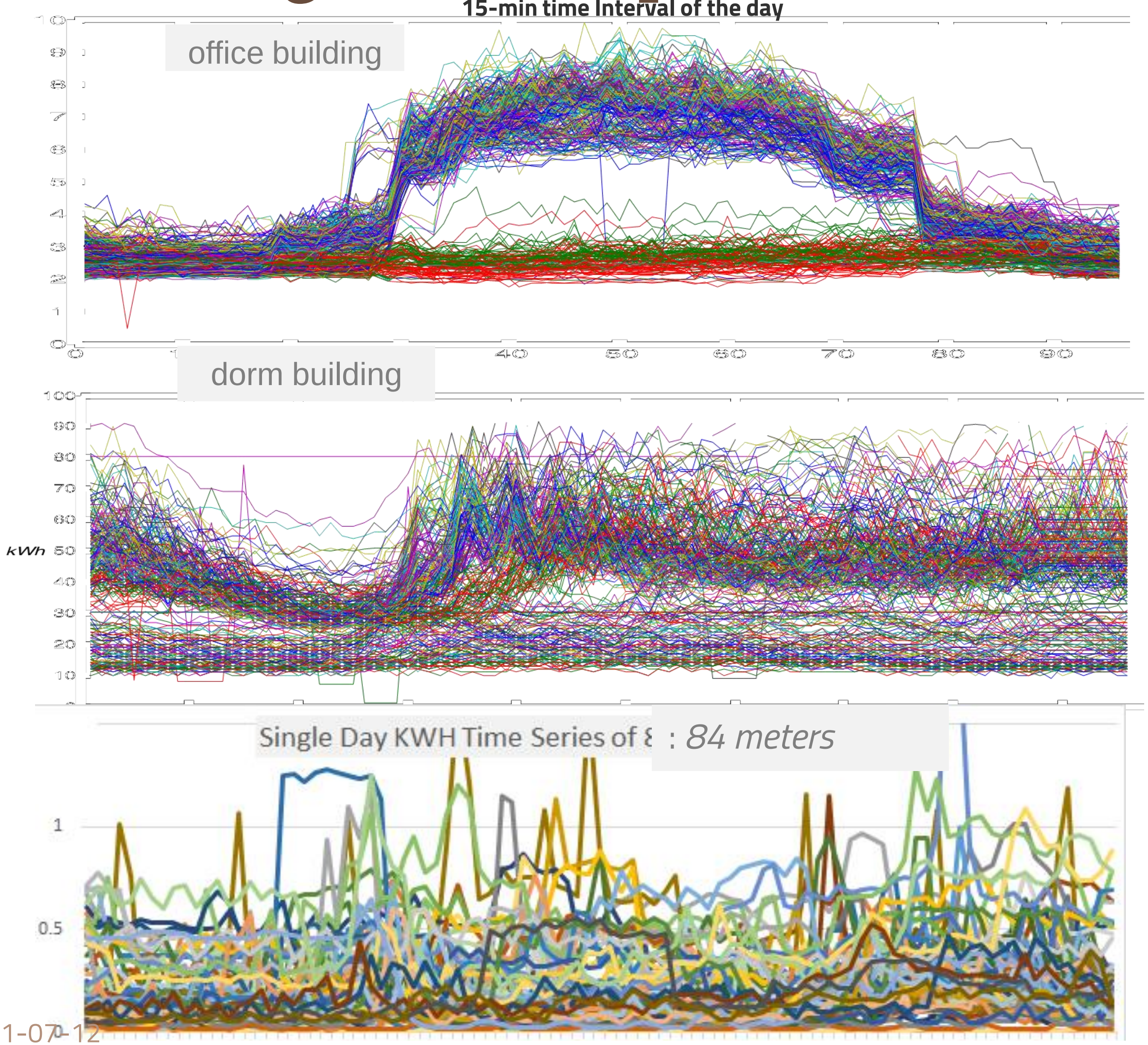


Adaptive Scaling on Cloud VMs

- Realtime Instrumentation
  - Message buffer size, CPU core utilization, Processing latency
- Adaptive Control
  - Cores allocated: *O(sec)* | VMs allocated: *O(min)*
  - Concurrent workers: *O(sec)*
- QoS Optimization
  - VM Cost (core-mins) | Latency/Throughput



Modelling Consumption Prediction



Modelling Consumption Prediction

- Time-series Models (ARIMA)
  - A time-ordered sequence of observations of a variable
  - Predict the future values of a variable based on previous observations
  - ✓ No domain knowledge required. Uni-variate model. Can capture trends & seasonality.
  - × Parameter estimation. Recent data.
- Data-driven Models (Regression Tree)
  - Decision tree with numerical output, built though partitioning on features to maximize distance
  - ✓ Making predictions just requires tree lookup. Tree can be interpreted from a domain perspective. Model is time-invariant (long horizon). Robust to missing data.
  - × Data collection & training overhead
- Baseline averaging models (ToW, CA ISO)

Curtailment Strategies

- Direct Load Control (HVAC-based)
  - Global Temperature Reset (GTR)
  - Variable Frequency Drive (VFD)
  - Duty Cycling
  - *Hot & Cold Calls*
- Voluntary Control
  - Request campus members/ residents to take action

USCSmartGrid
Living Laboratory

USC Energy Institute
Donald L. Paul, Ph.D.
Executive Director
William M. Keck Chair of Energy Resources

Hello Trojan,

This is to let you know that the USC campus-wide Save Power Event will be starting in the next hour. The event will last from **1pm-4pm today**, and participation is voluntary.

This is a reminder of the Save Power Event Program and some helpful suggestions to aid you in your energy conservation.

During the Save Power Event, you are being asked to voluntarily conserve power in your work stations for a few hours by choosing one or more power-saving actions, such as adjusting thermostats, using fewer lights, reducing machine usage, heat, and air conditioning. Please join your colleagues in this effort to prevent power outages and reduce energy waste.

Here are some suggestions for actions you may choose to take during the Save Power Event today:

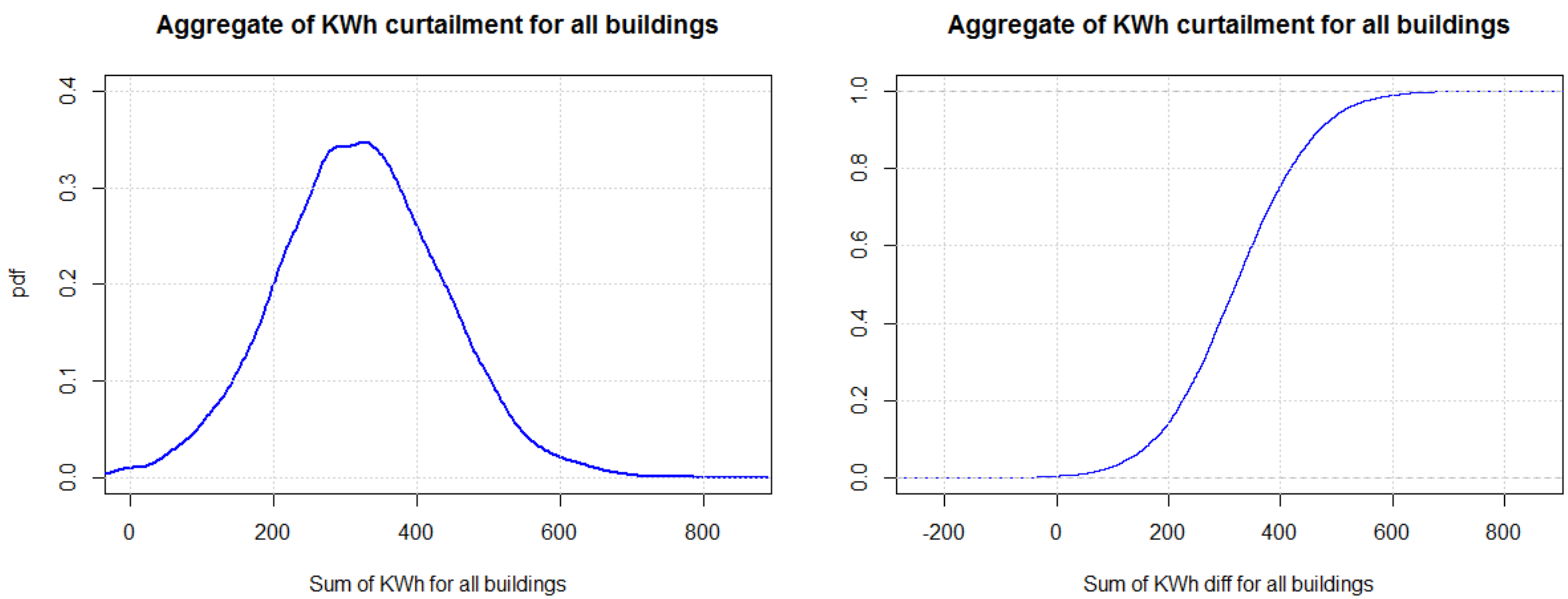
- Use e-copies instead of printing
- Turn off machines not in use
- Keep doors closed if it helps temperature regulation
- Wait to use office equipment (e.g. printers, copies, scanners)
- Manage temperature and lighting with windows and shades
- Turn off or unplug machines not in use
- Set your devices to Energy saver mode
- Regulate the thermostat (68 degrees when it is cold, 78 degrees when it is hot)
- Keep exterior doors and windows closed to conserve heat (when cold outside) and cooling (when it is hot outside).

It is appreciated that you are taking the time to conserve energy and reduce demand. Thank you!

If you would like more information about the program, please visit the website at http://smartgrid.usc.edu, or feel free to email spe-team@usc.edu.

Why we need non-naïve soln?

Probability and cumulative density functions for the power curtailments if DR event is performed on 27 campus buildings simultaneously.



- ~50% of events will result in 1h curtailment < (>) 300KWh
- Only 5% of the events will result in 1h curtailment > 500KWh

Optimization Problem

- Given: Period (t_0, t_n) & depth of curtailment (V)
- Select: A set of strategies to meet target
  - *Aligned* Piecewise Curtailed Load Estimation
    - » Curtailment period equals strategy period
    - » Large and small spatial scope

Given a set of floating-point values $\{A_{1jt}, A_{2jt}, \ldots, A_{Njt}\}$ for N~50,000, $0 \le j \le 2, 1 \le t \le 16$

where $A_{s0t} = 0$, and a positive value V. Find a combination of $\{j_1, j_2, \ldots j_N\}$ such that

$$\sum_{t=1}^{16} \left(\sum_{s=1}^{N} A_{sj_st} - V\right)^2$$ is minimum.

A Heuristic Solution

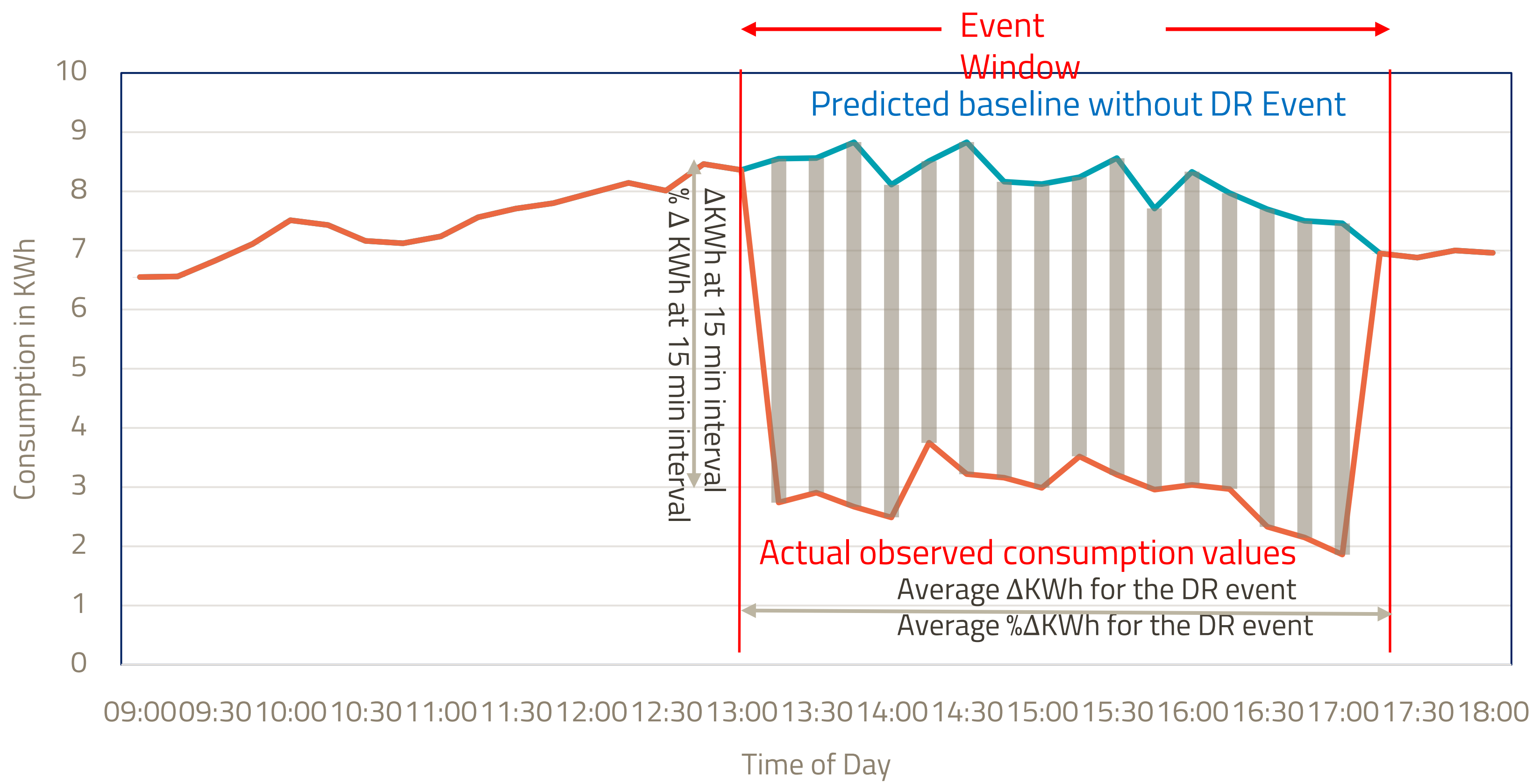
- Reduce to "change-making" problem
  - *Given a set of coins of different denominations, find the smallest number of coins required to meet a value*
- Greedy heuristic to create coins from building strategies
  - Ensure fairness, distribution
- Dynamic programming to create target from coin values

USC Microgrid as a Smart Grid Testbed

- *"City inside a city"*
- Largest single private institutional customer of LA DWP
  - Electric load of 28 MW in 2009-10
  - Annual consumption of 210 million kWh of electricity, 5.3 million therms of natural gas, and 280 million gallons of water
- Diversity
  - Dorms, Classrooms, Labs, Hospital buildings, Restaurants, Public transportation, Offices,...
  - Total enrollment of 33,408 students and 13,000 faculty and staff
    - Nearly 8000 international students (highest amongst all US univs)
  - 301 acres

***"A Living, Learning Laboratory"** to test and validate system architectures, algorithms, operational scenarios, and behavior*

Curtailment Estimate



Smart Cities: Power Grid USP

- Real-time analytics using Cloud elasticity
- Real-time prediction, decision making and Actuation
  - Demand prediction
  - Optimization

- Aman, S.; Simmhan, Y. & Prasanna, V. , **Holistic Measures for Evaluating Prediction Models in Smart Grids,** *IEEE Transactions on Knowledge and Data Engineering (TKDE) , Vol. 27 (2) , pp. 475-488 , 2015*
- Simmhan, Y.; Aman, S.; Kumbhare, A.; Liu, R.; Stevens, S.; Zhou, Q. & Prasanna, V. , **Cloud-Based Software Platform for Big Data Analytics in Smart Grids ,** *Computing in Science and Engineering , Vol. 15 (4) , pp. 38 - 47 , 2013*

Smart Cities

Optimizing Smart Water Distribution Networks

Acknowledgements: M.S. Mohankumar, R. Sundaresan, IISc

Objectives

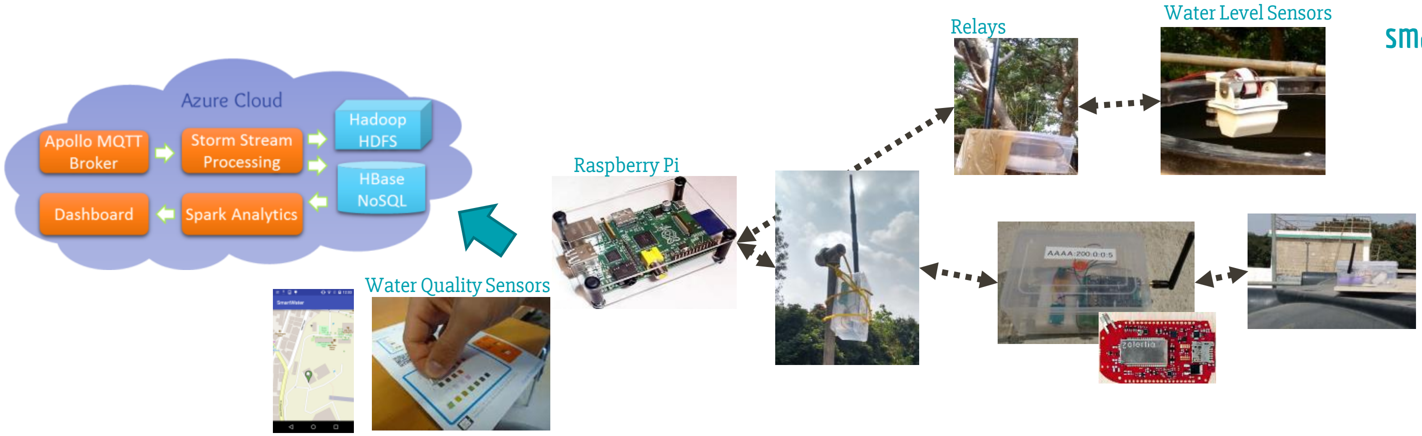
1. *Can we use IoT & Big Data technologies to make campus “smarter”?*
   - i.e. the “infrastructure”, not the people ☺
   - More efficient, reliable & safe resource delivery & management
   - Initial Case Study: **Water management**
2. *And in the process, understand the technology and improve on it?*

Smart Water Management

- Plan pumping operations for reliability
  - Avoid water running out/overflow
  - It can take 12 hrs to fill a large OHT
  - Water scarcity for several weeks in the year
- Provide safer water
  - Leakages, contamination from decades old N/W
- Reduce water usage for sustainability
  - IISc avg: 400 Lit/day, Global std: 135 Lit/day
  - Lack of visibility on usage footprint, sources
  - Rain water harvesting, Water recycling plant
- Lower the cost
  - Reduce water use & energy cost for pumping

IISc Smart Campus IoT Stack

- Hundreds of wireless motes & sensors
  - Realtime sensing of water level, quality, network health
  - Crowdsourced data collection from smart phones
  - *City-scale deployments will have even larger source and rates*
- Data analytics to drive online operational decisions
  - Water quality alert, water under/overflow, battery drain alert
  - *Smart power & transportation require even lower latency!*



smartx.cds.iisc.ac.in

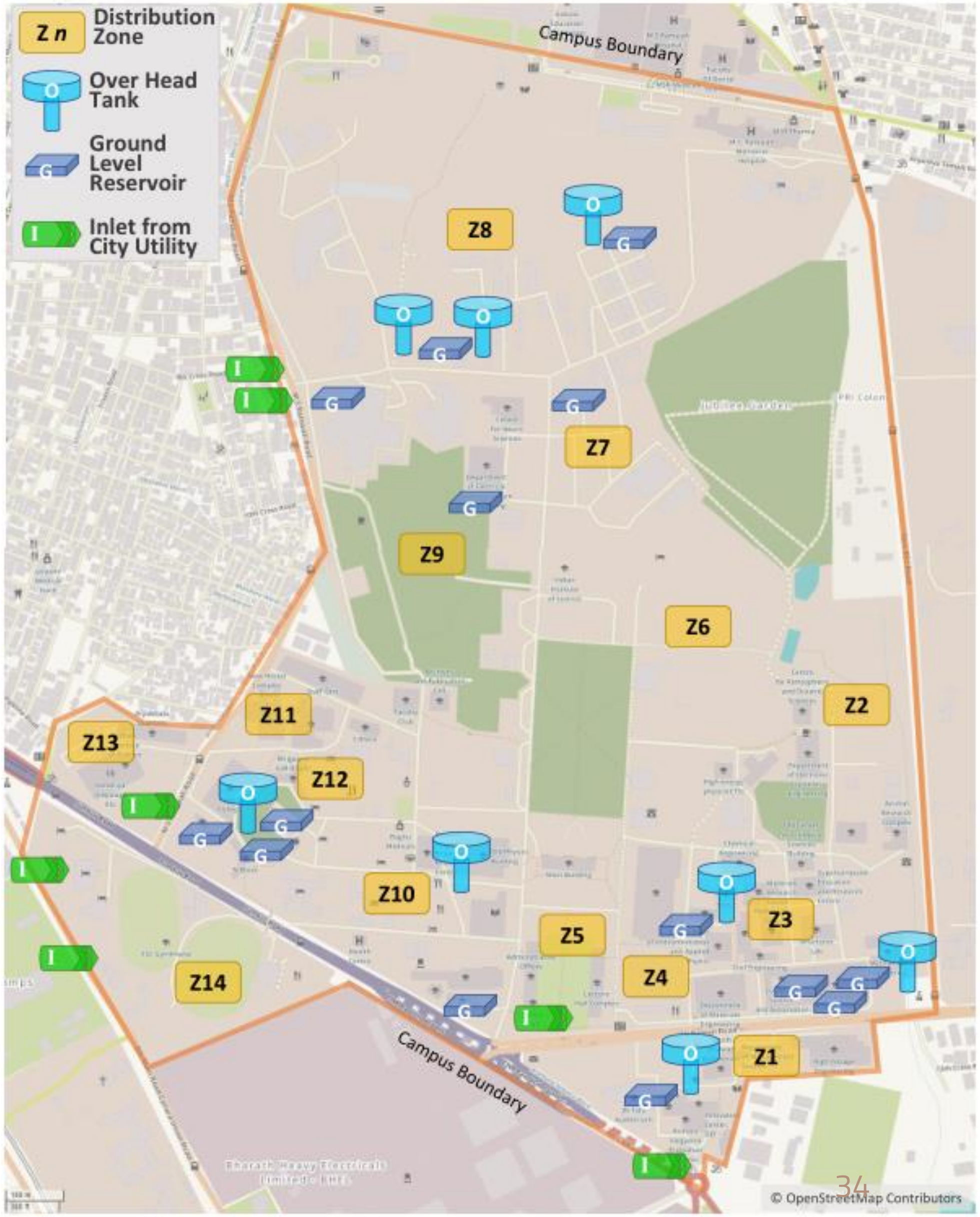
Towards a data-driven IoT software architecture for smart city utilities, Yogesh Simmhan, Pushkara Ravindra, Shilpa Chaturvedi, Malati Hegde, Rashmi Ballamajalu, *Software: Practice and Experience*, Volume48, Issue7, July 2018

Over Head Tank (OHT)
Ground Level Reservoir (GLR)
BWSSB Main Inlet

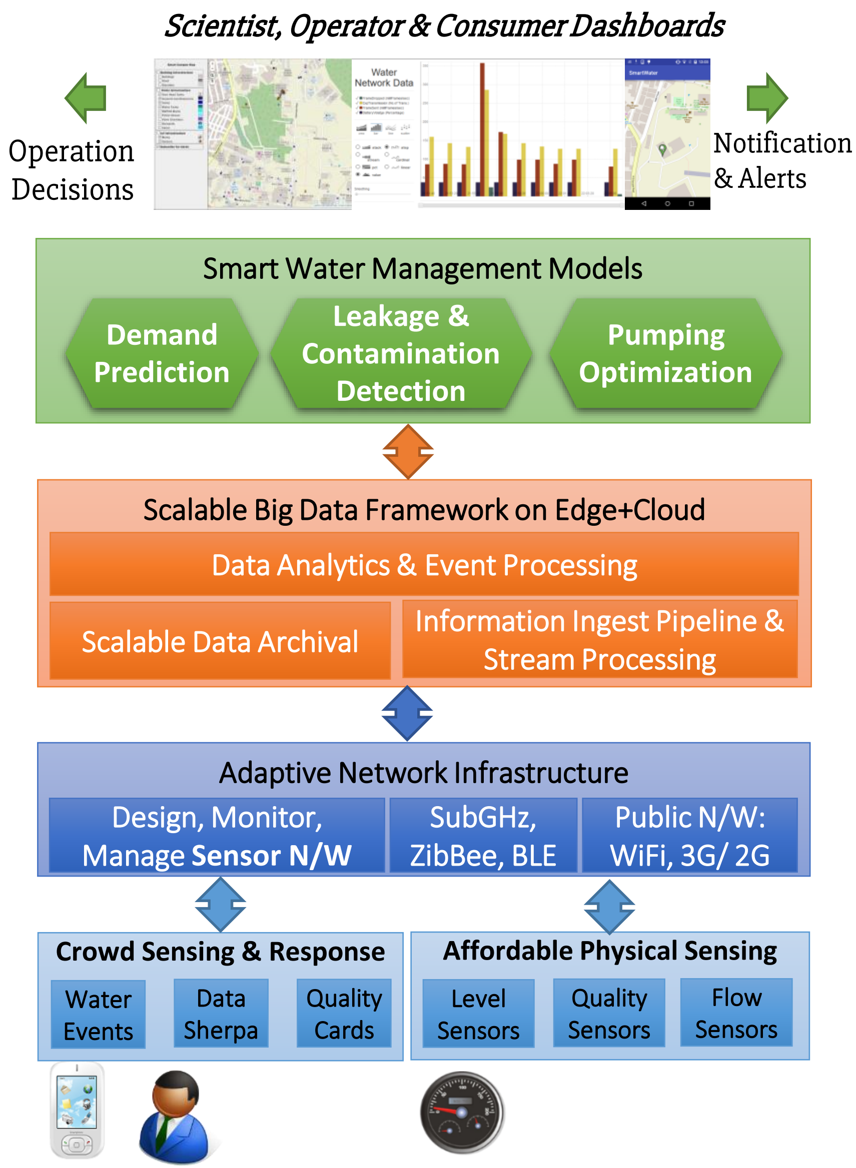
IISc Campus

- 440 Acres, 8 Km Perimeter
- 50 buildings: *Office, Hotel, Residence, Stores*
- 10,000 people
- Water Use: 4M Lit/Day
- 10MW Power Consumed

OHT	8
GLR	13
Inlet	4+3



End-to-End IoT Stack



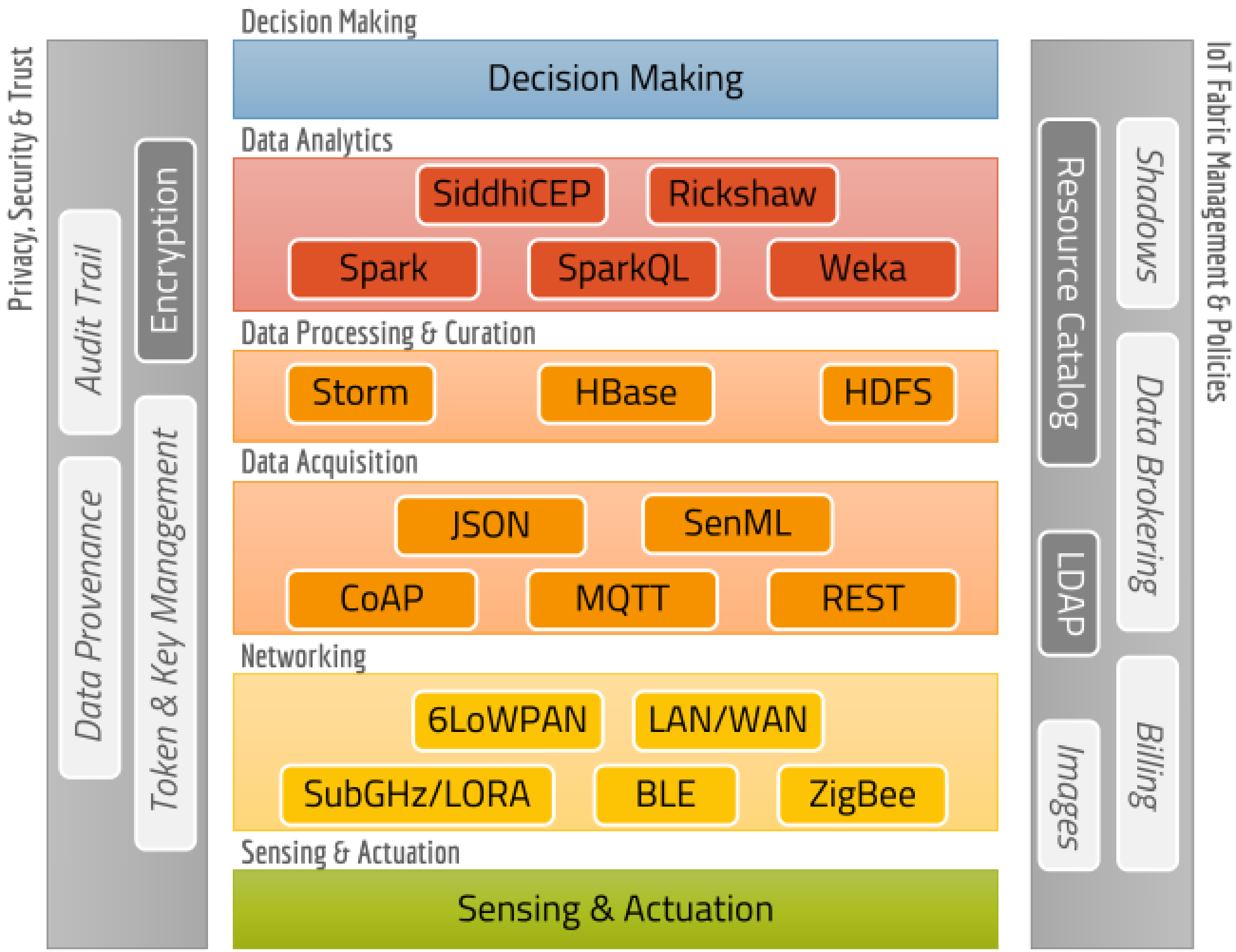
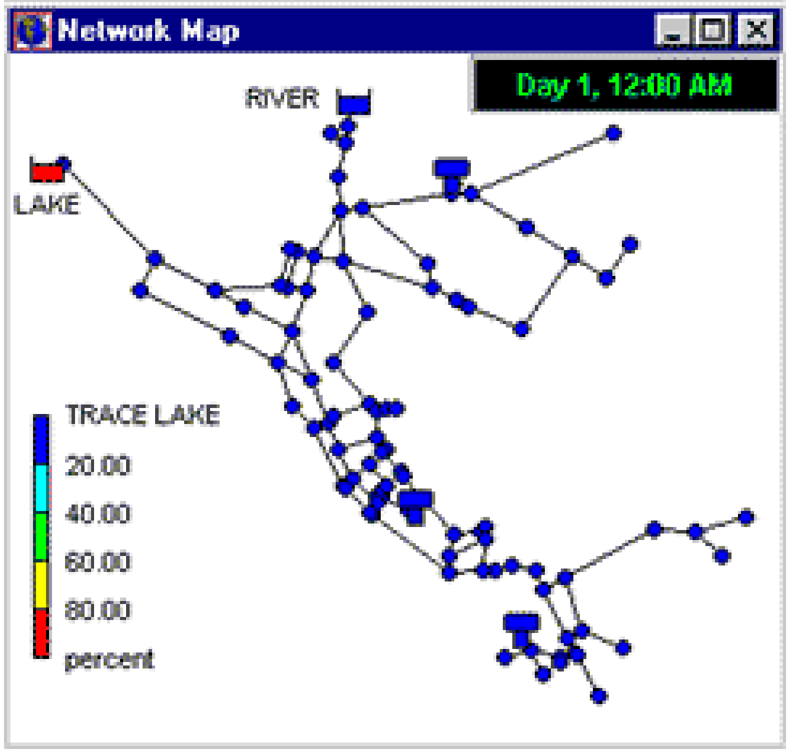


Figure 7. Protocols and standards used in the IoT architecture

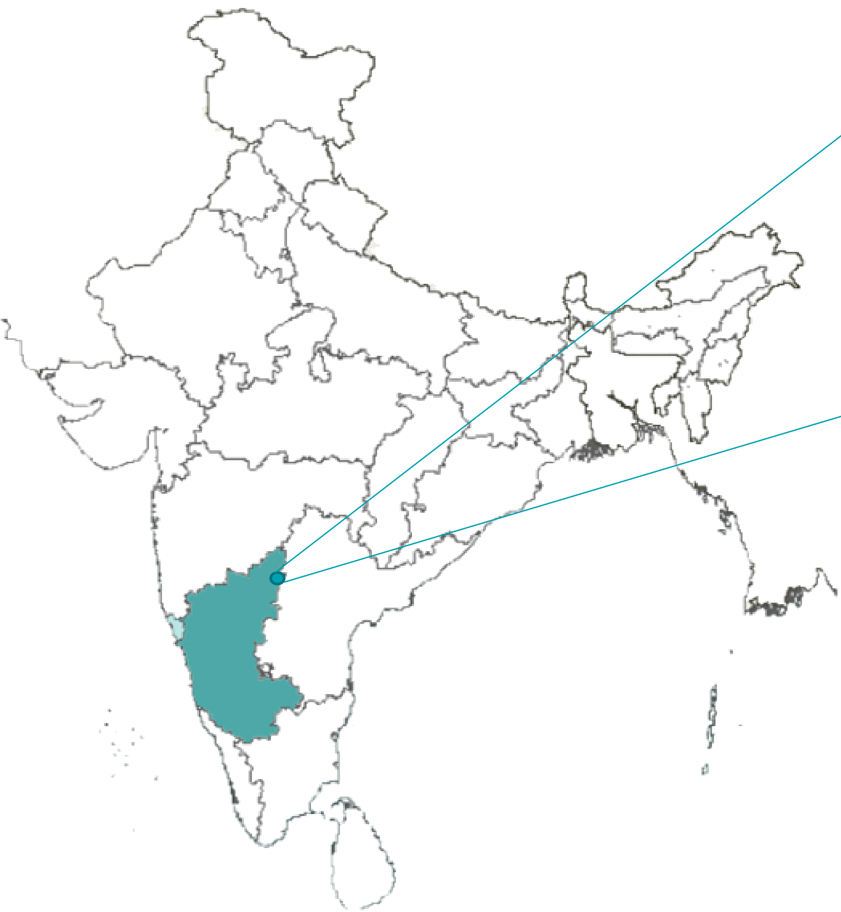
Scaling to the City

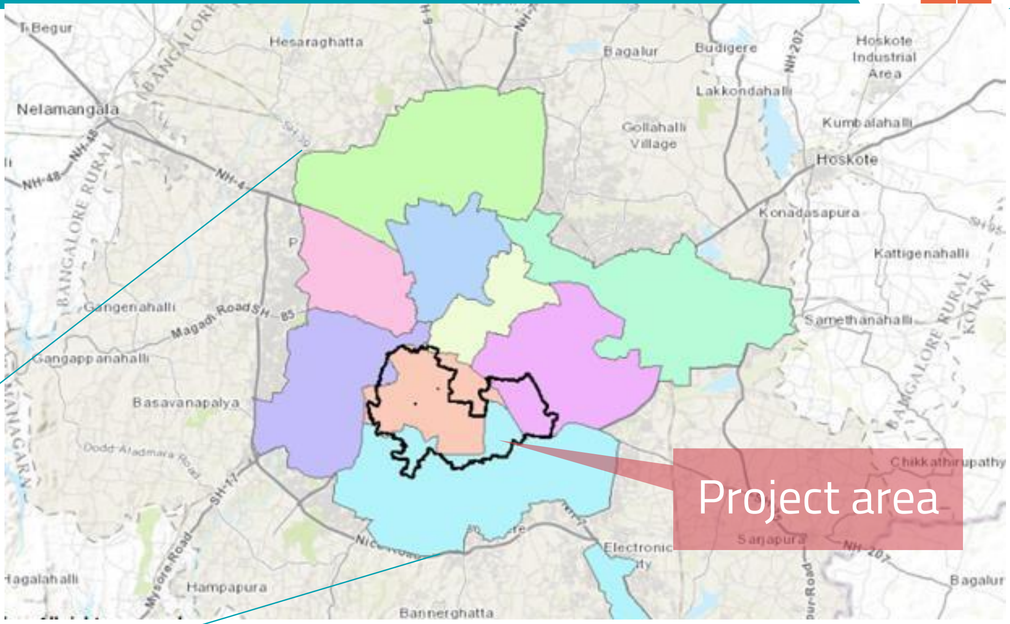
- Understand the complex water distribution network with existing measurement devices.
- Explain inequity in water distribution among the DMAs in the division
- Control and scheduling to achieve equitable distribution.
- Identify a candidate service station for a pilot test-bed and evaluate models/algorithms in achieving equity.
- Develop water quality models at service station level.
- Overall tool for effective management

How do we provide potable water in sufficient quantity, at adequate pressure and at an acceptable quality at a consumer's tap?



Bangalore Study Area





Source of water	Cauvery stage I, II, III, IV
Project area	52 sqkms
Population	1.52 Million
Water demand	278 MLD
Average supply	262 MLD
Ground water usage	11% + private bores
Total connections	0.16 Million
Total length of pipe network	1518 kms

13 Elevated service reservoirs (ESR)
Total storage capacity **11.83** Million liters (ML)

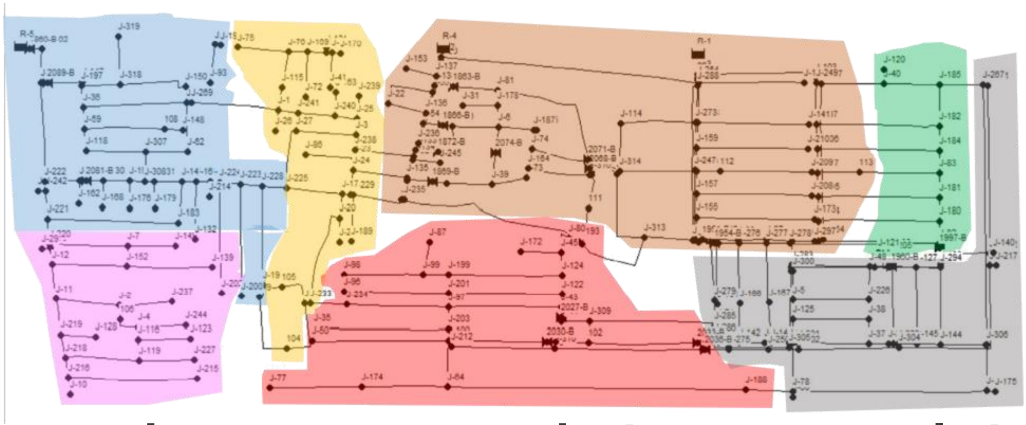
Sensor Placement for Water Quality Estimation

- Contaminants may be introduced in a *Water Distribution System (WDS)*
  - old infrastructure
  - accidental discharge
  - Malicious actors
- Installing water quality sensors in every junction within a large WDS is expensive
- *Given a fixed number of water quality sensors identify the optimal locations within the WDS to place these sensors*

Sensor Placement: Objectives

- **Mean time to detection (Z1)**
  - Elapsed time from the start of the contamination to the first detection of by a sensor
- **Mean estimated population exposed through ingestion (Z2)**
  - The number of people affected by the contamination events prior to detection by sensor
- **Mean volume of contaminated water consumed prior to detection (Z3)**
  - The mean volume of water is consumed through a contamination event before its detection.
- **Detection Likelihood (Z4)**
  - The total number of detected contamination events

Sensor Placement: Intuition

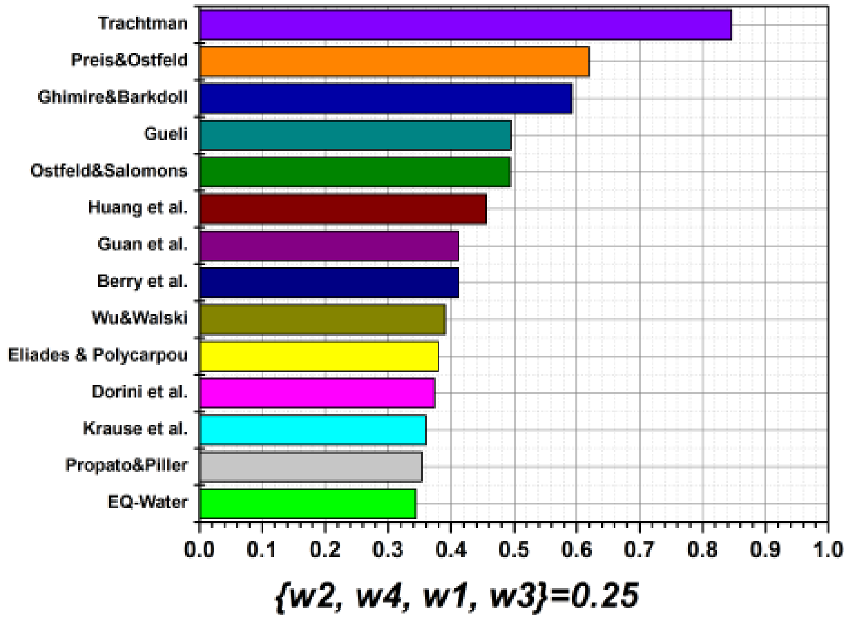


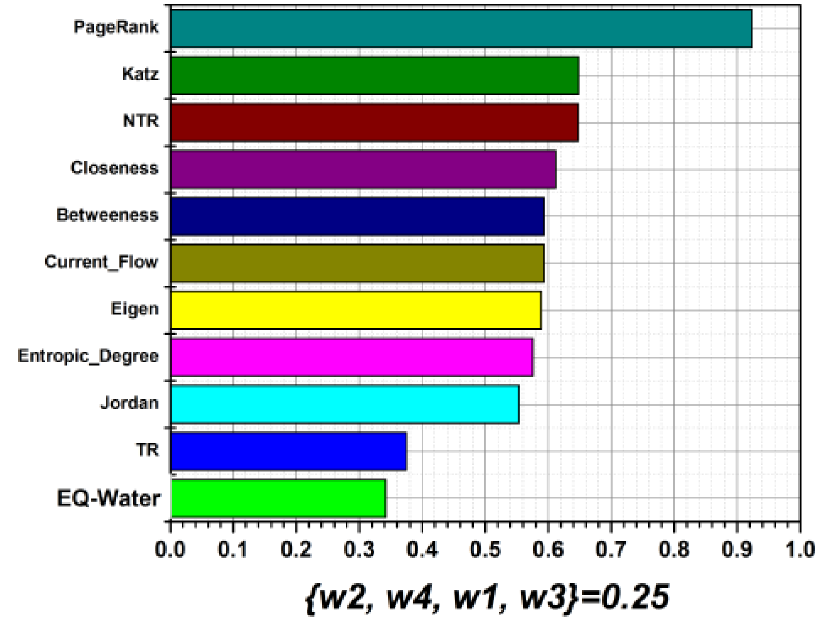
JP Nagar Network

- Cluster the network into multiple partitions, one per sensor
- Identify the ideal vertex within each partition for sensor placement
  - Use hydraulic models: Velocity and Pipe lengths
  - Use network models: Weighted Receivability

Results

- We outperform simulation and network based models





Smart Cities: Water Network USP

- Ad hoc sensor networks
- End to end software stack
  - Sensing
  - IoT protocols
  - Software Stack
  - Crowd sourcing
- Sensor placement problem
  - Deployment planning

- Simmhan, Y.; Ravindra, P.; Chaturvedi, S.; Hegde, M. & Ballamajalu, R. **Towards a Data-driven IoT Software Architecture for Smart City Utilities**, *Software: Practice and Experience, Vol. 48 (7), pp. 1390-1416*

Smart Cities

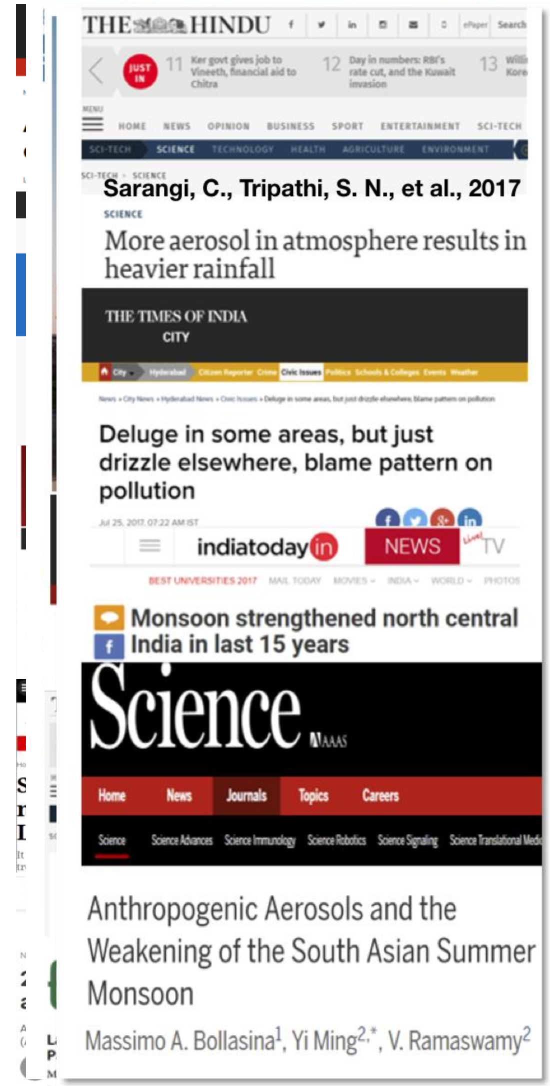
Low-cost Air Quality Monitoring using SATVAM

Acknowledgements: S. Tripathi, IIT-K, R. Zele, IIT-B

Toward Better Air Quality

- Outdoor air pollution leading environmental cause of cancer deaths
- Agricultural economy & Rainfall deeply affected by rising pollution
- India's Iconic Heritage Structures turning yellow due to pollution
- Affecting India's Solar Energy Goals

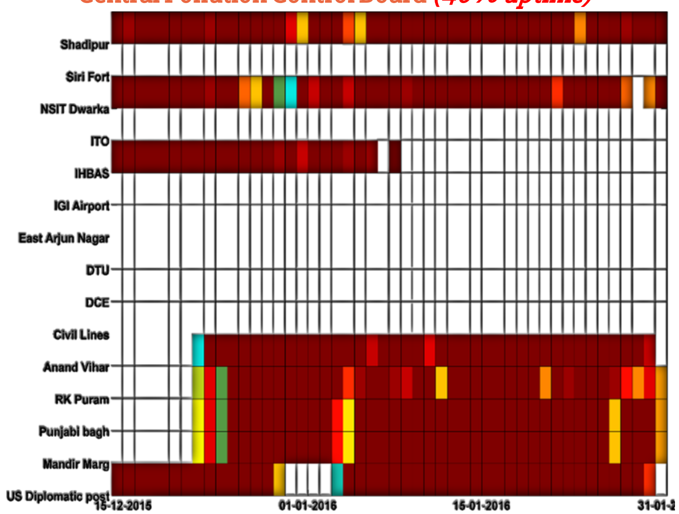
- *Severe adverse effects of Air Pollution on health of people as well as economy of the nation*



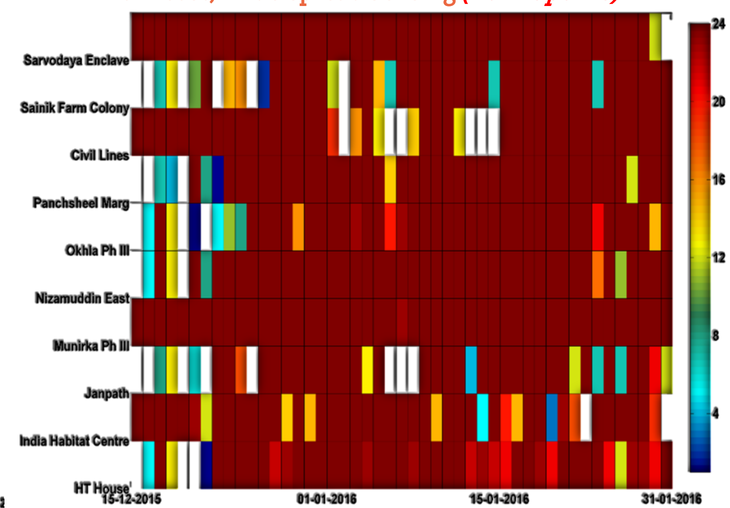
Technology & Societal Gaps

- Lack of systematic, large-scale, continuous urban sensing
- *Affordability* of sensors vs. *Accuracy* of observations
- Absence of 24x7 power and communication

Central Pollution Control Board *(40% uptime)*



Low-cost, Wide-spread Sensing *(80% uptime)*

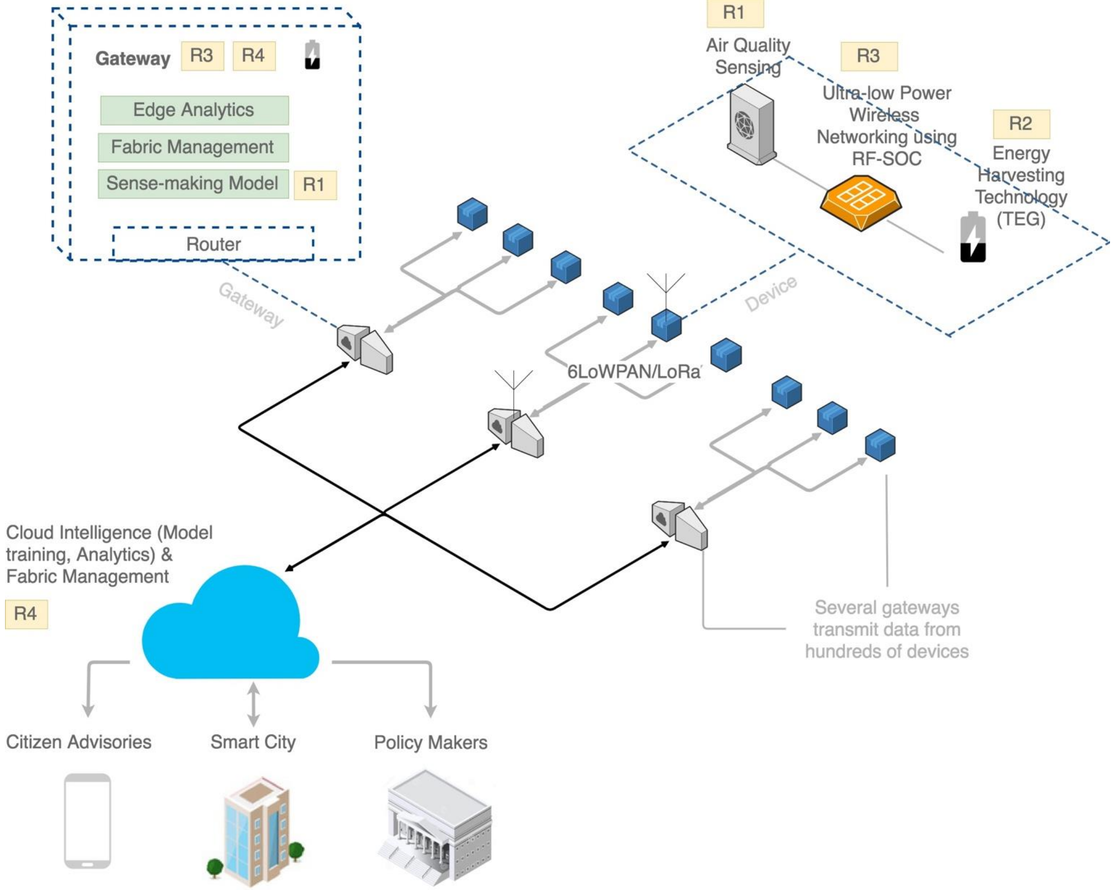


$PM_{2.5}$ Device Uptime During New Delhi Odd/Even Phase 1 (Chowdhury, Tripathi et al., 2017)

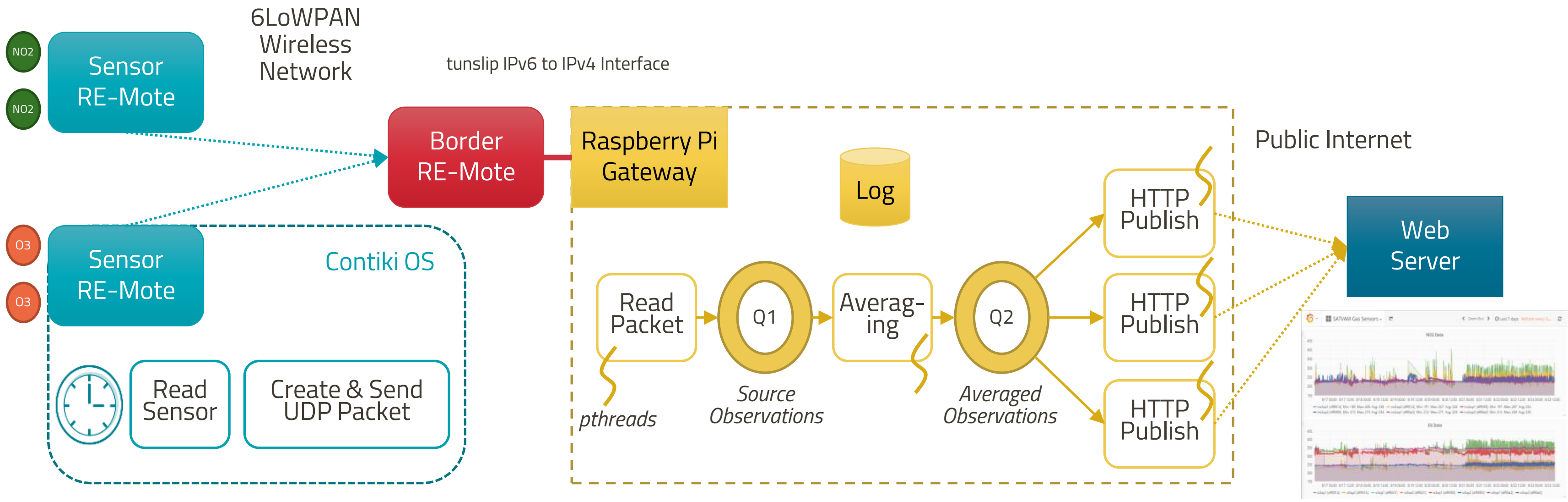
SATVAM: *Breathe Easy!*

- Streaming Analytics over Temporal Variables from Air quality Monitoring (SATVAM)
- Can we design affordable and accurate air quality monitoring systems for wide area deployment across cities?
  - **Sensing & Sense-Making**: Calibration, Edge analytics
  - **Energy-harvesting**: photovoltaic, concentrator
  - **Low-power networking & protocols**: 6LowPAN, Multi-hop
  - **Analytics on Edge & Cloud**: ML, temporal analytics, fabric health
- 5 years, Inter-disciplinary, Multi-institutional
- Supported by *Indo-US Science and tech. Forum (IUSSTF)*

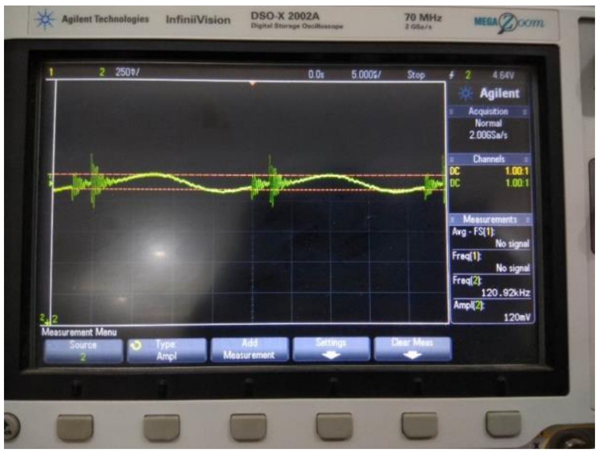
Architecture



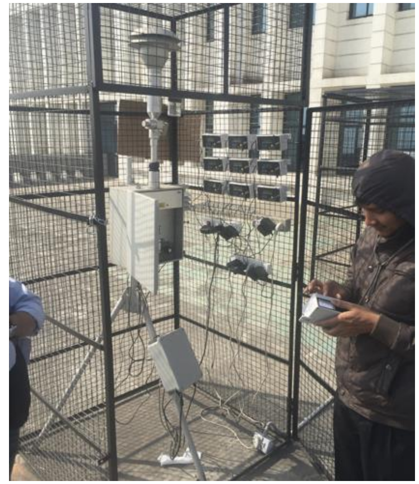
Edge Architecture



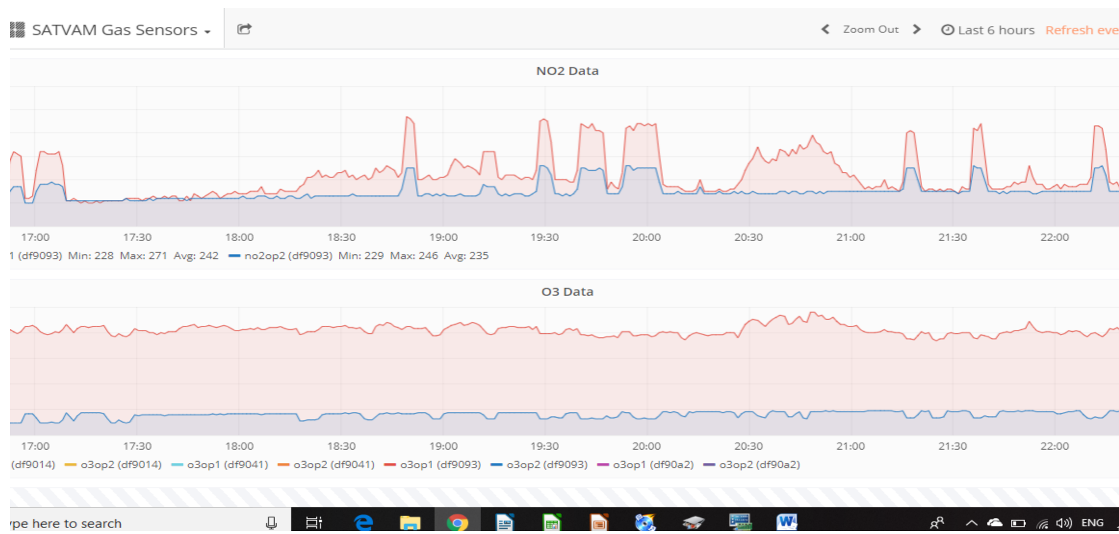
Deployment



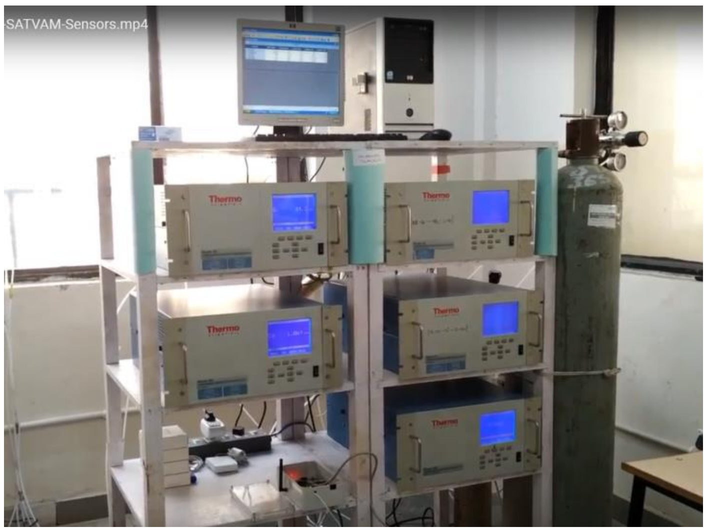
Lab Test Setup of SATVAM Sensor Nodes



Noise from Battery Pack to RE-MOTE



Affordable Air Quality Monitoring



- Can we replicate the quality of a high-end reference monitor (Costs about Rs.1 Crore) with a affordable air quality sensing device (~Rs.15k)?
- Allows us to deploy across the city, finer spatial scales

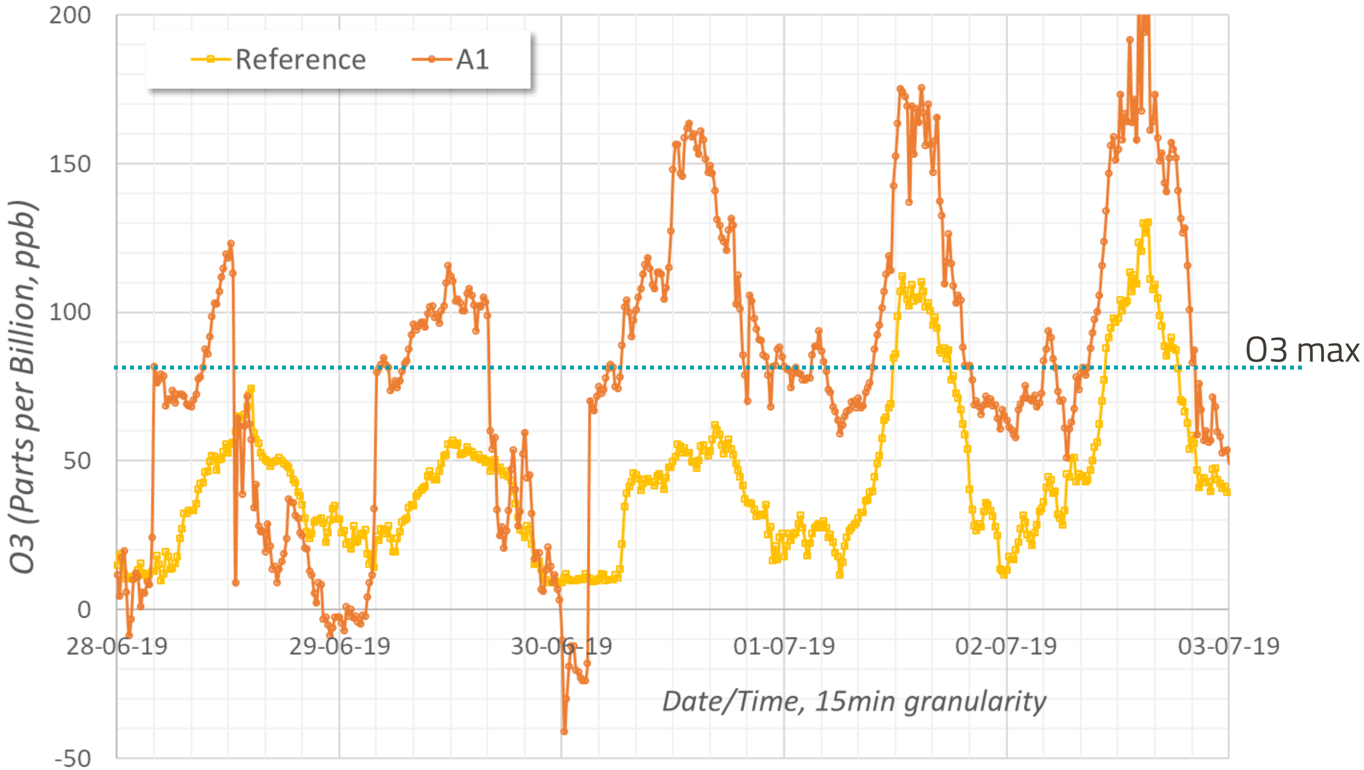
Low Quality Sensors, High Quality Data?

- Sensors produce two output measurements at either ends of a chamber, O1, O2
- These readings are used to model the actual quality value
  - $S = \mathbf{F}(O1, O2)$
- What is this function?
  - **We are "predicting" the air quality based on two "feature vectors"**
  - Sensor manufacturer offers 4 equations for different temperature ranges
  - Are these suitable for Indian conditions? Can we do better?

Analytics for Calibration

- **Simplicity** is important
  - Simple models rather than “deep” blackbox ones
  - Fewer variables
  - Explainable by scientists
- **Generalizability** is important
  - Usable across different sensors, at scale

Baseline Calibration for AQ Sensors is Poor

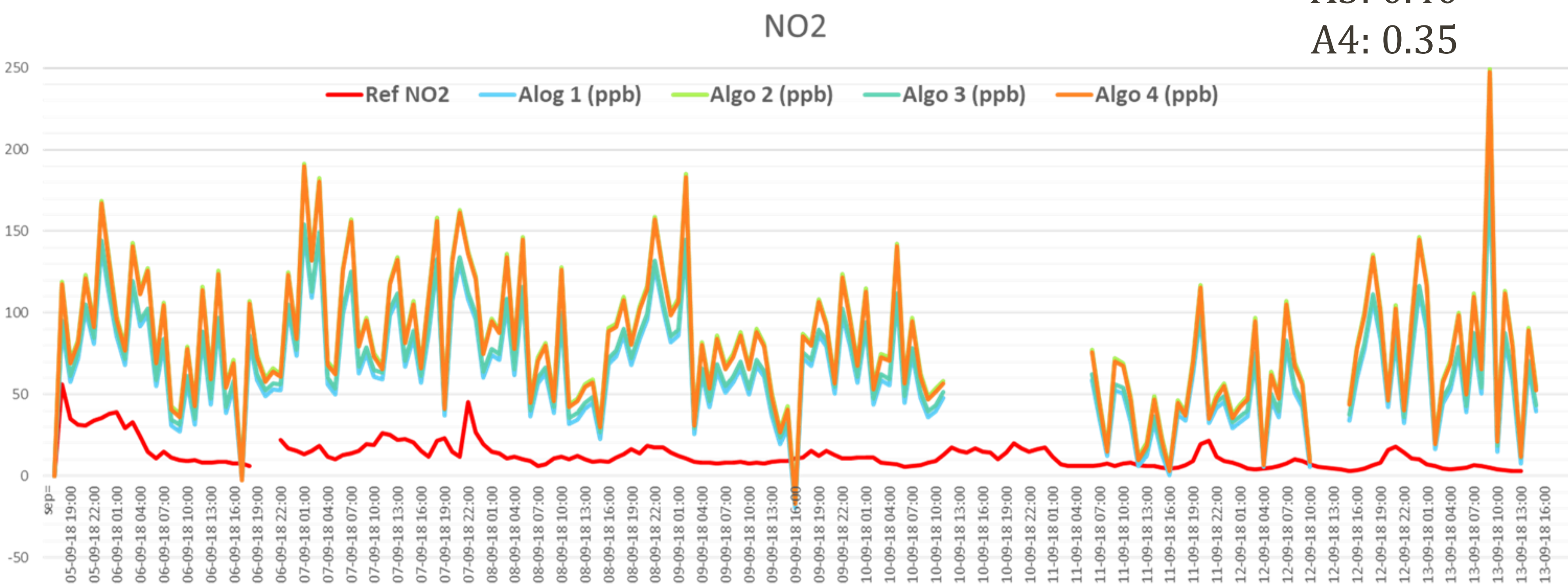


US EPA: Ozone below 85ppb is safe at sustained level

Baseline Calibration for AQ Sensors is Poor

R^2 Correlation

A1: 0.40
A2: 0.35
A3: 0.40
A4: 0.35

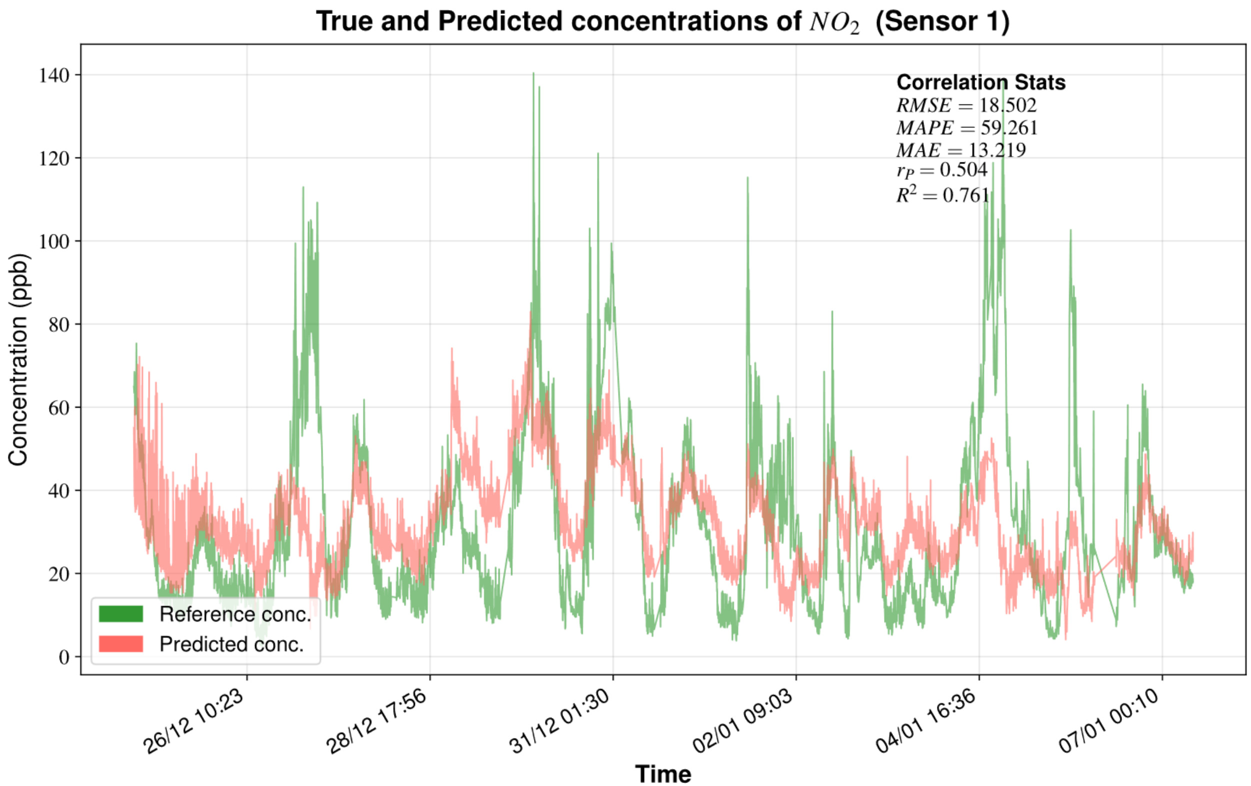


Results not so good! What may be the problem?

Is the data of a good quality?

- IoT has a lot of moving parts, complexity
- Are there reasons why the setup is not correct? Correct data is not being measured?
  - Hardware soldering, current leakage?
  - Problems with specific sensors?
  - Correlation between wireless transmission and sensing?
  - Location of the inlet pipes for reference monitor and our sensors?
- There is no point in trying to build a “better” model is your data is fundamentally flawed!
  - Problem is not new. LHC, other large instruments…

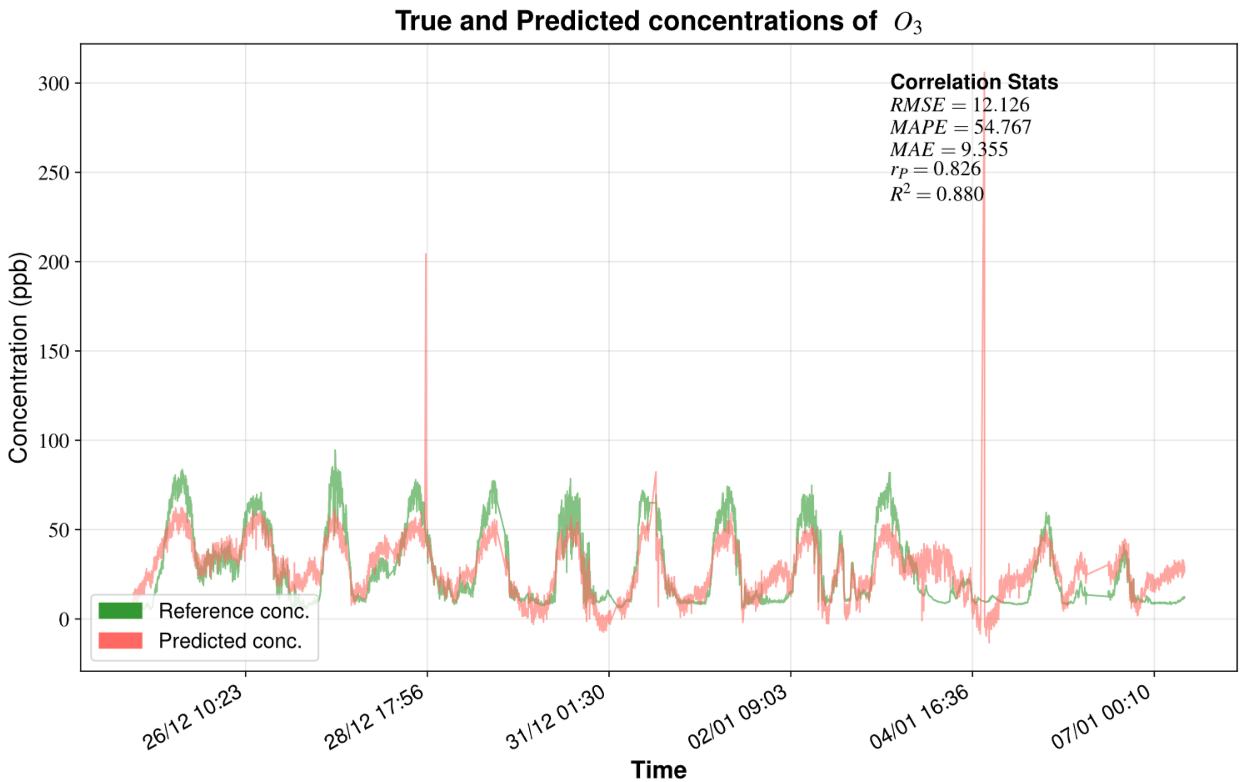
Using Simple Regression Models



R^2 Correlation

0.761

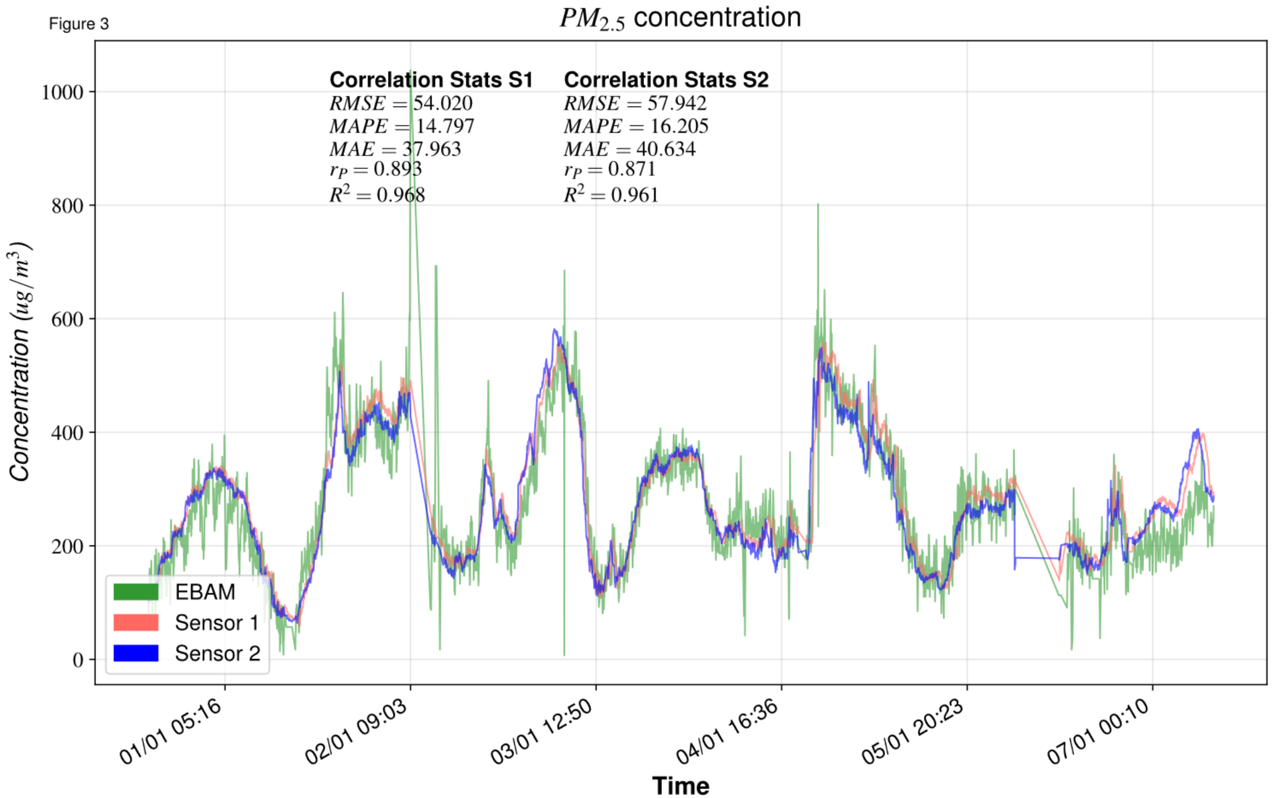
Using Simple Regression Models



R^2 Correlation

0.880

Using Simple Regression Models



R^2 Correlation

0.96

Regression Tree Model

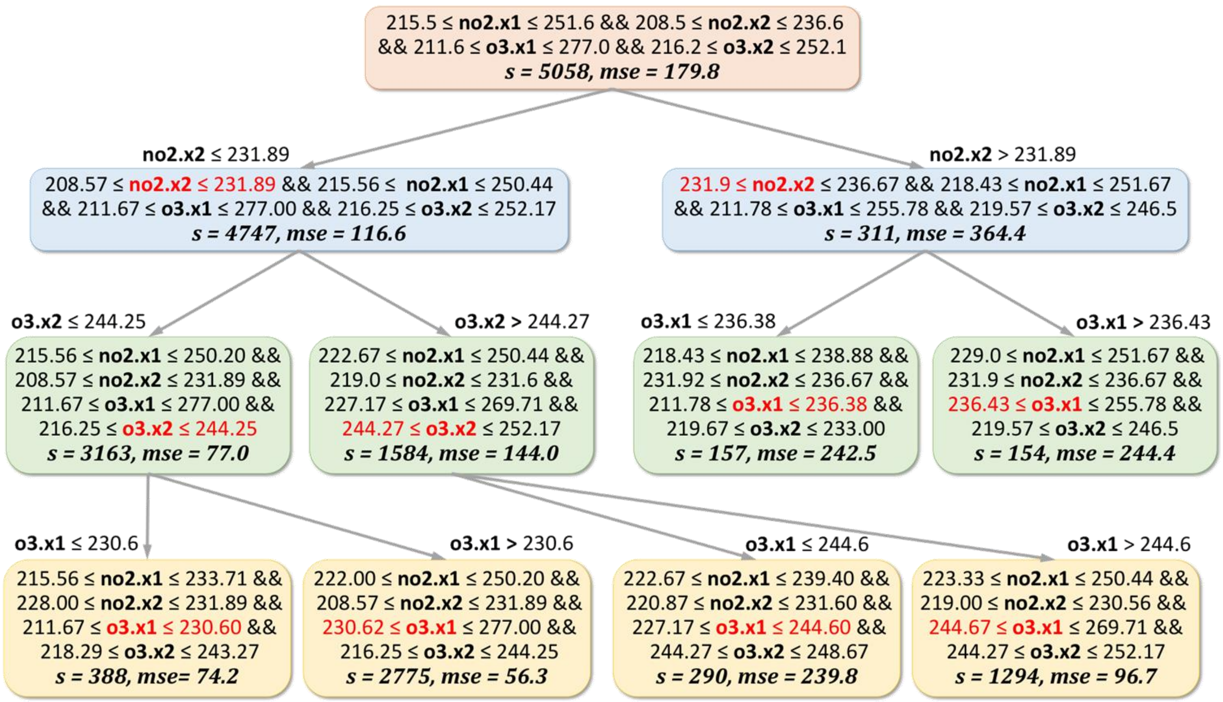
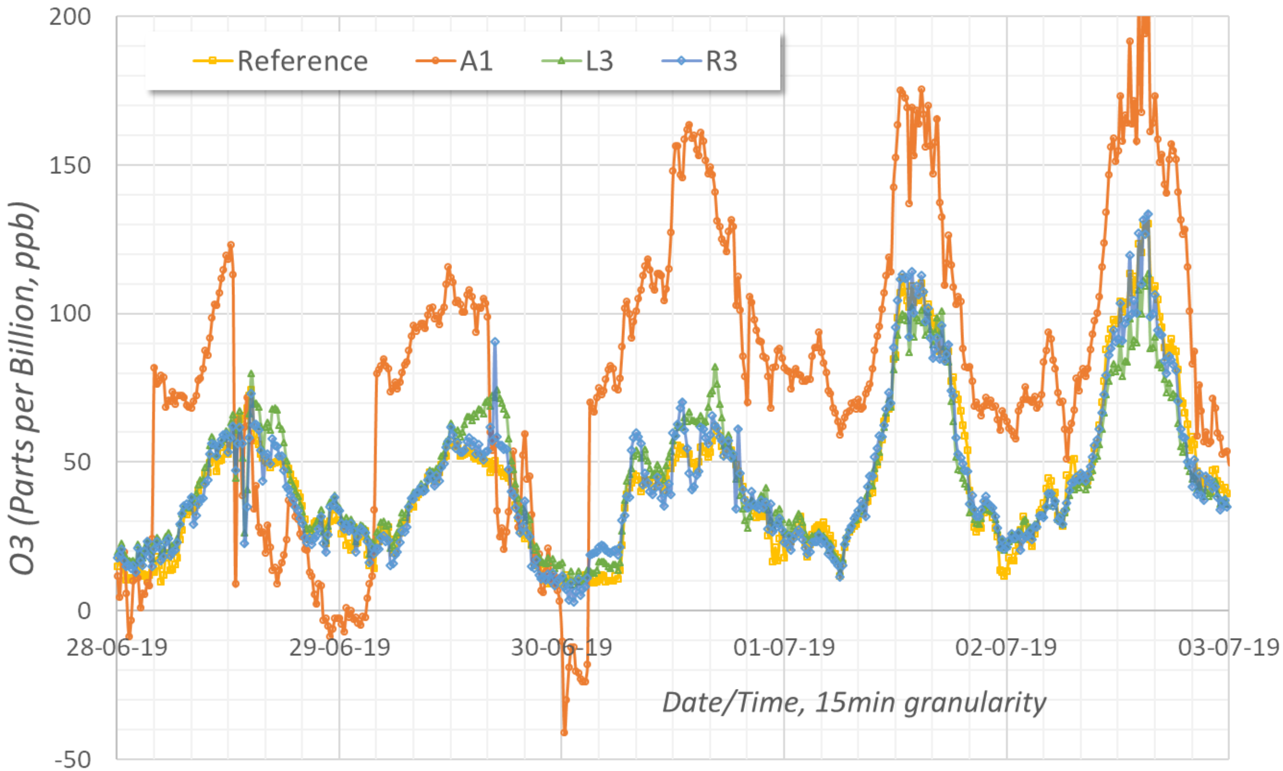


Figure 2: Regression tree of D=3, S=150, trained for NO_2

Proposed Models outperform Baseline



Smart Cities: Air Quality Monitoring USP

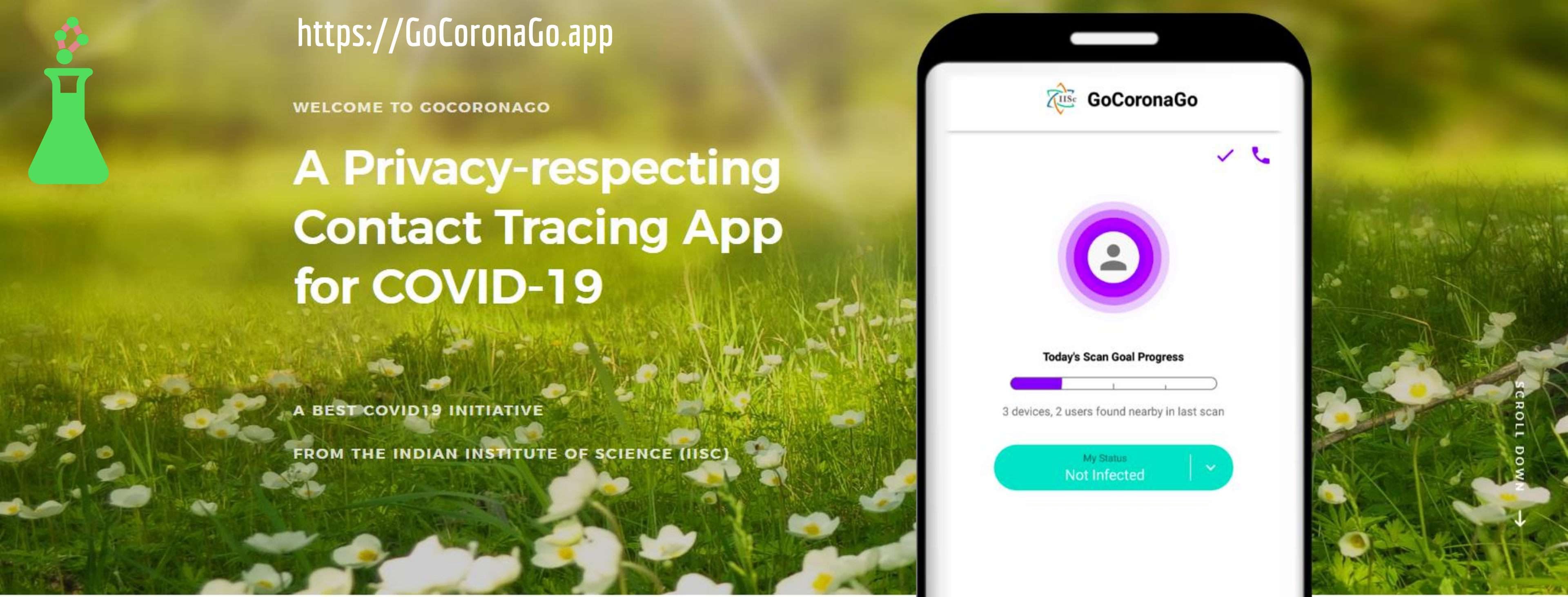
- Low-cost sensors
- Data analytics to address sensing errors
- Resilience of edge and WSN

- Yogesh Simmhan, Malati Hegde, Rajesh Zele, Sachchida N. Tripathi, Srijith Nair, Sumit K. Monga, Ravi Sahu, Kuldeep Dixit, Ronak Sutaria, Brijesh Mishra, Anamika Sharma and S. V. R. Anand, **SATVAM: Toward an IoT Cyber-Infrastructure for Low-Cost Urban Air Quality Monitoring** , *IEEE International Conference on eScience (eScience)* , 2019 , pp. 57-66
- Ravi Sahu, Ayush Nagal, Kuldeep Kumar Dixit, Harshavardhan Unnibhavi, Srikanth Mantravadi, Srijith Nair, Yogesh Simmhan, Brijesh Mishra, Rajesh Zele, Ronak Sutaria, Vidyanand Motiram Motghare, Purushottam Kar, and Sachchida Nand Tripathi **Robust statistical calibration and characterization of portable low-cost air quality monitoring sensors to quantify real-time O3 and NO2 concentrations in diverse environments**, *Atmospheric Measurement Techniques (AMT)*, 14, 37–52, 2021

Smart Health

COVID-19 Contact Tracing using GoCoronaGo

Acknowledgements: T. Rambha, IISc



- **Yogesh Simmhan (CDS) &**
- **Tarun Rambha (Civil and CiSTUP)**

GoCoronaGo: Privacy Respecting Contact Tracing for COVID-19 Management
J.IISc, 100(4), Springer
http://arxiv.org/abs/2009.04916

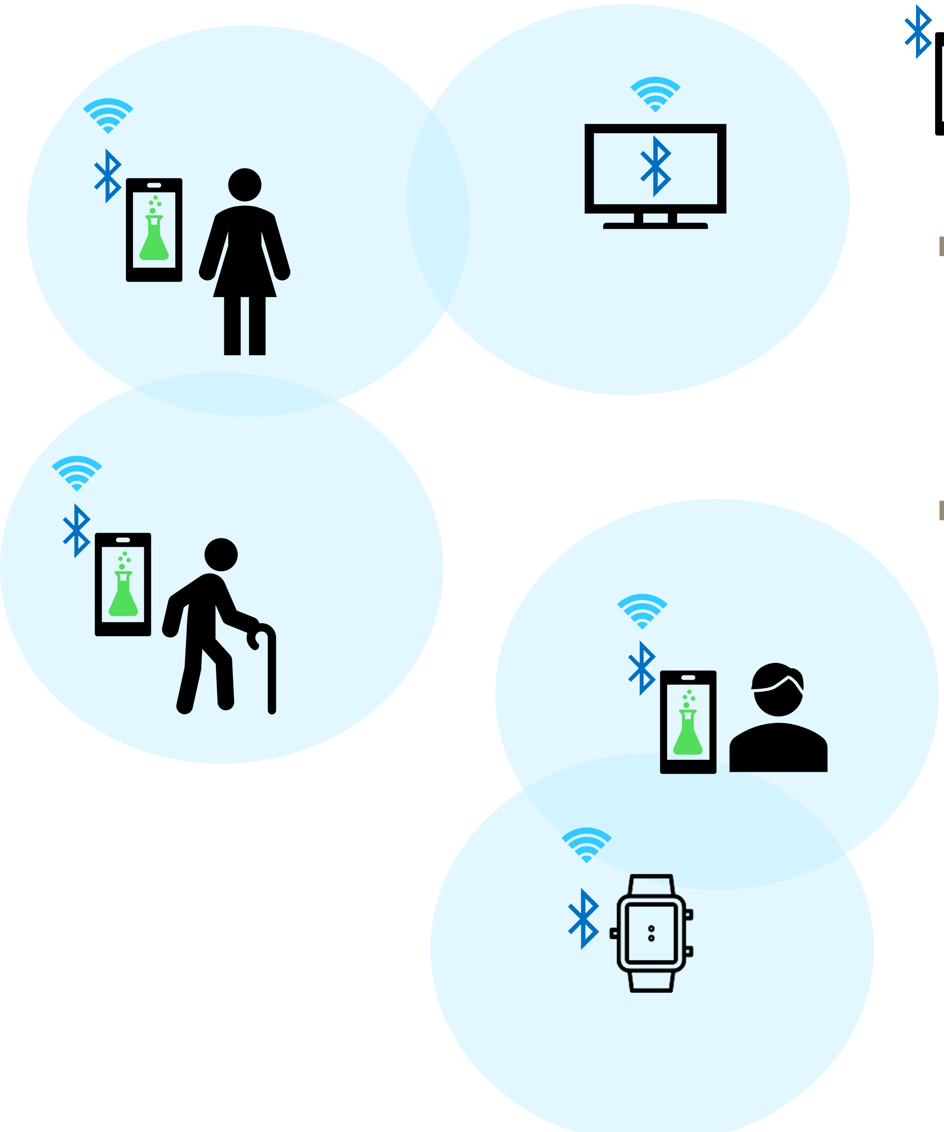
Supported by a DST NCPS RAKSHAK Grant

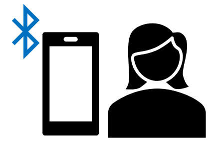
Welcome to the new Normal!

- Student health and safety is paramount at University campuses!
- Ensure health and safety of entire campus
  - *Can we be pro-active in managing COVID?*
  - *Can we follow best global practises and be a benchmark for other Institutions?*
  - *Can we have a pleasant and productive time on campus, while ensuring health and safety?*

COVID Management

BLE Digital Contact Tracing



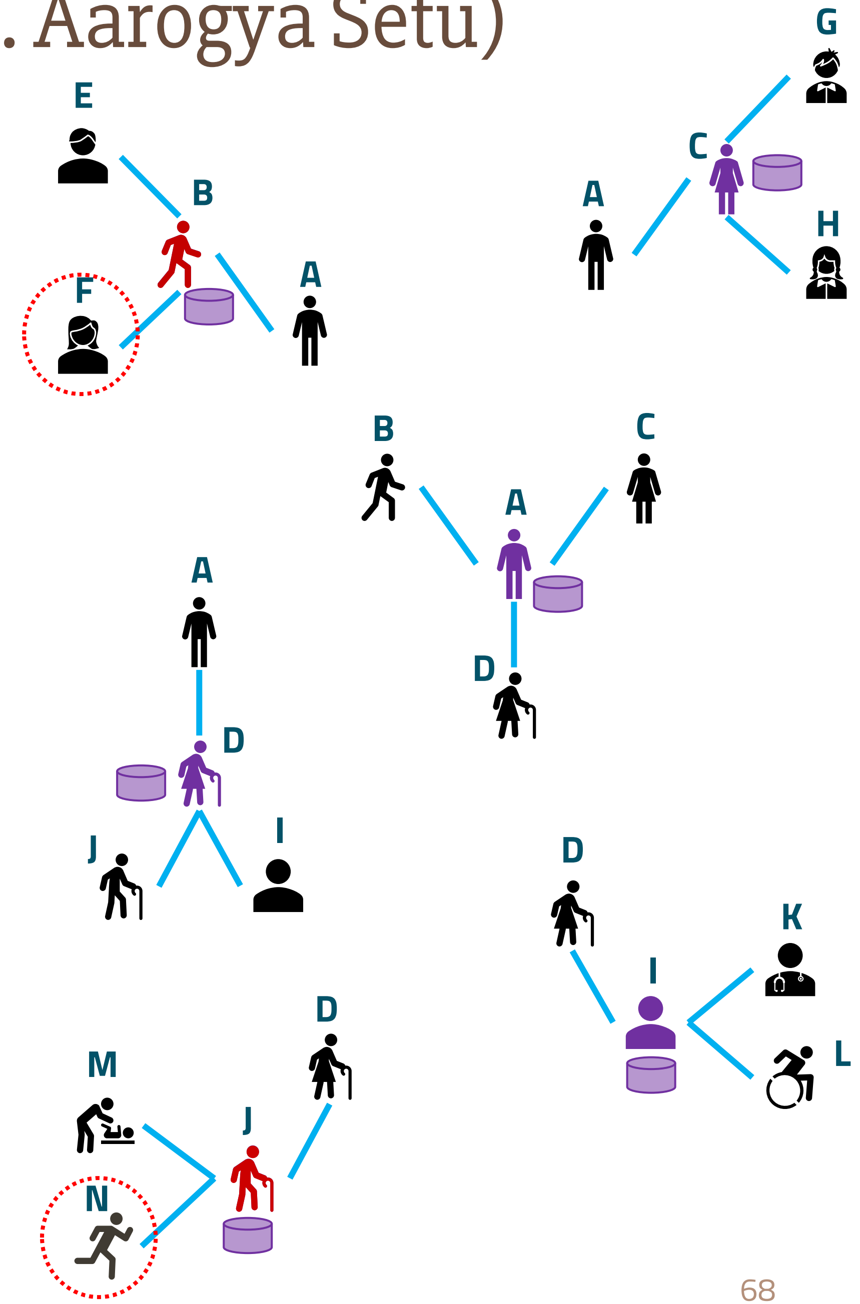


- Manual contact tracing is unreliable
  - Poor memory, unknown people

- Digital contact tracing
  - Apps scan for Bluetooth LE signals advertised by smartphone
  - Random ID of device is broadcast to ~15 feet
  - Others app devices scan for ID and record them
  - Happens regularly, say every 1 min
  - Record the *device ID*, *duration* of Contact, Bluetooth *signal strength*

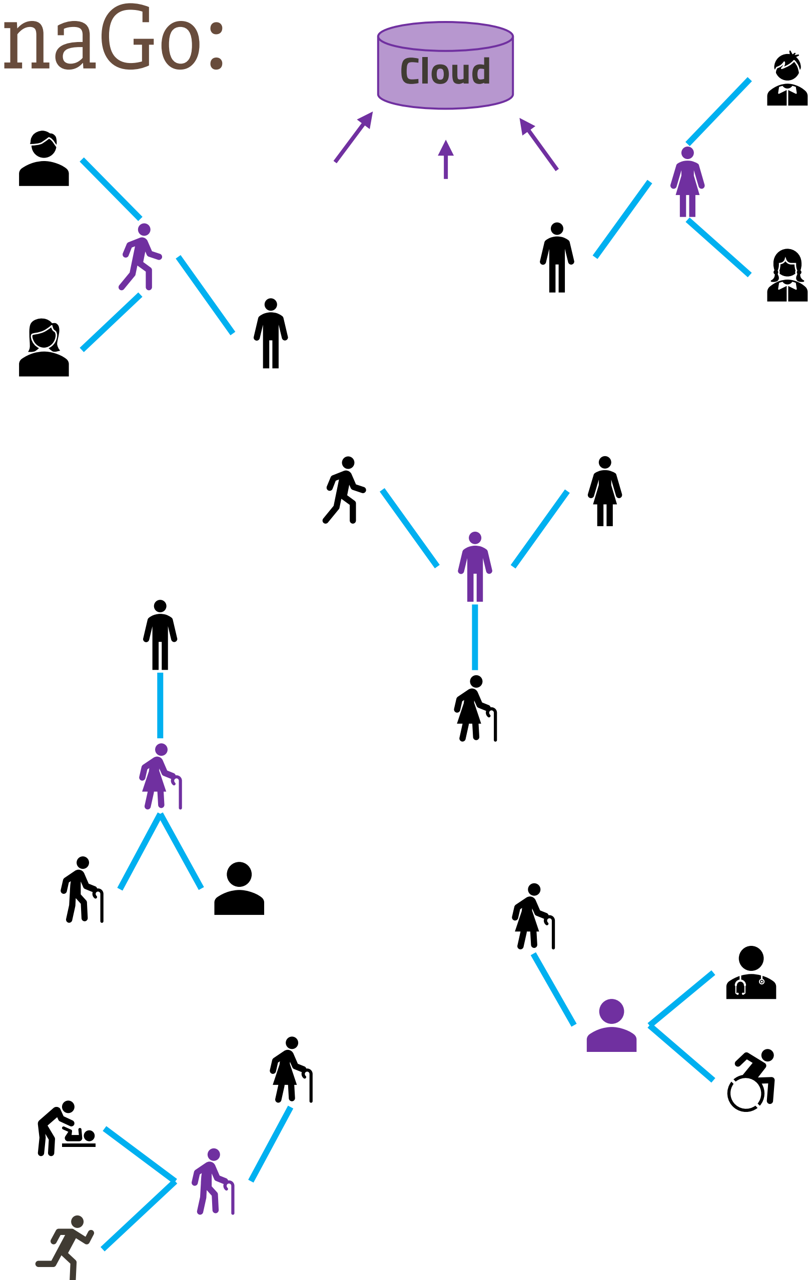
Classic Contact Tracing (e.g. Aarogya Setu)

- App records unique ID of proximate devices and the time in their *local database*
- COVID+ users share their unique ID on *government/public server*
- Apps download list of COVID+ IDs and check in their local database
  - If there is a match, they were a *primary contact* and they report themselves to the government
- Only primary contact tracing is possible automatically
- Individuals are informed after they come in contact – *Reactive*



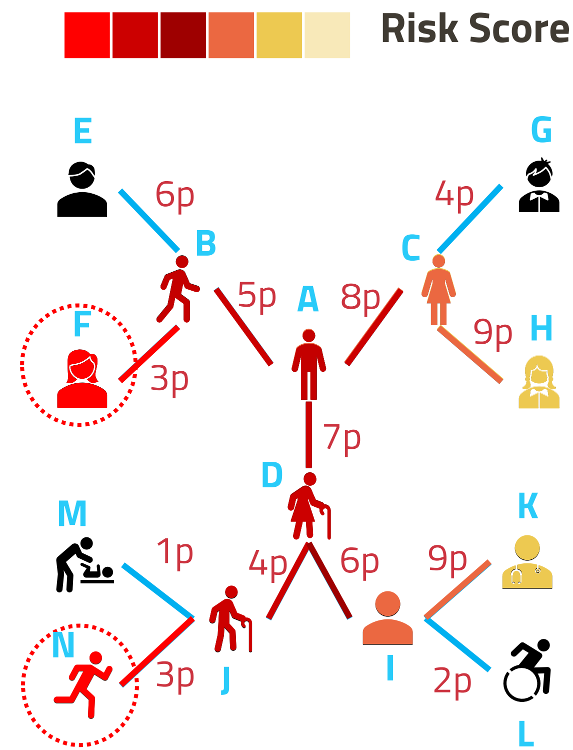
Contact Tracing in GoCoronaGo:
Beyond Primary Contacts

- GCG records the proximate devices and timestamps in a *backend cloud server*
- We link all this data to form a *global contact network*
  - ***At the Institutional Scale***
- We know *at what time and duration* they were close to each other
- This can be used for primary, secondary and tertiary contact tracing, *rapidly*
- It can identify high-risk individuals, before they get infected – *Proactively, in future*



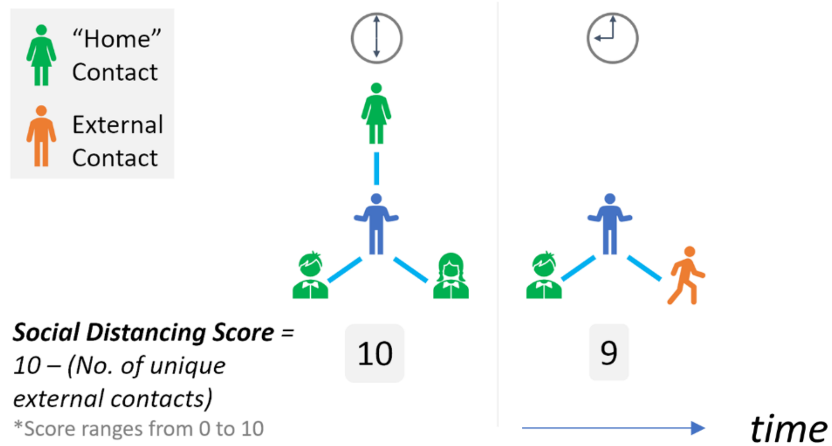
Contact Tracing in GoCoronaGo:
Beyond Primary Contacts

- GCG records the proximate devices and timestamps in a *backend cloud server*
- We link all this data to form a *global contact network*
  - ***At the Institutional Scale***
- We know *at what time and duration* they were close to each other
- This can be used for primary, secondary and tertiary contact tracing, *rapidly*
- It can identify high-risk individuals, before they get infected – *Proactively, <u>in future</u>*

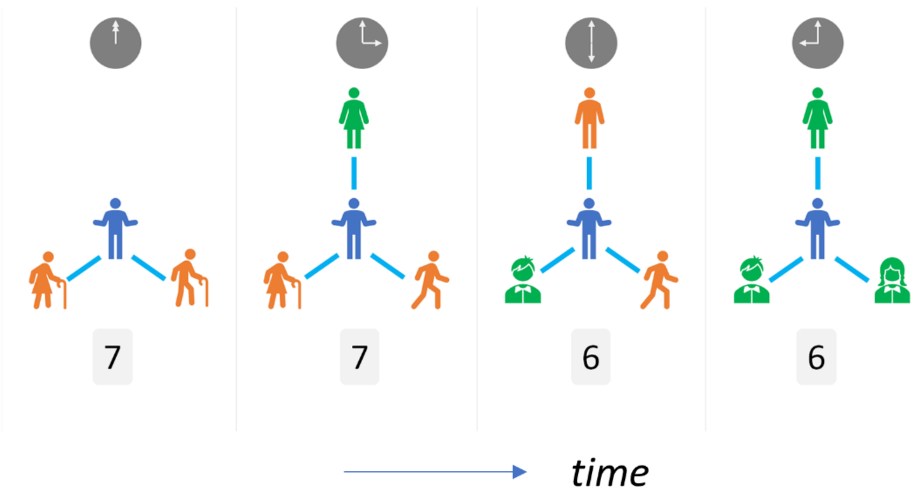


Analytics: Social Distancing

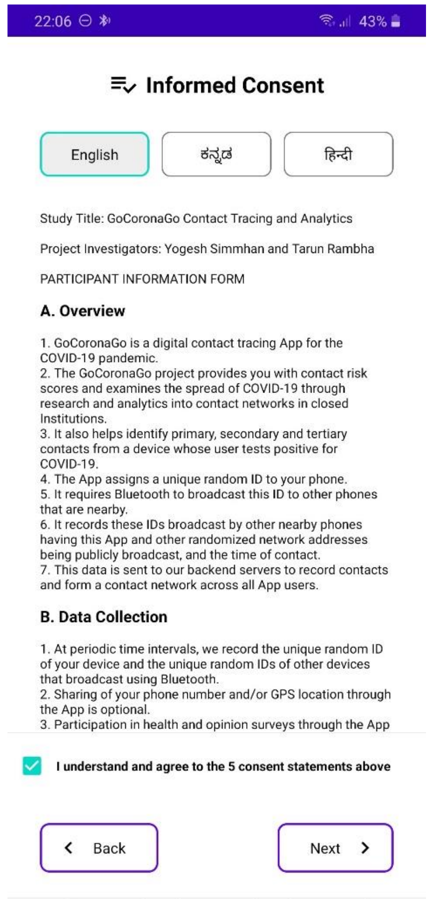
- Social Distancing scores computed and shared twice a day
- Account for "baseline" set of people (devices) they are normally with

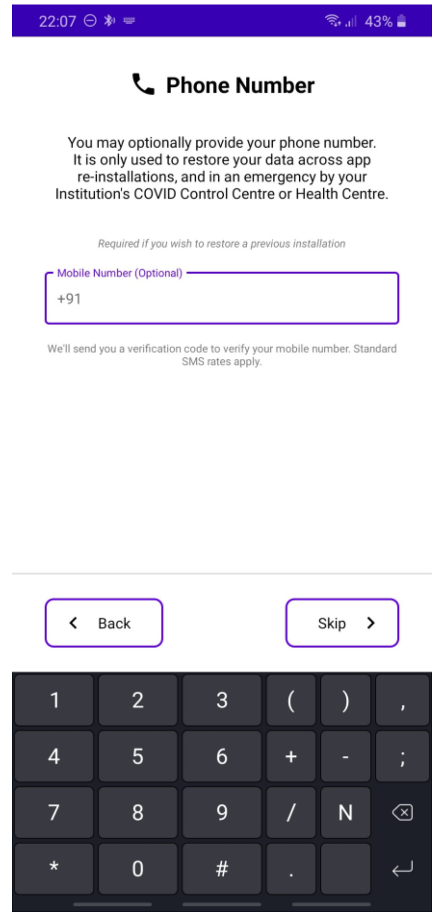


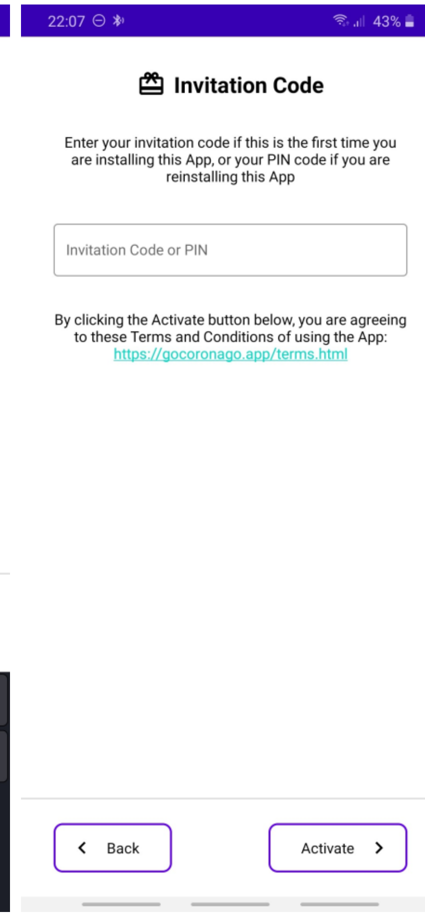
- Other network centrality and epidemiology models also possible with enough data, *in future*



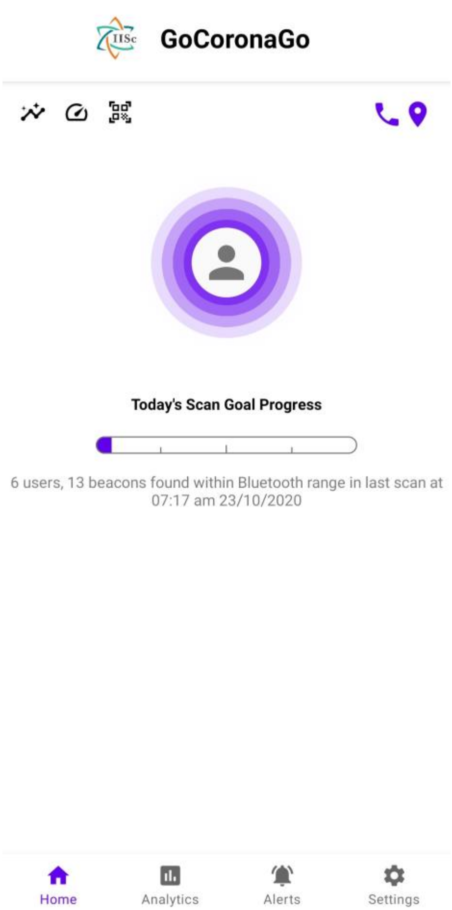
GoCoronaGo UI/UX: Installation

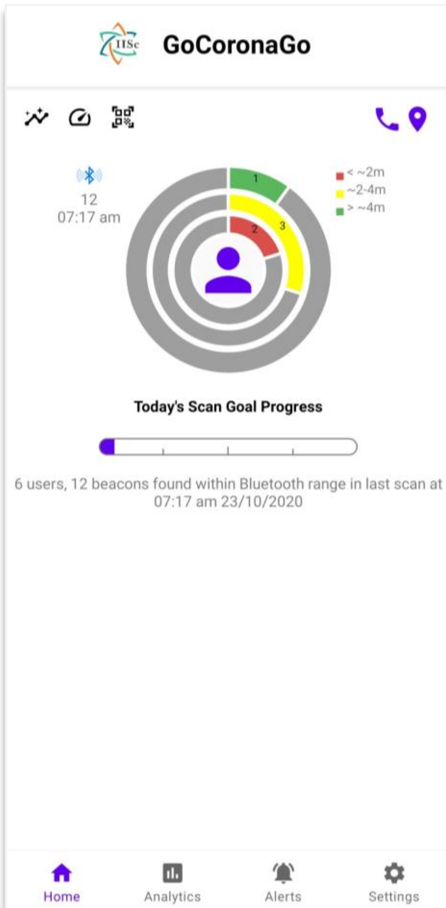


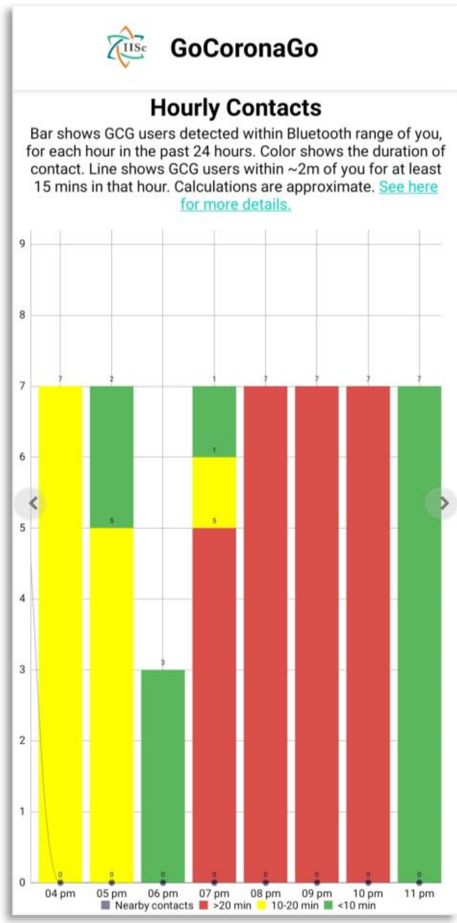


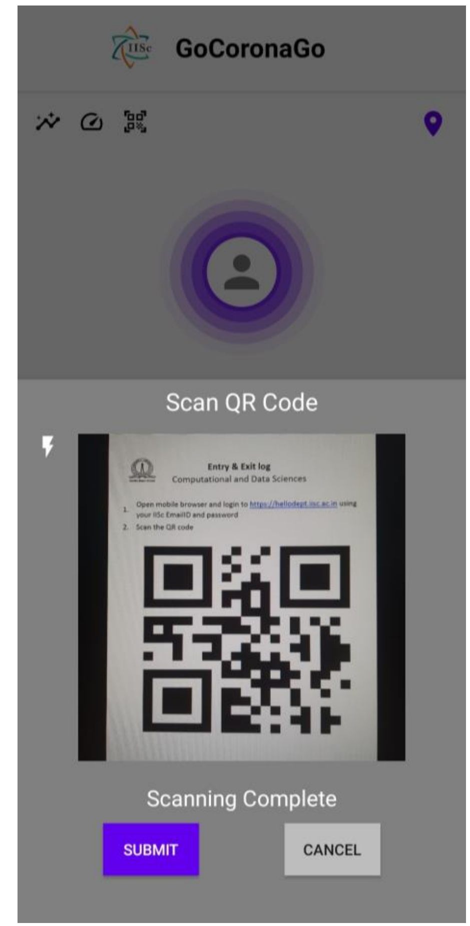


GoCoronaGo UI/UX: Main Screen



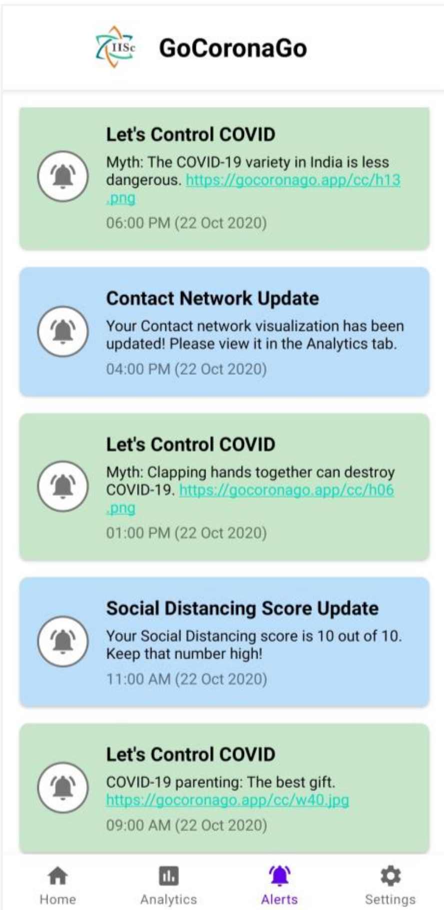


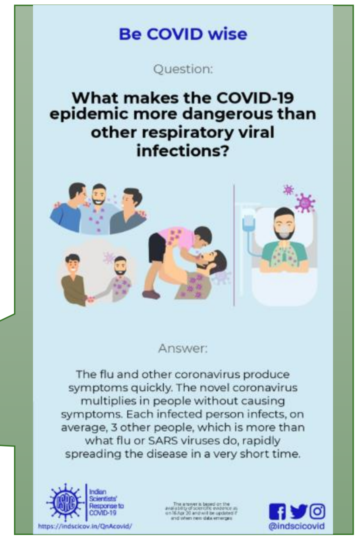


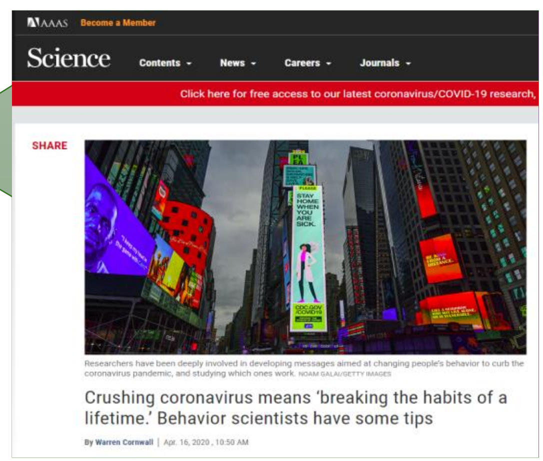


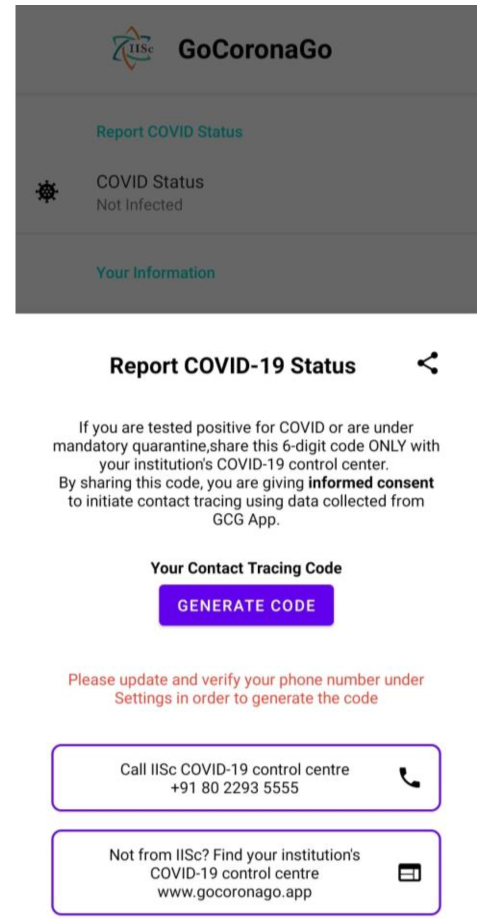
Thanks to ISRC and IISc/TIFR for outreach images!

GoCoronaGo UI/UX: Alerts & Contact Tracing



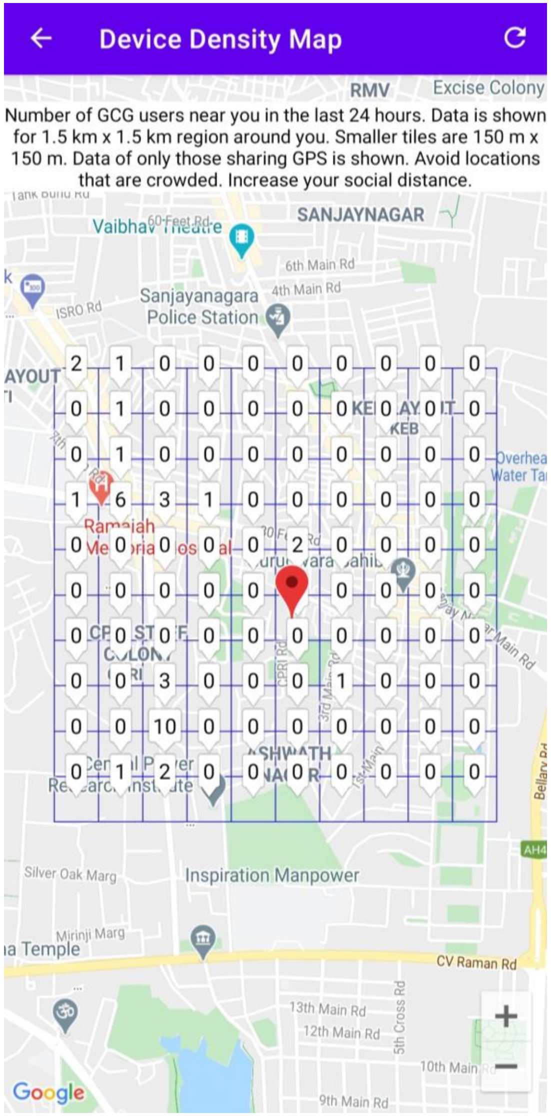


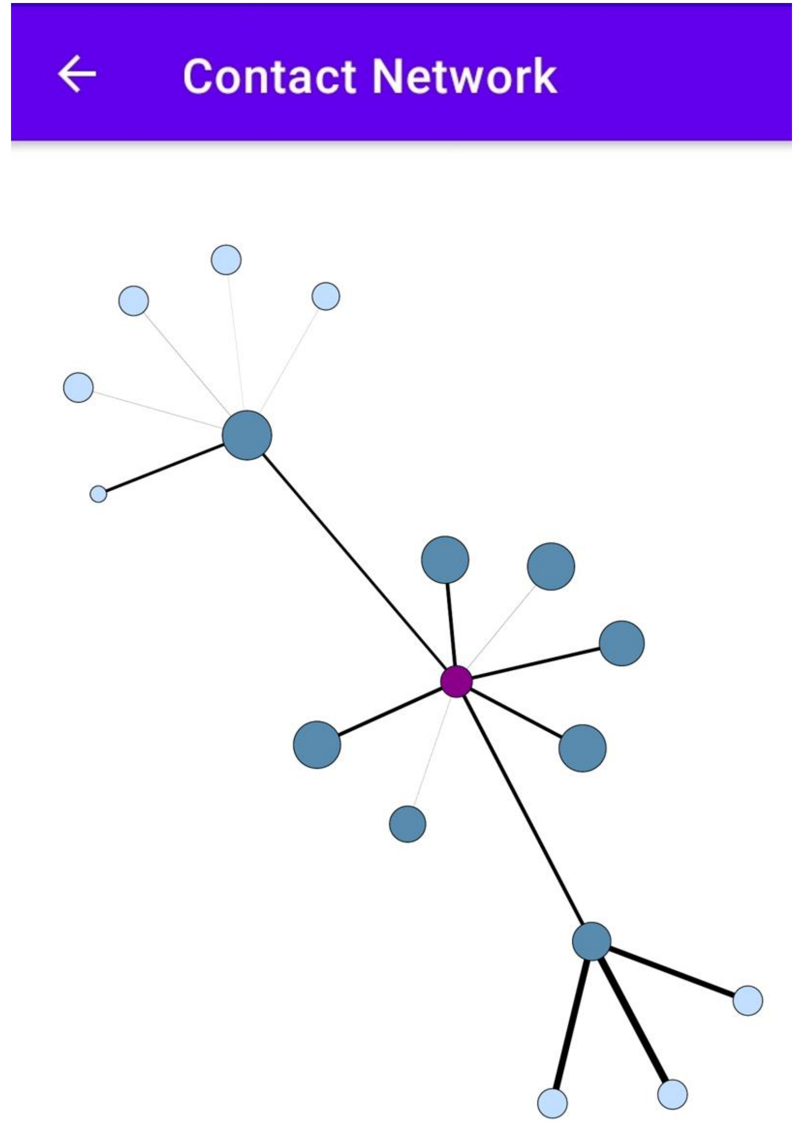




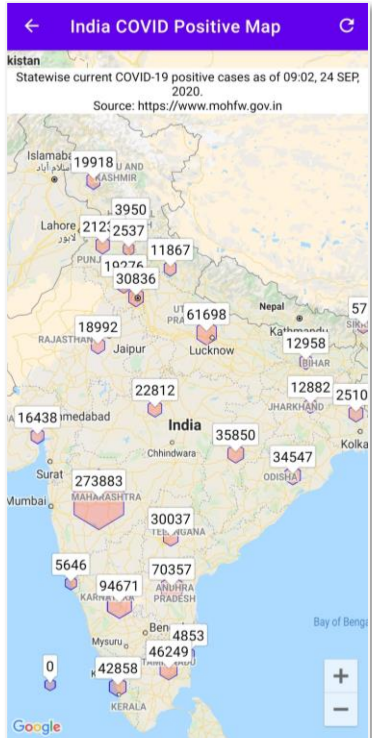
Thanks to ISRC and IISc/TIFR for outreach images!

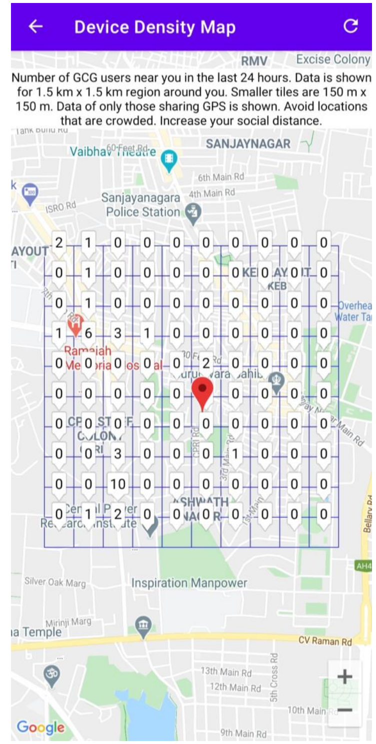
GoCoronaGo UI/UX: Custom Analytics

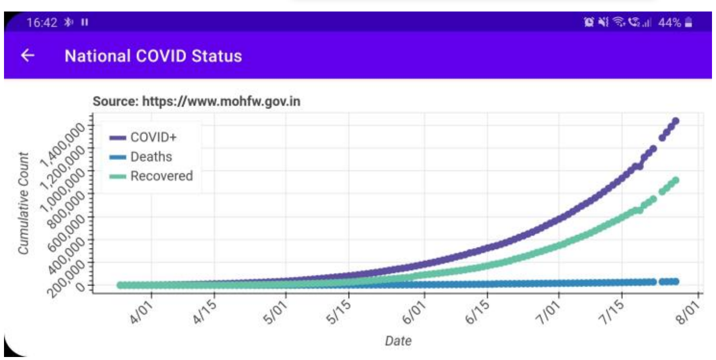


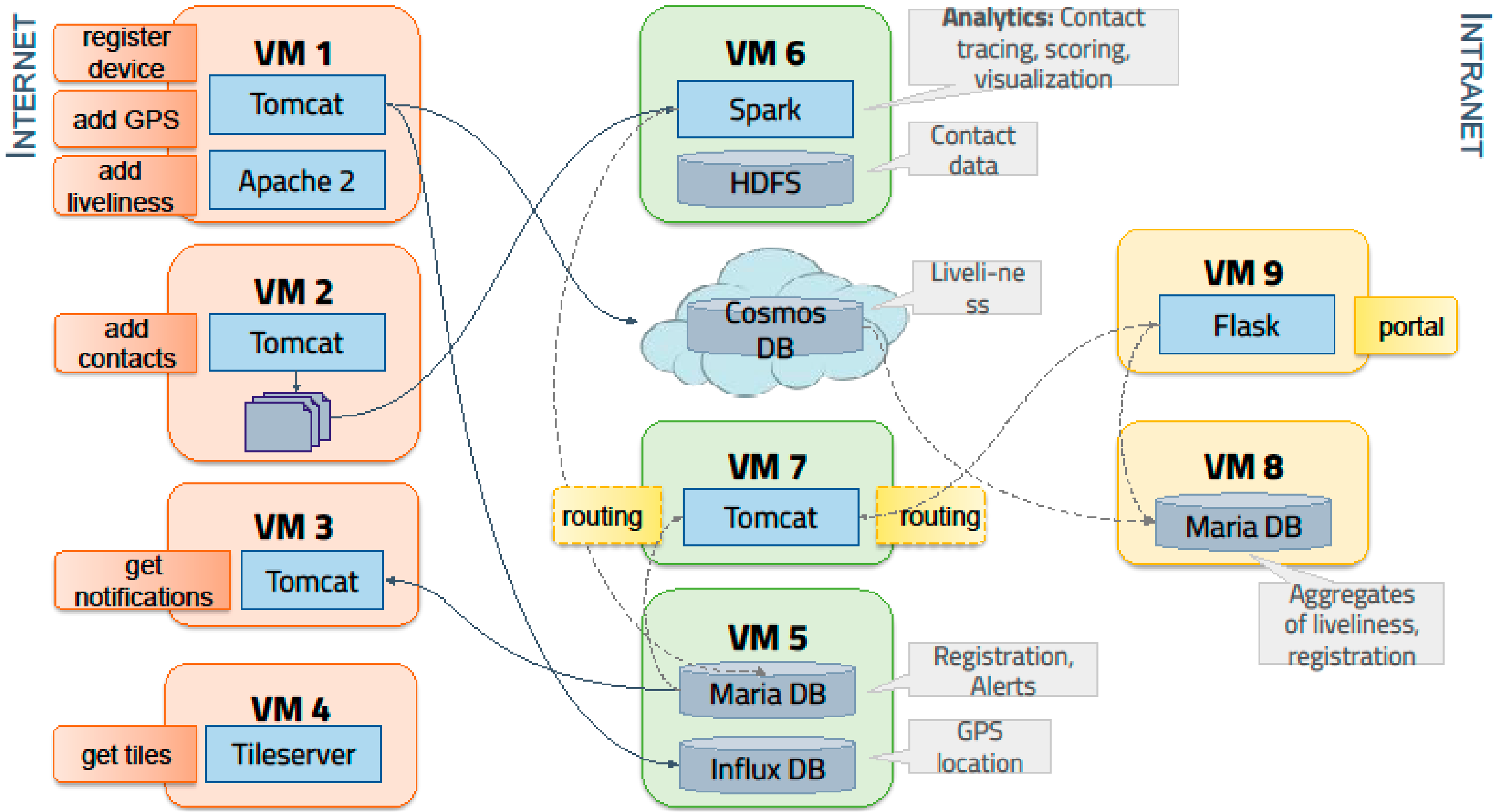


GoCoronaGo UI/UX: Visualizations








INTERNET
INTRANET
register device
add GPS
add liveliness
VM 1
Tomcat
Apache 2
add contacts
VM 2
Tomcat
get notifications
VM 3
Tomcat
get tiles
VM 4
Tileserver
VM 6
Spark
HDFS
Analytics: Contact tracing, scoring, visualization
Contact data
Cosmos DB
Liveli-ne ss
VM 7
routing
Tomcat
routing
VM 5
Maria DB
Influx DB
Registration, Alerts
GPS location
VM 9
Flask
portal
VM 8
Maria DB
Aggregates of liveliness, registration

Network Analytics

- Temporal breadth first search from COVID positive individual
  - Primary, secondary and tertiary contact tracing
- Centrality measures to identify higher-risk individuals
  - Take better precautions, test frequently
  - Feedback to users to reinforce positive behavior
- SIR models for understanding spread of pandemic
  - Identify individuals who should be vaccinated for most impact

Deployments

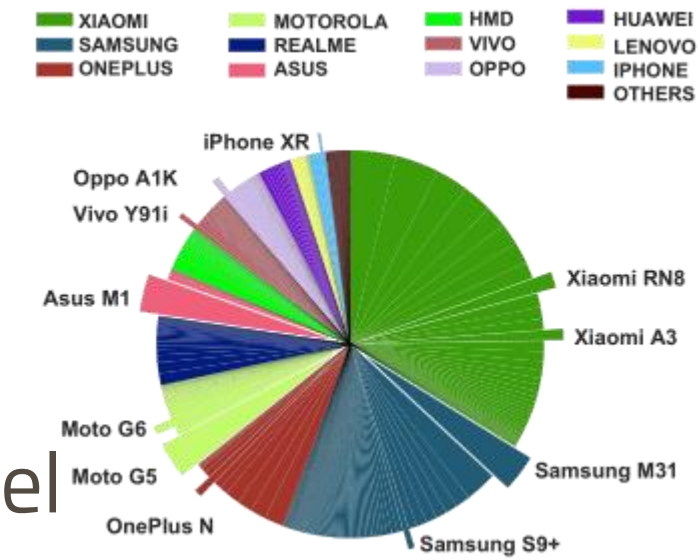
- IISc
- IIT-Jodhpur
- IIIT-H
- Total ~3000 installations
- Wider rollout as part of DST Campus RAKSHAK solution

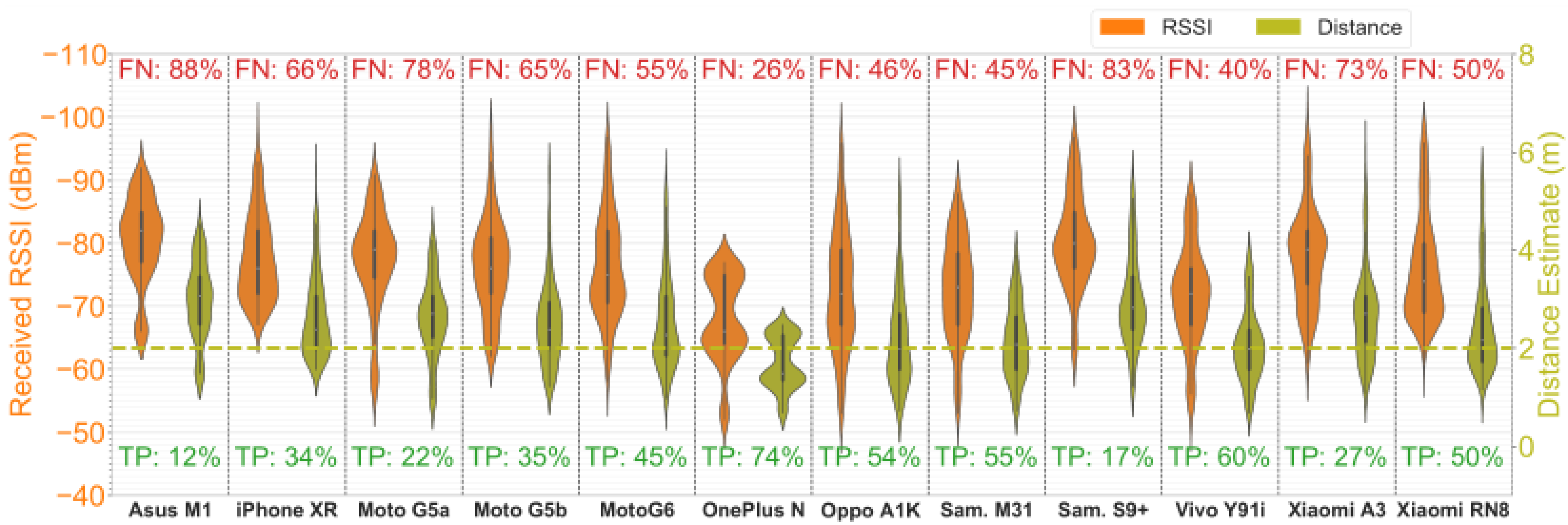
Bluetooth Ranging Study

- Estimate the distance between smartphones by observing the BLE signal strengths
  - Establish if there was a "contact" or not
  - Log-distance path loss model, $d = 10^{\frac{R_d - R_1}{-10\gamma}}$
    - » Not very accurate
  - Need a classification: Near (<1.5m), Medium (1.5-2.5m), Far (>2.5m)

Empirical Study

- Diverse phone models
- Baseline
  - Uncalibrated path loss model



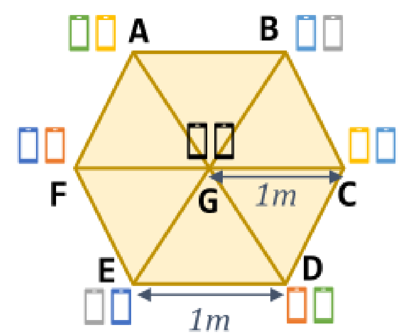


Data Collection

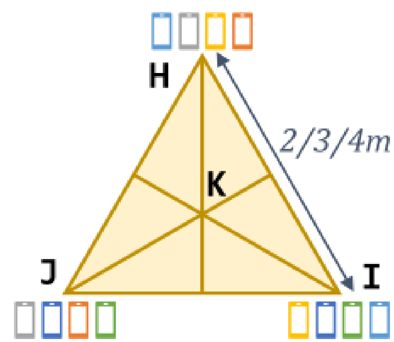
- 13 popular phone models
- Different distances
- Walled vs. Non-walled

Source \ Sink	Asus M1	iPhone XR	Moto G5a	Moto G5b	MotoG6	OnePlus N	Oppo A1K	Samsung M31	Samsung S9+	Vivo Y91i	Xiaomi A3	Xiaomi RN8
Asus M1	0	10	14	24	33	126	52	28	53	32	44	60
iPhone XR	63	0	336	203	134	63	136	141	267	198	134	66
Moto G5a	39	48	0	168	178	64	244	293	126	163	300	321
Moto G5b	46	32	150	0	64	64	180	127	315	257	129	205
MotoG6	49	20	160	60	0	65	185	194	317	201	203	136
OnePlus N	94	9	43	56	60	0	62	60	65	60	64	64
Oppo A1K	43	21	226	186	193	65	0	195	251	311	65	67
Samsung M31	39	21	258	125	184	61	198	0	193	158	133	202
Samsung S9+	42	41	99	251	261	63	232	164	0	53	241	247
Vivo Y91i	48	32	169	290	238	64	321	154	67	0	327	260
Xiaomi A3	50	20	228	119	180	65	56	124	298	266	0	66
Xiaomi RN8	60	9	277	172	116	64	65	190	305	195	67	0

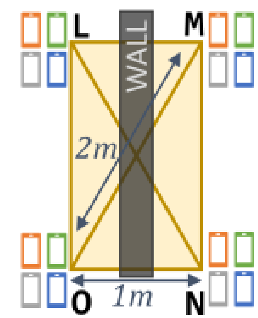
(a) # of data points collected at 2 m for all phone-pairs



(a) Hexagon, 1 m ***Yogesh: Show 2m arrow



(b) Triangle, 2 m, 3 m, 4 m



(c) Wall, 1 m, 2 m

Parametric Models for Distance Estimation

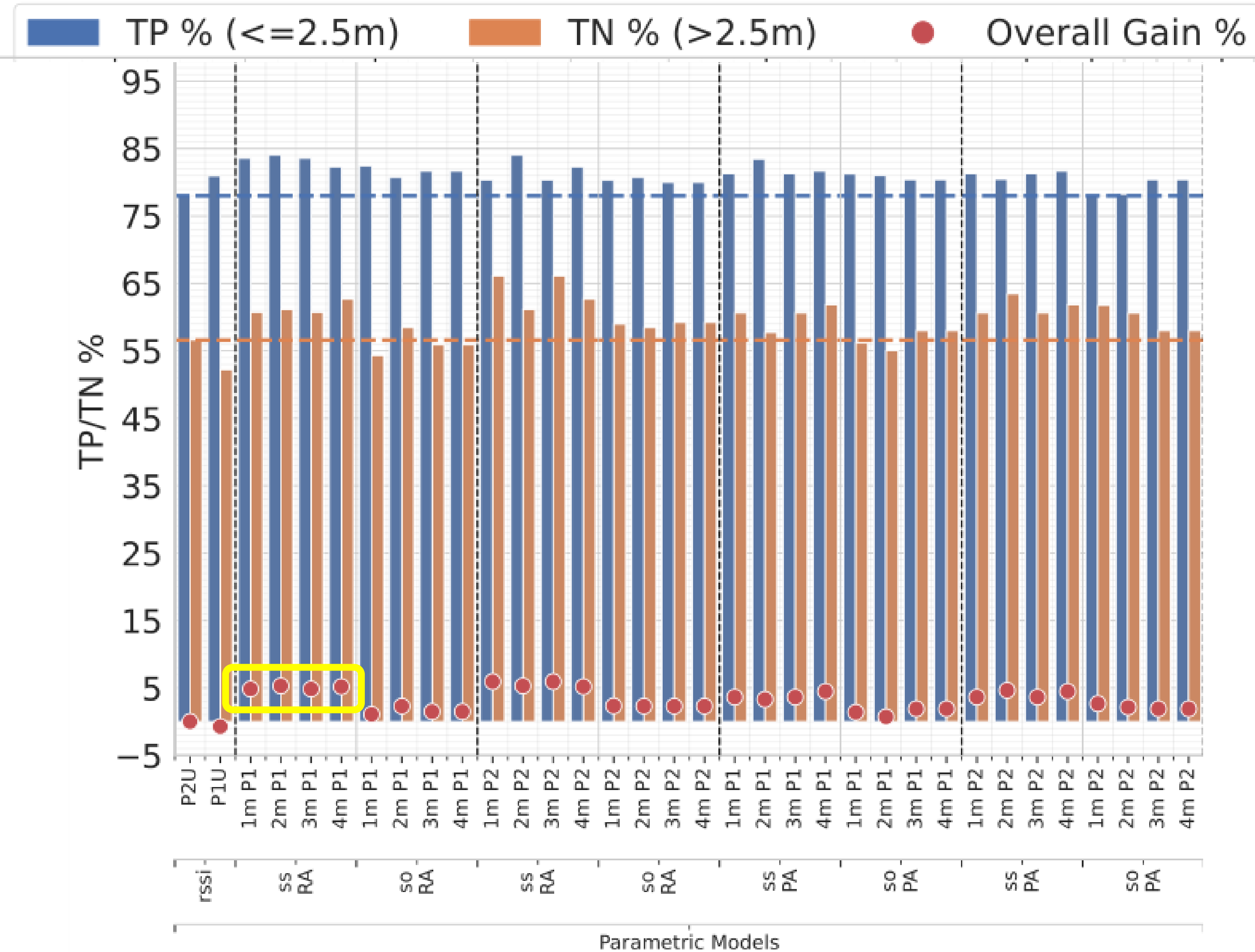
- Baseline path-loss model with curve fit to data we have collected $d = 10^{\frac{R_d - A}{-10 \times B}} + C$

- Calibrated models
  - Add correction factor to RSSI of specific phone models
  - Simple averaging, Power averaging

Non-parametric Models for Distance Classification

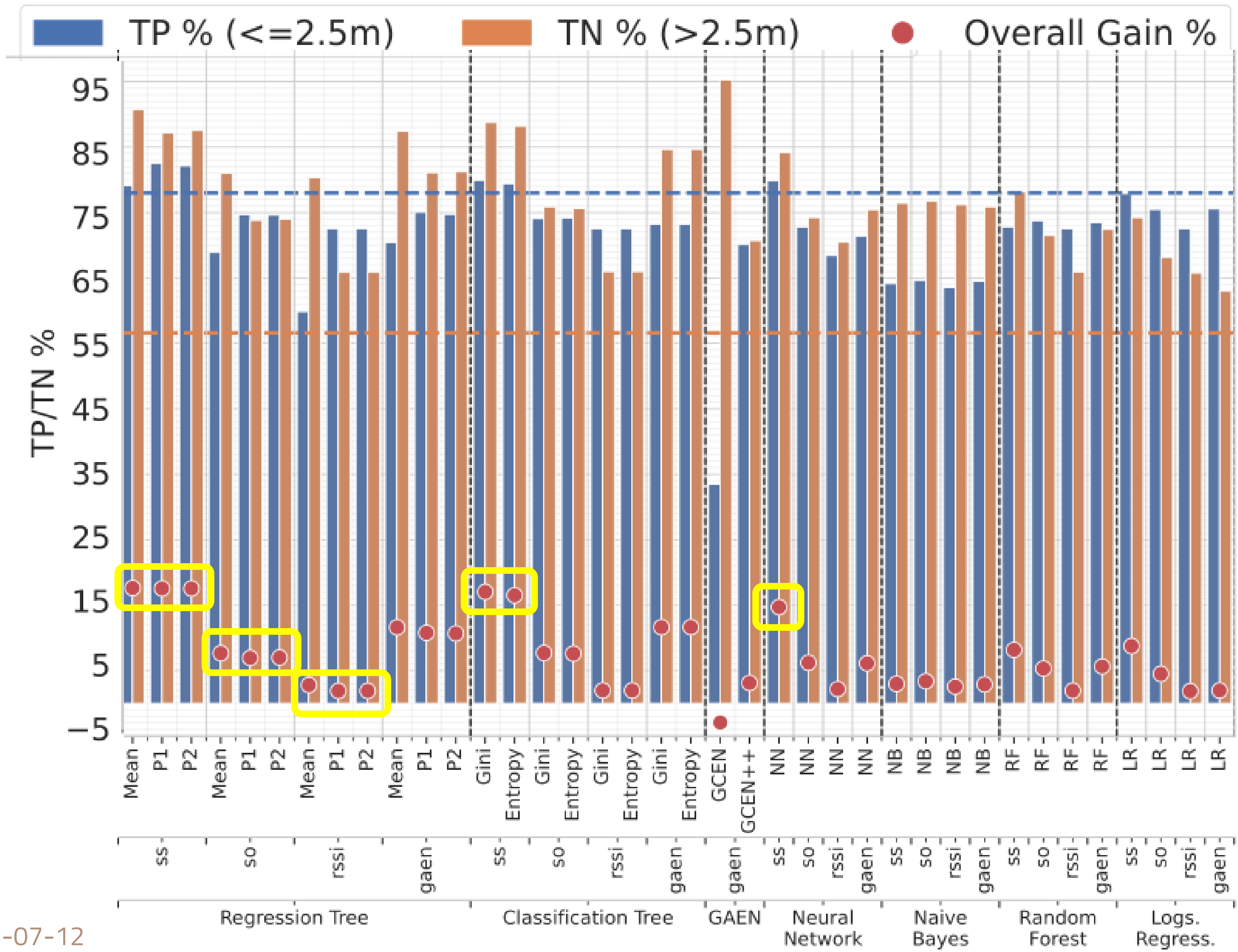
- Regression tree (Distance estimate)
  - RSSI, source phone model, sink phone model
  - Calibration factors from GCEN: txp and rssi_c
- Classification tree
  - Binary and ternary classification
  - Gini, Entropy for minimization
- Neural N/W Model
  - 3-layer, 100 neuron, ReLU
- Random Forest, Logistic Regression, Naïve Bayes, etc.

Results



TP % (<=2.5m)
TN % (>2.5m)
Overall Gain %
TP/TN %
95
85
75
65
55
45
35
25
15
5
−5
P2U
P1U
1m P1
2m P1
3m P1
4m P1
1m P2
2m P2
3m P2
4m P2
rssi
ss RA
so RA
ss PA
so PA
Parametric Models

Results


TP % (<=2.5m)
TN % (>2.5m)
Overall Gain %
TP/TN %
95
85
75
65
55
45
35
25
15
5
−5
Mean P1 P2 Mean P1 P2 Mean P1 P2 Mean P1 P2
Gini Entropy Gini Entropy Gini Entropy Gini Entropy
GCEN GCEN++
NN NN NN NN
NB NB NB NB
RF RF RF RF
LR LR LR LR
ss so rssi gaen ss so rssi gaen gaen ss so rssi gaen ss so rssi gaen ss so rssi gaen ss so rssi gaen
Regression Tree
Classification Tree
GAEN
Neural Network
Naive Bayes
Random Forest
Logs. Regress.

Smart Health: Contact Tracing USP

- Mobile app, rapid development, field deployment
- BLE ranging study
- Simulation study

- Yogesh Simmhan, Tarun Rambha, Aakash Khochare, Shriram Ramesh, Animesh Baranawal, John Varghese George, Rahul Atul Bhope, Amrita Namtirtha, Amritha Sundararajan, Sharath Suresh Bhargav, Nihar Thakkar and Raj Kiran, **GoCoronaGo: Privacy Respecting Contact Tracing for COVID-19 Management** , *Journal of the Indian Institute of Science*, Vol. 100, 2020, doi:10.1007/s41745-020-00201-5

Smart Mobility

Drones

Acknowledgements: S. Das, U. Missouri

UAVs in Smart Cities

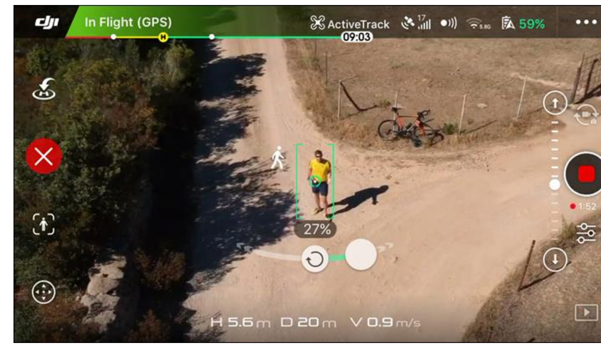
- Commercial UAVs are increasing deployed
  - Crowd Monitoring[1]
  - Air Quality Monitoring[2]
  - Package Delivery[3]
- UAVs typically carry a rich suit of
  - Sensors - GPS, Cameras, LiDAR
  - Compute - Pi-class or GPU accelerated
  - Networking - 5G/LTE connectivity
- Commercial UAV solutions coupled with rapid advances in Computer Vision enable several interesting applications

[1]https://www.financialexpress.com/defence/fight-against-coronavirus-drone-management-system-for-policing-and-crowd-control/1913384/
[2] https://balkaninsight.com/2021/02/18/polands-public-tries-to-clear-the-air/
[3] https://www.verizon.com/about/news/future-delivery-its-air

UAV Applications

Credits: Geospatialworld

Credits: China Daily



Credits: DJI

Construction Surveys[1]

- Batch processing application
- Processing can be done offline

Traffic Monitoring[2]

- Near realtime application
- Processing onboard is preferred

Person Tracking[3]

- Realtime application
- Onboard processing essential

[1] George, Shilpa, et al. "Towards drone-sourced live video analytics for the construction industry." *Proceedings of the 20th International Workshop on Mobile Computing Systems and Applications*. 2019.
[2] Kanistras, Konstantinos, et al. "A survey of unmanned aerial vehicles (UAVs) for traffic monitoring." *2013 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2013.
[3] Rohan, Ali, Mohammed Rabah, and Sung-Ho Kim. "Convolutional neural network-based real-time object detection and tracking for parrot AR drone 2." *IEEE access* 7 (2019): 69575-69584.

UAVs as an IoT System

- Sensing
  - On-board cameras, LIDAR, GPS, etc.
- Communication
  - To base-station (WiFi)
  - Through cellular
- Computation/Analytics
  - Basic navigation
  - Vision-based algorithms, routing & navigation
  - Collaboration between drones

Drone Fleet Operations

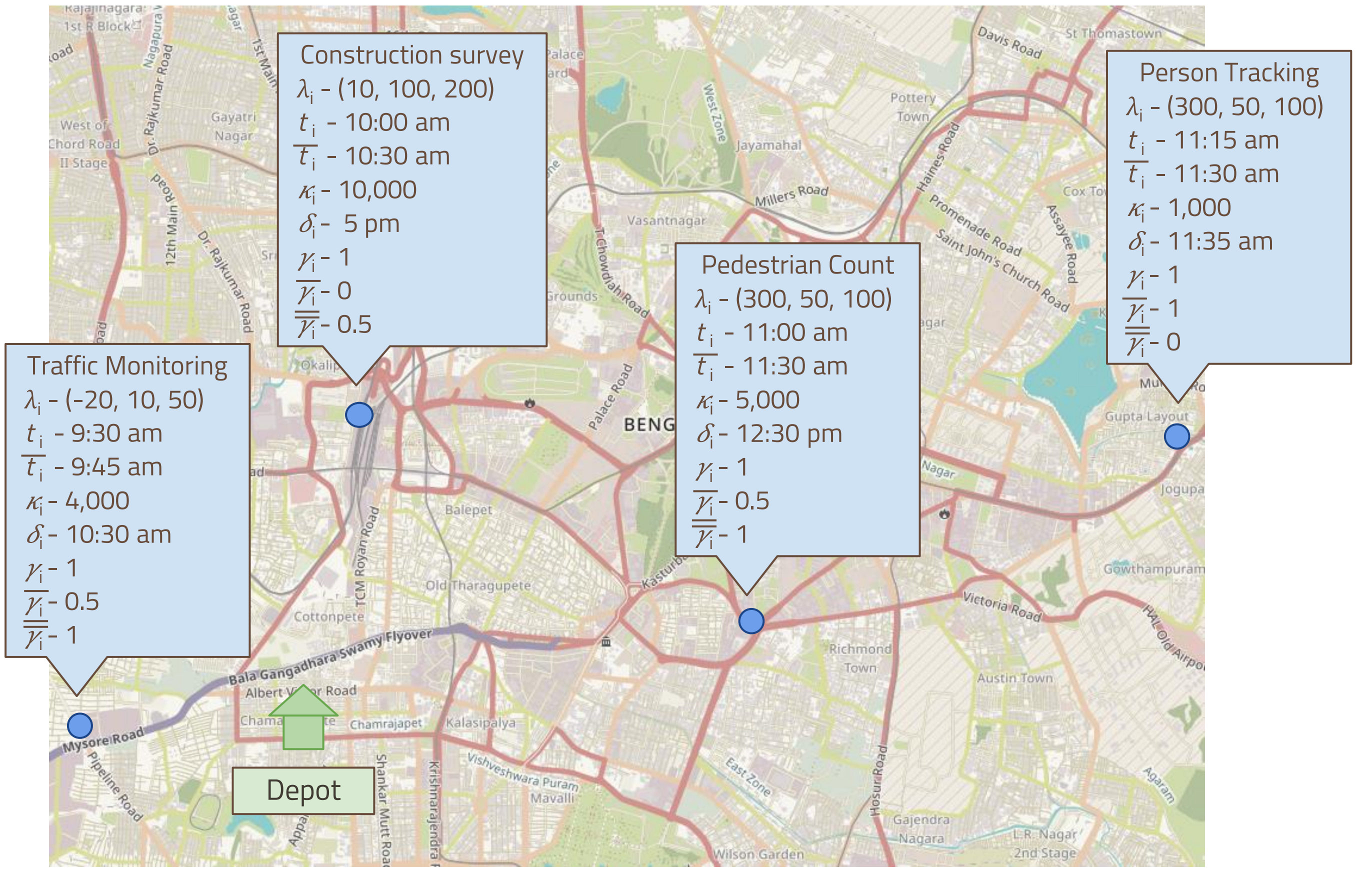
- Homogeneous fleet, Single depot for the entire fleet
- Each drone is equipped with
  - Camera
  - On-board persistent storage
  - Edge-class compute unit
- Customers submit requests for activities across a city
  - Visiting a waypoint
  - Performing observations during a time period
  - Optionally, performing analytics over the observation
- Co-scheduling of Routing, Sensing and On-board analytics

- Requests are accepted only at the start of a time window
  - All the requests are known for a time window
- Attempt to maximize the utility from requests served

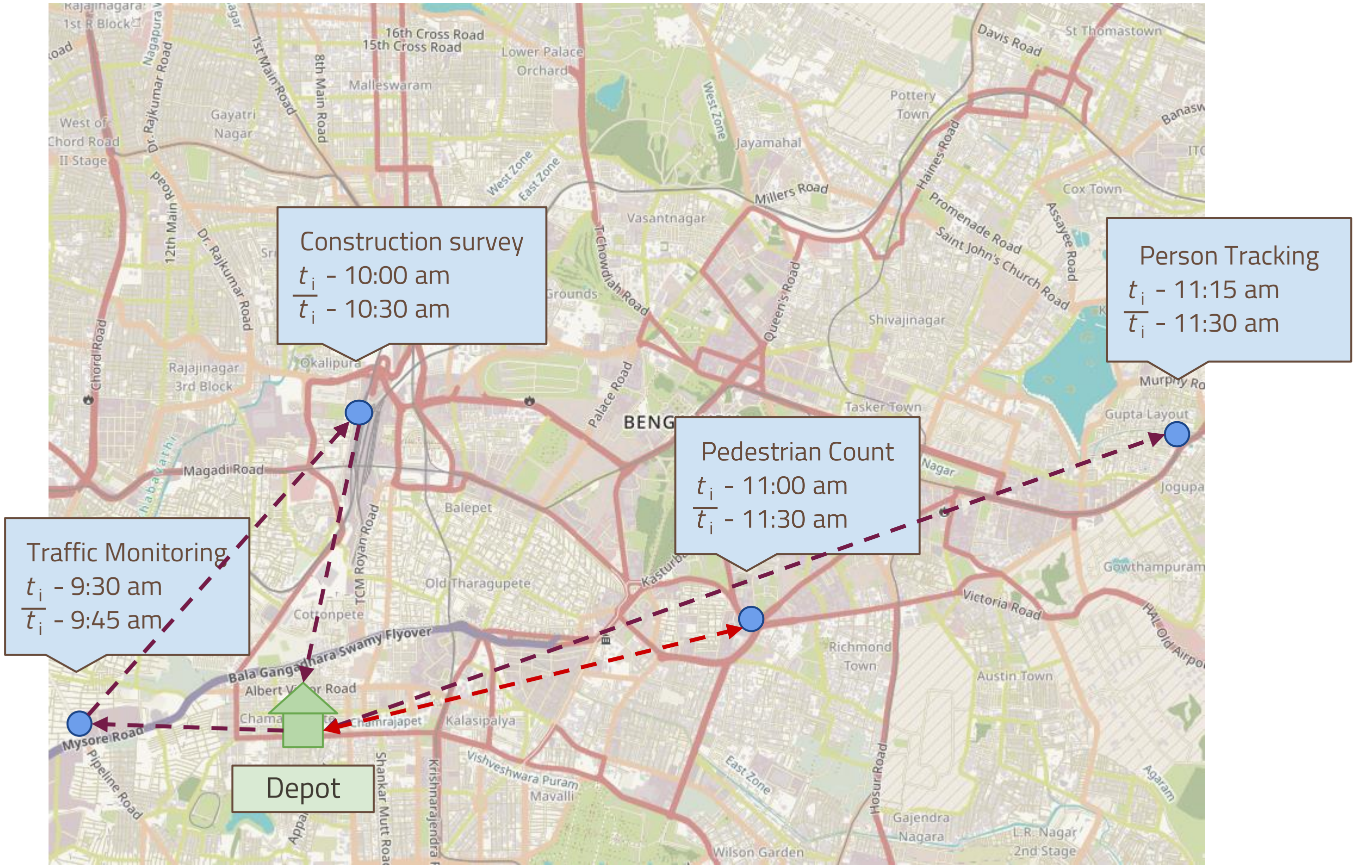
Mission Scheduling Problem (MSP)

Given a UAV depot in a city with a fleet of captive drones, and a set of observe and compute activities to be performed at locations in the city, each within a given time window and with associated utilities, the goal is to co-schedule the drones onto mission routes and the compute onto the drones, within the energy and compute constraints of the drones, such that the total utility achieved is maximized.

Example

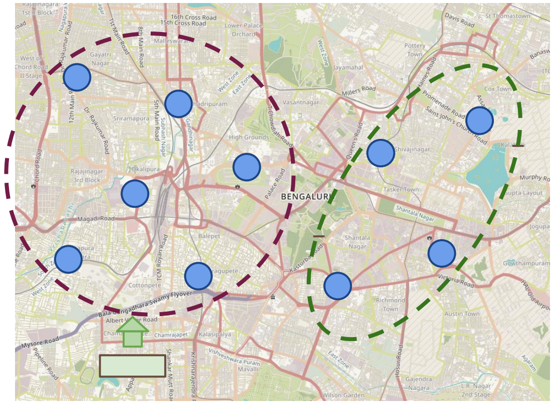

Construction survey
λ_i - (10, 100, 200)
t_i - 10:00 am
$\bar{t}_i$ - 10:30 am
κ_i - 10,000
δ_i - 5 pm
γ_i - 1
$\bar{\gamma}_i$ - 0
$\bar{\bar{\gamma}}_i$ - 0.5
Person Tracking
λ_i - (300, 50, 100)
t_i - 11:15 am
$\bar{t}_i$ - 11:30 am
κ_i - 1,000
δ_i - 11:35 am
γ_i - 1
$\bar{\gamma}_i$ - 1
$\bar{\bar{\gamma}}_i$ - 0
Pedestrian Count
λ_i - (300, 50, 100)
t_i - 11:00 am
$\bar{t}_i$ - 11:30 am
κ_i - 5,000
δ_i - 12:30 pm
γ_i - 1
$\bar{\gamma}_i$ - 0.5
$\bar{\bar{\gamma}}_i$ - 1
Traffic Monitoring
λ_i - (-20, 10, 50)
t_i - 9:30 am
$\bar{t}_i$ - 9:45 am
κ_i - 4,000
δ_i - 10:30 am
γ_i - 1
$\bar{\gamma}_i$ - 0.5
$\bar{\bar{\gamma}}_i$ - 1
Depot

Example Route



Job Scheduling Centric (JSC) algorithm

- Intuitively
  - Cluster activities -> minimizes travel time
  - Maximize utility within cluster
- Two phases of JSC
  - Clustering phase
  - Scheduling phase

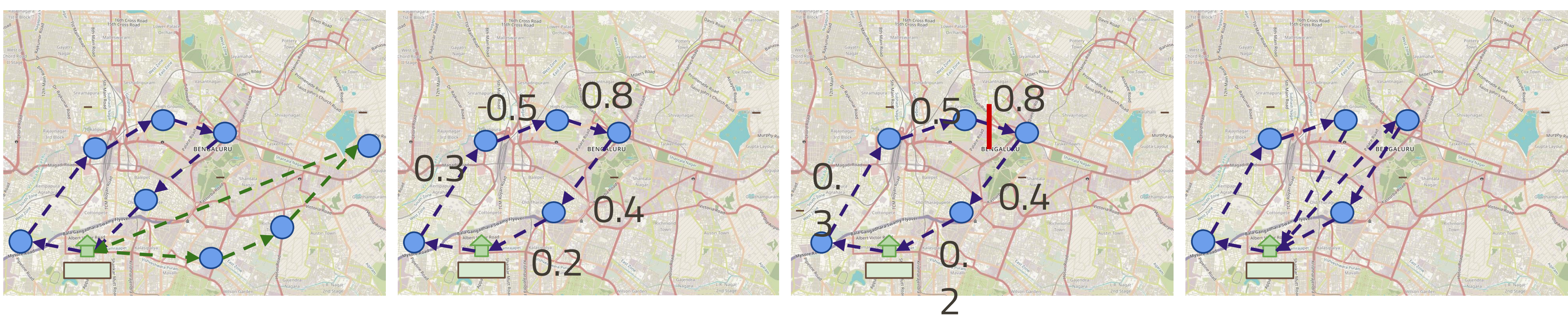


- Activities clustered using ST-DBSCAN[1]
- Each cluster is then assigned drones
  - # drones assigned ∝ cluster cardinality
  - Non-overlapping time window clusters
- may get assigned the same drones

[1] Birant, Derya, and Alp Kut. "ST-DBSCAN: An algorithm for clustering spatial–temporal data." *Data & knowledge engineering* 60.1 (2007): 208-221.

Vehicle Routing Centric (VRC) algorithm

- Intuitively
  - Find near-optimal waypoint routing
  - Maximize utility by scheduling batches
- The algorithm is split into three phases
  - Routing Phase
  - Splitting Phase
  - Scheduling Phase
- Routing Phase finds paths that satisfy the temporal constraints
  - Use a modified 2-OPT*[1] algorithm
  - Paths may not satisfy energy constraints



[1] Birant, Derya, and Alp Kut. "ST-DBSCAN: An algorithm for clustering spatial–temporal data." *Data & knowledge engineering* 60.1 (2007): 208-221.

Experimental Setup

- Performed benchmarks using a commercial-grade drone
- Drone specs
  - Top speed of 6 m/s (20 kmph)
  - 120 m peak altitude
  - 24000 mAh Li-Ion battery
  - Payload capacity of 3kg
  - Two cameras (forward and downward)
  - GPS and LiDAR Lite for navigation
  - NVIDIA Jetson TX2 onboard
    - 256 CUDA cores
    - 8 GB RAM

Drone used for benchmarking

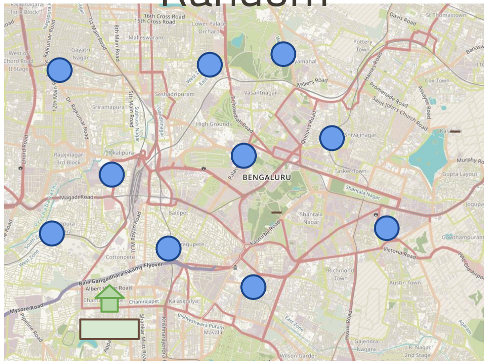
- We inferred the following drone parameters

Speed	Flying energy	Hovering Energy	Compute Energy	Total Energy
4 m/s	750 J/s	700 J/s	20 J/s	1350 kJ

Workload

- Evaluate two application workloads
  - Random (3.5 kms circle)
  - Depth First Search (3.5 km of Bangalore road network)
- The time horizon is set to 4 hrs
- Activity compute
  - SSD Mobilenet v2[1] (MNeT)
  - FCN Resnet18[2] (RNeT)
- A batch is 60 sec of video
- Number of activities varies between 10 to 200
- Number of drones varies between 5 to 50

Random

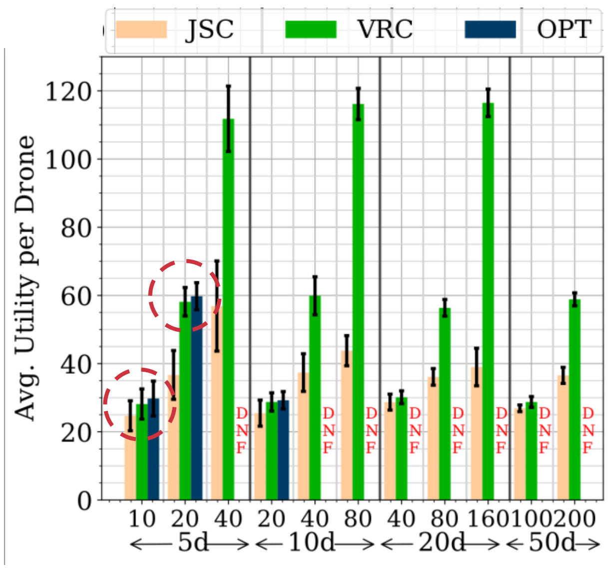


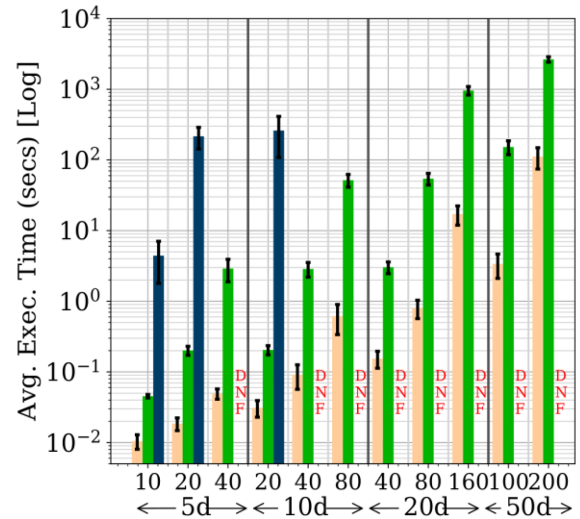
DFS

[1] Sandler, Mark, et al. "Mobilenetv2: Inverted residuals and linear bottlenecks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
[2] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.

Results

- Results for MNeT with DFS

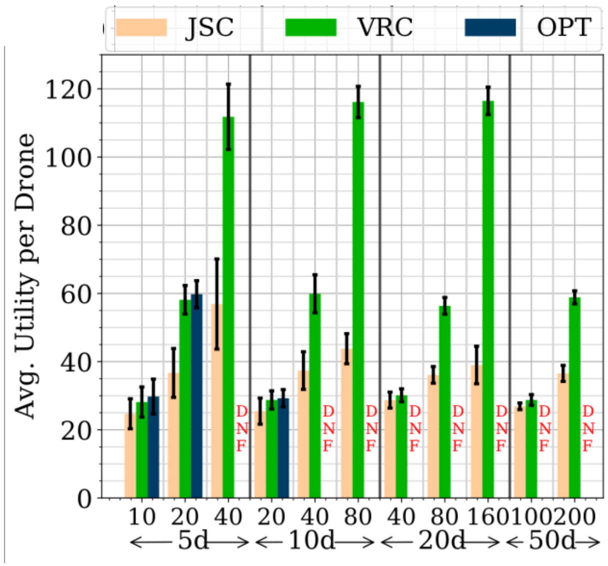


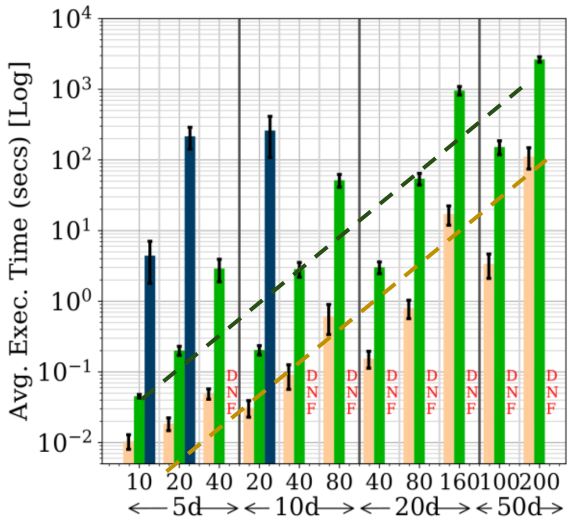


OPT offers the highest utility, if it completes executing, followed by VRC , and JSC

Results

- Results for MNeT with DFS

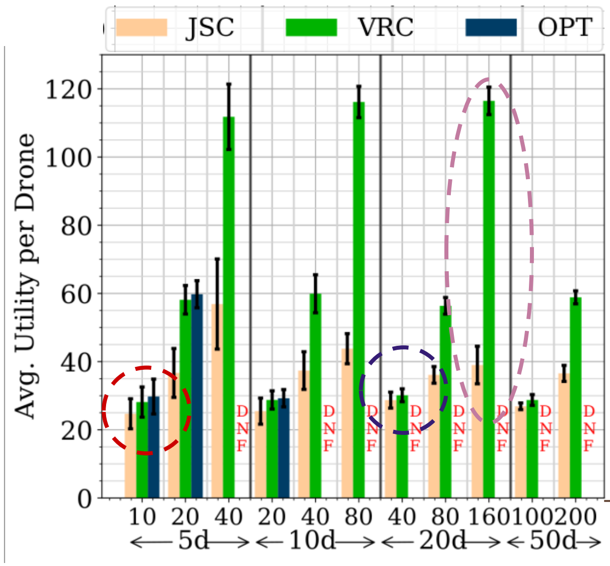


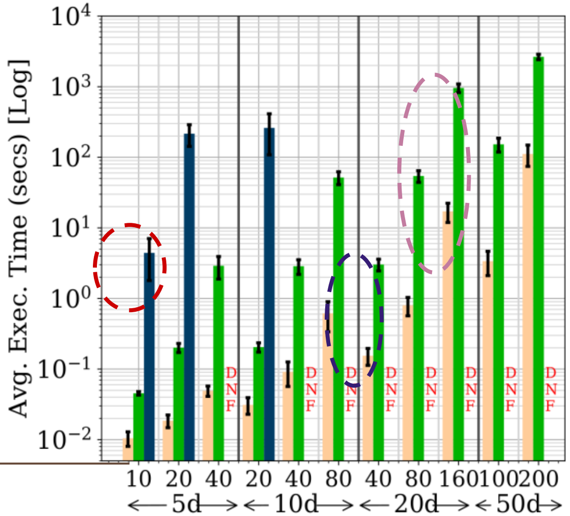


The execution times for VRC and JSC match their time complexity

Results

- Results for MNeT with DFS



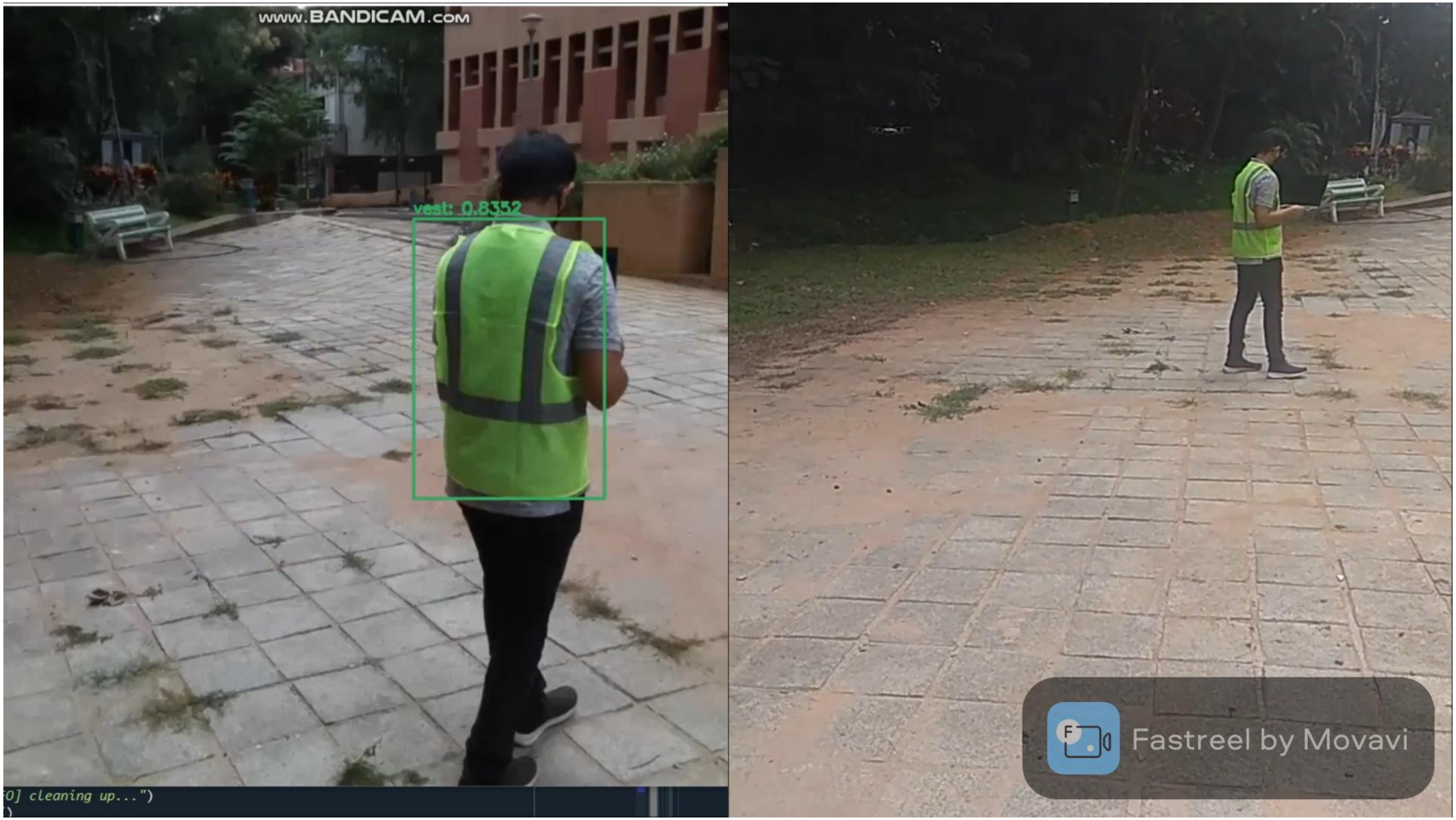


The choice of a good scheduling algorithm depends on the fleet size and activity count

Drones for Mobility Assistance

- Guidance for visually impaired runners to lead an active lifestyle
  - Prompt users on navigation
  - Alert about obstructions
  - Get assistance in emergency
- Extrapolate to "a marathon of the future"
  - Swarm of collaborative drones

Drone Buddy: Tracking and Following



Credits: Suman and Sougat, DREAM:Lab

Smart Mobility: Drones USP

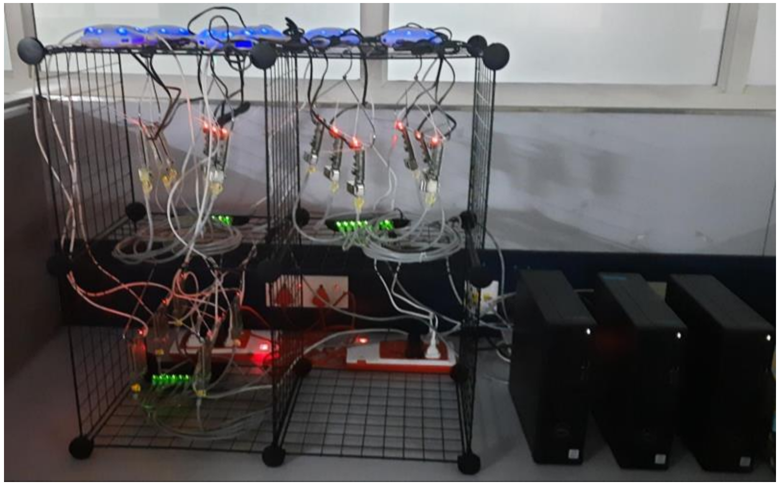
- Sensing + Compute
  - Route co-scheduling
- Online vision algorithms and decisions
  - Tracking person

- Aakash Khochare, Yogesh Simmhan, Francesco Betti Sorbelli and Sajal K. Das **Heuristic Algorithms for Co-scheduling of Edge Analytics and Routes for UAV Fleet Missions**, *IEEE International Conference on Computer Communications (INFOCOM)*, 2021

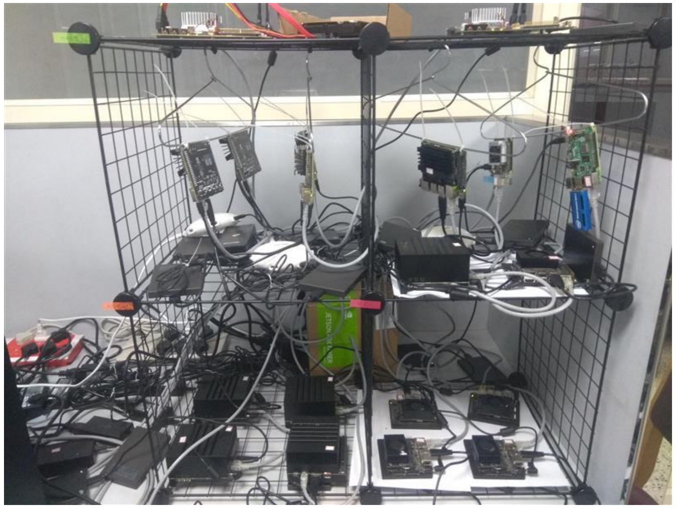
Validation Techniques

How do you validate IoT applications?

- Validation of
  - Physical system *(drone)*
  - Sensing/actuation
  - Communication system
  - Software systems
  - Analytics & decisions
- **Physical testbed**
  - Realistic
  - Hard to manage, scale
  - Less flexibility
  - High cost



Uruk 40-node Pi Cluster



Vega 20-node Jetson Cluster

Validation

- **Emulation**
  - Realistic behavior using real systems
  - Use virtual machines, containers for compute
  - Communication layer using mininet
  - Can run real code, send real data
  - Cost, flexibility, scale between physical and simulation

- Badiger, S.; Baheti, S. & Simmhan, Y. , VIoLET: A Large-scale Virtual Environment for Internet of Things, IInternational European Conference on Parallel and Distributed Computing (EuroPar), 2018

Validation

- **Simulation**
  - Discrete event simulators
  - Diverse network models: 4G LTE, wifi, ad hoc using NS3
  - Diverse physical sim: Vehicles (Sumo), Drones (Gazebo, ROS)
  - Edge and cloud resources (CloudSim)
  - Much more flexibility and scale
  - But less realistic
- Typically, need to combine simulators for different layers together

Thank you!

Tutorial details at:

https://github.com/dream-lab/acm-school-iot-2021