Hidden Figures

Fingerprinting Mobile Phones Through Passive Background App Traffic

Shiva Azizzadeh, Gunes Acar

Radboud Universiteit

One traveler, three open networks

07:58

onboard Wi-Fi, open
56 min on the network

12:34

guest Wi-Fi, open
31 min on the network

18:40

free Wi-Fi, open
63 min on the network

Three operators. Three randomized MAC addresses. Nothing that ties them together.

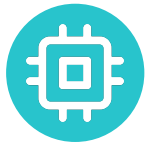
Fingerprinting

Actions & apps

Encrypted traffic still reveals user actions and the app.

Routines

Daily activity patterns can be inferred from metadata.



Devices

Transmission characteristics can survive the encryption.

Almost every one of these attacks needs the user to do something.

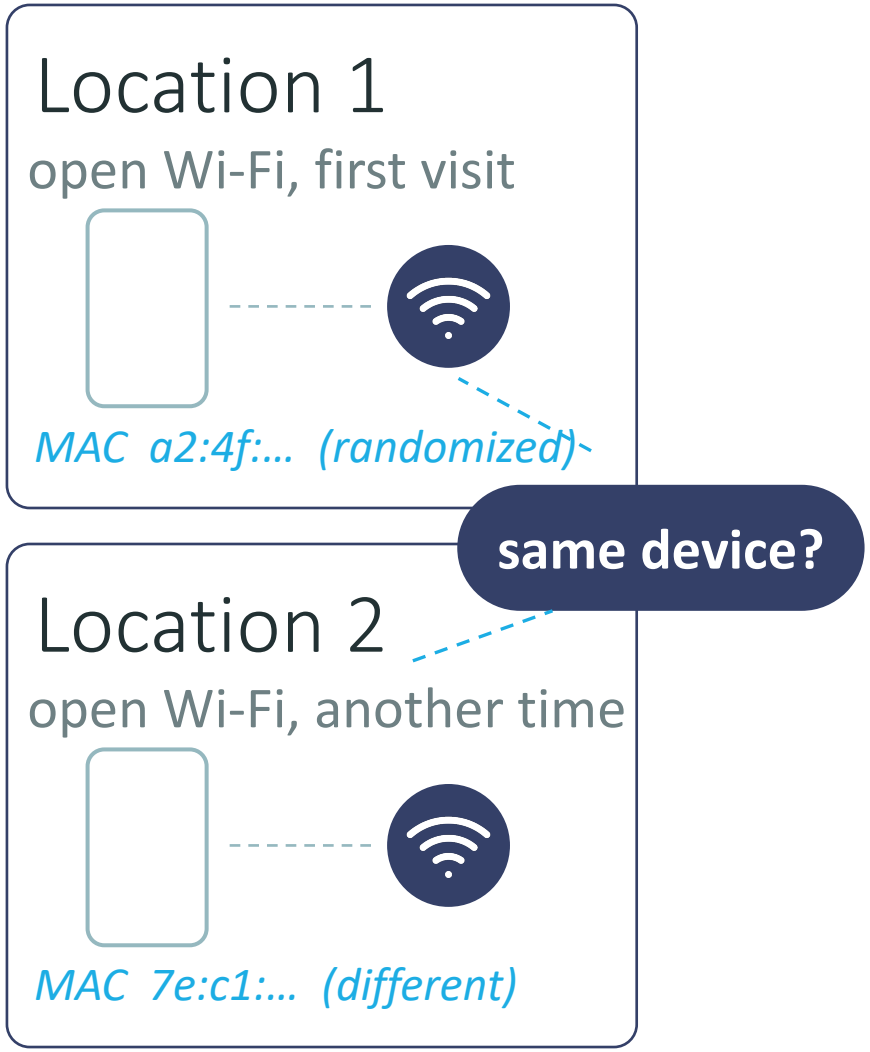
Research Question

Can two captures, recorded at different times and places, be linked to the same phone?

No Payload Content

No Persistent Identifiers

No User Interaction



Attack Setting

The attacker:

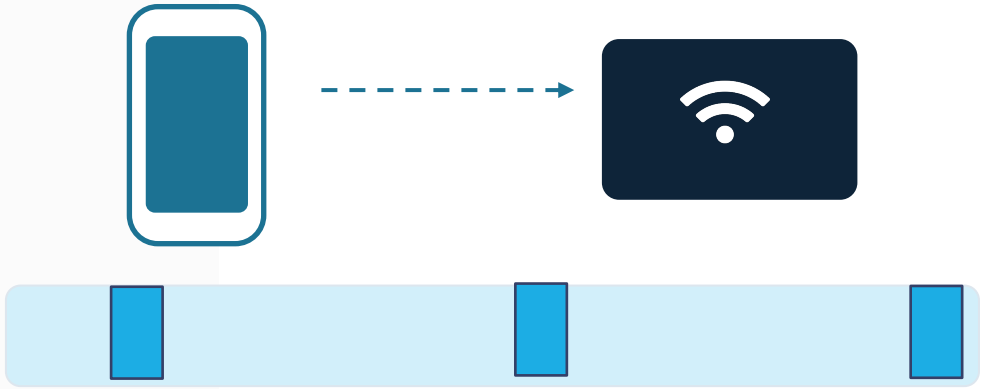
- Sniffs an open network or runs a rogue AP.
- Uses network-layer metadata only.

Our implementation needs:

- Generating realistic user profiles
- Phone connected to the access point

User Profiles

- Google Play Store popular apps and categories
- Random selection from weighted categories
- Random number of apps (mean=75)
- Average overlap of 16 apps between profiles



40p * 10r * 1h

POCO F2 Pro

Samsung Galaxy A14 5G

- Factory reset and disable pre-installed apps
- Enable current profile’s apps via adb
- Launch each app once so it completes a cold start
- Close everything, screen off, record for one hour

The Representation of a Trace

Raw capture

Every packet reaching the AP during the capturing window

Endpoints

Each destination IP resolved to its AS via the MaxMind database

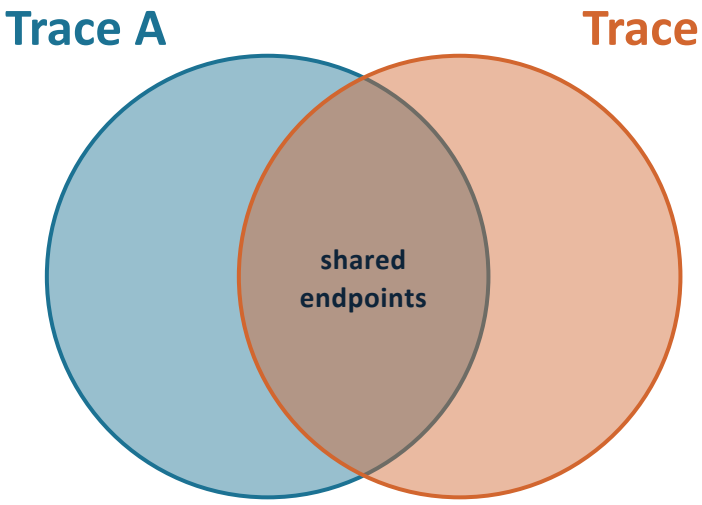
Filtered set

Filtered Ases considering the volume and uniqueness

Final representation

Unique ASes stay, non-unique ones are combined with the ports

Comparison of Two Traces



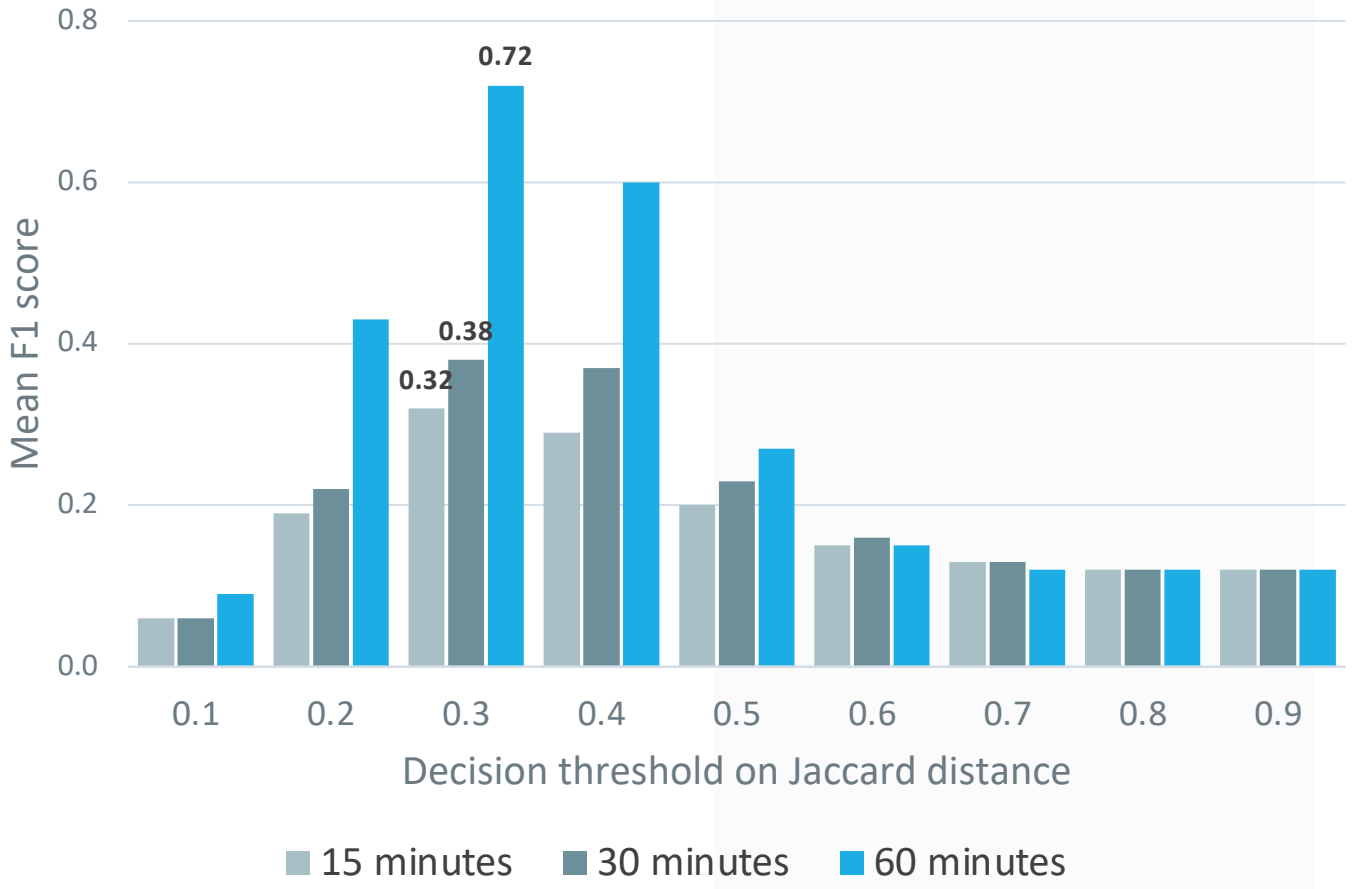
Jaccard distance = 1 − |A ∩ B| / |A ∪ B|

AS1	AS4	AS3:5222	AS6:123	AS9	AS11
AS1	AS4	AS3:5222	AS6:8883	AS9	AS11

AS1	AS3:5222	AS4	AS6:123	AS6:8883	AS9	AS11
1	1	1	1	0	1	1
1	1	1	0	1	1	1

|A ∩ B| = 5 |A ∪ B| = 7 Jac.Dist = 0.29

Decision Threshold



High and low threshold:

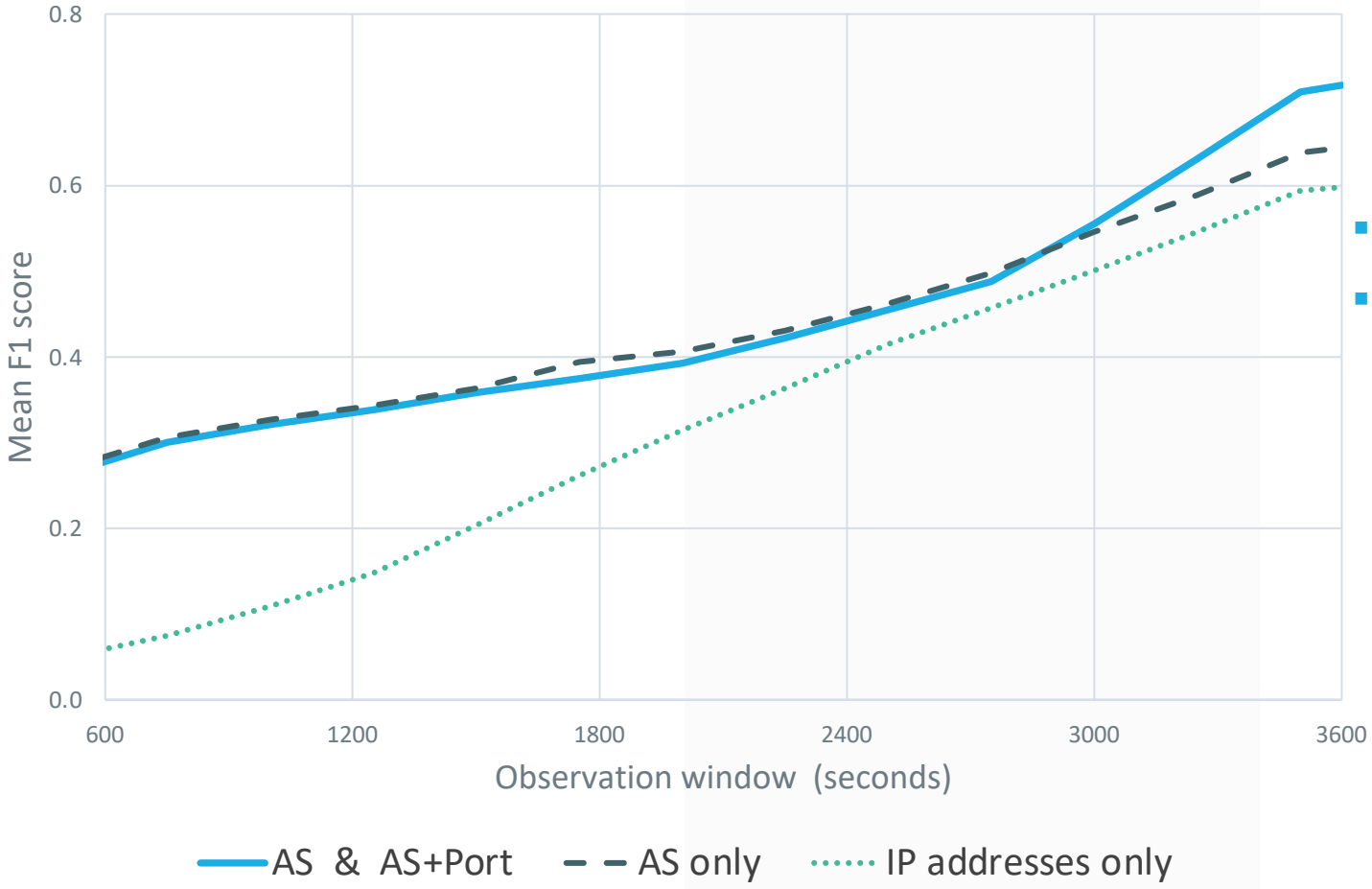
- Too low, too strict:

  Real matches rejected
- Too high: too relaxed:

  Unrelated traces accepted

Chosen threshold = 0.3

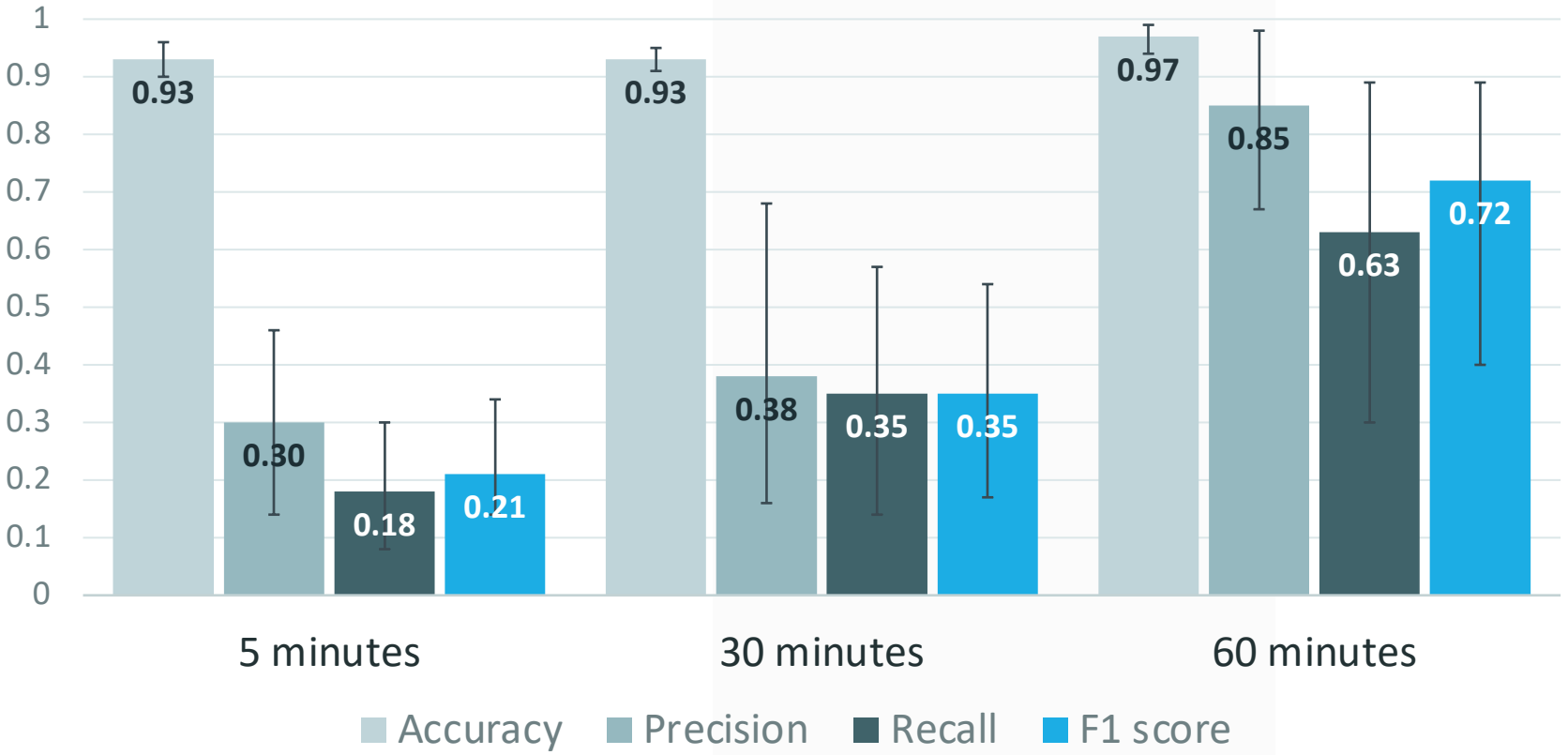
Capture Duration



Higher success rate when:

- Using AS+Port combination
- Longer capture

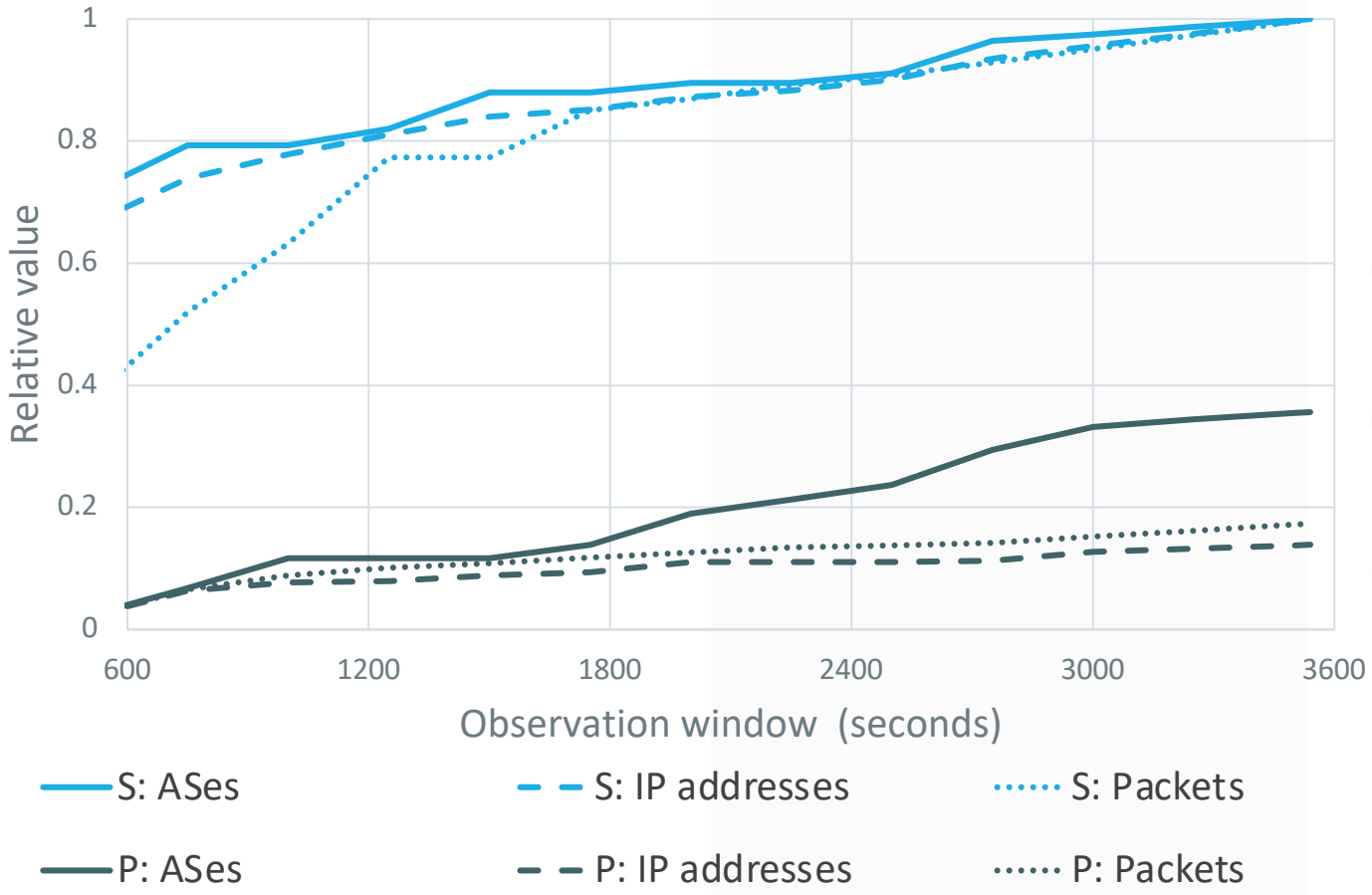
Evaluation Results



Poco Phone
Android 10

≈ 141 related pairs
≈ 2 360 unrelated pairs

Poco vs Samsung



Higher success rate in Samsung

- More background traffic
- More distinct endpoints in a trace
- Longer capture

Limitations

- Link-layer encryption

  The attack cannot run on a WPA2- or WPA3-protected network; but a large share of public Wi-Fi is still open.

- Number of generated profiles

  With millions of apps available, many real app sets are likely unique.

- Number of tested phones and OSes

  Shared Android behavior suggests the result carries over, but iOS is not tested.

Conclusion

Metadata alone is enough to distinguish the device and link the traces **across locations.**

No interaction needed

Time affects the success

Not tied to one phone

Thank you!

Questions?

Shiva.Azizzadeh@ru.nl