UCLA | Samueli School of Engineering

NSF #1740047

Distorting an Adversary's View in Cyber-Physical Systems

Christina Fragouli, Suhas Diggavi, Paulo Tabuada

University of California, Los Angeles

Labs: Algorithmic Research in Networked Information Flow (ARNI), Laboratory of Information Theory and Communication Systems (LICOS), The Cyber-Physical Systems Laboratory (CyPhyLab)

Reliable localization using landmarks [1]

Motivation

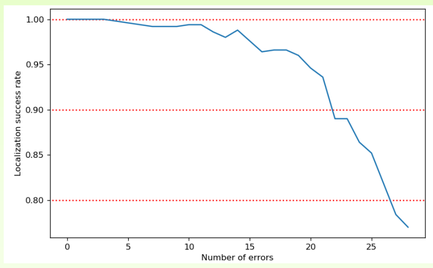
A fully autonomous vehicle should be able to self localize even in GPS-denied environments (e.g., when an adversary spoofs GPS)

Goal

- Leverage sensors together with a preloaded map of landmarks to self localize **quickly despite errors in landmark identification**
- Provide theoretical guarantees on the expected performance

Approach

- **Fast:** use unique sequences of plentifully available landmarks
- **Robust to errors:** connect to error correction in coding theory (where codeword = a group of landmark sequences ending at same location)
- **Efficient decoding:** using a Viterbi-like polynomial time algorithm
- **Probabilistic guarantees:** translate Hamming distance properties to probabilistic guarantees on the worst-case length of landmark sequences that need to be sensed for localization, given a number of expected errors
- **Numerical evaluation:** using landmarks extracted from map of Washington DC



Secure Time-Series Communication [2]

Dynamical Control System

$$X_{t+1} = A\,X_t + B\,U_t + wt$$
$$Y_t = C\,X_t + v_t$$

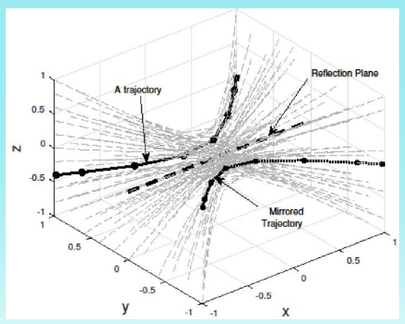
- We define two distortion measures for sequential data i.e., state transitions of a control system:
  - Expected Distortion: $D_E = \frac{1}{T}\sum_t \left(X_t - \hat{X}_t\right)^2$
  - Worst Case Distortion: $D_W = \min_t \min_{\tau_t(X_t,K)} \left(X_t - \hat{X}_t\right)^2$

Goal:

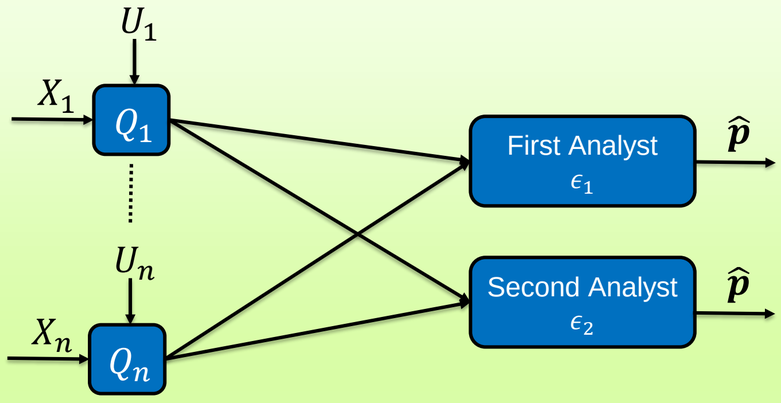
- Maximize Distortion by designing encoding functions τ_t

For Average Case:

- Mirror across hyperplanes for certain symmetric distributions
- Hyperplane can be picked with just one bit of key



Multi-Level Privacy [3]



Model:

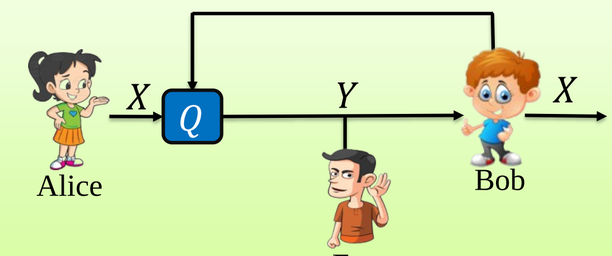
- n users, each has a sample $X_i \sim \boldsymbol{p}$
- d analysts want to estimate $\boldsymbol{p}$
- Each user i has a random key U_i

Goal:

- Design DP-mechanisms $\{Q_i : i \in [n]\}$: $Q_i = f(X_i, U_i)$
- To preserve privacy of each user
- To minimize the risk minimization of each analyst

$$r^{\ell}_{\epsilon,R,n,k} = \inf_{\hat{\boldsymbol{p}}} \inf_{Q_i} \sup_{\boldsymbol{p}} E[\ell(\hat{\boldsymbol{p}}(Y^n), \boldsymbol{p})]$$

Results [3]: Private-Recoverability



Necessary and Sufficient Conditions to Design Q

- The number of keys must be at least the same as the output size that must be at least the same as the input size: $|U| \geq |Y| \geq |X|$
- The entropy of the private key must satisfy: $H(U) \geq H(U^s_{min})$

$$U^s_{min} \sim q^s_{min} = \left[\frac{e^{\epsilon}}{s(e^{\epsilon}-1)+k}, \cdots, \frac{1}{s(e^{\epsilon}-1)+k}\right]$$

- The entropy of this key is less than what is required in one-time pad for perfect privacy

Privacy in Control over the Cloud [4]

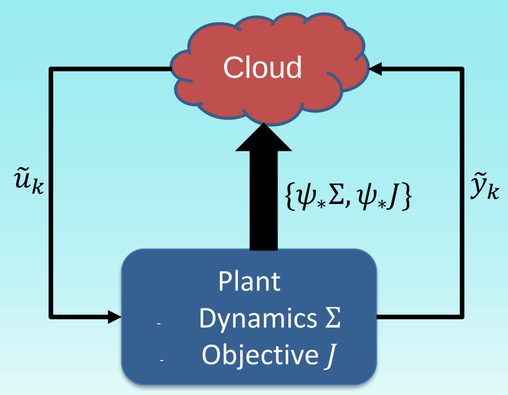
Motivation

- Control input can be calculated by minimizing an objective function (e.g., model predictive control)
- Control over the cloud requires communication of *private data* - vulnerable to *eavesdropping* attacks

Results

- Created a lightweight encoding scheme using isomorphisms of control systems
- Cloud is unable to learn the state, the dynamics, or the objective
- Proposed a measure of privacy (in terms of the dimension of uncertainty set)
- Quantified privacy loss with side knowledge

Minimize $\psi_*\mathrm{J}$ w.r.t dynamics $\psi_*\Sigma$



Secret key: ψ

Algorithm

1) **Handshaking**: Plant encodes Σ and J with ψ, and sends them to the cloud
2) **Plant operation:**
   - **Encoding:** Measure y_k and send encoded $\tilde{y}_k = \psi_* y_k$ to the cloud
   - **Optimization:** Cloud uses $\tilde{y}_k$ to find input $\tilde{u}_k$ minimizing $\psi_*\Sigma$ w.r.t dynamics $\psi_* J$, send $\tilde{u}_k$
   - **Decoding:** Decode $u_k = \psi_*^{-1}\tilde{u}_k$ and apply it to the actuators

References

[1] J. C. Rebanal, Y. H. Ezzeldin, C. Fragouli, P. Tabuada, " A coding approach to localization using landmarks, " GLOBECOM 2020

[2] G. Agarwal, M. Karmoose, S. Diggavi, C. Fragouli, P. Tabuada. "Distorting an adversary's view in cyber-physical systems", IEEE Transactions on Automatic Control 2021 vol. 66, issue 4 ,pp. 1588-1601

[3] A.M. Girgis, D. Data, K. Chaudhuri, C. Fragouli, S. Diggavi. "Successive Refinement of Privacy", IEEE Journal on Selected Areas in Information Theory, vol. 1, no. 3, pp. 745-759, Nov. 2020

[4] A. Sultangazin, P. Tabuada. "Symmetries and Privacy in Control Over the Cloud: Uncertainty Sets and Side Knowledge", IEEE Transactions on Automatic Control, vol. 66, no. 2, pp. 538-549, Feb. 2021

Publication output

The project has overall resulted in more than 30 publications in top tier journals and conferences

REU achievement

This project has supported 6 REU students that have worked on a variety of cross-cutting projects