UAV Routing for enhancing the performance of a classifier-in-the-loop

Suveer Kumar, Loy McGuire, Swaroop Darbha and Donald Sofge

June 14, 2023

Abstract

Some human-machine systems are designed so that machines (robots) gather and deliver to remotely located operators (humans) through an interface in order to aid them in classification. The performance of a human as a (binary) classifier-in-the-loop is characterized by probabilities of correctly classifying an object of type T as T and an object of type F as F. These two probabilities depend on (i) the dwell time, d, spent collecting information at a target. The information gain associated with collecting information at a target is then a function of dwell time d and discounted by the revisit time, R, i.e., the duration between consecutive revisits to the same target. The objective of the problem of routing for classification is to optimally route the vehicles and determine the optimal dwell time at each target so as to maximize the total discounted information gain while visiting every target at least once. In this paper, we make a simplifying assumption that the information gain is discounted exponentially by the revisit time; this assumption enables one to decouple the problem of routing with the problem of determining optimal dwell time at each target. We present an algorithm for solving this problem and provide corroborating numerical results.

1 Introduction

The design of human-machine systems requires a careful partition of the tasks to be performed by the human-in-the-loop and the machines, and the design of an associated human-machine interface. In this paper, we consider a human-machine system, where the human serves as a classifier-in-the-loop based on the information delivered to the human operator by the machines (vehicles) through an interface. The interface takes as input n Points of Interest (POIs) nominated by the human operator and computes the order in which they must be visited, the time to be spent at each POI along with the waypoints for the vehicles to follow. The vehicles persistently monitor the POIs by visiting them and dwelling at a POI while collecting information that is transmitted to the remotely located operator. Based on the information supplied by the vehicles, the primary task of the human is to assess/classify events happening at the n specified POIs and classify them as T (target) or F (false target). The probability of an operator correctly classifying events at a POI depends on the dwell time of the vehicle at that location and the revisit time, R, i.e., the time duration between consecutive revisits to the same location. A discounted information gain reflects the tradeoff between the information gained as a function of dwell time d and the revisit time, R. The problem considered in this paper is the determination of optimal sequence in which POIs must be visited and the dwell time at each POI that helps the operator maximize the sum of all discounted information gained from POIs, while ensuring that each of them is visited. While this problem is solved by the interface; in a centralized manner, distributed heuristics can be developed for this problem; however, it is not considered in this paper. Subsequent determination of waypoints for vehicles follows immediately and is communicated by the interface to the vehicles.

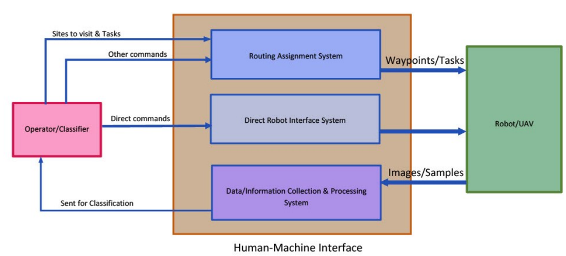

Figure 1: Human-Machine System with human as a classifier-in-the-loop

The novelty of this paper lies in proposing a model for accounting the dependence of UAV's dwell time on information gain and its subsequent use in their routing so as to enhance the operator's performance as a classifier-in-the-loop.

2 Literature Review

Make it concise

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, $d \in V$. At each POI, vehicle collects information and transmits it to a remotely located operator 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.

The topic of routing unmanned vehicles for autonomously collecting data has received significant attention in the literature [28, 10, 11, 9, 16, 19, 18, 17, 24, 23, 25, 5, 29]. The objective of routing depends on the mission, and the nature of the optimal solution depends further on the operational, motion, coordination and communication constraints. While the previous references addressed routing problems, none of them deal with routing vehicles to enhance the performance of an operator as a classifier-in-the-loop. Typically, the performance of an operator in the loop is specified by a confusion matrix, which specifies two conditional probability distributions – given that the event/object is of type $Z \in \{T, F\}$, the conditional probability distribution specifies the discrete probability distribution of correctly classifying the event/object[13]. These probability distributions depend on controllable operational parameters such as altitude, pose of the vehicle relative to the object and the time spent imaging the object/event, etc. The idea of vehicles enhancing the classification performance then rests on controlling or choosing these parameters to ensure that the conditional distributions are separated as far as possible, i.e, the mutual information gain is maximized. For persistent monitoring of the POIs, it is also necessary to minimize the time, R, between successive revisits to the POIs so that discounting the mutual gain by a function of the revisit time is reasonable and the corresponding objective for optimization is to maximize the discounted mutual information gain over all POIs through optimal routing and determination of optimal dwell time.

Information gain has been used in path planning in robotic applications that are distinct from what is considered in this paper. In simple terms, information gain [6] can be defined as the reduction in uncertainty of a single random variable due to another random variable. In this case for target classification, the probability that the POI is a target or not can be taken as the random variable for further analysis. The use of information gain theory in vehicle routing and motion planning has proved beneficial for several applications. Lee *et al*. [15] developed an enhanced ant colony optimization for the capacitated vehicle routing problem by using information gain to ameliorate the search performance when the good initial solution was provided by simulated annealing algorithm. Toit and Burdick [8] used the information gain theory in developing a partially closed-loop receding horizon control algorithm to solve the stochastic dynamic programming problem associated with dynamic uncertain environments (DUEs) robot motion planning. Kaufman *et al*. [14] presented a novel, accurate and computationally-efficient approach to predict map information gain for autonomous exploration where the robot motion is governed by a policy that maximises the map information gain within its set of pose candidates. Zaenker *et al*. [30] proposed a novel view motion planner for pepper plant monitoring while minimizing occlusions (a significant challenge in monitoring of large and complex structures), that builds a graph network of viable view poses and trajectories between nearby poses which is then searched by planner for graphs for view sequences with highest information gain. Paull *et al*. [26] used information gain approach in the objective function of sidescan sonars (SSS) and for complete coverage and reactive path planning of an autonomous underwater vehicle. Mostofi [21] proposed a communication-aware motion-planning strategy for unmanned autonomous vehicles, where each node considers the information gained through both its sensing and communication when deciding on its next move. They showed how each node can predict the information gained through its communications, by online learning of link quality measures and combining it with the information gained through its local sensing in order to assess the overall information gain.

Information gain finds its place in machine learning literature where it is being used for diverse feature ranking and feature selection techniques in order to discard irrelevant or redundant features from a given feature vector, thus reducing dimensionality of the feature space. Novakovic [22] applied Information Gain for the classification of sonar targets with C4.5 decision tree where the IG evaluation helped in increasing computational efficiency while improving classification accuracy by doing feature selection.

Information-theoretic methods have been used for computing heuristics for path-planning methods in autonomous robotic exploration where mutual information is calculated between the sensor's measurements and the explored map. Deng *etal.* [7] proposed a novel algorithm for the optimizing exploration paths of a robot to cover unknown 2D areas by creating a gradient-based path optimization method that tries to improve path's smoothness and information gain of uniformly sampled viewpoints along the path simultaneously. Julian *etal.* [12] proved that any controller tasked to maximise a mutual information reward function is eventually attracted to unexplored space which is derived from the geometric dependencies of the occupancy grid mapping algorithm and the monotonic properties of mutual information. Bai *etal.* [3] proposed a novel approach to predict mutual information using Bayesian optimisation for the purpose of exploring *apriori* unknown environments and producing a comprehensive occupancy map. They showed that information-based method provides not only computational efficiency and rapid map entropy reduction, but also robustness in comparison with competing approaches. Amigoni & Caglioti [2] presented a mapping system that builds geometric point-based maps of environments employing an information-based exploration strategy that determines the best observation positions by blending together expected gathered information (that is measured according to the expected *a posteriori* uncertainty of the map) and cost of reaching observation positions. Basilico & Amigoni [4] further extended this information-based exploration strategy for rescue and surveillance applications. In [1], the problem of routing a mobile agents for data aggregation in sensor networks is considered. Here, the main issue is the tradeoff between increasing information gain and power consumption among the source nodes that must be visited by the mobile agent and is accounted for in the cost of the edges.

The problem of routing vehicles for aiding an operator-in-the-loop for classification was first proposed by Montez [20]; however, this paper does not exploit the exponential discounting nature of mutual information gain to decouple the mixed-integer nonlinear program into a discrete optimization problem and a continuous optimization problem. This structure is exploited in this paper and an exact algorithm for a single vehicle routing is presented in this paper. In addition, extensions to the multiple vehicle case is presented with some heuristics along with the corroborating computational results.

3 Mathematical Formulation

3.1 Quantifying the Information Gained

Suppose a vehicle visits the i^{th} POI. Denote the set of classification choices as $C = \{T, F\}$. Each POI has a correct classification $X \in C$. The operator assigns a classification of $Z \in C$ to i^{th} POI after the visit. Let s_i represent the state of the i^{th} POI, 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 the i^{th} POI, 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) reduces to

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

If $P_i(s) = P_i(d_i)$, a function of d_i, then one can express mutual information gain I_i as an explicit function of the dwell time d_i. At this point, we observe the following properties about the mutual information gain function (see Figure 3.1):

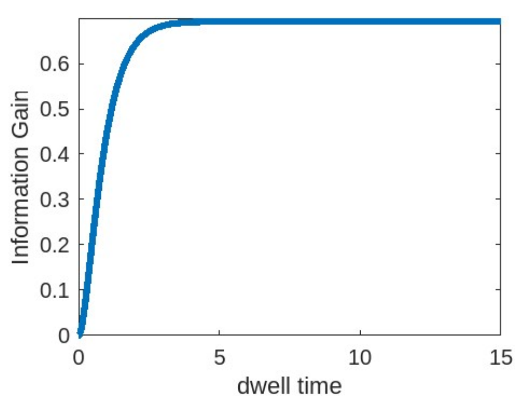

Figure 2: Information Gain vs Dwell Time at a POI ($\tau = 0.5$)

- The function $I_i(d_i)$ is monotonically increasing with d_i; essentially, the information gain increases with the time spent by a vehicle at the i^{th} POI. Hence, $\frac{\partial I_i}{\partial d_i} \geq 0$.
- Law of diminishing returns applies to the information gain, i.e., the marginal increase in information gain decreases with the dwell time; essentially, this implies that $\frac{\partial^2 I_i}{\partial d_i^2} \leq 0$.
- Information gained is always non-negative, i.e., $I_i(d_i) \geq 0$.

A consequence of these properties is that $I_i(d_i)$ is log-concave as $I_i(d_i)\frac{\partial^2 I_i}{\partial d_i^2} - (\frac{\partial I_i}{\partial d_i})^2 \leq 0$. It is also true that

$$J_0 = \sum_{i=1}^{n} I_i(d_i)$$

is also log-concave as

$$[\sum_{i=1}^{n} I_i(d_i)][\sum_{j=1}^{n} \frac{\partial^2 I_j}{\partial d_j^2}] - [\sum_{i=1}^{n} \frac{\partial I_i}{\partial d_i}]^2 \leq 0.$$

A consequence of this observation is that

$$J_s(d_1, d_2, \ldots, d_n) = e^{-\alpha_0(d_1+\ldots+d_n)} \sum_{i=1}^{n} I_i(d_i)$$

is log-concave, and hence, one may employ gradient ascent to $log(J_s(d_1, \ldots, d_n))$ to arrive at the optimum. In this paper, we model $P_i(s)$ as:

$$P_i(s) = P_i(d_i) = 1 - \frac{1}{2}e^{-d_i/\tau_i}\,, \tag{7}$$

where τ_i is a positive constant that represents the sensitivity to the time spent at the i^{th} POI. The plot below shows that the above three properties are satisfied by the information gain, $I_i(d_i)$. With this form of P_i, the information gain my be expressed as solely a function of d_i as follows:

$$P_i = 1 - \frac{1}{2}e^{-d_i/\tau_i}, \qquad (1 - P_i) = \frac{1}{2}e^{-d_i/\tau_i}$$
$$I_i(d_i) = (1 - \frac{1}{2}e^{-d_i/\tau_i})\log(1 - \frac{1}{2}e^{-d_i/\tau_i}) - \frac{1}{2}e^{-d_i/\tau_i}(\log 2 + d_i/\tau_i) + \log 2.$$

A sample plot of information gain corresponding to $\tau_i = 0.5\ units$ is given in Figure 3.1.

Since we want to incentivize the vehicles to visit all targets, we discount the information gain by the revisit time, R_i, for the i^{th} POI as follows:

$$\psi_i(d_i, R_i) = e^{-\alpha R_i} I_i(d_i),$$

where $\alpha > 0$ is a positive constant, R_i is the time duration between successive revisits to the i^{th} POI.

The objective of the optimization problem considered in this paper is to maximize the following function,

$$J_s(d_1, d_2, \ldots, d_n) = \sum_{i=1}^{n} \psi_(d_i, R_i),$$

through the choice of a route for the vehicles and the dwell time at each POI, while ensuring that each POI is visited.

3.2 Single Vehicle Case

In the case of a single vehicle, R_i is the same for every POI (say, it is R) if every other POI is visited exactly once between successive revisits; moreover $R = T + \sum_{i=1}^{n} d_i$, where T is time taken to tour the n POIs. If triangle inequality holds, this is true even if one may allow the same POI to be visited multiple times between consecutive revisits to another POI[10]. Note that $T \geq TSP^*$, where TSP^* is the minimum time taken to visit the n POIs before returning to the starting location. A consequence is the following:

$$\begin{aligned}
e^{-\alpha R} &\leq e^{-\alpha TSP^*} e^{-\alpha \sum_{i=1}^{n} d_i}, \\
\Longrightarrow J &\leq \sum_{i=1}^{n} e^{-\alpha TSP^*} e^{-\alpha \sum_{i=1}^{n} d_i} I_i(d_i), \\
&\leq \max_{d_1,\ldots,d_n} e^{-\alpha TSP^*} e^{-\alpha \sum_{i=1}^{n} d_i} \sum_{i=1}^{n} I_i(d_i) \\
&= e^{-\alpha TSP^*} \max_{d_1,\ldots,d_n} e^{-\alpha \sum_{i=1}^{n} d_i} \sum_{i=1}^{n} I_i(d_i)
\end{aligned}$$

If J^* is the optimum, clearly, it is achieved by minimizing T, and maximizing the log-concave function on the right hand side of the above inequality. In other words, the problem of optimal routing and the determination of optimal dwell time at each POIs is now decoupled.

3.2.1 Optimal Dwell Time

Let $\beta := e^{-\alpha TSP^*} > 0$. The objective is to find optimal values of $d_1, d_2, \ldots, d_n$ so as to maximize

$$J_1(d_1, \ldots, d_n) := \beta e^{-\alpha(d_1 + \ldots + d_n)} [\sum_{i=1}^{n} I_i(d_i)].$$

3.3 Multiple Vehicle Case

An additional complication arises in the multiple vehicle case – that of partitioning and assigning the POIs to be visited by each vehicle. If there are $m \geq 1$ vehicles, let the POIs be partitioned into m *disjoint* sets, namely $\mathcal{P}_1, \ldots, \mathcal{P}_m$, so that the i^{th} vehicle is tasked with visiting the POIs in $\mathcal{P}_i$. Let R_i be the revisit time associated with POIs assigned to i^{th} vehicle, and the associated tour cost for persistent monitoring per cycle be $TSP^*(\mathcal{P}_i)$. Associated with the i^{th} vehicle, the discounted information gained is given by

$$\max_{d_j, j \in \mathcal{P}_i} e^{-\alpha R_i} \sum_{j \in \mathcal{P}_i} I_j(d_j) = e^{-\alpha TSP^*(\mathcal{P}_i)} \max_{j \in \mathcal{P}_i} e^{-\alpha(\sum_{j \in \mathcal{P}_i} d_j)} \sum_{j \in \mathcal{P}_i} I_j(d_j).$$

Correspondingly, the objective is to maximize the discounted information gain over all possible partitions, sequences of visiting POIs by every vehicle and the dwell time at each POI:

$$J = \max_{\mathcal{P}_i,\, 1 \leq i \leq m} \max_{j \in \mathcal{P}_i} \sum_{i=1}^{m} e^{-\alpha TSP^*(\mathcal{P}_i)} e^{-\alpha(\sum_{j \in \mathcal{P}_i} d_j)} [\sum_{j \in \mathcal{P}_i} I_j(d_j)].$$

Since maximizing over partitions is a difficult combinatorial problem, we provide heuristics for the outer layer of optimization in the above optimization problem and use the single vehicle algorithm for the inner layer of optimization. We also include a two vehicle scenario so as to perform an exhaustive partitioning of the POIs and get an understanding of how well the partitioning heuristics perform as far as identifying the optimal partition.

4 Numerical Corroboration of Algorithms

4.1 Simulation Setup

Standard US datasets from TSPLIB [27] have been used to corroborated the routing algorithms developed in this article. These datasets provide the locations of POIs and identify the optimal tours for the single vehicle case; in this case, only optimal dwell times at each POI needs to be determined. In the case of multiple vehicles, we consider partitioning of the POIs from the datasets and use the single vehicle algorithm to compute the optimal sequences and dwell time; we use heuristics to refine the partitioning of POIs and present the results.

4.2 Numerical Results

Firstly, we have solved the problem for single vehicle case and extended it to multiple vehicle routing using the ideas developed in section 3 of the paper.

The optimal tour time was calculated using the TSP: Problem Based Solver in MATLAB. In Problem Based Approach, binary integer programming is used to solve the classical travelling salesman problem; which involves generating all possible trips *ie.* all distinct pairs of stops, calculating the distance for each trip and minimising the cost function *ie.* the sum of the trip distances over all possible trips. The decision variables associated with each trip are binary such that they are either 0 (when the trip is not on the tour) and 1 (when the trip is on the tour). In order to ensure that the tour includes every stop, a linear constraint is introduced that each stop is on exactly two trips (one arrival and one departure).

In figure 3 and figure 4, plots are shown for the optimal dwell time equation with only two POIs for simplicity. Here, $\tau_1 = \tau_2 = k$. It can be seen that there is an optimal dwell time where the dwell time objective function attains its maximum value.

We are presenting the case for 40 randomly generated cities for USA for both single and multiple vehicle (3 vehicles) routing. Problem-based solver for travelling salesman problem in MATLAB gives the minimum tour time for single vehicle routing whereas $K-means\ clustering$ was used to generate multiple clusters which were individually solved using problem based MATLAB solver to give minimum tour times for each cluster in multiple vehicle routing. MATLAB Problem-Based Non-Linear Optimization technique with $fmincon$ method has been used to find the optimal dwell time at each city by maximising the dwell time objective function. $fmincon$ is a gradient-based method designed to work on problems where the objective and constraint functions are both continuous and have continuous first derivatives, that finds a constrained minimum of a scalar function of several variables

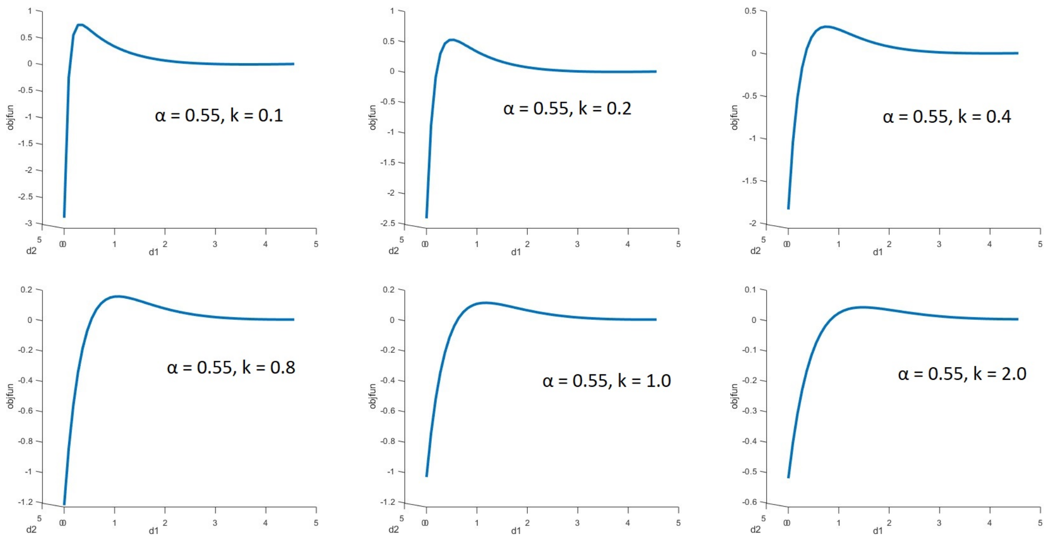

Figure 3: Dwell time variation w.r.t. k ($k = 0.1, 0.2, 0.4, 0.8, 1.0, 2.0$) with constant $\alpha = 0.55$

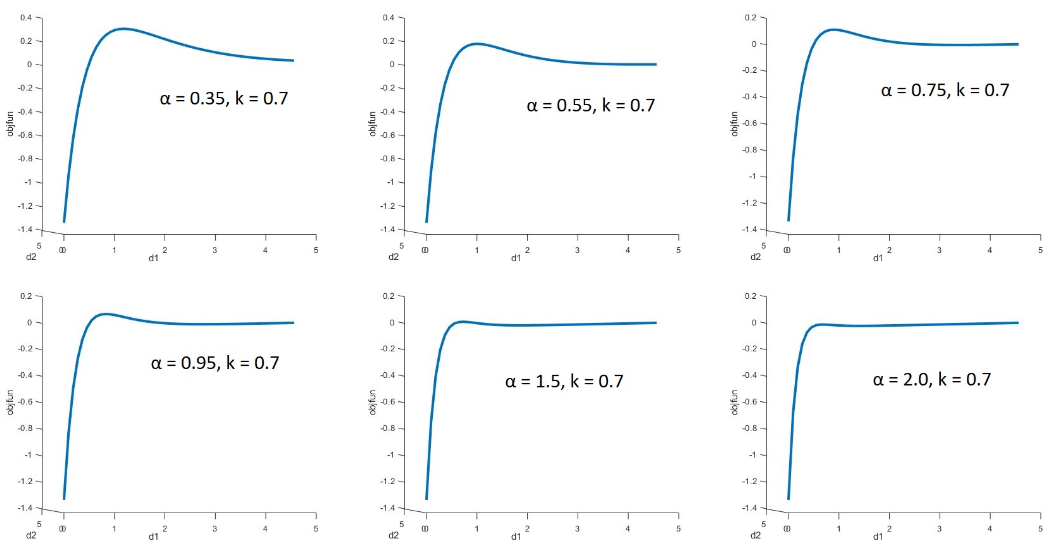

Figure 4: Dwell time variation w.r.t. α (α=0.35, 0.55, 0.75, 0.95, 1.5, 2.0) with constant $k = 0.7$

starting at an initial estimate. The total tour length for single vehicle routing is found to be 5.2613

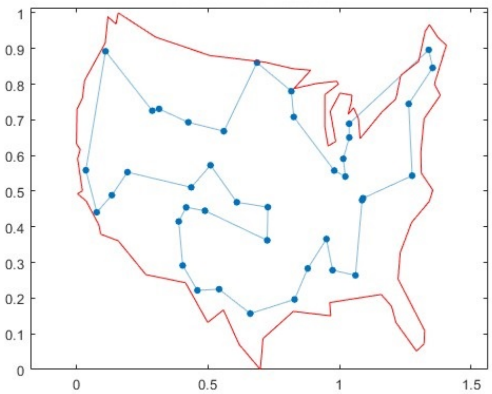

Figure 5: Single vehicle routing for 40 randomly chosen cities

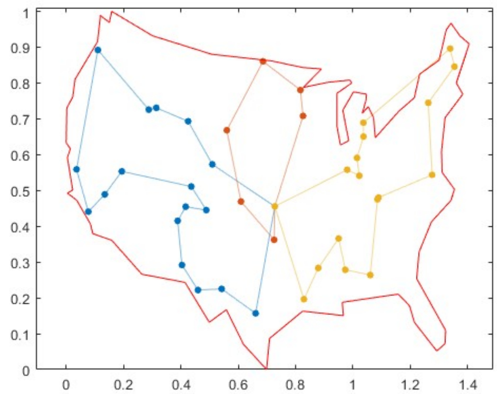

Figure 6: Three vehicle routing for 40 randomly chosen cities

units.

In case of multiple vehicle routing, either the optimum number of vehicles or a fixed number of vehicles can be taken to solve the problem. Since we are concerned with the information theory part of the problem, we have taken fixed number of vehicles, 3 in this case. $K-means\ clustering$ method has been used to find the optimal routes for the three vehicles. The individual tour lengths for multiple vehicle routing is found to be 2.310834 (rightmost tour yellow in color), 1.183055 (middle tour red in color) and 2.601053 (leftmost tour blue in color) units respectively.

Upon decomposing the objective function, it can be seen that the terms related to the total tour time and dwell time would separate. This would result in dwell time for each POI being dependent only on the values of α and τ_i. Thus even for multiple vehicle routing, the dwell time for each POI would remain the same. We will present the dwell time values for each POI later in this section. The values are presented for 40 cities case, taking the value of α to be 0.008 and τ to be $\sqrt{X-coo^2+Y-coo^2}$ where $X-coo$ and $Y-coo$ are the coordinates of individual nodes in 2D plane as shown in figure 5 and figure 6. The values of dwell times for each node will remain the same for both single and multiple vehicle cases.

Numerical simulations were performed on 20 instances for each case, *ie*. for n number of cities and m number of vehicles, 20 instances were randomly generated in MATLAB. Here, n ranged from 10 to 105 with a gap of 5 cities, and $m \in \{1, 3, 4, 5, 6\}$.

Running time was plotted for each case for single and multiple vehicle routing. It can be seen that average running time increases with the number of cities due to the nature of travelling salesman problem being NP-hard. It is known that the running time complexity for $K-means\ clustering$ varies linearly with the number of clusters, size of dataset and the number of iterations taken by the algorithm to converge. Hence, for smaller number of cities ($n \leq 50$), running time for single vehicle routing is lesser than for multiple vehicle routing case, but as the number of cities increases ($n \geq 50$) single vehicle routing becomes more time-consuming due to the NP-hard nature of TSP. $Intel(R)\ Core(TM)\ i7-8700\ CPU$ $@3.20GHz$, $3192Mhz$, 6 $Core(s)$, 12 Logical Processor has been used for running this on MATLAB.

5 Conclusions & Remarks

References

[1] Maryam Alipour and Karim Faez. On design of mobile agent routing algorithm for information gain maximization in wireless sensor networks. In *International Conference on Systems and Networks Communications*, 2011.

POI X-coo	POI Y-coo	Tour-Cluster	Dwell Time
1.3394	0.8963	yellow	2.0537
0.9737	0.2785	yellow	2.2025
1.0144	0.5909	yellow	2.2644
1.3543	0.8458	yellow	2.0756
0.9801	0.5578	yellow	2.2529
0.8794	0.2835	yellow	2.1431
1.0366	0.6892	yellow	2.2719
1.0372	0.6505	yellow	2.2711
1.0859	0.4751	yellow	2.2665
1.0224	0.541	yellow	2.2607
1.0608	0.2639	yellow	2.2409
1.0895	0.4802	yellow	2.2673
1.2632	0.7448	yellow	2.2048
0.9505	0.3659	yellow	2.2057
0.8293	0.1964	yellow	2.0822
1.2765	0.5436	yellow	2.2463
0.6607	0.1569	yellow	1.8853
0.8262	0.7081	red	2.239
0.817	0.7803	red	2.2535
0.6098	0.4689	red	1.9968
0.6865	0.8605	red	2.2438
0.7254	0.3622	red	2.0416
0.5604	0.6681	red	2.1002
0.7276	0.4553	red	2.0878
0.6607	0.1569	red	1.8853
0.4364	0.5108	blue	1.8755
0.0772	0.4408	blue	1.4998
0.036	0.5589	blue	1.7064
0.3889	0.4151	blue	1.7209
0.4253	0.6931	blue	2.044
0.6607	0.1569	blue	1.8853
0.4595	0.222	blue	1.6199
0.5423	0.2251	blue	1.7506
0.4039	0.2918	blue	1.5979
0.417	0.4546	blue	1.7967
0.5098	0.5728	blue	1.994
0.4887	0.4451	blue	1.8606
0.1939	0.5533	blue	1.7492
0.2881	0.7257	blue	2.0097
0.1347	0.4889	blue	1.6141
0.1104	0.8924	blue	2.1234
0.3141	0.7307	blue	2.0254

Table 1: Coordinate-wise dwell time for each POI for fig: 6

[2] Francesco Amigoni and Vincenzo Caglioti. An information-based exploration strategy for environment mapping with mobile robots. *Robotics and Autonomous Systems*, 58(5):684–699, 2010.

[3] Shi Bai, Jinkun Wang, Fanfei Chen, and Brendan Englot. Information-theoretic exploration with bayesian optimization. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1816–1822. IEEE, 2016.

[4] Nicola Basilico and Francesco Amigoni. Exploration strategies based on multi-criteria decision making for searching environments in rescue operations. *Autonomous Robots*, 31:401–417, 2011.

[5] Abhay Singh Bhadoriya, Christopher Montez, Sivakumar Rathinam Swaroop Darbha, David W Casbeer, and Satyanarayana G Manyam. Assisted shortest path planning for a convoy through a repairable network. *arXiv preprint arXiv:2204.00697*, 2022.

[6] Thomas M Cover and Joy A Thomas. *Elements of Information Theory*. Wiley, 2006.

[7] Di Deng, Runlin Duan, Jiahong Liu, Kuangjie Sheng, and Kenji Shimada. Robotic exploration of unknown 2d environment using a frontier-based automatic-differentiable information gain measure. In *2020 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, pages 1497–1503. IEEE, 2020.

[8] Noel E Du Toit and Joel W Burdick. Robot motion planning in dynamic, uncertain environments. *IEEE Transactions on Robotics*, 28(1):101–115, 2011.

[9] Sai Krishna Kanth Hari, Sivakumar Rathinam, Swaroop Darbha, and David W. Casbeer. Cooperative coverage with a leader and a wingmate in communication-constrained environments, 2022.

[10] Sai Krishna Kanth Hari, Sivakumar Rathinam, Swaroop Darbha, Krishna Kalyanam, Satyanarayana Gupta Manyam, and David Casbeer. Optimal uav route planning for persistent monitoring missions. *IEEE Transactions on Robotics*, 37(2):550–566, 2021.

[11] Sai Krishna Kanth Hari, Sivakumar Rathinam, Swaroop Darbha, Satyanarayana Gupta Manyam, Krishna Kalyanam, and David Casbeer. Bounds on optimal revisit times in persistent monitoring missions with a distinct and remote service station. *IEEE Transactions on Robotics*, 2022.

[12] Brian J Julian, Sertac Karaman, and Daniela Rus. On mutual information-based control of range sensing robots for mapping applications. *The International Journal of Robotics Research*, 33(10):1375–1392, 2014.

[13] Krishnamoorthy Kalyanam, Meir Pachter, Michael Patzek, Clayton Rothwell, and Swaroop Darbha. Optimal human–machine teaming for a sequential inspection operation. *IEEE Transactions on Human-Machine Systems*, 46(4):557–568, 2016.

[14] Evan Kaufman, Taeyoung Lee, and Zhuming Ai. Autonomous exploration by expected information gain from probabilistic occupancy grid mapping. In *2016 IEEE International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAR)*, pages 246–251. IEEE, 2016.

[15] Chou-Yuan Lee, Zne-Jung Lee, Shih-Wei Lin, and Kuo-Ching Ying. An enhanced ant colony optimization (eaco) applied to capacitated vehicle routing problem. *Applied Intelligence*, 32:88–95, 2010.

[16] Waqar Malik, Sivakumar Rathinam, and Swaroop Darbha. An approximation algorithm for a symmetric generalized multiple depot, multiple travelling salesman problem. *Operations Research Letters*, 35(6):747–753, 2007.

[17] Satyanarayana G. Manyam, David W. Casbeer, Swaroop Darbha, Isaac E. Weintraub, and Krishna Kalyanam. Path planning and energy management of hybrid air vehicles for urban air mobility. *IEEE Robotics and Automation Letters*, 7(4):10176–10183, 2022.

[18] Satyanarayana G Manyam, Sivakumar Rathinam, David Casbeer, and Eloy Garcia. Tightly bounding the shortest dubins paths through a sequence of points. *Journal of Intelligent & Robotic Systems*, 88:495–511, 2017.

[19] Satyanarayana G Manyam, Sivakumar Rathinam, and Swaroop Darbha. Computation of lower bounds for a multiple depot, multiple vehicle routing problem with motion constraints. *Journal of Dynamic Systems, Measurement, and Control*, 137(9), 2015.

[20] Christopher Montez, Swaroop Darbha, Christopher Valicka, and Andrea Staid. Routing of an unmanned vehicle for classification. In Tien Pham, Latasha Solomon, and Katie Rainey, editors, *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II*, volume 11413, page 1141319. International Society for Optics and Photonics, SPIE, 2020.

[21] Yasamin Mostofi. Decentralized communication-aware motion planning in mobile networks: An information-gain approach. *Journal of Intelligent and Robotic Systems*, 56:233–256, 2009.

[22] Jasmina Novakovic. Using information gain attribute evaluation to classify sonar targets. In *17th Telecommunications forum TELFOR*, pages 1351–1354. Citeseer, 2009.

[23] Paul Oberlin, Sivakumar Rathinam, and Swaroop Darbha. A transformation for a multiple depot, multiple traveling salesman problem. In *2009 American Control Conference*, pages 2636–2641. IEEE, 2009.

[24] Paul Oberlin, Sivakumar Rathinam, and Swaroop Darbha. Today's traveling salesman problem. *IEEE Robotics Automation Magazine*, 17(4):70–77, 2010.

[25] Karl J Obermeyer, Paul Oberlin, and Swaroop Darbha. Sampling-based path planning for a visual reconnaissance unmanned air vehicle. *Journal of Guidance, Control, and Dynamics*, 35(2):619–631, 2012.

[26] Liam Paull, Sajad Saeedi, Howard Li, and Vincent Myers. An information gain based adaptive path planning method for an autonomous underwater vehicle using sidescan sonar. In *2010 IEEE International Conference on Automation Science and Engineering*, pages 835–840. IEEE, 2010.

[27] Gerhard Reinelt. **TSPLIB**, 1995. Accessed on 22 May 2023.

[28] Swaroop Darbha Siva Rathinam, Raja Sengupta. A resource allocation algorithm for multivehicle systems with motion constraints. *IEEE Transactions on Automation Science and Engineering*, 4(1):98–104, 2007.

[29] Sai Yadlapalli, Sivakumar Rathinam, and Swaroop Darbha. 3-approximation algorithm for a two depot, heterogeneous traveling salesman problem. *Optimization Letters*, 6:141–152, 2012.

[30] Tobias Zaenker, Julius Rückin, Rohit Menon, Marija Popović, and Maren Bennewitz. Graph-based view motion planning for fruit detection. *arXiv preprint arXiv:2303.03048*, 2023.

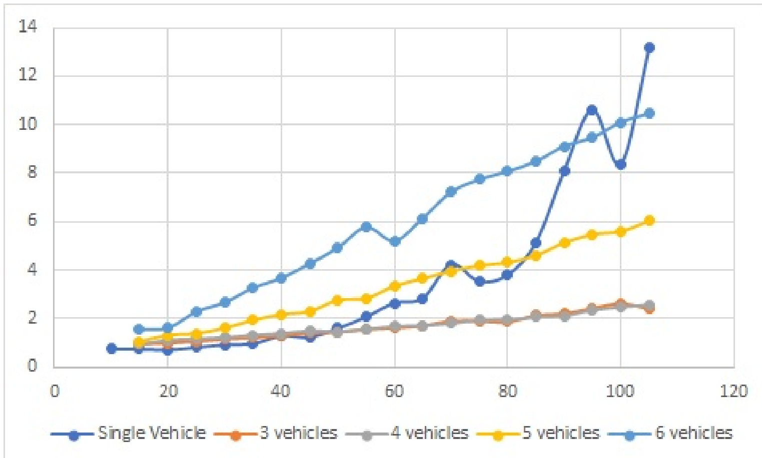

Figure 7: running time over number of cities for single & multiple vehicle routing

vehicle	Average Running Time
Single Vehicle	3.677111
Three vehicles	1.6715868
Four vehicles	1.6985218
Five vehicles	3.3510146
Six vehicles	5.8990271

Table 2: Average running time for each case