Abstract

Many modern information access problems involve highly complex patterns that cannot be handled by traditional keyword based search. Active Search is an emerging paradigm that helps users quickly find relevant information by efficiently collecting and learning from user feedback. We consider active search on graphs, where the nodes represent the set of instances users want to search over and the edges encode pairwise similarity among the instances. Existing active search algorithms are either short of theoretical guarantees or inadequate for graph data. Motivated by recent advances in active learning on graphs, namely the $\Sigma$-optimality selection criterion, we propose new active search algorithms suitable for graphs with theoretical guarantees and demonstrate their effectiveness on several real-world datasets.

Our problem setting is also similar to multi-armed bandits in that the number of positive instances found under a budget can be seen as the cumulative reward from pulling binary arms (without repetitions). We also discussed theoretical advantages for applying $\Sigma$-optimality as the exploration term for bandits on graphs.

1 INTRODUCTION

As the world gets increasingly digitized and electronically recorded, how to quickly identify relevant pieces of information becomes a major issue. Internet search engines are an integral part of modern life, serving as a probe into the diverse, complex and expanding space of human digital traces. Despite being successful in many information retrieval tasks, the keyword-based query mechanism in most search engines may fall short when targets are characterized by complex patterns or signatures beyond keywords. For example, financial transactions associated with illegal activities bear signatures involving multiple factors such as time, location, occupation of the account owner, etc. In the investigation of organizational misconduct, such as the Enron scandal, the important leads or evidences, oftentimes buried in a sea of diverse electronic and paper trails, usually involve information exchange among key individuals and their relationship. In these situations, keyword-based search may serve as a good starting point, but is certainly far from completing the task.

Such needs of more general search paradigms have recently motivated several efforts Garnett et al. (2012); Wang et al. (2013); Vanchinathan et al. (2013), most of which are related to the Active Search framework proposed by Garnett et al. (2012). It is an interactive search mechanism that begins with the user providing one or few target examples, referred to as seeds, such as past financial transactions that have been linked to illegal activities. Based on these seeds, an algorithm figures out what instance the user should examine next and presents it to the user, who then decides whether the presented instance is relevant or not. Upon receiving this feedback, the algorithm updates its search strategy accordingly and selects the next instance to present. The loop continues until the user quits, and the goal is to maximize the total number of relevant instances found.

As one can see, Active Search has close connections to some well-studied machine learning paradigms. At a first glance, Active Learning (Settles, 2010) seems the most related because they both ask for user feedback incrementally and adaptively. However, Active Learning aims at improving generalization performances with as few label queries as possible, while Active Search is evaluated by how many relevant instances it found along the way, and therefore must carefully balance exploitation and exploration. This trade-off relates Active Search to stochastic optimization in the multi-armed bandit setting (Robbins, 1985; Dani et al., 2008; Kleinberg et al., 2008; Bubeck et al., 2009), where the goal is to find the maximum of an unknown function.
using as few function evaluations as possible. However, Active Search deviates from this setting in that it selects instances without replacement and is competing with the best subset of instances rather than the single best.

We investigate Active Search when the instances are represented by a graph whose edges encode pairwise similarity among the instances. Many real-world data are of this type, such as web pages, citation networks, and e-mail correspondences. For data that are not naturally represented as graphs, a graph representation based on pairwise similarity can still be beneficial because it may reveal useful manifold structures (Tenenbaum et al., 2000; Belkin and Niyogi, 2001). Existing active search approaches (Wang et al., 2013; Garnett et al., 2012; Vanchinathan et al., 2013) either lack theoretical guarantees or ignore certain graph properties, thereby degrading empirical performances. By drawing ideas from recent advances in active learning on graphs (Ma et al., 2013), we proposed new active search algorithms with theoretical guarantees, and empirically demonstrate their advantages over existing methods. In particular, our new exploration criteria, motivated by the Σ-optimality criterion by design. Ma et al. (2013) for active learning on graphs, favor nodes with not only high uncertainty, but also high influence on the other nodes.

The rest of the paper is organized as follows. We describe related work in Section 2 and introduce the problem setup in Section 3. We then present our new methods in Section 4 along with theoretical guarantees, followed by experimental results in Section 5.

2 RELATED WORK

Wang et al. (2013) proposed an active search algorithm for graphs, building on label propagation and semi-supervised learning using Gaussian random fields (Zhu et al., 2003a,b). Despite decent empirical performances, this approach does not have any theoretical guarantee. Vanchinathan et al. (2013) proposed a Gaussian-Process (GP) based algorithm, GP-SELECT, for sequentially selecting instances with high user scores or ratings. This algorithm extends the popular GP-UCB algorithm (Cox and John, 1997; Auer, 2003) for stochastic optimization and inherits nice theoretical guarantees (Srinivas et al., 2012). When applied to graphs, however, it tends to select nodes at the boundary of the graph because they have small degrees and thus large predictive variances, leading to large exploration factors in the GP-UCB selection rule. Yet the labels of these nodes reveal little information about the label distribution over the whole graph.

Similar issues have been observed in active learning on graphs as well. In their experiments, (Ma et al., 2013) found that selection rules based on mutual information gain (Krause et al., 2008), which is closely related to per-node predictive variances, usually end up selecting nodes at the boundary of a graph. Ji and Han (2012) proposed a selection criterion based on not only a node’s predictive variance, but also its covariances with other nodes. Their criterion corresponds to standard V-optimality in experiment design. [Ma et al. 2013] further improved the state of the art by proposing the Σ-optimality criterion, which demonstrates greater robustness against outliers and better empirical performances than V-optimality. Motivated by these recent advances, we propose new active search algorithms that combine GP-UCB with Σ-optimality.

Valko et al. (2014) considered bandit problems where arms correspond to nodes on a graph and the reward is a smooth function over the graph. Their algorithm can be viewed as a special case of GP-UCB with a kernel defined by the inverse of a regularized graph Laplacian. To analyze the performance of their UCB-style algorithm, they propose the notion of effective dimension of a graph, which can be viewed as a measure of the spectral decay of the kernel, thereby determining the performance of the algorithm (Srinivas et al., 2012). We also use the effective dimension to analyze our proposed methods.

As shown later, we benefit from the analysis techniques of Contal et al. (2014) in proving theoretical guarantees of one of the proposed algorithms. They developed a GP optimization algorithm that achieves better theoretical guarantees than GP-UCB. By using their techniques, we are able to make a connection between exploration in the bandit optimization setting and the variance minimization principle in active learning.

3 PROBLEM SETUP

We are given a graph $G$ with known structure (edge connections). The edge connections are nonnegative and we use $A$ to represent the adjacency matrix of $G$, such that $A_{ij} ≥ 0, \forall i, j$. Let $V = \{v_1, \ldots, v_n\}$ denote the set of all nodes in $G$. From $A$ we can derive a graph Laplacian matrix, $L = D - A$, where $D = \text{diag}(A \cdot 1) = \text{diag}(\text{deg}(v_1), \ldots, \text{deg}(v_n))$.

Every node $v$ on our graph holds one data value (label) we denote as $f(v)$, which is not known at first but can be revealed upon player paying one unit price, as

$$y(v) = f(v) + \epsilon, \text{ where } \epsilon \sim \mathcal{N}(0, \sigma_v^2).$$

(1)

Common with bandit problems, querying a node also means collecting this node for reward evaluation based on its true data value. Our goal is to design a query strategy, which generates a random query sequence $v_t = (v_1, \ldots, v_t)^T$, to maximize the cumulative reward

$$F_T = \sum_{t=1}^{T} f(v_t).$$

(2)
The cumulative reward is always upper-bounded by the optimal strategy with unfair knowledge of the true data values on all the nodes. Let \( v_t^* = (v_t^1, \ldots, v_t^n) \) to be the optimal query sequence, our analysis focuses on the cumulative regret between our strategy and the unfairly optimal strategy,

\[
R_T = \sum_{t=1}^T f(v_t^*) - f(v_t).
\]

In our notations, bold letters indicate vectors or matrices, while light letters without subscripts mean functions and light letters with subscripts represent scalars or specific elements. \( t, \tau, \) and \( T \) are time indices, which when applied as subscripts, always mean the selection or model at that time step. Other letters as subscripts, such as \( i, j, n \), always mean the natural indices.

### 3.1 GAUSSIAN RANDOM FIELD PRIOR

We bring statistical assumptions to the generation process of true data values on the nodes, to simplify discussions. Assume that the labels are generated with respect to the adjacency matrix, in that there is always a bigger probability for connected nodes to share similar data values, as

\[
\log p(f) \simeq -\frac{N}{2} \sum_{i=1}^N \sum_{j=1}^N A_{ij} (f_i - f_j)^2 - \frac{1}{2} \sum_{j=1}^N \omega_j (f_j - \mu_0)^2,
\]

i.e., \( f \sim N(\mu_0 \cdot \mathbf{1}, \Sigma_0 = (\mathbf{L} + \omega_0 \mathbf{I})^{-1}) \),

where \( f \in \mathbb{R}^N \) is the vector concatenation of the data values of all nodes in the graph, \( \mu_0 \) is a prior mean, and \( \omega_0 > 0 \) is a regularization parameter. Define the initial covariance matrix as, \( \Sigma_0 = (\mathbf{L} + \omega_0 \mathbf{I})^{-1} \), and denote \( \Sigma_0 = \mathbf{L} + \omega_0 \mathbf{I} \).

The above generative model is also known as Gaussian random fields (GRFs).

### 3.2 POSTERIOR INFERENCE

Assume the nature draws one sample from the generative model, \( \mathbb{H} \), and we use query observations, \( \mathbb{I} \), to converge to that particular draw by performing posterior inference conditioned on the history,

\[
\mathbb{H}_t = \{(v_t, y_t)\}_{t=1}^T = \{v_t, y_t\},
\]

which allows us to update the posterior distribution as,

\[
\log p(f | \mathbb{H}_t) \simeq -\frac{1}{2} (f-\mu_0)^\top \Sigma_0^{-1} (f-\mu_0) - \sum_{\tau=1}^t \frac{(y_\tau - f_\tau)^2}{2\sigma_n^2}.
\]

Notice that the prior distribution and likelihood model form Gaussian conjugate pairs. Denote the posterior distribution as, \( f | \mathbb{H}_t \sim \mathcal{N}(\mu_t, \Sigma_t) \). First, we express \( \mu_t \) and \( \Sigma_t \) using the prior precision matrix, as

\[
\mu_t = \Sigma_t \Sigma_0 \mu_0 + \sum_{\tau=1}^t \frac{y_\tau \epsilon_\tau}{\sigma_n^2}, \quad \Sigma_t^{-1} = \Sigma_0^{-1} + \frac{1}{\sigma_n^2} \mathbf{H}_t.
\]

**Algorithm 1** \( \Sigma \)-optimality active search and its variants

**input** Adjacency matrix \( \mathbf{A} \), set of seed nodes \( S_0 \), query budget \( T \), class imbalance prior \( \mu_0 \), regularization \( \omega_0 \) (default 0.01), observation noise \( \sigma_n \), and tuning parameters, \( \alpha_t \).

**output** A set of \( T \) selected nodes.

1. Compute augmented Laplacian \( \tilde{\mathbf{L}}_0 = \mathbf{D} - \mathbf{A} + \omega_0 \mathbf{I} \)
2. Use \( S_0 \) to obtain the initial \( N(\mu_t, \Sigma_t) \) by (5) or (6)
3. for \( t = t_0 + 1, \ldots, T \), do
4. Solve \( v_t = \arg \max_{v \in \mathbb{V} \setminus S_{t-1}} \mu_t (v) + \alpha_t s_{t-1} (v) \) with \( s_{t-1} (v) \) defined via (10), (11), or (12)
5. Query the label \( y_t \) of \( v_t \)
6. Update \( \mu_t \) and \( \Sigma_t \) by (5) or (6)
7. \( S_t = S_{t-1} \cup \{v_t\} \)
8. end for
9. Return \( S_T \).

where \( \epsilon_{\tau v} = (0, \ldots, 0, 1, 0, \ldots, 0)^\top \) is an indicator vector of index \( v \), and \( \mathbf{H}_t \) is a diagonal matrix of index counts from \( v_t \), whose \( k \)-th diagonal element is \( \sum_{\tau=1}^t \epsilon_{\tau v}(v_k) \).

However, for convenience in later descriptions and to connect to Gaussian Process (GP) literature (Rasmussen and Williams, 2006), we also use the prior covariance matrix to express the posterior distribution, as,

\[
\mu_t (v) = \mu_0 (v) + \mathbf{c}_{v,v}^\top (\mathbf{C}_{v,v} + \sigma_n^2 \mathbf{I}^{-1})^{-1} (y_t - \mathbf{c}_{v,v}),
\]

\[
\mathbf{C}_{t} (v, v') = \mathbf{C}_0 (v, v') - \mathbf{c}_{v,v'}^\top (\mathbf{C}_{v,v} + \sigma_n^2 \mathbf{I}^{-1})^{-1} \mathbf{c}_{v,v'},
\]

where the matrices can all be defined in terms of the prior:

\[
\mathbf{c}_{v,v} = (\mathbf{C}_0 (v_1, v), \ldots, \mathbf{C}_0 (v_T, v))^\top
\]

\[
\mathbf{C}_{v,v} = (\mathbf{C}_0 (v_t, v_t))_{t=1}^\top
\]

\[
\mu_{v_t} = (\mu_0 (v_1), \ldots, \mu_0 (v_T))^\top.
\]

The above update rules also applies to any time interval that starts with \( t_0 \), by replacing prior models with the model at time \( t_0 \).

Define simple notations for correlation coefficients and standard deviations from the covariance matrix,

\[
\sigma_{t}^2 (v) = \rho_{t} (v, v') \sigma_{t} (v) \sigma_{t} (v'),
\]

which implies that \( \sigma_{t}^2 (v) = \mathbf{C}_{t} (v, v) \). Define \( \mathbf{c}_{t} (v) \) to be the column of \( \mathbf{C}_{t} \) corresponding to node \( v \).

### 4 METHOD

Our proposed active search algorithms are described in Algorithm 1. They resemble general exploration-exploitation style algorithms with GPs. Here we focus on binary functions that assign value 1 to relevant or target nodes, and 0 to all other nodes. The seed nodes are assumed to be all relevant. At iteration \( t \), Algorithm 1 selects the next node...
to query based on a deterministic selection rule of the form:

$$
\arg \max_{v \in V \setminus S_{t-1}} \mu_{t-1}(v) + \alpha_t \cdot s_{t-1}(v),
$$

where $\mu_{t-1}(v)$ is the usual exploitation term and $s_{t-1}(v)$ encourages exploration, with the two being balanced by a possibly iteration-dependent parameter $\alpha_t > 0$. The popular GP-UCB algorithm and its extension to Active Search, GP-SELECT (Vanschinathan et al., 2013), amount to setting $s_t(v)^2 = \sigma_t(v)^2$, the predictive/posterior variance of node $v$. Although this is a very reasonable choice in many situations, it may lead to undesirable exploration behaviors on graphs. Under our model assumption, this principle naturally connects with the predictive/posterior covariances of a node $v$ with other nodes, which is derived from the minimization of squared prediction error, known as $V$-optimality in experiment design. Ma et al. (2013) observed that $V$-optimality can still be undesirably sensitive to outliers, and propose the $\Sigma$-optimality criterion:

$$
\Delta^2(v) = \frac{(\sum_{v'} C_t(v, v'))^2}{C_t(v, v) + \sigma_n^2},
$$

which connects to the reduction of Bayes survey risk

$$
R^2_{\Delta} = \mathbb{E} \left( \sum_v f_t(v) - \sum_{v'} \mu_t(v') \right)^2 = \sum_{v, v'} C_t(v, v'),
$$

because

$$
R^2_{\Sigma} - R^2_{\Sigma + 1} = 1^T (C_t - C_{t+1}) 1 = \Delta^2_t(v_t),
$$

where the second equality is easily derived from (6). The $\Sigma$-optimality criterion (8) considers the sum of a node’s covariances with all other nodes, and therefore favors nodes that are well connected to other nodes, which are likely to provide rich information.

We propose three exploration criteria motivated by $\Sigma$-optimality. The first two are variants of (8) with an additional parameter $k$ that we will describe next:

**Thresholded Total Covariance:**

$$
s_t(v) = \min \left( k \sigma_t(v), \sum_{v' \in V} \frac{C_t(v, v')}{\sigma_t(v)} \right).
$$

By Lemma 4 in Appendix A, we have that $s_t(v) \geq \sigma_t(v)$ for both criteria, meaning that $s_t(v)$ maintains the UCB property. Note that if we keep the observation noise variance as in (8), this property might not always hold. As we will show in our theoretical analysis, we put a threshold in (10) against $k \sigma_t(v)$, where $k$ is a tuning parameter, in order to explicitly control the regret of the algorithm. As implied by Lemma 3 in Appendix A the Top-k Covariance criterion (11) is always upper-bounded by $k \sigma_t(v)$. It also captures the intuition that a node’s influence in its local neighborhood already provides much information. We refer to Algorithm 1 combined with these two criteria as GP-SOPT.TT and GP-SOPT.TOPK.

The third criterion uses the square-root of $\Sigma$-optimality without any modification:

**$\Sigma$-Optimality:**

$$
s_t(v) = \frac{\sum_{v'} C_t(v, v')}{\sqrt{C_t(v, v) + \sigma_n^2}}.
$$

As mentioned before, this criterion may not give an upper-confidence bound due to the observation noise variance. However, using analysis techniques recently proposed by Contal et al. (2014), we are able to derive some theoretical guarantees regarding the optimization regret of algorithms using this criterion. More generally, our analysis suggests a connection between the variance minimization principle in active learning and optimization regret in the bandit setting. We refer to Algorithm 1 combined with this criterion as GP-SOPT.

In the next two subsections we discuss in more details the properties of various exploration criteria, and present our theoretical analysis.

### 4.1 DISCUSSIONS

Our approach and two well-discussed criteria, information gain from Srinivas et al. (2012) and $V$-optimality in Ji and Han (2012), may also be connected via spectral functions of the posterior covariance matrix.

Suppose $C_t$ has decreasingly sorted eigen-values $\lambda^2_t = (\lambda^2_{t,(1)}, \ldots, \lambda^2_{t,(m)})^T$, and let $C_{t+1}$ and $\lambda^2_{t+1}$ to be their posterior counterparts after observing a node, $v$. A score based on spectral difference is then,

$$
s_t(v) = h^{-1} \left( \sum_{k=1}^n h(\lambda_{t,(j)}) - \sum_{k=1}^n h(\lambda_{t+1,(j)}) \right),
$$

for any monotone increasing function $h$. Notice that $C_t$ is positive semi-definite and $\lambda_{t,(j)}$ is monotone decreasing as $t$ increases.
Case 1. \( h(s) = -\log(s), s_t(v) = \sqrt{1 + \sigma_t^2(v) / \sigma_v^2} \). This heuristic is qualitatively similar to GP-UCB [Srinivas et al. 2012].

Case 2. \( h(s) = s^2, s_t(v) = \sqrt{\text{tr}(C_t) - \text{tr}(C_{t+1})} = \sqrt{\frac{\|c_t(v)\|^2}{\sigma^2_t(v) + \sigma_v^2}} \). This criterion resembles V-optimality, Jr and Han [2012].

Case 3. \( h(s) = s^p, p \to \infty, \lambda_{\max}(C_t) - \lambda_{\max}(C_{t+1}) \).

Our intuition is that \( \Sigma \)-optimality connects to this criterion. This is because its rank-one update rule is approximately, \( s_t(v) \approx \frac{1}{\lambda_{\max}(C_t)} \frac{c_t(v)}{\sqrt{\sigma_t(v)^2 + \sigma_v^2}} \), where \( q_t \) is the eigen-vector corresponding to the largest eigen-value of \( C_t \).

To compare these spectral functions, it seems that establishing biases to penalize more on a large eigen-value is beneficial, because it improves global robustness, as the posterior marginal variance of every node is upper-bounded by \( \lambda_{i,1} \).

Otherwise, the exploration may then have a tendency to shrink small eigen-values towards zeros, most effectively achieved via querying leaf nodes on the periphery nodes. \( \Sigma \)-optimality has the strongest bias among all three to penalize large eigen-values for the initial explorations. It also avoids a potential side-effect of Case 3 when \( t \) becomes large. Exact execution of Case 3 may over-simplify the true model as \( t \to \infty \).

4.2 REGRET ANALYSIS

We present an UCB-style analysis for GP-SOPT.TT and GP-SOPT.TOPK, and an analysis based on Contal et al. [2014] for GP-SOPT. We combine several results on GP optimization [Srinivas et al. 2012; Vanchinathan et al. 2013; Contal et al. 2014] and the spectral bandit analysis (Valko et al. 2014). As in these results, our regret bounds depend on the mutual information between \( f \) and observed values \( y_S \) at a set \( S \) of nodes:

\[
I(y_S; f) := H(y_S) - H(y_S | f),
\]

(14)

where \( H(\cdot) \) denotes the entropy. If \( f \) is drawn from a GP with observation noise distributed independently as \( N(0, \sigma_n) \), the mutual information has the following analytical form:

\[
I(y_S; f) = I(y_S; f_S) = \frac{1}{2} \log |I + \sigma_n^{-2} C_{vSvS}|.
\]

(15)

Let

\[
\gamma_T := \max_{S \in \mathcal{V} \mid |S| = T} \frac{1}{2} \log |I + \sigma_n^{-2} C_{vSvS}|,
\]

(16)

i.e., the maximum information about \( f \) gained by observing \( T \) function evaluations. The regrets of our algorithms depend on the growth rate of \( \gamma_T \), which can be linear in \( T \) for arbitrary graphs. However, real-world graphs often possess rich structures, such as clusters or communities, and practical measures of relevance are often highly correlated with these structures, resulting in slowly-growing \( \gamma_T \).

To formalize this intuition, we follow Valko et al. [2014] to consider the effective dimension:

\[
d_T^* := \max \left\{ i \mid \lambda_i \leq \frac{\sigma_n^{-2} T - i - 1}{\log\left(1 + \frac{q_i}{\sigma_n^2 \omega_0}\right)} \right\},
\]

(17)

where \( \lambda_i \) is the \( i \)th smallest eigenvalue of \( \tilde{L}_0 \) and \( \lambda_1 = \omega_0 \). The effective dimension is small when the first few \( \lambda_i \)'s are small and the rest increase rapidly, as is often the case for graphs with community or cluster structures. On the contrary, if all the eigenvalues are close to \( \omega_0 \), then \( d_T^* \) may be linear in \( T \). The following lemma bounds \( \gamma_T \) in terms of \( d_T^* \):

**Lemma 1.** Let \( T \) be the total number of rounds. Then

\[
\gamma_T \leq 2d_T^* \log \left(1 + \frac{T}{\sigma_n^2 \omega_0}\right).
\]

Proof. By Lemma 7.6 of Srinivas et al. [2012] and the fact that \( \lambda_i^{-1} \) is the \( i \)th largest eigenvalue of the kernel \( C_0 = \tilde{L}_0^{-1} \), we have

\[
\gamma_T \leq \max_{\{m_i\}_{i=1}^{T}, \{m_i\}_{i=1}^{T} \geq 0} \sum_{i=1}^{T} \sum_{m_i = 0}^{\infty} \log \left(1 + \frac{m_i}{\sigma_n^2 \lambda_i}\right).
\]

(18)

Then by applying the same argument that proves Lemma 6 of Valko et al. [2014], we obtain the desired result. \( \Box \)

4.2.1 Active Search Regret

Recall the cumulative regret of an active search algorithm is defined as

\[
R_T := \sum_{t=1}^{T} f(v_t^*) - f(v_t),
\]

(19)

where \( \{v_t\}_{t=1}^{T} \) is the sequence of unique nodes selected by the algorithm. For the two proposed UCB-style algorithms, GP-SOPT.TT [10] and GP-SOPT.TOPK [11], we give the following bounds on their cumulative regrets.

**Theorem 2.** Pick \( \delta \in (0, 1) \). Assume the vector of true node values, \( f \), has bounded quadratic norm, \( \|f\|_{\tilde{L}_0} = \sqrt{f^T \tilde{L}_0 f \leq B^2} \) and the observation noise \( \epsilon_t \) is zero-mean conditioned on the past and is bounded by \( \sigma_n \) almost surely. If GP-SOPT.TT and GP-SOPT.TOPK use GRF prior \([4]\) with zero-mean and graph Laplacian \( \tilde{L}_0 \), the observation noise model \( N(0, \sigma_n^2) \), and \( \alpha_t := \sqrt{2B + 300 \gamma_i \log^3(t/\delta)} \), then their cumulative regrets will satisfy

\[\Pr\left(\{R_T \leq k \sqrt{c_1 T \alpha T \gamma_T} \forall T \geq 1\} \right) \leq 1 - \delta,\]

This is similar to a bounded RKHS norm with kernel \( C_0 \) in Srinivas et al. [2012].
where the randomness is over the observation noise and $c_1 := \frac{s/\omega_0}{\log(1 + \sigma_0^2)}$. This implies that with high probability,

$$R_T = O\left(kv\sqrt{T}B\sqrt{d_T^* + d_T^+}\right).$$

This result is easily derived from the regret analysis of the GP-SELECT algorithm proposed by Vanchinathan et al. (2013) because the exploration terms used by GP-SOPT.TT and GP-SOPT.TOPK both satisfy $\sigma_i(v) \leq s_i(v) \leq k\sigma_i(v)$, thereby maintaining the UCB property. Although our regret bound is $k$ times worse than the GP-SELECT bound, the actual regret tends to behave more favorably as we observe in our experiments that after a few tens of rounds, $s_i(v)$ becomes smaller than $k\sigma_i(v)$ for almost all unqueried nodes, and the two proposed algorithm usually outperforms GP-SELECT. We give the proof in Appendix B for completeness.

### 4.2.2 Optimization Regret

Here we analyze the GP-SOPT algorithm, whose exploration term is exactly the risk reduction criterion in $\Sigma$-optimality without any modification. As shown earlier, that criterion may not give rise to an upper-confidence bound and we thus resort to a different type of analysis proposed recently by Contal et al. (2014), who studied sequential global optimization with Gaussian processes. Using their key result concerning a general GP-based sequential global optimization algorithm, we obtain a bound on the optimization regret:

$$R_T^* := \sum_{t=1}^T f(v^*) - f(v_t),$$

where $v^* \in \arg \max_{v \in V} f(v)$ and $\{v_t\}_{t=1}^T$ is the sequence of nodes picked by GP-SOPT with replacement.

Our bound depends on the maximum total risk reduction in $\Sigma$-optimality:

$$\Delta_T := \max_{v_T} 1^T(C_0 - C_T|v_T)1,$$

where $v_T$ denotes a sequence of $T$ nodes and $C_T|v_T$ denotes the posterior covariance conditioned on $v_T$.

**Theorem 3.** Pick $\delta \in (0, 1)$. Assume the true function $f$ is drawn from a GP with zero mean and covariance $C_0 = (\mathcal{L} + \omega_0 I)^{-1}$, and the observation noise is distributed independently as $N(0, \sigma^2)$. Let $T$ denote the total number of rounds and $\sigma_0^2$ max := $1/\omega_0 + \sigma_0^2$. The optimization regret of GP-SOPT, GP-SOPT.TT, and GP-SOPT.TOPK using $\alpha_t = \frac{2 \log(2/\delta)\sigma_{\max}^2}{\sqrt{T}\Delta_T}$ all satisfy that for all $T > 1$, 

$$P\left(R_T^* \leq 2\sqrt{2\log(2/\delta)\sigma_{\max}^2 \sqrt{T}\Delta_T}\right) \geq 1 - \delta.$$  

Further, for graphs that contain large clusters, $\Delta_T$ may introduce a significant improvement. Since one may obtain good surveying results after sampling a few cluster centers, $\Delta_T$ usually grows much slower after a few iterations.

### 5 EXPERIMENTS

We conduct experiments on three graph data sets that were studied by Wang et al. (2013) and a version of the Enron e-mail data by Priebe et al.

#### 5.1 Three Graph Data Sets of Wang et al. (2013)

We briefly summarize the data sets below.

**5000 Populated Places.** The nodes of this graph are 5000 concepts in the DBpedia ontology marked as populated places. Each place is supported by a Wikipedia page, and an undirected edge is created between two places if either one of their two Wikipedia pages links to the other. There may be multiple edges between two places. The DBpedia ontology divides populated places into five categories: administrative regions, countries, cities, towns, and villages. The 725 administrative regions are selected as our target class while all the others are considered to be in null class.

**Citation Network.** This dataset consists of 14,117 papers in top Computer Science venues available on citeseer. The graph is created by adding an undirected edge between two papers if either one cites the other. The 1844 NIPS papers are chosen as our target class.

**Wikipedia Pages on Programming Languages.** A total of 5,271 Wikipedia pages related to programming languages are the nodes of this graph, and an undirected edge exists between two pages if they are linked together. Wang et al. (2013) performed topic modeling and chose the 202 pages related to objective oriented programming as our target class.

As demonstrated by Wang et al. (2013), the three graphs and their target label distributions exhibit qualitative differences and thus serve as good benchmarks. The citation network has many small components and target nodes appear in many of them, while the Wikipedia graph has large hubs and most target nodes reside in one of them. The graph of populated places lies in between these two extremes, with components of various sizes containing target nodes.

---

1 Notice its subscript difference with [19].

2 www.dbpedia.org
On all of the three data sets we compare two of the proposed methods: GP-SOPT.TT and GP-SOPT against GP-SELECT (GP-UCB without replacement) and the active search algorithm (AS-on-Graph) by Wang et al. (2013).

We only evaluate GP-SOPT.TOPK on the 5000 populated places data due to its heavy computation. For each dataset we perform 5 independent runs, each with a randomly chosen target node as the seed. For the proposed methods and GP-SELECT, the main tuning parameters are the exploration-exploitation trade-off parameter \( \alpha \), and the observation noise variance \( \sigma^2 \). For GP-SOPT.TT and GP-SOPT.TOPK there is additionally the thresholding parameter \( k \). We consider the following values for them. Populated Places: \( \alpha \in \{4, 2, 1, 0.1, 0.01, 0.001\} \), \( \sigma^2 \in \{1, 0.5, 0.25, 0.1\} \) and \( k \in \{200, 400, 800\} \). Wikipedia: \( \alpha \in \{0.1, 0.01, 0.001\} \), \( \sigma^2 \in \{1, 0.5, 0.25, 0.1\} \) and \( k \in \{200, 400, 800\} \). Citation Network: \( \alpha \in \{1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\} \), \( \sigma^2 \in \{1, 0.5, 0.25, 0.1\} \) and \( k \in \{400, 800, 1600\} \). Although in theory \( \alpha \) should be iteration-dependent, we find that a fixed value often performs well in practice. On all data sets we set the kernel regularization parameter \( \omega_0 = 0.01 \). The AS-on-Graph algorithm has several parameters, and we only tune the exploration-exploitation trade-off parameter \( \alpha \). It is set to 0.1 on Populated Places and Citation Network, and 0.0001 on Wikipedia, which are the best performing values. Other parameters are set based on Wang et al. (2013).

Results are in Figure 1 where we plot the recall, i.e., the fraction of targets found by the algorithms, versus the fraction of the whole data set queried. More specifically, for each algorithm we obtain its mean recall curve over the top 15% (except for AS-on-Graph) parameter combinations in each experiment, as judged by the area under the recall curve. We then plot the median, maximum and minimum over the five runs in Figure 1.

The three proposed methods clearly outperform AS-on-Graph and GP-SELECT on Populated Places, while all methods perform equally well on Wikipedia. We think this has to do with the underlying graph structure and target distribution. As mentioned before, target nodes in the Populated Places graph are spread over sub-graphs of various sizes, and therefore exploration strategies do make a difference. We observe that the proposed methods tend to select high-degree nodes in the first few iterations, thereby gaining much information, while GP-SELECT initially selects low-degree nodes. In contrast, most target nodes in the Wikipedia graph reside in one large component, and therefore less exploration is needed. In fact, the best values for \( \alpha \) are very small, suggesting that an exploitation-only strategy is good enough for this data. On Citation Network, most methods perform well except that GP-SELECT performs quite poorly in one run. This may again indicate GP-SELECT is less robust against low-degree nodes.

### 5.2 Enron E-mails

We experimented on the Enron e-mail data set\(^4\) with topics assigned by Priebe et al.,\(^5\) based on the annotations by Berry and Browne. We further processed the dataset into a format suitable for active search experiments as detailed below. Each e-mail \( i \) is represented by a unique Unix time stamp \( t_i \), a unique sender index and the set of receiver (excluding self-copying) indices, which are collectively denoted as \( U_i \).

Between e-mails \( i \) and \( j \), we created an edge with the following weight:

\[
A_{ij} := \exp \left(-\frac{(t_i - t_j)^2}{\tau^2}\right) \cdot |U_i \cap U_j|/\sqrt{|U_i||U_j|},
\]

where \( \tau = 12 \) weeks in seconds and \( |U_i| \) denotes the size of \( U_i \). We thus measure pairwise similarity among e-mails by the product of nearness in time and degree of overlap between users involved. The resulting e-mail graph has 20,112 nodes, and we chose the subset of 803 e-mails that are assigned topic 16 in LDC topic\(^4\) which is related to the downfall of Enron, to be the target class in this experiment.

Due to the size of the dataset, we only compared three methods: GP-SOPT.TT, GP-SELECT and AS-on-Graph in three independent runs each initialized with a target

\(^4\)Available at [http://cis.jhu.edu/~parky/Enron/execs.email.linednum.ldctopic](http://cis.jhu.edu/~parky/Enron/execs.email.linednum.ldctopic)

\(^5\)Available at [http://cis.jhu.edu/~parky/Enron](http://cis.jhu.edu/~parky/Enron)
node chosen uniformly at random. We also limited the tuning parameters to be the following fixed values across the three runs: \((k, \alpha, \sigma^2, \omega_0) = (800, 0.001, 0.05, 0.01)\) for GP-SOPT.TT, \((\alpha, \sigma^2, \omega_0) = (0.01, 0.05, 0.01)\) for GP-SELECT, and \(\alpha = 0.001\) for AS-on-Graph. These values were chosen based on a coarse parameter search to be indicative of the performance of each method on this data set. Results are in Figure 2 which shows GP-SOPT.TT is more stable across initial seeds than the other methods, and outperforms AS-on-Graph significantly at early iterations.

Acknowledgement

This work was funded in part by DARPA grant FA87501220324.

A Predictive Covariance Matrix

**Lemma 4.** For augmented graph Laplacian, the posterior covariance matrix, \(C_t(v, v') \geq 0, \forall v, v'\).

**Proof.** Let \(\ell_k = \sum_{\tau=1}^t e_{v_k}(v_k)\) to be the count of queries on node \(k\); further define its diagonal matrix, \(H = \text{diag}(h_1, \ldots, h_n)\). We rewrite (5) as,

\[
(C_t)^{-1} = (C_0)^{-1} + \sigma_n^{-2}H = D - A + \omega_0I + \sigma_n^{-2}H
\]

Define \(D_t = D + \omega_0I + \sigma_n^{-2}H\), we have

\[
C_t = (D_t - A)^{-1} = D_t^{-\frac{1}{2}} \left( \sum_{k=0}^{\infty} \left( D_t^{-\frac{1}{2}} AD_t^{-\frac{1}{2}} \right)^k \right) D_t^{-\frac{1}{2}},
\]

where the right hand side is always nonnegative.

The convergence of \(\|D_t^{-\frac{1}{2}} AD_t^{-\frac{1}{2}}\|_2 < 1\) is as follows.

Define the components for the posterior as \(D_t = \text{diag}(d_i^{(t)}, \ldots, d_n^{(t)})\) with \(d_i^{(t)} = \sum_{k=1}^{t} d_k^{(t)}\). Also, define for the prior model \(D = \text{diag}(d_1^{(0)}, \ldots, d_n^{(0)})\) with \(d_i^{(0)} = \sum_{k=1}^{n} d_k^{(0)}\).

The following holds for any \(v \in \mathbb{R}^n\),

\[
v^T D_t^{-\frac{1}{2}} A D_t^{-\frac{1}{2}} v = \sum_{ij} \frac{v_i v_j a_{ij}}{\sqrt{d_i^{(t)}} \sqrt{d_j^{(t)}}} \leq \sqrt{\left( \sum_{ij} \frac{v_i^2 a_{ij}}{d_i^{(t)}} \right) \left( \sum_{ij} \frac{v_j^2 a_{ij}}{d_j^{(t)}} \right)} = \sum_i v_i^2 \frac{d_i}{d_i^{(t)}} \leq \|v\|_2^2.
\]

Further, both equalities cannot hold simultaneously, because for the first equality to hold, it is required that \(\frac{v_i^2 a_{ij}}{d_i^{(t)}} \propto \frac{v_j^2 a_{ij}}{d_j^{(t)}}\), i.e., \(v_j^2 \propto d_j^{(t)}\), \forall v_i, \forall v_j\) in the same connected component, which then dictates that,

\[
\sum_i v_i^2 \frac{d_i}{d_i^{(t)}} = \sum_i \left( \frac{d_i^{(t)}}{d_i^{(t)}} \|v\|_2^2 \right) \frac{d_i}{d_i^{(t)}} = \frac{d(0)}{d(t)} \|v\|_2^2 < \|v\|_2^2.
\]

**Lemma 5.** The diagonal elements in \(C_t\) is always no smaller than the off-diagonal elements, i.e., \(\sigma_t(v)^2 = C_t(v, v) \geq C_t(v, v'), \forall v, v'\).

**Proof.** Without loss of generality, let \(v\) be the last index of \(C_t = (\mathcal{L}_0 + \sigma_n^{-2}H)^{-1}\). For simplicity, let \(\mathcal{L}_t = \mathcal{L}_0 + \sigma_n^{-2}H\) and it has the following matrix partition,

\[
\mathcal{L}_t = \begin{pmatrix}
\mathcal{L}_{\bar{v} \bar{v}} & \mathcal{L}_{\bar{v} v} \\
\mathcal{L}_{v \bar{v}} & \mathcal{L}_{vv}
\end{pmatrix},
\]

where \(\bar{v}\) is the complement of \(v\). From Woodbury matrix inversion lemma, we have

\[
C_t = \mathcal{L}_t^{-1} = \begin{pmatrix}
M & 0 \\
\mathcal{L}^{-1}_{\bar{v} \bar{v}} & -\frac{1}{m} \mathcal{L}_{\bar{v} v} \mathcal{L}_{v \bar{v}}^{-1} \mathcal{L}_{v v}
\end{pmatrix},
\]

where \(m = \mathcal{L}_{\bar{v} v} \mathcal{L}_{v v}^{-1} \mathcal{L}_{v \bar{v}} = \mathcal{L}_{\bar{v} v} \mathcal{L}_{\bar{v} \bar{v}}\) and \(M = \mathcal{L}_{\bar{v} \bar{v}}^{-1} + \frac{1}{m} \mathcal{L}_{\bar{v} v} \mathcal{L}_{v v} \mathcal{L}_{v \bar{v}}^{-1} \mathcal{L}_{v v}\). To show that \(C_t(v, v) \geq C_t(v, v')\), we need to verify that \((-\mathcal{L}_{\bar{v} v} \mathcal{L}_{v v})v' \leq 1\).

In fact, since \(\mathcal{L}_t\) is diagonally dominant, we have \(\mathcal{L}_{11} \geq 0\). Take its first \(n - 1\) rows to get \(\mathcal{L}_{\bar{v} v} \cdot 1_{n-1} + \mathcal{L}_{v v} \geq 0\). Notice \(\mathcal{L}_{\bar{v} v}\) is also a valid augmented graph Laplacian. By Lemma 4 we could left multiply the element-wise nonnegative matrix \(\mathcal{L}_{\bar{v} v}^{-1}\) to both sides to obtain, \(1_{n-1} + \mathcal{L}_{\bar{v} v}^{-1} \mathcal{L}_{v v} \geq 0\), which completes our proof for any \(v' \in \bar{v}\).

B Active Search Regret Bound

We start by stating the following result.

**Theorem 6 (Theorem 6, Srinivas et al. (2012)).** Let \(\delta \in (0, 1)\). Assume the observation noises are uniformly bounded by \(\sigma_n\) and \(f\) has RKHS norm \(B\) with kernel
Lemma 8. Conclude on the high-probability of Theorem 2 the following bound holds:

\[ \forall t, r_t := f(v_t^*) - f(v_t) \leq 2\alpha_t k\sigma_{t-1}(v_t), \]

where \( v_t^* \) is the node with the \( t \)-th globally largest function value and \( v_t \) is node selected at round \( t \).

Proof. At round \( t \) there are two possible situations. If \( v_t^* \) was picked at some earlier round, the definition of \( v_t^* \) implies that there exists some \( t' < t \) such that \( v_t^* \) has not been picked yet. According to our selection rule, the fact that \( s_t(v) \geq \sigma_t(v) \), and Theorem 6 holds the following:

\[ \mu_{t-1}(v_t) + \alpha_t \sigma_{t-1}(v_t) \geq \mu_{t-1}(v_t^*) + \alpha_t \sigma_{t-1}(v_t^*) \]

\[ \geq \mu_{t-1}(v_t^*) + \alpha_t \sigma_{t-1}(v_t^*) \geq f(v_t^*) \geq f(v_t). \]

Thus we always have

\[ f(v_t^*) \leq f(v_t) + \alpha_t \sigma_{t-1}(v_t) \]

\[ \leq f(v_t) + \alpha_t \sigma_{t-1}(v - t) + \alpha_t \sigma_{t-1}(v_t) \]

\[ \leq f(v_t) + 2\alpha_t k\sigma_{t-1}(v_t). \]

Lemma 8 (Lemma 5.4, Srinivas et al. (2012)). Let \( \alpha_t \) be defined as in Theorem 2 and \( c_1 \) be defined as in Theorem 2. Conclude on the high-probability of Theorem 4 the following holds:

\[ \forall T \geq 1, \sum_{t=1}^{T} r_t^2 \leq \alpha_T k^2 c_1 T \sqrt{T} \sum_{t=1}^{T} r_t \leq \alpha_T k^2 c_1 \gamma T. \]

Finally, the Cauchy-Schwarz inequality gives \( R_T \leq \sqrt{T \sum_{t=1}^{T} r_t^2} \leq k \sqrt{T c_1 \alpha_T T}. \)

C Optimization Regret Bound

Under the assumptions in Theorem 3 we use a stronger theorem that follows similar proof techniques as Theorem 1 of Contal et al. (2014) as the following.

Theorem 9. With probability at least \( 1 - \delta \), the optimization regret satisfies for \( \forall T > 1, \forall y \geq 0 \) and any \( C_{T_1} \geq \sup T \sum_{t=1}^{T} \sigma_{t-1}(v_t) \)

\[ R_T^2 \leq \sum_{t=1}^{T} \alpha_t (s_{t-1}(v_t) - s_{t-1}(v^*)) \]

\[ + \sqrt{2 \log(2/\delta)} C_{T_1} + \sqrt{2 \log(2/\delta)} \sum_{t=1}^{T} \sigma_{t-1}^2(v^*) \sqrt{T \Delta T}. \]

Proof. The sum of error in reward is a martingale, \( M_T = \sum_{t=1}^{T} (f(v_t) - f(v^*)) - (\mu_t(v_t) - \mu_t(v^*)) \). Let \( [M_T]_T = \text{Var}(M_t - M_{t-1}) \leq \sum_{t=1}^{T} \sigma_{t-1}^2(v_t) + \sigma_{t-1}^2(v^*). \)

By combining (4.1) and (4.2) in Bercu et al. (2008) while explicitly using fixed \( C_{T_1} > \sum_{t=1}^{T} \sigma_{t-1}^2(v_t) \) and \( y \),

\[ \mathbb{P}(y < M_T \geq x) \leq \exp \left\{ -\frac{x^2}{2y} \right\} \]

\[ \mathbb{P}(M_T > y, M_T \geq x + \sum_{t=1}^{T} x_t \sigma_{t-1}^2(v^*) C_{T_1}) \leq \exp \left\{ -\frac{x'^2 y}{2C_{T_1}^2} \right\} \]

Solving for RHS has \( \frac{x}{2} \) yields \( x = \sqrt{2 \log(2/\delta)y} \) and \( x' = \sqrt{2 \log(2/\delta)} C_{T_1} \). Finally, plug in \( y = C_{T_1} \).

The choice of \( C_{T_1} = 8(C_1 \gamma T + 1) \) yields the same result as Theorem 1 of Contal et al. (2014). However, we resort to a tighter bound by realizing the following,

\[ s_{t-1}(v) := \sum_{v} C_{t-1}(v, v') \sqrt{1/(\sigma_{t-1}(v)^2 + \sigma^2)} \]

To derive the desired bound from (24), we observe that

\[ s_{t-1}(v) \geq \frac{\sigma_{t-1}(v)^2}{\sqrt{1/(\sigma_{t-1}(v)^2 + \sigma^2)}} \geq \frac{\sigma_{t-1}(v)}{\sigma_{\max}} \]

because \( \sigma_{\max} \leq \sigma_0(v)^2 \leq 1/\omega_0 \).

Using the Cauchy-Schwarz inequality,

\[ \sum_{t=1}^{T} s_{t-1}(v_t) \leq \left( T \sum_{t=1}^{T} s_{t-1}(v_t)^2 \right)^{1/2} \]

\[ \leq \max \sqrt{T} 1/(C_T - C_0) 1 \leq \sqrt{T \Delta T}. \]

This can help us decide a different choice of \( C_{T_1} \), as

\[ C_{T_1} = \sigma_{\max} \sqrt{T \Delta T} \]

\[ \geq \sigma_{\max} \sum_{t=1}^{T} s_{t-1}(v) \geq \sum_{t=1}^{T} \sigma_{t-1}^2(v) \]

Now we find \( \alpha_t \) that guarantees,

\[ -\sum_{t=1}^{T} \alpha_t s_{t-1}(v_t) + \sqrt{2 \log(2/\delta)} \sum_{t=1}^{T} \sigma_{t-1}^2(v^*) \sqrt{T \Delta T} \]

as, \( \alpha_t = \sqrt{2 \log(2/\delta)} \sigma_{\max} \sqrt{T \Delta T} \) which simplifies the r.h.s. of (24),

\[ \sqrt{2 \log(2/\delta)} \sum_{t=1}^{T} s_{t-1}(v_t) + \sqrt{2 \log(2/\delta)} \sigma_{\max} \sqrt{T \Delta T}. \]
References


Ming Ji and Jiawei Han. A variance minimization criterion to active learning on graphs. In *AISTAT*, 2012.


Yifei Ma, Roman Garnett, and Jeff Schneider. Σ-optimality in active learning on Gaussian random fields. In *NIPS*, 2013.


