Where the Sidewalk Ends: Privacy of Opportunistic Backhaul

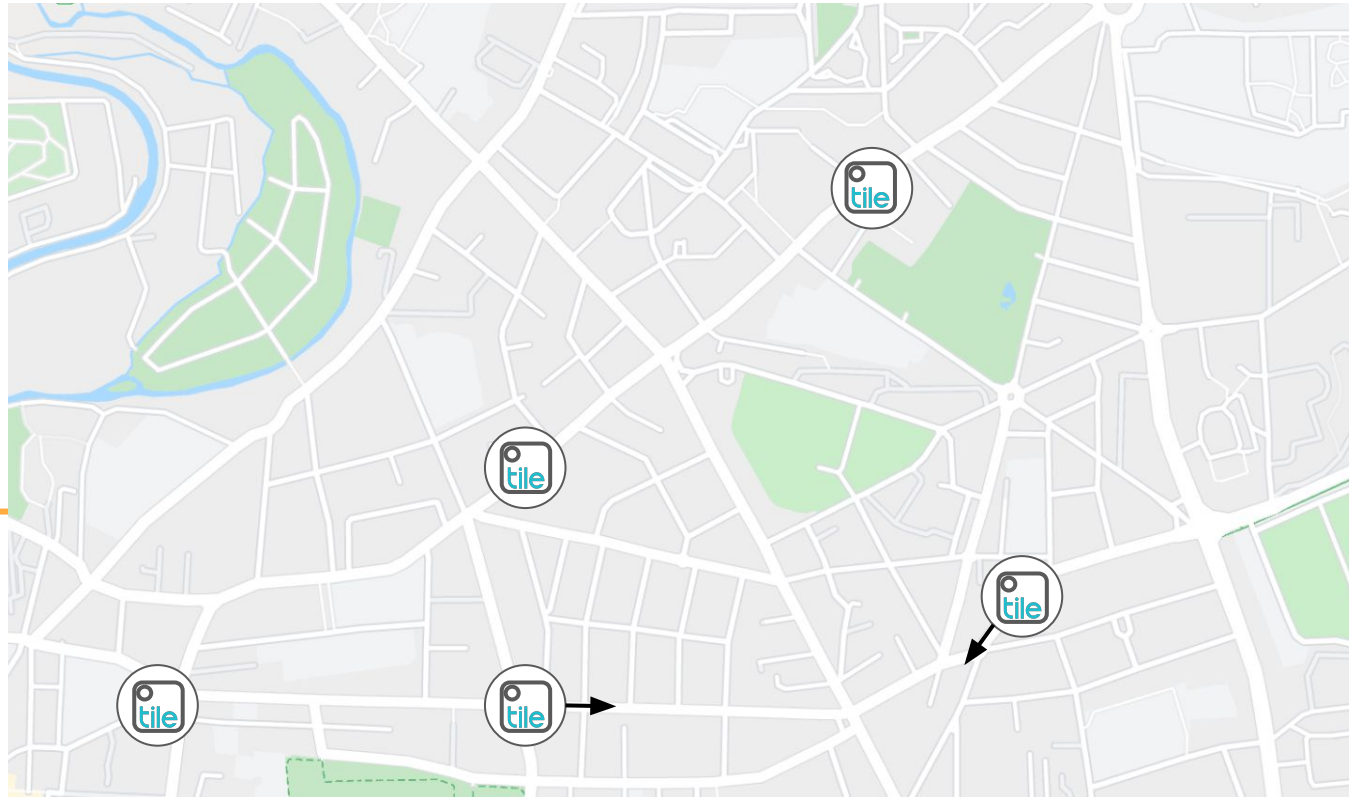
Tess Despres, Shishir Patil, Alvin Tan,
Jean-Luc Watson, Prabal Dutta
UC Berkeley

EuroSec '22 | April 5th, 2022

Opportunistic networks have significantly evolved

Tile

2012



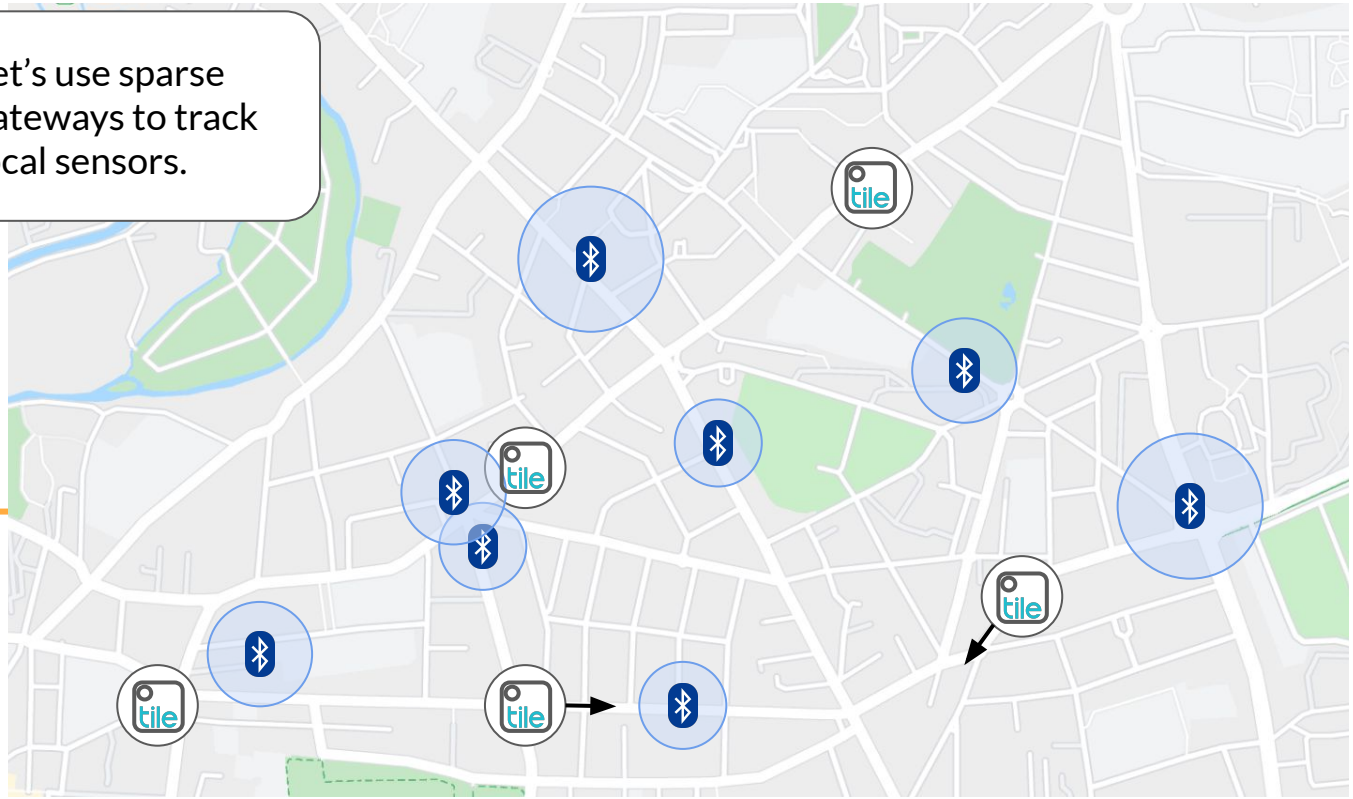
Opportunistic networks have significantly evolved

Let's use sparse gateways to track local sensors.

Tile

2012

BLE-based app

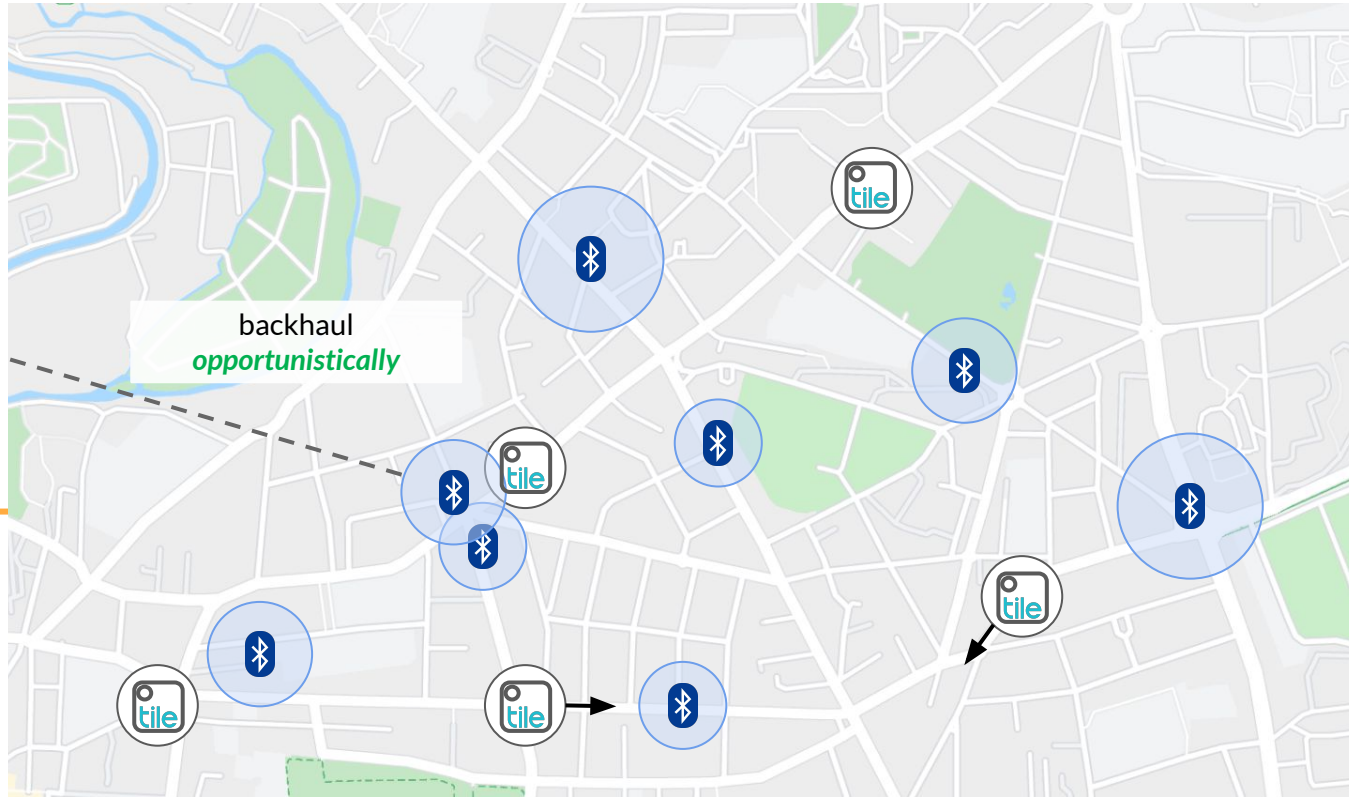


Opportunistic networks have significantly evolved

Tile

2012

BLE-based app



Opportunistic networks have significantly evolved

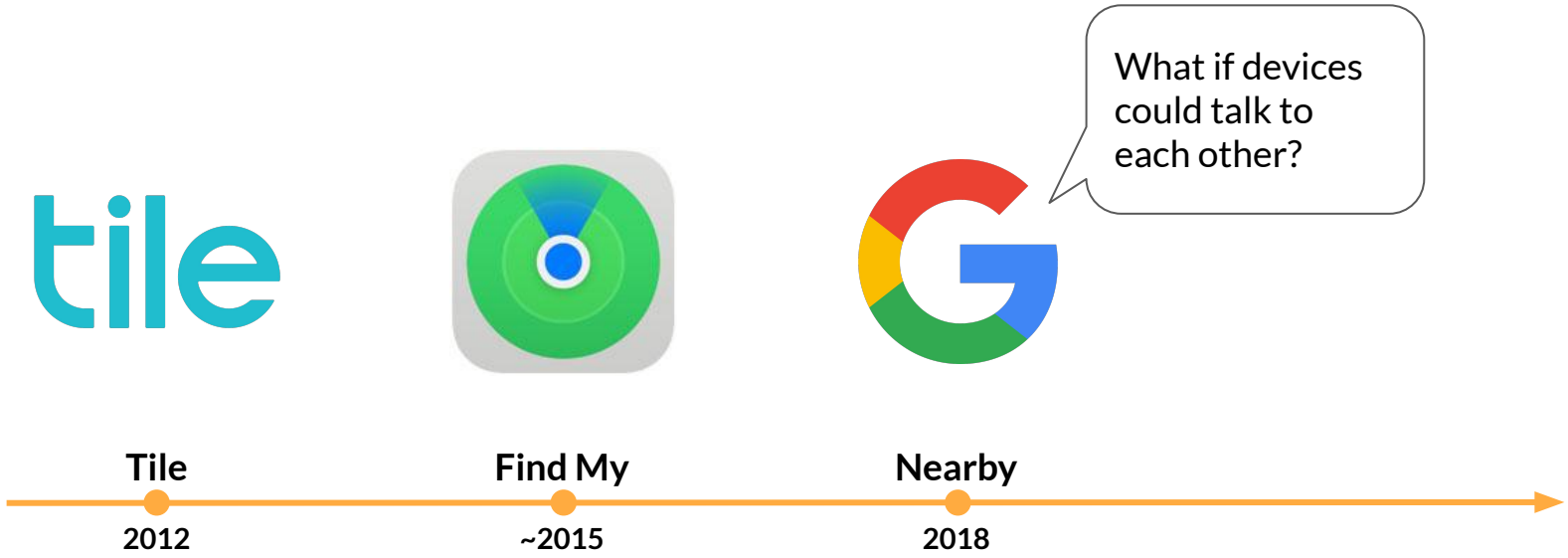
Tile — 2012

BLE-based app

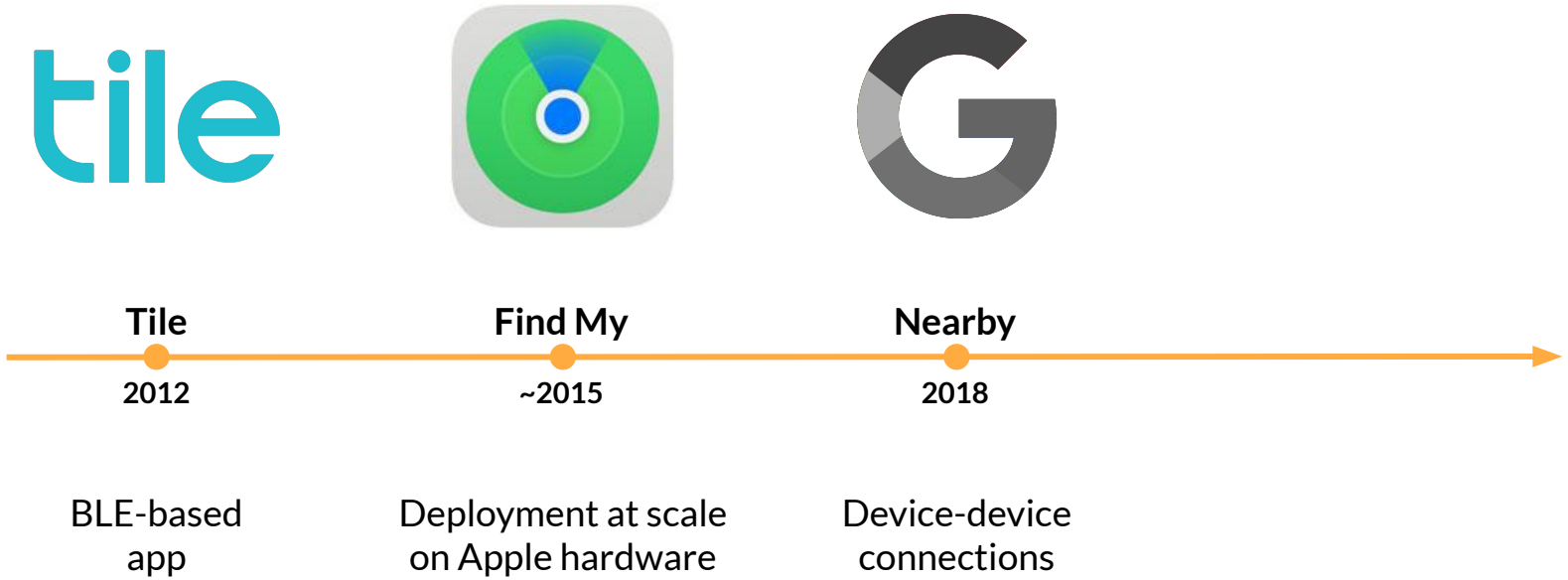
Find My — ~2015

Deployment at scale on Apple hardware

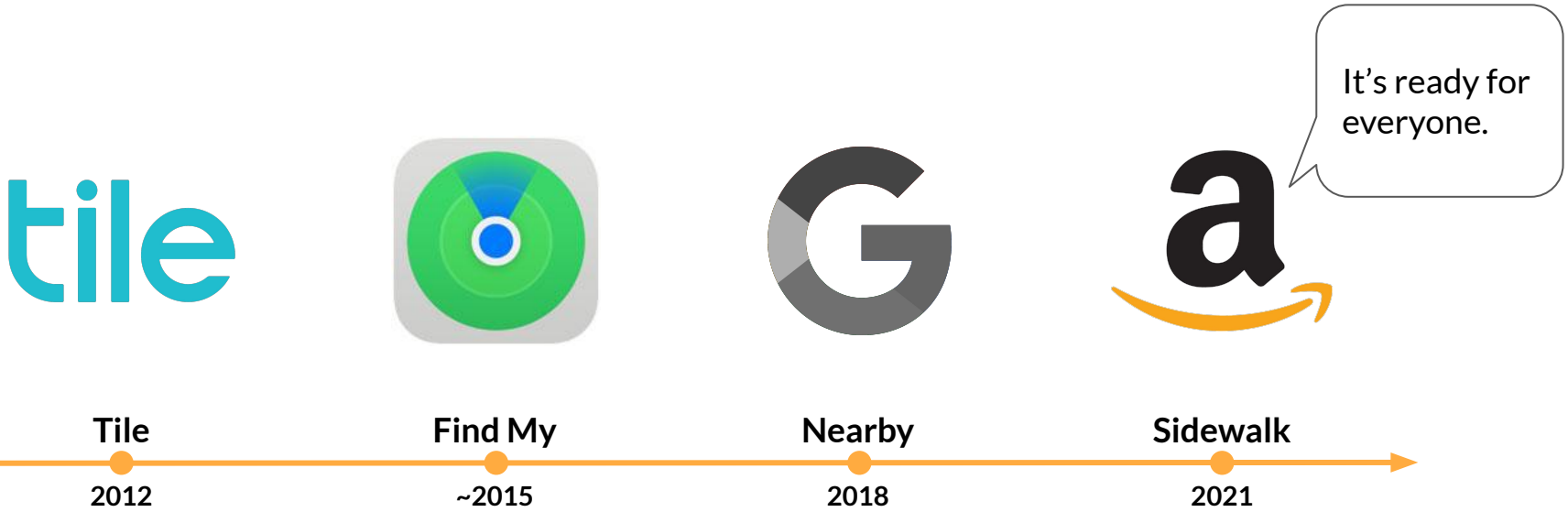
Opportunistic networks have significantly evolved



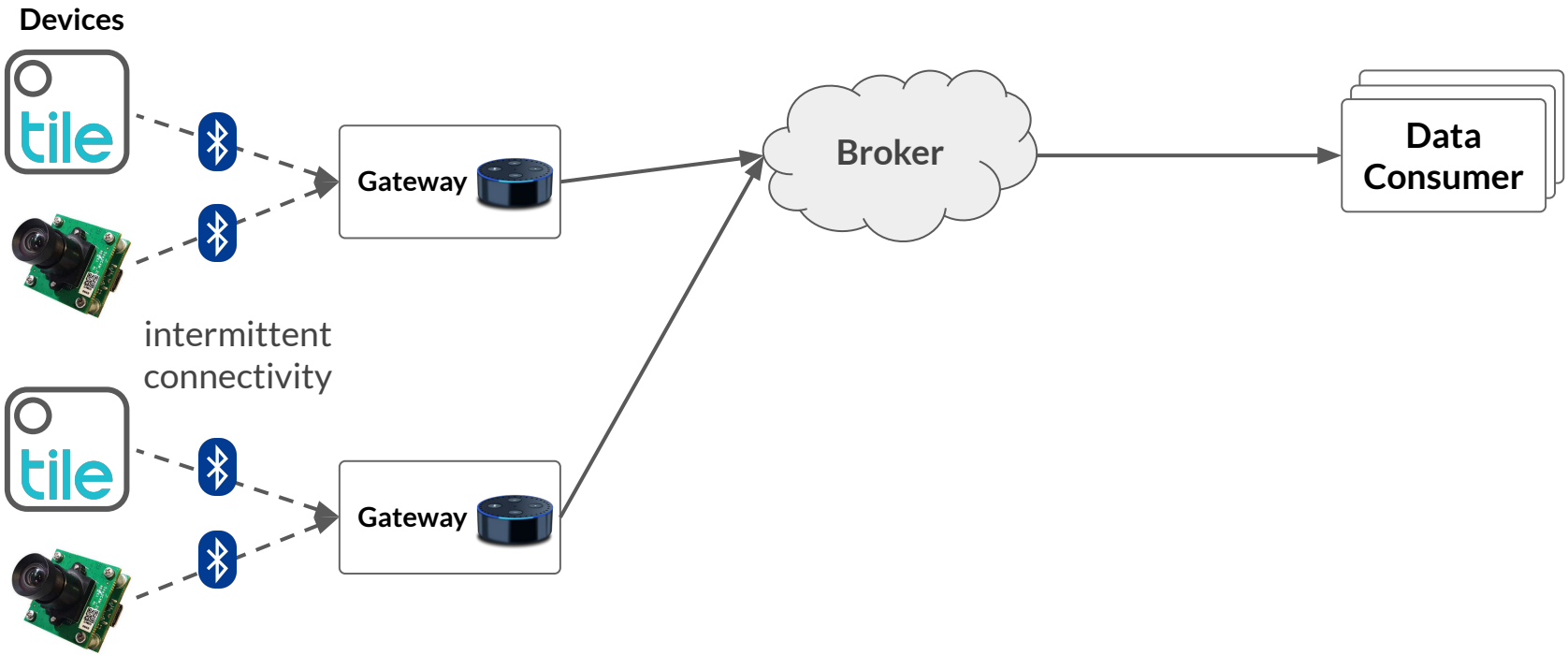
Opportunistic networks have significantly evolved



Opportunistic networks have significantly evolved



An opportunistic network



What are the real privacy concerns?

WIRED SIGN IN

DAVID NIELD GEAR MAY 11, 2021 3:18 PM

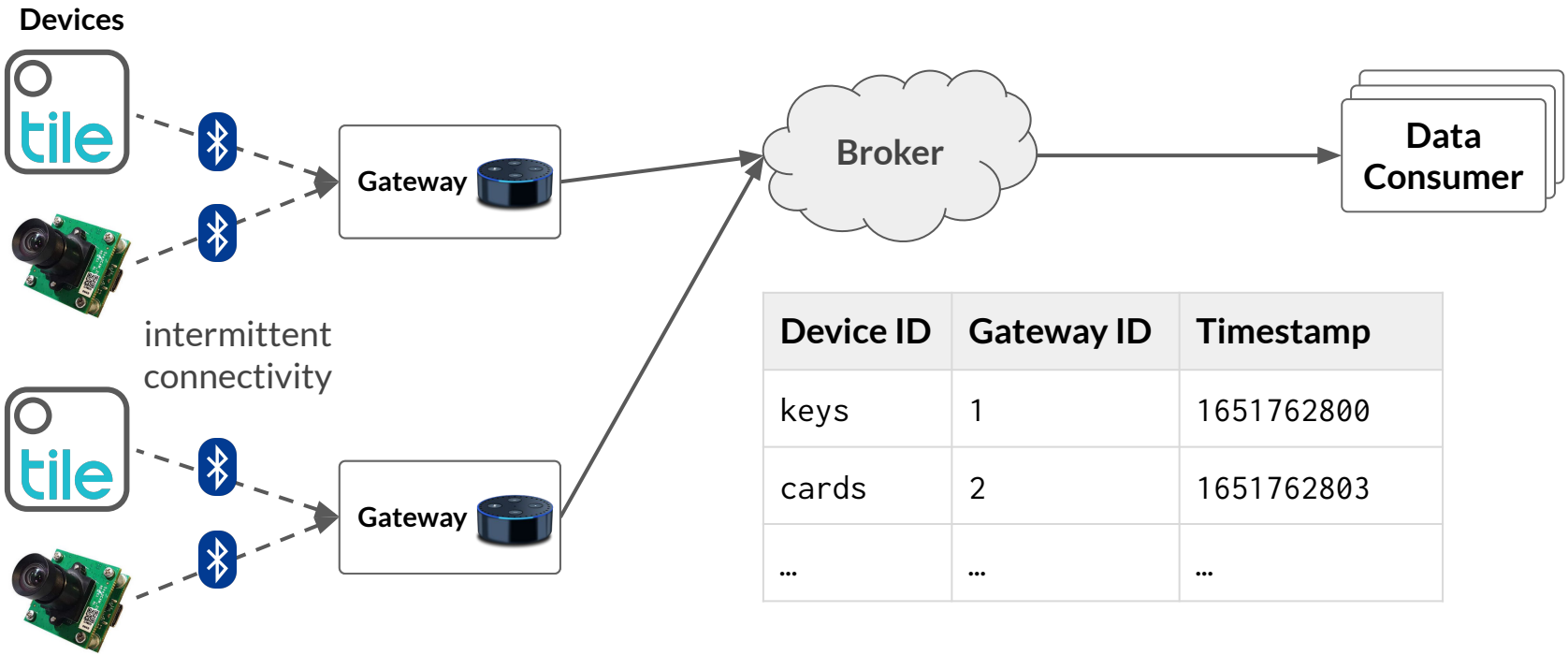
How Amazon Sidewalk Works—and Why You May Want to Turn It Off

T | Wirecutter

Amazon Sidewalk Will Share Your Internet With Strangers. It's Not As Scary As It Sounds.

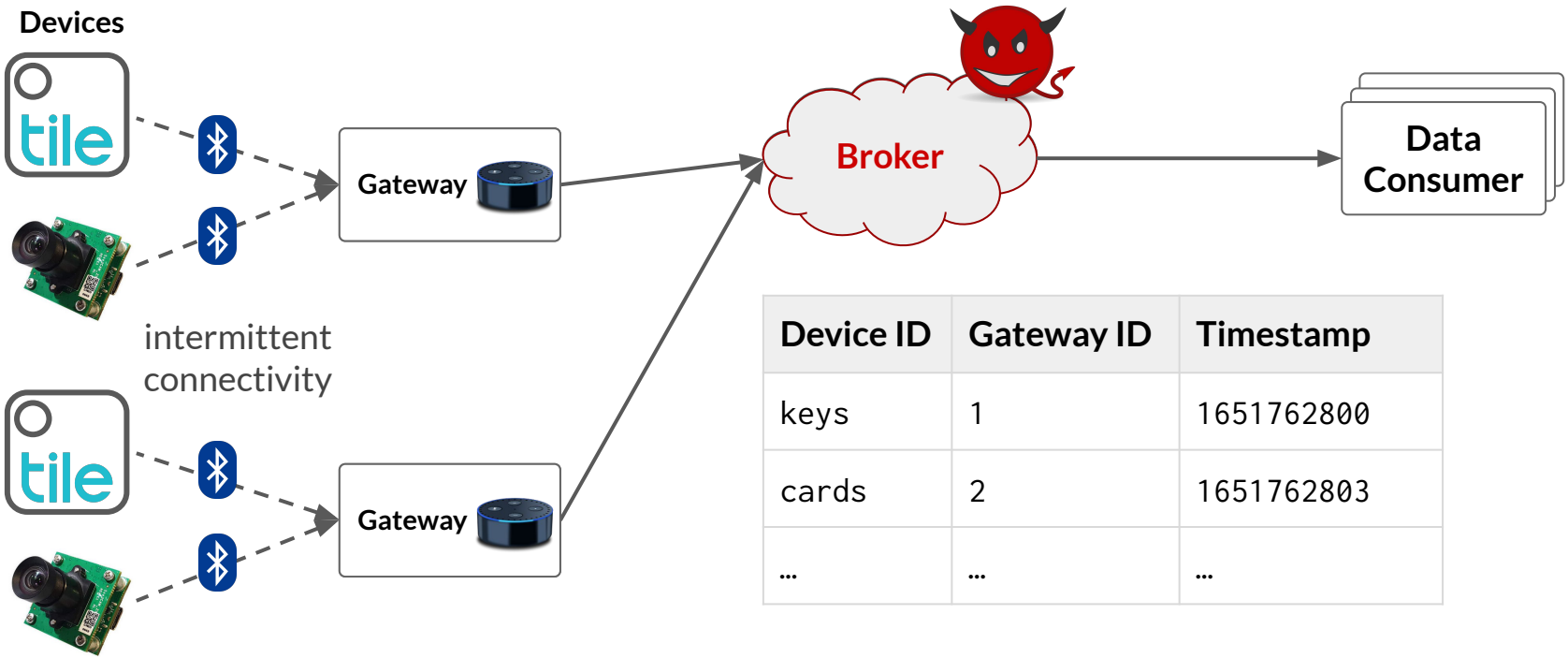
PUBLISHED JUNE 7, 2021

Broker-collected metadata



Device ID	Gateway ID	Timestamp
keys	1	1651762800
cards	2	1651762803
…	…	…

Threat model



Device ID	Gateway ID	Timestamp
keys	1	1651762800
cards	2	1651762803
…	…	…

Sidewalk cannot avoid collecting metadata

Reliance on data retention policy

- Device IDs rotate frequently but the broker knows the PRG seeds they are derived from

Device ID	Gateway ID	Timestamp
keys	1	1651762800
cards	2	1651762803
…	…	…

[1] https://m.media-amazon.com/images/G/01/sidewalk/final_privacy_security_whitepaper.pdf

Sidewalk cannot avoid collecting metadata

Reliance on data retention policy

- Device IDs rotate frequently but the broker knows the PRG seeds they are derived from

Bidirectional communication

- The broker can identify which recent gateways a device used by their persistent ID

Device ID	Gateway ID	Timestamp
keys	1	1651762800
cards	2	1651762803
…	…	…

[1] https://m.media-amazon.com/images/G/01/sidewalk/final_privacy_security_whitepaper.pdf

Sidewalk cannot avoid collecting metadata

Reliance on data retention policy

- Device IDs rotate frequently but the broker knows the PRG seeds they are derived from

Bidirectional communication

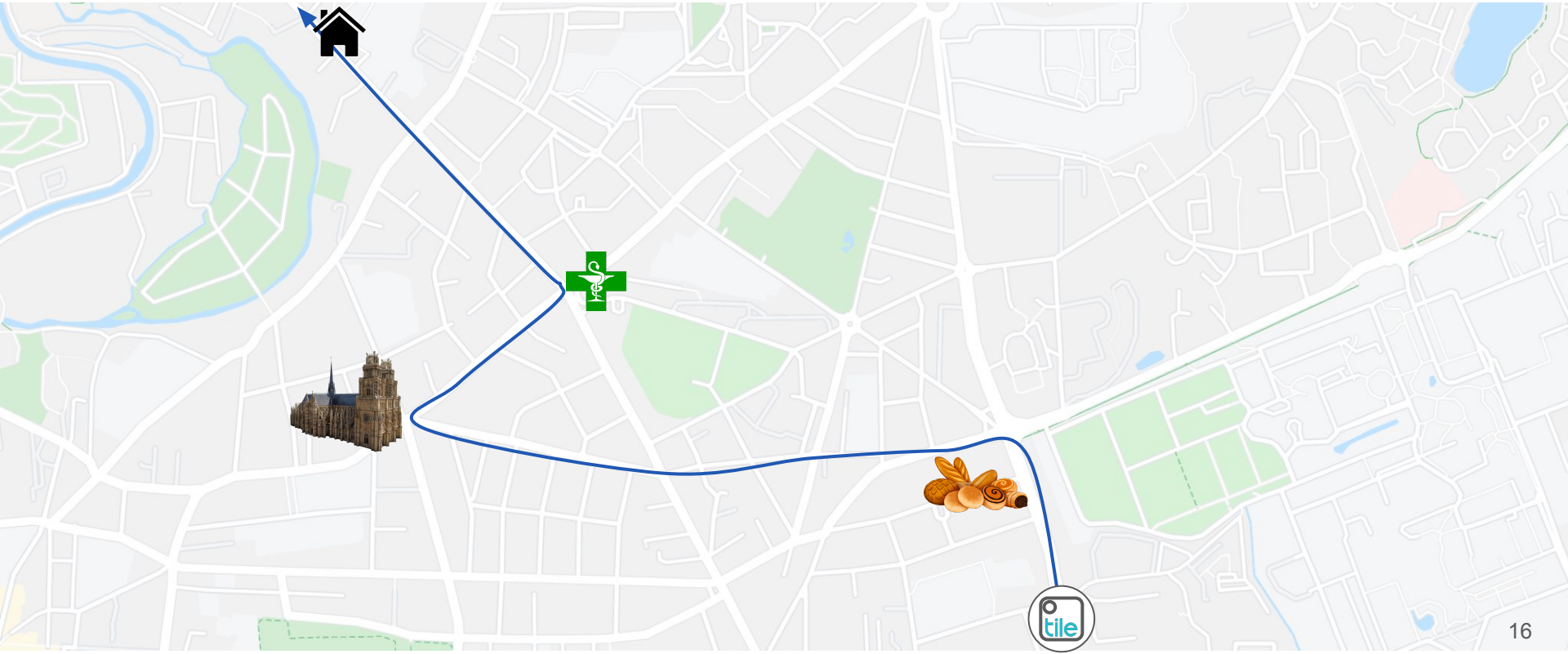
- The broker can identify which recent gateways a device used by their persistent ID

Device ID	Gateway ID	Timestamp
keys	1	1651762800
cards	2	1651762803
…	…	…

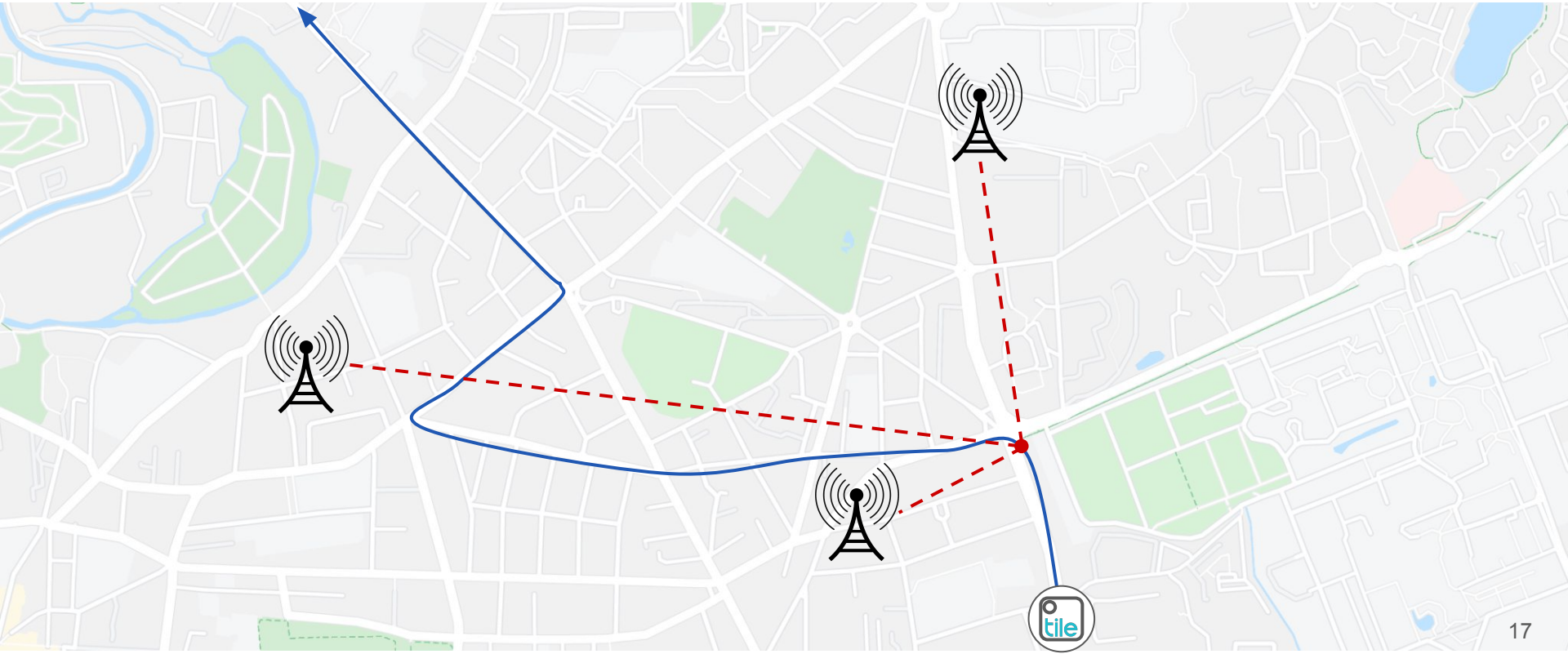
Mobility Trace
Estimation

[1] https://m.media-amazon.com/images/G/01/sidewalk/final_privacy_security_whitepaper.pdf

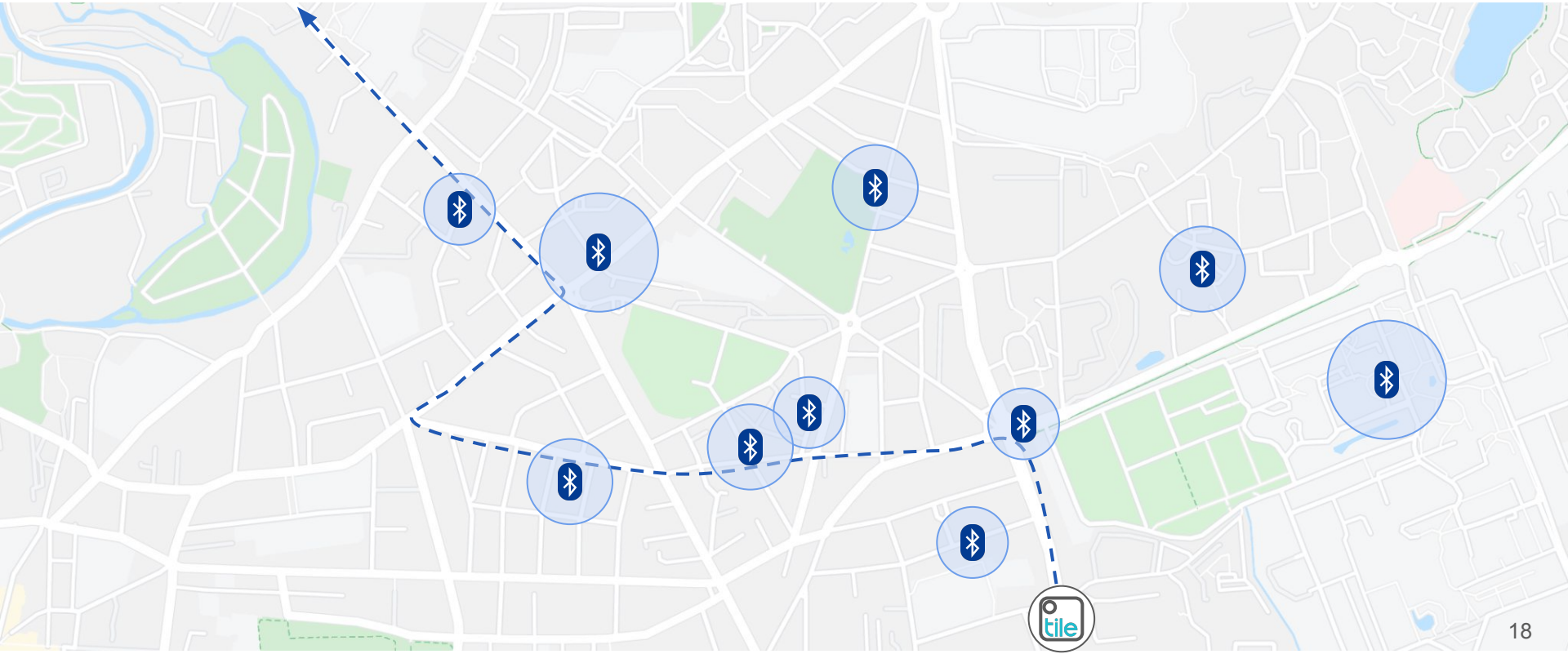
Mobility traces are extremely invasive



Cell towers make triangulation easy



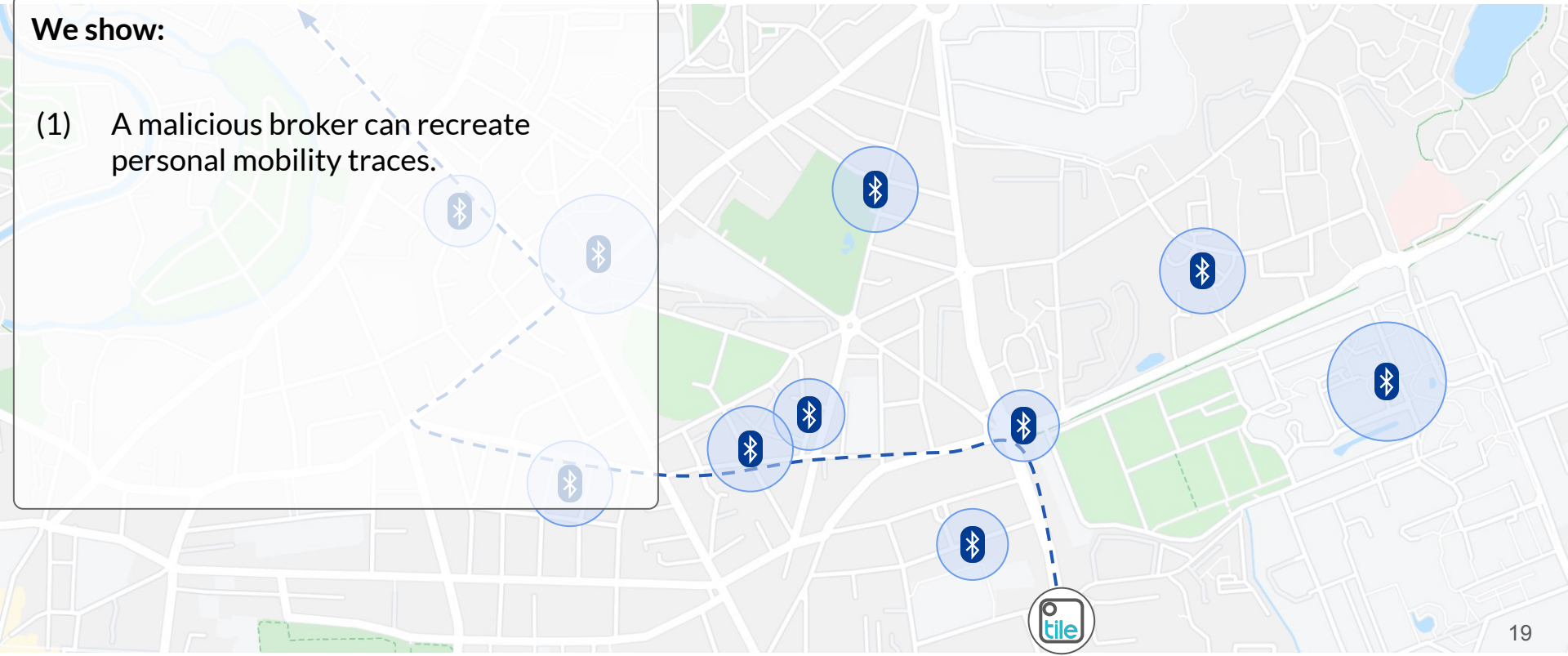
Is the same true in an opportunistic network?



Is the same true in an opportunistic network?

We show:

(1) A malicious broker can recreate personal mobility traces.

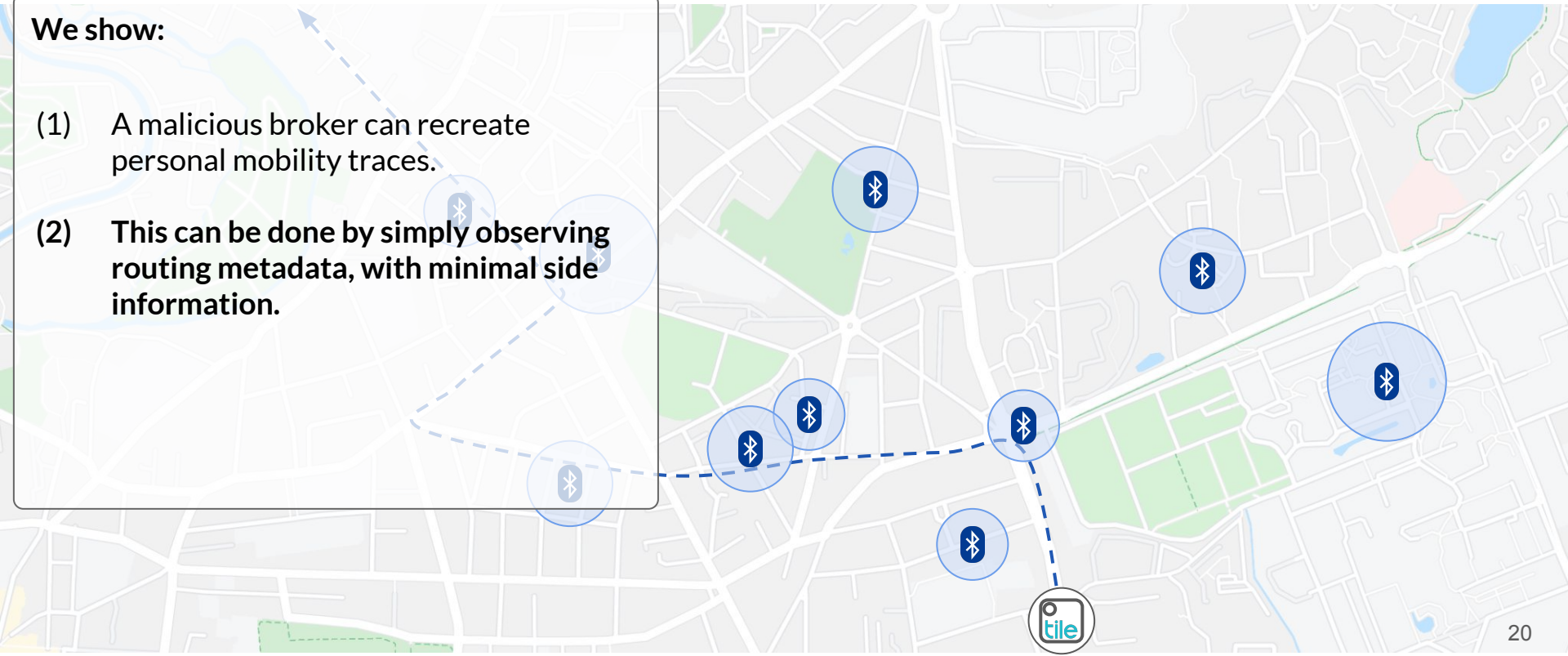


Is the same true in an opportunistic network?

We show:

(1) A malicious broker can recreate personal mobility traces.

(2) This can be done by simply observing routing metadata, with minimal side information.



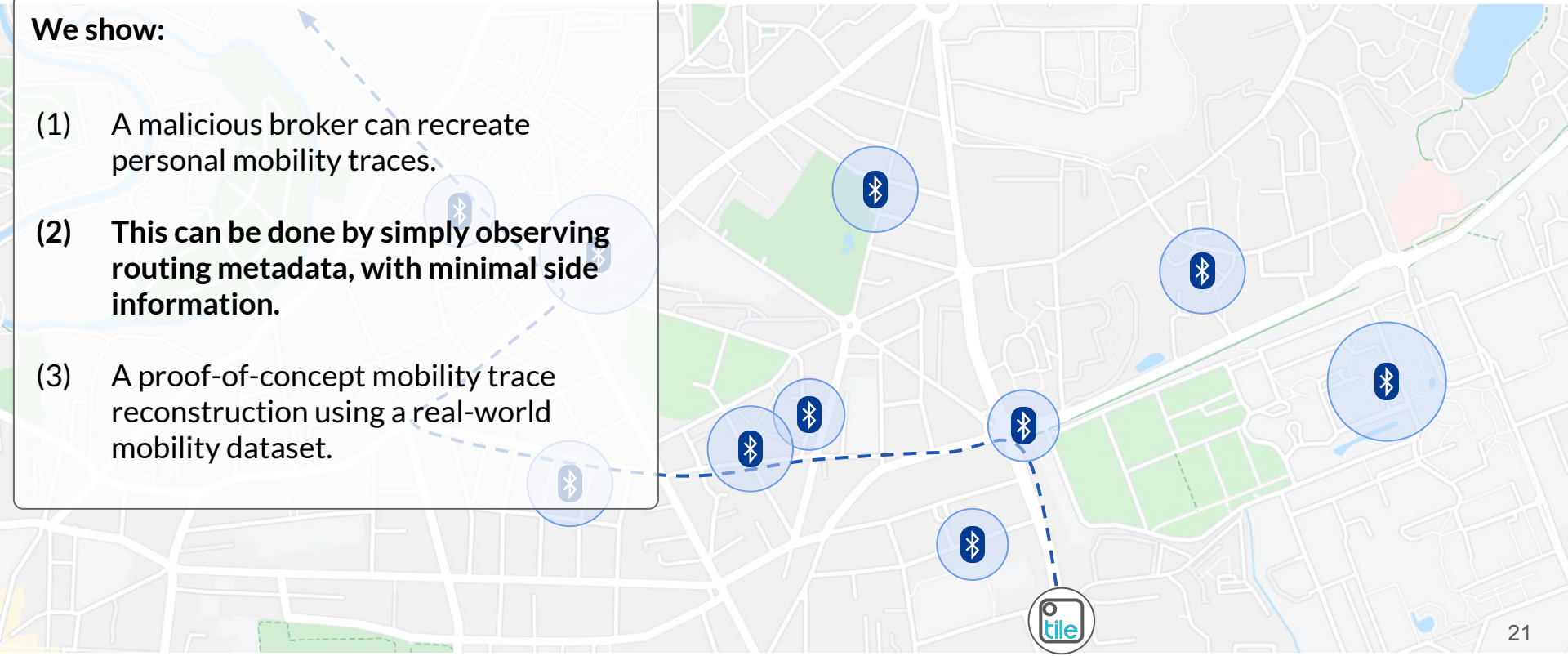
Is the same true in an opportunistic network?

We show:

(1) A malicious broker can recreate personal mobility traces.

(2) This can be done by simply observing routing metadata, with minimal side information.

(3) A proof-of-concept mobility trace reconstruction using a real-world mobility dataset.



Overview

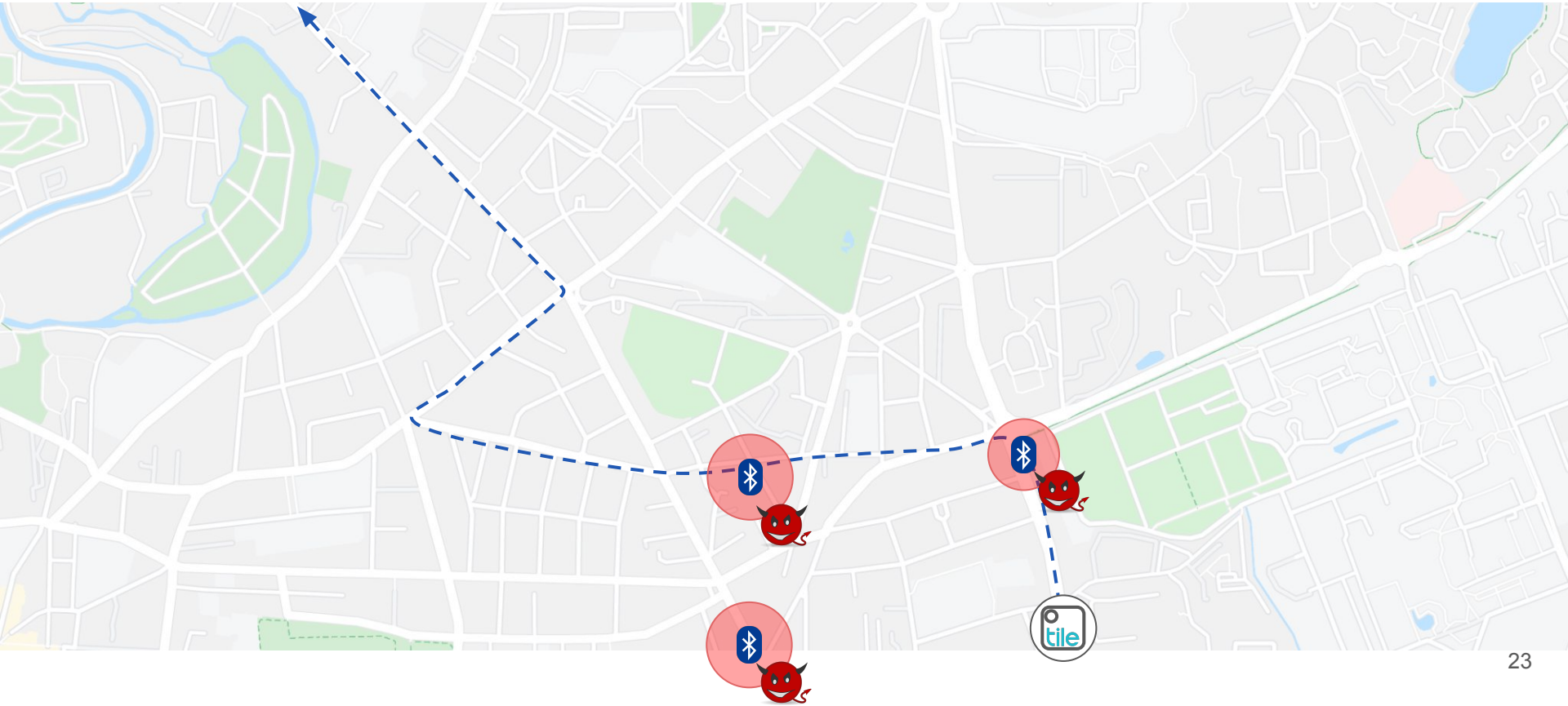
Introduction to opportunistic networks

Reconstructing mobility traces from routing metadata

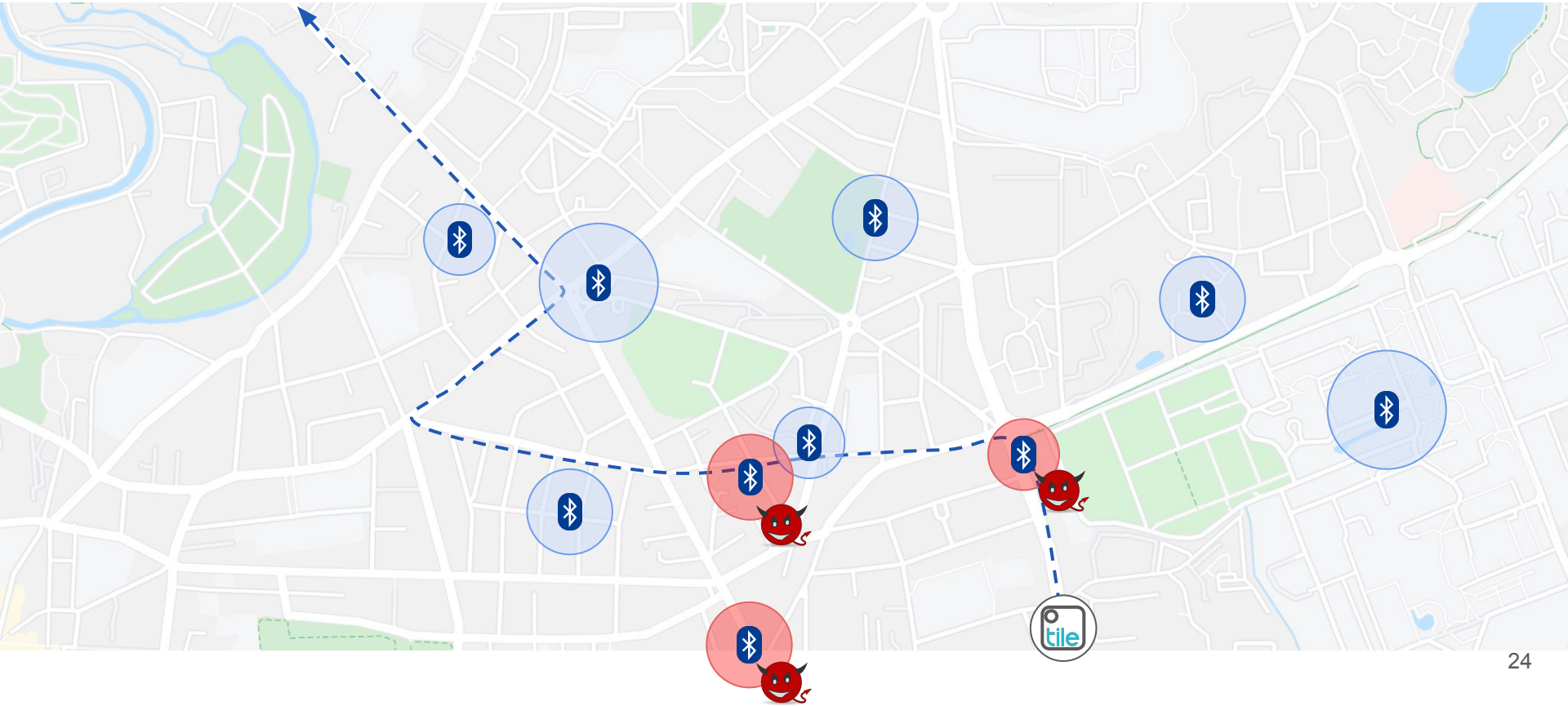
Simulation results

Future steps to addressing backhaul privacy

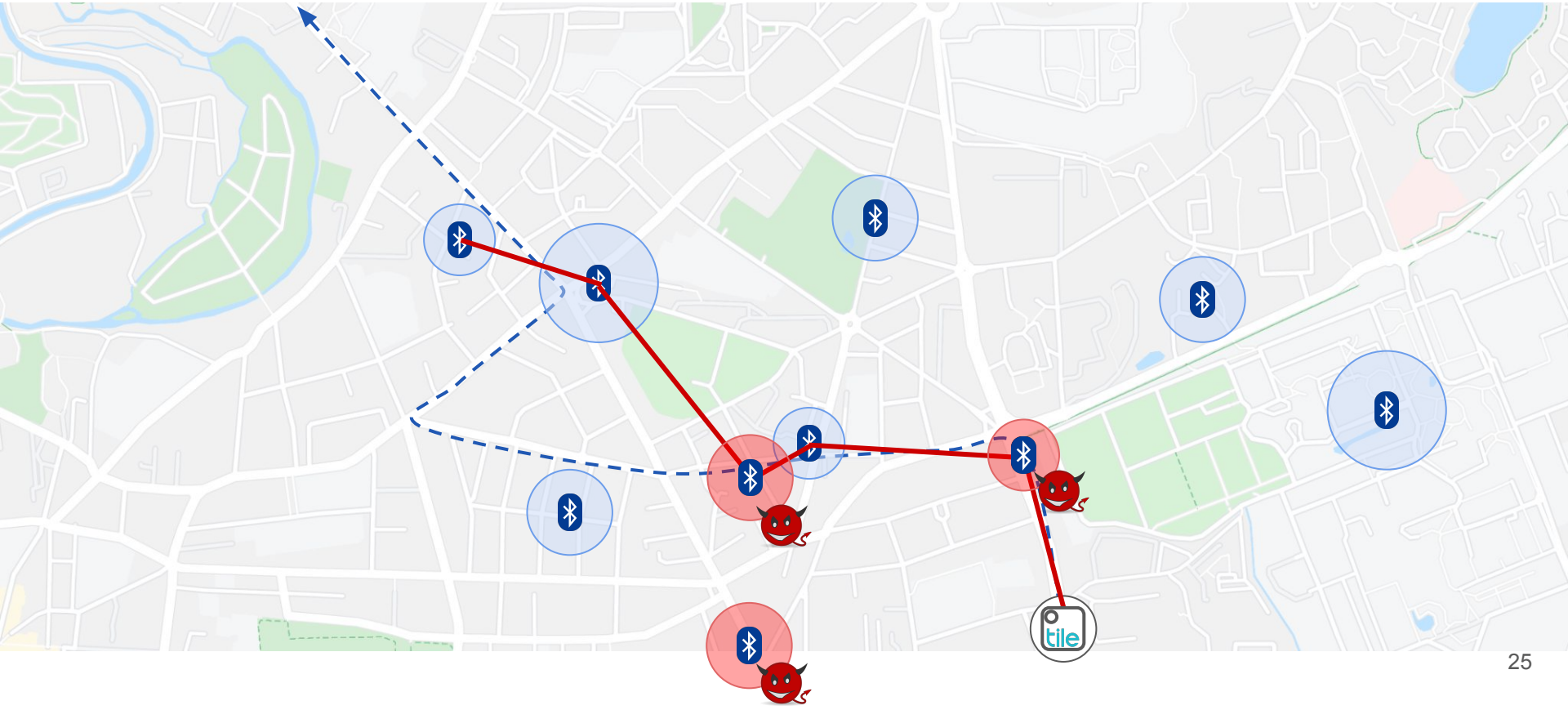
We reconstruct mobility traces using third-party gateways with minimal side-location information.



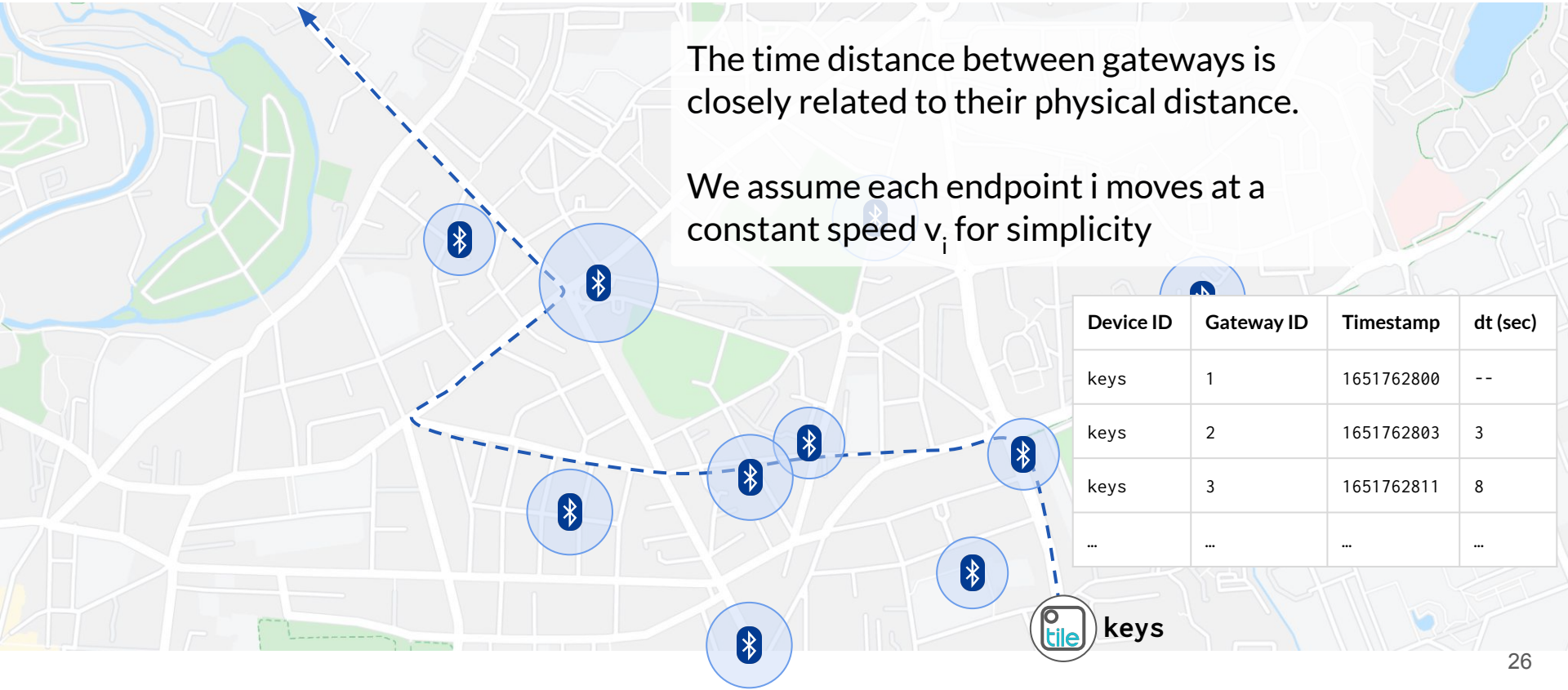
We reconstruct mobility traces using third-party gateways with minimal side-location information.



We reconstruct mobility traces using third-party gateways with minimal side-location information.



Each endpoint generates a physically-defined sequence of network interactions as it moves.



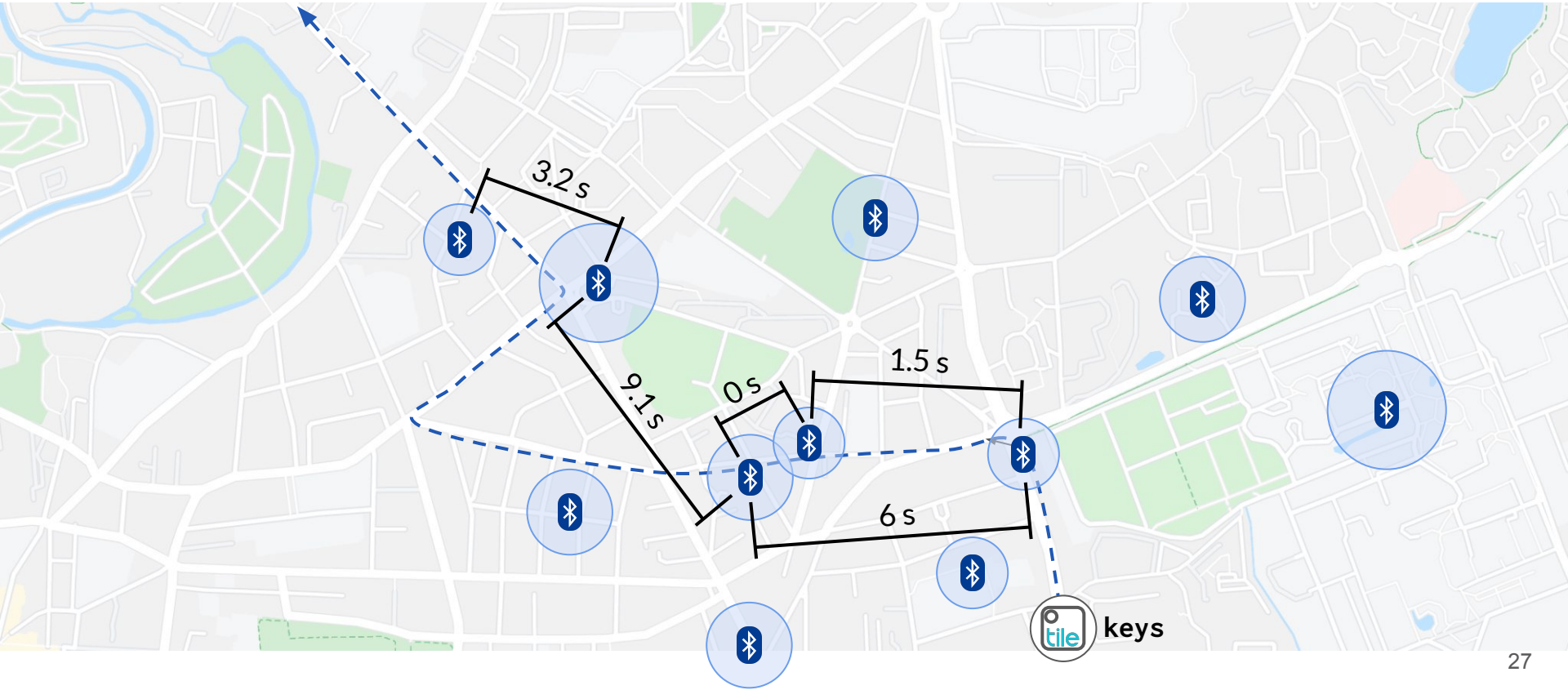
The time distance between gateways is closely related to their physical distance.

We assume each endpoint i moves at a constant speed v_i for simplicity

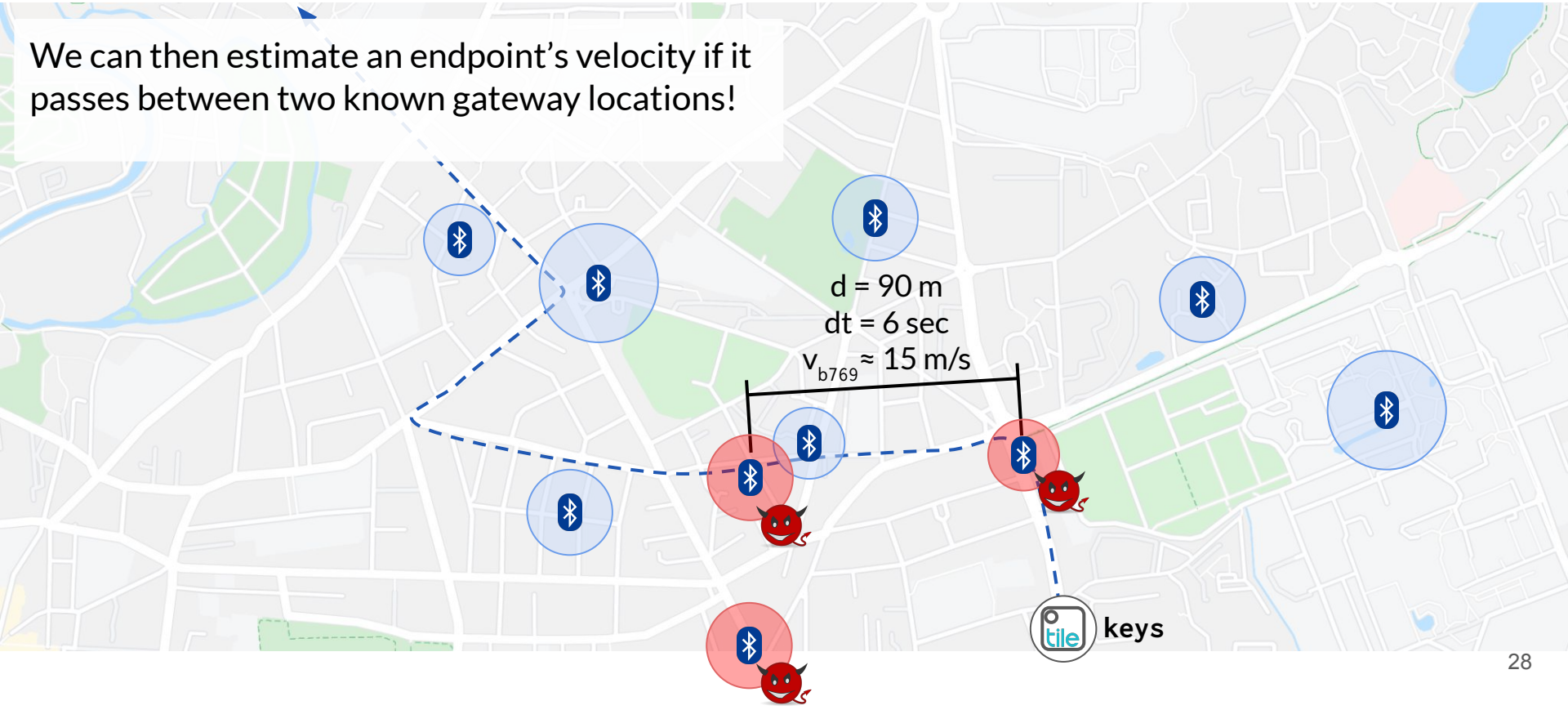
Device ID	Gateway ID	Timestamp	dt (sec)
keys	1	1651762800	--
keys	2	1651762803	3
keys	3	1651762811	8
…	…	…	…

keys

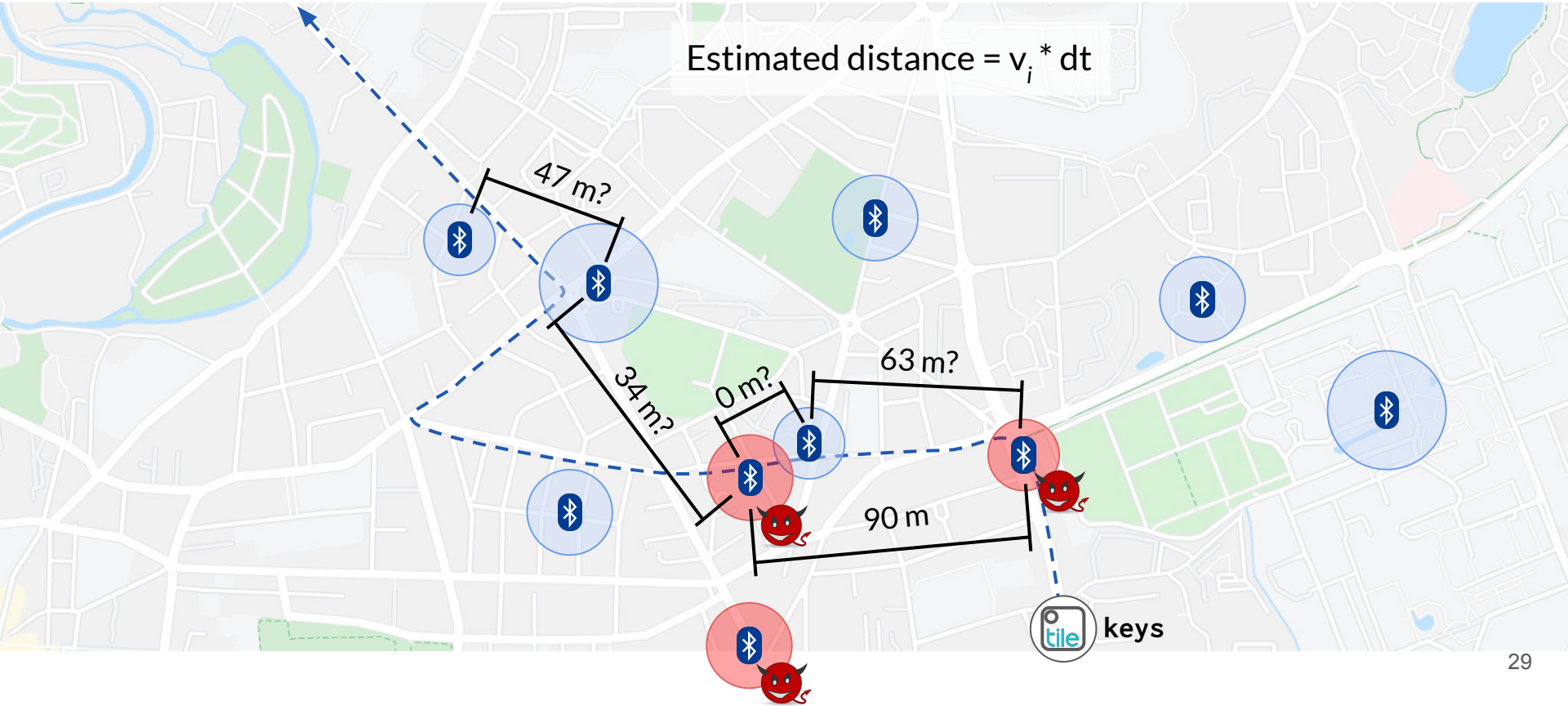
Differences in time correspond to relative distances, but do not immediately give absolute distances.



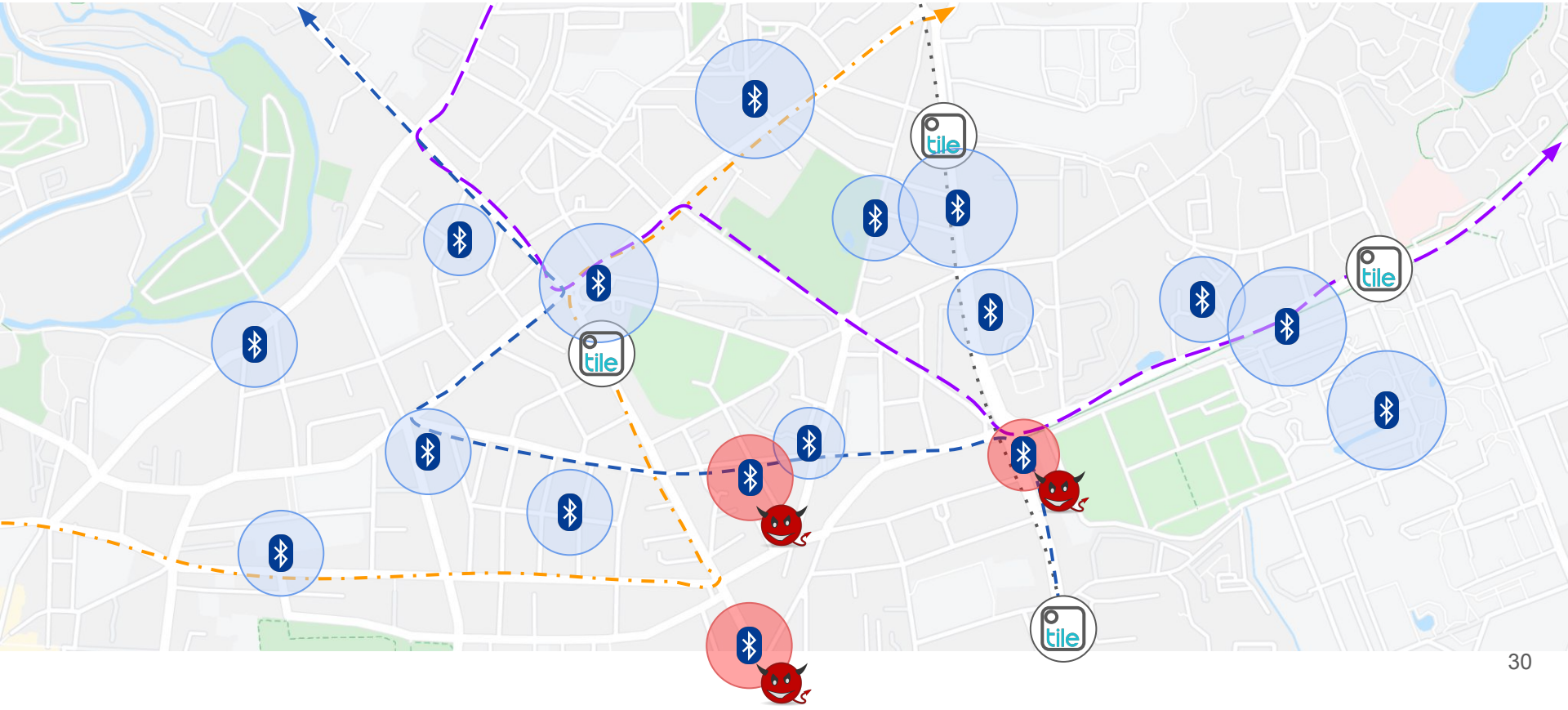
We require *some* information on gateway locations to infer absolute distances from time differences.



With this location information, we can estimate distances between gateways and then triangulate.



Accuracy increases with more network interactions.



Overview

Introduction to opportunistic networks

Reconstructing mobility traces from routing metadata

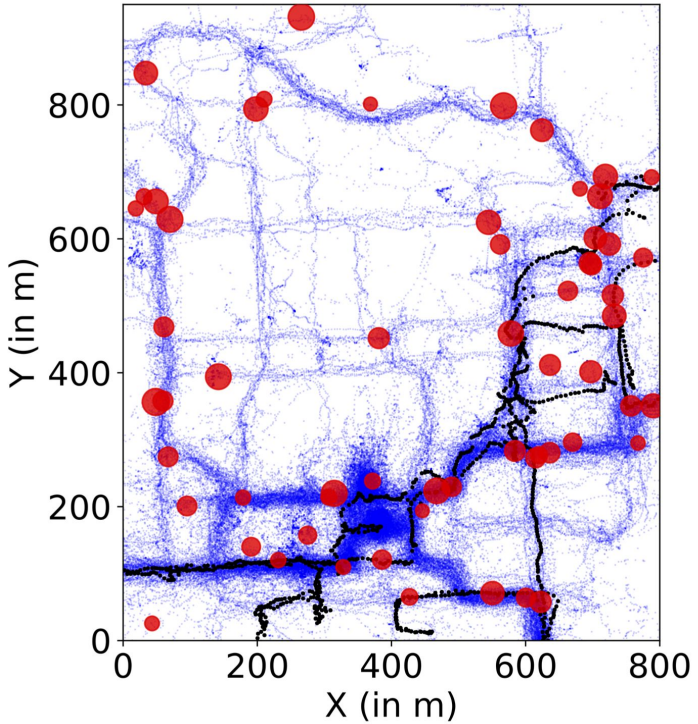
Simulation results

Future steps to addressing backhaul privacy

We simulate a deployment of stationary gateways and mobile endpoints and record their interactions.

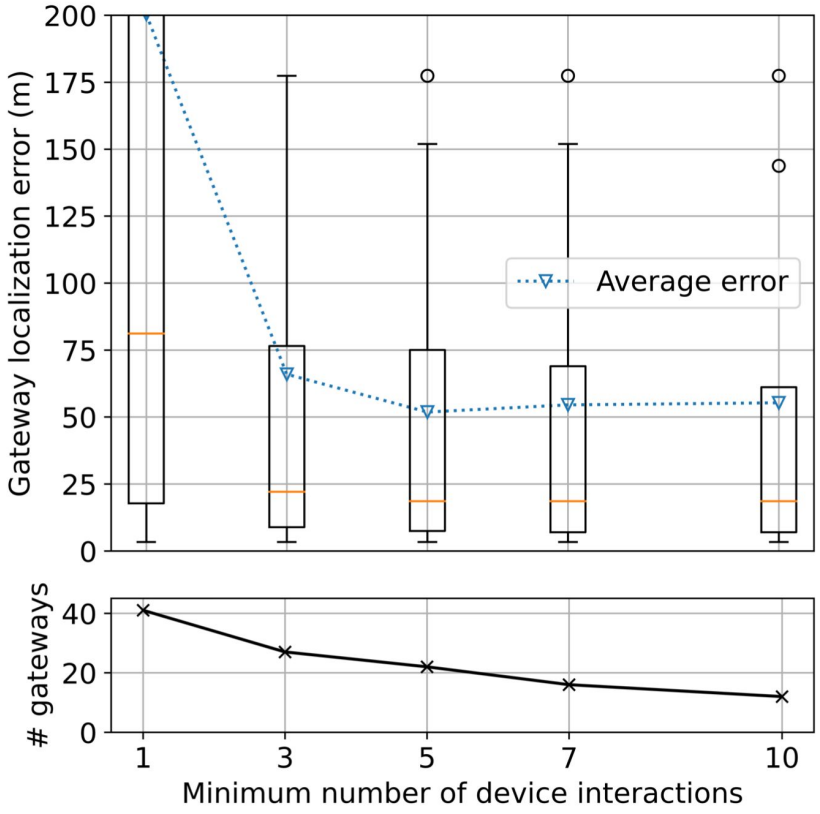
Simulation setup:

- GeoLife mobility traces (blue)
  - Example trace (black)
- Simulated BLE gateways (red)

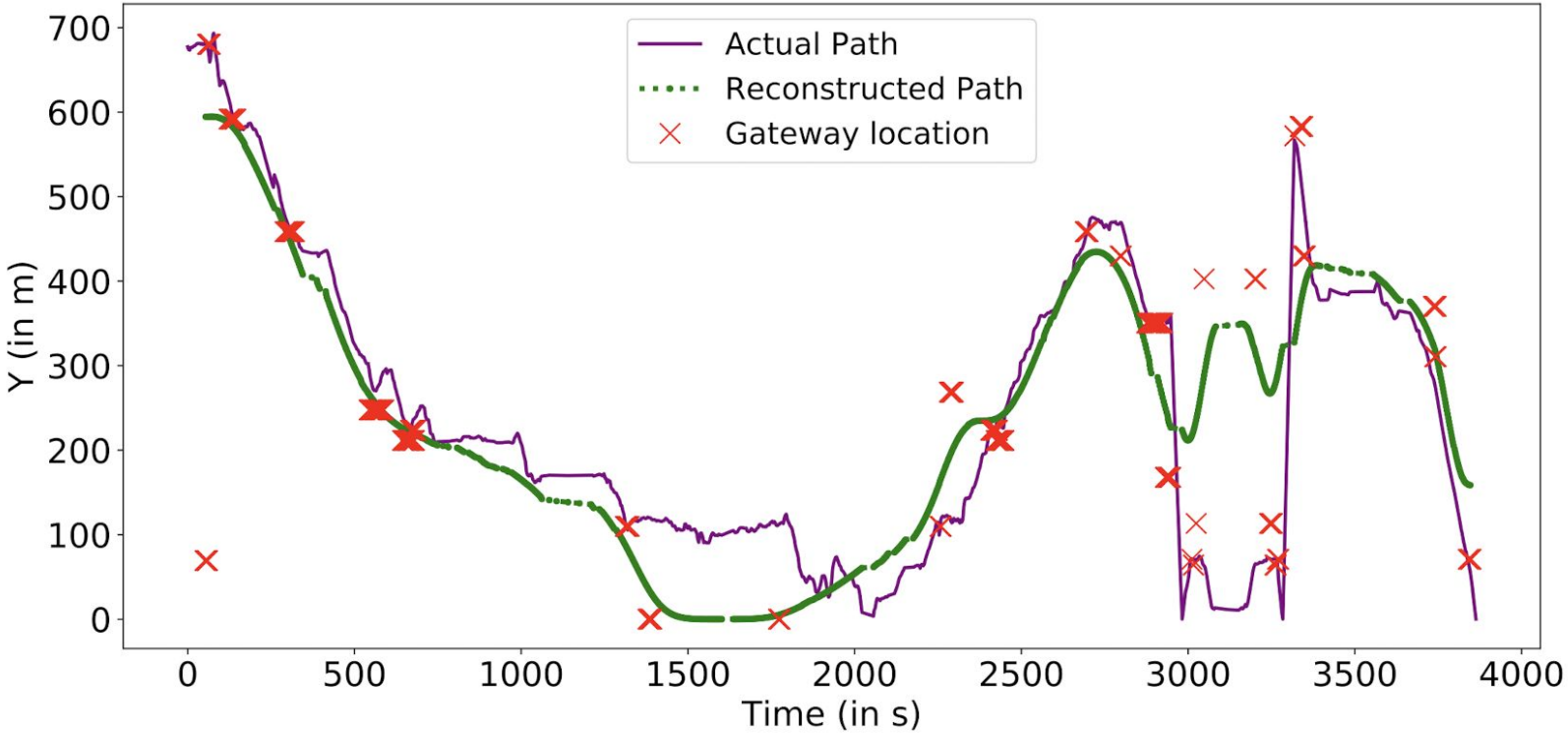


Gateways can be localized with ~50 m average error.

- Only need 3-5 interactions for reasonable accuracy
- Accuracies likely increase at scale
- Outliers likely diminish at scale



Mobility traces can be reconstructed with ~44 m average error.



Overview

Introduction to opportunistic networks

Reconstructing mobility traces from routing metadata

Simulation results

Future steps to addressing backhaul privacy

Core issue: metadata correlation

payload: **22e29...**
dest: **45d0...**
sensor: **A**
gateway: **B**
time: **100203...**

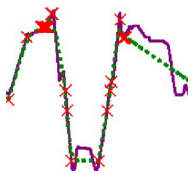
Alvin reads the data at **45d0....**

Core issue: metadata correlation

payload: **22e29...**
dest: **45d0...**
sensor: **A**
gateway: **B**
time: **100203...**

Alvin reads the data at **45d0....**

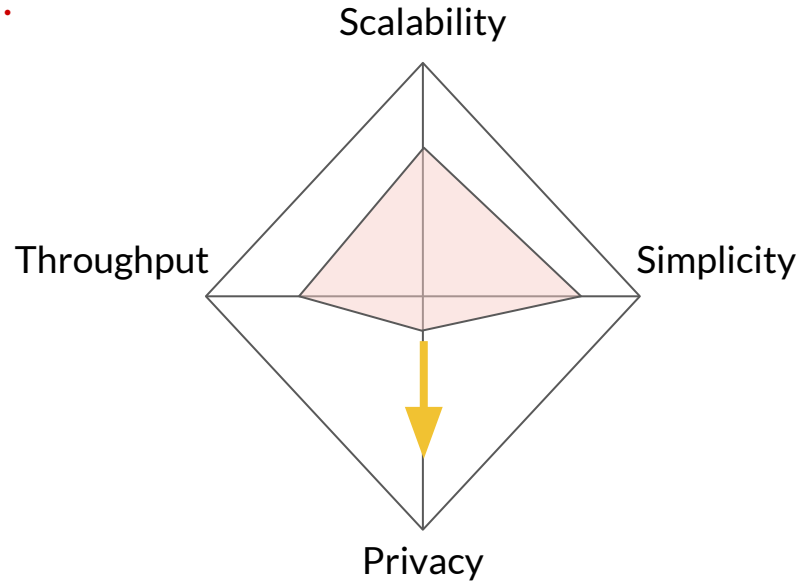
home and work address, medical visits, commute times and route …



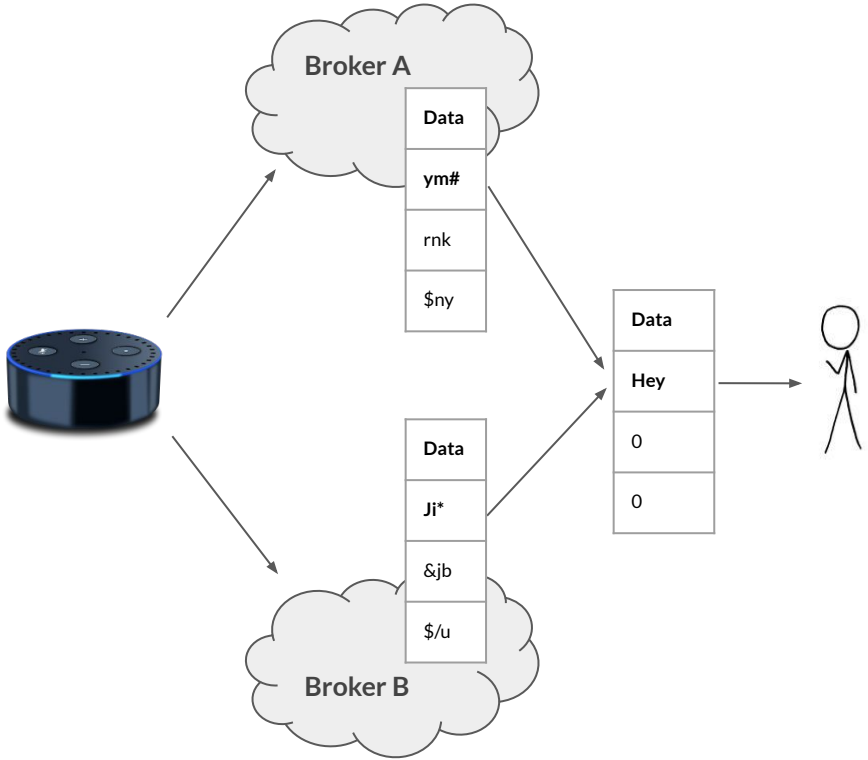
Core issue: metadata correlation

payload: **22e29...**
dest: **45d0...**
sensor: **A**
gateway: **B**
time: **100203...**

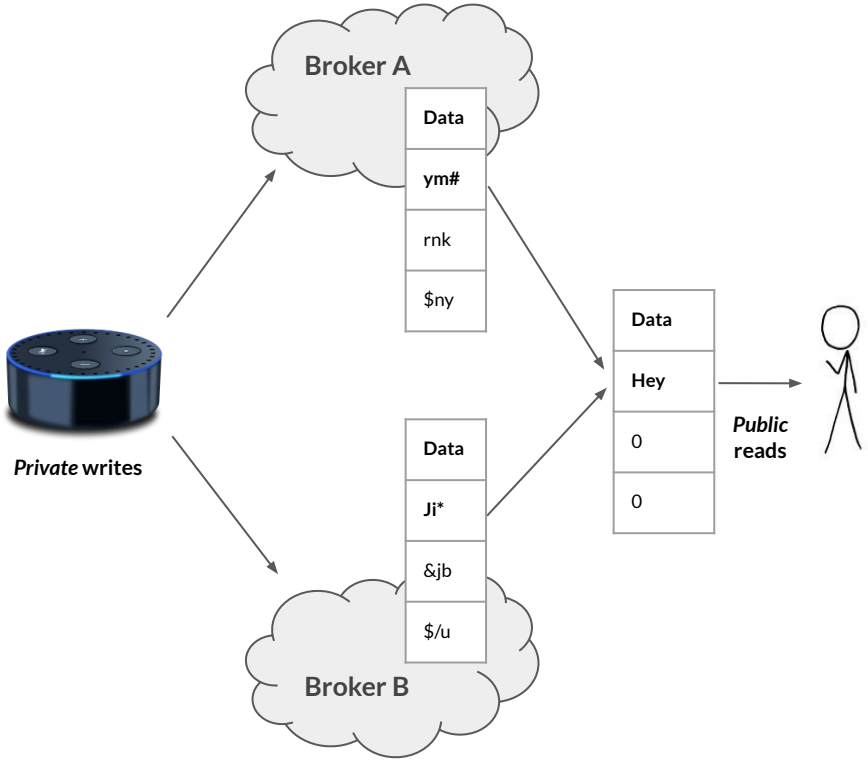
Alvin reads the data at **45d0....**



The Path Forward: Private Information Retrieval (PIR)



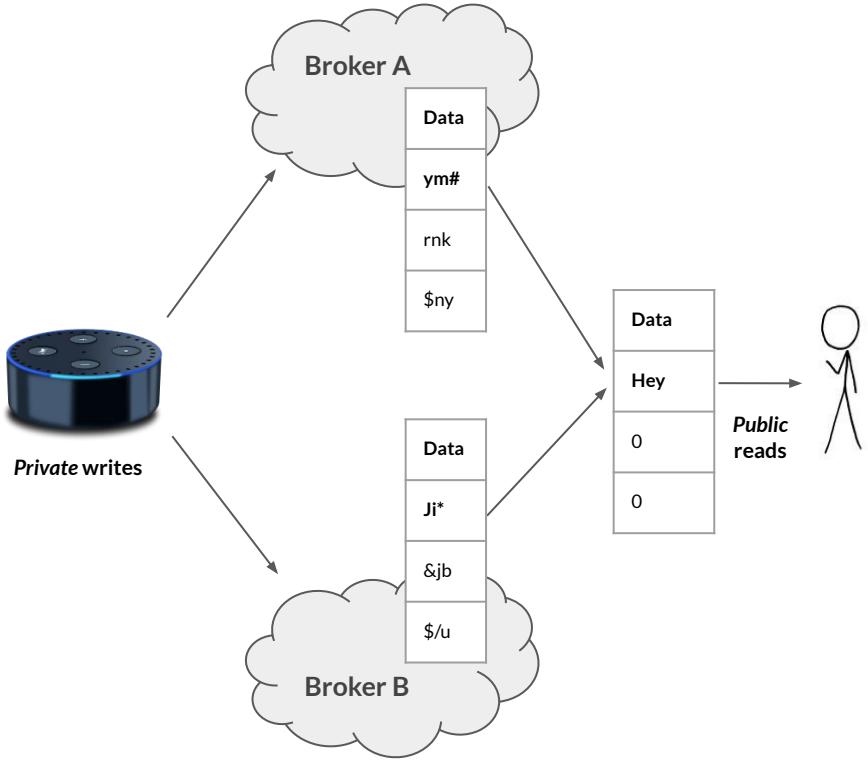
The Path Forward: Private Information Retrieval (PIR)



Breaks correlation link between gateways and individual payloads

- Gateways write secret-shared payloads to broker(s)
- Allows network to charge data consumers

The Path Forward: Private Information Retrieval (PIR)



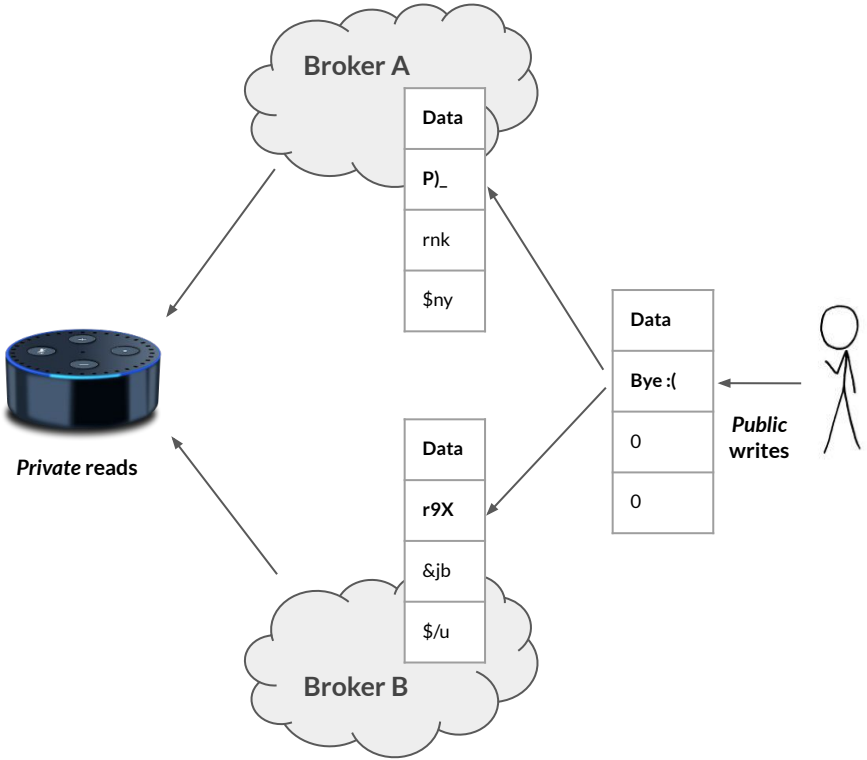
Breaks correlation link between gateways and individual payloads

- Gateways write secret-shared payloads to broker(s)
- Allows network to charge data consumers

Distributed trust assumption

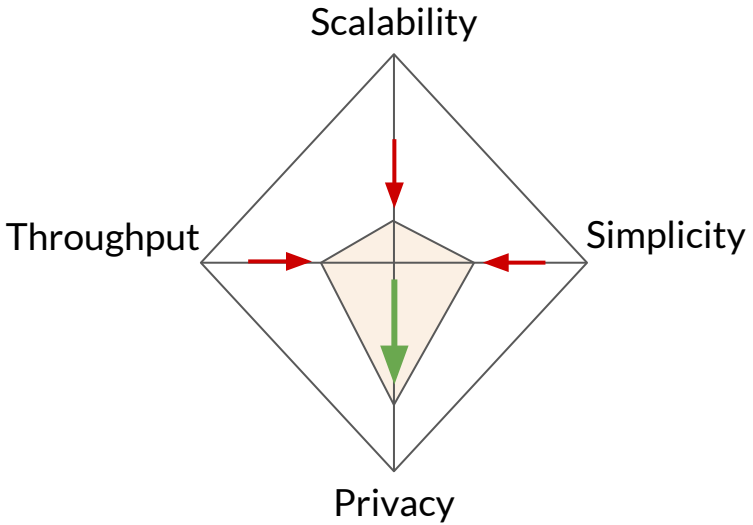
- Truly large-scale backhaul system requires cooperation between hardware manufacturers and broker operating network.

The Path Forward: Private Information Retrieval (PIR)

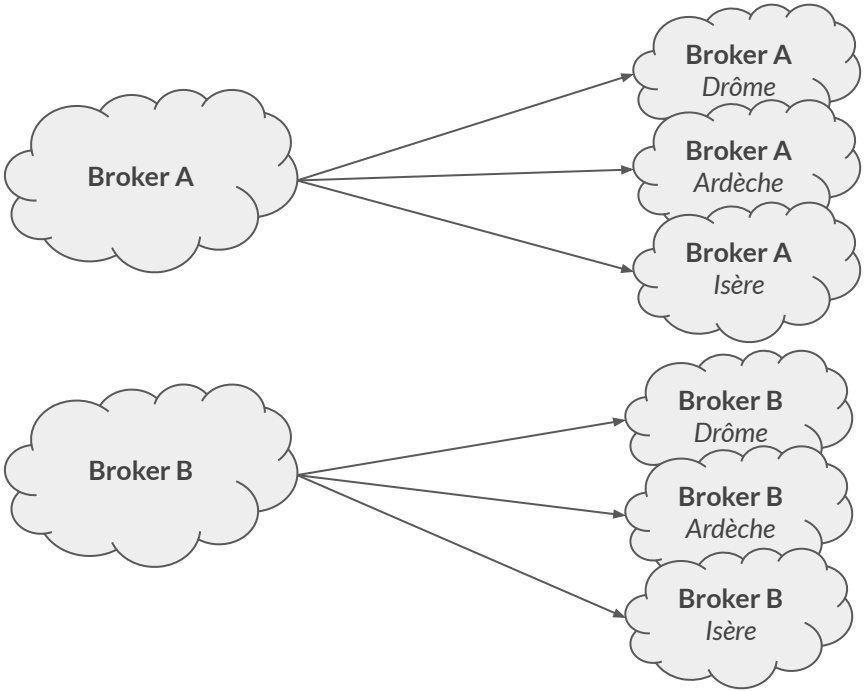


Bidirectional communication no longer leaks device-gateway proximity

- Gateways can *pull* payloads privately from brokers rather than receiving *pushed* data

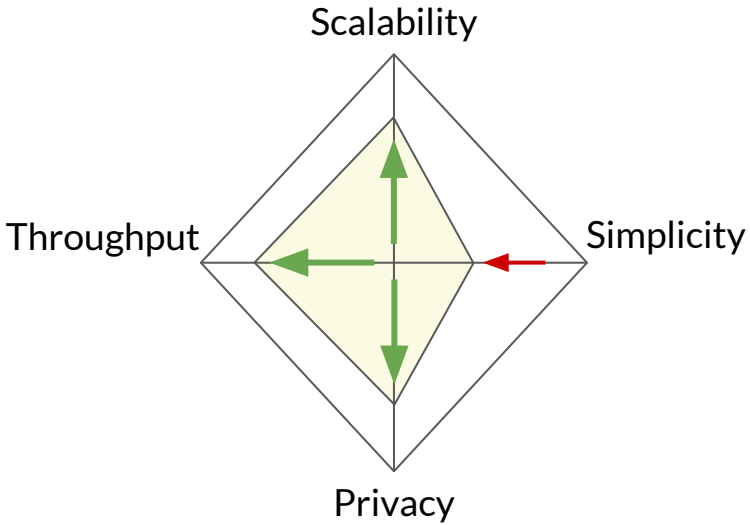


Addressing scalability

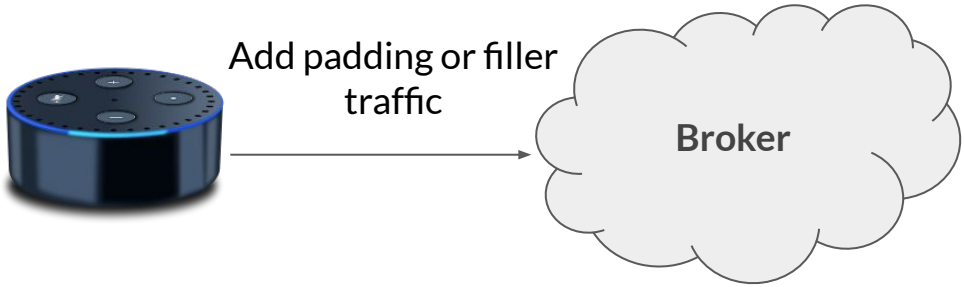


Global anonymity guarantees are unnecessary

- Gateways simply need to provide privacy on a regional/urban level

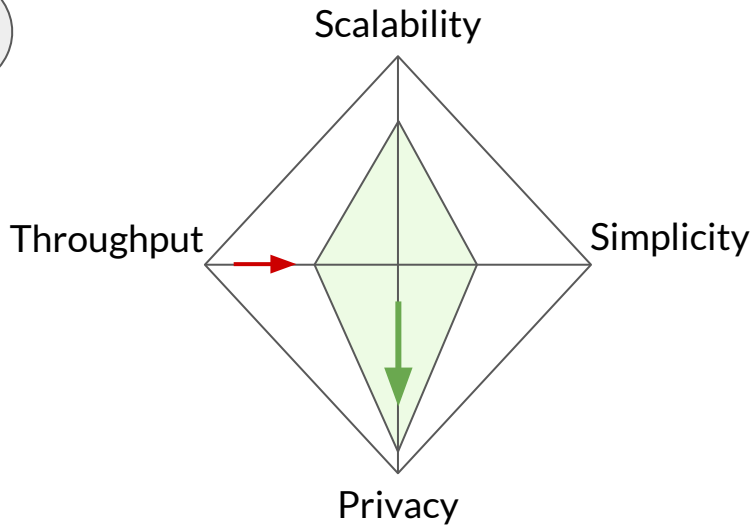


Strict(er) location privacy

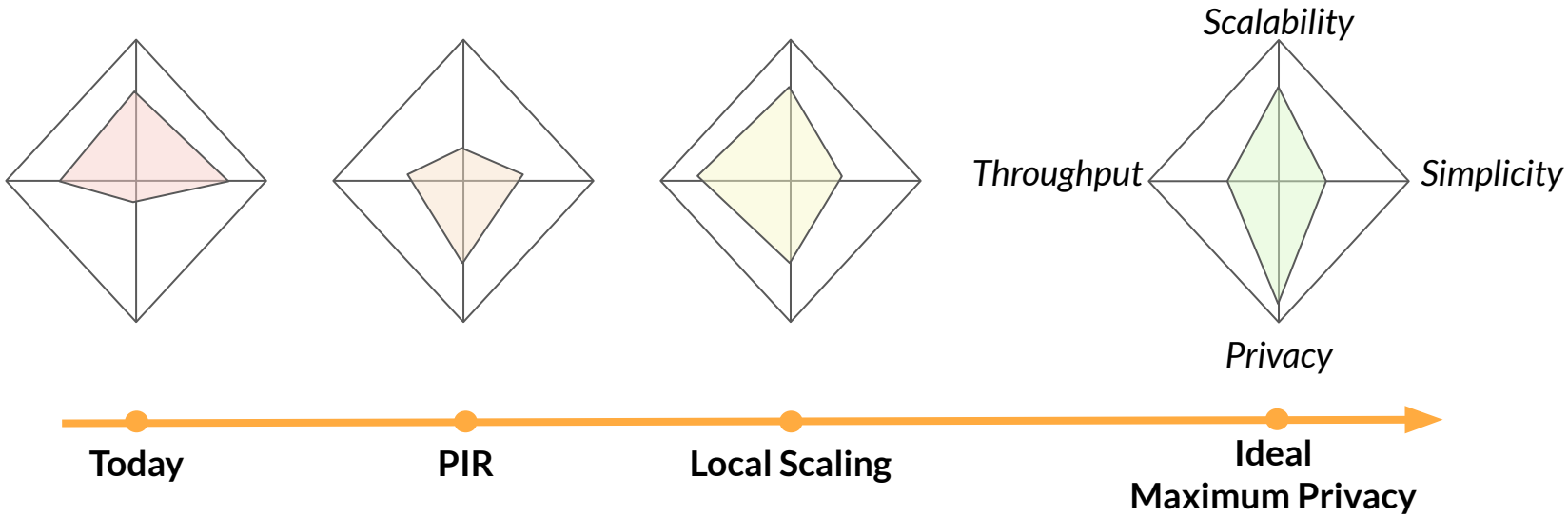


Private access alone can't hide timing side channels

- Although difficult, a broker could fingerprint locations by write patterns.



Where the Sidewalk Ends: Privacy of Opportunistic Backhaul



Tess Despres, Shishir Patil, Alvin Tan, Jean-Luc Watson, Prabal Dutta
jlw@berkeley.edu