We are repeatedly testing Thompson Sampling for multi-armed bandit problems. Before a change time \( k \), the pre-change parameters are \( \theta_0 \), and the post-change parameters are \( \theta_q \). Assume that the change is at \( k \) and \( q < K \). Before the change, we have \( \sum_{i=1}^{k} f_i(\theta_0, x_i) > \sum_{i=k+1}^{K} f_i(\theta_0, x_i) \). After the change, we have \( \sum_{i=1}^{k} f_i(\theta_q, x_i) > \sum_{i=k+1}^{K} f_i(\theta_q, x_i) \). This paper investigates the problem of online monitoring high-dimensional streaming data in resource-constrained environments, where one has limited capacity in data acquisition, 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 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-Roberts-Pollak procedure to develop efficient algorithms for adaptively sampling and global online monitoring.

**Problem Statement and Background**

- **K-dimensional data** \( X_k = (X_{1,k}, \ldots, X_{K,k}) \)
- \( q \) out of \( K \) (\( q < K \)) local components of \( X_k \) can be observed
- Before a change time \( \nu \), we follow \( f(\theta_0, x) \); after \( \nu \), we follow \( f(\theta_q, x) \)
- Assume \( \theta_0 - \theta_q \) is sparse
- We are repeatedly testing \( H_0: \nu = \infty \) vs \( H_1: \nu = 1, 2, \ldots \)

**Proposed Method**

At the high level, we propose to follow the Thompson Sampling algorithm that samples 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-Roberts-Pollak procedure to develop efficient algorithms for adaptively sampling and global online monitoring.

**Algorithm 1 Thompson Sampling**

*Input:* A prior for the reward distributions, \( D = \emptyset \)

For each arm \( i = 1, \ldots, N \) do
- For each arm \( i = 1, \ldots, K \), sample \( \theta_i \) from \( P(\theta_i|D) \) distribution.
- Play arm \( a(t) = \arg \max \{ \theta_i(t) \} \) and observe reward \( r(t) \).
- Append \( r(t), a(t) \) to \( D \) and update posterior distribution \( P(\theta_i|D) \).

end for

**Proposed TSSRP Algorithm**

- Set \( R_0 = 0 \), \( S_0 = 1 \) for all \( k = 1, 2, \ldots, m \). Random sample \( q \) variables as the initial layout \( S \)

In each round \( t = 1, 2, \ldots \), 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 \( R_k \) based on equation (1) and update \( h_k \) by \( h_k = \max h_k \). Otherwise, update \( R_k \) based on equation (2) and keep \( h_k \) stays the same as \( h_k \).
2. For each data stream, sample \( R_k \) from the distribution \( G \).
3. Calculate the randomized local statistic \( S_k = R_k + R_k \cdot h_k \).
4. Order the local statistics \( S_k \) (\( k = 1, 2, \ldots, m \)) from the largest to the smallest, and let \( s_k \) denote the variable index of the order statistics \( S_k \).
5. If the local statistics reach the stopping rule defined in (3), break the loop and raise an alarm. Otherwise, update sensor layout \( = (s_s(1), \ldots, s_s(q)) \) and proceed to the next iteration.

There are three steps in our proposed adaptive sampling strategy:

- **Local Statistics**
  - When observable
    \[
    R_k = \frac{f_i(X_{i,k})}{f_i(X_{i,k})} \cdot R_{k-1} + 1
    \]
  - When unobservable
    \[
    R_k = R_{k-1} + 1
    \]
- **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
    \[
    T = \inf \{ t \geq 1 : \sum_{k=1}^{t} R_{s_s(k)} \geq d \}
    \]

**Theoretical Properties**

**Proposition 1** Under the case that \( \nu = \infty \), for \( \forall \nu \in [n] \), \( \forall \nu > 0 \), there exist \( N \) such that \( P(k \in S_0) > 0 \).

**Proposition 2** Under the case that \( \nu < \infty \), for \( \forall \mu \in C \), we have \( P(k \in S_0, \nu > 0) > \mu(k \in Q_0) \) for any \( \mu \).

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

**Simulation**

**Independent Multivariate Gaussian Variables**

Setting: \( K = 100, q = 10, r = 3 \), pre-change \( N(0,1) \), post-change \( N(2,1) \). \( E(T_A) = 1000 \)

The results are obtained from 10000 replications.

**Real Data**

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**