Bandit Algorithms for Data-Driven Resolution/Field-of-View Tradeoffs in Multi-Mode Sensing and Intelligence Collection

Jefferson Huang, PhD

Assistant Professor

Operations Research Department

Naval Postgraduate School

2020 Emerging Techniques Forum

7 December – 10 December

Military Operations Research Society

Plan of the Talk

Bottom Line Up Front (BLUF)

We propose an algorithm that can provide data-driven guidance on when to use a

low-resolution/high-field-of-view sensor,

versus a

high-resolution/low-field-of-view sensor.

Outline:

I. **Motivation:** Cyber Intrusion Model

II. **Motivation:** Routing Imaging Assets

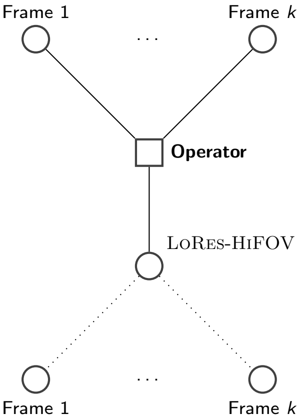
III. **Background:** Multi-Armed Bandit Problems

IV. A Bandit Model of the Resolution/Field-of-View Tradeoff

V. A Data-Driven Algorithm for Dual-Mode Sensing

VI. Efficient Employment of the LoRes-HiFOV Mode

VII. Conclusions & Extensions



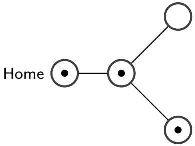
Motivation: Cyber Intrusion Model

D. Kronzilber (NPS Master's Thesis, 2017) proposed a model for the optimal infiltration of, and intelligence collection from, a computer network.

- The intelligence yield from each computer is random, with unknown mean.

Question: Given a set of computers that have been infiltrated, how can the infiltrator simultaneously:

- **Learn** which computers have the highest mean intelligence yield?
- **Maximize** the rate of intelligence collection?



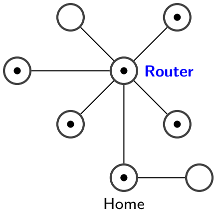
Example: Infiltrated network nodes are denoted by a •

In addition to nodes representing computers, Kronzilber suggested the inclusion of **routers** as special nodes.

- Using a router, the infiltrator can observe **small/partial bits of intelligence** from all computers connected to it.
- The router provides a wider **field-of-view (FOV)**, at the cost of additional noise.

Resolution/FOV Tradeoff

How can the infiltrator effectively employ the router, versus individual infiltrated computers?



Example: Infiltrated network nodes are denoted by a •

Motivation: Routing Imaging Assets

An operator is tasked with collecting imagery intelligence over an **area of interest**.

Collected image frames are fed into an image processing module that assigns a score reflecting the **intelligence value** of the image.

- lower-resolution image $\implies$ noisier intelligence score

Two **imaging modes** are available:

LoRes-HiFOV: Collect from all frames at once, at low resolution

HiRes-LoFOV: Collect from a single frame, at high resolution

Resolution/FOV Tradeoff

How can the operator effectively employ LoRes-HiFOV, versus HiRes-LoFOV?

Background: Multi-Armed Bandit Problems

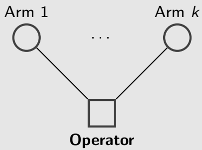
In the old days, a "one-armed **bandit**" referred to a lever-operated slot machine.

Multi-Armed Bandit (MAB)

There are k reward sources, referred to as **arms**.

- Rewards are random.
- The **mean reward** μ_a for each arm a is unknown.

In each **round** $t = 1, \ldots, n$, the operator can collect from exactly one arm.



The operator's **objective** is to minimize their **regret** R_n; letting X_t be the reward earned in round t,

$$R_n = \mathbb{E}\left[\sum_{t=1}^{n}\left(\max_a \mu_a - X_t\right)\right]$$

Exploration/Exploitation Tradeoff

In each round, the operator must balance two considerations:

- **Exploration:** Learn about the mean rewards.
- **Exploitation:** Maximize the reward earned.

A classic approach involves assigning iteratively updated **indices** $\mathcal{I}_a(t)$ to each arm a. They reflect, as of the start of round t,

- the average reward $\bar{X}_a(t)$ earned from arm a and
- the number of times $N_a(t)$ arm a has been collected from

In round t, the operator collects from arm with the **highest index**, i.e.,

$$a^* = \arg\max_{a=1,\ldots,k} \mathcal{I}_a(t)$$

Example: Upper Confidence Bound (UCB) Algorithm

$$\mathcal{I}_a(t) = \bar{X}_a(t) + \sqrt{\frac{2\ln(n)}{N_a(t)}}$$

A Bandit Model of the Resolution/Field-of-View Tradeoff

Introduce an additional arm, called LoRes-HiFOV.

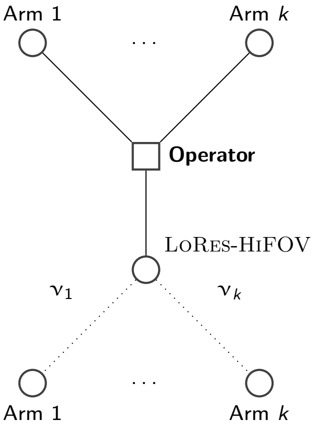
- Collecting from LoRes-HiFOV means that, **for each arm**

$$a = 1, \ldots, k$$

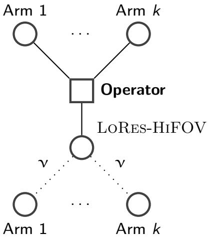
the operator collects from arm a, with probability ν_a.

Assumptions for this Talk

- The ν_a's are identically equal to ν.
- The operator knows ν.



A Data-Driven Algorithm for Dual-Mode Sensing



Notation:

- $\bar{X}_a(t)$ = average reward from arm a up to round t
- $N_a(t)$ = number of collections from arm a prior to round t
- $N_{\mathrm{LH}}(t)$ = number of LoRes-HiFOV collections
- $\mathcal{A}(t)$ = set of **active arms** at the start of round t
- $\mathcal{I}_a(t) = \bar{X}_a(t) + \sqrt{\frac{2\ln(n)}{N_a(t) + \nu N_{\mathrm{LH}}(t)}}$ = **index** of arm a
- $\mathcal{I}_{\mathrm{LH}}(t) = \nu|\mathcal{A}(t)|\sqrt{\frac{2\ln(n)}{\nu N_{\mathrm{LH}}(t)}}$ = **index** of LoRes-HiFOV

Bandit Algorithm for Dual-Mode Sensing

1. **Initialization:**
   - Collect from LoRes-HiFOV until each arm has been collected from at least once.
   - **Eliminate** all arms $a \in \{1, \dots, k\}$ where

$$\mathcal{I}_a(t) < \max_{a=1,\dots,k} \left\{ \bar{X}_a(t) - \sqrt{\frac{2\ln(n)}{N_a(t) + \nu N_{\mathrm{LH}}(t)}} \right\}$$

2. **Main:** For each round $t = 1, \dots, n$,
   - If $\mathcal{I}_{\mathrm{LH}}(t) > \max_{a=1,\dots,k} \mathcal{I}_a(t)$, collect from LoRes-HiFOV.
   - Otherwise, collect from arm

$$a^* = \arg\max_{a=1,\dots,k} \mathcal{I}_a(t)$$

   - **Eliminate** all arms $a \in \{1, \dots, k\}$ where

$$\mathcal{I}_a(t) < \max_{a \in \mathcal{A}(t)} \left\{ \bar{X}_a(t) - \sqrt{\frac{2\ln(n)}{N_a(t) + \nu N_{\mathrm{LH}}(t)}} \right\}$$

   - Update $\mathcal{A}(t)$

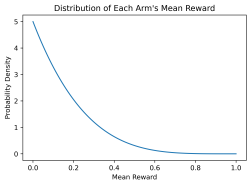
Efficient Employment of the LoRes-HiFOV Mode

The LoRes-HiFOV arm should be especially beneficial when there are many arms, and relatively few are good.

Example (0-1 Rewards)

Each arm yields a reward of 1 (e.g., "Useful Intelligence") or 0.

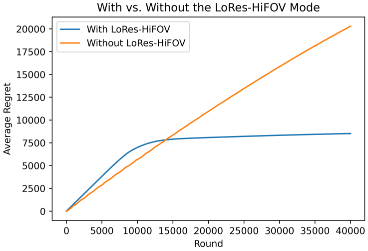
- The **mean rewards** μ_a vary according to a Beta distribution with parameters $\alpha = 1$ and $\beta = 5$:

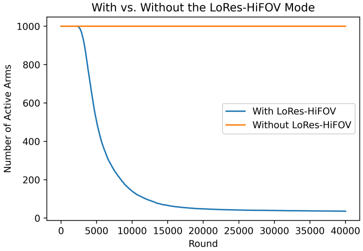


- $k = 1000$ arms.
- $n = 40,000$ rounds.
- $\nu = 0.02$

The average performance of our algorithm over 100 simulation replications, with and without the LoRes-HiFOV mode, is shown on the right.

- Without LoRes-HiFOV = UCB Algorithm with Arm Elimination





With LoRes-HiFOV, the less desirable arms are quickly screened.

Conclusions & Extensions

Conclusions:

- We proposed a model and an index-based algorithm for judiciously trading off resolution versus FOV.
- Empirical results indicate that our **index-based algorithm** can efficiently employ the LoRes-HiFOV mode to quickly screen less desirable frames.

Some Potential Extensions:

- Don't know the ν_a's.
- Different noise models for LoRes-HiFOV
- More than two resolution/FOV options
- Detecting **changes** in mean intelligence values.

Contact Information

Jefferson Huang, PhD

Assistant Professor
Operations Research Department
Naval Postgraduate School

✍ **Web:** http://faculty.nps.edu/jefferson.huang/
✉ **Email:** jefferson.huang@nps.edu