Active Search and Bandit Methods for Complex Actions and Rewards

Yifei Ma

References & collaborators

Jeff Schneider, Roman Garnett, Aarti Singh, Tzu-Kuo Huang.

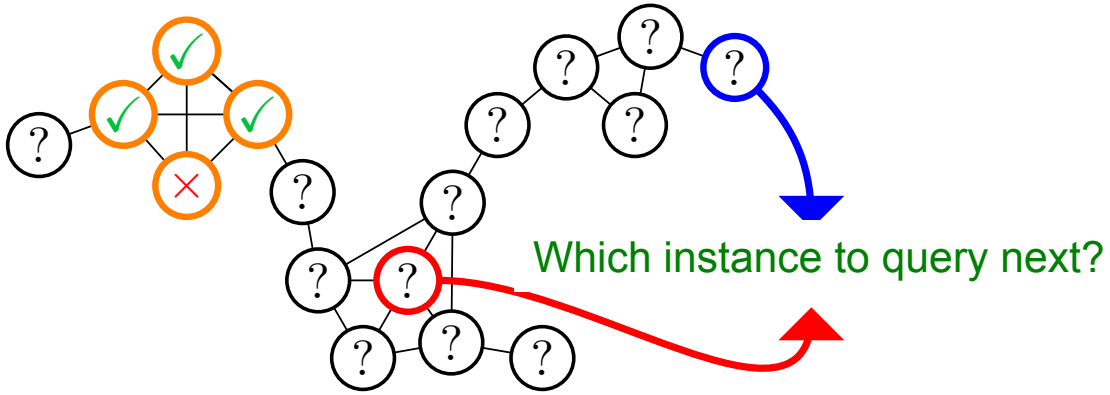
Auton Lab

ML MACHINE LEARNING DEPARTMENT

Carnegie Mellon University

Active Search

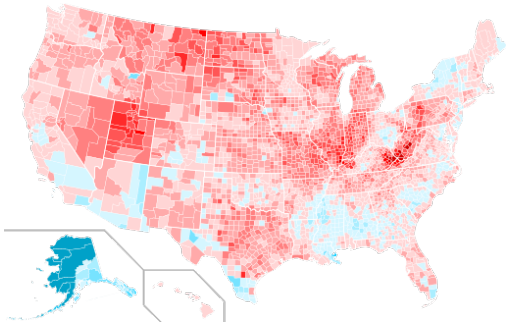
Given a set of instances and features, find all positive instances as quickly as possible, by querying label values from user/environment.



Applications

Pollution detection

Search and rescue



Social polling

Email investigation

Roadmap

Active search	Point rewards	Region rewards
Point action	Σ-Optimality for Active Learning on Graphs [Ma et al. 2013] Active Search on Graphs Using Sigma-Optimality [Ma et al. 2014]	Active Area Search [Ma et al. 2014] Active Pointillistic Pattern Search [Ma and Sutherland et al. 2015]
Group action	Active Search for Sparse Rewards with Region Constraints (in progress)	Theory of everything ?

Active Search On Graphs Using Σ-Optimality

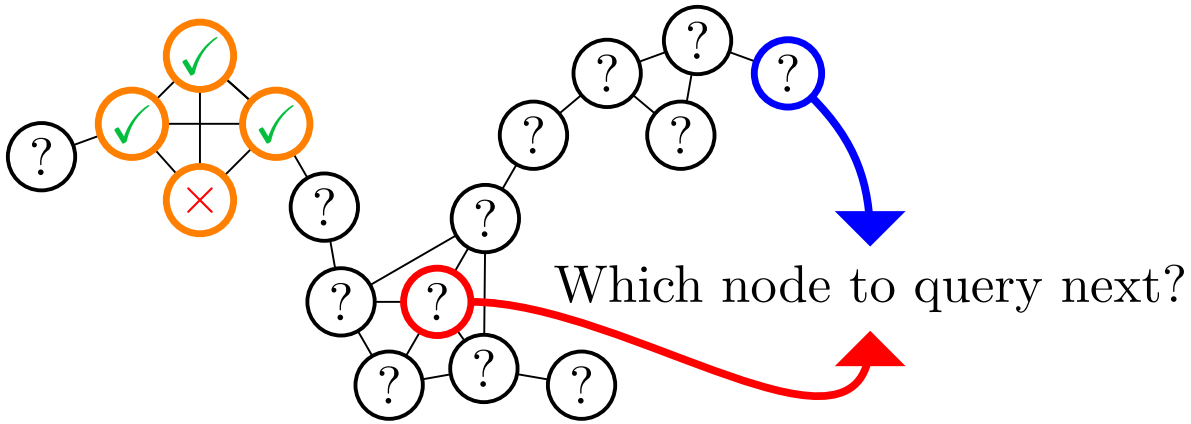
Assume: smooth function on given graph

(Gaussian Markov random fields)

Task: find all ✓ nodes interactively

Approach:

Exploit & explore



Main contribution:

Better exploration (◯ , favoring cluster centers)

than literature (◯ , favoring leaf nodes)

Yifei Ma, Roman Garnett, Jeff Schneider. Σ-Optimality for Active Learning on Gaussian Random Fields. NIPS 2013.
Yifei Ma, Tzu-Kuo Huang, Jeff Schneider. Active Search and Bandits on Graphs Using Sigma-Optimality. UAI 2015.

Ideas of Σ-Optimality

Multi-armed bandits (select by upper bounds)

previous $\arg\max_i \mu_t(i) + \alpha_t \sigma_t(i)$

ours $\arg\max_i \mu_t(i) + \alpha_t \Sigma_t(i)$

Where

$\mu_t(i) =$ posterior marginal mean

$\sigma_t(i) =$ posterior marginal std

$\Sigma_t(i) = \sigma_t(i) + \sum_{j \neq i} \rho_t(i,j)\sigma_t(j)$

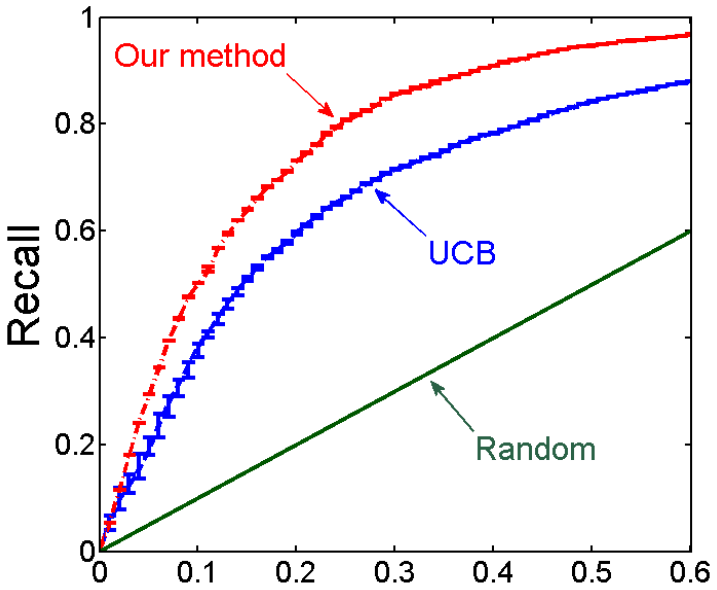
Motivation of "Σ"

Marginal std is large on leaf nodes

"Σ" promotes cluster centers

Reduces variance of global mean surveying

Near-optimality from submodularity



Active search:
Fraction of nodes queried

Yifei Ma, Roman Garnett, Jeff Schneider. Σ-Optimality for Active Learning on Gaussian Random Fields. NIPS 2013.
Yifei Ma, Tzu-Kuo Huang, Jeff Schneider. Active Search and Bandits on Graphs Using Sigma-Optimality. UAI 2015.

Active Pointillistic Pattern Search

Point observations

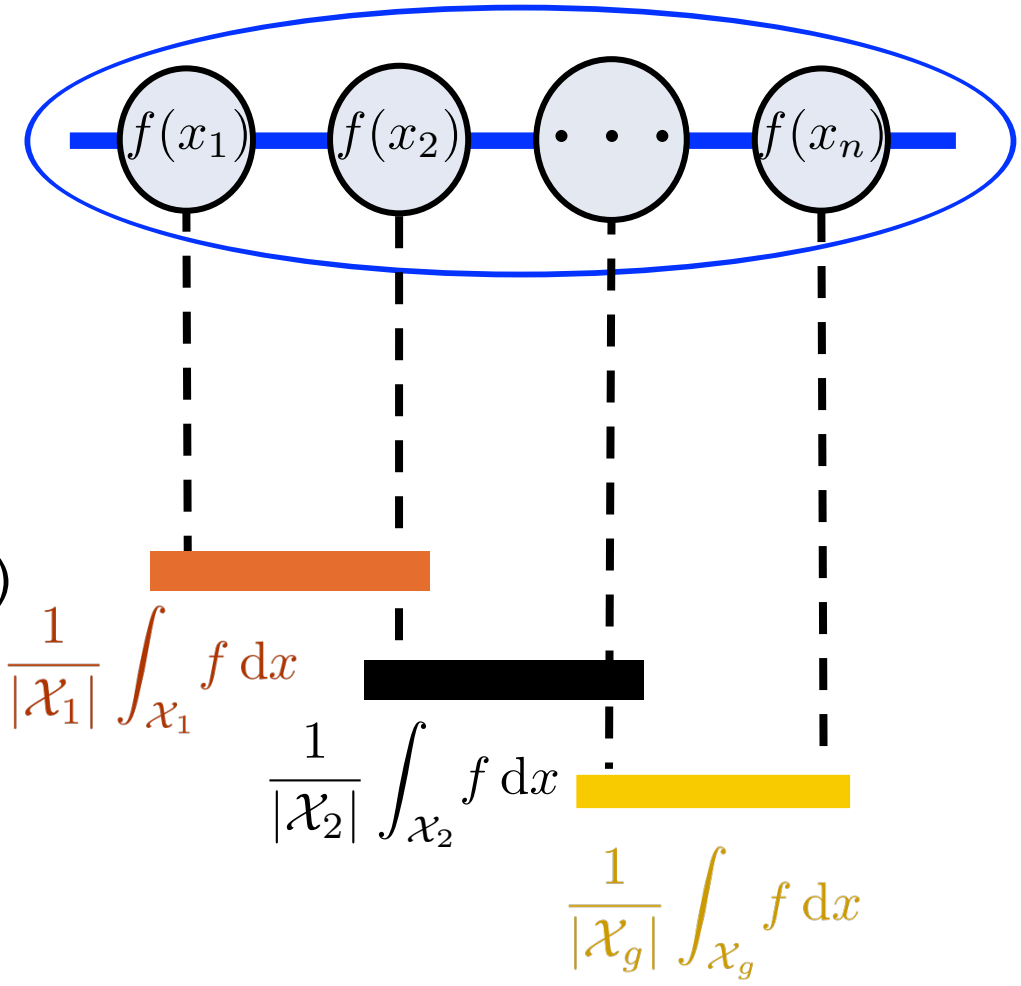
- On the upper level
- Pay to observe

Region rewards

- On the lower level
- Region integral > threshold
- (can replace by a classifier)

Input

- Smooth priors
- Region definitions
- Thresholds / classifiers



Yifei Ma, Roman Garnett, Jeff Schneider. Active Area Search via Bayesian Quadrature. AISTATS 2014

Yifei Ma and Dougal J. Sutherland, Roman Garnett, Jeff Schneider. Active Pointillistic Pattern Search. AISTATS 2015

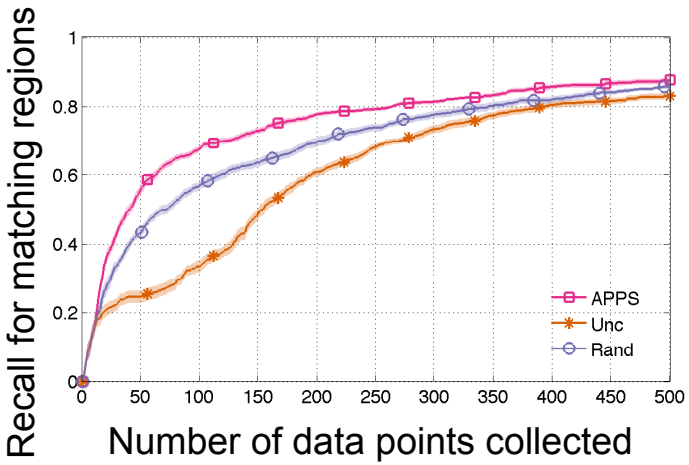
Our Method

Select to **maximize** 1-step look-ahead **expected reward**

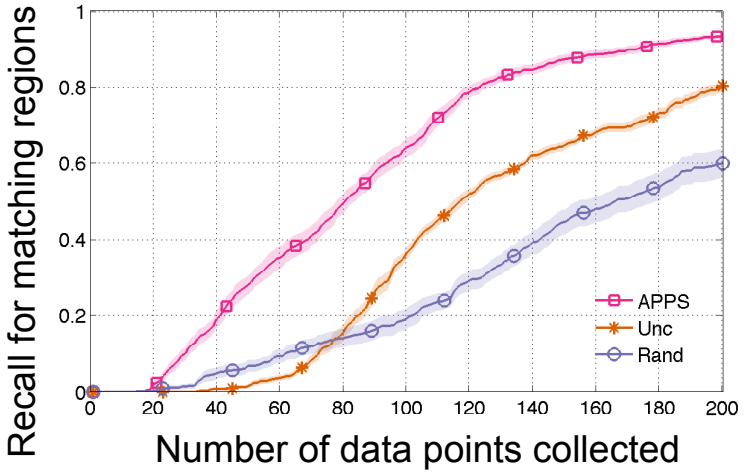
$$x_t = \arg\max_{x_t} \int p_{t-1}(y_t|x_t) \cdot \sum_{g \in \mathcal{G}_t} \mathbf{1}(\text{reward}_g \mid x_{1:t}, y_{1:t}) \, \mathrm{d}y_t$$

x, y point observations; $g \in \mathcal{G}$ regions for rewards

Unbiased **precision**: points in a region are chosen only to reduce variance



Electoral polling for district opinions



Identifying fluid flow vortex patterns by observing point vectors

Yifei Ma, Roman Garnett, Jeff Schneider. Active Area Search via Bayesian Quadrature. AISTATS 2014

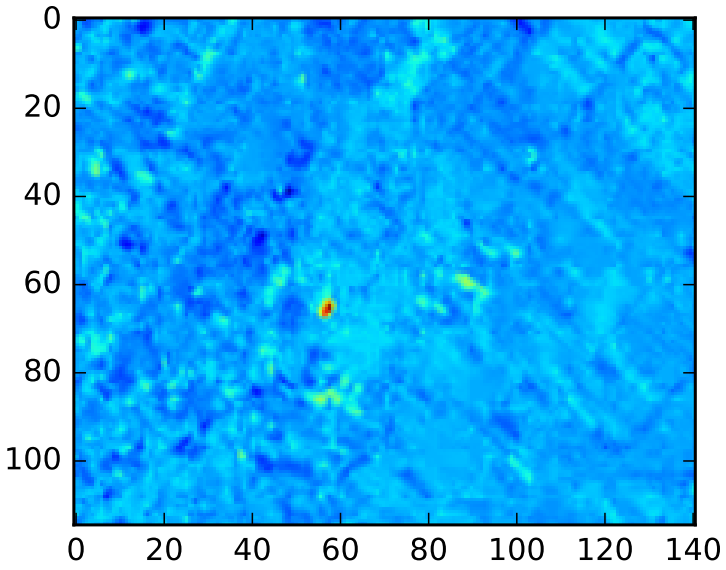
Yifei Ma and Dougal J. Sutherland, Roman Garnett, Jeff Schneider. Active Pointillistic Pattern Search. AISTATS 2015

Active Sensing with Region Sensing Constraints

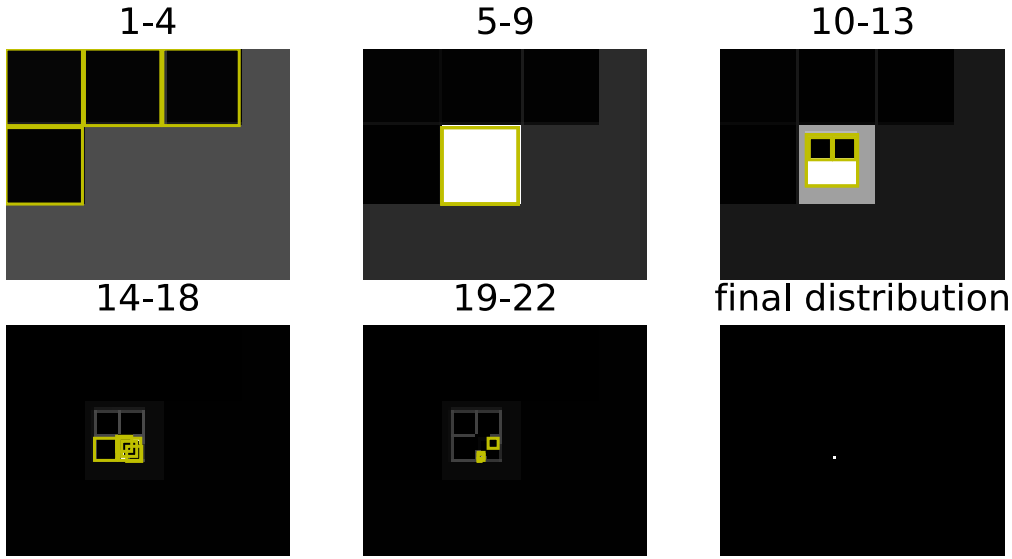
Action: **observe the mean value on any contiguous region with noise**

Objective: **find all non-zero entries of a sparse vector.**

Intuition: **coverage vs. signal decay**



(a) True point values



(b) search sequence

Main Contributions

Method: maximize the Information Gain of an observation

Sample complexity for 1-sparse signals:

$$\tilde{O}\left(\frac{n}{\mu^2}+\frac{1}{\epsilon}\right)$$

μ : signal strength; $\Omega\left(\frac{n}{\mu^2}\right)$ required for active sensing;

$\Omega(n)$ any passive sensing under the region-contiguity constraints.

Can extend to K-sparse signals and nonlinear signal decay functions

work in progress

Thank you!

Active search	Point rewards	Region rewards
Point action	Σ-Optimality for Active Learning on Graphs [Ma et al. 2013] Active Search on Graphs Using Sigma-Optimality [Ma et al. 2014]	Active Area Search [Ma et al. 2014] Active Pointillistic Pattern Search [Ma and Sutherland et al. 2015]
Group action	Active Search for Sparse Rewards with Region Constraints (in progress)	Theory of everything ?

- More flexible group/region definitions
- Categorical/multi-label outcomes
- Ordinal regression
- Imitation/reinforcement learning
- Epoch parallel active search