Sequential Change Detection with Differential Privacy

Liyan Xie* and Ruizhi Zhang†

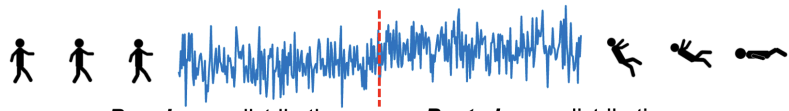
*Department of Industrial and Systems Engineering, University of Minnesota,
†Department of Statistics, University of Georgia

The 9th International Workshop in Sequential Methodologies

Outline

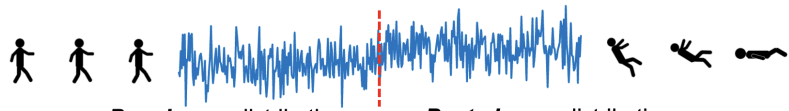
- Motivation: Private Online Change Detection
- Proposed: Differentially-Private CUSUM Test (DP-CUSUM)
- Theoretical Guarantees: False Alarm and Detection Delay
- Extensions and Numerical Results

What is online change detection?



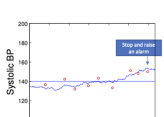
$$x_1, \ldots, x_{\tau-1} \overset{iid}{\sim} f_0, \quad x_\tau, x_{\tau+1}, \ldots \overset{iid}{\sim} f_1.$$

What is online change detection?



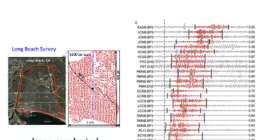
$$x_1, \ldots, x_{\tau-1} \overset{iid}{\sim} f_0, \quad x_\tau, x_{\tau+1}, \ldots \overset{iid}{\sim} f_1.$$

- Application examples



Bloop pressure monitoring

Anomaly detection



Seismic event monitoring

Privacy threats in modern data streams

- Growing aggregation and use of personal data
  - Google Maps: traffic monitoring, location tracking
  - YouTube: recommendations
  - Wearable devices: health data

Half a billion Facebook users' information posted on hacking website, cyber experts say

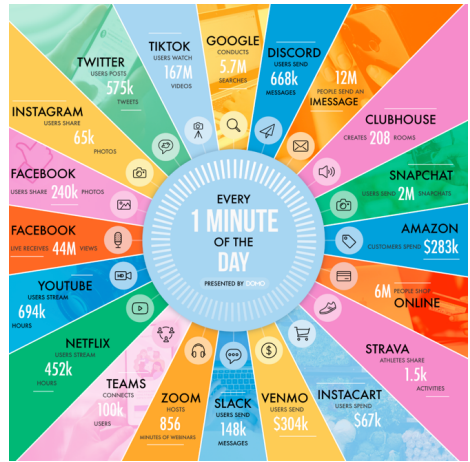
By Donie O'Sullivan, CNN Business
2 min read · Updated 7:01 AM EDT, Mon April 5, 2021

Share

Home > Security > Cybersecurity

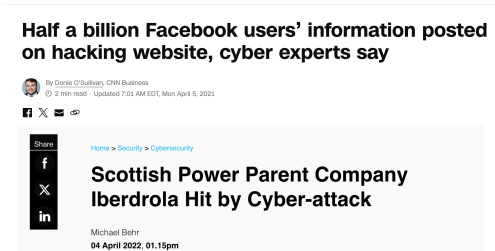
Scottish Power Parent Company Iberdrola Hit by Cyber-attack

Michael Behr
04 April 2022, 01.15pm



Privacy threats in modern data streams

- Growing aggregation and use of personal data
  - Google Maps: traffic monitoring, location tracking
  - YouTube: recommendations
  - Wearable devices: health data



Half a billion Facebook users' information posted on hacking website, cyber experts say

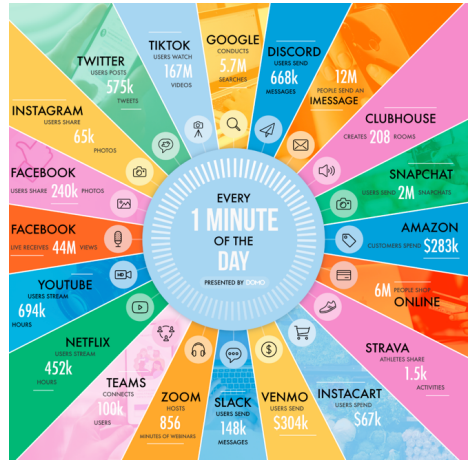
By Donie O'Sullivan, CNN Business
2 min read · Updated 7:01 AM EDT, Mon April 5, 2021

Share

Home > Security > Cybersecurity

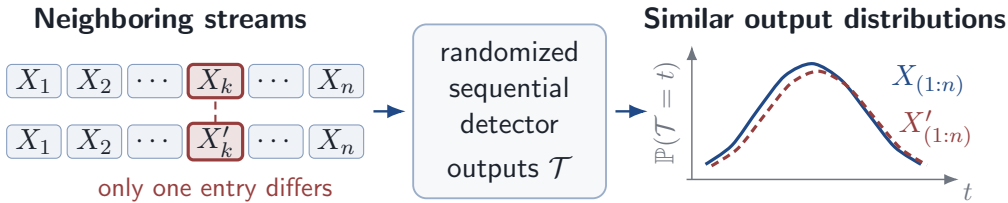
Scottish Power Parent Company Iberdrola Hit by Cyber-attack

Michael Behr
04 April 2022, 01.15pm

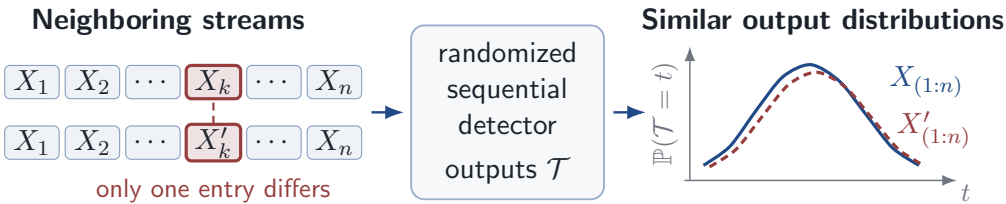


Research Question: How to detect distributional change in an online and **private** way?

What should privacy protect?



What should privacy protect?



Privacy Goal: Changing **one individual's sample** should only slightly change the **distribution of the alarm time** $\mathcal{T}$. Thus, the released detection output should not reveal whether X_k or X'_k was present.

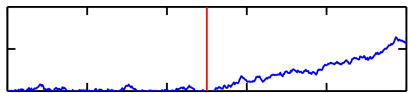
The optimal detector without privacy consideration

Data model:

$$x_1, \ldots, x_{\tau-1} \overset{iid}{\sim} f_0, \quad x_\tau, x_{\tau+1}, \ldots \overset{iid}{\sim} f_1.$$

- The optimal CUSUM detection procedure:

$$S_t = \max(0, S_{t-1}) + \ell(X_t), \quad \ell(X) = \log \frac{f_1(X)}{f_0(X)} \text{ (log-likelihood ratio).}$$



- Stopping rule:

$$T(b) = \inf\{t : S_t \geq b\}.$$

Page, Ewan S. "Continuous inspection schemes." Biometrika, 1954.

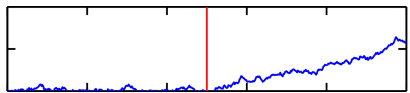
The optimal detector without privacy consideration

Data model:

$$x_1, \ldots, x_{\tau-1} \overset{iid}{\sim} f_0, \quad x_\tau, x_{\tau+1}, \ldots \overset{iid}{\sim} f_1.$$

- The optimal CUSUM detection procedure:

$$S_t = \max(0, S_{t-1}) + \ell(X_t), \quad \ell(X) = \log \frac{f_1(X)}{f_0(X)} \text{ (log-likelihood ratio).}$$



- Stopping rule:

$$T(b) = \inf\{t : S_t \geq b\}.$$

- **Non-private**: Release of S_t reveals the raw data X_t.

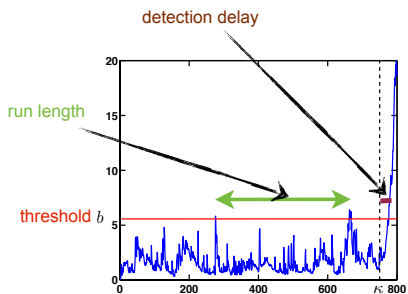
Page, Ewan S. "Continuous inspection schemes." Biometrika, 1954.

The optimal detector without privacy consideration

Data model:

$$x_1, \ldots, x_{\tau-1} \overset{iid}{\sim} f_0, \quad x_\tau, x_{\tau+1}, \ldots \overset{iid}{\sim} f_1.$$

CUSUM procedure minimizes the detection delay, subject to average run length $\geq \gamma$.

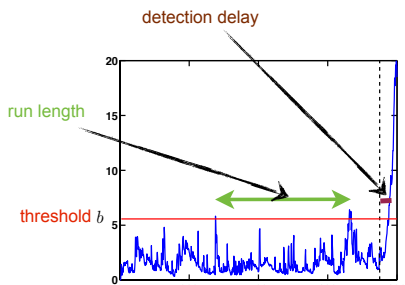


The optimal detector without privacy consideration

Data model:

$$x_1, \ldots, x_{\tau-1} \overset{iid}{\sim} f_0, \quad x_\tau, x_{\tau+1}, \ldots \overset{iid}{\sim} f_1.$$

CUSUM procedure minimizes the detection delay, subject to average run length $\geq \gamma$.



Goal: Design a private detection procedure without degrading the delay significantly.

Outline

- Motivation: Private Online Change Detection
- Proposed: Differentially-Private CUSUM Test (DP-CUSUM)
- Theoretical Guarantees: False Alarm and Detection Delay
- Extensions and Numerical Results

Differential privacy in sequential detection

Definition (ϵ-DP Sequential Detection)

A randomized sequential change detection procedure with stopping time $\mathcal{T}$ is said to be ϵ-differentially private (ϵ-DP), if for every pair of neighboring data streams $X_{(1:n)}, X'_{(1:n)}$ (differing in at most one entry), the distribution over the randomized stopping time $\mathcal{T}$ satisfies the differential privacy constraint,

$$\mathbb{P}_{\mathcal{T}}(\mathcal{T} = n \mid X_{(1:n)}) \leq e^{\epsilon}\, \mathbb{P}_{\mathcal{T}}(\mathcal{T} = n \mid X'_{(1:n)}),\ \forall\, n \geq 1,$$

where the probability is taken over the randomness in $\mathcal{T}$.

Differential privacy in sequential detection

Definition (ϵ-DP Sequential Detection)

A randomized sequential change detection procedure with stopping time $\mathcal{T}$ is said to be ϵ-differentially private (ϵ-DP), if for every pair of neighboring data streams $X_{(1:n)}, X'_{(1:n)}$ (differing in at most one entry), the distribution over the randomized stopping time $\mathcal{T}$ satisfies the differential privacy constraint,

$$\mathbb{P}_{\mathcal{T}}(\mathcal{T} = n \mid X_{(1:n)}) \leq e^{\epsilon}\, \mathbb{P}_{\mathcal{T}}(\mathcal{T} = n \mid X'_{(1:n)}),\ \forall\, n \geq 1,$$

where the probability is taken over the randomness in $\mathcal{T}$.

- No restriction on the data series length.
- Smaller ϵ implies more privacy.
- CUSUM is obviously not ϵ-DP.

DP-CUSUM: Main Idea

- Recall CUSUM statistic:

$$S_t = \max(0, S_{t-1}) + \ell(X_t), \quad \ell(X) = \log \frac{f_1(X)}{f_0(X)} \text{ (log-likelihood ratio)}$$

DP-CUSUM: Main Idea

- Recall CUSUM statistic:

$$S_t = \max(0, S_{t-1}) + \ell(X_t), \quad \ell(X) = \log \frac{f_1(X)}{f_0(X)} \text{ (log-likelihood ratio)}$$

- Add Laplace noise:

$$\tilde{S}_t = S_t + Z_t, \quad Z_t \sim \text{Lap}\left(\frac{2\Delta}{\epsilon}\right), \ \Delta = \sup_{x,y} |\ell(x) - \ell(y)| \text{ (sensitivity)}$$

DP-CUSUM: Main Idea

- Recall CUSUM statistic:

$$S_t = \max(0, S_{t-1}) + \ell(X_t), \quad \ell(X) = \log \frac{f_1(X)}{f_0(X)} \text{ (log-likelihood ratio)}$$

- Add Laplace noise:

$$\tilde{S}_t = S_t + Z_t, \quad Z_t \sim \text{Lap}\left(\frac{2\Delta}{\epsilon}\right), \Delta = \sup_{x,y} |\ell(x) - \ell(y)| \text{ (sensitivity)}$$

- Randomize the threshold:

$$\mathcal{T}(b) = \inf\{t : \tilde{S}_t \geq b + W\}, \quad W \sim \text{Lap}\left(\frac{2\Delta}{\epsilon}\right)$$

DP-CUSUM: Main Idea

- Recall CUSUM statistic:

$$S_t = \max(0, S_{t-1}) + \ell(X_t), \quad \ell(X) = \log \frac{f_1(X)}{f_0(X)} \text{ (log-likelihood ratio)}$$

- Add Laplace noise:

$$\tilde{S}_t = S_t + Z_t, \quad Z_t \sim \text{Lap}\left(\frac{2\Delta}{\epsilon}\right), \Delta = \sup_{x,y} |\ell(x) - \ell(y)| \text{ (sensitivity)}$$

- Randomize the threshold:

$$\mathcal{T}(b) = \inf\{t : \tilde{S}_t \geq b + W\}, \quad W \sim \text{Lap}\left(\frac{2\Delta}{\epsilon}\right)$$

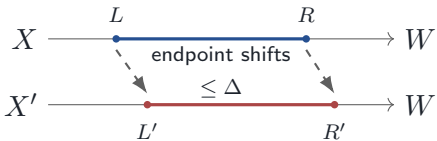
- The stopping rule $\mathcal{T}(b)$ is ϵ-DP. Still **recursive and** $O(1)$ **per step**.

Why DP-CUSUM is sequential ϵ-DP

- Fix t, the event "alarm at t" is
  $\{\mathcal{T} = t\} \iff \{L < W \le R\}$.
  - $L = M_{t-1} - b$,
    $M_{t-1} = \max_{j<t}(S_j + Z_j)$.
  - $R = Z_t + S_t - b$.

- Condition on the earlier noises $Z_1, \dots, Z_{t-1}$, neighboring streams only move the two boundaries:

$$|L' - L| \le \Delta, \quad |R' - R| \le \Delta.$$



$$\frac{f_Z(z-\Delta z) f_W(w-\Delta w)}{f_Z(z) f_W(w)} \le e^{\epsilon/2} e^{\epsilon/2} = e^{\epsilon}$$

$$\downarrow$$

$$\mathbb{P}(\mathcal{T} = t \mid X_{(1:t)}) \le e^{\epsilon} \mathbb{P}(\mathcal{T} = t \mid X'_{(1:t)}).$$

DP-CUSUM algorithm summary

Algorithm: DP-CUSUM Procedure

Input: Data stream $\{X_t\}$, privacy parameter ϵ, sensitivity Δ, threshold b

Output: Stopping time $\mathcal{T}$

Sample threshold noise $W \sim \text{Lap}\left(\frac{2\Delta}{\epsilon}\right)$. ▸ Laplace noise to threshold

Initialize $t \leftarrow 0$, $S_0 \leftarrow 0$, and $\tilde{S}_0 \leftarrow -\infty$.

while $\tilde{S}_t < b + W$ **do**

  $t \leftarrow t + 1$ and observe X_t.

  Update $S_t = \max(0, S_{t-1}) + \ell(X_t)$.

  Sample $Z_t \sim \text{Lap}\left(\frac{2\Delta}{\epsilon}\right)$. ▸ Laplace noise to detection statistic

  Set $\tilde{S}_t = S_t + Z_t$.

end while

Output $\mathcal{T} = t$; declare that a change occurred before $\mathcal{T}$.

Same recursive CUSUM update; only one fresh noise draw per time step.

Theoretical guarantees

Theorem (Average run length; under no-change regime)

Assume Δ is bounded. For any $\epsilon > 0$ and threshold $b > 2$, define $h(\epsilon, \Delta) = \min\{\frac{\epsilon}{2\Delta}, 1\}$. Then we have

$$\mathbb{E}_\infty[\mathcal{T}(b)] \geq \frac{e^{h(\epsilon,\Delta)b-2}}{4(b+1)^2}.$$

Theoretical guarantees

Theorem (Average run length; under no-change regime)

Assume Δ is bounded. For any $\epsilon > 0$ and threshold $b > 2$, define $h(\epsilon, \Delta) = \min\{\frac{\epsilon}{2\Delta}, 1\}$. Then we have

$$\mathbb{E}_\infty[\mathcal{T}(b)] \geq \frac{e^{h(\epsilon,\Delta)b-2}}{4(b+1)^2}.$$

Theorem (Detection delay; post-change regime)

Assume Δ is bounded. We have for any $b > 0$,

$$\mathbb{E}_0[\mathcal{T}(b)] \leq \frac{b}{I_0} + \frac{4\Delta}{I_0^{3/2}\epsilon}\sqrt{b} + C,$$

where I_0 is the KL divergence of the post- and pre-change distributions, and C is a constant that only depends on Δ, ϵ, I_0.

Choosing the threshold

- Target false-alarm requirement:

$$\mathbb{E}_\infty[\mathcal{T}] \geq \gamma.$$

- The ARL bound gives an analytical calibration rule:

$$\frac{e^{h(\epsilon,\Delta)b-2}}{4(b+1)^2} \geq \gamma, \qquad h(\epsilon,\Delta) = \min\left\{\frac{\epsilon}{2\Delta}, 1\right\}.$$

- Asymptotically,

$$b_\gamma = \frac{\log\gamma}{h(\epsilon,\Delta)}(1+o(1)).$$

- Stronger privacy means smaller ϵ, smaller h, and therefore a larger threshold.

Privacy-delay tradeoff

- Relationship between ARL γ and Detection delay:

$$\mathbb{E}_0[\mathcal{T}(b_\gamma)] \leq \frac{\log \gamma}{h(\epsilon, \Delta) I_0}(1 + o(1)), \quad h(\epsilon, \Delta) = \min\{\frac{\epsilon}{2\Delta}, 1\}.$$

Privacy-delay tradeoff

- Relationship between ARL γ and Detection delay:

$$\mathbb{E}_0[\mathcal{T}(b_\gamma)] \leq \frac{\log \gamma}{h(\epsilon, \Delta) I_0}(1 + o(1)), \quad h(\epsilon, \Delta) = \min\{\frac{\epsilon}{2\Delta}, 1\}.$$

- From CUSUM method, we know the **optimal** delay vs. ARL γ is

$$\inf_{\mathcal{T}: \mathbb{E}_\infty[\mathcal{T}] \geq \gamma} \mathbb{E}_0[\mathcal{T}] = \frac{\log \gamma}{I_0}(1 + o(1)).$$

Privacy-delay tradeoff

- Relationship between ARL γ and Detection delay:

$$\mathbb{E}_0[\mathcal{T}(b_\gamma)] \leq \frac{\log \gamma}{h(\epsilon, \Delta) I_0}(1 + o(1)), \quad h(\epsilon, \Delta) = \min\{\frac{\epsilon}{2\Delta}, 1\}.$$

- From CUSUM method, we know the **optimal** delay vs. ARL γ is

$$\inf_{\mathcal{T}: \mathbb{E}_\infty[\mathcal{T}] \geq \gamma} \mathbb{E}_0[\mathcal{T}] = \frac{\log \gamma}{I_0}(1 + o(1)).$$

- The DP-CUSUM procedure is still optimal when $\epsilon \geq 2\Delta$ (weaker privacy);
- The delay of DP-CUSUM may increase as ϵ decreases (stronger privacy).

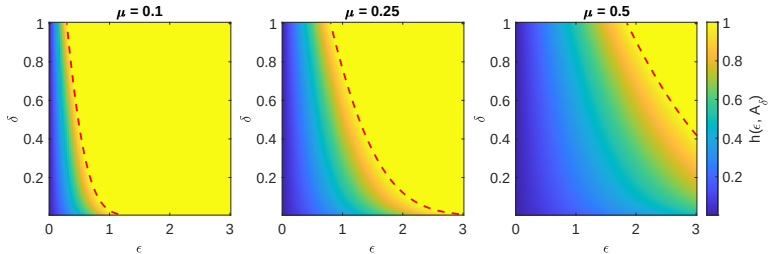
Outline

- Motivation: Private Online Change Detection
- Proposed: Differentially-Private CUSUM Test (DP-CUSUM)
- Theoretical Guarantees: False Alarm and Detection Delay
- Extensions and Numerical Results

Extension to unbounded sensitivity

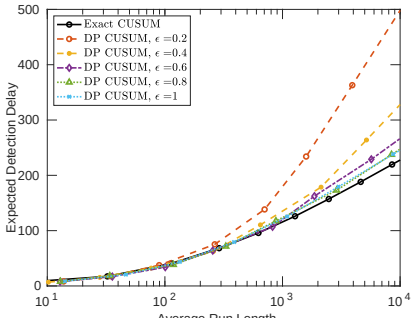
For unbounded log-likelihood ratio: $\Delta = \sup_{x,y} |\ell(x) - \ell(y)| = \infty$.

- Pre-specified a tolerance parameter $\delta \in (0, 1)$.
- Set $A_\delta = \inf\{t : \max_{i=0,1} \mathbb{P}_{X \sim f_i} (2|\ell(X)| \geq t) \leq \delta/2\}$.
- Use A_δ instead of Δ when adding Laplace noise Lap $\left(\frac{2A_\delta}{\epsilon}\right)$.
- This yields a relaxed (ϵ, δ)-DP guarantee.

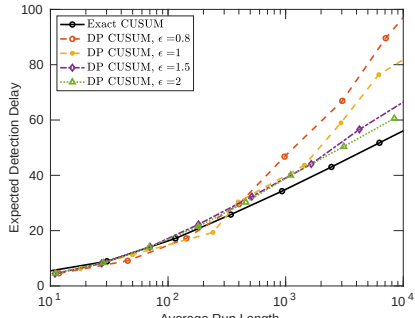


Heatmaps of the effective privacy factor $h(\epsilon, A_\delta)$ for $N(0, 1) \to N(\mu, 1)$.
Dashed red line: $h(\epsilon, A_\delta) = 1$.

Simulation results: bounded LLR



(a) $\mathsf{Lap}(0,1) \to \mathsf{Lap}(0.2,1)$

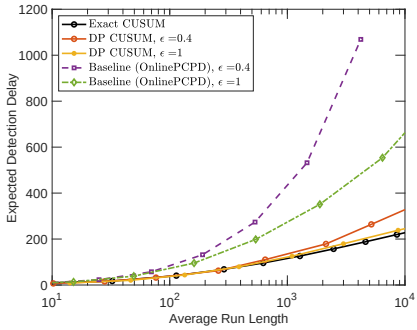


(b) $\mathsf{Lap}(0,1) \to \mathsf{Lap}(0.5,1)$

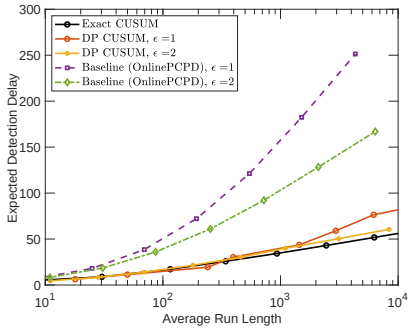
Figure: Average detection delay versus average run length under Laplace distributions. Results are averaged over 10,000 trials.

Comparison with OnlinePCPD: bounded LLR

- OnlinePCPD uses a sliding window with $w = 700$ and recomputes windowed likelihood statistics: $O(w)$ per time step.
- It adds larger noise: $Z_t \sim \text{Lap}(8\Delta/\epsilon)$ and $W \sim \text{Lap}(4\Delta/\epsilon)$.
- DP-CUSUM uses $\text{Lap}(2\Delta/\epsilon)$ noise and keeps the recursive CUSUM update.



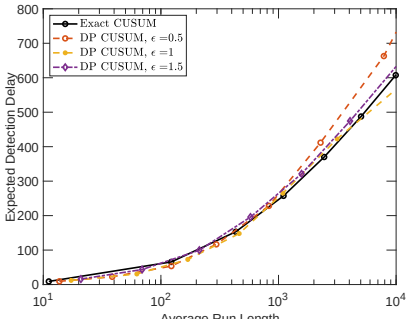
(a) $\text{Lap}(0, 1) \to \text{Lap}(0.2, 1)$



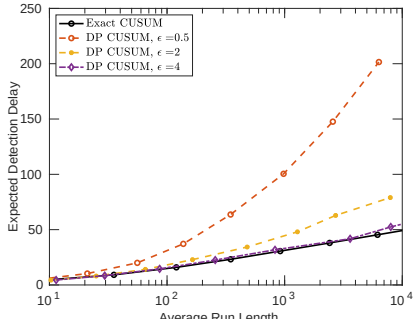
(b) $\text{Lap}(0, 1) \to \text{Lap}(0.5, 1)$

R. Cummings, et al. "Differentially private change-point detection," NeurIPS, 2018.

Simulation results: unbounded LLR



(a) $N(0,1) \to N(0.1,1)$

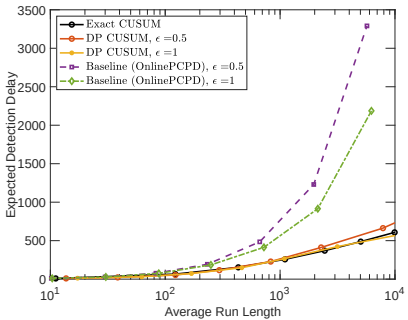


(b) $N(0,1) \to N(0.5,1)$

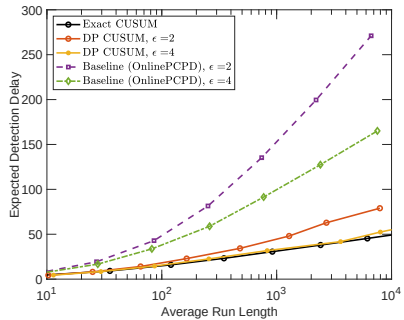
Figure: Average detection delay versus average run length under Normal distributions. Results are averaged over 10,000 trials.

Comparison with OnlinePCPD: unbounded LLR

- Both procedures use the relaxed sensitivity A_δ with $\delta = 0.1$ and the same privacy parameters.
- DP-CUSUM has smaller delay than OnlinePCPD, especially at larger ARL.



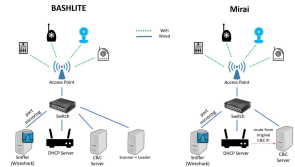
(a) $N(0,1) \to N(0.1,1)$

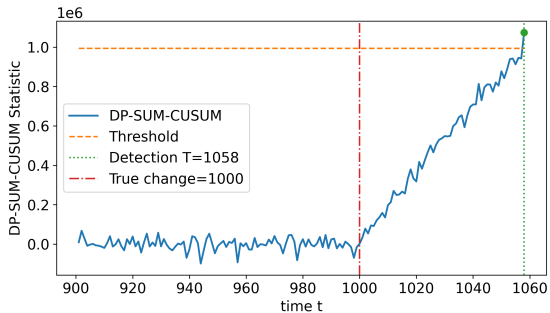


(b) $N(0,1) \to N(0.5,1)$

Real-data example

- *N-BaIoT* dataset (Internet of Things)
  - 9 nodes: thermostat, webcam, cameras, etc.
  - Each node: 115-dimensional data stream of network traffic statistics
  - From benign traffic to attacks





Y. Meidan, et al. "N-baiot—network-based detection of IoT botnet attacks using deep autoencoders," IEEE Pervasive Computing, 2018.

Takeaways

- First set of theoretical results for online DP change detection:
  - explicit ARL and WADD bounds,
  - asymptotic optimality under weak privacy.
- Simple mechanism:
  - recursive, computationally efficient,
  - minimum modification on CUSUM test,
  - smaller noise added compared to baseline.
- Open question and future directions:
  - Is DP-CUSUM optimal under stronger privacy $\epsilon < 2\Delta$?
  - Extensions to other data structures or settings.
  - Extensions to other privacy measures.

Liyan Xie and Ruizhi Zhang (2026) "Sequential Change Detection with Differential Privacy." IEEE Transactions on Information Theory.