Poirot: Private Contact Summary Aggregation

Chenghong Wang[1], David Pujol[1], Yanping Zhang[1], Johes Bater[1], Matthew Lentz[1,4], Ashwin Machanavajjhala[1], Kartik Nayak[1], Lavanya Vasudevan[1,2,3] and Jun Yang[1]

[1]Computer Science Department, Duke University

[2]Department of Family Medicine and Community Health, Duke University

[3]Duke Global Health Institute

[4] VMware Research

Poirot: In a Nutshell

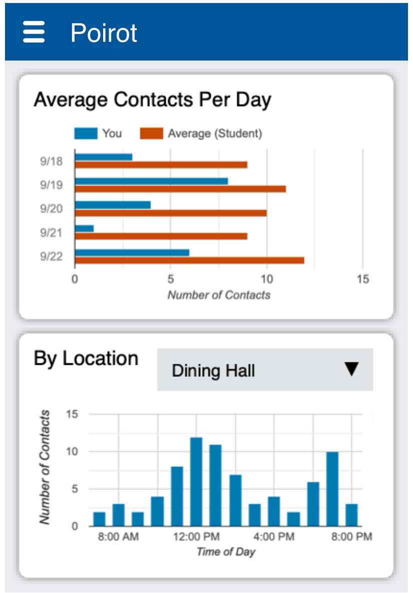
Physical distancing between individuals is key to preventing the spread of a disease such as COVID-19

We want:

- Functionality: Measure physical interactions through “contact events”
- Privacy: Ensure that the resulting data cannot be linked back to an individual

How will Poirot be used?

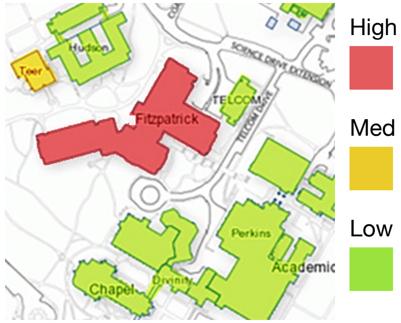
Individuals

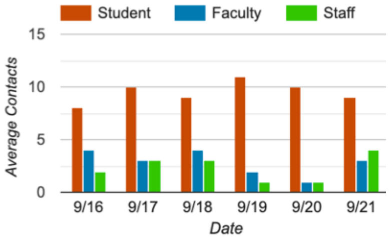


How many contacts do I have on a daily basis?

When is it safest for me to visit a given building?

Administrators





Which buildings require policy changes?

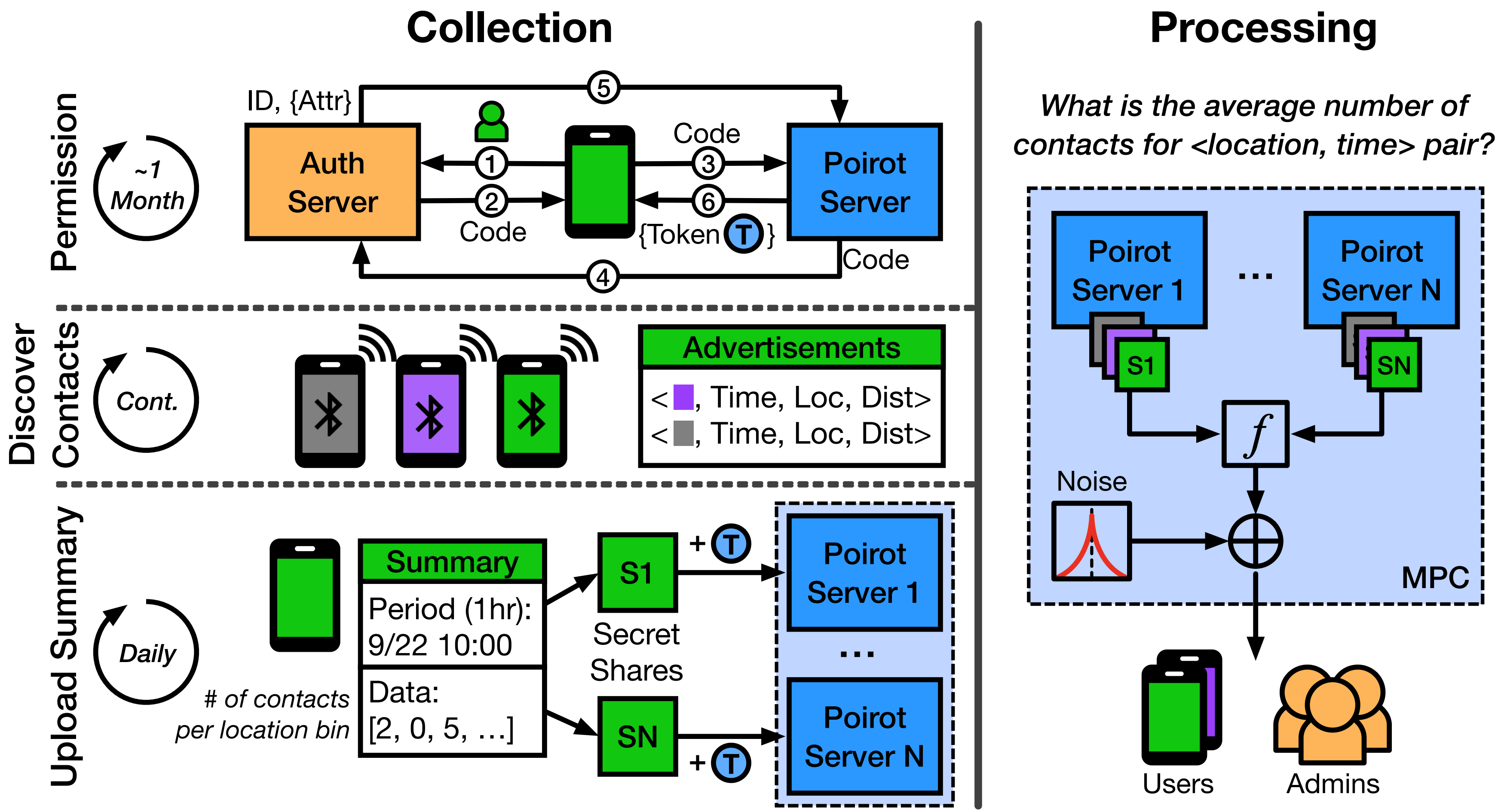
Are certain groups at higher risk?

Provide actionable information to individual users and decision makers in a privacy-preserving manner.

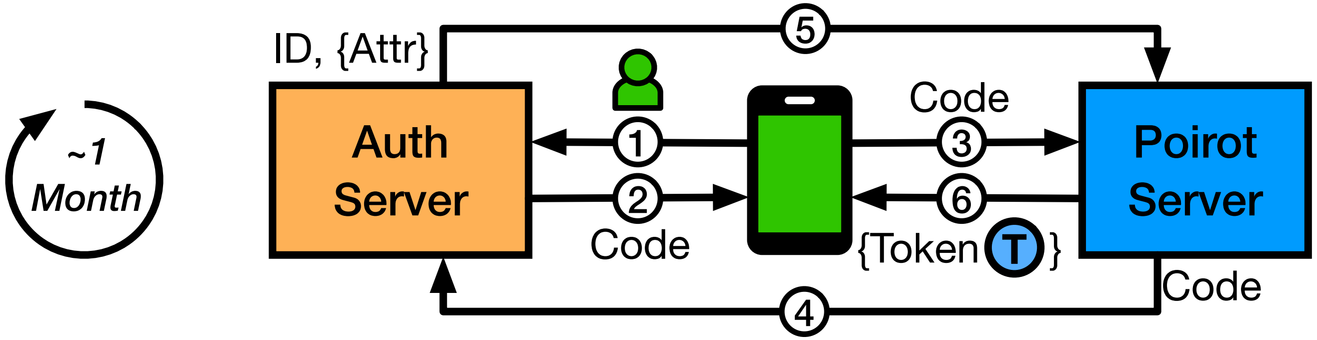
Threat Model

Users	semi-honest	Learns their own #contacts with locations and times plus differentially-private aggregated statistics
Admins	untrusted	Untrusted administrators: learn differentially-private aggregate statistics
Auth Server	semi-honest	Learns the set of participating users
Poirot Server N	semi-honest, assume non-collusion	Learns the set of participating users + some metadata.

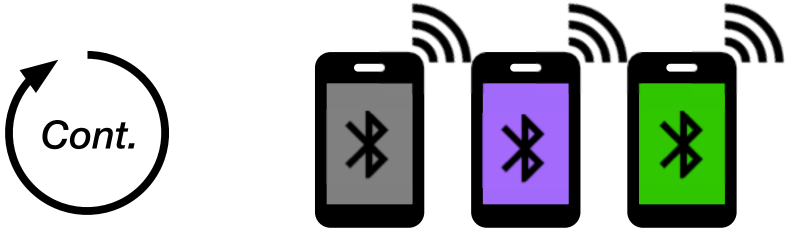
Poirot Design



Poirot Design -> Data Collection-> Private Permissioning

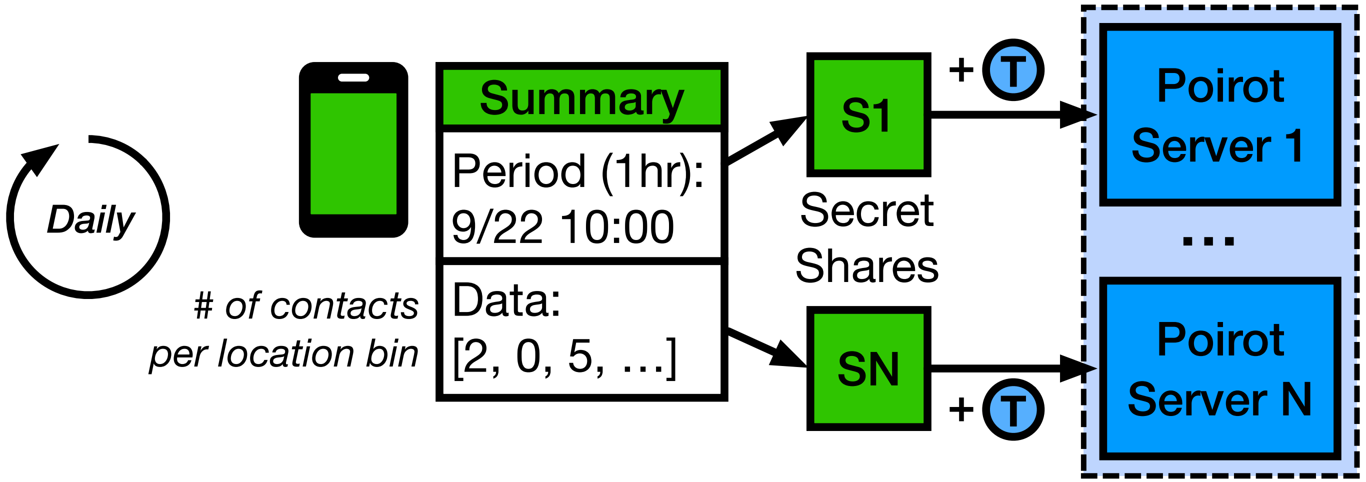


Poirot Design-> Data Collection-> Discover Contacts



Advertisements
< 🟪, Time, Loc, Dist>
< ◼, Time, Loc, Dist>

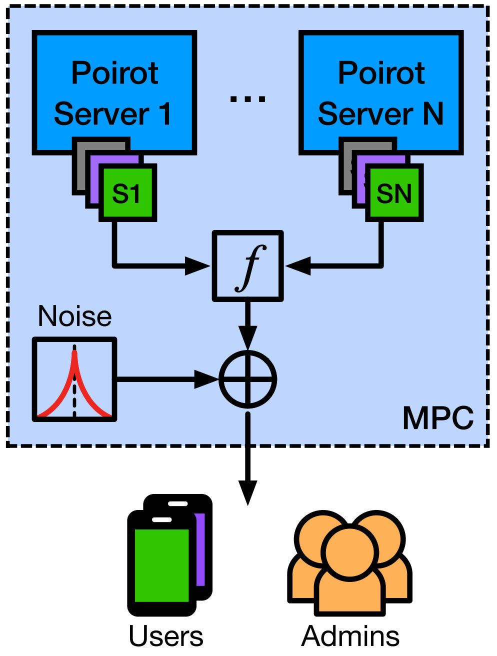
Poirot Design -> Data Collection-> Upload Summary



Servers only learn metadata about contact summaries

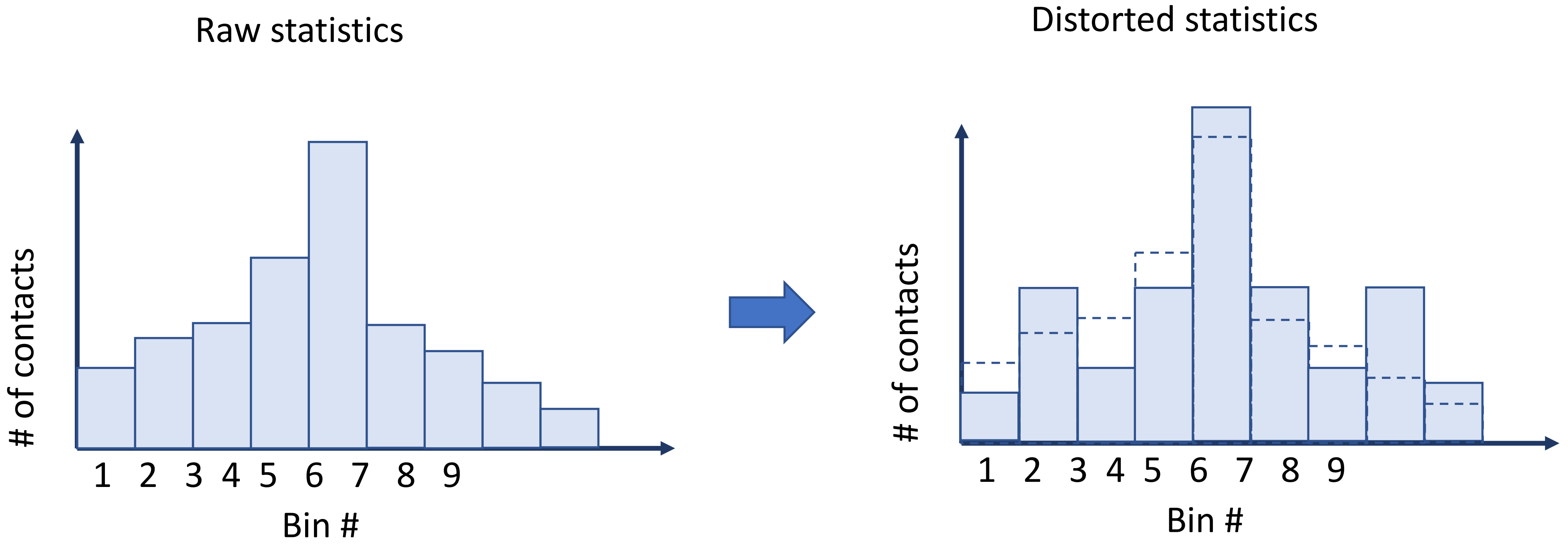
Poirot Design-> Data Processing

What is the average number of contacts for <location, time> pair?



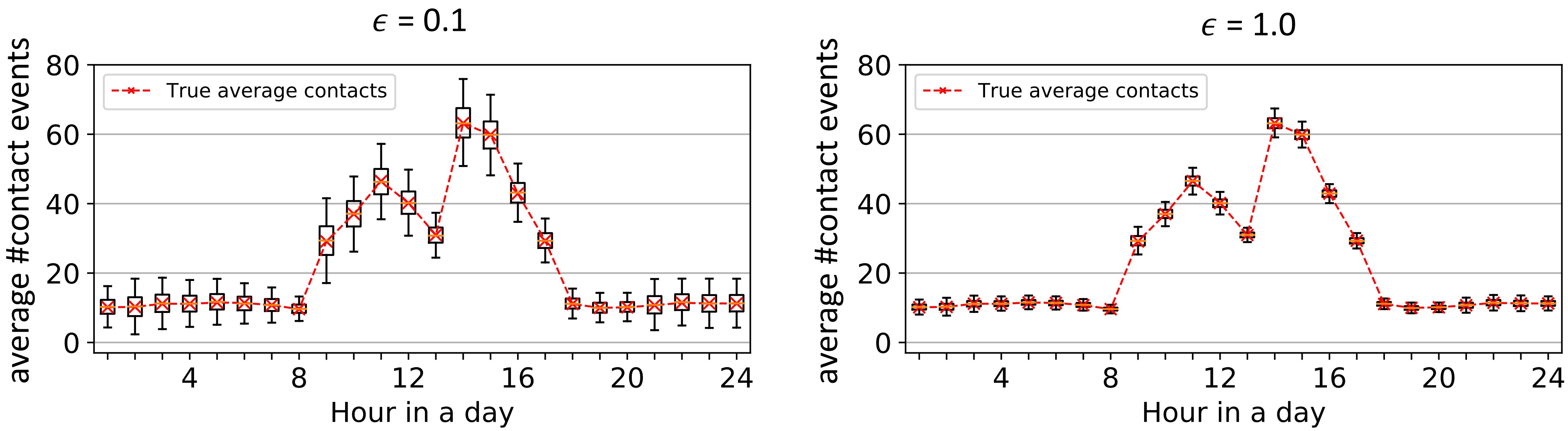
Multiparty Computation (MPC) allows computing on secret-shared data, Differential Privacy ensures released statistics do not reveal individual’s data

Poirot Design-> Data Processing



Poirot-> Evaluation-Accuracy

- Dataset: Copenhagen Network Study dataset



Poirot-> Evaluation-Performance

Case	# of Locations	Time	User Population	App execution time (ms)	Server execution time (s)
Duke	256	24	20K	366.1	94.4
NC	100	1	10M	6.0	776.1
Copenhagen	1	24	705	1.68	0.015

Conclusion

- Provide accurate information about physical interactions.
- Guarantees individual’s contact privacy
- Our system scales to large, realistic deployment scenarios.

https://poirot.cs.duke.edu/

Duke
UNIVERSITY