Bandit Models of Cyber Intrusion

Jefferson Huang, PhD

Assistant Professor

jefferson.huang@nps.edu

Operations Research Department

Naval Postgraduate School

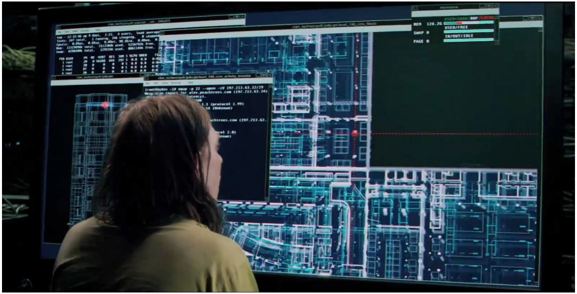
INFORMS Annual Meeting

Seattle, WA

October 20, 2019

Network Intrusion Scenario

Task: Collect useful information from computers on a network.



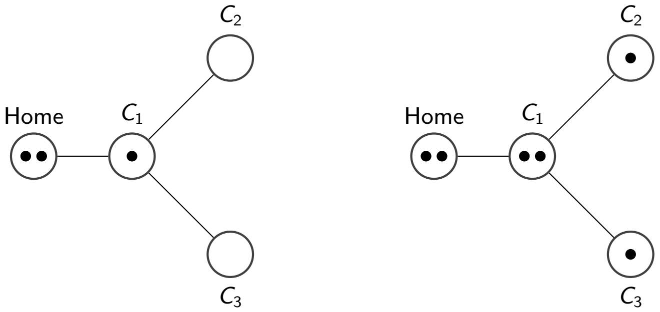
Initially have access to a "home node", from which you can:

- try to collect information from that node, or
- try to infiltrate other computers connected to that node.

Question: How should information collection be balanced with infiltration?

Network Intrusion Scenario

Example:



- 2 dots = **infiltrated** computer (can try to collect info.)
- 1 dot = **accessible** computer (can try to infiltrate)
- No dot = **inaccessible** computer.

Linear Network

The home node is initially at one end of a line of K computers.

Example:

- Trying to collect info from C_k yields a unit of info with probability μ_k.
- Each infiltration attempt is successful with probability s.

Objective: Maximize the average number of info units collected over a discrete and finite number T of **decision epochs**.

Linear Network: "Full" Knowledge

"Full" Knowledge Assumption: The values $\mu_1, \ldots, \mu_K$ and s are known.

- An optimal collect/infiltrate policy can be computed via dynamic programming.

Definition

Let $\mu^*(k) := \max_{i=1,\ldots,k} \mu_i$, and define the operators $\mathcal{T}_C, \mathcal{T}_I$ on functions $f : \{1, \ldots, K\} \to \mathbb{R}$ by

$$\mathcal{T}_C f(k) := \mu^*(k) + f(k)$$

and

$$\mathcal{T}_I f(x) := sf(k+1) + (1-s)f(k).$$

Optimality Equations

Letting $V_0 \equiv 0$, for $t = 1, \ldots, T$ the value function V_t satisfies

$$V_t(k) = \begin{cases} \max\{\mathcal{T}_C V_{t-1}(k), \mathcal{T}_I V_{t-1}(k)\}, & k = 1, \ldots, K-1 \\ \mathcal{T}_C V_{t-1}(K), & k = K. \end{cases}$$

Linear Network: "Full" Knowledge

Definition

A collect/infiltrate policy is a **threshold policy** if

$$\text{collect info at epoch } t \implies \text{collect info at epoch } t+1.$$

Theorem

If $T = 3$, then there is an optimal threshold policy.

Conjecture

For any horizon T, there is an optimal threshold policy.

- To prove the conjecture, it suffices to show that for $k = 1, \ldots, K-1$,

$$V_t(k+1) - V_t(k)$$

is non-decreasing in t.

Linear Network: "Partial" Knowledge

"Partial" Knowledge Assumption: The chance of successful infiltration s is known, but the values $\mu_1, \ldots, \mu_K$ are unknown.

Idea

Consider policies consisting of **two phases**:

1. Devote the first T_I epochs to attempting to infiltrate new computers.
2. During the remaining epochs, collect info from the infiltrated computers using a bandit algorithm (e.g., UCB).

Theorem (Dor Kronzilber (MAJ, IDF), Master's Thesis, NPS, 2017)

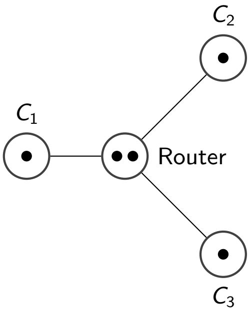
To achieve a regret of

$$O(\sqrt{T \log(T)}),$$

it suffices to let $T_I = O(\sqrt{T/\log(T)})$ *and to use UCB.*

Using Routers Under "Partial" Knowledge

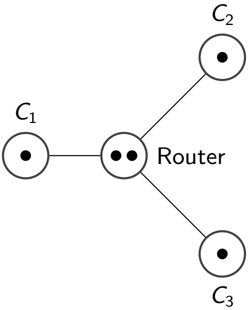
Having access to a **router** node enables you to simultaneously get filtered intelligence (e.g., "snippets") from all computers connected to it.



Question: Suppose you have access to a router that is connected to K infiltrated computers. How should you optimally extract information from those K computers over T decision epochs?

- lots of filtered information **vs.** targeted un-filtered information

Using Routers Under "Partial" Knowledge



Filtered Information: Suppose that whenever the router is used, the following occurs for each connected infiltrated computer C_k:

- with probability η_k, C_k responds as if you had tried to collect info from it;
- with probability $1 - \eta_k$, C_k responds as if you had not tried to collect info from it.

Using Routers Under "Partial" Knowledge

Idea

Consider policies consisting of **two phases**:

1. Use the router during the first T_R decision epochs to select a subset of the connected infiltrated computers.
2. Collect info from the selected subset of computers using a bandit algorithm (e.g., UCB).

Subset selection can be done based on confidence intervals for the μ_k's

- Lykouris, T., E. Tardos, and D. Wali. "Graph regret bounds for Thompson sampling and UCB." arXiv, May 23, 2019.

Theorem

Suppose the number K of connected infiltrated computers is fixed. To achieve a regret of

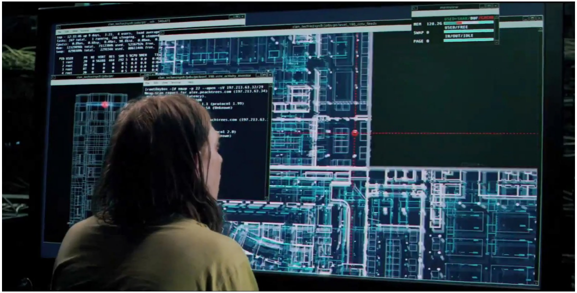
$$O\left(\frac{\log(T)}{\min_k \eta_k} + \sqrt{T\log(T)}\right),$$

it suffices to let $T_R = O(\log(T)/\min_k \eta_k)$ and to use UCB.

Summary and Extensions

Summary:

- Sequential network intrusion model, from attacker's point of view.
- Results for linear network.
- Results on using routers that provide filtered batch observations.



Extensions: network topologies, fatal detections, multiple "players", ...