Trustworthy Video Analytics

Yuan Hong

Assistant Professor
Cybersecurity Program Director
Director, DataSec Lab
Department of Computer Science
Illinois Institute of Technology

CSIRO, Data61
August 26, 2021

ILLINOIS TECH

Overview

- **<u>Security</u>** in Video Analytics (**Slides Removed**)
  - ➢ Universal 3-Dimensional (U3D) Perturbations for Black-Box Attacks on Video DNNs – **Oakland'22**

- **<u>Privacy</u>** in Video Analytics
  - ➢ A Flexible Platform for Video Analytics with Differential Privacy (VideoDP) – **PETS'20**

ILLINOIS TECH

Video Privacy

- Huge amounts of sensitive information in videos (e.g., vehicle plates, human faces/bodies, and name tags) may raise privacy concerns

Video Surveillance

Traffic Monitoring

Video Privacy Techniques

- **Computer Vision based Protection**

  [1] Privacy-Preserving Action Recognition using Coded Aperture Videos

  [2] Pre-Capture Privacy for Small Vision Sensors

  Informal privacy guarantee

- **Cryptographic Protocols**

  [3] Privacy-Preserving Outsourcing Computation of Feature Extractions Over Encrypted Image Data

  [4] PrivacyCam: A Privacy Preserving Camera using uCLinux on the Blackfin DSP

  Limited to specific applications (e.g., action recognition or SIFT)

  Expensive computation

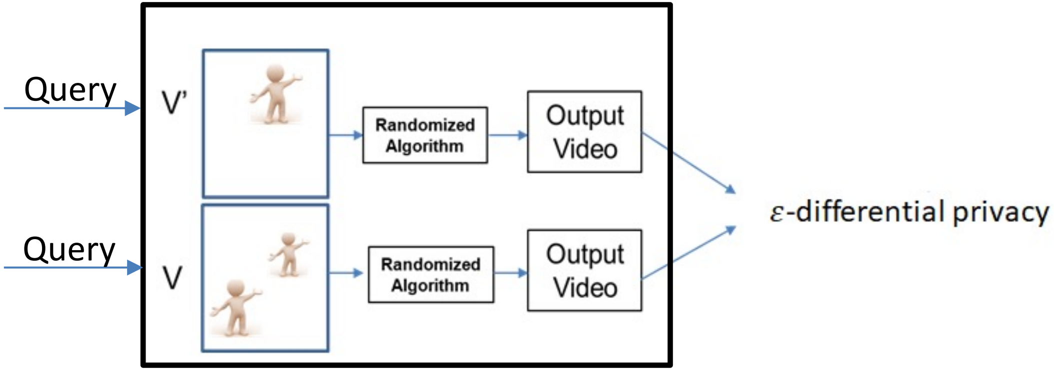
[1] Wang, et al. "Privacy-preserving action recognition using coded aperture videos." *CVPR 2019*.
[2] Pittaluga and Koppal. "Precapture privacy for small vision sensors." *IEEE TPAMI 2017*.
[3] Hu, et al. "Securing SIFT: Privacy-preserving outsourcing computation of feature extractions over encrypted image data." *IEEE Transactions on Image Processing 2016*.
[4] Chattopadhyay and Boult. "PrivacyCam: a privacy preserving camera using uCLinux on the Blackfin DSP." *CVPR 2007*.

Differential Privacy for Videos

Differential Privacy in Videos (protecting each visual element)

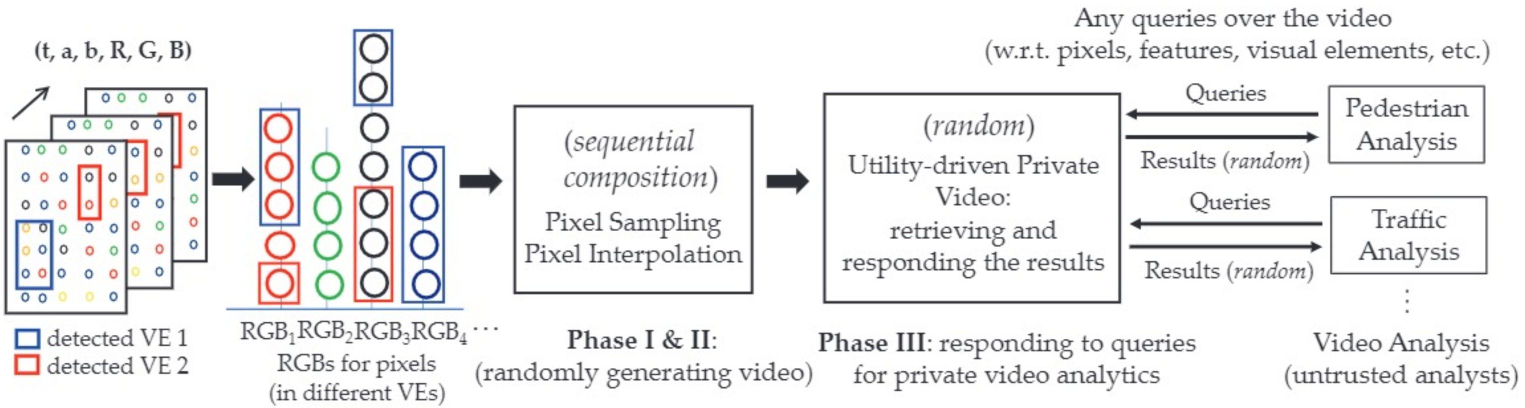
Randomized mechanism A provides ϵ-differential privacy if for any two input videos V and V' that differ in any visual element γ (e.g., object or human), and for any output $O \in range(A)$, we have

$$e^{-\varepsilon} \leq \frac{\Pr[A(V) = O]}{\Pr[A(V') = O]} \leq e^{\varepsilon}$$



Note: 1. Background scene can be a visual element (if requested for protection)
2. Similar to generic differential privacy notion, it can be relaxed as (ε, δ)-DP

VideoDP Framework



Video Pre-processing | Utility-driven Private Video | Video Queries

- Detecting and tracking all objects (assign an ID for every object in all the frames)
- Generating utility-driven private video:
  - Sampling pixels for the video with differential privacy
  - Interpolating unsampled pixels (post-processing DP results)
- Privately querying the randomly generated video (e.g., traffic monitoring, and street surveillance): post-processing DP results

Phase I: Pixel Sampling

1. All the RGBs $\theta_r, r \in [1, n]$ follow **sequential composition** in the sampling process and satisfy

$$\sum_{r=1}^{n} \epsilon(\widetilde{\theta}_r) = \epsilon$$

Impractical to *allocate privacy budget for every RGB (n can be as large as* 256^3*) – negligible budgets*

2. Pixels are categorized into Case (1), (2), (3) with their RGBs

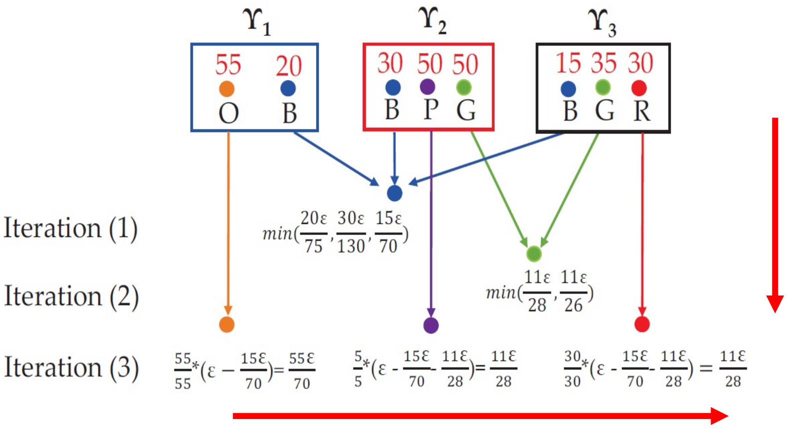
   - Case (1): the RGBs in any visual element but not the background
   - Case (2): the RGBs in the background but not any of the visual element
   - Case (3): the RGBs in the background and at least one visual element

3. Case (3): assign privacy budgets to a subset of RGBs and derive sampled pixel counts for them for satisfying differential privacy

4. For each RGB, randomly sample pixels with DP to generate a raw output video (with sparse pixels)

Phase I: Budget Allocation

1. Derive the optimal k RGBs in each visual element (maximizing utility)
2. Partition the visual element into k multi-scales and choose the top frequent RGB in each scale
3. The criteria for allocating budget:
   - Privacy budgets based on the count distributions of RGBs in different VEs
   - Fully utilizing the privacy budget ϵ



Allocating all the privacy budgets

Not exceeding the privacy bound

Phase I: Counts Computation

RGB	Video	VE$_j$	Optimal Counts
RGB 1	c_1	c_1^j	x_1
RGB 2	c_2	c_2^j	x_2
…	…	…	

Sampling Pixels for Each RGB → …

$$Pr[\mathcal{A}(V(\widetilde{\theta}_r)) = O(\widetilde{\theta}_r)] = 1/\binom{c_r}{x_r}$$

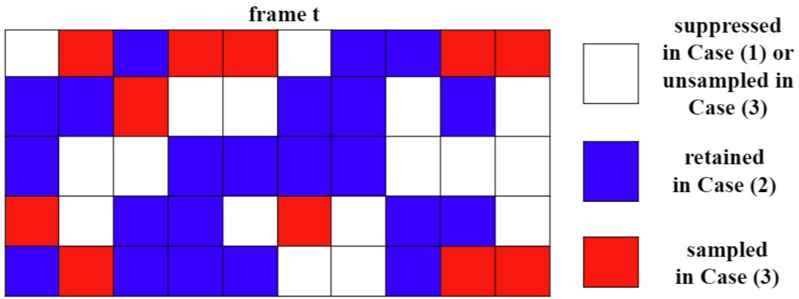
$$Pr[\mathcal{A}(V'(\widetilde{\theta}_r)) = O(\widetilde{\theta}_r)] = 1/\binom{c_r - c_r^j}{x_r}$$

$$\Longrightarrow e^{-\epsilon(\widetilde{\theta}_r)} \leq \binom{c_r}{x_r}/\binom{c_r - c_r^j}{x_r} \leq e^{\epsilon(\widetilde{\theta}_r)}$$

$$\max\{x_r | \forall j \in [1, n], \binom{c_r}{x_r}/\binom{c_r - c_r^j}{x_r} \leq e^{\epsilon(\widetilde{\theta}_r)}\}$$

The left-hand side is monotonic on x_r

Phase II: Pixel Interpolation



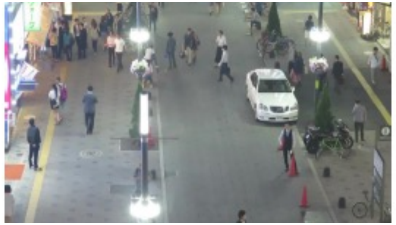
- **Post-processing** the raw output video (sampled with DP)
  - ➢ All pixels in Case (1) (pixels with a unique RGB) are suppressed
  - ➢ All pixels in Case (2) (pixels in the background) are retained
  - ➢ Pixels in Case (3) are partially sampled
  - ➢ Estimating the RGBs for missing pixels (suppressed or unsampled) using Bilinear interpolation

Experimental Datasets

Multiple Object Tracking (MOT) dataset with pedestrians and vehicles
UCSD Anomaly Detection (UAD) dataset with crowded pedestrians
Boxy Vehicle Detection (BVD) dataset with crowded vehicles

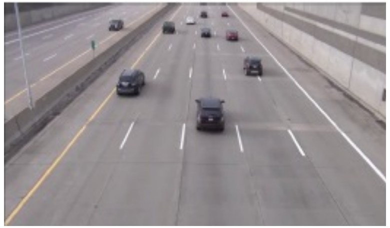
Table 1. Characteristics of Experimental Datasets

Datasets	Avg. Resolution	Video #	Avg. Frame #
MOT	1920×1080	15	846
UAD	740×480	24	180
BVD	2464×2056	5	1200



MOT16-04

UADTest-001



BVD-Highway

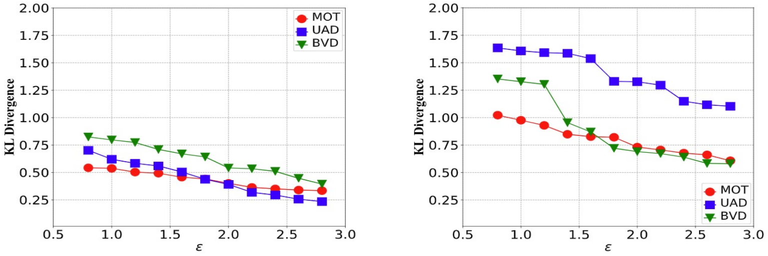
Evaluations

- RGB Count Distribution
  - **KL-divergence** measures the **count distribution** difference of the RGBs in the input and private videos
- RGB Values at Pixel Level
  - **Mean square error (MSE)** measures the difference between all the RGB values in the input and private videos
- Detection and Tracking Accuracy in the Private Video
  - **Precision** and **Recall** in the **entire video**
  - VE **detection accuracy** in **each frame**
- Case study with specific queries in applications
  - **Small sensitivity** and **large sensitivity**
  - Benchmarking with the **PINQ** platform

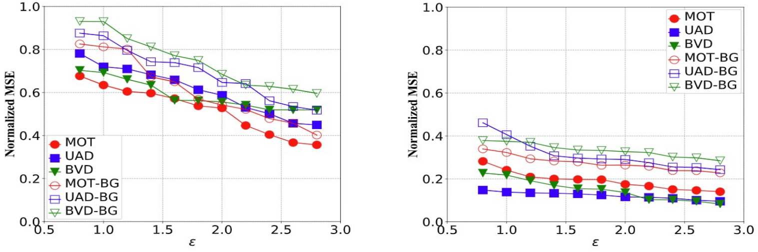
ILLINOIS TECH

Pixel-Level and Detection/Tracking Utility

- KL-divergence and MSE



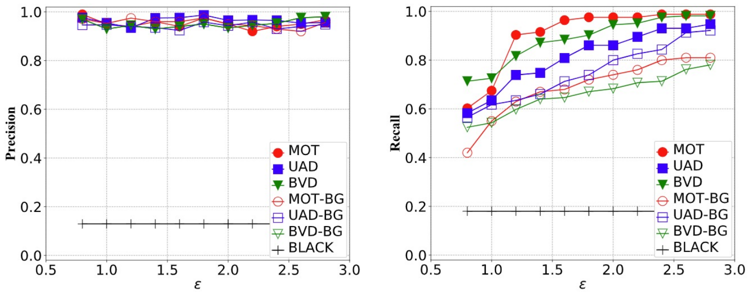
(a) KL vs ϵ (b) KL vs ϵ (Background VE)



(c) MSE vs ϵ (after Phase I) (d) MSE vs ϵ (after Phase II)

- Detection and Tracking

(VE counts precision and recall)

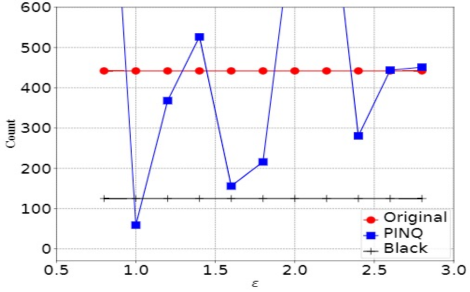


(a) Precision vs ϵ (b) Recall vs ϵ

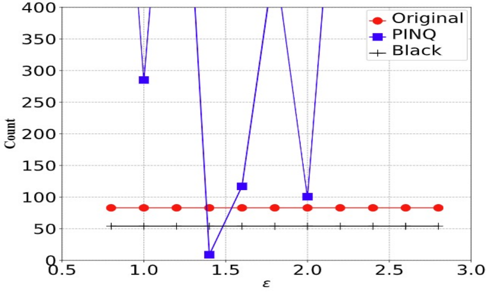
Fig. 5. Visual Elements Detection and Tracking

ILLINOIS TECH

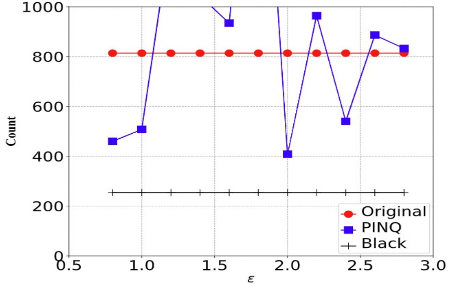
Detection and Tracking (vs. PINQ and Black)



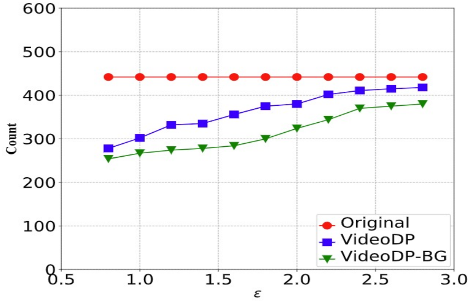
(a) PINQ and Black (MOT)



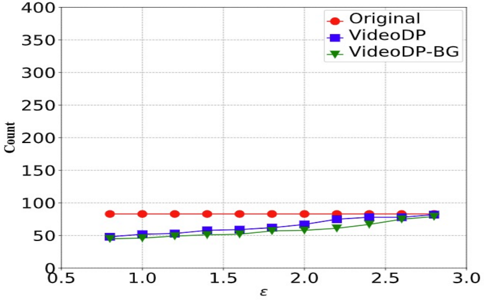
(c) PINQ and Black (UAD)



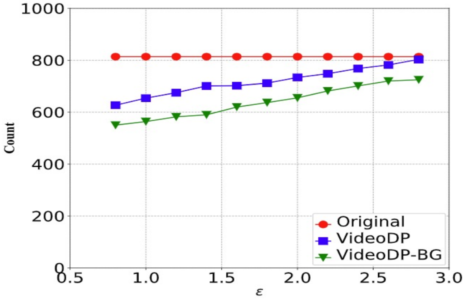
(e) PINQ and Black (BVD)



(b) VideoDP (MOT)



(d) VideoDP (UAD)

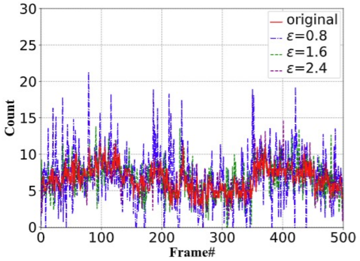


(f) VideoDP (BVD)

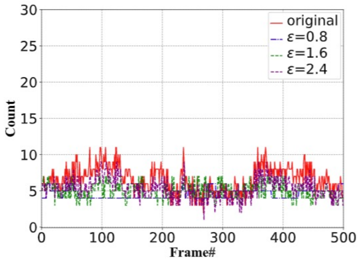
ILLINOIS TECH

Case Studies

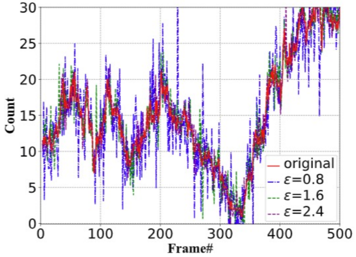
- VE Density (Small Sensitivity)



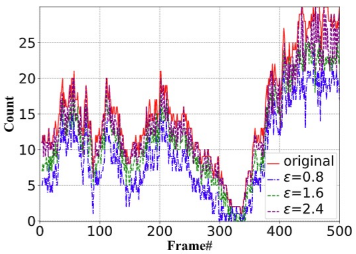
(a) Vehicle (PINQ)



(b) Vehicle (VideoDP)

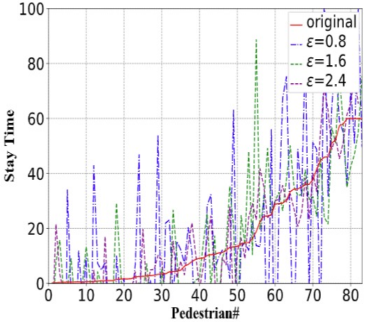


(c) Pedestrian (PINQ)

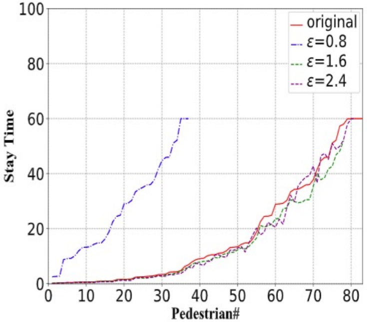


(d) Pedestrian (VideoDP)

- Pedestrian Stay Time (Large Sensitivity)



(a) Pedestrian (PINQ)



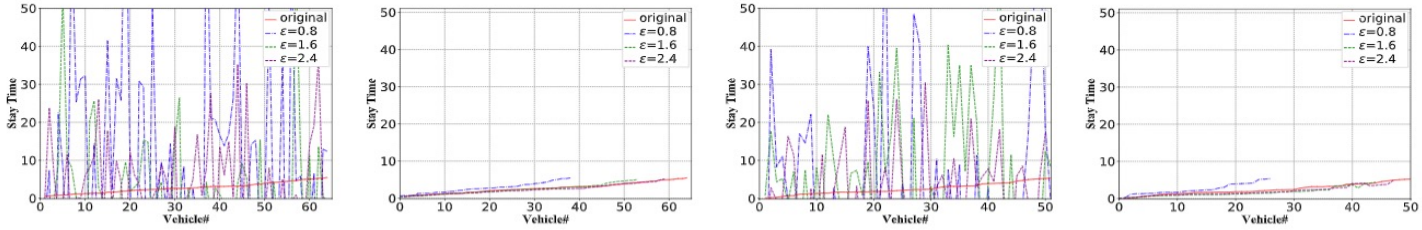
(b) Pedestrian (VideoDP)

Fig. 7. Pedestrian Stay Time in PED

ILLINOIS TECH

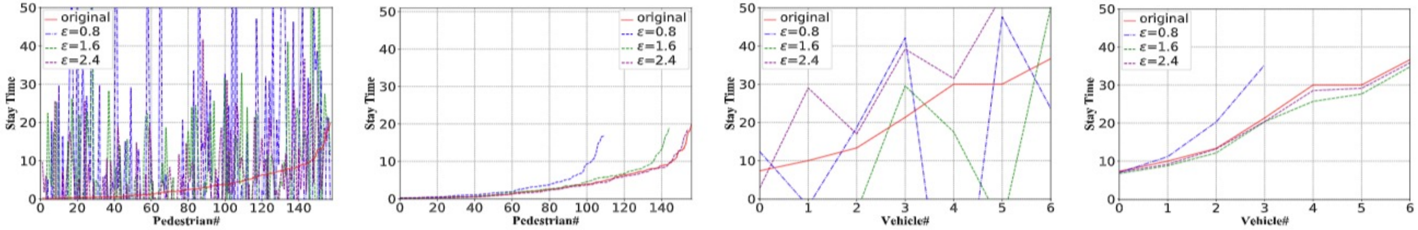
Case Studies (Cont'd)

- Vehicle Stay Time (Large Sensitivity)



(a) Upstream (PINQ) (b) Upstream (VideoDP) (c) Downstream (PINQ) (d) Downstream (VideoDP)

Fig. 8. Vehicle Stay Time in VEH



(a) Pedestrian (PINQ) (b) Pedestrian (VideoDP) (c) Vehicle (PINQ) (d) Vehicle (VideoDP)

Fig. 9. Pedestrian and Vehicle Stay Time in PV

Representative Frames

(a) Original (b) $\epsilon = 0.8$ (Phase I) (c) $\epsilon = 1.6$ (Phase I) (d) $\epsilon = 0.8$ (Phase II) (e) $\epsilon = 1.6$ (Phase II)

Fig. 13. Representative Frames in the Random Output Video of PED (available for differentially private queries/analysis)

(a) Original (b) $\epsilon = 0.8$ (Phase I) (c) $\epsilon = 1.6$ (Phase I) (d) $\epsilon = 0.8$ (Phase II) (e) $\epsilon = 1.6$ (Phase II)

Fig. 14. Representative Frames in the Random Output Video of VEH (available for differentially private queries/analysis)

Thank You!
Questions?

Yuan Hong
Email: yuan.hong@iit.edu
Webpage: http://cs.iit.edu/~yhong/

Acknowledge of Support

My Students (Contributed to the Works)

Shangyu Xie
Han Wang