ACTIVE ANOMALY DETECTION WITH SWITCHING COST

Da Chen[1], Qiwei Huang[1], Hui Feng[1], Qing Zhao[2], Bo Hu[1]

[1]Research Center of Smart Networks and Systems, School of Information Science and Engineering, Fudan University, Shanghai 200433, China

[2]School of Electrical and Computer Engineering, Cornell University

ABSTRACT&INTRODUCTION

The problem of anomaly detection among multiple processes is considered within the framework of sequential design of experiments. The objective is an active inference strategy consisting of a selection rule governing which process to probe at each time, a stopping rule on when to terminate the detection, and a decision rule on the final detection outcome. The performance measure is the Bayes risk that takes into account of not only sample complexity and detection errors, but also costs associated with switching across processes. While the problem is a partially observable Markov decision process to which optimal solutions are generally intractable, a low-complexity deterministic policy is shown to be asymptotically optimal and offer significant performance improvement over existing methods in the finite regime.

Incorporating switching cost into Bayes risk is motivated by a number of applications. For example, in many robotics applications, relocating the robot (or other autonomous decision makers such as UAVs) incurs considerable cost in terms of energy or delay. Another example is medical diagnostics, where frequent and fast switching across drugs and medical procedures carries high risk and side effects.

PROBLEM FORMULATION

- Consider the problem of detecting a target among M cells. At each time, only one cell can be probed. Let H_m denote the hypothesis that the target is in cell m.

- Noisy observations from a probed cell:
  - ▹ If the cell contains a target: $y(t) \sim g(y)$
  - ▹ If the cell is empty,: $y(t) \sim f(y)$

- The test statistic is the log-likelihood ratio (LLR) of each cell m denoted as:

$$l_m(n) \triangleq \log \frac{g(y_m(n))}{f(y_m(n))}.$$

  - ▹ If $y_m(t) \sim g(y)$: $\mathbf{E}_m(l_m(n)) = D(g||f) > 0$
  - ▹ If $y_m(t) \sim f(y)$: $\mathbf{E}_m(l_m(n)) = -D(f||g) < 0$

- Sum of the Log Likelihood Ratio (sum LLRs) can be regarded as the score of whether the region has a target:

$$S_m(n)=\sum_{t=1}^{n} l_m(t) 1_m(t)=\sum_{t=1}^{n}\log\frac{g(y_m(t))}{f(y_m(t))}1_m(t)=S_m(n-1)+\log\frac{g(y_m(n))}{f(y_m(n))}1_m(n)$$

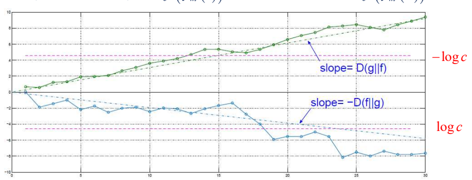

Fig.1 The change of sum of the log likelihood ratio

- Find an anomaly detection policy $\Gamma = (\tau, \delta, \phi)$
  - ▹ Selection Rule: $\phi(n) \in \{1,2,\dots,M\}$
  - ▹ Stopping Rule: τ
  - ▹ Decision Rule: $\delta \in \{1,2,\dots,M\}$

- To minimize the Bayes risk: $\inf_{\Gamma} \mathbf{R}(\Gamma)$

$$\mathbf{R}(\Gamma) = P_e(\Gamma) + c\mathbf{E}(\tau \mid \Gamma) + s\mathbf{E}(\tau_s \mid \Gamma)$$

Error probability | Observation cost | Switching cost

ALGORITHM

➢ THE DBS POLICY

The DBS policy partition the problem space into two regions:

$$\text{Case I: } D(g \,||\, f) + \Delta \geq \frac{D(f \,||\, g)}{M-1},$$
$$\text{Case II: } D(g \,||\, f) + \Delta < \frac{D(f \,||\, g)}{M-1}.$$

where

$$\Delta \triangleq \frac{s(M-2)D(g \,||\, f)D(f \,||\, g)}{-c(M-1)\log c}.$$

which is the offset caused by the switching cost, Δ only affects the performance of the DBS policy in the finite regime.

▹ **In Case I**: probes the cell most likely to be the target.

$$\phi(n) = m^1(n)$$
$$\tau = \inf\left\{n : S_{m^1(n)}(n) > -\log c\right\}$$
$$\delta = H_{m^1(\tau)}$$

where $m^1(n) = \arg\max_m S_m(n)$ is the index of the cell with the largest sum LLRs.

▹ **In Case II** : probes the cell most likely to be empty and eliminate them one by one.

$$\phi(n) = \tilde{m}^{-1}(n)$$
$$\tau = \inf\left\{n : |\mathcal{B}(n)| = M-1\right\}$$
$$\delta = \mathcal{M} \setminus \mathcal{B}(\tau)$$

where $\mathcal{B}(\tau)$ is the set of cells that can be declared as empty at time n, and $\tilde{m}^{-1}(n) = \arg\min_{m \notin \mathcal{B}(n)} S_m(n)$ is the index of the cell with the smallest sum LLRs among all cells that have not been declared.

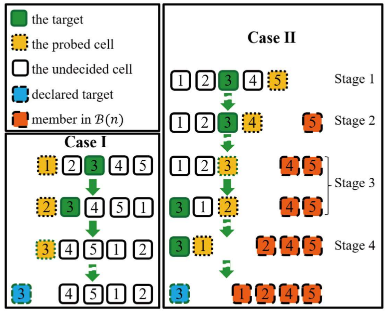

Fig.2 Illustration of the DBS policy (M=5)

➢ ASYMPTOTIC OPTIMALITY OF THE DBS POLICY

Theorem 1:Let R^* and $R(\Gamma)$ denote the Bayes risks in the DBS policy and any other policy Γ respectively. Under the assumption that s=O(c), we have:

$$R^* \sim \frac{-c\log c}{I^*(M,L)} \sim \inf_{\Gamma} R(\Gamma) \text{ as } c \to 0$$

where

$$I^*(M,L) \triangleq \begin{cases} \frac{D(g \,||\, f)}{L}, & \text{if Case I,} \\ \frac{D(f \,||\, g)}{(M-L)}, & \text{if Case II.} \end{cases}$$

the notion $f \sim g$ as $c \to 0$ refers to $\lim_{c\to 0} f/g = 1$

NUMERICAL RESULTS

➢ Simulation 1: The DBS policy in M=5, s=2c. The observation follow Poisson distribution $f \sim Pois(\lambda_f)$, $y \sim Pois(\lambda_y)$, where $\lambda_f = 1, \lambda_g = 0.1$

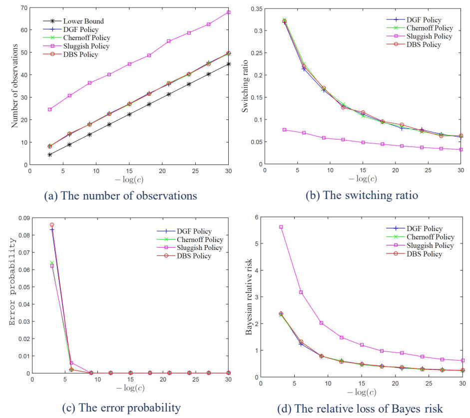

Fig.3 The DBS policy in M=5, s=2c, $\lambda_f = 1, \lambda_g = 0.1$

We can obtain that $D(g||f) \approx 1.4$, $D(f||g)/(M-1) \approx 0.35$, the DBS policy is in Case I for all value of c. The DBS policy, the Chernoff policy and the DGF policy perform similarly because they all observe the cell with the largest sum LLRs.

➢Simulation 2: The DBS policy in M=5, s=2c. The observation follow Poisson distribution $f \sim Pois(\lambda_f)$, $y \sim Pois(\lambda_y)$, where $\lambda_f = 1.5, \lambda_g = 0.001$

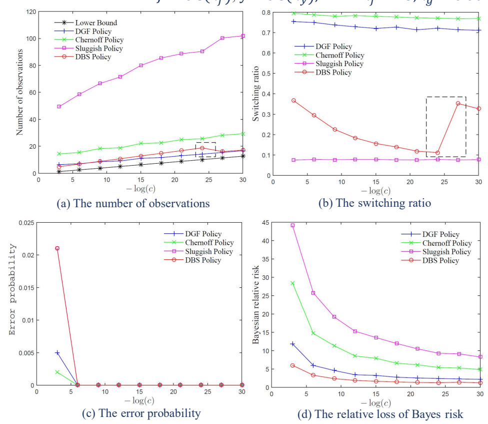

Fig.4 The DBS policy in M=5, s=2c, $\lambda_f = 1.5, \lambda_g = 0.001$

We can obtain that $D(g||f) \approx 1.49$, $D(f||g)/(M-1) \approx 2.37$, the dashed rectangular area of Fig.4(a) and Fig.4(b) showed that the observation delay and the switching ratio change suddenly when $-\log c = 24$. The reason is that DBS policy is in Case I when $-\log c < 24.2$. Although the policy will result in more observations, which can reduce the number of switching. The DBS policy will change to Case II when $-\log c > 24.2$. which will increase the number of switching and reduce observation delay. It is showed that the DBS policy is optimal among algorithms.

CONCLUSION

The problem of anomaly detection with switching cost is studied. We propose a low-complexity deterministic test for the above active hypothesis testing problem with switching cost. Referred to as the Deterministic Bounded Switching (DBS) policy, the proposed policy explicitly specifies the probing action at each time with little computation. Specifically, the policy is based on a key criterion that integrates all parameters affecting the Bayes risk: the number M of cells, the switching cost s, the observation cost c, and the rates at which the target cell and normal cells can be identified as given by the Kullback-Liebler (KL) divergences between the corresponding observation distributions. This criterion partitions the problem space into two cases. In one case, the DBS policy probes the cell most likely to be the target. In the other, DBS probes cells that are likely to be normal and eliminates them one at a time to reduce the number of switching. The DBS policy is simple, intuitively appealing, yet enjoys asymptotic optimality and strong performance in the finite regime as demonstrated in the simulation examples. Future directions include extensions to cases with multiple targets and simultaneous probing of multiple cells.

REFERENCE

[1] Herman Chernoff, "Sequential design of experiments," The Annals of Mathematical Statistics, vol. 30, no. 3, pp. 755–770, 1959.

[2] Stuart A Bessler, "Theory and applications of the sequential design of experiments, k-actions and infinitely many experiments. part i. theory," Tech. Rep., Stanford Univ CA Applied Mathematics and Statistics Labs, 1960.

[3] Mohammad Naghshvar, Tara Javidi, et al., "Active sequential hypothesis testing," The Annals of Statistics, vol. 41, no. 6, pp. 2703–2738, 2013.

[4] Sirin Nitinawarat and Venupogal V Veeravalli, "Controlled sensing for sequential multihypothesis testing with controlled markovian observations and nonuniform control cost," Sequential Analysis, vol. 34, no.1, pp. 1–24, 2015.

[5] Kobi Cohen and Qing Zhao, "Active hypothesis testing for anomaly detection," IEEE Trans. Inf. Theory, vol.61, no. 3, pp. 1432–1450, 2015.

[6] Nidhin Koshy Vaidhiyan and Rajesh Sundaresan, "Active search with a cost for switching actions," in Information Theory and Applications Workshop (ITA). IEEE, 2015, pp. 17–24.

[7] Da Chen, Qiwei Huang, Hui Feng, and Bo Hu, "Active anomaly detection with switching cost," https://arxiv.org/abs/1810.11800, 2018.

ACKNOWLEDGEMENT

This work of Fudan team was supported by the National Key Research and Development Program of China (No.213).

The work of Qing Zhao was supported in part by the National Science Foundation under Grant CCF-1815559 and the European Unions Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No 754412 during her visit at Chalmers University, Sweden.