Location Verification of Crowd-Sourced Sensors

Christopher Kitras[*], Carter Pollan[*], Kyle Myers[*], Camille Wirthlin Tischner[†], Philip Lundrigan[*]

[*]Brigham Young University
[†]Utah State University

Outline

- Background Motivation
- Current State of the Device Registration
- New Device Registration Process
- Proximity Validation Tests
- Change of Location Detection (CoLD)
- Change of Location Tests
- Conclusion

Outline

- **Background Motivation**
- Current State of the Device Registration
- New Device Registration Process
- Proximity Validation Tests
- Change of Location Detection (CoLD)
- Change of Location Tests
- Conclusion

BYU


2017 © Reuters

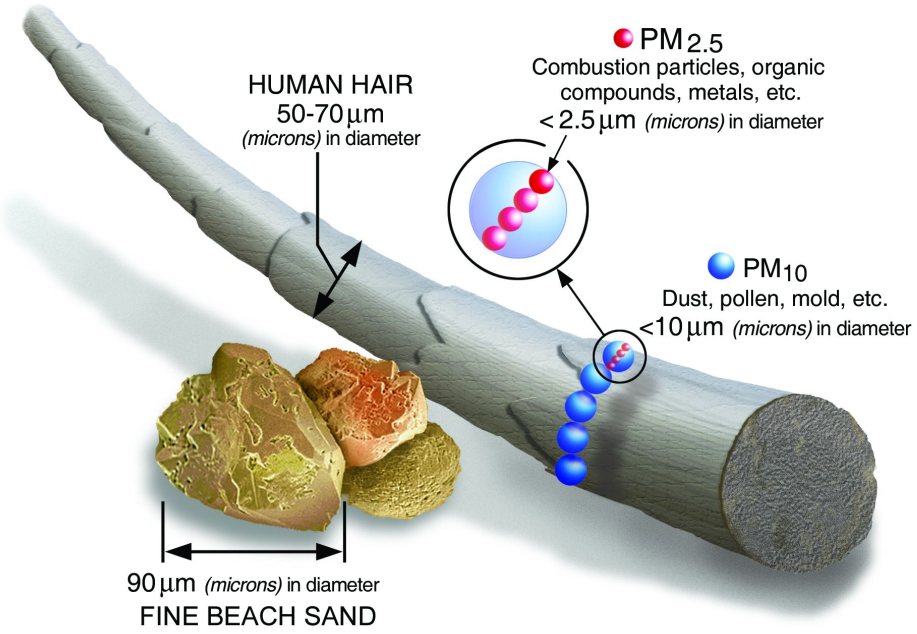

2022 © The New York Times


2023 © Anadolu Agency


2023 © The New York Times

Why Air Quality?

- Climate change has exacerbated air quality crises
- PM2.5 is sediment of a diameter of ≤2.5μm
- PM2.5 enters directly into the bloodstream due to its small size.
- Growing focus on monitoring PM2.5 to track impact

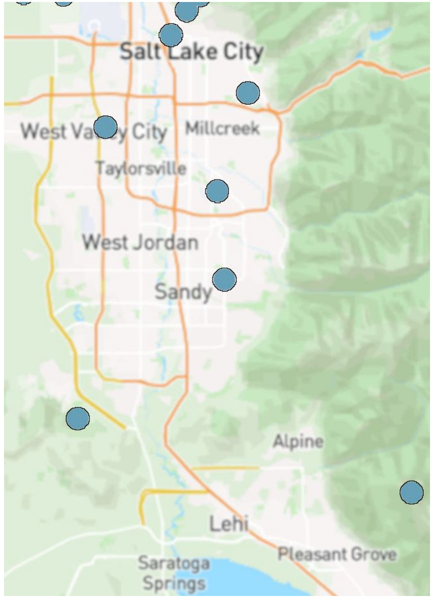


Source: U.S. EPA

BYU

Government Sensor Density

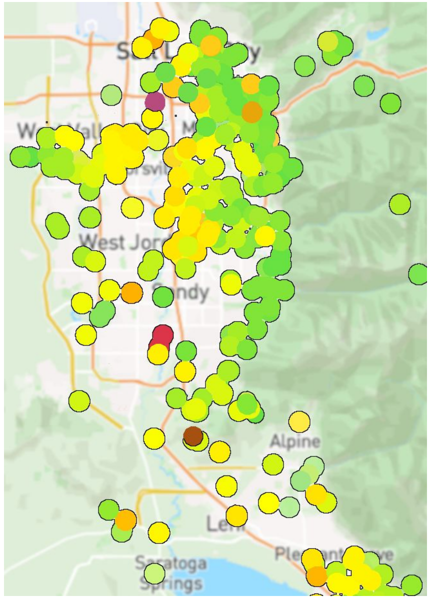
- Air quality monitoring is regulated in the United States[3]
- Mandated air quality stations deployed and monitored by government
- Must be calibrated on a frequent basis by trained personnel
- Very expensive to deploy and maintain (i.e. $10,000+)[1]



BYU

Citizen Science Sensor Density

- Citizen Science (i.e. Crowd Sourced sensors) are sold by companies and deployed by enthusiasts/users
- Sensors are calibrated in firmware by the company (i.e. baked in correction factor)
- Low cost: ~$230/unit[2]
- This leads to greater sensor density



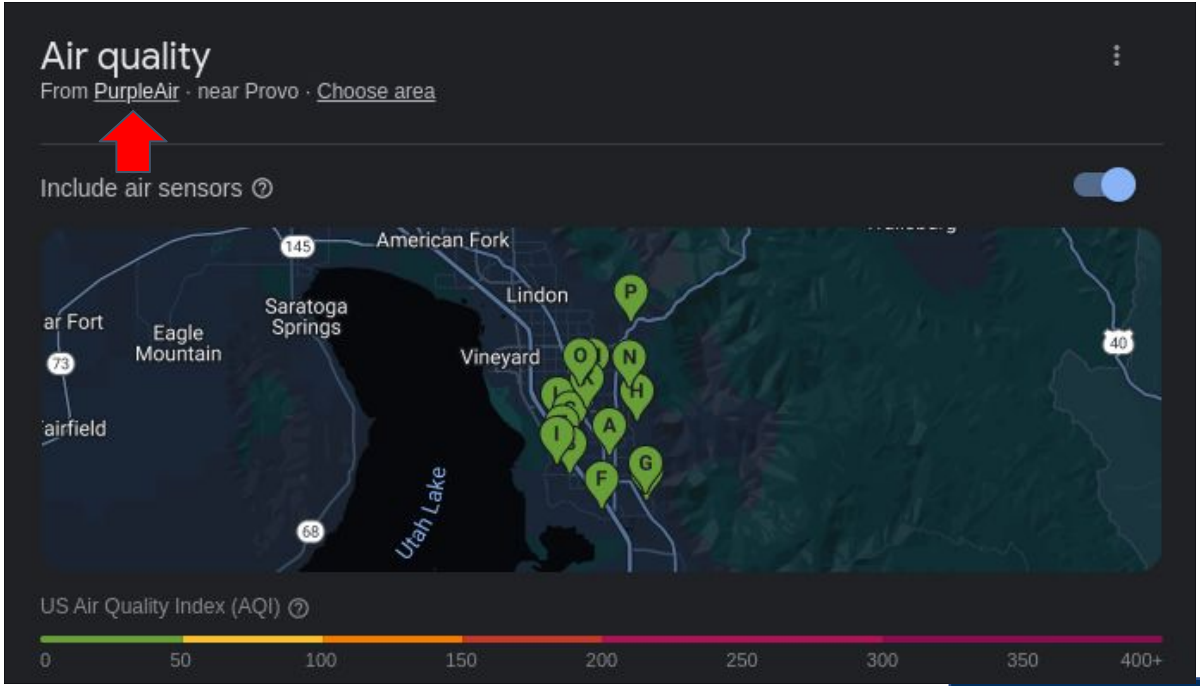
BYU


PurpleAir
Map
Sensors
Support
Login
Lynnwood
Edmonds
Mountlake Terrace
Woodway
Shoreline
Lake Forest Park
Seattle
PA-II-SD AIR QUALITY SENSOR
www.purpleair.com

BYU

Sensor Data in the Wild

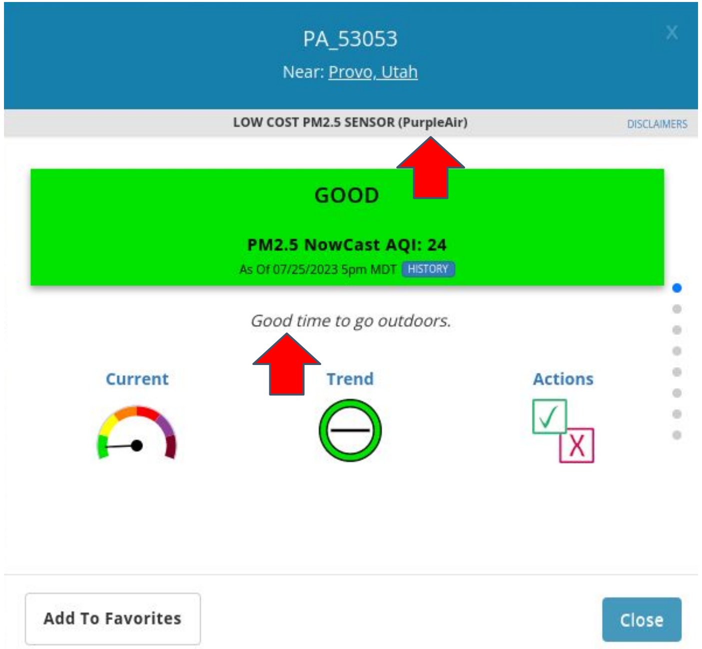
- Citizen science data is becoming more trusted
- A simple Google search shows air quality from local sources



BYU

Sensor Data in the Wild

- AirNow also shows citizen science-based PM2.5 readings
- Citizen science data is everywhere!



BYU

Problem

Device-Id (MAC)*

Printed on the device label just above the bar code. Please include the colons (:)

xx:xx:xx:xx:xx:xx

Associated Email *

This email address would have been used in the device purchase or other communication with PurpleAir. *(A copy of this sensor registration will be e-mailed to this address.)*

Associated email address

Installed*

Outside | Inside

Location Name*

The name that appears on the map

Visibility*

Public (everyone) | Private (only me)

☑ Set a location on the map

Map Location*
(drag the marker to adjust)

Latitude 21.2758001

Longitude -157.8251292

BYU

Problem

- With this lack of location verification, anyone from **absent-minded, well-meaning users to malicious actors** intent on ruining the integrity of the system's data could falsely **place a sensing device anywhere on the map**.

- People are making **important health decisions** on data that cannot be trusted

- How can we prove that a sensing **device is installed in its registered location without extra hardware?**

BYU

Related Work

Some previous efforts to pinpoint location of a device:

1. GPS: requires extra hardware, finicky outside of certain situations, i.e building cover, etc.
2. WiGle: WiFi fingerprinting database. Not as useful in rural locations. Not great for real-time verification
3. IP Geolocation databases Geolocate and GeoIP2: not very granular, dependant on ISP conformity and population density

BYU

Our Solution

- We aimed to create a solution that:
    - verifies a device's location **without extra hardware**
    - **detects any changes** in the device's location
    - scales to be deployed on any system **without requiring a platform-specific application**
- These design goals **prevent** the need for **recalling and retrofitting devices** with localization hardware, prevent device relocation after verification, and ensure accessibility to users with unsupported smartphone models.

BYU

Solution

We assume that a viable solution will ensure:

1. **Proximity of a registering device** with trusted geolocation services to a WiFi device
2. **Detect any change of location** of the WiFi device after a verified registration

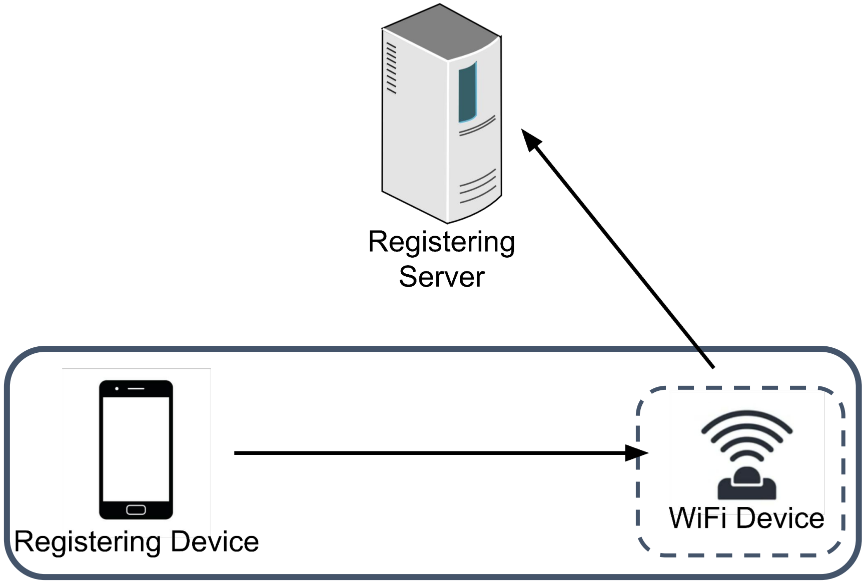
BYU

Outline

- Background Motivation
- **Current State of the Device Registration**
- New Device Registration Process
- Proximity Validation Tests
- Change of Location Detection (CoLD)
- Change of Location Tests
- Conclusion

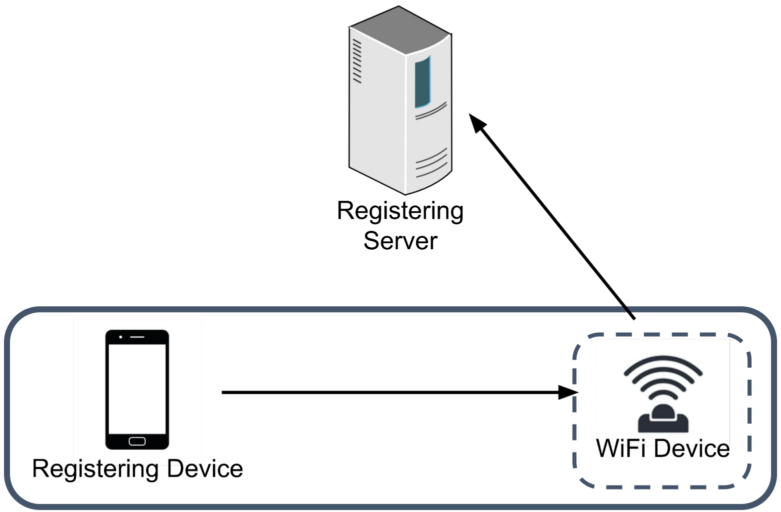
BYU

Registration Model



Registration Model

- WiFi Device (sensor) establishes access point
- Registering Device (phone) connects to WiFi Device and provides network credentials
- Location registration is done via user input or device installer
- There usually little to no verification of this process



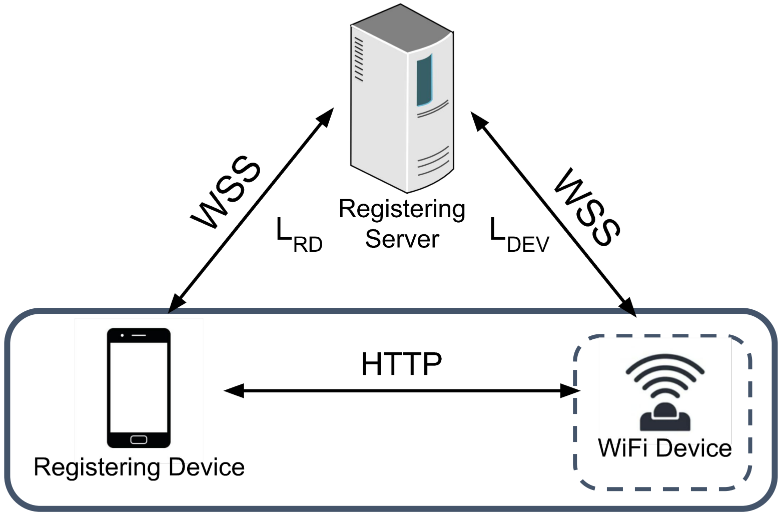
BYU

Outline

- Background Motivation
- Current State of the Device Registration
- **New Device Registration Process**
- Proximity Validation Tests
- Change of Location Detection (CoLD)
- Change of Location Tests
- Conclusion

New Registration Model

- Use WebSockets to measure the latencies between phone and registering server (L_{RD}) and the sensor and registering server (L_{DEV}).
- Registration token (T_{REG}) is shared between all nodes to ensure integrity
- Define a tolerance between latencies (L_{TOL}) and ensure $|L_{RD} - L_{DEV}| \leq L_{TOL}$

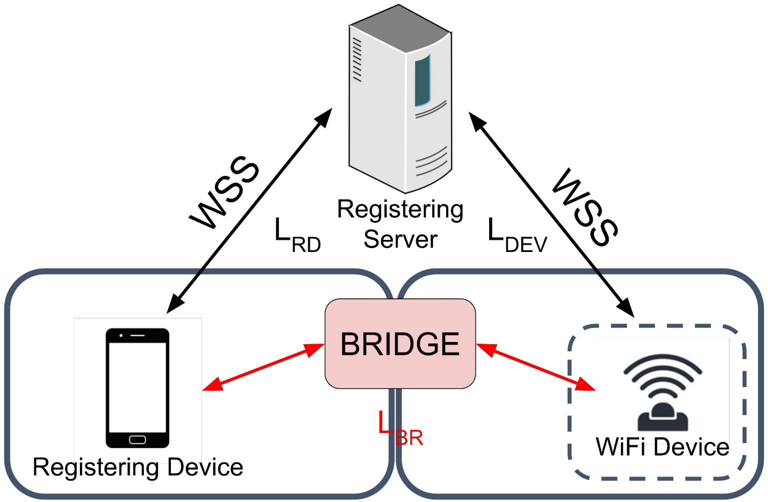


BYU

Adversarial Model

A supposed attacker:

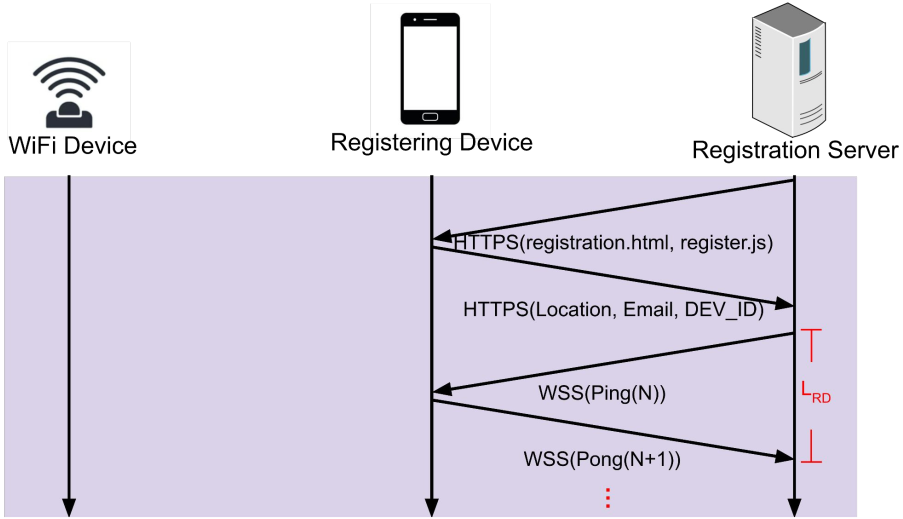
- Has complete control over their network, local packets, firmware on sensor, and software on phone
- Can perform man-in-the-middle attacks on packets in their network
- Can relay packets through different devices (i.e. a bridge) to give appearance of different location of origin



BYU

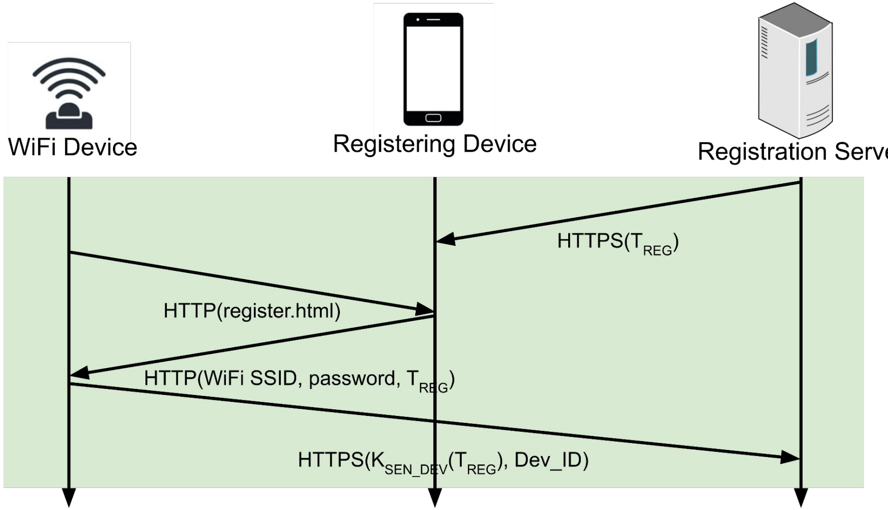
Registration Flow

Pre-existing credentials are shared from phone to Registration Server and L_{RD} is derived



Registration Flow

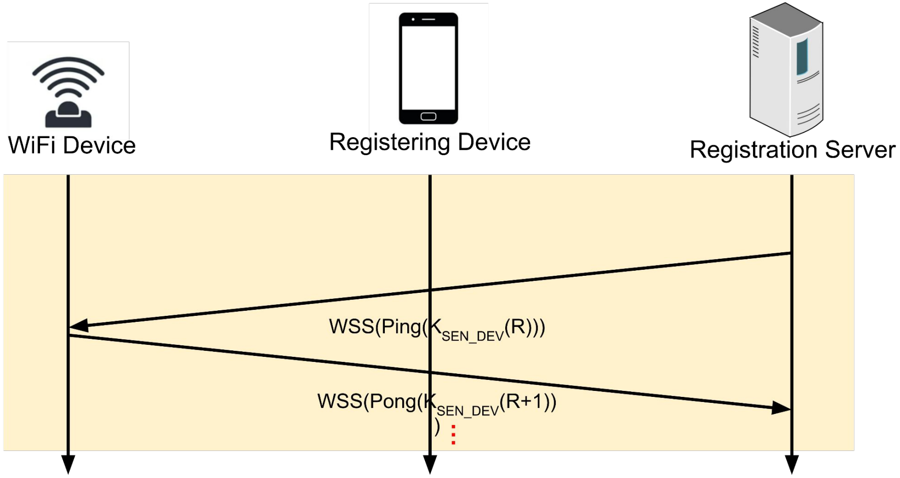
Registration server assigns T_{REG} to phone who passes this to the WiFi Device



BYU

Registration Flow

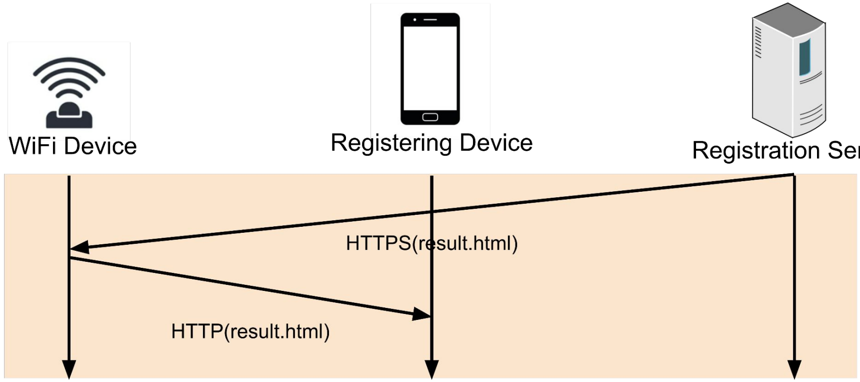
After receiving network credentials, sensor and Registration server derive L_{DEV}



BYU

Registration Flow

Registration server checks $|L_{RD} - L_{DEV}| \leq L_{TOL}$ and accepts or rejects registration session



BYU

Some Development Challenges

- Solution must run in a browser!
- Changing a window from the registration page to the WiFi Device AP
    - RFC 1918
    - No redirecting from broader to smaller network type
- Minimize ping times with WebSockets to avoid overhead of repeated HTTPS requests

Although these requirements may seem strict and obstructive of creative solutions, compliance to them ensures that **anyone from any web browser can carry out the new registration process**.

BYU

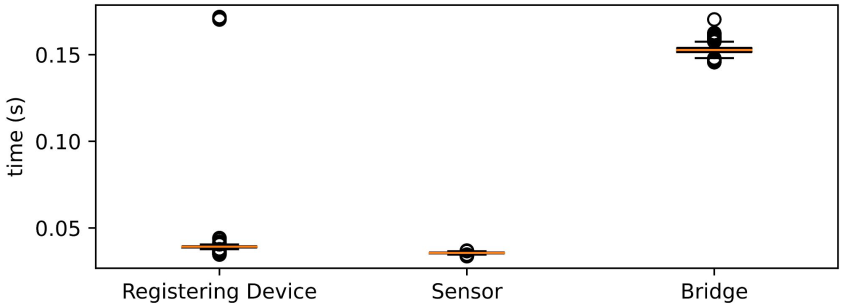
Outline

- Background Motivation
- Current State of the Device Registration
- New Device Registration Process
- **Proximity Validation Tests**
- Change of Location Detection (CoLD)
- Change of Location Tests
- Conclusion

BYU

Comparing Latencies

- Measured L_{RD} and L_{DEV} over the span of a day
- $L_{RD} - L_{DEV} \approxeq 5ms$
- Measured latency of a bridged setup (L_{BR})
- $L_{BR} - L_{RD} \approxeq 125ms$
- Set L_{TOL} such that $|L_{RD} - L_{DEV}| \leq L_{TOL} \leq L_{BR}$



BYU

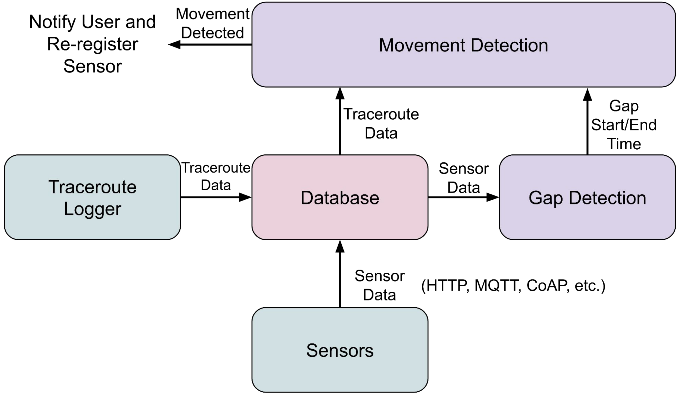
Outline

- Background Motivation
- Current State of the Device Registration
- New Device Registration Process
- Proximity Validation Tests
- **Change of Location Detection (CoLD)**
- Change of Location Tests
- Conclusion

BYU

CoLD Algorithm

- Gather sensor data and poll sensors for traceroute data
- Detect unexpected gaps caused by loss of internet connection/power greater than defined threshold T_{GAP}
- Upon a gap $\geq T_{GAP}$ we take a sample of trusted traceroute data (1 week) and a sample of new traceroute data
- If samples are 90%+ alike, the gap is ignored, else the data is flagged and the sensor is marked for re-registration



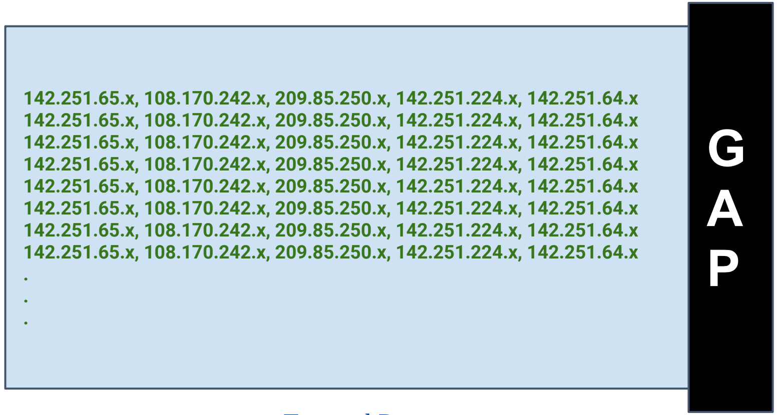
BYU

Experiment Procedure

G
A
P

BYU

Experiment Procedure



Trusted Data

Experiment Procedure

142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 142.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 142.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 142.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 142.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 142.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 142.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 142.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 142.251.64.x
.
.
.

Trusted Data

GAP

142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 209.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 209.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 209.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 209.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 209.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 209.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 209.251.64.x
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 209.251.64.x
.
.
.

Questionable Data

Experiment Procedure

142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 142.251.64.x, x.x.x.x, y.y.y.y, z.z.z.z

Trusted Data

142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 209.251.64.x, a.a.a.a, b.b.b.b, c.c.c.c

Questionable Data

BYU

Experiment Procedure

142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 142.251.64.x, x.x.x.x, y.y.y.y, z.z.z.z

Trusted Data

142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 209.251.64.x, a.a.a.a, b.b.b.b, c.c.c.c

Questionable Data

50%

BYU

Experiment Procedure

Acceptance threshold in our system is 90%+

Trusted Data	Questionable Data
142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 142.251.64.x, x.x.x.x, y.y.y.y, z.z.z.z	142.251.65.x, 108.170.242.x, 209.85.250.x, 142.251.224.x, 209.251.64.x, a.a.a.a, b.b.b.b, c.c.c.c

50%

BYU

Outline

- Background Motivation
- Current State of the Device Registration
- New Device Registration Process
- Proximity Validation Tests
- Change of Location Detection (CoLD)
- **Change of Location Tests**
- Conclusion

Experiment Procedure

In three different geographical regions we did the following:

1. Run framework normally for at least 1 week
2. Simulate a gap event
3. Fetch trusted data and a sample of questionable data
4. Compare data samples and assign a score
5. Create a confusion matrix to compare accuracy of scoring future data to past data

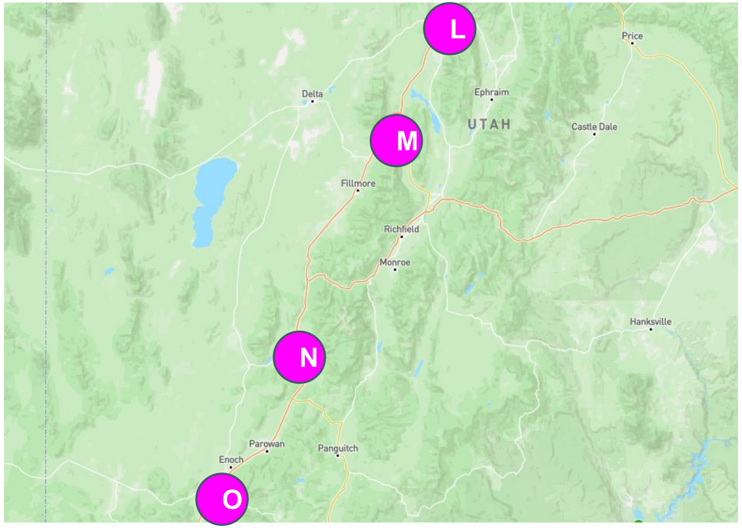
	L	M	N	O
L	99.1	0	0	0
M	0	99.9	0	0
N	0	0	99.9	23.9
O	0	0	1.8	98.0

(a) Rural Area

BYU

Rural Area Test

- Nodes are ~40 miles (64 km) apart
- Compare current node with other node's traceroute data
- Average of ~99.23% same node recognition
- Highest recognition in N to O with 23.9%



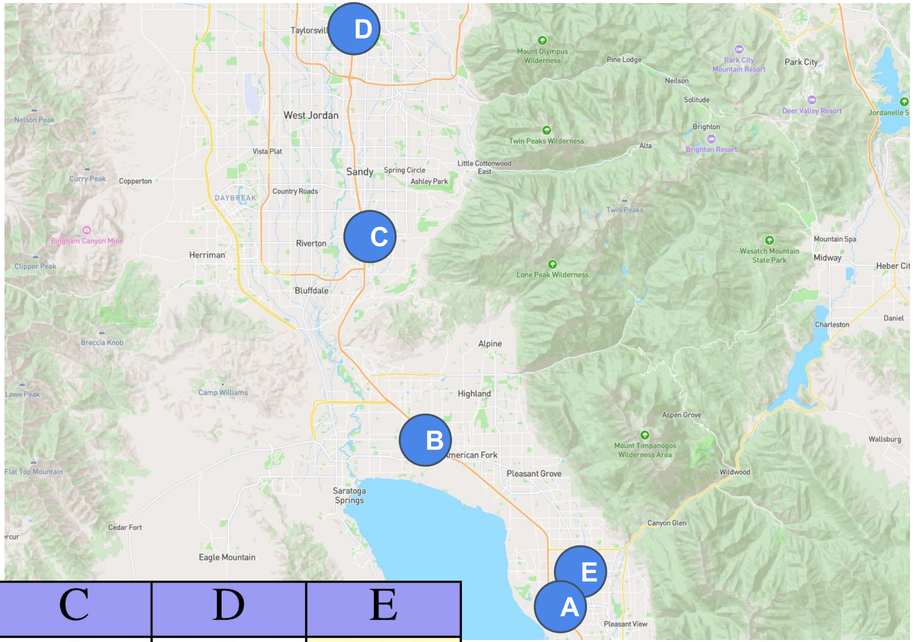
	L	M	N	O
L	99.1	0	0	0
M	0	99.9	0	0
N	0	0	99.9	23.9
O	0	0	1.8	98.0

(a) Rural Area

BYU

Inter-City Test

- Nodes are ~8 miles (13 km) apart
- Average of ~98.76% same node recognition
- Highest recognition in A to E with 14.6%



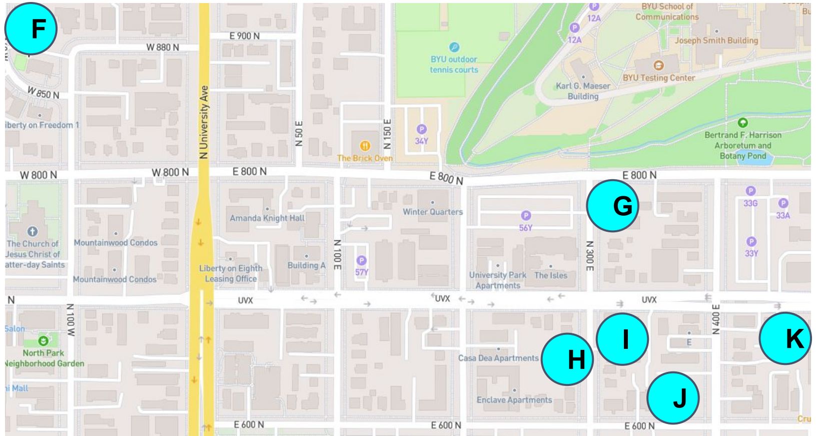
	A	B	C	D	E
A	99.9	0	0	0	14.6
B	0	96.39	0	0	0
C	0	0	99.7	0	0
D	0	0	0	98.0	0
E	11.3	0	0	0	99.8

(b) Inter-City

BYU

Intra-City Test

- Nodes are a few city blocks apart
- Average of ~98.47% same node recognition
- Highest recognition across several pairings with a 66.6%



	F	G	H	I	J	K
F	99.9	0	66.6	0	0	66.6
G	0	96.3	0	0	0	0
H	66.6	0	99.9	0	0	66.6
I	0	0	0	94.9	0	0
J	0	0	0	0	99.9	0
K	66.6	0	66.6	0	0	99.9

(c) Intra-City

BYU

Conclusion

- Created a solution that **detects location and change of location**
- **No need of retrofitting** sensors with more hardware
- Experiments indicate a high rate of success with self identifying across:
  - Distant cities
  - Neighboring cities
  - Same city
- Solution can run on **any registering device** with a browser and localization engine
- Framework provides the necessary key for **automatic, low-cost location verification** for citizen science devices

BYU

Questions?

BYU

References

1. Gryech I, Ben-Aboud Y, Guermah B, Sbihi N, Ghogho M, Kobbane A. MoreAir: A Low-Cost Urban Air Pollution Monitoring System. Sensors (Basel). 2020 Feb 13;20(4):998. doi: 10.3390/s20040998. PMID: 32069821; PMCID: PMC7071408.
2. https://www2.purpleair.com/products/list
3. https://www.epa.gov/laws-regulations/summary-clean-air-act (Clean Air Act, 1970)

BYU