SAND2020-4877C

Routing of an Unmanned Vehicle for Classification

Christopher Montez[a], Swaroop Darbha[a], Christopher Valicka[b], and Andrea Staid[b]

[a]Texas A & M University, 400 Bizzell St., College Station, USA

[b]Sandia National Laboratories, Albuquerque, USA

March 2020

ABSTRACT

Routing problems for unmanned vehicles are frequently encountered in civilian and military applications and have been studied extensively as a result. One such routing problem of interest is the problem on constructing a tour that maximizes the total information gained in the tour in an attempt to classify the points of interest that have been visited in the tour. The information gained at each point of interest was modeled using the Kullback-Leibler divergence (also referred to as mutual information) where the probability of correctly classifying the point of interest was taken to be time-dependent. A mixed-integer programming (MIP) formulation of this problem was constructed and two standard heuristics (a modified two-step greedy algorithm and a standard 2-OPT algorithm) were combined in an attempt to produce high quality solutions to this MIP. Simulations were run with various conditions for the nature of the information gain and position of the points of interest. It was found combining these two heuristics produced near-optimal solutions in nearly all of the trials for up to 10 points of interest.

Keywords: Routing, classification, 2-OPT, greedy, information gain, mutual information, heuristic

1. INTRODUCTION

Routing problems for unmanned vehicles are frequently encountered in civilian and military applications. Although routing problems such as the traveling salesman problem have been studied extensively and have heuristics that produce near-optimal solutions for many problem instances, there still remains difficulties in constructing optimal or near-optimal routes for more complex routing problems. One such routing problem of interest is the problem of constructing a tour that maximizes the total information gained in the tour in an attempt to classify the points of interest that have been visited in the tour. This problem can be naturally found in military applications where an unmanned vehicle is sent out into the field to examine various points of interest and classify them as targets or non-targets. In an emergency relief setting where an unmanned vehicle is deployed to various sites, this problem is encountered when determining if immediate aid at a site is needed. This problem is examined and a heuristic that produces high quality solutions in numerous problem instances is presented in this paper.

The structure of this paper is as follows: In Section 2, the mathematical formulation classification routing problem is presented. The information gained by the unmanned vehicle visiting a point of interest is quantified using the Kullback-Leibler divergence. The a MIP formulation of the problem is then presented. In Section 3, the heuristics used to produce high quality solutions are presented. The first heuristic is a modified two step greedy algorithm. The second heuristic is a standard 2-OPT algorithm that has been modified to require the unmanned vehicle start and return to a depot that has been specified in advance. These two heuristics are then applied in series to produce the recommended heuristic for this problem. In Section 4, computational results comparing the heuristics used are presented and Section 5 gives a brief summary of the paper.

2. PROBLEM FORMULATION

Consider a set of n points of interest (POIs) to be visited by a single vehicle to gain information about each POI. Let $G = (V, E)$ be an undirected graph where V is the set of POIs to be visited and E is the set of edges with weights corresponding to known non-negative traveling costs between the POIs. These costs need not be the Euclidean distances between POIs and may depend on other factors such as environmental considerations. It will be assumed the cost of traveling from t to u is the same as the cost to travel from u to t. The vehicle is to depart and return to a single depot, $d \in V$. When the vehicle visits a POI, information is gained about that

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

POI in order to aid its classification. The means of gaining this information by the vehicle is not considered in this paper. Each POI is to be classified as either T (target) or F (not a target) with the exception of the depot which will have no classification. It will assumed there is no *a priori* information about the POIs. It will also be assumed the probability of correctly classifying the POI is the same whether the POI has a true classification of T or F. That is, it is equally difficult to classify the POI, regardless of what it really is. The objective is then to construct a path in G that visits each POI once and maximizes the total information gained.

2.1 Quantifying the Information Gained

Suppose the k-th visit by the vehicle is to $i \in V$ at time t. Denote the set of classification choices as $C = \{T, F\}$. Each POI has a correct classification $X \in C$. The vehicle assigns a classification of $Z \in C$ to i after the visit. Let s_i represent the state of the POI, i, the vehicle (or operator) sees/measures upon visiting. Denote the conditional probabilities of correctly classifying i as T or F given the state s_i as

$$P_t(s_i) = P(Z = T \,|\, X = T,\, s_i) \text{ and} \tag{1}$$

$$P_f(s_i) = P(Z = F \,|\, X = F,\, s_i), \tag{2}$$

respectively. The information gained by visiting each POI will be quantified using the Kullback-Leibler divergence (also referred to as the mutual information or information gain). The mutual information for $i \in V$ between the two classification variables X and Z will be denoted as $I_i(X, Z)$. The mutual information is defined to be

$$I_i(X, Z) := H(X) - H(X|Z)\,, \tag{3}$$

where $H(X)$ and $H(X|Z)$ are the entropy and conditional entropy, respectively. From the definitions of $H(X)$ and $H(X|Z)$, we have

$$I_i(X, Z) = \sum_{x,z \in C} P(X = x, Z = z) \log \frac{P(X = x, Z = z)}{P(X = x)P(Z = z)}\,. \tag{4}$$

Denote the *a priori* probability a POI is a target, $P(X = T)$, as p. It can then be shown Equation (4) can be rewritten as

$$\begin{aligned} I_i(X, Z) = & \, pP_t(s) \log \left(\frac{P_t(s)}{pP_t(s) + (1-p)(1-P_f(s))} \right) \\ & + (1-p)(1-P_f(s)) \log \left(\frac{1-P_f(s)}{pP_t(s) + (1-p)(1-P_f(s))} \right) \\ & + p(1-P_t(s)) \log \left(\frac{1-P_t(s)}{p(1-P_t(s)) + (1-p)P_f(s)} \right) \\ & + (1-p)P_f(s) \log \left(\frac{P_f(s)}{p(1-P_t(s)) + (1-p)P_f(s)} \right). \end{aligned} \tag{5}$$

It will be assumed the *a priori* probability a POI is a target is 0.5. That is, there is effectively no known information about the POIs before sending out the vehicle to investigate and so each POI is equally likely to be either a target or not a target. Additionally, it will be assumed it is equally difficult to correctly classify a POI, i, as a target or not a target. That is, $P_t(s) = P_f(s) = P_i(s)$ for any state s_i. Then Equation (5) is reduced to

$$I_i(X, Z) = P_i(s) \log P_i(s) + (1 - P_i(s)) \log(1 - P_i(s)) + \log 2\,. \tag{6}$$

For computational purposes, $P_i(s)$ will be taken to have the form

$$P_i(s) = P_i(t) = 0.5(1 + e^{-t/\tau_i})\,, \tag{7}$$

where τ_i is a non-negative constant that represents the sensitivity to the time taken to visit i. This form was chosen due to the fact $P_i \to 0.5$ as $t \to \infty$. That is, if the vehicle takes a sufficiently long amount of time to visit i, then the vehicle (or operator) will have to effectively guess whether i is a target or not due to the lack of information available. Conversely, at $t = 0$ we have $P_i = 1$. That is, if the vehicle can instantaneously arrive to i before any information can be lost, then the classification of i is guaranteed to be correct. Additionally, the curvature of the classification probability can be modified by varying the value of τ_i. In this formulation of P_i, the only state considered is the time the vehicle collects information at the POI (which is taken to be the time of arrival in this paper) and so the information gained, I_i, also only depends on the time. In practice, the probability of correctly classifying the POI (and consequently the information gained) will depend on other state information (aspect angle, altitude, etc.) when collecting information. Experimental data collected before sending out the vehicle can be used to construct such a probability model.[1]

2.2 MIP Formulation

We now present a MIP formulation of the problem. Let the binary variable $x_v^{i,j}$ denote v-th visit of the vehicle from i to j, where $x_v^{i,j} = 1$ if the departs from i and arrives at j for the v-th visit and $x_v^{i,j} = 0$ otherwise. Let f_v^i denote the time elapsed at i at the v-th visit in the tour. Denote the cost of traveling from i to j in $G = (V, E)$ as $c(i, j)$. If an edge does not exist in G, it will be taken to have infinite cost. The number of nodes, $|V|$, will be denoted as N and the information gain at the depot will be taken to be identically zero. The MIP formulation is then the following:

$$\max \sum_{i \in V} \sum_{j \in V} I_j(f_v^j) x_v^{i,j} \tag{8}$$

$$\text{subject to} \quad \sum_{j \in V} x_1^{d,j} = 1, \tag{9}$$

$$\sum_{l \in V} x_N^{l,d} = 1, \tag{10}$$

$$x_v^{i,i} = 0, \quad \forall i \in V, \quad v = 1, 2, ..., N, \tag{11}$$

$$\sum_{i \in V} \sum_{j \in V} x_v^{i,j} = 1, \quad v = 1, 2, ..., N, \tag{12}$$

$$\sum_{l \in V} x_{v-1}^{l,j} = \sum_{j \in V} x_v^{i,j}, \quad \forall i \in V, \quad v = 2, 3, ..., N, \tag{13}$$

$$\sum_{v=1}^{N} \sum_{j \in V} x_v^{i,j} \geq 1, \quad \forall i \in V, \tag{14}$$

$$f_v^i - f_{v-1}^i = \sum_{i \in V} \sum_{j \in V} x_v^{i,j} c(i, j), \quad \forall i \in V, \quad v = 2, 3, ..., N, \tag{15}$$

$$f_1^i = \sum_{j \in V} x_1^{d,j} c(d, j) \quad \forall i \in V, \tag{16}$$

$$f_v^i \geq 0 \quad \forall i \in V, \quad v = 1, 2, ..., N, \tag{17}$$

$$x_v^{i,j} \in \{0, 1\}, \quad \forall i, j \in V, \quad v = 1, 2, ..., N. \tag{18}$$

The objective (8) is the sum of the information gain, $I_j(f_v^i)$, collected in the tour with f_v^i representing the time elapsed. This time elapsed is used in calculating the information gained through $P_i(s)$ in Equation (7). Constraints (9) through (14) restrict the path to a Hamiltonian circuit with the final visit in the tour going to the designated depot, $d \in V$. Constraints (15) through (17) detail how the time elapsed is calculated and restricts the time to non-negative values. In general, solving this MIP is very difficult and so a heuristic was used to find high quality solutions.

Algorithm 1 Modified Two-Step Greedy Algorithm in Words

Initialize partial tour as $T = \{1\}$ (where node 1 is the depot) and an unvisited node list as *unvisited*. Denote d_{ij} as time from i to j. Initialize time elapsed as $t = 0$.

```
while length(unvisited) > 1 do
  denote q last node in tour
  for all unvisited nodes j do
    for all unvisited nodes k ≠ j do
      Calculate potential information loss for taking q → k → j instead of q → j
      Store potential information loss as loss_k
    end for
    Find minimum potential reward loss for j
      minimum loss_j = min_k loss_k
    Calculate the penalty (information loss/time) for j
      penalty_j = minimum loss_j / d_qj
  end for
  Find largest penalty and set the corresponding node as j*
    maximum penalty = max_j {penalty_j} = penalty_j*
  Remove j* from unvisited and add to tour, T
  Add d(q, j*) to the time elapsed, t
end while
Add final remaining node to the partial tour, T
Return to depot by adding 1 to the end of T, completing the tour
```

3. HEURISTICS

In this section, two heuristics are presented and then combined to produce the heuristic recommended for this problem. The first heuristic is a modified two-step greedy (M2SG) algorithm. The second heuristic is a standard 2-OPT algorithm commonly used in literature that has been modified to start and end at a specified depot. The M2SG comes from work done by Erkut and Zhang[2] for constructing tours in a maximum collection problem with time dependent rewards. In their work, they presented a "penalty heuristic" to construct a relatively high quality tour. The standard 2-OPT algorithm was then applied to the constructed tour to create a higher quality tour. This work followed a similar procedure.

The M2SG algorithm is constructed as follows: Suppose the vehicle is at $q \in V$ in a partially completed tour of G and there are still several (at least two) POIs left to visit before completing the tour. Denote the set of unvisited POIs as $U \subset V$ (proper subset as the vehicle always starts at the depot) and consider $j \in U$ and $k \in U \setminus \{j\}$. The potential amount of information lost for taking the route $q \to k \to j$ instead of $q \to j$ can be computed directly using Equation (6). The potential information loss for all $k \in U \setminus \{j\}$ can be computed and the minimum potential information loss can be found. Divide this minimum potential information loss by the time to travel from q to j and denote this value as the penalty associated with j. The penalty associated with j is used to quantify the time rate of information loss associated with j. This process can then be repeated for all $j \in U$ and the largest penalty (rate of information loss) corresponding to $j^* \in U$ can be found. This POI is then considered to have the highest priority amongst all potential POIs left to be visited and so the vehicle visits this POI next. This process is repeated until there is only one POI remaining. This remaining POI is added to the end of the partial tour and the vehicle then returns to the depot, completing the tour. Algorithm 1 presents the M2SG algorithm using pseudocode.

Suppose the vehicle requires a significant amount of time to measure/gather information at a POI. Denote this time delay at $q \in V$ as $v(q)$. This time delay can be accounted for in the M2SG algorithm by adding this time in the calculation of the potential information loss and dividing by $d(q, j^*) + v(q)$ in the calculation of the penalty associated with j. This M2SG algorithm is flexible and can be adjusted for other forms of P_i in Equation (6). The only step that would be affected is the calculation of the potential information loss for taking $q \to k \to j$ instead of $q \to j$.

The second heuristic used is the standard 2-OPT algorithm, a local search heuristic that has been widely used for traveling salesman problem (TSP) applications. The 2-OPT algorithm starts with a given tour that can be arbitrarily chosen (so long as it is feasible) or found using another algorithm such as a nearest neighbor algorithm. The algorithm then looks at two non-intersection arcs in the tour and swaps them if the tour after swapping arcs has a smaller total distance. This process is repeated until swapping no longer reduces the tour's total distance. For this problem, the 2-OPT algorithm was slightly modified so that the tour always begins and ends at the depot, $d \in V$. This is done by restricting the candidate arcs to not include the depot POI. For this problem, the tour used to initialize the 2-OPT algorithm is the tour that has been constructed using the M2SG algorithm.

4. COMPUTATIONAL RESULTS

In this section, computational results for the heuristics in Section 3 are presented. First, the notation and MATLAB setup is presented. Three cases are considered for the setup: (1) relatively low time sensitivity, (2) relatively high time sensitivity, and (3) a mixture of time sensitivities. The relative scale for the time sensitivity depends on the ratio of τ_i and the cost matrix entries. That is, if the cost to travel from i to j is high when compared to the time sensitivity, τ_j, then the probability of correctly classifying the POI will be heavily affected. Thus, the exact numerical values of τ_i and the entries of the cost matrix are not important. Rather, their relative ratios are what dictates the sensitivity of the solution quality.

4.1 Notation and Setup

Denote the number of POIs considered in a problem instance as n. The number of POIs considered was varied from $n = 5$ to $n = 10$. For each n, 100 problem instances were generated. Each problem instance had a randomly generated cost matrix with entries ranging from 5 to 15. The choice of this range is arbitrary and will only have an affect on the choice of the range for τ_i. For the case of all POIs having low time sensitivity, each τ_i was randomly chosen from the interval $[100, 250]$. For the case of all POIs having high time sensitivity, each τ_i was randomly chosen from the interval $[30, 50]$. For the case of a mixture of time sensitivities among the POIs, each τ_i was randomly chosen from the interval $[50, 100]$. For a different range of cost matrix entries, these τ_i ranges will vary accordingly. For each n, four tours were constructed. The first tour was constructed using the M2SG algorithm. The second tour was constructed using the standard 2-OPT algorithm. The initial tour used for the 2-OPT algorithm for all instances was a tour connecting the POIs in sequential order and returning back to the depot ($1 \to 2 \to ... \to n \to 1$). The third tour was constructed by taking the tour constructed using the M2SG algorithm as the initial tour in the 2-OPT algorithm (referred to as M2SG + 2OPT). The final tour was constructed using brute force to find the optimal tour for n POIs. The total information gain for each tour was computed and compared to the optimal total information gain found using brute force. The results are presented in the following subsection.

4.2 Results

The average percentage of the optimal value for each heuristic method's solution was computed for each n after 100 randomly generated problem instances (see 4.1 for full details) and tabulated for the cases of low time sensitivity, high time sensitivity, and mixed time sensitivity. The optimal value for each problem instance was computed using brute force. Table 1 shows the results for the case of all POIs having low time sensitivity. It can be seen that on average the 2-OPT algorithm outperforms the M2SG algorithm. Additionally, applying the 2-OPT algorithm to the tour generated by the M2SG algorithm appears to increase the solution quality in some cases.

Next, Table 2 shows the results for the case of all POIs having relatively high time sensitivity. As was the case for low time sensitivity, the 2-OPT algorithm consistently outperformed the M2SG algorithm. However, once again in some instances applying the 2-OPT algorithm to the M2SG tour increased the quality of the tour solution in some instances.

Finally, Table 3 shows the results for the case where the POIs have relatively mixed time sensitivities. Once again, the 2-OPT algorithm outperformed the M2SG algorithm on average. However, applying the 2-OPT algorithm to the M2SG tour improved the quality of the solution more often on average than the previous cases.

Table 1. Average percent of the optimal solution of tour solutions and average computation times for the case of all POIs having relatively low time sensitivity ($100 \leq \tau_i \leq 250$) for 500 randomly generated problem instances.

	M2SG		**2-OPT**		**M2SG + 2-OPT**	
n	%	ms	%	ms	%	ms
5	82.65	0.097	98.99	0.082	100	0.174
6	77.62	0.136	98.58	0.129	100	0.266
7	73.84	0.181	98.28	0.190	93.61	0.374
8	69.98	0.248	97.95	0.278	99.17	0.519
9	66.77	0.325	97.53	0.378	99.97	0.690
10	63.57	0.399	97.03	0.506	99.58	0.896

Table 2. Average percent of the optimal solution of tour solutions and average computation times for the case of all POIs having relatively high time sensitivity ($30 \leq \tau_i \leq 50$) for 500 randomly generated problem instances.

	M2SG		**2-OPT**		**M2SG + 2-OPT**	
n	%	ms	%	ms	%	ms
5	58.27	0.102	97.36	0.088	89.59	0.184
6	51.25	0.145	96.20	0.135	100	0.277
7	45.52	0.203	94.59	0.209	96.47	0.412
8	42.09	0.262	93.99	0.281	97.42	0.543
9	39.13	0.328	92.89	0.374	85.85	0.697
10	36.84	0.404	91.73	0.512	89.93	0.898

Table 3. Average percent of the optimal solution of tour solutions and average computation times for the case of all POIs having relatively mixed time sensitivity ($50 \leq \tau_i \leq 100$) for 500 randomly generated problem instances.

	M2SG		**2-OPT**		**M2SG + 2-OPT**	
n	%	ms	%	ms	%	ms
5	71.06	0.098	98.38	0.084	100	0.177
6	64.28	0.140	97.72	0.127	100	0.269
7	58.69	0.197	96.80	0.202	100	0.398
8	54.33	0.251	96.07	0.280	99.95	0.522
9	51.10	0.327	95.93	0.380	92.48	0.695
10	48.25	0.399	95.10	0.496	93.66	0.887

5. CONCLUSION

The problem of routing an unmanned vehicle for classification of points of interest (POIs) was examined. The vehicle has a probability of correctly classifying each POI as either a target or not a target. The mutual information was then used to quantify the information gained by the vehicle visiting each POI at some time t. With this, the problem was formulated as an optimization problem where the vehicle is to visit each POI once and maximize the total information gained in the tour. A modified two-step greedy (M2SG) algorithm, a standard 2-OPT algorithm, and a combination of M2SG and 2-OPT were used to construct tours and were compared to the optimal solution obtained through brute force for several hundred problem instances. It was found the 2-OPT algorithm outperformed the M2SG algorithm on average. Additionally, it was found using the M2SG algorithm and inputting the newly constructed tour into the 2-OPT algorithm yielded nearly optimal solutions in many cases. This includes cases where the information gain at each POI is relatively sensitive to the time for the vehicle to reach the POIs. The computation time for applying both algorithms was found to be on the order of 1 millisecond for relatively small problem instances (less than 10 POIs).

ACKNOWLEDGMENTS

This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

REFERENCES

[1] Kalyanam, K., Pachter, M., Patzek, M., Rothwell, C., and Darbha, S., "Optimal human–machine teaming for a sequential inspection operation," *IEEE Transactions on Human-Machine Systems* **46**(4), 557–568 (2016).

[2] Erkut, E. and Zhang, J., "The maximum collection problem with time-dependent rewards," *Naval Research Logistics* **43**, 749–763 (1996).