Generalised Context Privacy

26th February 2024, Manchester University

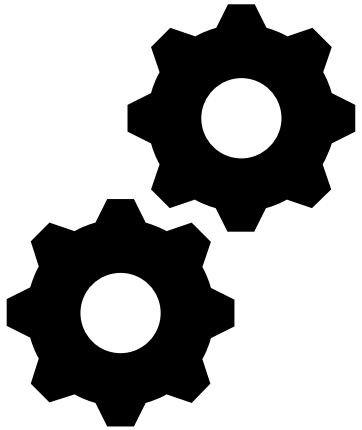
Matthew Bradbury

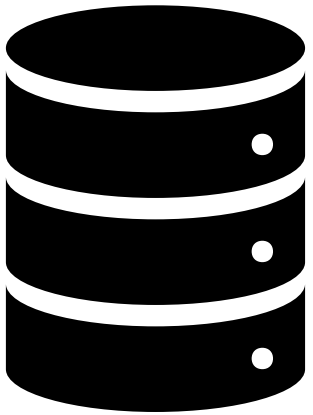
[EP/X040038/1]

Content vs Context Privacy

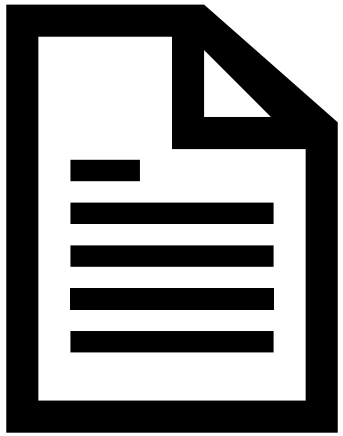
Lancaster University

Context privacy threat via adversary observations





Content privacy threat to data in motion



Cyber Physical System performing actions in an environment

Information stored in a database

Content privacy threat to data at rest

Information being revealed to relevant parties

Content privacy threat to data release

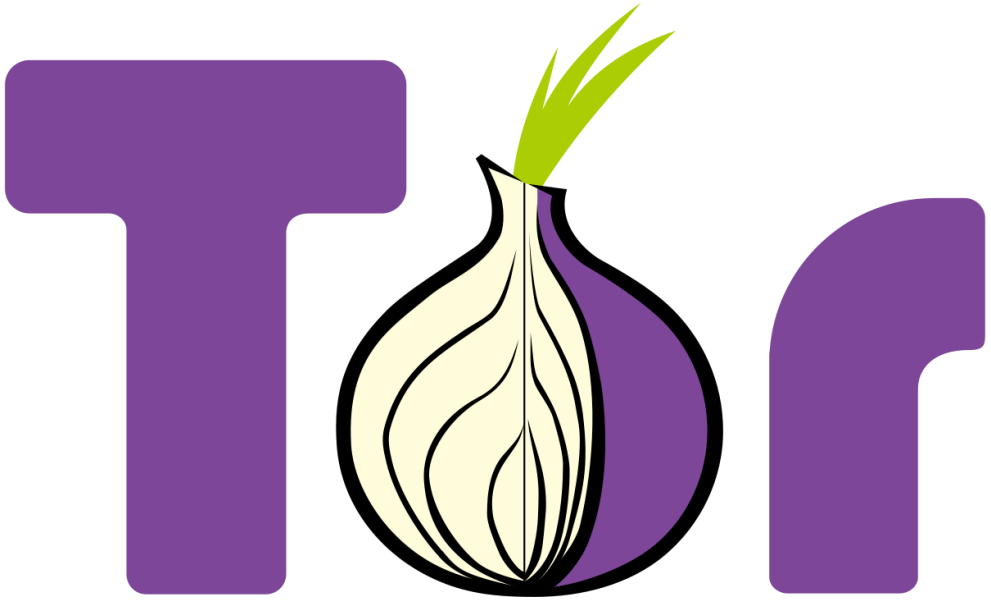
Lancaster University

Domain-specific techniques exist

Onion Routing

Lancaster University

- Onion routing provides source/destination privacy over the internet
- Obscures path messages take with multiple layers of encryption
- Threat model: Adversaries at (potentially multiple) points in the network and can monitor communications



©TOR Project https://www.torproject.org/

M. G. Reed, P. F. Syverson and D. M. Goldschlag, "Anonymous connections and onion routing," in *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 4, pp. 482-494, May 1998, doi: 10.1109/49.668972.

Wireless Sensor Networks

Lancaster University

- Large networks of devices with:
  - Low power – Two AA batteries to last multiple years
  - Low resources – 10s MHz CPU / 10s KiB RAM / 100s KiB ROM
  - Potentially no stable storage
  - Various sensors / actuators
  - Low data rate communications – 250 kbps
- Useful when access to infrastructure is limited or costly

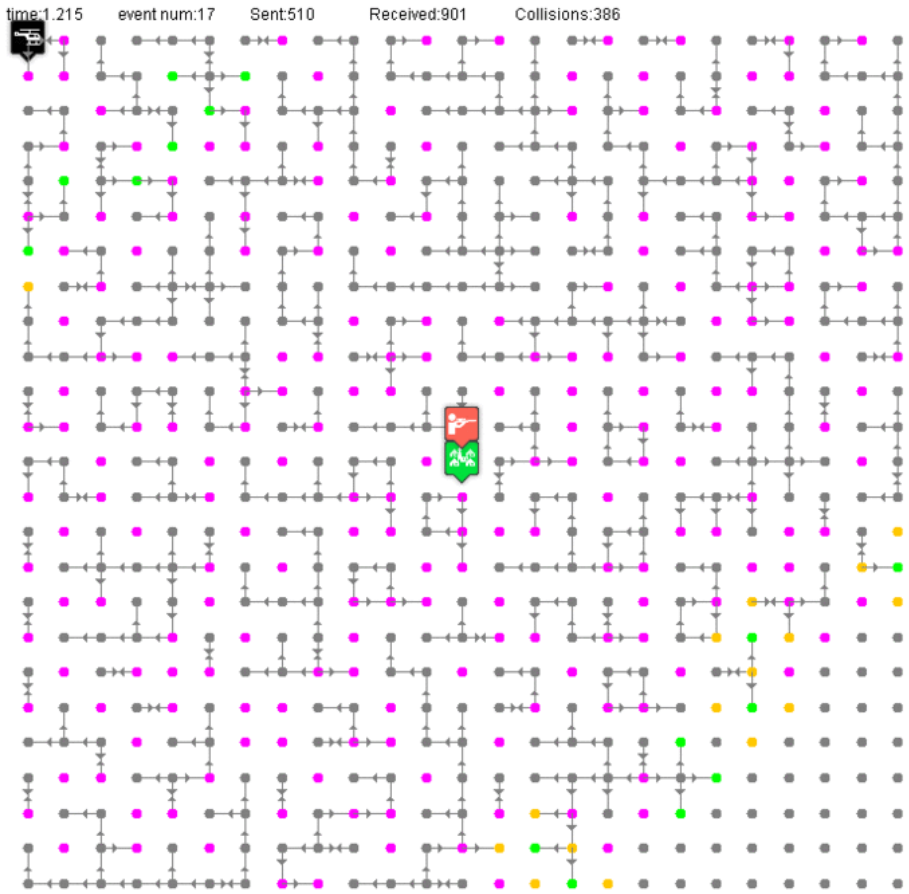
nRF52840 Dongle ©Nordic Semiconductor

https://research.csiro.au/robotics/cow-tracking/ ©CSIRO

Source Location Privacy

Lancaster University

- Messages are routed from a valuable asset to a base station
- Messages are encrypted
- Context information – the direction from which a message is received – allows locating the source
- Assume: Base station's location is known by adversary

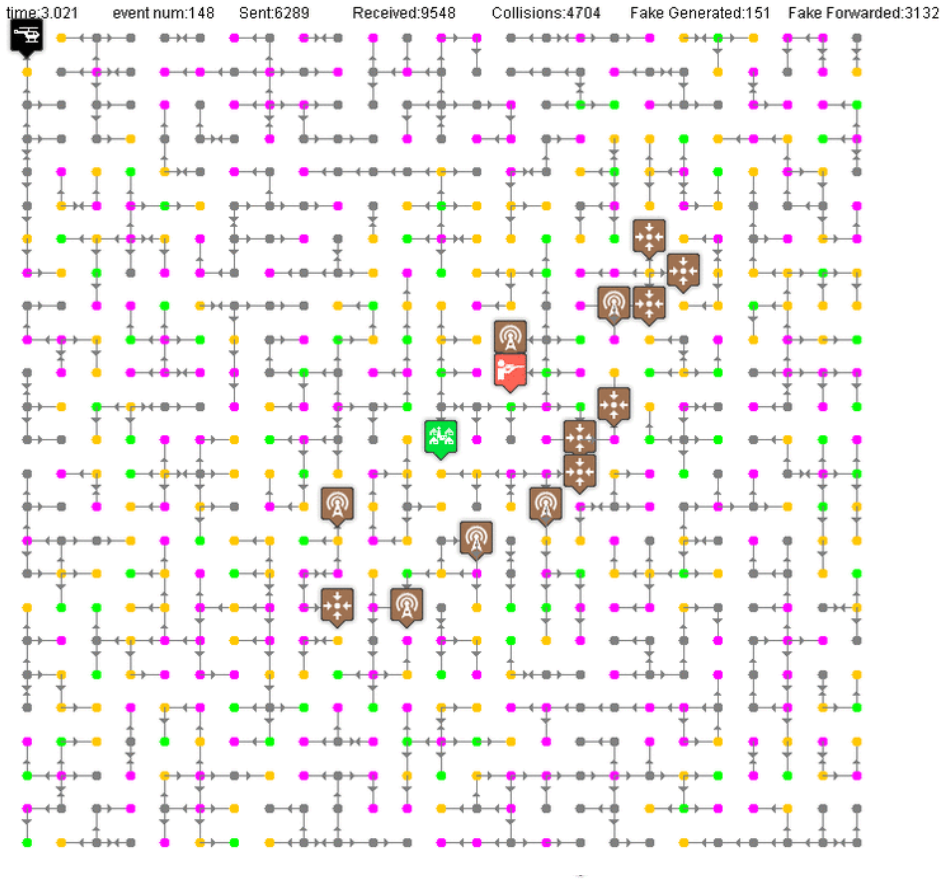


Kamat, P., Zhang, Y., 2005. Enhancing Source-Location Privacy in Sensor Network Routing. *25th IEEE International Conference on Distributed Computing Systems*. Columbus, OH, USA. (6-10 June), 599-608. DOI:10.1109/ICDCS.2005.31.

Existing Techniques – Fake Sources

Lancaster University

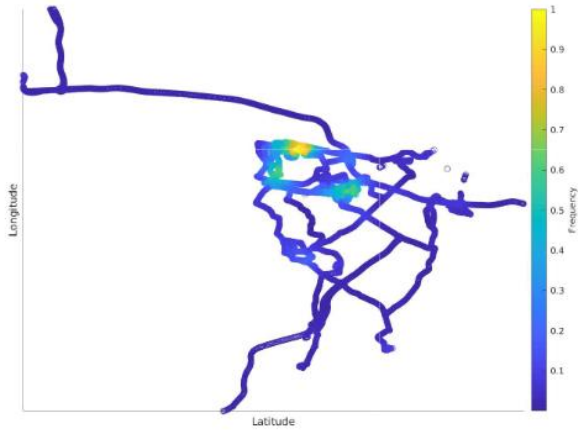
- Fake sources generate fake messages
- Fake messages indistinguishable from normal messages
- Lure the adversary in a different direction to the real source



Bradbury, M., Jhumka, A. and Leeke, M. 2018. Hybrid Online Protocols for Source Location Privacy in Wireless Sensor Networks. *Journal of Parallel and Distributed Computing*. 115, (May 2018), 67–81. DOI:10.1016/j.jpdc.2018.01.006.

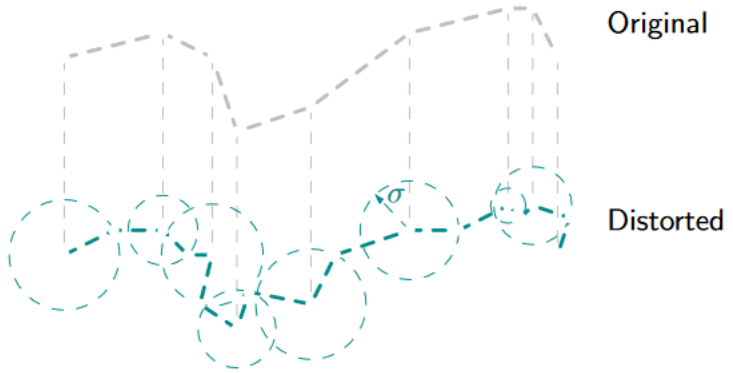
Many other context privacy threats

- Connected Vehicles – facilitating tracking and potentially unwanted pattern of life analysis



M. Bradbury, P. Taylor, U. I. At., C. Maple, and N. Griffiths. 2020. Privacy Challenges with Protecting Live Vehicular Location Context. IEEE Access 8 (Nov. 2020), 207465–207484. https://doi.org/10.1109/ACCESS.2020.3038533

- Mouse movements – used to predict demographics of users



L. A. Leiva, I. Arapakis, and C. Iordanou. 2021. My Mouse, My Rules: Privacy Issues of Behavioral User Profiling via Mouse Tracking. In Proceedings of the 2021 Conference on Human Information Interaction and Retrieval. 51–61. https://doi.org/10.1145/3406522.3446011

Lancaster University

There is a need for context privacy in new domains

Example: Water Treatment Plant

Lancaster University

- Adversary will observe the plant to better understand how to attack it
- Obscuring the activities taken by the plant reduce the ability of the adversary to attack it
- Hide causal link between actions
  - Wireless PLC controls release of chlorine to kill bacteria in water
  - Adversary can learn that the wireless signal leads to chlorine release
  - Change actions to obscure this cause-effect

Example: Drone Surveillance

Lancaster University

- Autonomous drones used to perform surveillance of an area (e.g., farmland)
- Where a drone is and when it performs surveillance is valuable information
- Indicates areas of interest where attacks should be focused
- Add redundant surveillance to obscure areas of interest

©Consortiq

Example: Home Appliances

- Resources consumed in home can reveal important information
  - Are you home?
  - What devices do you own?
  - When do you use them?
- Perturb their activity to obscure this
- Move the threat actor to your smart meter
- Existing: Use energy harvesting and storage to obscure activity

©Alamy

O. Tan, D. Gunduz, and H. V. Poor. 2013. Increasing Smart Meter Privacy Through Energy Harvesting and Storage Devices. IEEE Journal on Selected Areas in Communications 31, 7 (2013), 1331–1341. https://doi.org/10.1109/JSAC.2013.130715

Lancaster University

Developing context privacy techniques for novel situations is slow

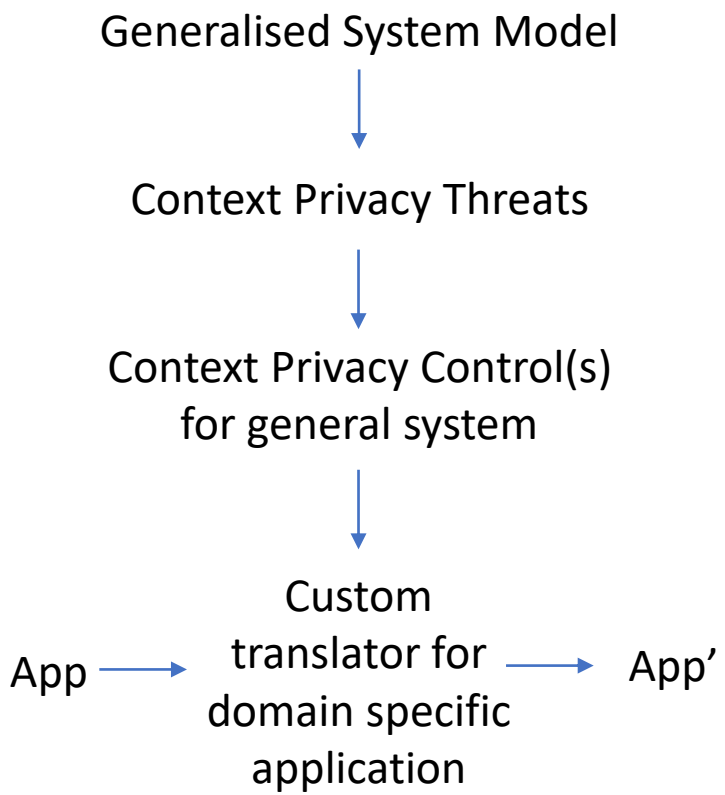
Generalised Context Privacy

Lancaster University

- We have solutions to many domain-specific problems
- They are not easily translatable to new systems
- Time is needed to develop solutions to novel systems

To improve:

- Develop general context privacy solutions to an arbitrary system once
- Develop domain-specific translators as needed when a new context privacy threat is identified



Generalised Context Privacy: Aims

Lancaster University

1. How to quantify information loss from an arbitrary cyber-physical system?
2. How can the sequence of actions have controls applied to reduce information loss?
    1. While maintaining system availability
    2. While minimising the cost of the controls
3. How can the system be changed such that it performs actions with a bounded information loss?

Threat Model

Lancaster University

Goal: Adversary is directly observing the system to learn information about it

- Aims of the system
- Approaches it is taking to achieve those aims
- Detect changes in behaviour that could be useful signals to launch other attacks

Capabilities: Adversary is passive (i.e., does not interact with the system)

Assumption: Adversary perfectly makes observations

- All states / actions of a system are correctly observed
- None are omitted

Assumption: This is a new system that has not been previously observed

- Previously deployed systems have already revealed information

Approaches to reduce privacy loss

Lancaster University

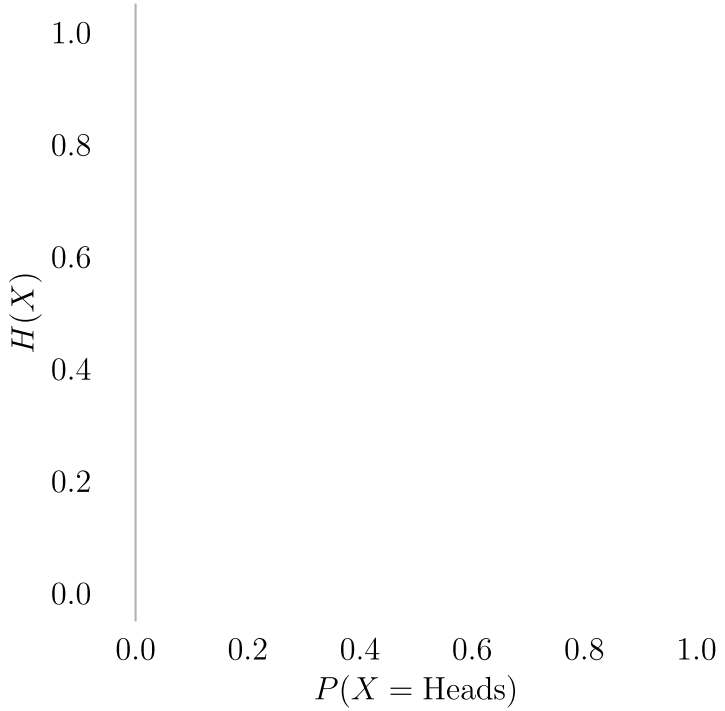
Typically three categories:

1. Make the sensitive action commonplace
2. Introduce noise to the actions the system takes / states the system is in
3. Limit the observability of the system to the adversary

Lancaster University

Quantifying Context Privacy

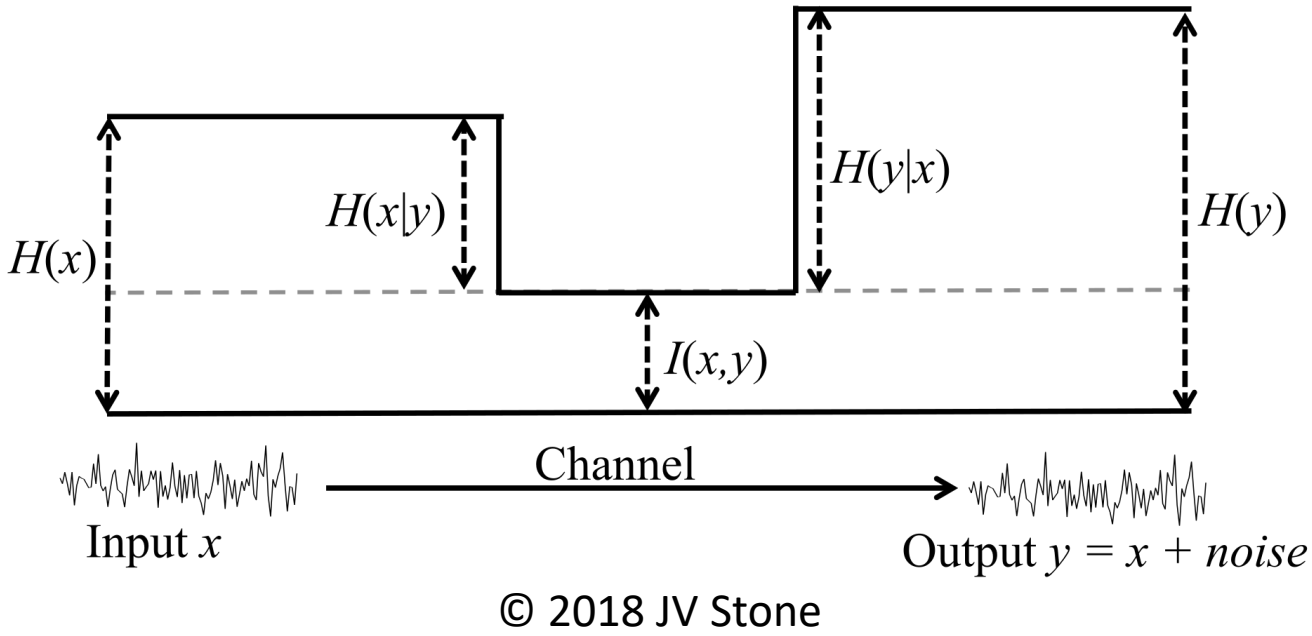
Brief Information Theory Refresher



- Entropy (rv. X)
  $\mathrm{H}(X) = -\sum_{x \in \boldsymbol{X}} \Pr(X = x) \log_2 \Pr(X = x)$
  Measure of uncertainty in random variable X
- Conditional entropy
  $\mathrm{H}(X|Y) = -\sum_{x \in \boldsymbol{X}} \sum_{y \in \boldsymbol{Y}} \Pr(X = x, Y = y) \log_2 \frac{\Pr(X=x, Y=y)}{\Pr(Y=y)}$
  Measure of uncertainty in X given Y
- Use base 2 so result is in the units of Shannon bits

Quantification of Privacy Loss

- Normal communication channel
  - How much information can you convey across a noisy channel?
- Context Privacy
  - How much noise needs to be added to increase uncertainty of observer?



Using Directed Information

Lancaster University

Directed Information: How much information is conveyed from one process to another?

$$I(J_{0:t} \to Y_{0:t}) = \sum_{i=0}^{t} H(J_{0:i}|Y_{0:i-1}) - H(J_{0:i}|Y_{0:i})$$

- $J_{0:t}$ (rv. system state/actions)
- $Y_{0:t}$ (rv. adversary observations)

What is needed to calculate this:

- System model: $\Pr(J_{0:t} = T_{0:t})$ probability of a system trace $T_{0:t}$
- Adversary model: $\Pr(Y_{0:t} = O_{0:t})$ probability of an adversary making obs. $O_{0:t}$
- Joint distribution: $\Pr(J_{0:t} = T_{0:t}, Y_{0:t} = O_{0:t})$

- System Trace – Sequence of states and actions $T_{0:t} \triangleq (S_0, A_1, S_1, \dots, A_t, S_t)$
- Observation Sequence $O_{0:t} \triangleq (O_0, O_1, \dots, O_t)$
- Joint dist. States and Actions $J_{0:t} \triangleq (E_0, N_1, \dots, E_t, N_t)$
- Rv. System state at time i E_i
- Rv. System action at time i N_i

Information Gain

Lancaster University

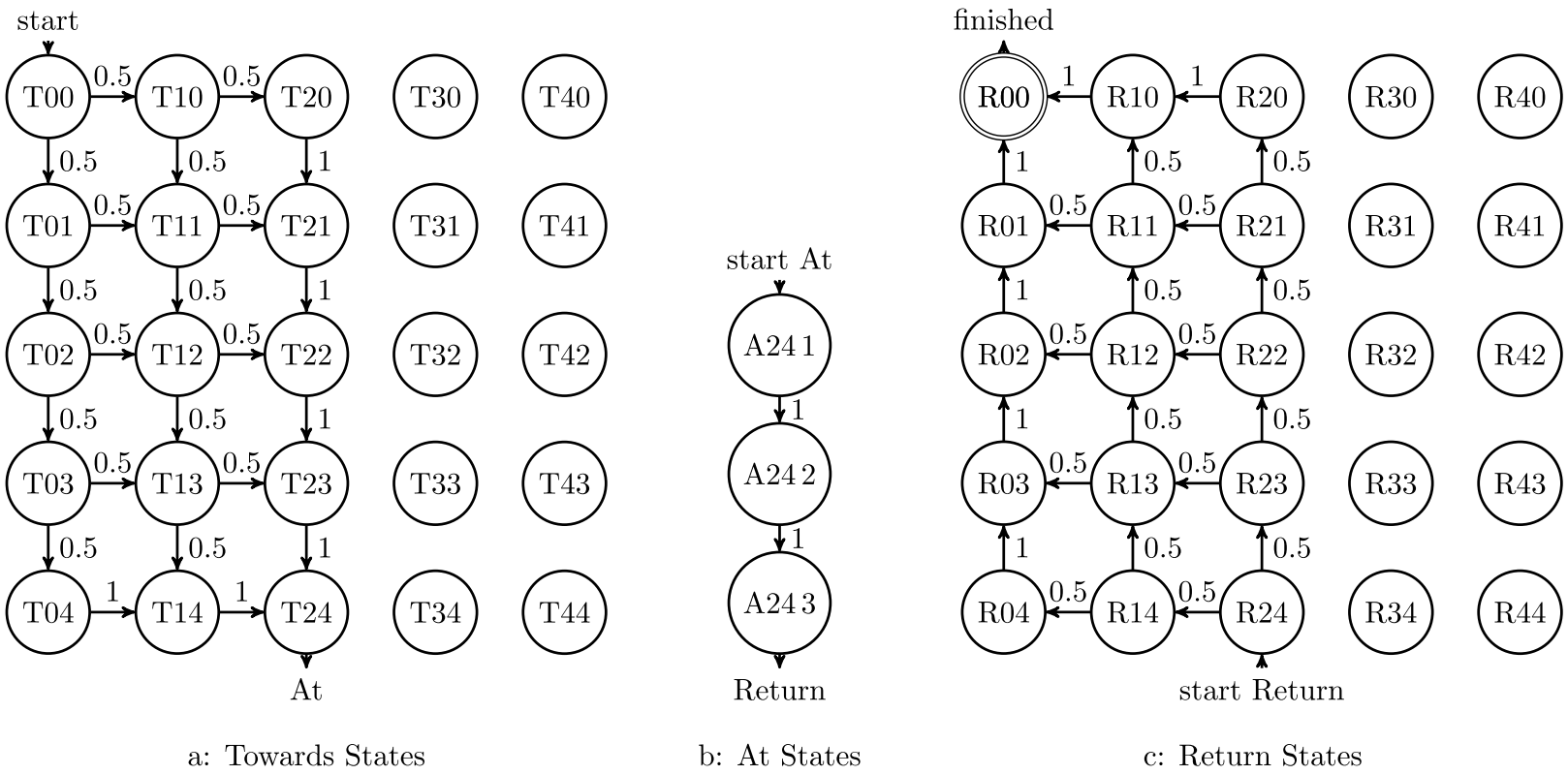
Self-information: *surprise* in observing outcome.

Used to quantify information gain by adversary when making observation

$$I(O_{0:t}) = -\log_2 \Pr(Y_{0:t} = O_{0:t})$$

- $Y_{0:t}$ (rv. adversary observations)
- Adversary Observation Sequence
  $O_{0:t} \triangleq (O_0, O_1, \ldots, O_t)$

Example system – Last mile drone delivery



a: Towards States b: At States c: Return States

Figure 1: Markov chain for one drone with actions taken between states omitted. Each state leads to one or more actions and each action leads to a single state in this example. For example, at state T00 two actions can be taken T00-South and T00-West. These actions lead to states T01 and T10 respectively.

How to obtain adversary beliefs?

Challenges:

- How many traces to train on? What is the impact of varying them?
- How to handle beliefs on observations that cannot be made on this system?

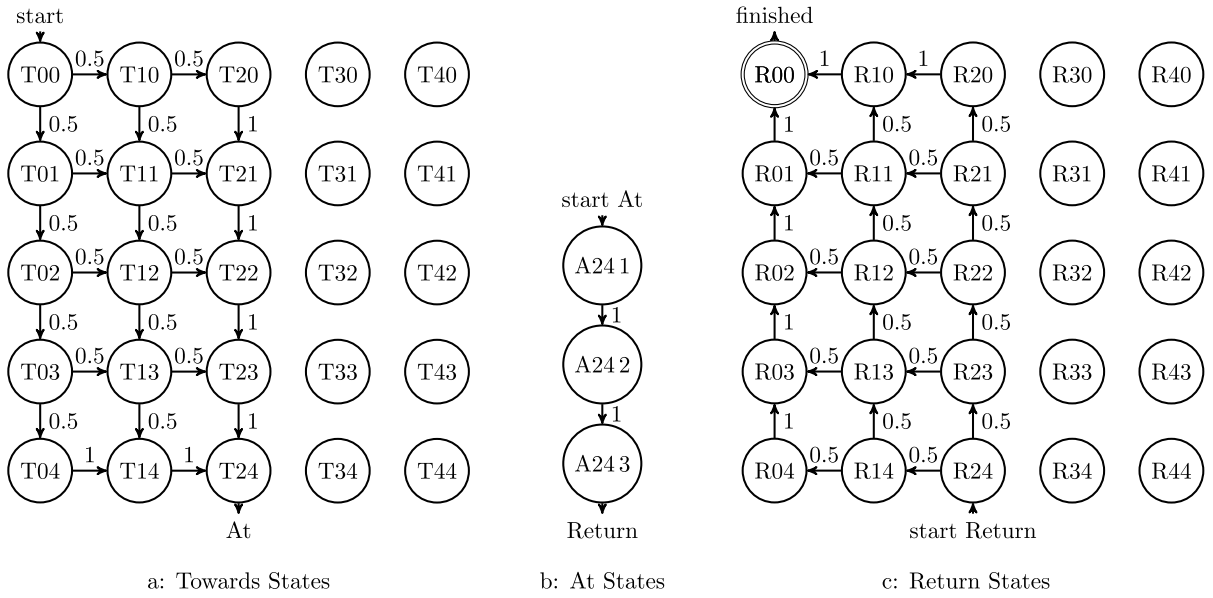
Why a Hidden Markov Model?

Lancaster University

- Can train on arbitrary sequence of observations
  - Hidden states map to actions / states of the system
  - Have a finite number of discrete hidden states and discrete observation
- Gives probability of an adversary making observations $\Pr(Y_{0:t} = O_{0:t})$
- Potential for other ML models to act as adversary belief model

Test the quantification

- Non-sensitive system
  - Drone goes to (2, 4) and returns to (0, 0)
  - Adversary has observed and built a model on
- Sensitive action
  - Drone goes to (4, 2) and returns to (0, 0)
  - Adversary not previously observed such behaviour

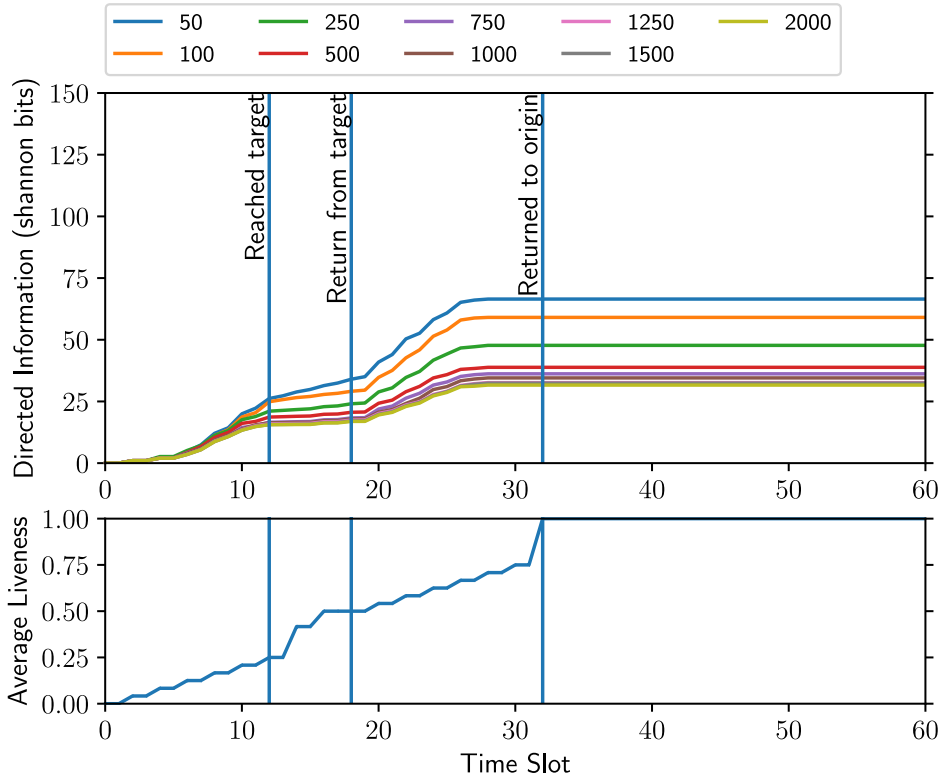


a: Towards States b: At States c: Return States

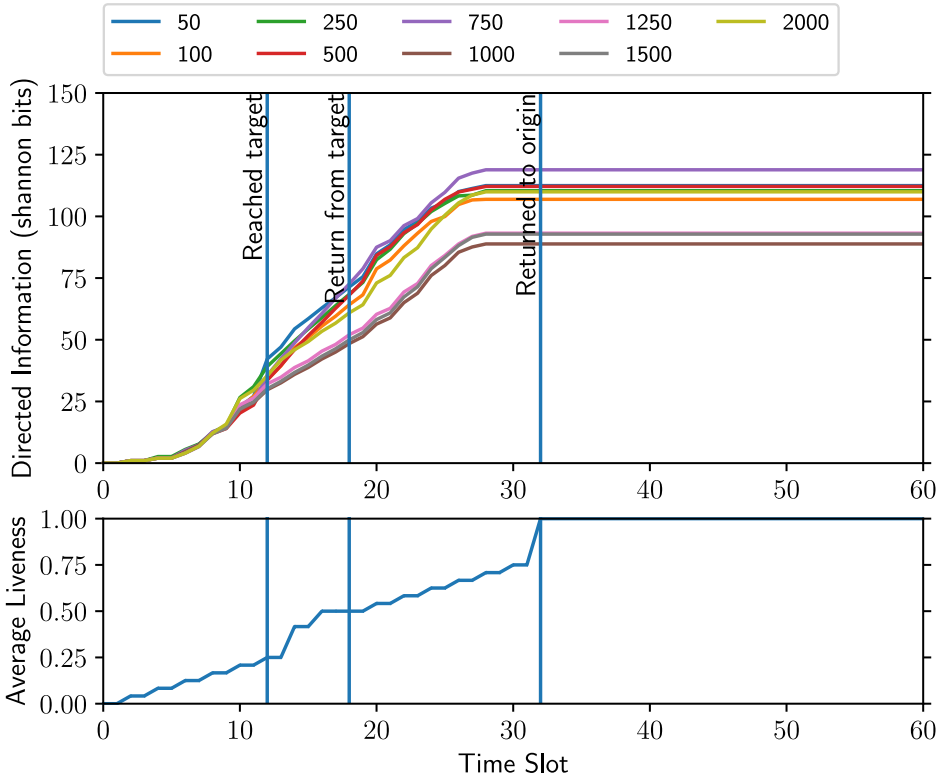
How much information is conveyed?

Lancaster University

Target is (2,4)



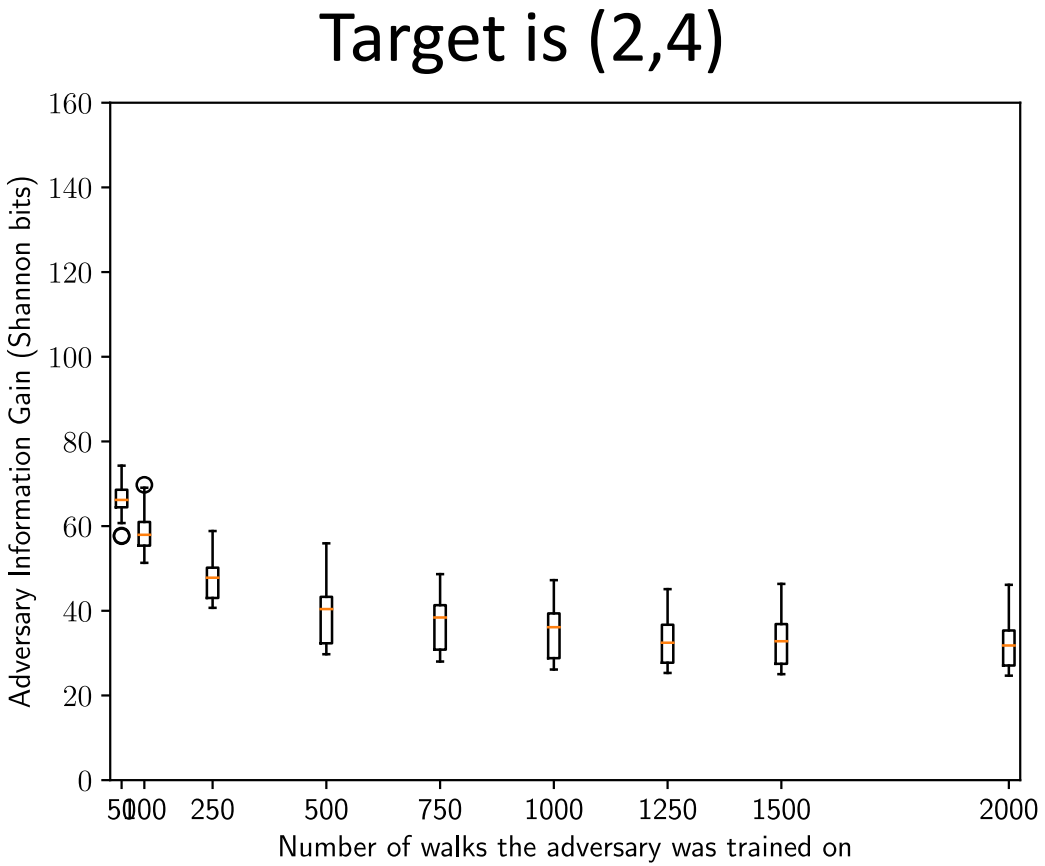
Target is (4,2)

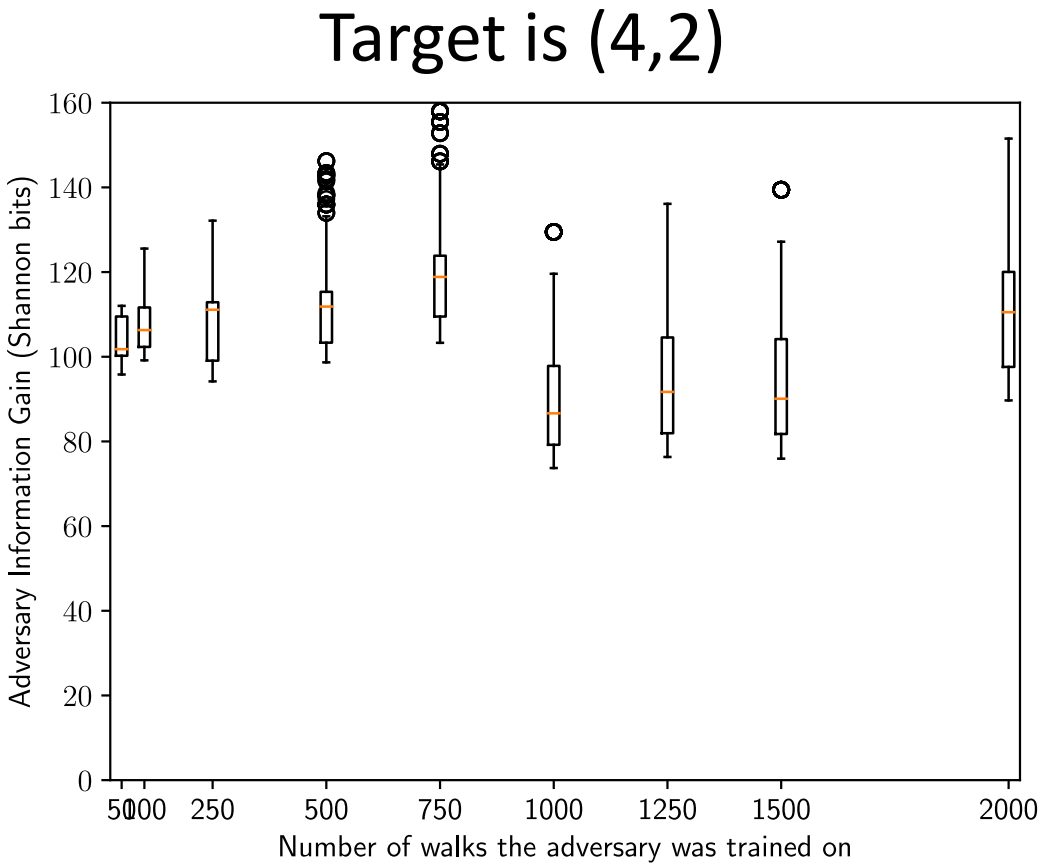


Adversary trained on observations going to (2,4)

Adversary self-information

Lancaster University

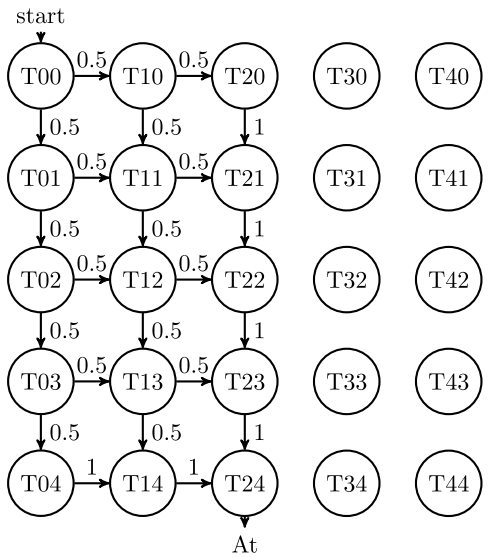




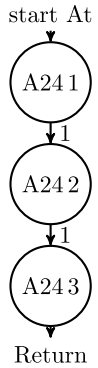
Adversary trained on observations going to (2,4)

Manual System Transformation

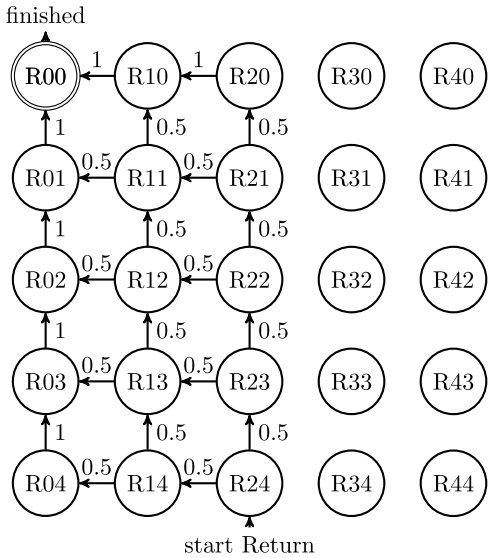
- Make the sensitive action commonplace
- Instead of going directly to a target, go via a different location



a: Towards States



b: At States

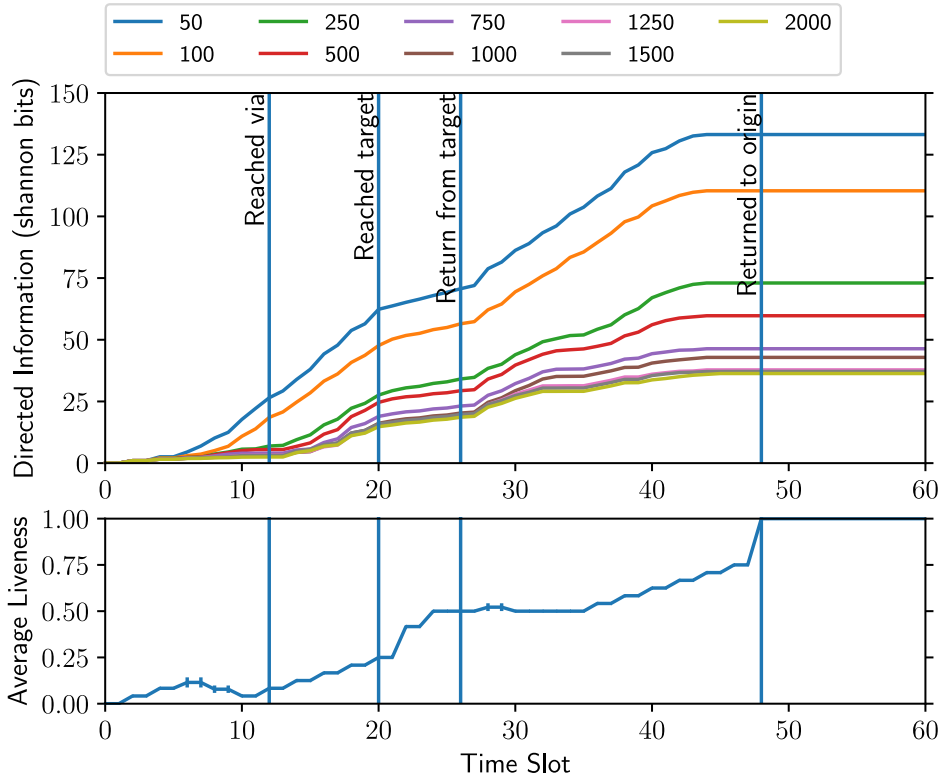


c: Return States

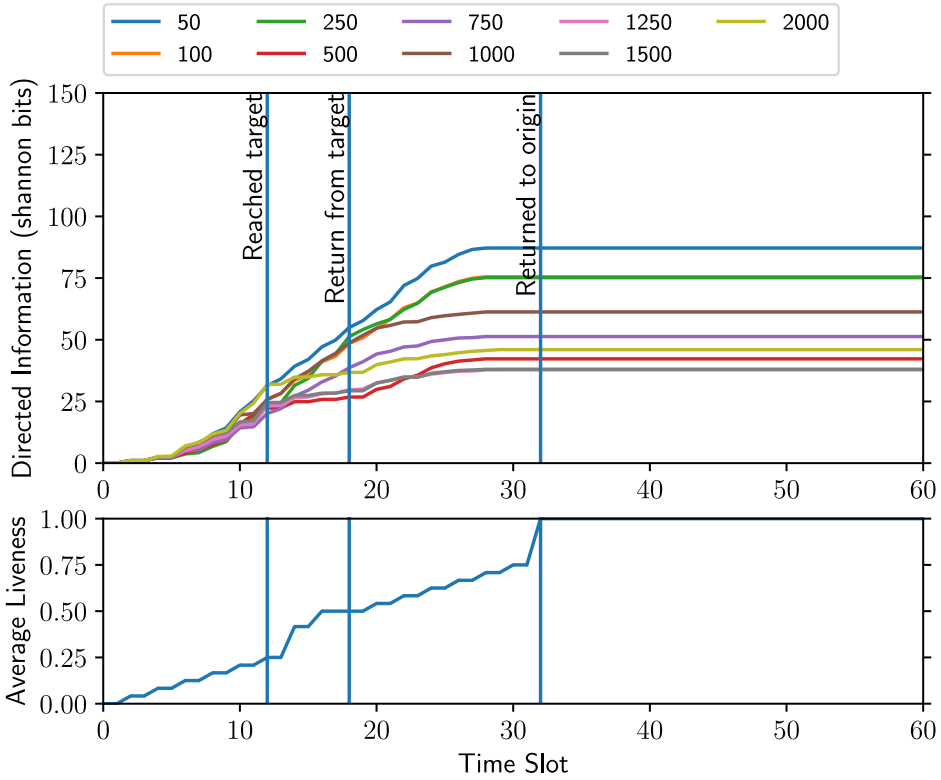
How much information is conveyed when the system is changed?

Lancaster University

Target is (2,4) via (4,2)



Target is (4,2)

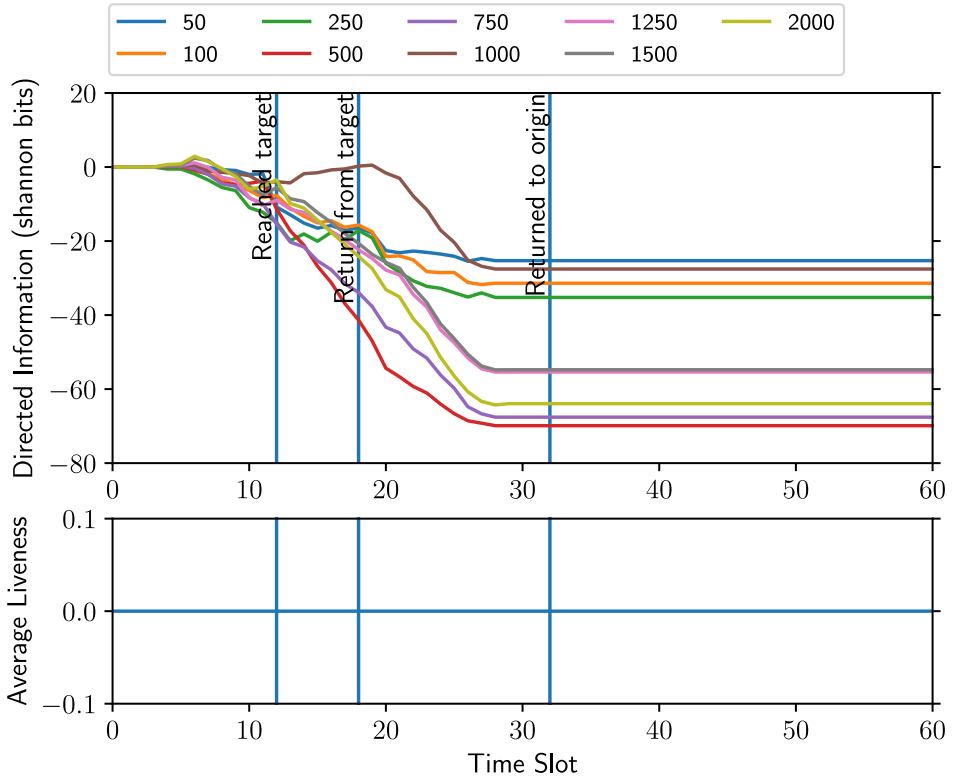


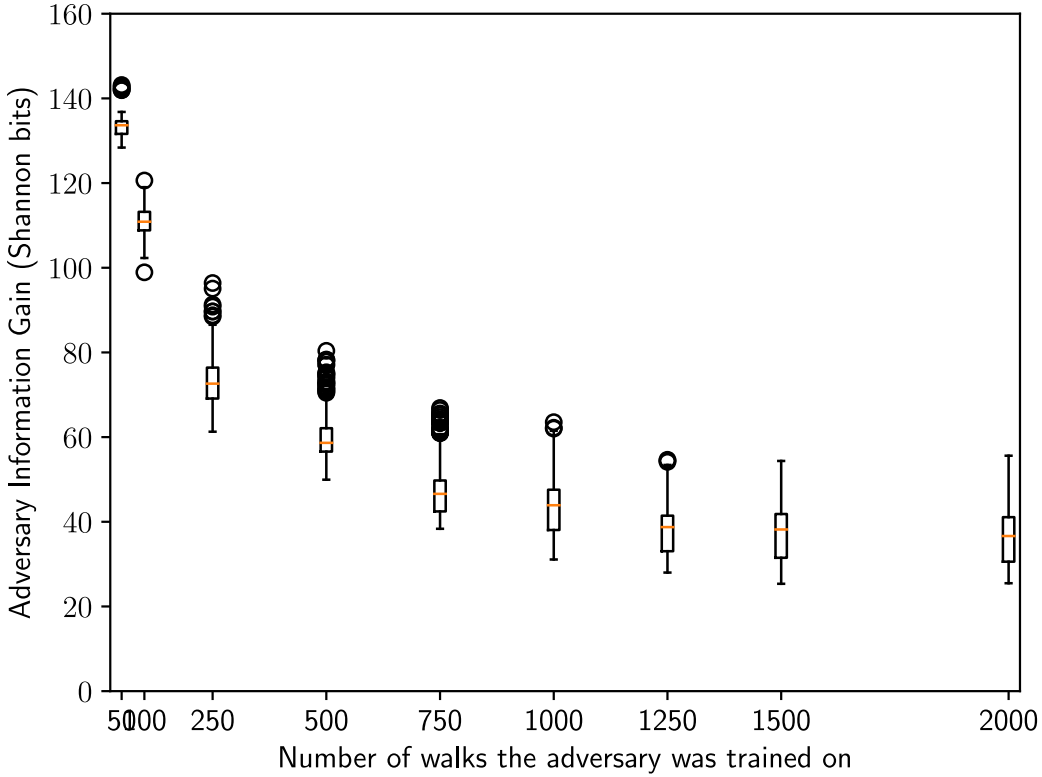
Adversary trained on observations going to (2,4) via (4,2)

How much information is conveyed when the system is changed?

Lancaster University

Difference: (4,2) when trained on (2,4) and (2,4) via (4,2)





Challenges

Lancaster University

Using Directed Information

- Need to have a model of system and adversary that provide the three probability distributions
- Computation of Directed Information is expensive (factorial over states, actions and observations) – Need special cases or use estimators
- Not always easy to interpret

Using Markov chains

- Systems may not be Markovian

J. Jiao, H. H. Permuter, L. Zhao, Y. -H. Kim and T. Weissman, "Universal Estimation of Directed Information," in *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6220-6242, Oct. 2013, doi: 10.1109/TIT.2013.2267934.

Why not differential privacy?

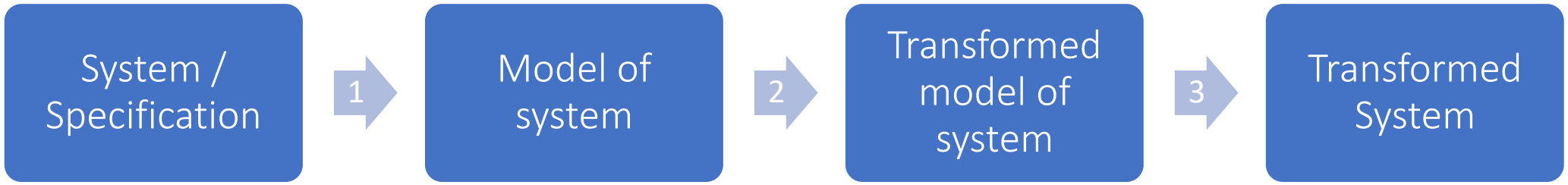
Lancaster University

- Differential Privacy (DP): Provide bounds on information release from a database
- Problem space fits tweaked information theory perspective
  - Problem: System revealing information to an observing adversary
  - Information theory: How much noise should be added to a communication channel to reduce the information conveyed?
- Harder to link problem space to DP
  - What is the database?
  - Cyber-physical system will change between observations
  - Limits to how noise can be added to cyber-physical systems (feasibility, safety, liveness, …)
  - Data manipulations techniques insufficient, action/state also need noise

Lancaster University

Arbitrary System Transformation

Process to transform system

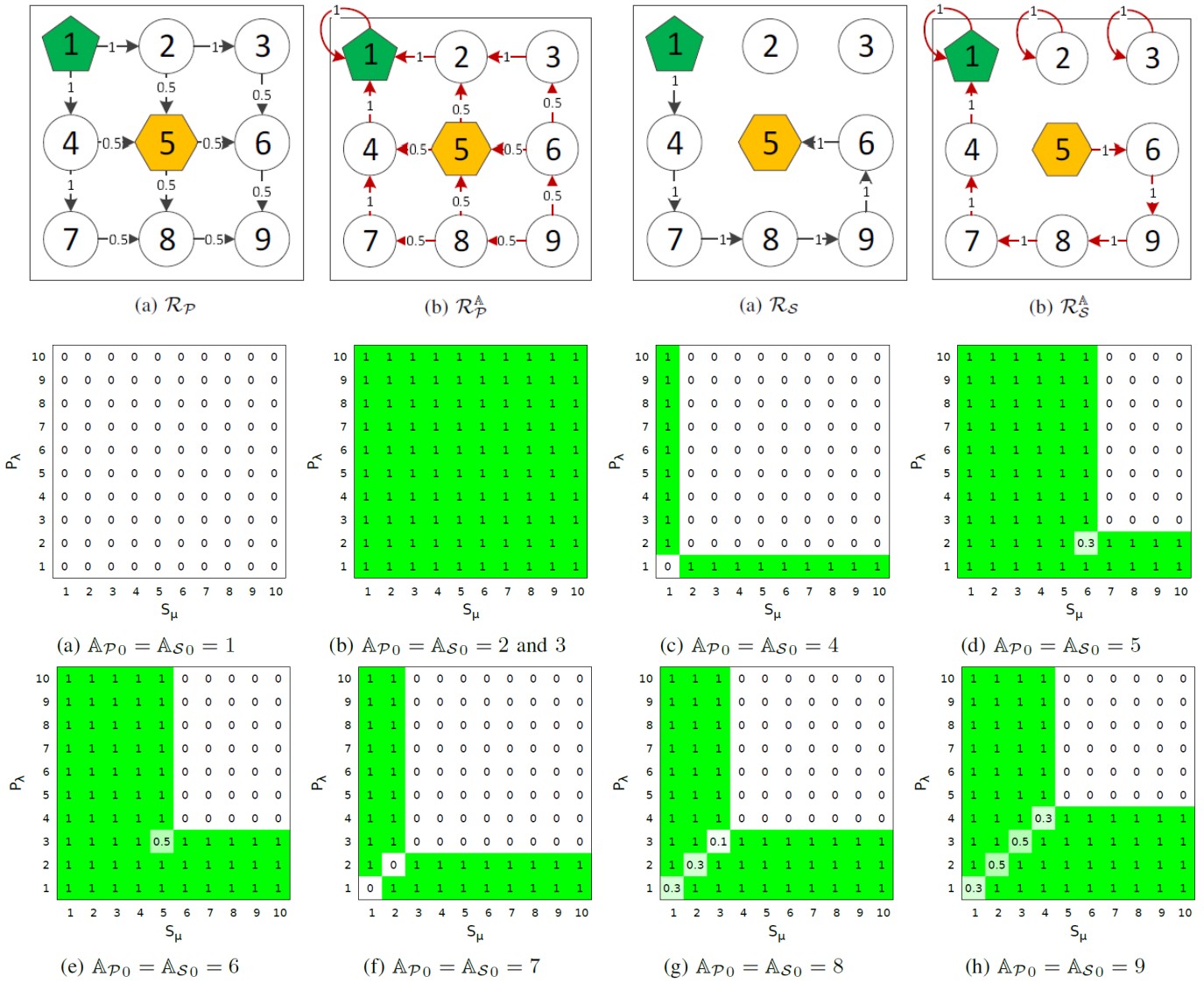


- Focusing on automating step 2 – get system with lower information loss
- Step 1 is reasonably automatable
- Step 3 is challenging to automate

Enforce divergence within time limit

- Requires that adversary does not observe a sensitive action within some time limit
- System P: does sensitive actions
- System S: does sensitive actions, but has been transformed to diverge from P for a safety period
- Jensen-Shannon Divergence:

$$\mathrm{JSD}(P_\lambda \| S_\mu) = H\left(\frac{P_\lambda + S_\mu}{2}\right) - \frac{1}{2}\left(H(P_\lambda) + H(S_\mu)\right)$$

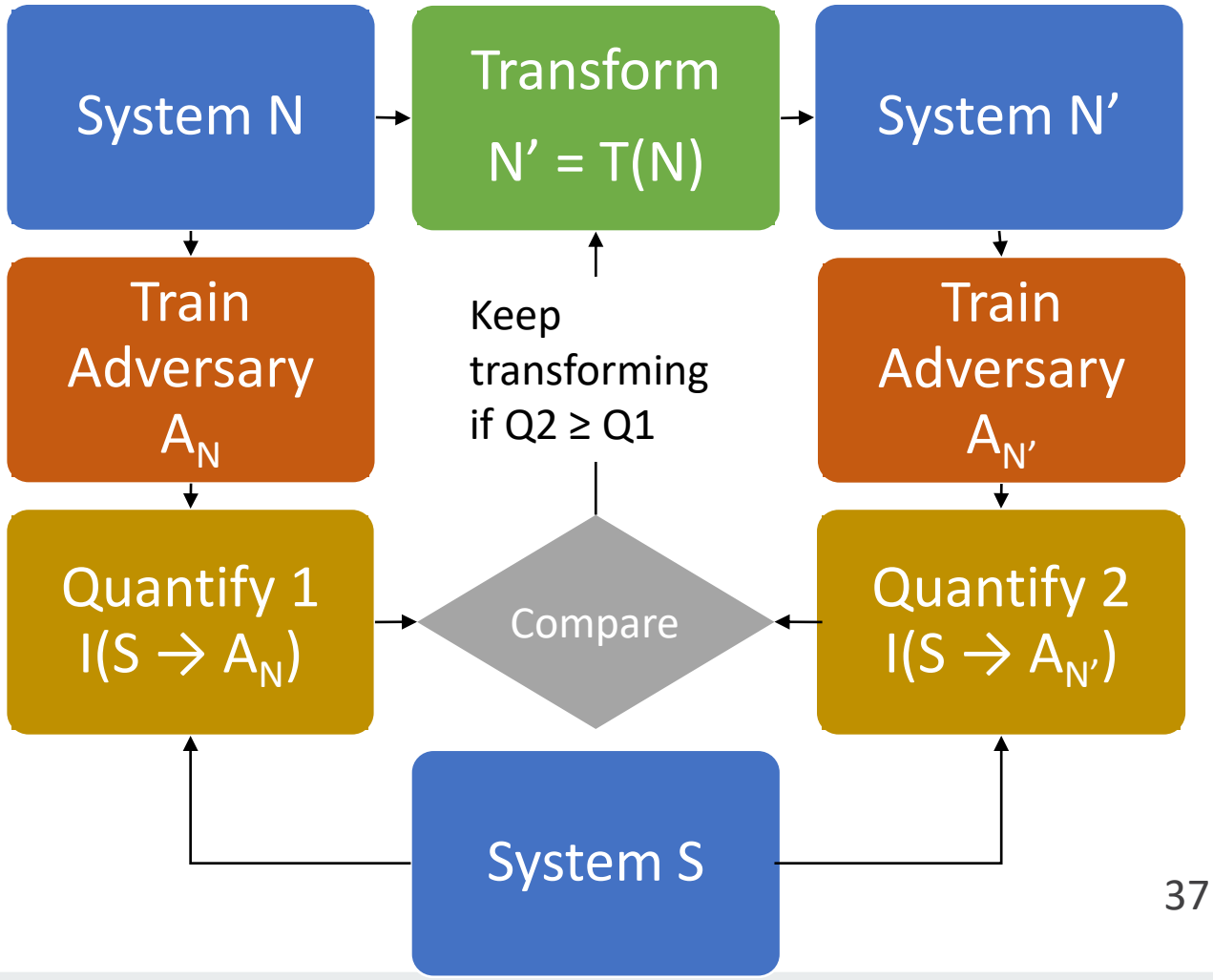


M. Bradbury and A. Jhumka. Quantifying Source Location Privacy Routing Performance via Divergence and Information Loss. *IEEE Transactions on Information Forensics and Security*, 17:3890–3905, 2022. doi:10.1109/TIFS.2022.3217385.

Solving the General Problem

Lancaster University

- System N: does non-sensitive actions
- System S: does sensitive actions
- Problem: Transform N into N', such that when adversary observes S, less information to conveyed to an Adversary:
  $I(S \rightarrow A_{N'}) < I(S \rightarrow A_N)$
- Potential problem:
  $I(N' \rightarrow A_{N'}) \geq I(N \rightarrow A_N)$
- Reminder: Assuming a new system



Unsuccessful Attempts

Lancaster University

Various attempts that have not worked:

- Markov Decision Processes
  - Reinforcement Learning to obtain the transformed N'
- Linear Programming
  - Find optimal N', with liveness constraints that minimises privacy loss

Issue: Developing suitable reward / objective functions

- Quantification requires an adversary is trained on the system being generated by the technique

In Progress Approach: Genetic Algorithm

Lancaster University

- Allow for arbitrary fitness functions
  - Including training an adversary belief model
- Convert Markov chain probability transitions to genes
- When converting back normalise to ensure valid Markov chain
- Issues:
  - Very slow (due to training adversary belief on every system generated)
  - Mutation can easily lead to loss of liveness
  - Potential lack of freedom in drone example to find a transformation

Challenges

Lancaster University

Transformation

- Need to ensure system remains live
- Transformation can be computationally expensive
- Very large systems may be hard to transform
- Encoding sufficient flexibility into the model is important
  - May limit exploration of options to change actions taken by the system

Conclusions

Lancaster University

- Many existing context privacy solutions
- Novel systems / environments likely to need new context privacy techniques quickly - Existing approach to develop techniques is too slow
- Instead:
  1. Solve the problem in general (Quantification, Technique Design)
  2. Design domain-specific translators
  3. Test translated technique in real-world environment

Lancaster University

Thank you for attending, any questions?