A Multi-Armed Bandit Approach for Online Monitoring High-Dimensional Data in Resource Constrained Environments

Wanrong Zhang, Yajun Mei

H. Milton Stewart School of Industrial and Systems Engineering Georgia Institute of Technology

Overview

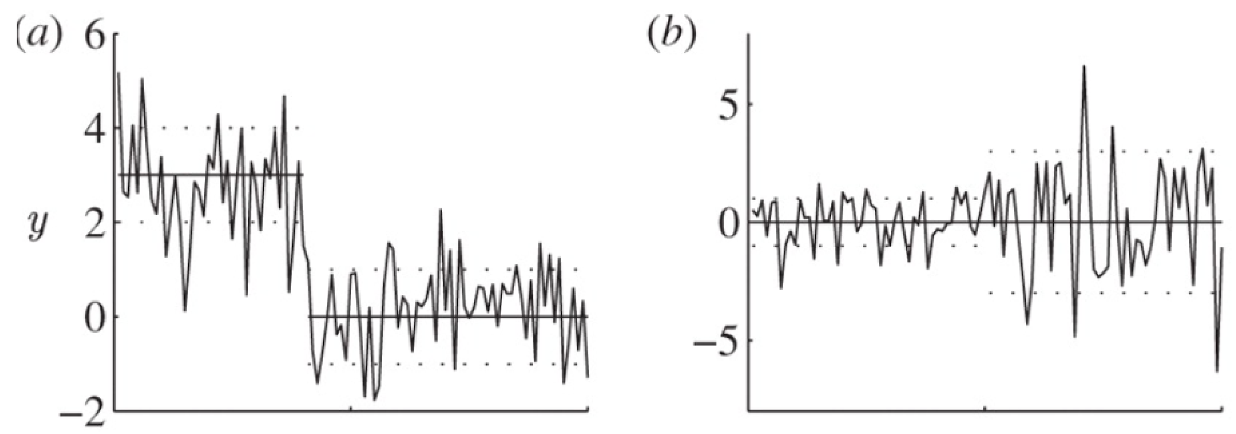
This paper investigates the problem of online monitoring high-dimensional streaming data in resource constrained environments, where one has limited capacity in data acquisi- tion, transmission or processing, and thus can only observe or utilize partial, not full, data for decision making.

- We propose a multi-armed bandit approach to adaptively sampling useful local compo- nents of data, and our method, termed Thompson-Sampling-Shiryaev-Roberts-Pollak (TSSRP) algorithm, is to combine the Thompson Sampling algorithm in the multi-armed bandits problem with the Shiryaev-Roberts-Pollak procedure in the sequential change-point detection literature.
- Our proposed TSSRP algorithm is able to balance between exploiting those ob- served local components that maximize the immediate detection performance and exploring new local components that might accumulate new information to improve future detection performance.

Problem Statement and Background

Problem Formulation.

- K-dimensional data $\mathbf{X}_t = (X_{1,t}, \cdots, X_{K,t})$,
- q out of K ($q < K$) local components of $\mathbf{X}_t$ can be observed
- Before a change time ν, $\mathbf{X}_t$ i.i.d. follows $f(\theta_0, x)$; after ν, $\mathbf{X}_t$ i.i.d. follows $f(\theta_1, x)$
- Assume $\theta_1 - \theta_0$ is sparse
- We are repeatedly test $H_0 : \nu = \infty$ vs $H_1 : \nu = 1, 2, \ldots$



Thompson Sampling for Multi-Armed Bandit

Algorithm 1 Thompson Sampling

Input: A prior for the reward distributions, $\mathcal{D} = \emptyset$;

for $t = 1, \ldots, N$ **do**

  For each arm $i = 1, \ldots, K$, sample $\theta_{i,t}$ from the $\mathcal{P}(\theta_i|\mathcal{D})$ distribution.

  Play arm $a(t) := \text{argmax}_i\, \theta_i(t)$ and observe reward $r(t)$.

  append $(r(t), a(t))$ to $\mathcal{D}$ and update posterior distribution $\mathcal{P}(\theta_i|\mathcal{D})$.

end for

Proposed Method

At the high level, we propose to follow the Thompson Sampling algorithm that samples those local components or local data streams that have the largest (randomized) posterior distributions of local changes having occurred, and then take a limiting Bayes approach as in Shiryaev-Robert-Pollak procedure to develop efficient algorithms for adaptively sampling and global online monitoring.

Algorithm 2 Thompson-Sampling-Shiryaev-Roberts-Pollak (TSSRP) algorithm

Parameters: interested smallest magnitude of mean shift u_{min}, the number of local statistics to be sumed r, the number of sensors q, a prior distribution G and stopping threshold d.

Input: m data streams

Algorithm: Set $R_{k,0} = 0$, $h_{k,t} = 1$ for all $k = 1, 2, \ldots, m$. Random sample q variables as the initial layout S

In each round $t \leftarrow 1, 2, \ldots$ do the following until reaching the stopping condition:

(1) Calculate the local statistics based on the current sensor layout S. If $k \in S$, update $R_{k,t}$ based on equation (1) and update $h_{k,t}$ by multiply $\frac{f_1(X_t)}{f_0(X_t)}$. Otherwise, update $R_{k,t}$ based on equation (2) and $h_{k,t}$ stays the same as $h_{k,t-1}$.

(2) For each data stream, sample $R'_{k,0}$ from the distribution G.

(3) Calculate the randomized of local statistic $R'_{k,t} = R_{k,t} + R'_{k,0} * h_{k,t}$.

(4) Order the local statistics $R'_{k,t}$ ($k = 1, 2, \ldots, m$) from the largest to the smallest, and let $l_{(k),t}$ denote the variable index of the order statistics $R'_{(k),t}$.

(5) If the local statistics reach the stopping rule defined in (3), break the loop and raise an alarm. Otherwise, update sensor layout $= \{l_{(1),t}, \ldots, l_{(q),t}\}$ and proceed to the next iteration.

There are three steps in our proposed adaptive sampling strategy.

- **Local Statistics**
  When observable
  $$R_{k,t} = \frac{f_1(X_{k,t})}{f_0(X_{k,t})}(R_{k,t-1} + 1) \tag{1}$$
  when unobservable
  $$R_{k,t} = R_{k,t-1} + 1 \tag{2}$$
- **Adaptive Sensors Allocation** Randomize the local statistics. Note that the randomized value can be computed recursively by random sample the initial value.
- **Global Stopping Time** When the sum of top-r original local statistics exceeds the threshold
  $$T = \inf\{t \geq 1 : \sum_{k=1}^{r} R_{(k),t} \geq d\} \tag{3}$$

Theoretical Properties

Proposition 1 Under the case that $\nu = \infty$, for $\forall k \in [n]$, $\forall t > 0$, there $\exists t' > t$ such that $\mathbb{P}(k \in S_t) > 0$.

Proposition 2 Under the case that $\nu < \infty$, for $\forall k \in C$, we have $\mathbb{P}(k \in S_t, \forall t > t_0 | k \in Q_{t_0}) > 0$ for any $t_0 > \nu$.

Theorem [Average Run Length] Define the stopping time as $T_A = \inf\{t : \sum_{k=1}^{r} R_{(k),t} \geq A\}$, then $\mathbb{E}_\infty(T_A) \geq \frac{A}{r}$.

Simulation

Independent Multivariate Gaussian variables

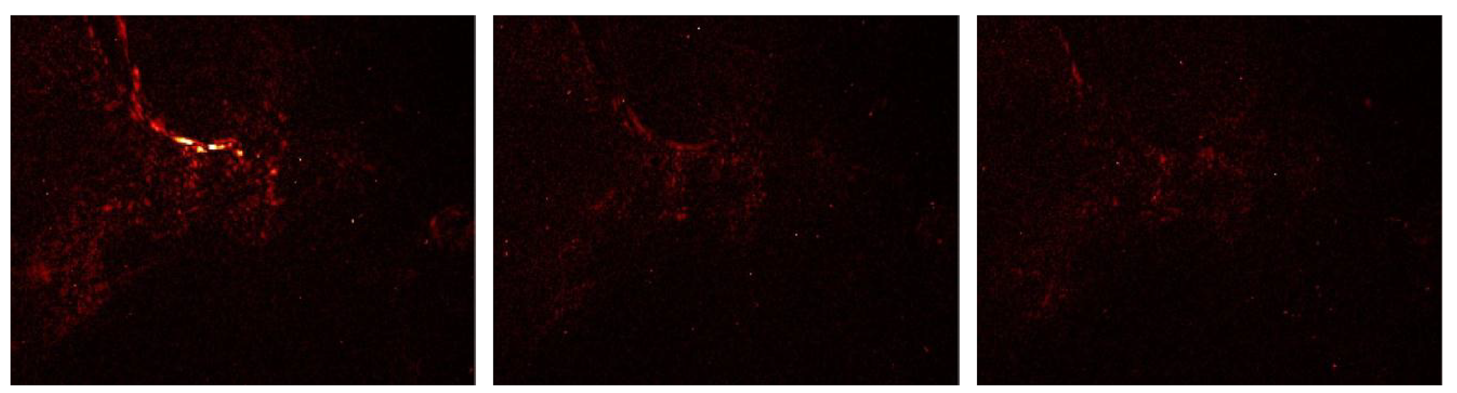
Setting: $K = 100$, $q = 10$, $r = 3$, pre-change $N(0, 1)$, post-change $N(2, 1)$, $\mathbb{E}_\infty T_A = 1000$

(All the results are obtained from 10000 replications.)

The number of shifts	1	2	3	4	5
TSSRP(U[0,1])	12.98(0.195)	9.715(0.122)	8.443(0.101)	7.724(0.084)	7.182(0.076)
TSSRP(deterministic)	12.975(0.198)	9.669(0.118)	8.534(0.102)	7.672(0.083)	7.150(0.075)
TRAS($\delta = 0.03$)	24.442(0.431)	17.15(0.271)	14.046(0.228)	12.126(0.185)	10.437(0.155)
TRAS($\delta = 0.05$)	23.85(0.383)	17.28(0.255)	14.007(0.193)	12.389(0.174)	10.932(0.145)
TRAS($\delta = 0.1$)	29.078(0.310)	21.780(0.216)	18.602(0.169)	16.452(0.148)	14.782(0.136)
The number of shifts	6	7	8	9	10
Proposed1	6.615(0.063)	6.428(0.058)	6.151(0.054)	5.991(0.051)	5.898(0.048)
NonrandomSR	6.624(0.063)	6.312(0.059)	6.144(0.055)	5.875(0.050)	5.871(0.046)
TRAS($\delta = 0.03$)	9.607(0.140)	8.945(0.127)	8.296(0.105)	7.972(0.097)	7.361(0.079)
TRAS($\delta = 0.05$)	13.483(0.223)	9.207(0.110)	8.695(0.098)	8.136(0.089)	7.81(0.078)
TRAS($\delta = 0.1$)	13.809(0.122)	12.944(0.110)	12.180(0.100)	11.718(0.095)	11.320(0.092)

Figure 1: Detection delay for 100 data streams case by sum of top 3 local statistics

Real Data



(a) (b) (c)

Figure 2: Snapshot of Solar Flare

We use the solar flare data in [3]. The results comparison is as follows:

- TSSRP: t=192, t=219
- Optimal TRAS: t=190, t=221
- Full observation: t=191, t=217

Selected References

[1] Kaibo Liu, Yajun Mei, and Jianjun Shi. An adaptive sampling strategy for online high-dimensional process monitoring. Technometrics, 57(3):305-319, 2015.

[2] Shipra Agrawal and Navin Goyal. Analysis of thompson sampling for the multi-armed bandit problem. In Conference on Learning Theory, pages 39-1, 2012.

[3] Yao Xie, Jiaji Huang, and Rebecca Willett. Change-point detection for high-dimensional time series with missing data. IEEE Journal of Selected Topics in Signal Processing, 7(1):12-27, 2013.