Reconstruction Attack on Differential Private Trajectory Protection Mechanisms

ACSAC'22
Erik Buchholz, *UNSW Sydney & CSCRC*
Alsharif Abuadbba, *CSIRO's Data61 & CSCRC*
Shuo Wang, *CSIRO's Data61 & CSCRC*
Surya Nepal, *CSIRO's Data61 & CSCRC*
Salil S. Kanhere, *UNSW Sydney & CSCRC*

**CSCRC = Cyber Security Cooperative Research Centre*

Acknowledgement

The authors would like to thank **UNSW**, the **Commonwealth of Australia**, and the **Cybersecurity Cooperative Research Centre Limited** for their support.

Annual Computer Security Applications Conference (ACSAC)
December 5 – 9, 2022

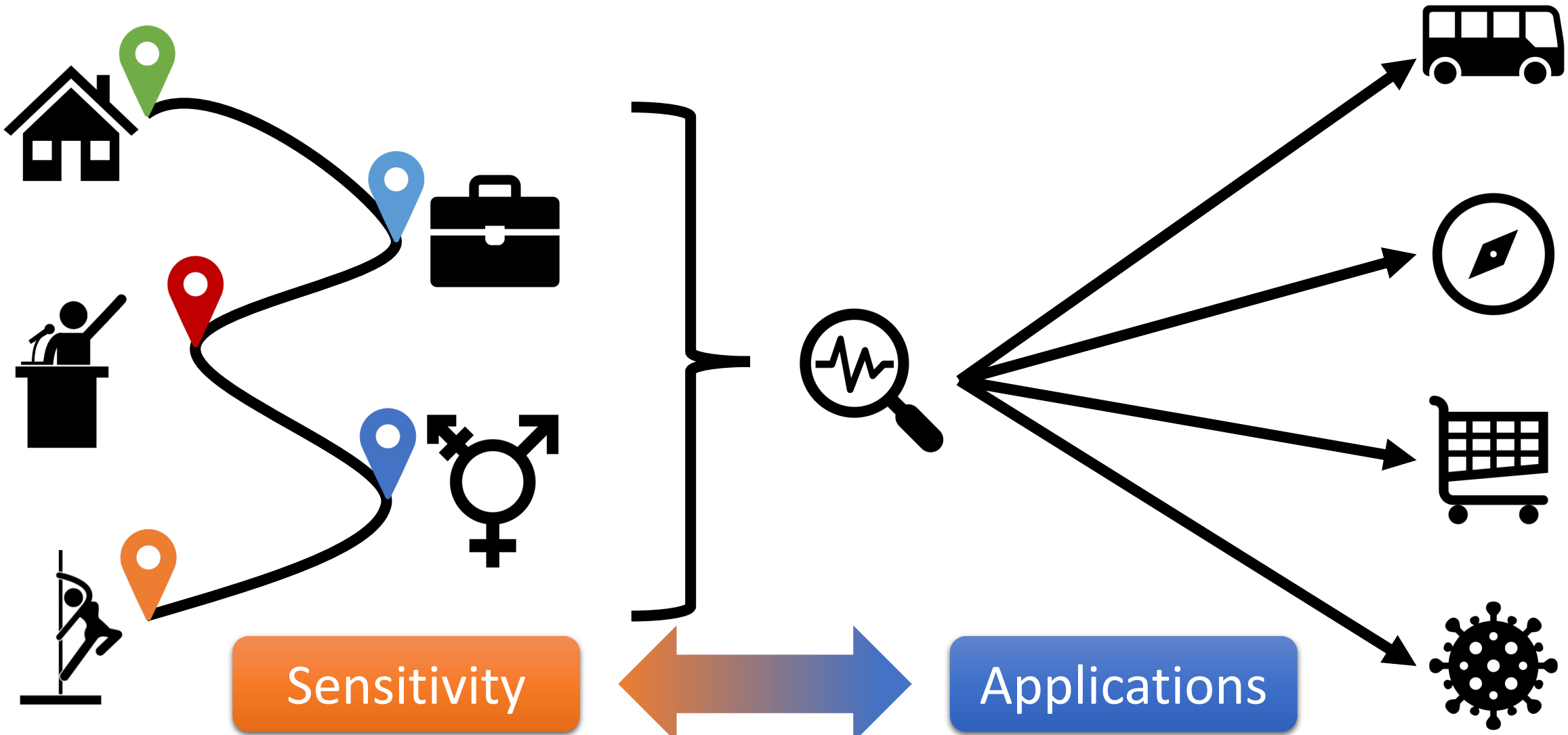
UNSW SYDNEY

UNSW Institute for Cyber Security

CSIRO DATA 61

Trajectory Publication

- 4 locations might identify 95% of humans [1]
- Redditor identified Muslim taxi drivers [2]



[1] Y.-A. de Montjoye, C. A. Hidalgo, M. Verleysen, and V. D. Blondel, "Unique in the Crowd: The privacy bounds of human mobility," Scientific Reports, vol. 3, no. 1, pp. 1–5, Dec. 2013, doi: 10.1038/srep01376.

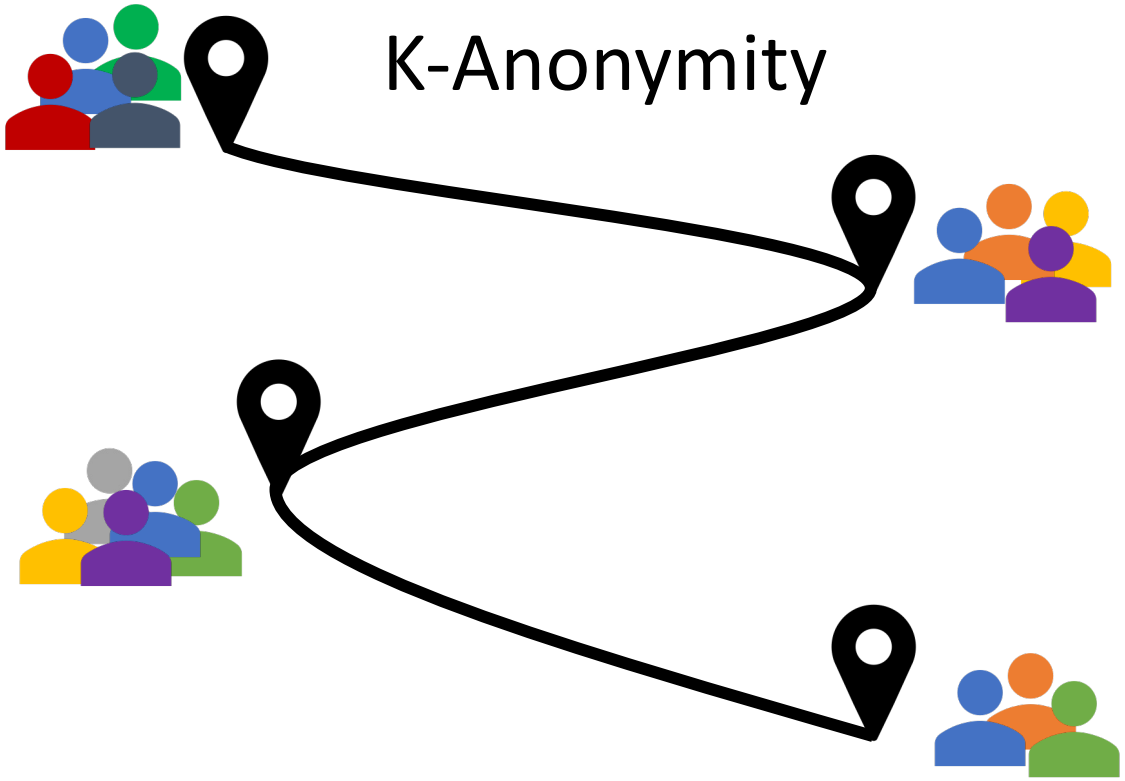
[2] L. Franceschi-Bicchierai, "Redditor cracks anonymous data trove to pinpoint Muslim cab drivers," 2015. https://mashable.com/archive/redditor-muslim-cab-drivers (accessed Sep. 28, 2021).

UNSW Institute for Cyber Security

UNSW SYDNEY

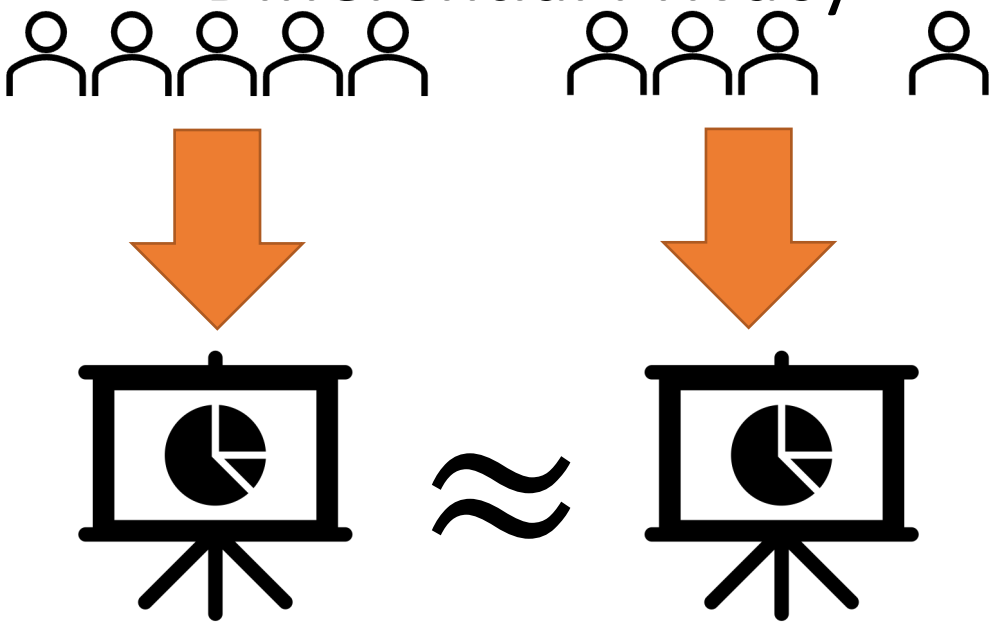
Trajectory Protection

K-Anonymity



- + Intuitive Parametrization
- + Simple(r) to achieve
- - No theoretical guarantees
- → Vulnerable to (background attacks)

Differential Privacy



- + Strong theoretical guarantees
- + Independent of background knowledge
- - Unintuitive parameters (ε, δ)

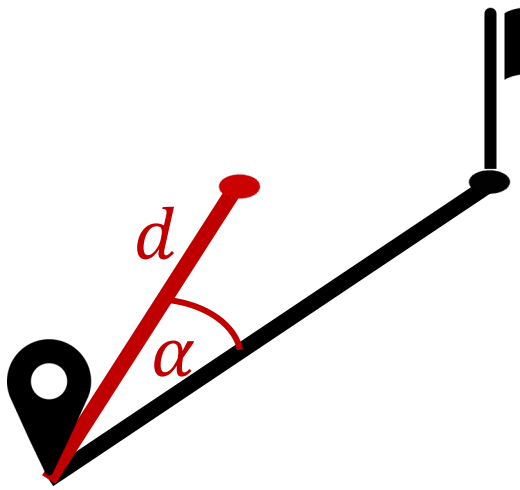
→ **De-facto privacy standard**

UNSW Institute for Cyber Security

UNSW SYDNEY

One example: Sampling Distance and Direction (SDD) mechanism [1]

Note: Still used as baseline/ state-of-the-art in 2020 [9, 10]



Distance d
Direction α

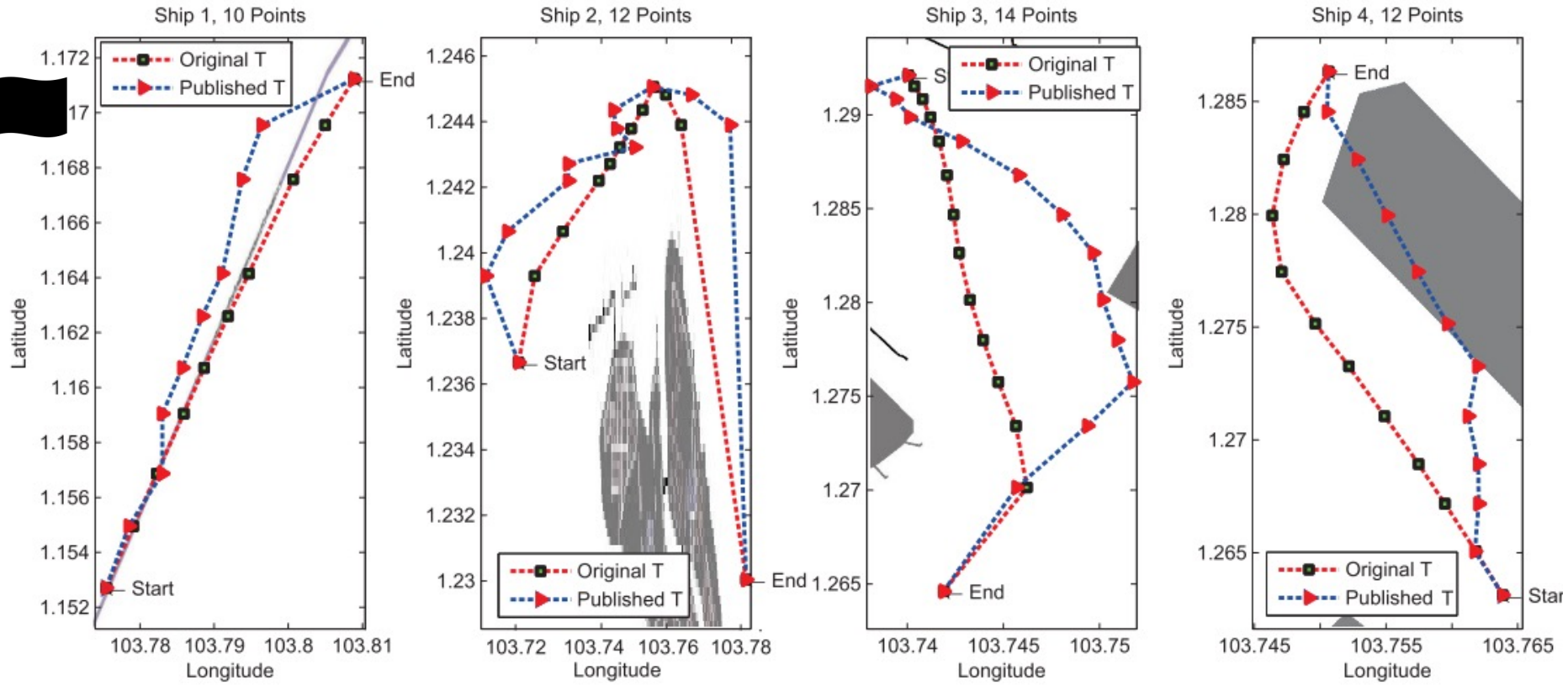


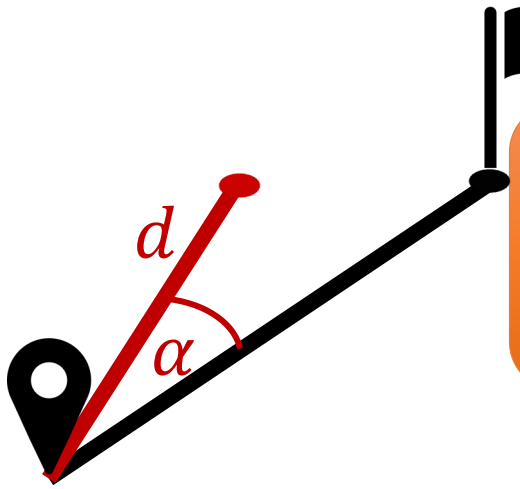
Figure Source: [1]

Figure 4: Original and published trajectories of 4 ships in Singapore Straits with $\varepsilon = 0.1$.

[1] K. Jiang, D. Shao, S. Bressan, T. Kister, and K.-L. Tan, "Publishing trajectories with differential privacy guarantees," in Proceedings of the 25th International Conference on Scientific and Statistical Database Management - SSDBM, New York, New York, USA, 2013, p. 1. doi: 10.1145/2484838.2484846.

One example: Sampling Distance and Direction (SDD) mechanism [1]

Note: Still used as baseline/ state-of-the-art in 2020 [9, 10]



Distance d
Direction α

Research Question:

Can an adversary (partly) reconstruct trajectories from a differential private release?

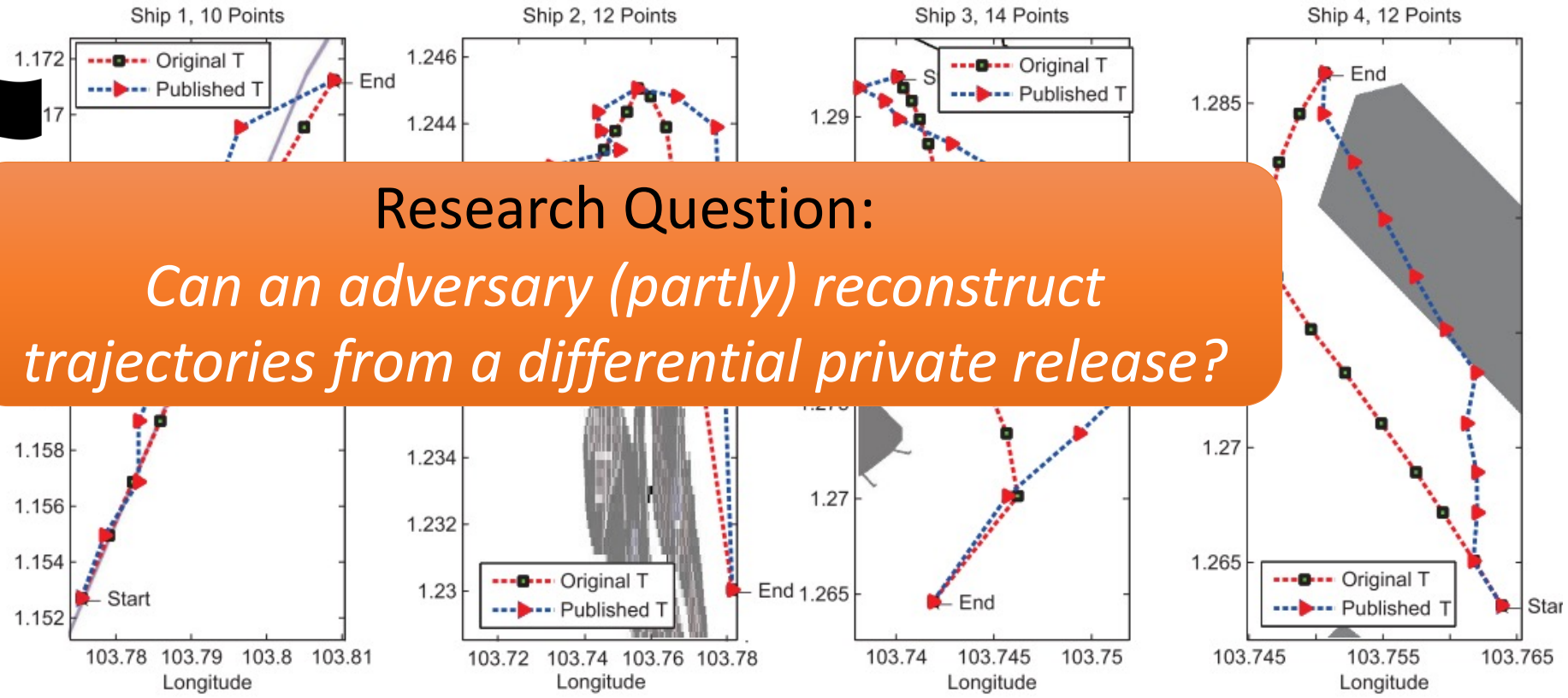


Figure Source: [1]

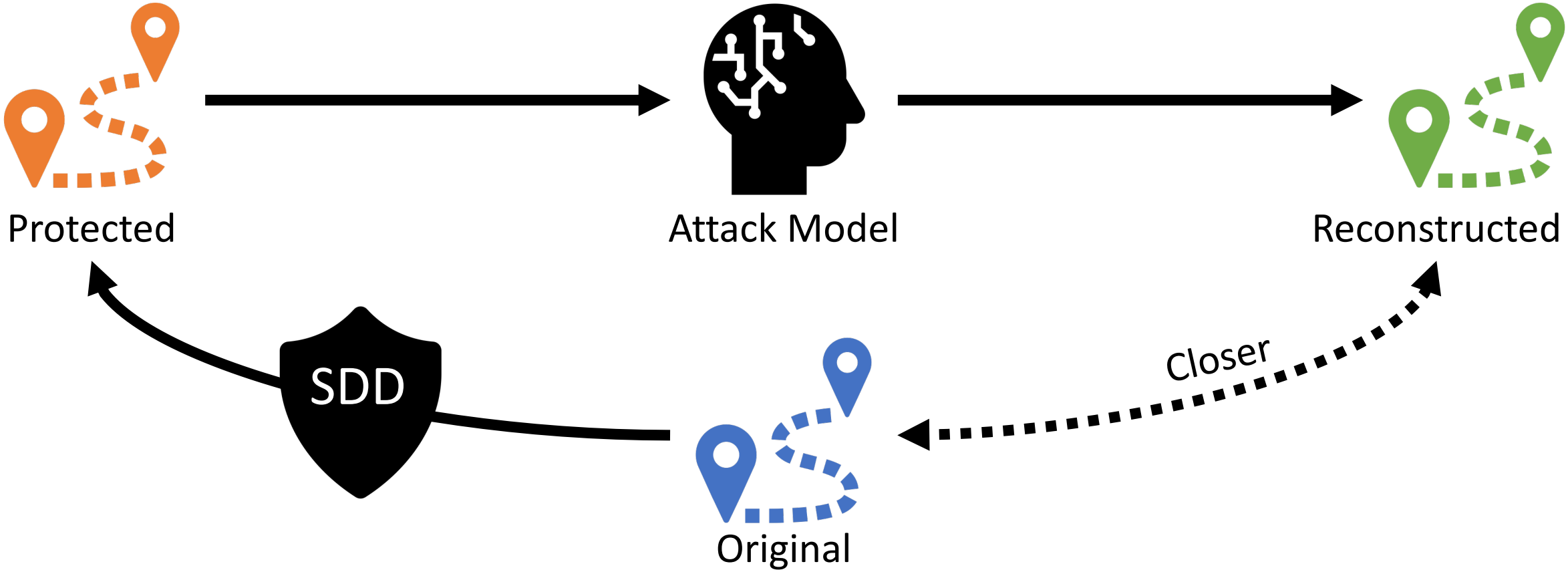
Figure 4: Original and published trajectories of 4 ships in Singapore Straits with $\varepsilon = 0.1$.

[1] K. Jiang, D. Shao, S. Bressan, T. Kister, and K.-L. Tan, "Publishing trajectories with differential privacy guarantees," in Proceedings of the 25th International Conference on Scientific and Statistical Database Management - SSDBM, New York, New York, USA, 2013, p. 1. doi: 10.1145/2484838.2484846.

UNSW Institute for Cyber Security

UNSW SYDNEY

Attack Idea

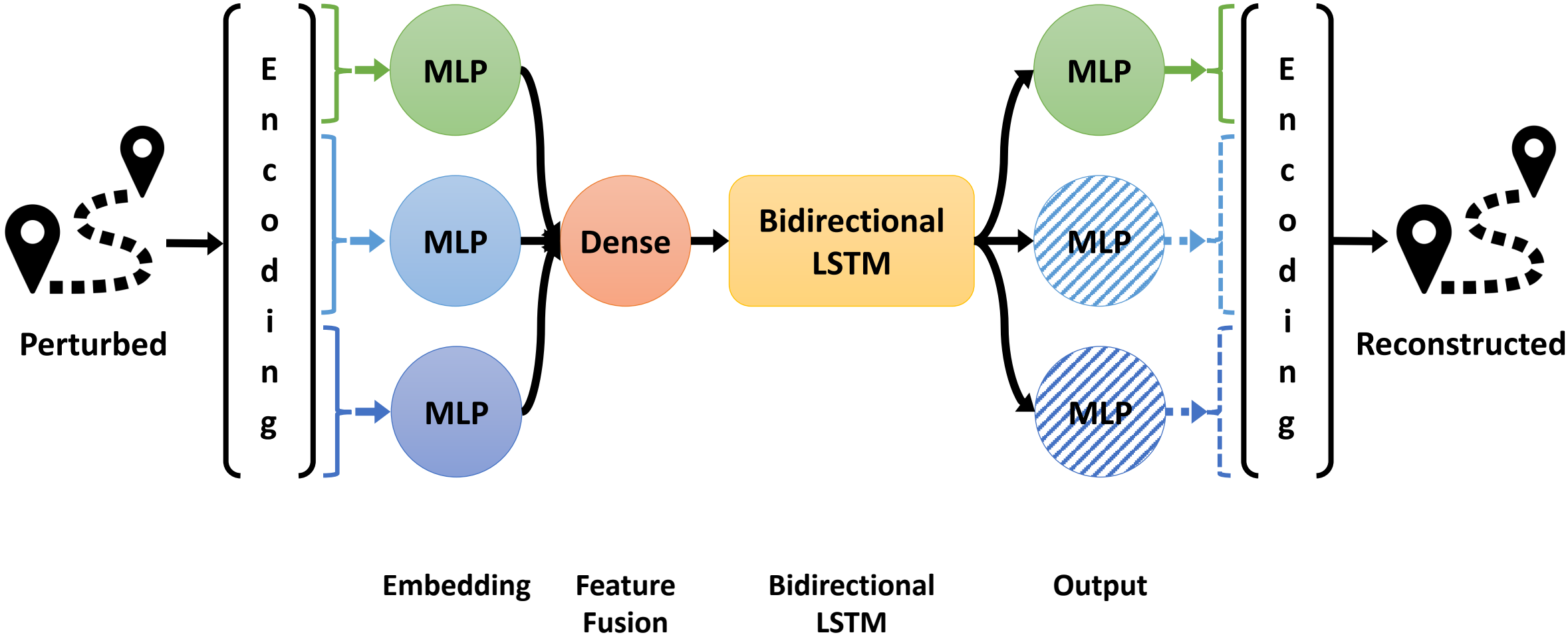


Idea: Reconstruct trajectories from a supposedly anonymized/protected release through a deep learning model.

UNSW Institute for Cyber Security

UNSW SYDNEY

Model



UNSW Institute for Cyber Security

UNSW SYDNEY

Evaluation

Pre-Processing:

- Outlier Removal (SDD requires upper bound on speed)
- Splitting of trajectories based on long breaks
- Latitude and Longitude measured from central reference point

Datasets:

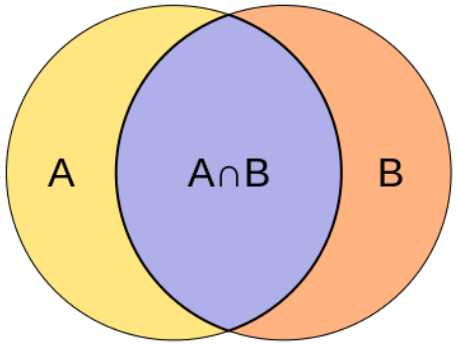
- **T-Drive:** Taxi trajectories only. Beijing area.
  - $163'006$ trajectories; $10 \leq length \leq 100; v \leq 90\ km/h$
- **GeoLife:** All transportation types. Larger geographical area.
  - $90'146$ trajectories; $10 \leq length \leq 200; v \leq 100\ km/h$

Protection Mechanisms:

- **CNoise:** Independent Laplace noise added to each coordinate
- **SDD:** Better utility through exponential mechanism

Metrics:

- **Euclidean Distance:** Standard trajectory similarity metric
- **Hausdorff Distance:** Standard trajectory similarity metric
- **Jaccard Index:** Representation of activity space *(Intersection over Union)*

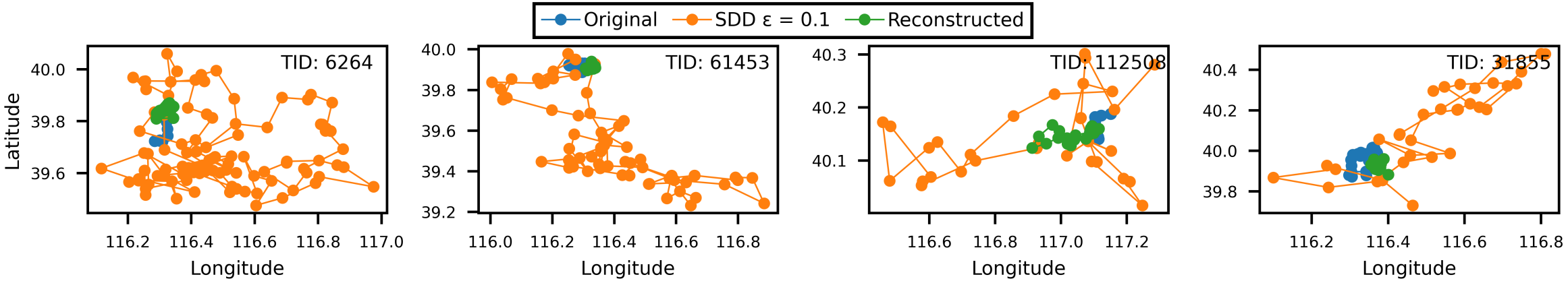


UNSW Institute for Cyber Security

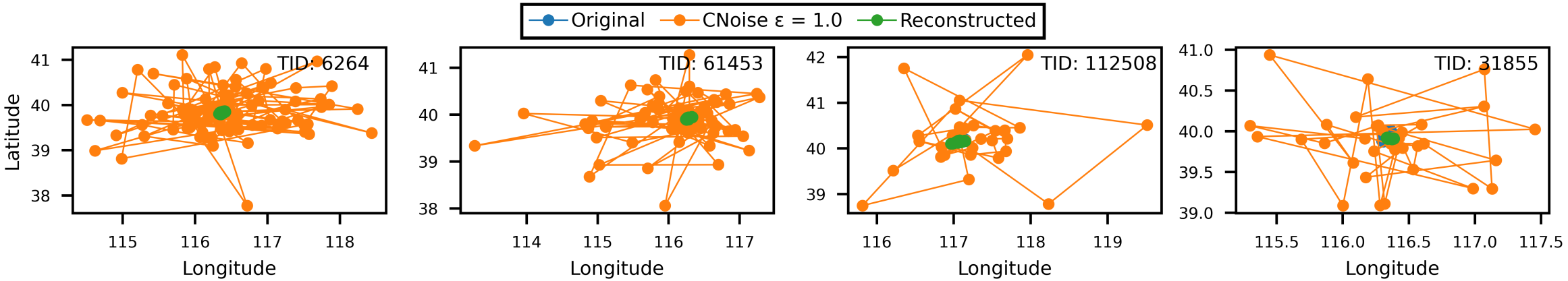
UNSW SYDNEY

Example Reconstruction

- Randomly chosen examples for $SDD\ with\ \varepsilon = 0.1$ from T-Drive



- Randomly chosen examples for $CNoise\ with\ \varepsilon = 1.0$ from T-Drive

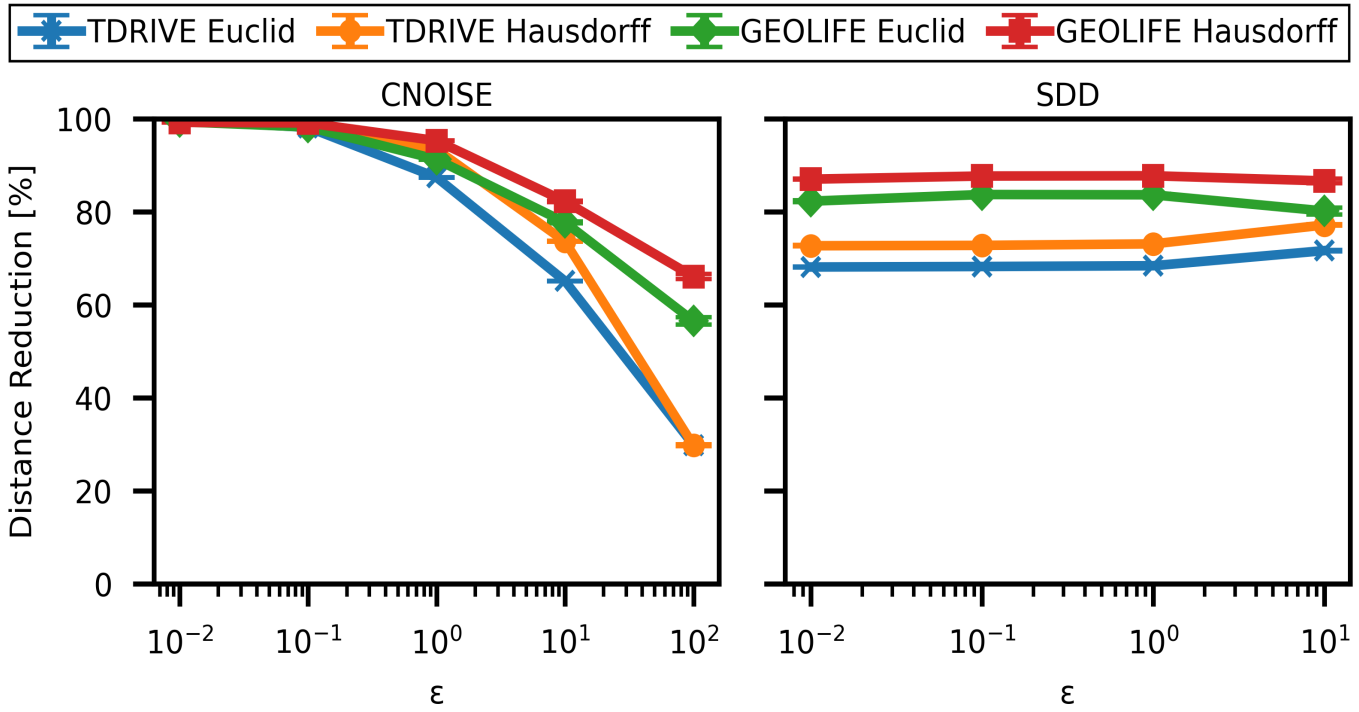


UNSW Institute for Cyber Security

UNSW SYDNEY

Results

- For $\varepsilon \leq 1$ **over 68% reduced distances** through reconstruction
- Found **security-privacy trade-off**
  - → A higher level of privacy (i.e., smaller ε/more perturbation) yields a higher reconstruction access

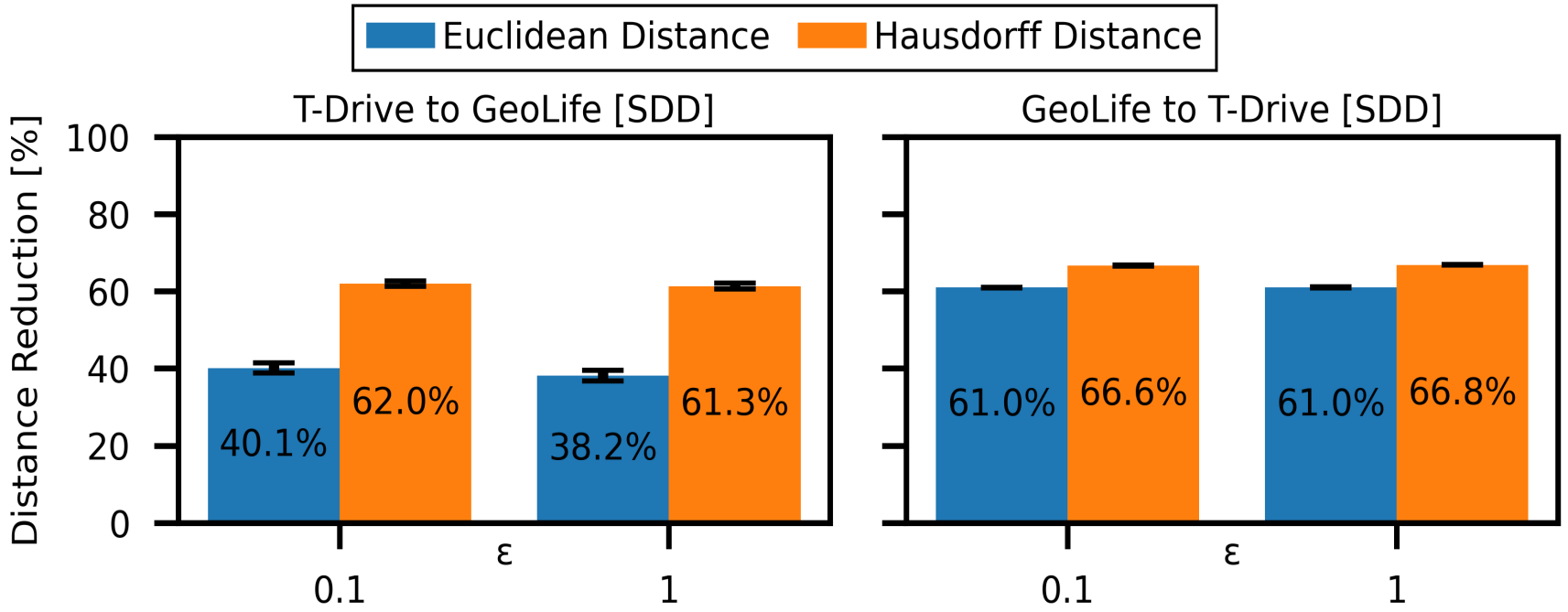


UNSW Institute for Cyber Security

UNSW SYDNEY

Transfer of Datasets

- Up to **67% distance reduction**
- → Attack represents threat for **real-world adversaries** and **state-of-the-art** protection mechanisms (vs Laplace noise)

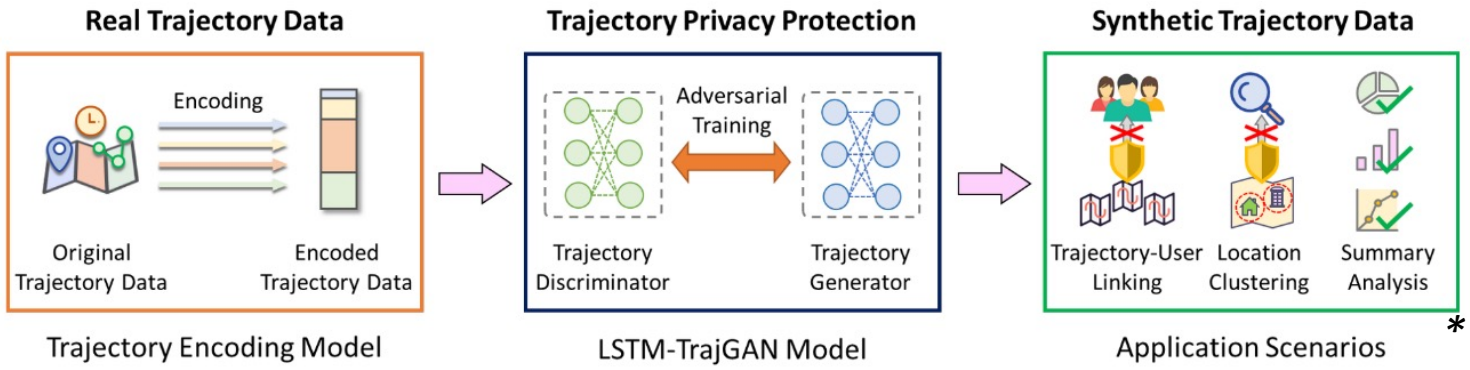


UNSW Institute for Cyber Security

UNSW SYDNEY

Related Work

- One existing attack on differential private trajectory publication mechanisms: iTracker [1]
  - Only considers standard Laplace noise protection
  - No implementation available (contacted authors)
- Model Baseline: LSTM-TrajGAN [2]
  - Uses a GAN to generate synthetic trajectories
  - Provides very good utility compared to other approaches
  - But no differential privacy guarantees (yet)



** Figure from [2]*

[1] M. Shao, J. Li, Q. Yan, F. Chen, H. Huang, and X. Chen, "Structured Sparsity Model Based Trajectory Tracking Using Private Location Data Release," IEEE Transactions on Dependable and Secure Computing, vol. 18, no. 6, pp. 2983–2995, 2020, doi: 10.1109/TDSC.2020.2972334.

[2] J. Rao, S. Gao, Y. Kang, and Q. Huang, "LSTM-TrajGAN: A Deep Learning Approach to Trajectory Privacy Protection," Leibniz International Proceedings in Informatics, vol. 177, no. GIScience, pp. 1–16, 2020, doi: 10.4230/LIPIcs.GIScience.2021.I.12.

Conclusion

- Current DP protection mechanisms yield *unauthentic perturbation*
- These differences can be exploited for *reconstruction attacks*
- → Results in *reduced level of privacy protection*

Improved privacy-preserving publication mechanisms have to be developed!

Artifacts: Functional

Paper

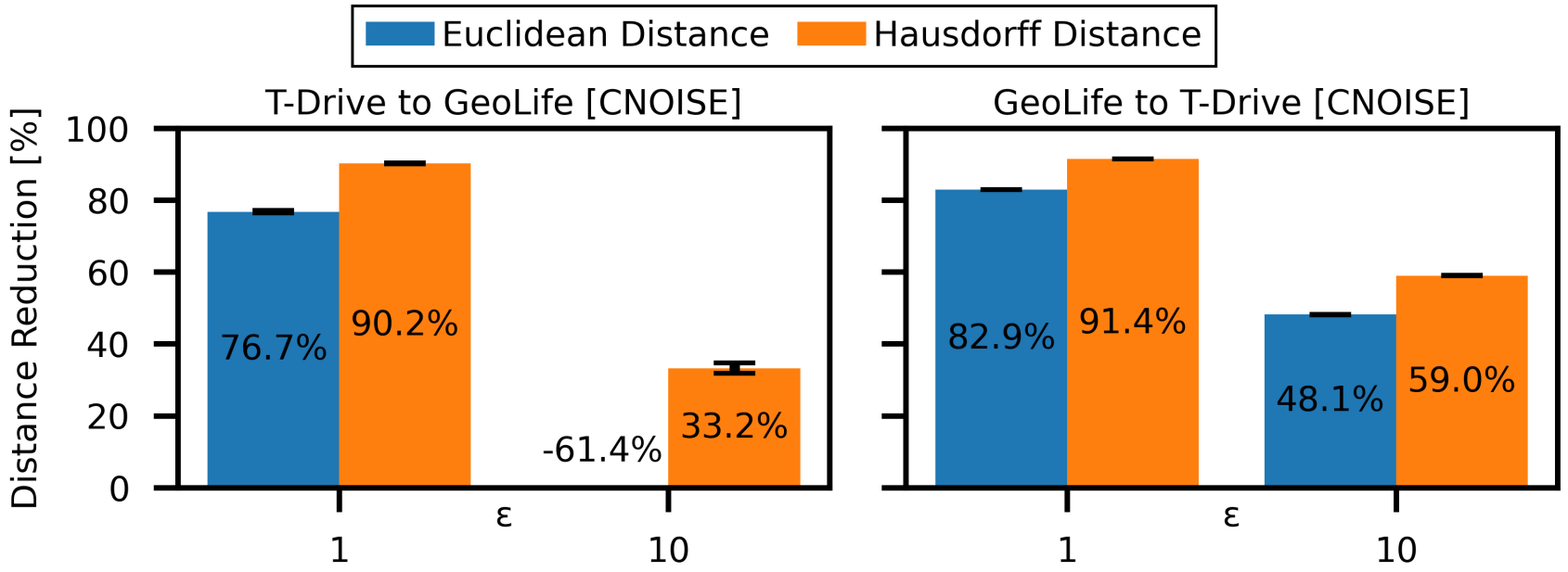
Acknowledgement

The authors would like to thank **UNSW**, the **Commonwealth of Australia**, and the **Cybersecurity Cooperative Research Centre Limited** for their support.

UNSW Institute for Cyber Security

UNSW SYDNEY

Backup: Transfer CNoise



UNSW Institute for Cyber Security

UNSW SYDNEY

Backup: Transfer ε

ID	Mechanism	ε Train	ε Test	Euclidean	Hausdorff
27	CNoise	1.0	10.0	24.3%	46.2%
28	CNoise	10.0	1.0	72.5%	79.3%
29	SDD	0.1	1.0	68.4%	73.1%
30	SDD	1.0	0.1	68.3%	72.8%

UNSW
Institute for
Cyber Security

UNSW
SYDNEY

Transfer Mechanism

ID	Train	Test	ε	Euclidean	Hausdorff
31	CNoise	SDD	1.0	27.7 %	44.9%
32	SDD	CNoise	1.0	53.0 %	70.3%

UNSW
Institute for
Cyber Security

UNSW
SYDNEY

Backup: Runtime

- Reconstruction of one trajectory

- GeoLife, SDD $\varepsilon = 0.1$: $[\mathbf{51.3}; \mathbf{52.1}]\boldsymbol{ms}$ is 99% conf. interval
- T-Drive, SDD $\varepsilon = 0.1$: $[\mathbf{44.8}; \mathbf{45.6}]\boldsymbol{ms}$ is 99% conf. interval

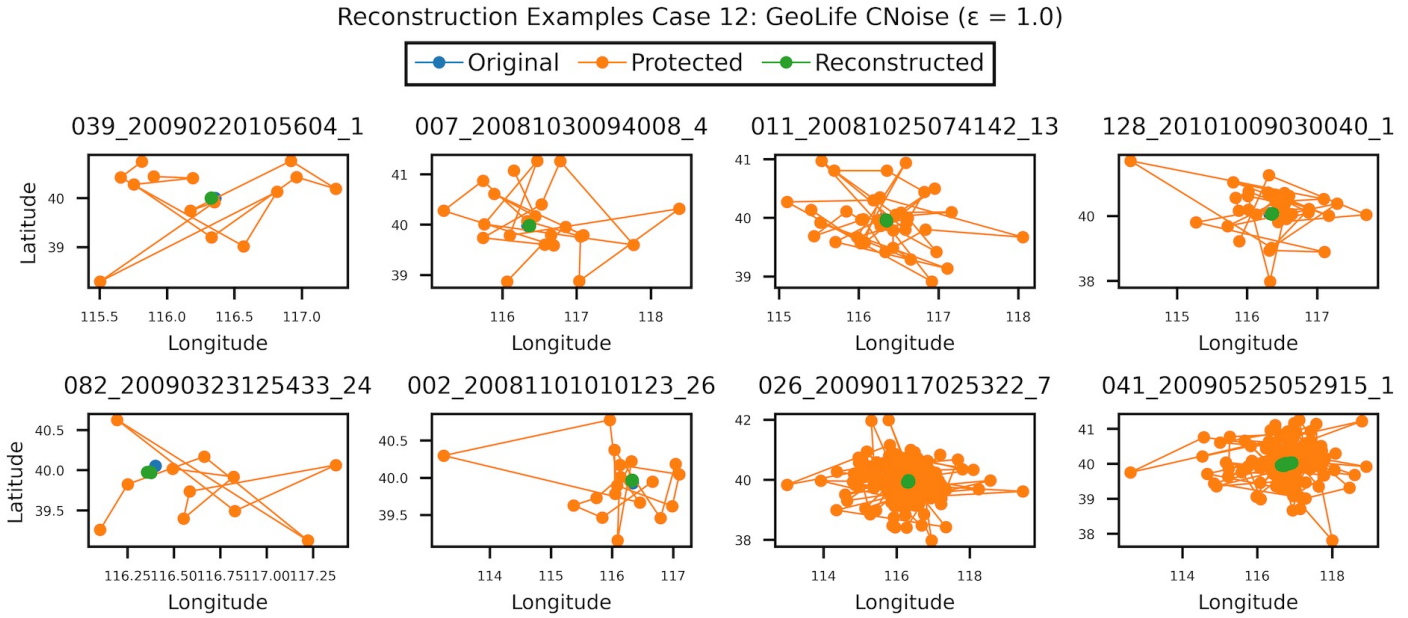
- Ubuntu 20.04 LTS
  - 2x Intel Xeon Silver 4208; 128GB RAM
  - NVIDIA Tesla T4 with 16 GB RAM (4 GPUs available, only one used)

UNSW
Institute for
Cyber Security

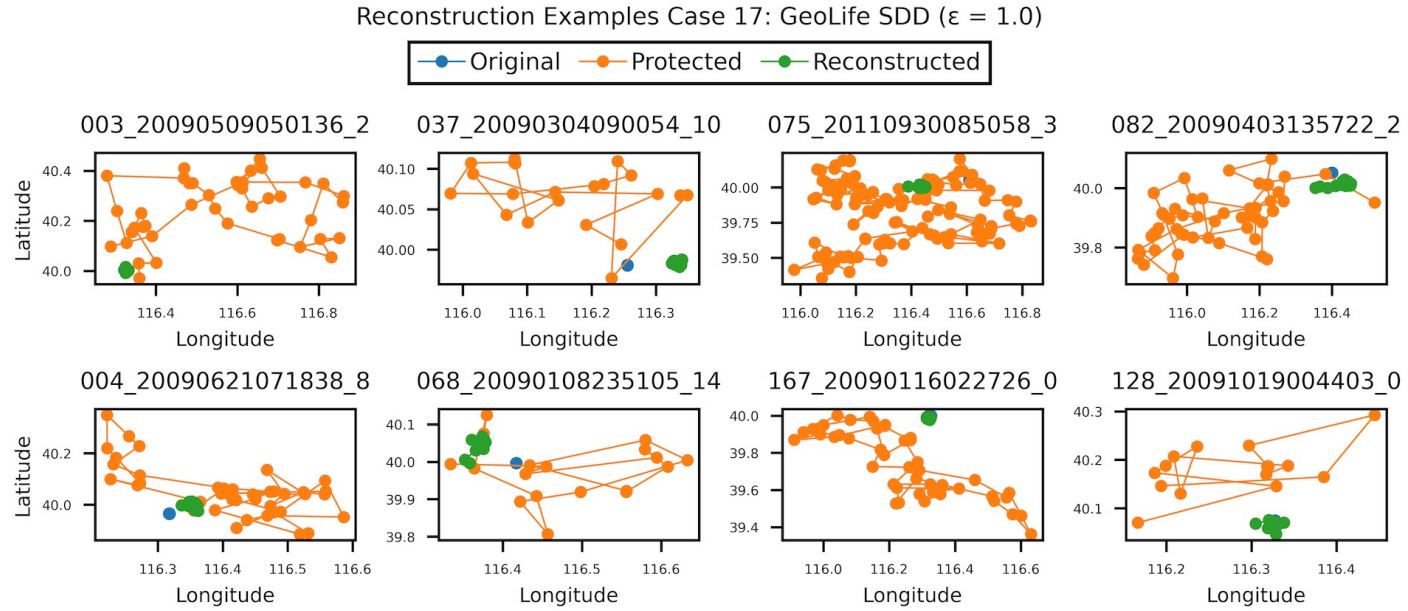
UNSW
SYDNEY

Backup: Example GeoLife

$GeoLife\ with\ CNoise\ \varepsilon = 1.0$



$GeoLife\ with\ SDD\ \varepsilon = 1.0$



UNSW Institute for Cyber Security

UNSW SYDNEY