Auton Lab
Carnegie Mellon

Active Search and Bandits on Graphs Using Sigma-Optimality

Yifei Ma* and Tzu-Kuo Huang** and Jeff Schneider*

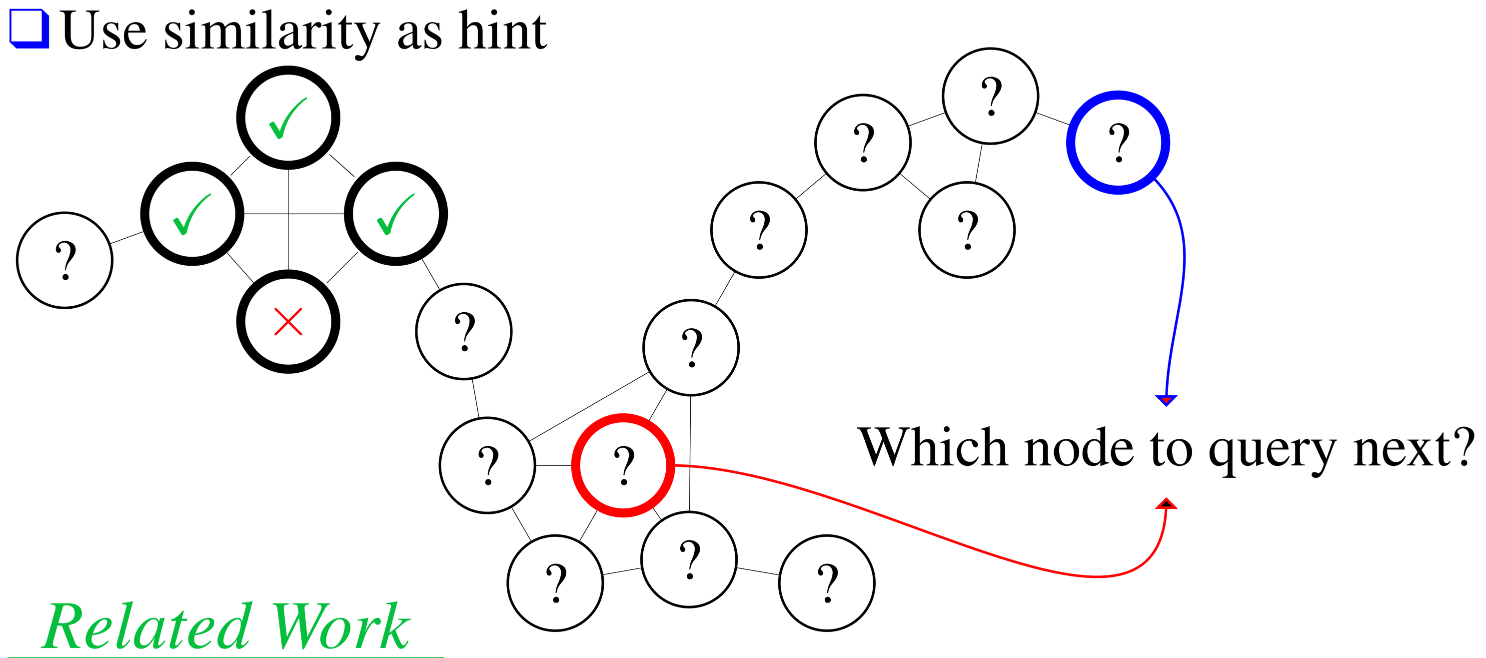
*Auton Lab, School of Computer Science, Carnegie Mellon University. **Microsoft Research

email: yifeim@cs.cmu.edu, tkhuang@microsoft.com, schneide@cs.cmu.edu

Microsoft Research

Problem Setup

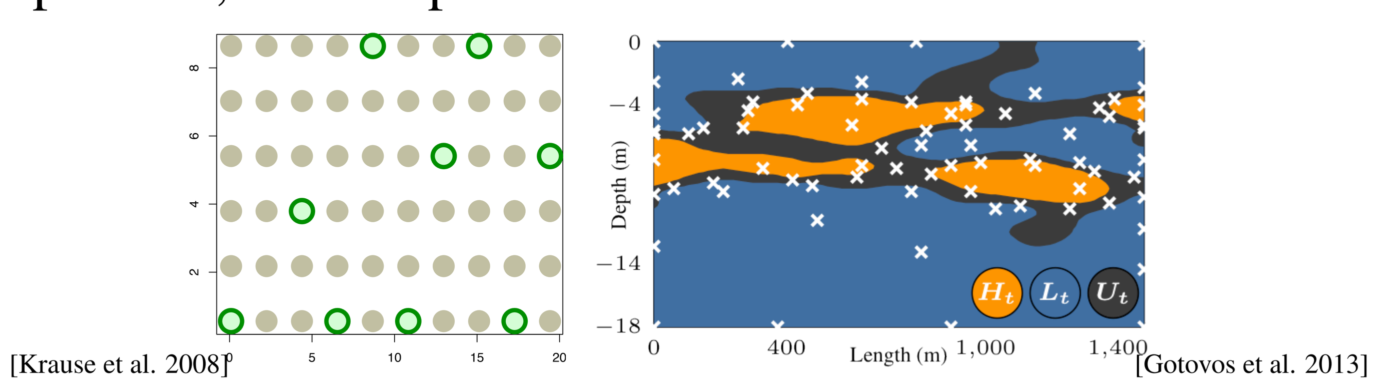
- Given an undirected graph: edges given, nodes unlabeled
- Search for (i.e., query) all positive nodes
- Feedback provided as each node is queried
- Use similarity as hint



Related Work

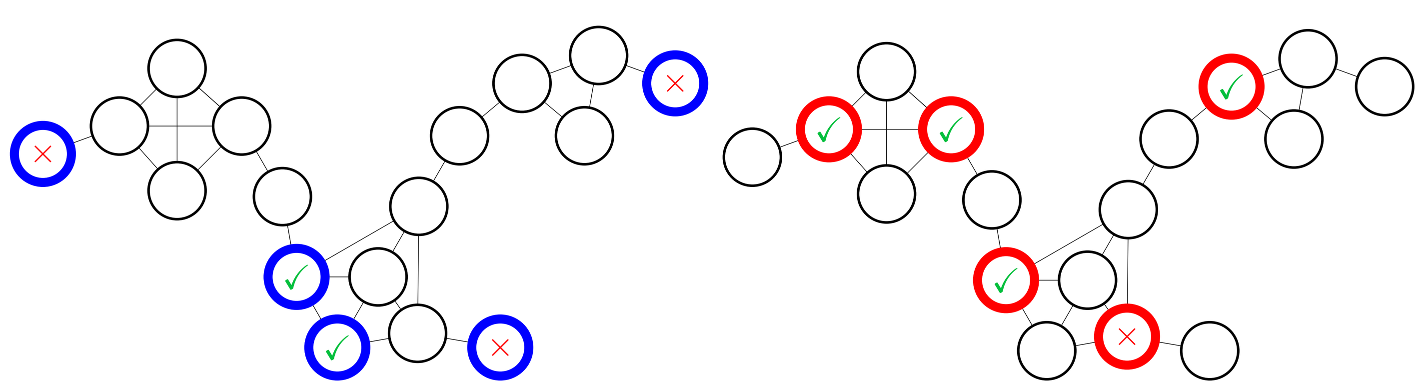
- Query strategy needs to balance:
  - Exploitation: query nodes that are mostly likely targets
  - Exploration: get the most information about label distribution
- GP-UCB [Srinivas et al. 2010; Gotovos et al. 2013]
  - Bayes assumption $f \sim \mathcal{GP}(\mu, c)$
  - Decision rule $v_{t+1} \leftarrow \arg\max_v \mu_t(v) + \alpha_t \sigma_t(v)$.

    Exploitation (post mean), Tradeoff, Exploration (post std)
  - In practice, often explore boundaries first ...



- Spectral-UCB [Valko et al. 2014]
  - Graph kernel from regularized Laplacian, $c(v, v') = \mathbf{1}_v^\top \tilde{\mathcal{L}}^{-1} \mathbf{1}_{v'}$
  - Algorithm unchanged; theoretical guarantees improved
  - Still explores boundaries first ...

Main Contributions



Choices in previous work   Choices by our algorithm

- New exploration term: favor cluster centers over boundaries
  - Measured by improvement of Σ-optimality [Ma et al. 2013]
- High probability regret bounds; compare with
  - [Srinivas et al. 2010; Valko et al. 2014; Contal et al. 2014]
- Empirical results

Bayes Model (GRF) [Zhu et al. 2003]

- Every node v_i has value $f(v_i)$; GRF prior in vector form $\mathbf{f}$,

$$\mathbf{f} \sim \mathcal{N}\Big(\boldsymbol{\mu}_0 = \mu_0 \cdot \mathbf{1},\ \mathbf{C}_0 = \tilde{\mathcal{L}}^{-1} = \big(\mathbf{D} - \mathbf{A} + \omega_0 \mathbf{I}\big)^{-1}\Big), \quad (1)$$

$$\Leftrightarrow \quad \log p_0(\mathbf{f}) \simeq -\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{A_{ij}(f_i - f_j)^2}{2} - \sum_{j=1}^{N} \frac{\omega_0 (f_j - \mu_0)^2}{2}$$

where μ_0, ω_0 are hyper-parameters, $\mathbf{D}$ degree diagonal matrix. Define usual $C(v, v') = \rho(v, v')\sigma(v)\sigma(v')$.

- Objective: cumulative reward in T rounds, $\sum_{t=1}^{T} f_{v_t}$, by active search with noisy feedback $y(v_t) = f(v_t) + \epsilon_t$, $\epsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_n^2)$

Our Methods

Algorithm 1 GP-SOPT and its variants

input μ_0, $\mathbf{A}$, ω_0, σ_n, α_t, T; if warm start, $\{v_\tau, y(v_\tau)\}_{\tau=1}^{t_0}$

Obtain initial $\mathcal{N}(\boldsymbol{\mu}_0, \mathbf{C}_0)$ // (1)

for $t = t_0, \ldots, T-1$, **do**

  Update to conjugate posterior $\mathcal{N}(\boldsymbol{\mu}_t, \mathbf{C}_t)$

  $v_{t+1} \leftarrow \arg\max_{v \in V \setminus S_t} \mu_t(v) + \alpha_t\, s_t(v)$ // (2, 3, or 4)

  Query label $y(v_{t+1})$; include $S_{t+1} \leftarrow S_t \cup \{v_{t+1}\}$

end for

return S_T.

- Exploitation uses posterior mean $\mu_t(v)$.
- Exploration uses **GP-SOPT** (vanilla Σ-optimality)

$$s_t(v) = \frac{\sum_{v' \in V} C_t(v, v')}{\sqrt{\sigma_t^2(v) + \sigma_n^2}} \quad (2)$$

or the following variants,
GP-SOPT.TT (threshold) **GP-SOPT.TOPK** (top-k terms)

$$\min\Big(k\sigma_t(v), s_t(v)\Big) \quad (3) \qquad \max_{B \subset V, |B| = k} \frac{\sum_{v' \in B} C_t(v, v')}{\sqrt{\sigma_t^2(v) + \sigma_n^2}} \quad (4)$$

- Trade-off α_t uses a fixed value in practice; has theoretically optimal choices under lenient assumptions.

Discussions: Σ-Optimality as Exploration

- **Σ-optimality** originally motivated by survey risk minimization,

$$\mathrm{Var}_t(\bar{\mathbf{f}} - \bar{\boldsymbol{\mu}}_t) = \mathrm{Var}_t\Big(\frac{1}{n}\mathbf{1}^\top \mathbf{f} - \frac{1}{n}\mathbf{1}^\top \boldsymbol{\mu}_t\Big) = \frac{1}{n^2}\mathbf{1}^\top \mathbf{C}_t \mathbf{1}$$

has natural tendency to go to cluster centers

- **GP-SOPT** exploration from one-step decrease of Σ-optimality,

$$s_t^2(v) = \mathbf{1}^\top \mathbf{C}_t \mathbf{1} - \mathbf{1}^\top \mathbf{C}_{t+1} \mathbf{1}$$

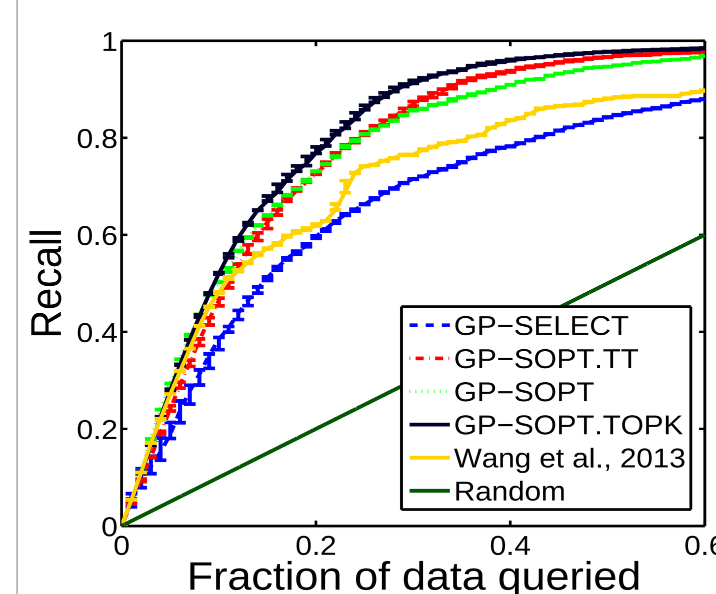
- Another explanation: as $\sigma_n \to 0$, $s_t(v) = \sum_{v' \in V} \rho_t(v, v')\sigma_t(v')$; requires posterior correlation with many uncertain nodes

High Probability Regret Bounds

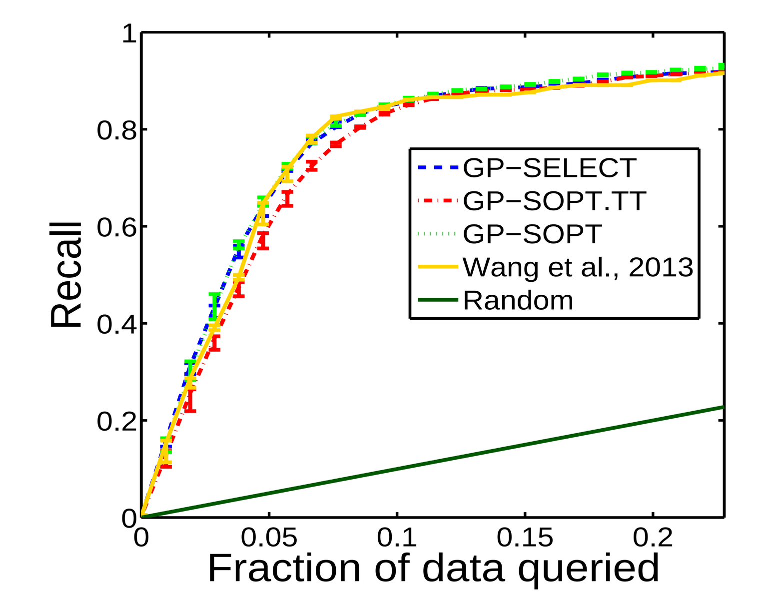
Define Regret	$R_T = \max_{v_t^*, \text{non-repeat}} \sum_{t=1}^{T} f(v_t^*) - f(v_t)$
Define Information	$\gamma_T = \max_{\lvert S \rvert \le T} \mathcal{I}(\mathbf{y}_S; f)$
Assume	$\sqrt{\mathbf{f}^\top \tilde{\mathcal{L}} \mathbf{f}} \le B$, proper α_t $\gamma_T \le d_T^* \log\Big(1 + \frac{T}{\sigma_n^2 \omega_0}\Big)$,
GP-SOPT.TT/TOPK	$\tilde{O}\Big(k\sqrt{T}\big(B\sqrt{d_T^*} + d_T^*\big)\Big)$, any T.
Compare With	$\tilde{O}\Big(\sqrt{T}\big(B\sqrt{d_T^*} + d_T^*\big)\Big)$, [ref 5]

Experiments

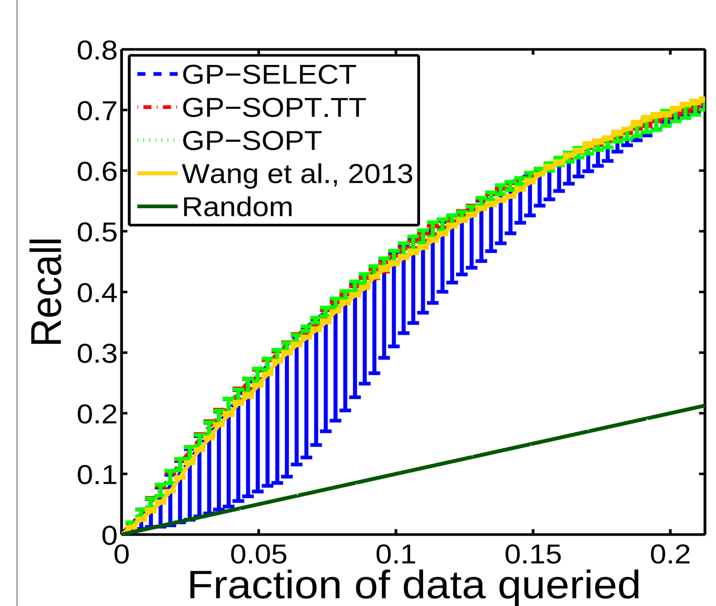
- Populated Places. 725 targets (administrative regions) in 5,000 nodes. Targets spread over components of varying sizes
- Wikipedia Pages on Programming Languages. 202 targets (object-oriented programming pages) in 5,271 nodes. Most targets reside in one large hub.
- Citation Network. 1844 targets (NIPS papers) in 14,117 nodes. Targets appear in many small components
- Enron E-mails. 803 targets (related to downfall of Enron) in 20,112 nodes



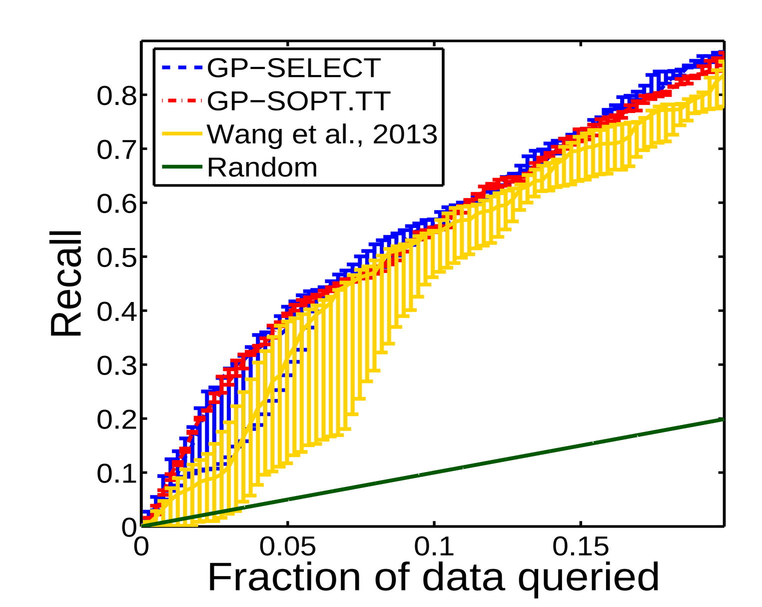
Populated places



Wiki pages



Citation network



Enron e-mails

Significant improvement over existing methods when exploration matters; more robust against outliers.

1. Emile Contal, Vianney Perchet, and Nicolas Vayatis. **GP-MI.** ICML 2014.
2. Andreas Krause, Ajit Singh, and Carlos Guestrin. **Sensor placement.** JMLR 2008.
3. Yifei Ma, Roman Garnett, and Jeff Schneider. **Σ-optimality.** NIPS 2013.
4. Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. **GP-UCB.** TIT'12
5. Michal Valko, Rémi Munos, Branislav Kveton, Tomáš Kocák. **Spectral Bandits.** ICML'14
6. Xuezhi Wang, Roman Garnett, and Jeff Schneider. **Active search on graphs.** SIGKDD 2013.
7. Xiaojin Zhu, Zoubin Ghahramani, John Lafferty, et al. **Semi-supervised GRF.** ICML 2003.