Cheap Bandits

Manjesh K Hanawal
Dept. of ECE Boston University

Joint work with
Venkatesh Saligrama, (BU,USA)
Michal Valko (INRIA, France)
Remi Munos (DeepMind, UK)

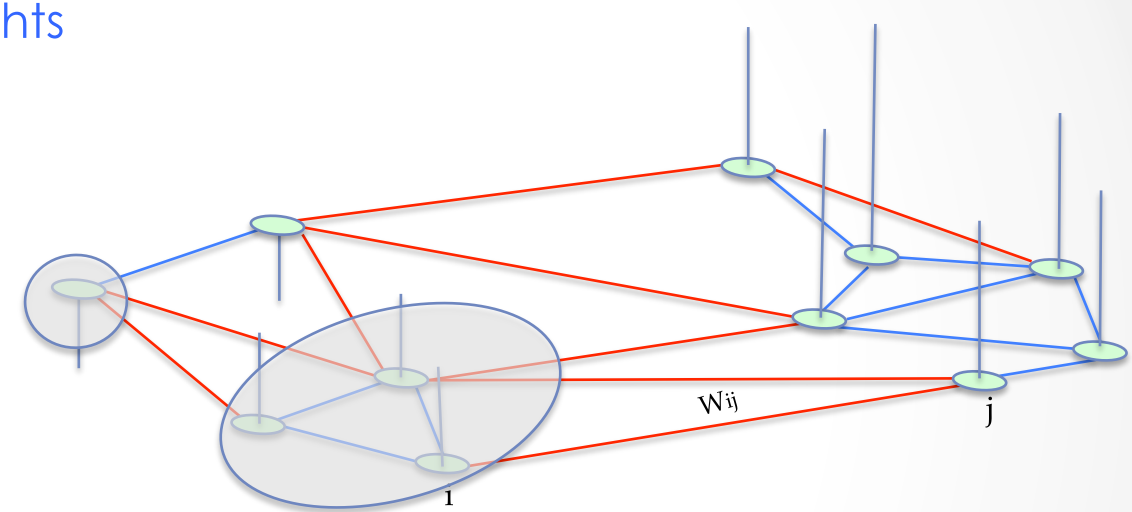
ICML 2015, Lille, France

Problem setting

- Undirected Graph: G=(V,E,W)
  - N Nodes, $W=\{w_{ij}\}$: Weights
- Signal on Graph
  - Reward Function

    $$f : V \longrightarrow \mathbb{R}$$

  - Smooth Function



- Locate maxima

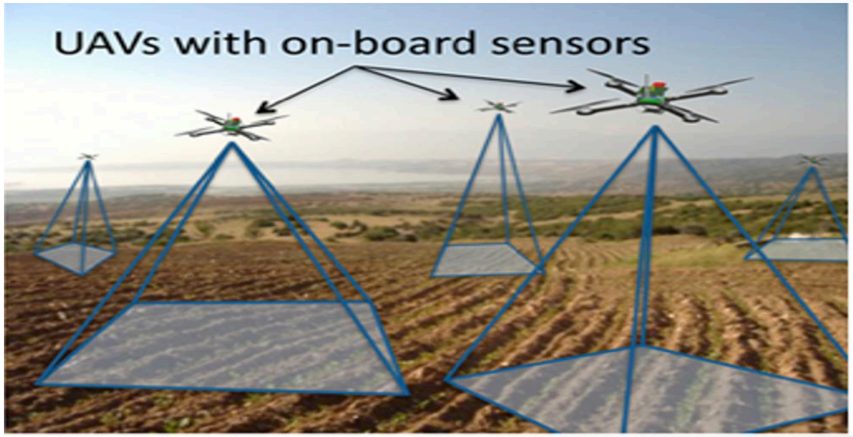
  $$u^* = \arg\max_{u \in V} f(u)$$

- Actions:
  - Noisy Cluster Averages; Differentiated Costs
- Goal: In min Time (T << N) locate u*; Min Cost?

Application Scenarios

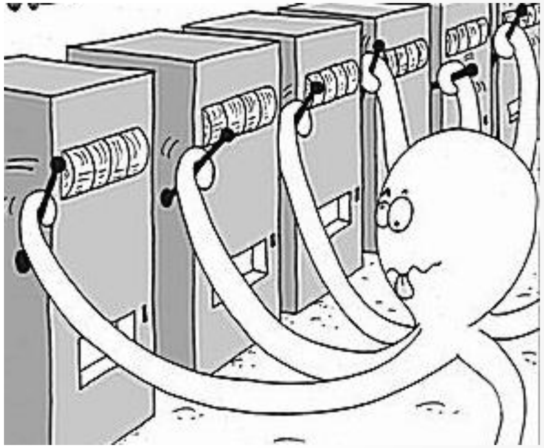
- Surveillance/Geography
  - Forest Cover Dataset: labeled samples on $30m^2$ region
  - **Nodes**: Regions of forest; **Edge weights**: feature similarity;
  - **Rewards**: Density of species. Locate highest density.
  - **Actions**: Zoom-in to a node (high cost); Zoom-out (low cost).

- Sensor networks:
- Radar search:
- Online advertisements:

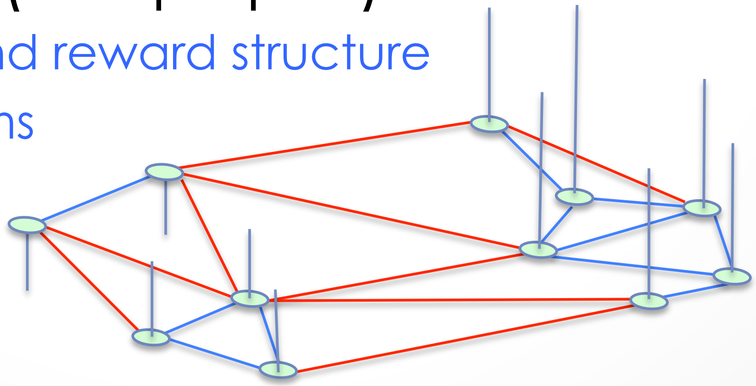


Bandit Setting

- N-arm Bandit [Robbins'72, Lai-Robbins85]
  - N Independent Rewards/arms
    - Each arm ~ action
  - N-nodes ~ no coupling between nodes
  - Need T >> N.
    - Multiple looks per node



- We want T << N (this paper)
  - Exploit graph and reward structure
  - Very large # arms



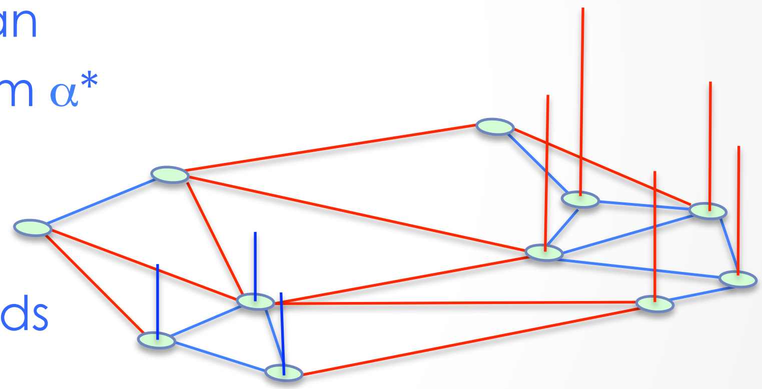
Reward is Linear and Smooth

- Linear Reward
  - Fourier decomposition

$$f = Q\alpha^*$$

  - Q: Eigenvectors of the graph Laplacian
  - Linearly Param Bandit: unknown param α^*

- Smooth Reward
  - Neighboring nodes have similar rewards



$$(u,v) \in E \implies f(u) \approx f(v)$$

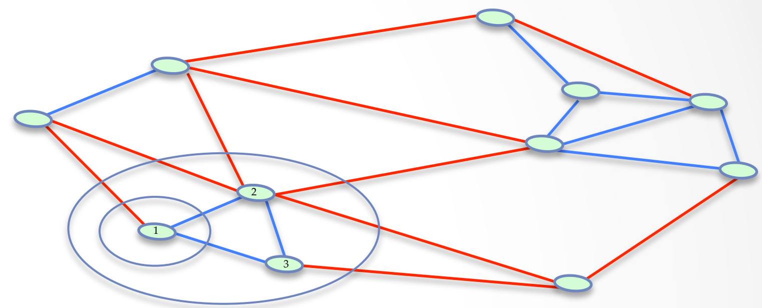
$$\|\mathcal{L}f\|_2^2 = \sum_{u,v} w_{uv}(f(u) - f(v))^2 \leq c$$

[Valko et. al. ICML'14]

Actions: Sample Node or Group

- Actions consists of subset of simplex:
  - Sample a node, u:

$$s(v) = \delta(u - v)$$



$$[1 \;\; 0 \;\; 0 \;\; 0 \;\; \dots \;\; 0]$$

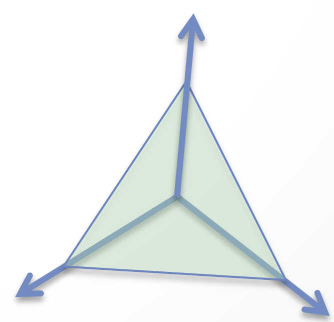
  - Sample a group of nodes $A \subset V$

$$s(v) = \frac{1}{|A|} \sum_{u \in A} \delta(u - v)$$

$$\left[\frac{1}{3} \;\; \frac{1}{3} \;\; \frac{1}{3} \;\; 0 \;\; \dots \;\; 0\right]$$

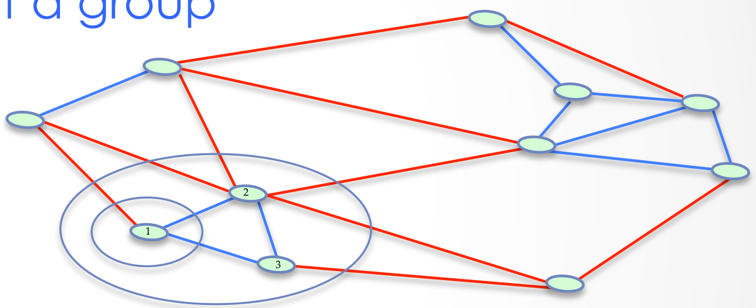
  - General
    - Any Probability Mass Function

$$\mathcal{S} = \Delta^N$$



Cost of Actions

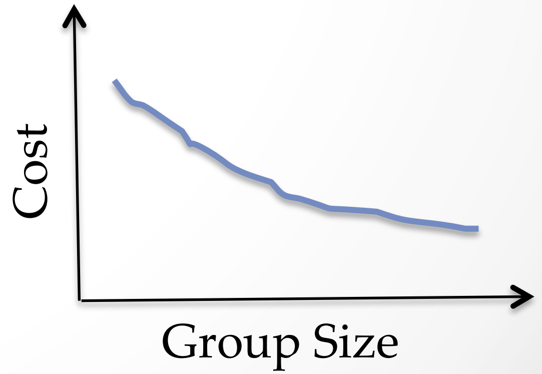
- Cost of actions:
  - Costly: Zoom-in to observe a particular node
  - Cheap: Zoom-out to observe average of a group



- Cost Model

$$C(s) = \sum_{(u,v)\in E} (s(u) - s(v))^2 = \|\mathcal{L}s\|_2^2$$

- Why this model?
  - Larger the group size smaller the cost
  - Probing Nodes has high cost
  - In Fourier domain: Energy of s



Regret and Cost

- Policy(π): In round t, pick an action s_t
- Observe reward

$$r_t(s_t) = \langle s_t, f \rangle + \epsilon_t = \sum_u s_t(u) f(u) + \epsilon_t$$

- Cumulative Regret

$$R_T(\pi) = Tf(u^*) - E\left[\sum_{t=1}^{T} r_t(s_t)\right]$$

- Cumulative Cost

$$C_T(\pi) = \sum_{t=1}^{T} C(s_t)$$

Objective: Cost vs Regret

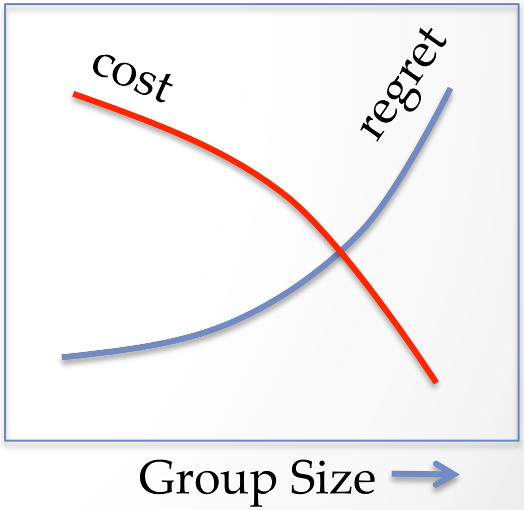
- Minimize Cost subject to 'optimal' Regret

$$\begin{aligned} \min_{\pi, \mathcal{S}} \quad & C_T(\pi) \\ \text{subject to} \quad & R_T(\pi) \leq R_T^* \end{aligned}$$

- Best admissible policies

$$R_T^* = \min_{\pi, S} R_T(\pi)$$

- Conflicting goals:
  - Node actions give better estimates, but costly
  - Group actions give poor estimates, but cheaper



Optimal Regret with lower cost

What is a good Regret Constraint? Lower Bound

- No smoothness constraint (c → ∞)
  - Finite set of actions

$$R_T(\cdot) = \Omega(\sqrt{NT}) \quad \text{(Chu et. al. AISTATS'11)}$$

- Smooth Functions (This paper)

Proposition: For Smooth function on graphs with effective dimension d

$$R_T(\cdot) = \Omega(\sqrt{dT}) \quad \text{where} \quad d \ll N$$

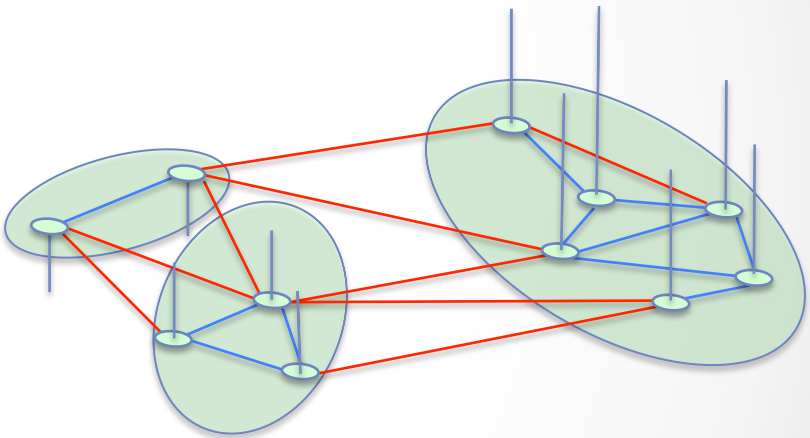
  - Effective Dimension [Valko et.al. ICML'14]

$$d = \max\left\{ i \mid \lambda_i(i-1) \leq \frac{T}{log(T+1)} \right\}$$

Intuition: Lower Bound

- Effective Dimension related to Graph Clusters
  - d clusters
    - # Disconnected clusters or
    - # sparse clusters

$$d = \max\left\{ i \mid \lambda_i(i-1) \leq \frac{T}{log(T+1)} \right\}$$



Need to sample at least one node per cluster

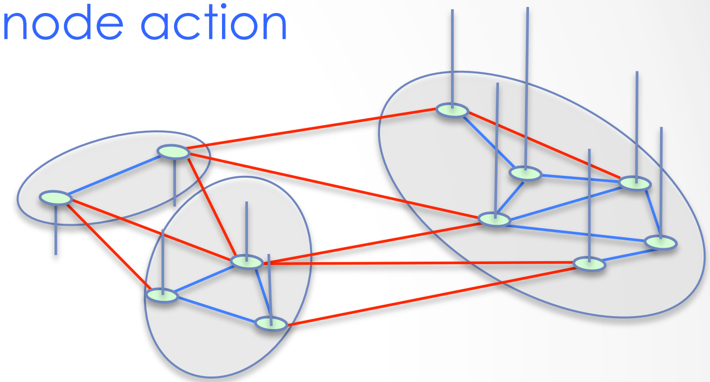
$$\min_{\pi, \mathcal{S}} \quad C_T(\pi)$$

$$\text{subject to} \quad R_T(\pi) \leq \mathcal{O}(\sqrt{dT})$$

Key Intuition: Locally Smooth Rewards

- Smoothness condition implies local smoothness
  - Group actions are good approximation to node action

$$u \in A \implies f(u) \sim \frac{1}{|A|} \sum_{v \in A} f(v) + \text{const}$$

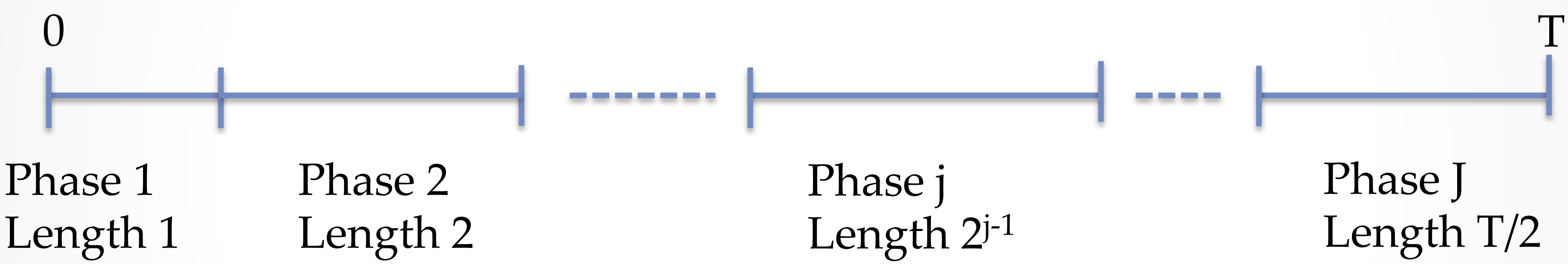


Proposition: Let **f** be a smooth function on a graph with effective dimension d. Then,

$$|\boldsymbol{f}(i) - \frac{1}{\mathcal{N}_i} \sum_{j \in \mathcal{N}_i} \boldsymbol{f}(j)| \leq \frac{c'd}{\lambda_{d+1}}$$

CheapUCB: Algorithm

- Inspired by SpectralUCB Algorithm [Valko et. al. ICML14]
- SpectralUCB uses only node actions, cannot control cost
- CheapUCB uses both node actions and group actions



- **Phases:** Split the T into $J=\lfloor \log T \rfloor$ phases
- **Length:** Phase j=1,2,…J is of 2^{j-1} rounds
- **Select action:** In phase j select groups of size J-j+1 optimistically using UCB

Zoom-in slowly using progressively costly actions

Algorithm Performance

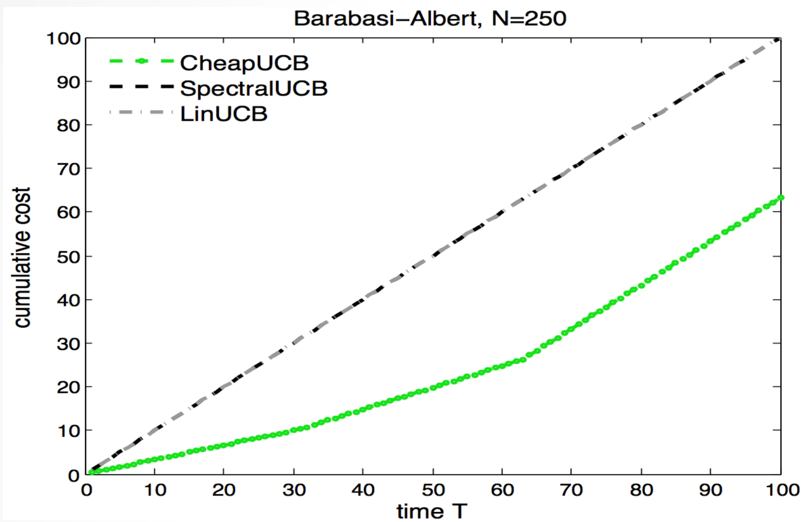
Algorithm	Regret bound	Cost
SpectralUCB (ICML'14)	$\mathcal{O}(d\sqrt{T})$	T
CheapUCB (This paper)	$\mathcal{O}(d\sqrt{T})$	¾ T

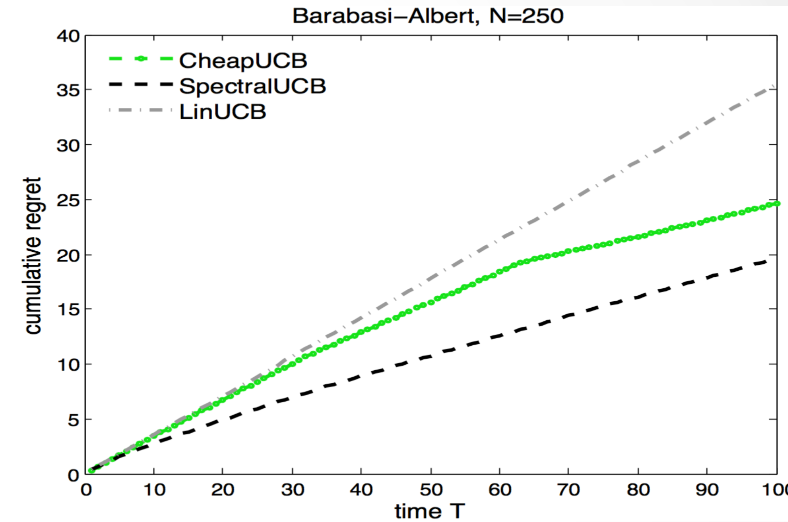
CheapUCB provides good regret guarantee and also provides *O(T)* cost saving

CheapUCB achieves at least 25% reduction in cost!!

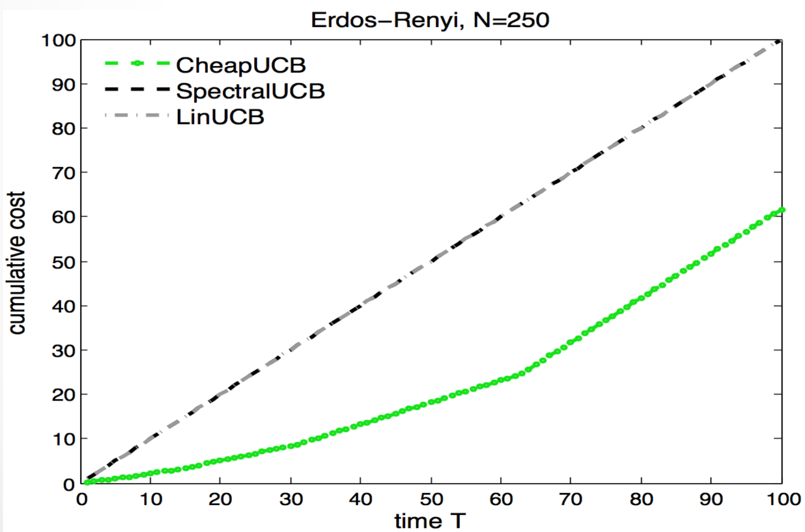
Network Experiments

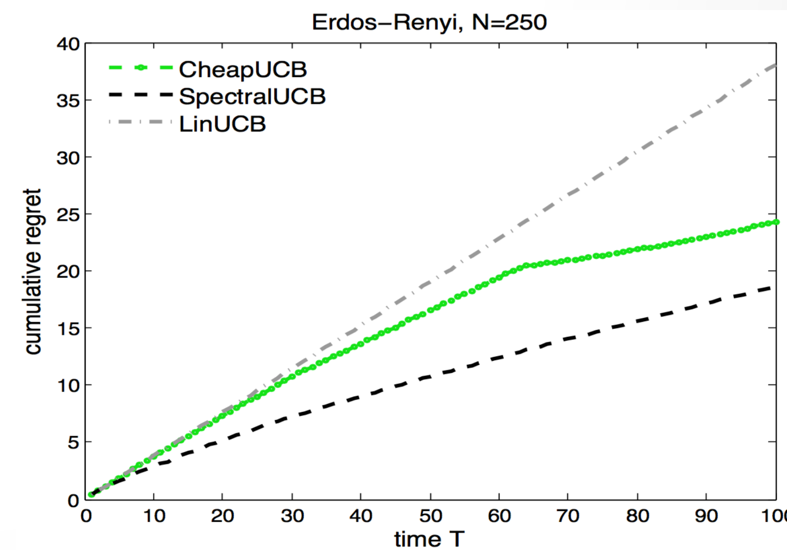
Barabasi-Albert





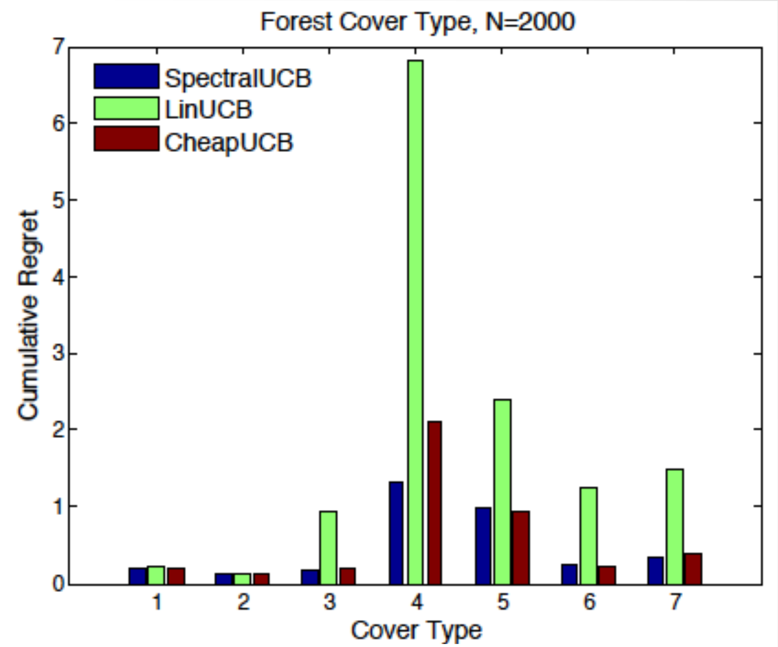
Erdos-Renyi

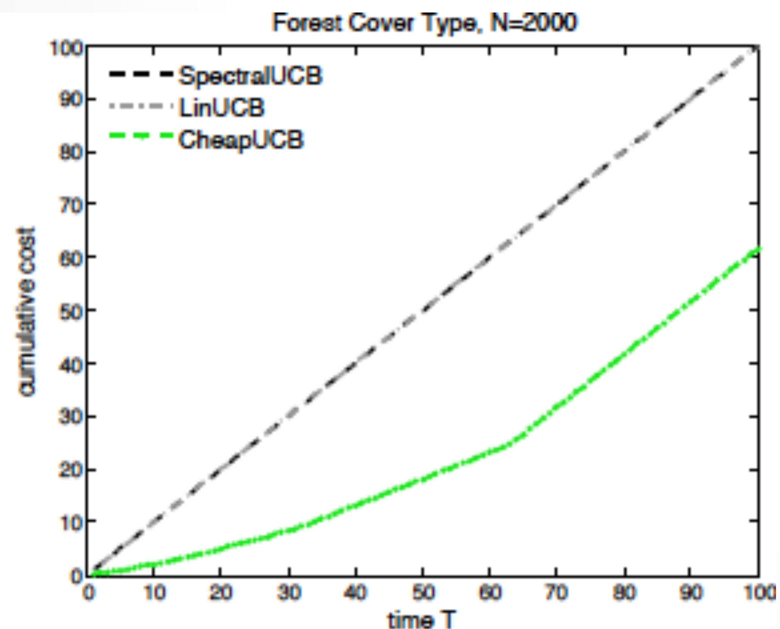


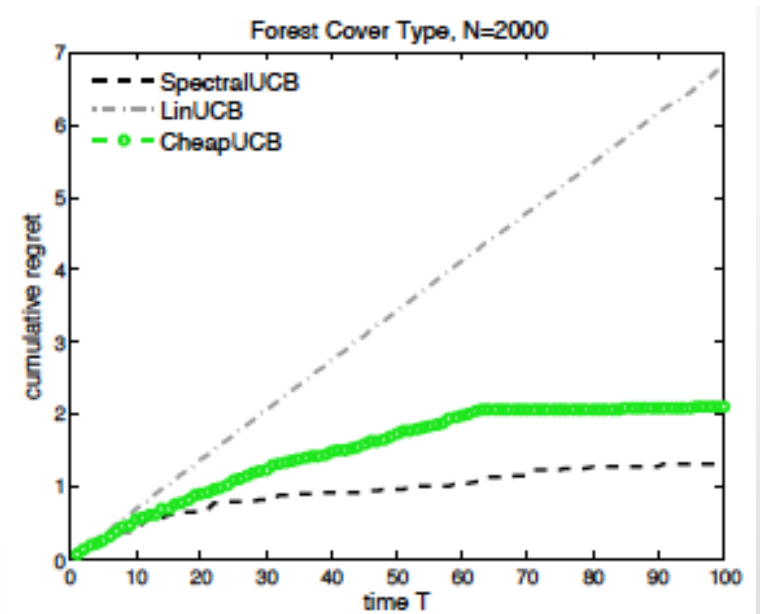


Forest Cover Dataset

- 50000 Samples; 7 Species
- 30m^2 regions; 2000 clusters
- Nodes: regions; Edges: Feature similarity
  - Connect K-NN
- Reward: Density of Desired Species
  - Continuous Classifier Output

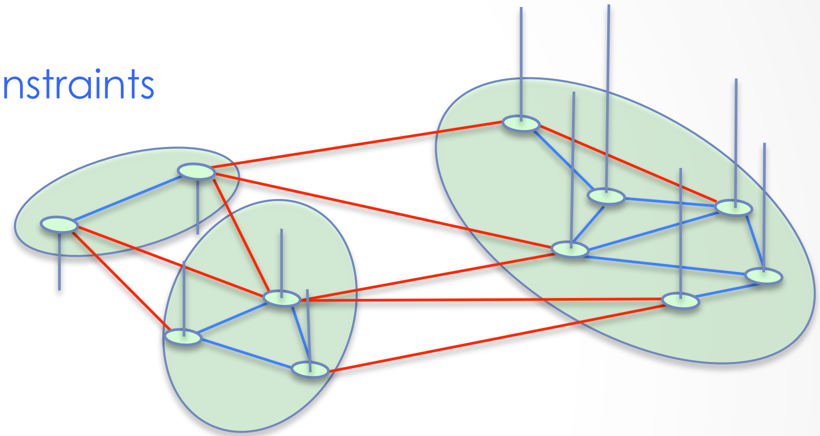






Conclusions

- Cheap Bandit Formulation
  - Optima of Smooth signals on graphs
  - Minimize cost under optimal regret constraints



- Probes/Actions
  - Actions: Sample a node or a group
  - Cost of actions
- Effective Dimension governs regret
  - Time << N, depends on statistical dimension
- Expand actions beyond node actions to reduce cost
  - CheapUCB algorithm
  - Reduces cost by at least by 25%

Thank You!!

Please visit our poster today: 2A