An Improved Algorithm for Optimal Subset Selection in Chain Graphical Models

Qi Qi, *Student Member, IEEE,* Yi Shang, *Senior Member, IEEE,* and Hongchi Shi, *Senior Member, IEEE*

***Abstract*—The VoIDP algorithm is the first optimal algorithm for efficiently selecting the subset of observations in chain graphical models [10]. The original VoIDP algorithm has a mistake in the process of recovering the optimal selections, and fails to produce correct outputs. In this paper, we present an improved version of the algorithm; which fixes the mistakes and verifies the solutions in experiments. Further more, we discuss some recent works in the area of subset selection problems, and present a simplified solution for computing the maximum expected total reward for a sub chain under certain circumstances.**

I. Introduction

A typical problem in real world applications is the optimization of information gathering. Wireless sensor networks, for example, is a powerful tool for monitoring spatio-temporal phenomena. However, its limited power source makes sensing expensive. It is a trade off between obtaining more and useful information, versus making less observations. Scheduling a sensor to turn on to observe, then to turn off to save energy is a very big optimization problem.

Chain graphical models such as Hidden Markov Models (HMM) can be trained using data time series from sensors. The observation variables of 24 time points roll over onto the chain, if each hour in a day is treated as a time point for observation. When a selection is made at a time point for observation, the distributions of observation variables after this point will become certain to some extent. The sensor scheduling problem then turns into optimizing a subset of observations. The selection in the chain graphical model is to minimize the uncertainty overall. The VoIDP algorithm is the first optimal algorithm for efficiently selecting a subset of observations in chain graphical models [10]. It is a dynamic programming approach to optimize the value of information. However, during our evaluation of this algorithm for the subset selection problem in chain graphical models, we found that following the exact algorithm could not give desirable solutions. We were hence motivated to improve on it.

We identified a critical overlook in the original VoIDP algorithm; which causes the failure. We will present the improved version of VoIDP algorithm in Section III. In Section IV, we evaluated and verified the improved VoIDP algorithm and its solutions empirically. In Section V, we discuss a situation where the computation in the algorithm can be simplified. We will give a brief review in Section VI regarding some interesting works recently published in the area of optimizing the information gathering. We will start in the following Section to give a brief description of the optimization problem. For convenience, same notation will be used as does in [10].

Qi Qi and Yi Shang are with the Department of Computer Science, University of Missouri, Columbia, MO 65211 (emails: qiqi@mail.mizzou.edu, shangy@missouri.edu).

Hongchi Shi is with the Department of Computer Science, Texas State University-San Marcos, San Marcos, Texas 78666 (email: hs15@txstate.edu).

II. Problem Statement

The Algorithm 1 in [10] implements a dynamical programming approach to efficiently select an optimal subset of the variables to observe in chain graphical models. The algorithm is named as "VoIDP" because of its use of Dynamic Programming to optimize the information value. In a chain graphical model such as HMM, each observation variable has a distribution over the hidden states. When a variable is observed, the value will affect the distributions of the following variables along the chain. As more observations are obtained, the distributions over all the variables on the chain will become more certain. Also, the problem of optimizing the selections of observations across the entire chain can be cast in the following subset selection problem...

Suppose in a collection of random variables $\mathcal{X}_{\mathcal{V}} = (\mathcal{X}_1, \ldots, \mathcal{X}_n)$. A subset of the variables, $\mathcal{X}_{\mathcal{A}} = (\mathcal{X}_{i_1}, \ldots, \mathcal{X}_{i_k})$ is observed as $\mathrm{x}_{\mathcal{A}}$. The posterior distribution over all variables $P(\mathcal{X}_{\mathcal{V}} \mid \mathcal{X}_{\mathcal{A}} = \mathrm{x}_{\mathcal{A}})$ can be computed and transformed to a total reward $R(P(\mathcal{X}_{\mathcal{V}} \mid \mathcal{X}_{\mathcal{A}} = \mathrm{x}_{\mathcal{A}}))$. Since the future observations of $\mathcal{X}_{\mathcal{A}}$ are unknown, the expected total reward is there for used to measure the quality of the subset selection.

Hence, the subset selection problem is to select a subset $\mathcal{A}^* \subseteq \mathcal{V}$ which maximizes,

$$\mathcal{A}^* = \mathrm{argmax}_{\mathcal{A}} \sum_{\mathrm{x}_{\mathcal{A}}} P(\mathcal{X}_{\mathcal{A}} = \mathrm{x}_{\mathcal{A}}) R(P(\mathcal{X}_{\mathcal{V}} \mid \mathcal{X}_{\mathcal{A}} = \mathrm{x}_{\mathcal{A}})) \ [10]$$

The expected total reward $R(\mathcal{A})$ is also the sum of all the expected local rewards, where an expected local reward $R_j(\mathcal{X}_j \mid \mathcal{X}_{\mathcal{A}})$ equals to $\sum_{\mathrm{x}_{\mathcal{A}}} P(\mathcal{A} = \mathrm{x}_{\mathcal{A}}) R_j(\mathcal{X}_j \mid \mathrm{x}_{\mathcal{A}})$. A local reward $R_j(\mathcal{X}_j \mid \mathrm{x}_{\mathcal{A}})$ depends on $P(\mathcal{X}_j \mid \mathcal{X}_{\mathcal{A}} = \mathrm{x}_{\mathcal{A}})$, which is the marginal distribution of variable $\mathcal{X}_j$ conditioned on observations $\mathcal{X}_{\mathcal{A}} = \mathrm{x}_{\mathcal{A}}$. The authors in [10] choose residual entropy to measure the uncertainty of the marginal distribution of variable $\mathcal{X}_j$, and the objective of the optimization problem thus becomes one as how to minimize the total residual entropy, which is equivalent to maximize the expected total reward.

The probabilistic inference techniques in chain graphical models benefit the evaluation of the local rewards. For example, in Section IV a HMM based on sensors' temperature time series has n time steps. The observation on each time step is determined by a certain number of hidden states, and the hidden variables $\mathcal{X}_i$ form a chain conditional on the

978-1-4244-8126-2/10/$26.00 ©2010 IEEE

Authorized licensed use limited to: Hainan University. Downloaded on October 10,2021 at 01:24:41 UTC from IEEE Xplore. Restrictions apply.

observations. To evaluate the marginal distribution of $\mathcal{X}_j$, if only the observations made no later than the time step i are considered, it is referred to as the *filtering* case. Otherwise, if observations made anywhere along the entire chain can be taken into account, this situation is referred to as the *smoothing* case. The conditional independence property of the graphical model simplifies the evaluation of $P(\mathcal{X}_j \mid \mathcal{X}_{\mathcal{A}})$, which also implies that the expected total reward $R(\mathcal{A})$ for the entire chain can be decomposed into the expected rewards for sub chains [10].

In papers [7], [10], besides the subset selection problem briefly described above, the authors also addressed the *conditional planning problem*, which is not the focus of this paper.

III. Improved Efficient Algorithm for Optimal Subset Selection in Chain Graphical Models

In this section, we will first give a brief description of the original VoIDP algorithm as appearing in papers [10], [7], then, discuss the reason for and present the improved version of this algorithm.

A. Original VoIDP Algorithm

In the subset selection problem, the target is to decide a subset of the variables to observe before any observation is made in order to predict the overall observation most accurately based on the observed values of the selected variables. In the running example, before a sensor is deployed, we want to find a number of time points out of 24 to pre-schedule its sensing for a day.

The authors in paper [10] developed VoIDP, the first optimal algorithm for efficient subset selection in chain graphical models in both filtering and smoothing cases. For convenience, we have attached its pseudo code shown in figure 1. The algorithm implements a dynamic programming approach that is inspired by the reward decomposition property briefly discussed in Section II. It also considers some other factors in the subset selection process, such as operating within a limit budget B, the cost β_j of making observations, and associated penalties C_j applied to the expected total reward. The authors proved the time complexity of this algorithm given budget B in terms of evaluations of expected local rewards is $(\frac{1}{6}n^3 + O(n^2))B$.

The main body of the algorithm is to compute a number of tables of $L_{a:b}(k)$ which is denoted as the optimal expected total reward that can be achieved for the sub-chain $a:b$ with the budget k. And $L_{0:n+1}(B)$, therefore, denotes the optimal expected total reward for the entire chain with full budget B, while $L_{a:b}(0)$ is the total reward without any additional observations. The $\Lambda_{a:b}(k)$ stores the choice that realizes $L_{a:b}(k)$. The choices could be either the index of next variable to select or -1, which means no variable should be selected. In the innermost loop, $sel(j)$ is the expected total reward for the sub chain $a:b$ obtained by observing at j, and $sel(-1)$ is the reward if no observation is made. The optimal solution of subset selection is obtained by tracing the quantities in $\Lambda_{a:b}(k)$.

When we evaluated the algorithm, however, it failed to give desirable outputs (see table I) in the experiment. And we were thus motivated to improve on it. Following is an improved version of the VoIDP algorithm.

Input: Budget B, rewards R_j, costs β_j and penalties C_j
Output: Optimal selection $\mathcal{A}$ of observation times
begin
  for $0 \le a < b \le n+1$ **do** compute $L_{a:b}(0)$; **end**
  for $k = 1$ **to** B **do**
    for $0 \le a < b \le n+1$ **do**
      $sel(-1) := L_{a:b}(0)$;
      for $j = a+1$ **to** $b-1$ **do** $sel(j) := R_j(\mathcal{X}_j|\mathcal{X}_j)$
        $-C_j + L_{a:j}(0) + L_{j:b}(k-\beta_j)$; **end**
      $L_{a:b}(k) = \max_{j\in\{a+1,\ldots,b-1,-1\}} sel(j)$;
      $\Lambda_{a:b}(k) = \mathrm{argmax}_{j\in\{a+1,\ldots,b-1,-1\}} sel(j)$;
    end
  end
  $a := 0$; $b := n+1$; $k := B$; $\mathcal{A} := \emptyset$;
  while $j \neq -1$ **do**
      $j := \Lambda_{a:b}(k)$;
      if $j \ge 0$ **then** $\mathcal{A} := \mathcal{A} \cup \{j\}$; $k := k - \beta_j$; **end**
      end
  end
end

Fig. 1. VoIDP algorithm for optimal subset selection (Krause and Guestrin, [10])

B. Improved VoIDP Algorithm

The core part of VoIDP algorithm is to recursively compute the optimal expected total reward $L_{a:b}(k)$ for the sub chain $a:b$ using the budget k. The base case is simply $L_{a:b}(0)$, and the recursion for $L_{a:b}(k)$ is either $L_{a:b}(0)$ or $\max_{a<j<b,\beta_j\le k}\{sel(j)\}$. It means that we can choose not to spend any more of the budget to reach the base case, or we can select the optimal observation at j, which depends on the obtained expected total rewards. In our experiments, we let reward penalty C_j be zero, and let selection cost β_j be one. In this situation, the computation of $L_{a:b}(k)$ can actually be further simplified. We will discuss this in more details in Section V.

According to the reward decomposing property (see in [10]), selecting an observation will divide the computation of expected total reward of the chain into expected total reward computations along the two sub chains separated by the observation. This is reflected in the equation (1) of computing $sel(j)$ which is the expected total reward for chain $a:b$ when making an observation at j.

$$sel(j) := R_j(X_j \mid X_j) - C_j + L_{a:j}(0) + L_{j:b}(k-\beta_j) \qquad (1)$$

The total reward of observing at j for the chain $a:b$ is the sum of the reward of observing j at itself, the optimal total reward achieved for the sub-chain $a:j$ without any spending and the optimal total reward achieved for the sub-chain $j:b$

Authorized licensed use limited to: Hainan University. Downloaded on October 10,2021 at 01:24:41 UTC from IEEE Xplore. Restrictions apply.

with the budget $k-\beta_j$ and minus the reward penalty C_j. In this way, all the selected observations will fall into one side $j:b$, and on the other side $a:j$ there is no selection yet. The variable selection candidate j separates these into two sides. After all the tables $L_{a:b}(k)$ and $\Lambda_{a:b}(k)$ are computed, we can trace back to find all the optimal selections in $\Lambda_{a:b}(k)$. One key point here is that after we find an optimal selection at j, the entry for locating the next optimal selection should be at $\Lambda_{j:b}(k-\beta_j)$. The pseudo code of the improved VoIDP algorithm is illustrated in figure 2. For convenience, we use the same notations as in figure 1.

```
Input : Budget B, rewards R_j, costs β_j and penalties C_j
Output : Optimal selection A of observation times
begin
  for 0 ≤ a < b ≤ n+1 do compute L_{a:b}(0); end
  for k = 1 to B do
     for 0 ≤ a < b ≤ n+1 do
        sel(−1) := L_{a:b}(0);
        for j = a+1 to b−1 do
           sel(j) := R_j(X_j|X_j) − C_j + L_{a:j}(0)
              +L_{j:b}(k − β_j);
        end
        L_{a:b}(k) = max_{j∈{a+1,...,b−1,−1}} sel(j);
        Λ_{a:b}(k) = argmax_{j∈{a+1,...,b−1,−1}} sel(j);
     end
  end
  a := 0; b := n+1; k := B; A := ∅;
  while k > 0 do
      j := Λ_{a:b}(k);
      if j ≥ 0
         A := A ∪ {j};
         a := j;
         k := k − β_j;
      end
      if j < 0
         break;
      end
  end
end
```

Fig. 2. Improved VoIDP algorithm for optimal subset selection

In the improved version, the main change is reflected in the second part. After $L_{a:b}(k)$ and $\Lambda_{a:b}(k)$ are all computed, we need to find out the optimal selection from the series of $\Lambda_{a:b}(k)$ tables. Initially, it will start from the entire chain with the full budget B and an empty selection set $\mathcal{A}$. The first selection will thus be $\Lambda_{0:n+1}(B)$. If it is set to j, then the next selection should be from $\Lambda_{j:b}(k-\beta_j)$ instead of $\Lambda_{0:b}(k-\beta_j)$ as in the original VoIDP (see figure 1). This slight change, however, leads to a dramatically improved outputs (see table I). As discussed in the last paragraph, the way of tracing back the optimal selections is actually determined by how they were calculated. With this crucial change in the process of recovering optimal selections, the improved version of VoIDP produces desirable outputs. We have verified the effectiveness of those optimal selections through experiments. The results will be presented in the next section.

IV. Experiments

In this section, we first compare the selection outputs of the original VoIDP algorithm with those of our improved version. Then, we evaluate the optimal selections by the improved VoIDP against the ones generated by the greedy heuristic and uniform spacing methods. We use the temperature time series data set, which was also used in paper [10]. The data set was collected from a network of wireless sensors deployed in the Intel Berkeley Research Lab [1].

A. Optimal Observation Selection

One of the research problems in wireless sensor networks is that how a sensor should be scheduled for sensing in order to both save its power and, in the meantime, obtain the most informative observation possible. The goal of the optimal subset selection, for example, in this running setting is to select k out of 24 time points of a day for a sensor to monitor the indoor temperature, of which the readings will be the most informative.

We followed the experimental setting for temperature time series in paper [10], though the data set we used would not be from the exactly same sensors as they chose. All the data were pre-processed for missing samples and discretized into 10 bins of 2 degrees Kelvin. We got 45 sample time series combined from the data collected by three adjacent sensors (#3, #4, and #6) lasting 19 days. We used them to train a HMM that also had four latent states representing from 12 am - 7 am, 7 am - 12 pm, 12 pm - 7 pm and 7 pm - 12 am. All the input rewards used both in the original VoIDP and the improved version were computed from this trained chain graphical model under the filtering case, with assumed unit cost and zero reward penalty for making any observations.

Table I shows the comparison of the outputs supposed to be the optimal selections of observational time points from both algorithms and their relevant rewards. The higher reward and the better quality of selection of the improved algorithms are quite evident. As can be seen from the table, the original VoIDP algorithm repeatedly selects the first time point after budget 5, which is apparently a waste of the budget. We will further evaluate the solutions given by the improved VoIDP algorithm.

B. Performance Comparison

Since the original VoIDP algorithm does not produce satisfying outputs, we will not evaluate it in the following experiments. To examine the performance of the improved VoIDP algorithm, we also use the greedy heuristic and uniform spacing methods for comparison in our experiments.

Basically the greedy method makes a selection each time greedily from one of the sub-chains divided by the previously

Authorized licensed use limited to: Hainan University. Downloaded on October 10,2021 at 01:24:41 UTC from IEEE Xplore. Restrictions apply.

TABLE I

OPTIMAL OBSERVATION SELECTIONS BY ORIGINAL VOIDP AND THE IMPROVED VOIDP (IN THIS EXAMPLE WE LET UNIT COST AND ZERO PENALTY WHEN SELECTING ANY OBSERVATIONS).

Budget	Optimal Observation Outputs (time point ranges from 1 to 24)			
	Original VoIDP		Improved VoIDP	
	outputs	reward	reward	outputs
1	6	−32.2979	−32.2979	6
2	5, 6	−31.2704	−29.1210	5, 14
3	4, 5, 6	−30.3201	−26.8453	4, 10, 17
4	3, 4, 5, 6	−29.4919	−25.1229	3, 7, 12, 18
5	1, 3, 4, 5, 6	−28.5762	−23.7130	1, 5, 9, 14, 19
6	1, 1, 3, 4, 5, 6	−28.5762	−22.5522	1, 5, 8, 12, 16, 20
7	1, 1, 1, 3, 4, 5, 6	−28.5762	−21.6115	1, 5, 8, 11, 14, 17, 21
8	1, 1, 1, 1, 3, 4, 5, 6	−28.5762	−20.8333	1, 4, 6, 9, 12, 15, 18, 21

selected time points. For example, assuming unit selection cost, if $k=1$, then the best observation will be selected by greedy heuristic across the entire chain in terms of expected total reward. If $k=2$, it will pick a best observation respectively for each of the sub chains divided by the optimal observation selected in the $k=1$ case. It then will choose the second optimal observation among the two observations. This process can be deduced into the $k=n$ case. The uniform spacing heuristic evenly distributes the k observations across the entire chain.

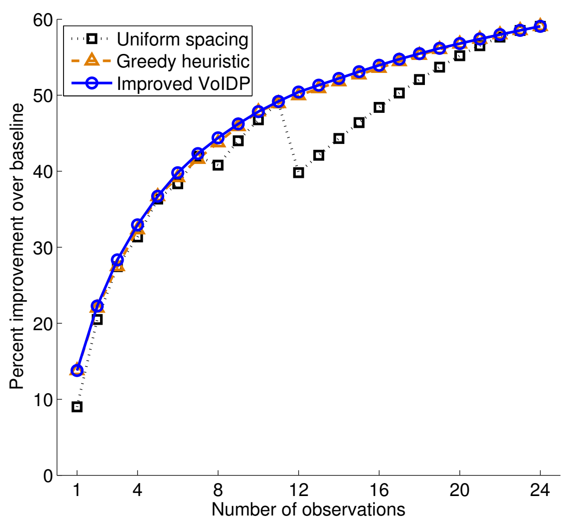

Fig. 3. Baseline performance comparison: The relative improvement of the uniform spacing method, the greedy heuristic, and the improved VoIDP algorithm over the baseline reward which is the expected total reward for the entire chain without any observations.

In figure 3, all the performance results are compared against the baseline, which is the expected total reward for the entire chain without any observation. The performance is measured as an increase of the expected total reward, which is equivalent to decrease of expected entropy for the entire chain. It shows that the optimal selections given by the improved VoIDP algorithm outperform those by both the heuristics. To give a better picture of how much the improved VoIDP algorithm outperforms the greedy heuristic, we then compare their relative improvements against the uniform spacing. The result is illustrated by figure 4. Here, performance is measured as an increase of expected total reward, with the uniform spacing as the baseline.

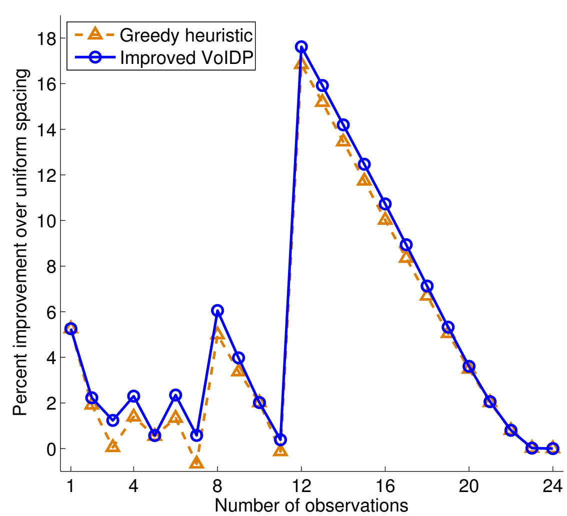

Fig. 4. Relative performance comparison between the greedy heuristic and the improved VoIDP: The relative improvement of the greedy heuristic and the improved VoIDP algorithm over the performance of the uniform spacing method.

As shown in figure 4, the difference of performance improvement between the optimal selections given by the improved VoIDP algorithm and those by greedy heuristic is obvious, when fewer number of observations are selected. It can be seen that if k is less than one third of all possible observations, the optimal gain by the improved VoIDP algorithm is more than one percent over that by the greedy heuristic. And the gain remains even when k reaches about two thirds of all possible observations. After that, the optimal subset

Authorized licensed use limited to: Hainan University. Downloaded on October 10,2021 at 01:24:41 UTC from IEEE Xplore. Restrictions apply.

and the subset selected by the greedy heuristic are almost identical. These results empirically verify the effectiveness of the optimal selections produced by the improved VoIDP algorithm.

V. Discussion

The VoIDP algorithm was claimed in [10] as the first optimal algorithm for nonmyopically computing and optimizing value of information in chain graphical models. This algorithm appears at least in [7], [10], [5] and remains in the same form. We evaluated the algorithm for subset selection problem as appearing in [10] and found the issue illustrated by table I. We think it necessary to improve on the algorithm.

The computation of $L_{a:b}(k)$ (see notations in III-A) in the improved VoIDP algorithm can be further simplified. If there was no penalty towards the total reward of making any observations, which means $C_j = 0$, then $sel(j) := R_j(X_j \mid X_j) + L_{a:j}(0) + L_{j:b}(k - \beta_j)$. The computation of $L_{a:b}(k)$ would become the following, because in this case any $sel(j)$ would be bigger than $L_{a:b}(0)$.

$$L_{a:b}(k) = \max_{j:a<j<b,\beta_j \leq k} \{R_j(X_j \mid X_j) + L_{a:j}(0) + L_{j:b}(k - \beta_j)\} \tag{2}$$

In other words, if there were no reward penalties, then the expected total reward for sub chain $a:b$ after making any additional observations would always be larger than that of making no additional observations. In this situation, the algorithm (see figure 2) does not need to compute $sel(-1)$ in the inner loop, and hence $L_{a:b}(k)$ does not need to compare with $L_{a:b}(0)$.

VI. Related Work

The original VoIDP algorithm as an efficient tool for selecting observation to maximize the value of information in chain graphical models was first introduced in [7], and was further presented in [10]. Although the optimization problem can be effectively solved for chain graphical models, it is much harder for more general graphical models. The authors proved that the problem of subset selection for even discrete polytree is computational intractable. There are some other approaches suggested for selecting observations in graphical models, but the authors of [10] argued that either some of them, such as the greedy methods, do not have theoretical performance guarantees, or others are running in exponential worst-case time, although they could be applied to more general graphical models. Besides developing algorithms to schedule a single sensor, the authors in [7], [10] also studied scheduling multiple sensors whose measurements are correlated, in which case the graphical model becomes more general, consisting of multiple chains.

The optimization problem of selectively gathering information with a variety of objectives exists in many tasks of real world applications. When, for example, deploying a wireless sensor network to monitor a spatio-temporal phenomenon, we want to choose locations and time points to deploy and schedule the sensors in order to maximize the information gains, and in the meantime minimize its communication costs. A doctor may want to have a most effective diagnostic plan designed at a minimal cost for a patient. Nowadays, the Internet provides a vast amount of information, but people would like to spend a small amount of time to read the most important news or useful information. Several efficient algorithms have been developed to address such problems.

In spatial monitoring such as in [15], [19], the sensor placement problem can be modeled using Gaussian Processes with a mutual information criterion, which is a submodular function. Submodularity, an intuitive diminishing returns property, can be exploited to develop faster, strongly polynomial time combinatorial algorithms with provable theoretical performance guarantees ([20], [4], [6], [13]). It turns out that many observation selection problems ([16], [8], [11], [18], [3]) can utilize this important structural property to develop efficient and near optimal algorithms incorporating greedy heuristic. However, in a more complex setting where another criterion besides the informativeness needs to be considered such as communication cost, greedy algorithms perform arbitrarily badly [12]. The authors of [12] presented a nonmyopic algorithm *pSPIEL* which can near-optimally trade off between information and communication cost. Another nonmyopic algorithm *Saturate* [14] was designed to minimize the uncertainty that could be exploited by adversaries. In [21], [17], it is shown that submodular functions are applicable to optimization of informative paths for multiple robots. Submodular functions have inspired researchers not only to develop efficient algorithms but also to study theoretical foundation of solving complex combinatorial problems. The authors in [2] introduced an algorithmic framework for studying combinatorial problems with multi-agent submodular cost functions and presented an approximate algorithm with theoretic lower bound.

There is another alternative approach to selection problems. Other than choosing according to a model (open-loop) before any observations are obtained, sequential planning (closed-loop) decides on the next selection based on previously observed values. In paper [9], the authors compared a sequential algorithm sequentially optimizing mutual information in Gaussian Processes with the model based selection approach, and quantified the advantage of the sequential strategy. A conditional planning based algorithm for selecting observations in chain graphical models was also presented in [7], [10].

VII. Conclusions

We have presented an improved version of VoIDP algorithm and discussed the reason for doing so. It is a slight-change but significant improvement to the original VoIDP algorithm, because it is critical for producing the desired optimal selection. The performance of the improved algorithm has been empirically verified. We also discussed a case when there are no reward penalties, the expected total reward of making any observations will always be larger than that of making no observations. This is used to simplify the computation of

Authorized licensed use limited to: Hainan University. Downloaded on October 10,2021 at 01:24:41 UTC from IEEE Xplore. Restrictions apply.

the optimal expected total reward for a sub chain. According to the review of recent works, submodular functions have spawned great interest in selection optimization in the past. Optimizing monotonic submodular functions strongly speeds up combinatorial algorithms. We are currently focusing our research in this direction.

Acknowledgement

We would like to thank Yi Zhang for providing valuable inputs in discussion of probabilistic inference in Hidden Markov Models.

References

[1] Amol Deshpande, Carlos Guestrin, Samuel R. Madden, Joseph M. Hellerstein, and Wei Hong. Model-driven data acquisition in sensor networks. In *VLDB '04: Proceedings of the Thirtieth international conference on Very large data bases*, pages 588–599. VLDB Endowment, 2004.

[2] Pushkar Tripathi Gagan Goel, Chinmay Karande and Lei Wang. Approximability of combinatorial problems with multi-agent submodular cost functions. In *FOCS '09: 50th Annual IEEE Symposium on Foundations of Computer Science*. IEEE, 2009.

[3] Michel X. Goemans, Nicholas J. A. Harvey, Satoru Iwata, and Vahab Mirrokni. Approximating submodular functions everywhere. In *SODA '09: Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 535–544, Philadelphia, PA, USA, 2009. Society for Industrial and Applied Mathematics.

[4] Satoru Iwata, Lisa Fleischer, and Satoru Fujishige. A combinatorial strongly polynomial algorithm for minimizing submodular functions. *J. ACM*, 48(4):761–777, 2001.

[5] Andreas Krause. *Optimizing sensing: theory and applications*. PhD thesis, Carnegie Mellon University, 2008.

[6] Andreas Krause and Carlos Guestrin. Near-optimal nonmyopic value of information in graphical models. In *Proceedings of the Proceedings of the Twenty-First Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-05)*, pages 324–331, Arlington, Virginia, 2005. AUAI Press.

[7] Andreas Krause and Carlos Guestrin. Optimal nonmyopic value of information in graphical models: efficient algorithms and theoretical limits. In *IJCAI'05: Proceedings of the 19th international joint conference on Artificial intelligence*, pages 1339–1345, San Francisco, CA, USA, 2005. Morgan Kaufmann Publishers Inc.

[8] Andreas Krause and Carlos Guestrin. Near-optimal observation selection using submodular functions. In *AAAI'07: Proceedings of the 22nd national conference on Artificial intelligence*, pages 1650–1654. AAAI Press, 2007.

[9] Andreas Krause and Carlos Guestrin. Nonmyopic active learning of gaussian processes: an exploration-exploitation approach. In *ICML '07: Proceedings of the 24th international conference on Machine learning*, pages 449–456, New York, NY, USA, 2007. ACM.

[10] Andreas Krause and Carlos Guestrin. Optimal value of information in graphical models. *J. Artif. Int. Res.*, 35(1):557–591, 2009.

[11] Andreas Krause and Carlos Guestrin. Optimizing sensing: From water to the web. *Computer*, 42(8):38–45, 2009.

[12] Andreas Krause, Carlos Guestrin, Anupam Gupta, and Jon Kleinberg. Near-optimal sensor placements: maximizing information while minimizing communication cost. In *IPSN '06: Proceedings of the 5th international conference on Information processing in sensor networks*, pages 2–10, New York, NY, USA, 2006. ACM.

[13] Andreas Krause, B. McMahan, Carlos Guestrin, and A. Gupta. Robust submodular observation selection. *J. Mach. Learn. Res.*, 9:2761–2801, 2008.

[14] Andreas Krause, Brendan McMahan, Carlos Guestrin, and Anupam Gupta. Selecting observations against adversarial objectives. In J.C. Platt, D. Koller, Y. Singer, and S. Roweis, editors, *Advances in Neural Information Processing Systems 20*, pages 777–784. MIT Press, Cambridge, MA, 2008.

[15] Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *J. Mach. Learn. Res.*, 9:235–284, 2008.

[16] Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, and Natalie Glance. Cost-effective outbreak detection in networks. In *KDD '07: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 420–429, New York, NY, USA, 2007. ACM.

[17] Alexandra Meliou, Andreas Krause, Carlos Guestrin, and Joseph M. Hellerstein. Nonmyopic informative path planning in spatio-temporal models. In *AAAI'07: Proceedings of the 22nd national conference on Artificial intelligence*, pages 602–607. AAAI Press, 2007.

[18] Bilge Mutlu, Andreas Krause, Jodi Forlizzi, Carlos Guestrin, and Jessica Hodgins. Robust, low-cost, non-intrusive sensing and recognition of seated postures. In *UIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology*, pages 149–158, New York, NY, USA, 2007. ACM.

[19] Zhuang Peng, Qi Qi, Shang Yi, and Shi Hongchi. Model-based traffic prediction using sensor networks. In *Consumer Communications and Networking Conference, 2008. CCNC 2008. 5th IEEE*, pages 136–140, 2008.

[20] Alexander Schrijver. A combinatorial algorithm minimizing submodular functions in strongly polynomial time. *J. Comb. Theory Ser. B*, 80(2):346–355, 2000.

[21] Amarjeet Singh, Andreas Krause, Carlos Guestrin, William Kaiser, and Maxim Batalin. Efficient planning of informative paths for multiple robots. In *IJCAI'07: Proceedings of the 20th international joint conference on Artifical intelligence*, pages 2204–2211, San Francisco, CA, USA, 2007. Morgan Kaufmann Publishers Inc.

Authorized licensed use limited to: Hainan University. Downloaded on October 10,2021 at 01:24:41 UTC from IEEE Xplore. Restrictions apply.