ON MAXIMISING THE TOTAL INFORMATION GAIN IN A VEHICLE ROUTING PROBLEM

A Thesis

by

SUVEER KUMAR

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Swaroop Darbha
Committee Members,	Sivakumar Rathinam
	Dezhen Song
Head of Department,	Guillermo Aguilar

August 2023

Major Subject: Mechanical Engineering

ABSTRACT

Information gain-based approach is frequently used for exploration in human-machine systems where multiple robots aid a remotely located operator in classification. The amount of information and the marginal increment in information gain depends on the time spent (dwell time) by a robot at the POI. If there are multiple POIs to be monitored, not all of them may be simultaneously monitored. In such a case, the information gain must be discounted by the duration between successive revisits to the same POI. Based on the discounted information gain, a robot can adaptively choose the dwell time for each POI to aid the operator in better classification.

This thesis develops a mathematical formulation for maximizing the total discounted information gain when monitoring multiple POIs using a human-machine system. In this framework, an interface typically takes multiple POIs as input from the human operator (who often serves as a classifier-in-the-loop) and computes the order in which they should be visited, and the dwell time of robots sent for monitoring at each POI.

The underlying technical problem consists of determining the optimal assignment of POIs to visit for each robot, the sequence of POIs to visit by each robot, and the dwell time at each robot. For the single robot case, the problem simplifies to the determination of last two sets of variables. In this thesis, the log-concavity of total discounted information gain is exploited to show that the optimal routing for the single robot reduces to the determination of optimal tour (using TSP solution), and optimal dwell time through a gradient ascent or equivalent approaches. In the multiple robot case, this thesis presents a partitioning heuristic for POIs based on k-means clustering; once the clusters are determined, a robot is assigned to a cluster to which it belongs, and the resulting problem reduces to the single robot case. Numerical simulations presented corroborate the algorithms developed in this thesis.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a thesis committee consisting of Professor Swaroop Darbha [Advisor and Chair], Professor Sivakumar Rathinam [Committee Member I] of the Department of Mechanical Engineering and Professor Dezhen Song [Committee Member II] of the Department of Computer Science & Engineering at Texas A& M University, College Station.

The idea and theoretical development of this project was pioneered by Professor Swaroop Darbha. All other work conducted for the thesis was completed by the student with guidance from the advisor.

Funding Sources

No additional source of funding was provided for this work.

NOMENCLATURE

UAV	Unmanned Aerial Vehicle
VRP	Vehicle Routing Problem
DUE	Dynamic Uncertain Environment
SSS	Side Scan Sonar
IG	Information Gain
POI	Point Of Interest
T	Target
F	Not A Target
TSP	Travelling Salesman Problem

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

This thesis considers a human-machine system where the human serves as a classifier-in-the-loop, and makes decisions based on the information delivered to the human operator by the robots (in this case, UAVs) through an interface. The interface takes n points of interest (POI) as input from the human operator, computes the order in which they should be visited, the time to be spent by each robot at each assigned POI (dwell time) in the sequence. The UAVs persistently monitor the POIs by visiting them and dwelling at each POI while collecting information that is then transmitted to the remotely located human operator. Based on this information transmitted by the UAV, the primary task of human operator is to classify the n specified POIs as T (target) or F (false target). The probability of correctly classifying a POI depends upon the dwell time of the UAV at that location. At each POI, the information gained is the K-L distance between the two conditional probability distributions: the first one is conditioned on the POI being a T and indicates the probabilities of the operator classifying it as a T or a F; similarly, the second one is conditioned on the POI being a F. Since persistent monitoring is desired, a penalty is imposed on the information gain for excessive time duration between revisits to the same POI (which will henceforth be referred to as revisit time for the POI). The discount factor associated with a POI decreases exponentially with the corresponding revisit time. Summing up, the problem considered in this thesis is to maximize the total discounted information gain through the determination of the

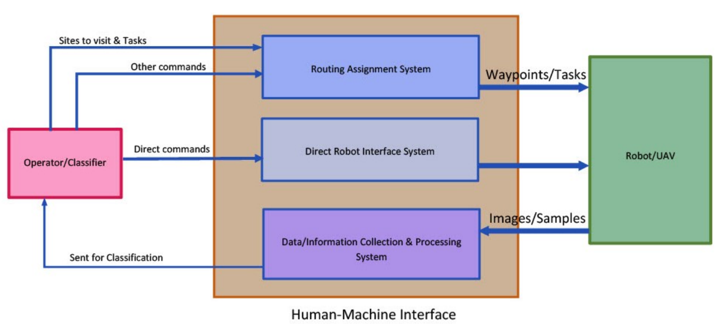

Figure 1.1: Human-Machine System with human as a classifier-in-the-loop

optimal sequence in which POIs must be visited and the dwell time at each POI, while ensuring that each of them is visited.

1.1 Literature Review

Automatic Target Recognition (ATR) systems have been shown in literature to perform at an acceptable level in relatively benign environments (such as low clutter); however, medium to high background clutter introduces unacceptable levels of false alarms for ATR systems. Moreover, variability in targets and environmental conditions significantly degrade their performance [1]. Human classification under these conditions is admittedly better and is the premise for the use of human as a classifier-in-the-loop for human-machine systems considered in this thesis.

An information gain based approach has been adopted in this thesis for aiding classification by the operator. The performance of any classifier can be characterized by a Receiver-Operating Curve (ROC); it is also called confusion matrix. If there are p classification categories for an object or a POI, the operator, based on the observations/measurements, may or may not classify correctly. The $(i, j)^{th}$ entry of the confusion matrix indicates the probability of an operator (mis)classifying the object/POI of i^{th} category as an object/POI of j^{th} classification category. The entries of the confusion matrix must be determined experimentally; they can depend on multiple operational factors such as the altitude or pose of the vehicle taking the image of the object, time duration spent at an object/POI etc. In the case of binary classification, there are only two classification categories, T and F; only binary classification is considered in this thesis. Correspondingly, there are only two rows in the confusion matrix, each corresponding to operator's probability of correctly classifying the object/POI. The first row corresponds to the object/POI being of type T and the second corresponds to the object/POI being of type F. Since the probabilities depend on controllable operational variables, a natural question arises as to what the optimal values of the operational variables should be to aid the operator to the maximum extent in classification. In this thesis, we consider the dwell time at each POI as the only controllable operational variable. Clearly, if the distance between the two conditional probability distributions is maximized, the operator's performance as a classifier will improve. The distance between probability distributions, referred to as information gain [2], is a key metric, and depends on the controllable operational variables in this application.

UAVs are treated as data gathering robots/platforms in this thesis. The tracking and recognition system of Unmanned Aerial Vehicles (UAVs) has the advantages of low cost, greater affordability, easy usage, zero casualty, good concealment, high flexibility, and small volume, which can make up for the lack of satellite acquisition of near ground and low-altitude information [3]. Besides the military applications in covert tracking & reconnaissance of the target [4], UAVs can be used in civil areas such as monitoring crop growth [5], rescue and disaster relief, and tracking and hunting criminals when they escape [6].

Path planning for a UAV is central for enabling autonomy and concerns with finding a path for the UAV from the starting location to the goal point in such a way that the assigned tasks are carried out efficiently. Israr $et\ al.$ [7] summarized various promising motion planning techniques and algorithms for determining the optimum path (consume less time and energy) for UAVs such that the performance constraints are satisfied and collisions get avoided. In a typical UAV application, the number of degrees of freedom range from two to four, and has differential constraints such as limited speed and maximum acceleration. As the number of degrees of freedom for a UAV increases, the motion planning problem becomes increasingly complex.

Unmanned Vehicle Routing for autonomously gathering information has received significant attention in the literature [8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20] where 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. The previous references addressed the problems related to routing but very little was described about the human operator involved.

The use of information gain in vehicle routing and motion planning has proved beneficial for several applications. Lee $et\ al.$ [21] developed an enhanced ant colony optimization for the capacitated vehicle routing problem by using information gain to enhance the search performance when the good initial solution was provided by simulated annealing algorithm. Toit and Burdick [22] 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 $et\ al.$ [23] 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 $et\ al.$ [24] 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 $et\ al.$ [25] 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 [26] 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 has been used in 3D data acquisition and geometry reconstruction as well, areas which have great applications in the field of computer vision & robotics. Potthast & Sukhatme [27] proposed a method that utilizes a belief model of the unobserved space to estimate the expected information gain of each possible viewpoint in the next best view problem for occluded environment. The proposed belief model allows a more precise estimation of the visibility of occluded space and a more accurate prediction of the potential information gain of new viewing positions. Palazzolo & Stachniss [28] presented a novel vision based autonomous exploration on Micro Aerial Vehicles (MAVs) using information gain. Their approach iteratively samples candidate viewpoints and greedily selects the optimal one based on a utility function aimed at maximising the expected information gain while minimising the cost of acquiring the new measurement during exploration. Stachniss $et\ al.$ [29] presented an integrated approach to exploration, mapping, and localisation by computing the expected information gain (change of entropy) of a highly efficient Rao-Blackwellized particle filter to evaluate an action, that guides a robot from its current location to a goal location. Paul *et al.* [30] proposed an algorithm for Autonomous eXploration to Build A Map (AXBAM) of an unknown 3D complex steel bridge structure using a 6 DOF robot manipulator, that considers the trade-off between the predicted environment information gain available from a sensing viewpoint and the manipulator joint angle changes required to position a sensor at that viewpoint. In this approach, information is gained from multiple viewpoints that is fused to obtain a detailed 3D map. Quin *et al.* [31] built upon this algorithm by considering the information gain from only a small set of poses (vector of joint angles) neighbouring the robot's current pose for exploration of complex 3D environments. Zhang *et al.* [32], while developing a new roadmap for computing a robotic sensor path in order to classify multiple fixed targets located in an obstacle-populated workspace, observed that the paths obtained from the information theoretic function criterion exhibited a classification efficiency several times higher than that of existing search strategies. They quantified the value of information by the expected entropy reduction which was computed from the Bayesian Network (BN) conditional probability tables (CPTs) and from the prior information, such as, prior sensor measurements and environmental conditions. Denzler & Brown [33] demonstrated the benefits of using information theory in an object recognition application using an active camera for sequential gaze control and viewpoint selection. They used reduction of uncertainty in the state estimation process as the optimality criterion, rather than an estimator-specific metric (e.g., minimum mean squared error) by claiming that the state estimation becomes more reliable if the uncertainty and ambiguity in the estimation process can be reduced. Their technique explicitly takes into account the *a priori* probabilities governing the computation of the mutual information, which is then used to form a sequential decision process by treating the *a priori* probability at a certain time step in the decision process as the *a posteriori* probability of the previous time step.

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 [34] 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 $et\ al.$ [35] proposed a novel algorithm for the optimising 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 view-points along the path simultaneously. Julian $et\ al.$ [36] 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 $et\ al.$ [37] proposed a novel approach to predict mutual information using Bayesian optimisation for the purpose of exploring $a\ priori$ 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 [38] 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 [39] further extended this information-based exploration strategy for rescue and surveillance applications.

1.2 Vehicle Routing Problem

Vehicle Routing Problem [40] can be described as the problem of designing least-cost delivery routes from a depot to a set of geographically scattered customers, subject to side constraints. Though several variations of this problem exist due to which this problem is a class in itself, an archetypical version of VRP can be defined as follows. Let *G=(V,A)* be a graph where *V={0,...,n}* be the set of vertices representing cities with the depot located at vertex 0, and $A = (i, j) : i, j \in$

$V, i \neq j$ be the set of edges. Now, with every arc (i,j) is associated a non-negative *distance* matrix $C = (c_{ij})$, where c_{ij} can be interpreted for travel cost or travel time assuming both are directly proportional to one another. When C is symmetrical, it is safe to assume that the graph is undirected means that edges do not have directionality in their nature. Let A then be replaced by E which is a set of undirected edges for this graph. Let's suppose that we have m vehicles for this purpose such that $m_L \leq m \leq m_U$ and that each vehicle is identical with same capacity D and has a fixed cost f for its use. The VRP consists of designing a set of least-cost vehicle routes in such a way that each city in V is visited exactly once by exactly one vehicle, all vehicles start and end at the depot and some side-constraints are satisfied. The most common side-constraints include capacity constraints, bound on the number of cities in every route, total time restrictions, time windows, precedence relations $ie.$ one city been visited prior to another one etc.

Thus, we have an objective function to minimize for the total cost, with the restrictions on tour geometry, maximum number of vehicles deployed and other side-constraints. Let's assume that depot is at vertex 0 and that x_{ij} is a binary variable such that it attains the value 1 when the vehicle departs from vertex i and arrives at vertex j directly, for $i, j \in V$.

$$\text{min} \sum_{i=0}^{n} \sum_{j=0}^{n} c_{ij} x_{ij}, \tag{1.1}$$

$$\text{s.t.} \sum_{\substack{j=1 \\ j \neq i}}^{n} x_{ij} = 1, \tag{1.2}$$

$$\sum_{\substack{i=0 \\ i \neq h}}^{n} x_{ih} - \sum_{\substack{j=0 \\ j \neq h}}^{n} x_{hj} = 0, \tag{1.3}$$

$$\sum_{j=0}^{n} x_{0j} \leq m \tag{1.4}$$

Constraints 1.2 ensure that all vertices or nodes or cities are visited exactly once. Constraints 1.3 ensure that if a vehicle arrives at a node $h \in V$, then it must depart from this node. Constraint 1.4 limits the maximum number of routes to m, the number of vehicles.

1.2.1 Routing for Classification

The problem of routing vehicles for aiding an operator-in-the-loop for classification was first proposed by Montez [41]; 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 in this thesis.

1.3 Technical Challenges

The problem considered in this thesis requires overcoming the following challenges:

1. **Partitioning of POIs for assigning them to each robot/UAV:** This is a combinatorially challenging problem, especially when the number of robots and POIs increase.

2. **Sequencing & finding the dwell time at each POI:** This is a mixed-integer optimization problem, where one must determine the continuous variables, namely the dwell time at each POI, and the binary variables determining the edge selected for each vehicle route. The routing problem is a known NP-hard problem.

The approach taken in this thesis is to exploit the log-concavity of the objective function constructed for optimization, and construct efficient algorithms for a single-robot case and build heuristics for partitioning so that the top-level problem can be broken into multiple single-robot problems.

1.3.1 Organization of thesis

The assumptions and formulation of the problem considered in this thesis are given in section 2; while the corroborating numerical results are provided in section 3. Concluding remarks and suggestions for future research are given in section 4.

2. PROBLEM FORMULATION

The following assumptions are the basis for formulating the problem considered in this thesis:

- POIs model locations where activities take place that may need to be classified. The duration of the activity needing attention is reasonably large that it is at least the maximum revisit time for any target. This assumption ensures that the activity to be classified may not avoid detection by UAVs.

- Longer dwell time at a POI provides a better idea of the activity and hence, leads to a better classification.

- Confusion matrix is a reasonable representation of the operator as a classifier-in-the-loop and is available for an operator *a priori*.

2.1 Mathematical Formulation

2.1.1 Setup

Let us consider a set of n points of interest (POIs) to be visited by a single vehicle to gain information about each POI. We can use the same set-up as before, i.e., $G = (V, A)$ be the undirected graph and that each vehicle has to depart and return to a single depot (vertex 0). When the vehicle visits a POI, information is gained about that POI in order to classify it as either T (target) or F (not a target) with the exception of depot which will have no classification. Let us assume that there is no *a priori* information about the POIs and that the probability of correctly classifying the POI is the same whether the POI has a true classification of T or F, which means it is equally difficult to classify the POI, regardless of what it really is. The objective is then to construct a path in G such that the vehicle visits each POI once and maximises the total information gained.

2.1.2 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

$$P_t(s_i) = P(Z = T \,|\, X = T,\, s_i) \text{ and} \tag{2.1}$$

$$P_f(s_i) = P(Z = F \,|\, X = F,\, s_i), \tag{2.2}$$

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

$$I_i(X, Z) := H(X) - H(X|Z)\,, \tag{2.3}$$

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

$$I_i(X, Z) = \sum_{x,z \in C} P(X = x, Z = z) \log \frac{P(X = x, Z = z)}{P(X = x)P(Z = z)}\,. \tag{2.4}$$

Denote the *a priori* probability a POI is a target, $P(X = T)$, as p. It can then be shown Equation (2.4) can be rewritten as

$$\begin{aligned} I_i(X, Z) = {} & pP_t(s) \log\left(\frac{P_t(s)}{pP_t(s) + (1-p)(1-P_f(s))}\right) \\ & + (1-p)(1-P_f(s)) \log\left(\frac{1-P_f(s)}{pP_t(s) + (1-p)(1-P_f(s))}\right) \\ & + p(1-P_t(s)) \log\left(\frac{1-P_t(s)}{p(1-P_t(s)) + (1-p)P_f(s)}\right) \\ & + (1-p)P_f(s) \log\left(\frac{P_f(s)}{p(1-P_t(s)) + (1-p)P_f(s)}\right). \end{aligned} \tag{2.5}$$

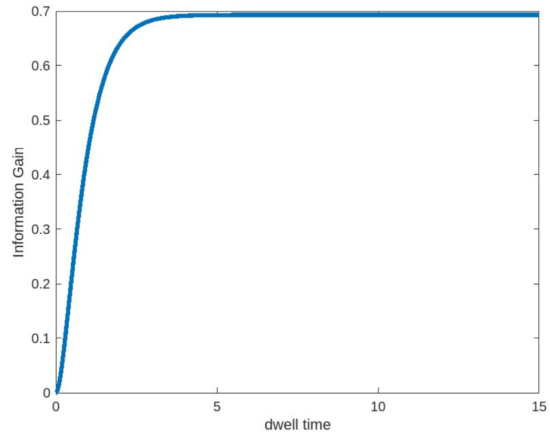

Figure 2.1: Information Gain vs Dwell Time at a POI ($\tau = 0.5$)

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 (2.5) reduces to

$$I_i(X, Z) = P_i(s) \log P_i(s) + (1 - P_i(s)) \log(1 - P_i(s)) + \log 2\,. \tag{2.6}$$

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 2.1):

- 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

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

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:

$$P_i(s) = P_i(d_i) = 1 - \frac{1}{2}e^{-d_i/\tau_i}\ , \tag{2.7}$$

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 2.1.

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.

2.1.3 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[9]. 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:

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

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.

2.1.3.1 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

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

2.1.4 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.

3. NUMERICAL RESULTS

We now illustrate the above information gain theory with the help of an example. We have solved this optimization problem in MATLAB. For the purpose of convenience, special constraints such that zero information gain from depot has been removed and depot is being treated just like any other node (or vertex or POI). Also we have assumed that the POI nearest to the centroid of all POIs is chosen as the depot. Firstly, this problem has been solved for single vehicle routing and then it has been extended for multiple vehicle routing.

While solving this problem, firstly the total tour time was calculated using the TSP: Problem Based method in MATLAB. In Problem Based approach, binary integer programming is used to solve the classical Travelling Salesman Problem. It involves generating all possible trips $i.e.$ all distinct pairs of stops, calculating the distance for each trip and minimizing the cost function $i.e.$ the sum of the trip distances for each trip in the hour. 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 3.6 and figure 3.7, plots are shown for $e^{-\alpha(d_1+...+d_n)} \sum_{i=1}^{n} I_i(d_i)$ for simplest case with just two cities. It can be seen that there is an optimum dwell time where the function $e^{-\alpha(d_1+d_2)} \sum_{i=1}^{2} I_i(d_i)$ attains the maximum value. Here, $k = \tau_1 = \tau_2$.

We are presenting the case for 40 randomly generated cities of 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 Optimisation technique with $fmincon$ method has been used to find the optimal dwell time at each city by maximising $e^{-\alpha(d_1+...+d_n)} \sum_{i=1}^{n} I_i(d_i)$. $fmincon$ 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 starting at an initial estimate.

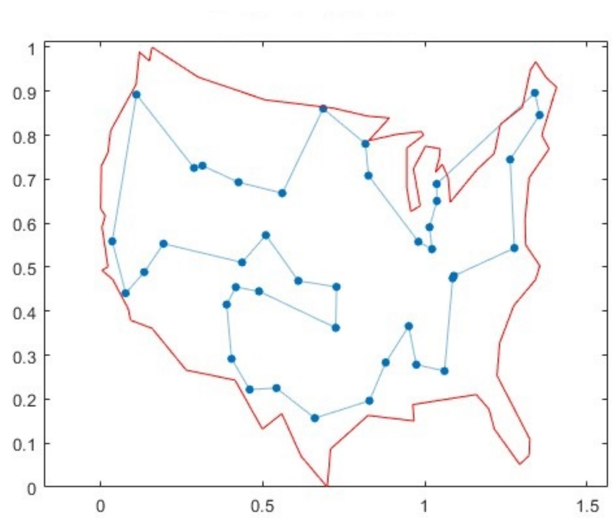

Figure 3.1: Single vehicle routing for 40 randomly chosen cities

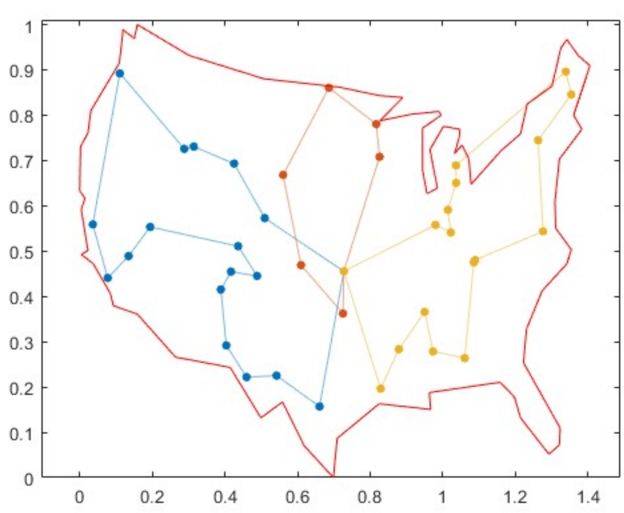

Figure 3.2: Three vehicle routing for 40 randomly chosen cities

The total tour length for single vehicle routing is found to be 5.2613 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. 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 3.1 and figure 3.2. The values of dwell times for each node will be different for single and multiple vehicle cases.

Numerical simulations were performed on 20 instances for each case, $i.e.$ 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.

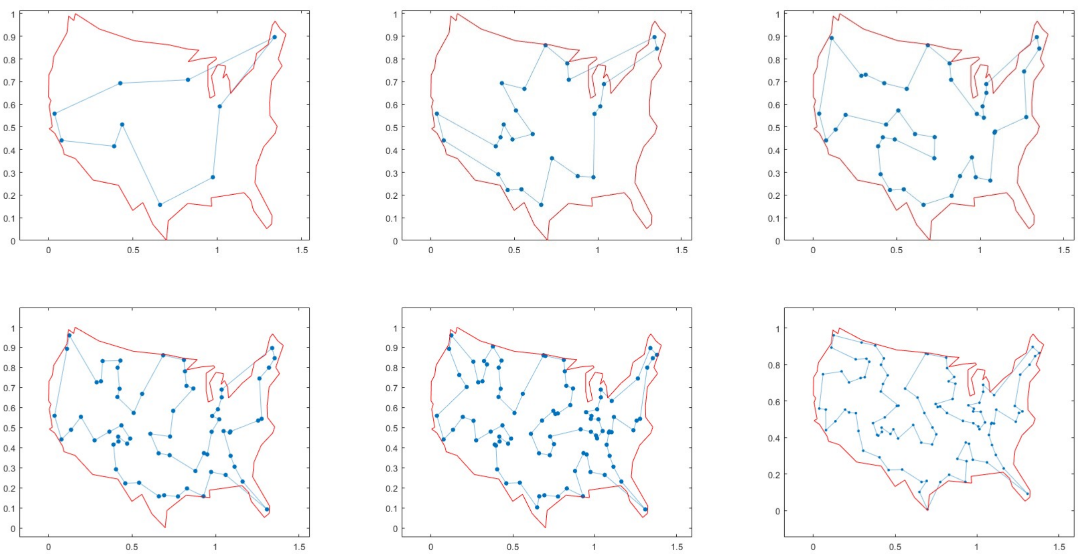

Figure 3.3: Single Vehicle Routing for 10, 25, 40, 65, 90, 105 cities

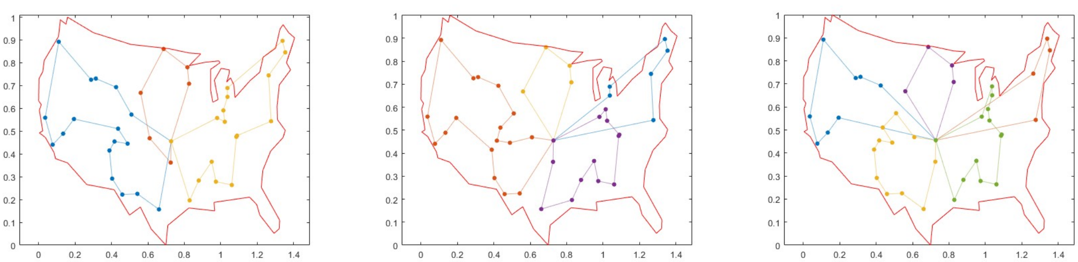

Figure 3.4: Multiple Vehicle Routing for 40 cities with 3, 4, 5 vehicles

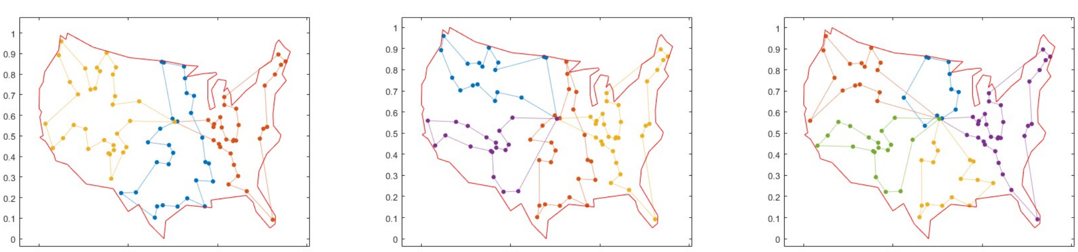

Figure 3.5: Multiple Vehicle Routing for 90 cities with 3, 4, 5 vehicles

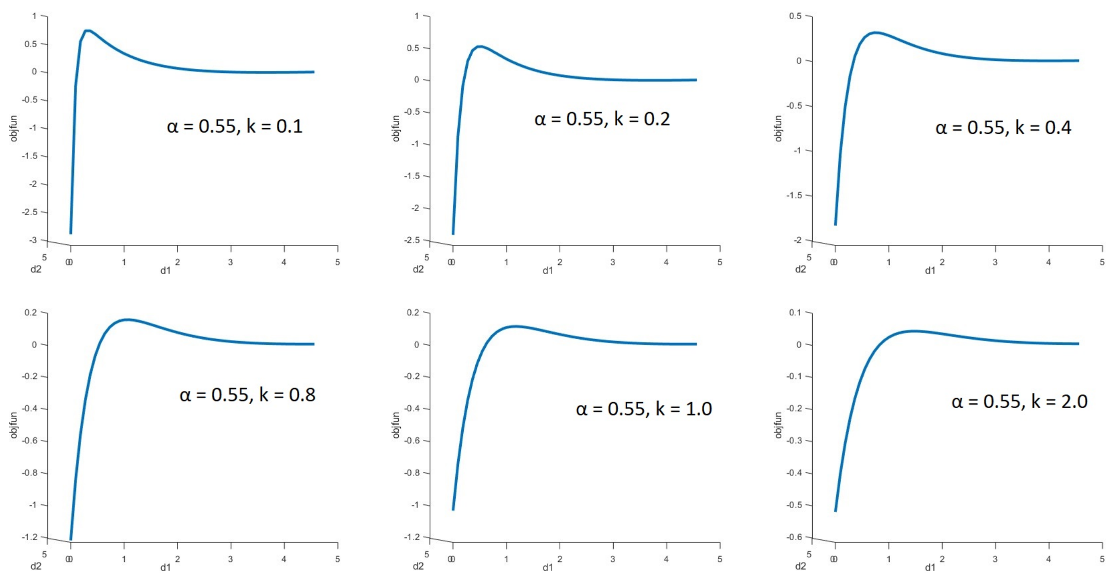

Figure 3.6: Variation of $e^{-\alpha(d_1+d_2)} \sum_{i=1}^{2} I_i(d_i)$ w.r.t. k ($k = 0.1, 0.2, 0.4, 0.8, 1.0, 2.0$) with constant $\alpha = 0.55$

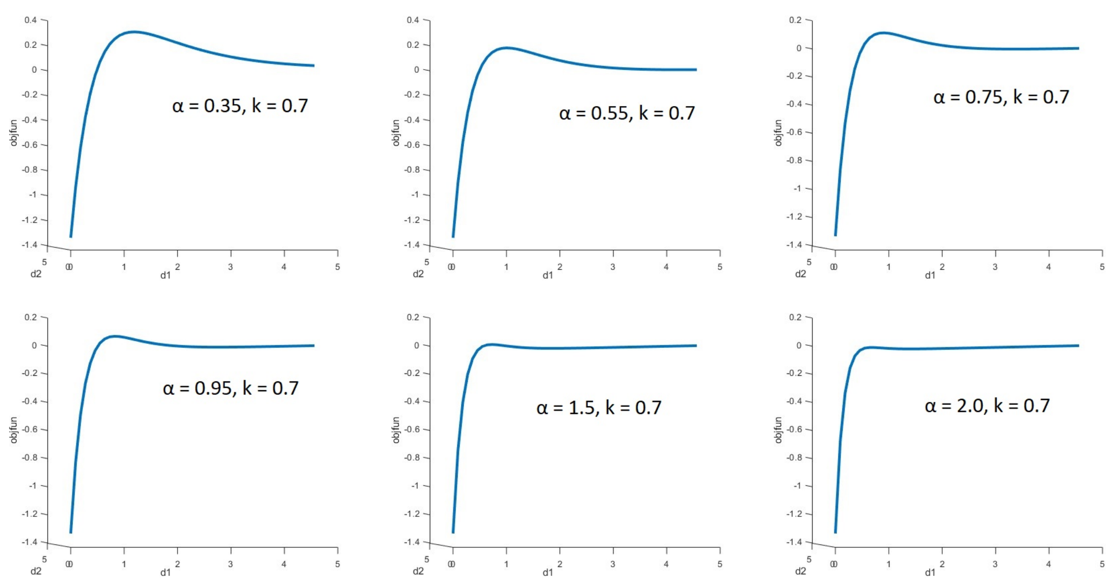

Figure 3.7: Variation of $e^{-\alpha(d_1+d_2)} \sum_{i=1}^{2} I_i(d_i)$ w.r.t. α (α=0.35, 0.55, 0.75, 0.95, 1.5, 2.0) with constant $k = 0.7$

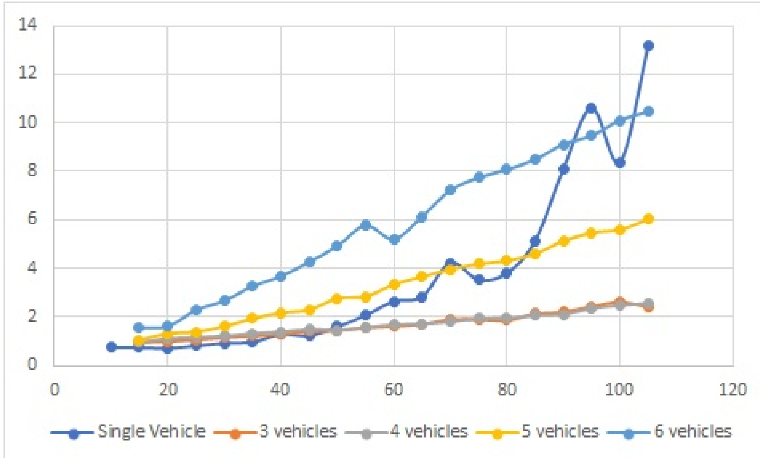

Figure 3.8: Running time over number of cities for single & multiple vehicle routing

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 3.1: Average running time for each case

POI X-coo	POI Y-coo	Tour-Cluster	Single Vehicle Dwell Time	Three Vehicle Dwell Time
1.3394	0.8963	yellow	2.0537	4.0893
0.9737	0.2785	yellow	2.2025	3.2351
1.0144	0.5909	yellow	2.2644	3.5117
1.3543	0.8458	yellow	2.0756	4.0734
0.9801	0.5578	yellow	2.2529	3.4361
0.8794	0.2835	yellow	2.1431	3.0658
1.0366	0.6892	yellow	2.2719	3.6217
1.0372	0.6505	yellow	2.2711	3.5906
1.0859	0.4751	yellow	2.2665	3.5298
1.0224	0.541	yellow	2.2607	3.4839
1.0608	0.2639	yellow	2.2409	3.3776
1.0895	0.4802	yellow	2.2673	3.5383
1.2632	0.7448	yellow	2.2048	3.9242
0.9505	0.3659	yellow	2.2057	3.2456
0.8293	0.1964	yellow	2.0822	2.9199
1.2765	0.5436	yellow	2.2463	3.8232
0.6607	0.1569	yellow	1.8853	2.5291
0.8262	0.7081	red	2.239	4.2832
0.817	0.7803	red	2.2535	4.3928
0.6098	0.4689	red	1.9968	3.365
0.6865	0.8605	red	2.2438	4.317
0.7254	0.3622	red	2.0416	3.4935
0.5604	0.6681	red	2.1002	3.6775
0.7276	0.4553	red	2.0878	3.637
0.6607	0.1569	red	1.8853	3.0759
0.4364	0.5108	blue	1.8755	2.4385
0.0772	0.4408	blue	1.4998	1.8577
0.036	0.5589	blue	1.7064	2.1649
0.3889	0.4151	blue	1.7209	2.1874
0.4253	0.6931	blue	2.044	2.7464
0.6607	0.1569	blue	1.8853	2.4552
0.4595	0.222	blue	1.6199	2.0334
0.5423	0.2251	blue	1.7506	2.2339
0.4039	0.2918	blue	1.5979	2.0005
0.417	0.4546	blue	1.7967	2.3076
0.5098	0.5728	blue	1.994	2.6497
0.4887	0.4451	blue	1.8606	2.4132
0.1939	0.5533	blue	1.7492	2.2317
0.2881	0.7257	blue	2.0097	2.6793
0.1347	0.4889	blue	1.6141	2.0247
0.1104	0.8924	blue	2.1234	2.9155
0.3141	0.7307	blue	2.0254	2.7097

Table 3.2: Coordinate-wise dwell time for each POI for fig: 3.1 and fig: 3.2

4. CONCLUSIONS

Through this paper, we investigated how Kullback-Leibler divergence (or Mutual Information) can be used to quantify the information gained about the classification status of a POI and can be used to decide the dwell time at each POI.

Calculating the total information gain in vehicle routing boils down to an objective function which can be separated variably into two functions: one which aims at minimising the total tour time and another which aims at maximising the product of information gain with an exponential function. This is when the probability of correctly classifying a POI as a target (T) or not a target (F) is dependent solely on the dwell time of vehicle at that POI during its visit. However, in real application, this probability will depend on other state information that can be experimentally determined before sending the vehicle for exploration.

Our present work is aimed at human-machine systems where the information gained by a UAV is transmitted to a remote human operator for classifying POIs as target or not a target. A key area for future work is the adaptation of this approach in persistent monitoring by UAVs especially in rescue and surveillance applications. This information-theoretic approach can also be beneficial in multiple target tracking by UAV when time resource has to be effectively allocated to its responsible targets.

REFERENCES

[1] J. A. Ratches, “Review of current aided/automatic target acquisition technology for military target acquisition tasks,” *Optical Engineering*, vol. 50, no. 7, p. 072001, 2011.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. 2006.

[3] G. V. Konoplich, E. O. Putin, and A. A. Filchenkov, “Application of deep learning to the problem of vehicle detection in uav images,” in *2016 XIX IEEE International Conference on Soft Computing and Measurements (SCM)*, pp. 4–6, IEEE, 2016.

[4] R. Wise and R. Rysdyk, “Uav coordination for autonomous target tracking,” in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, p. 6453, 2006.

[5] B. Li, R. Liu, S. Liu, Q. Liu, F. Liu, and G. Zhou, “Monitoring vegetation coverage variation of winter wheat by low-altitude uav remote sensing system,” *Transactions of the Chinese Society of Agricultural Engineering*, vol. 28, no. 13, pp. 160–165, 2012.

[6] A. Prabhakaran and R. Sharma, “Autonomous intelligent uav system for criminal pursuit—a proof of concept,” *The Indian Police Journal*, vol. 68, no. 1, pp. 1–20, 2021.

[7] A. Israr, Z. A. Ali, E. H. Alkhammash, and J. J. Jussila, “Optimization methods applied to motion planning of unmanned aerial vehicles: A review,” *Drones*, vol. 6, no. 5, p. 126, 2022.

[8] S. D. Siva Rathinam, Raja Sengupta, “A resource allocation algorithm for multivehicle systems with motion constraints,” *IEEE Transactions on Automation Science and Engineering*, vol. 4, no. 1, pp. 98–104, 2007.

[9] S. K. K. Hari, S. Rathinam, S. Darbha, K. Kalyanam, S. G. Manyam, and D. Casbeer, “Optimal uav route planning for persistent monitoring missions,” *IEEE Transactions on Robotics*, vol. 37, no. 2, pp. 550–566, 2021.

[10] S. K. K. Hari, S. Rathinam, S. Darbha, S. G. Manyam, K. Kalyanam, and D. Casbeer, "Bounds on optimal revisit times in persistent monitoring missions with a distinct and remote service station," *IEEE Transactions on Robotics*, 2022.

[11] S. K. K. Hari, S. Rathinam, S. Darbha, and D. W. Casbeer, "Cooperative coverage with a leader and a wingmate in communication-constrained environments," 2022.

[12] W. Malik, S. Rathinam, and S. Darbha, "An approximation algorithm for a symmetric generalized multiple depot, multiple travelling salesman problem," *Operations Research Letters*, vol. 35, no. 6, pp. 747–753, 2007.

[13] S. G. Manyam, D. W. Casbeer, S. Darbha, I. E. Weintraub, and K. Kalyanam, "Path planning and energy management of hybrid air vehicles for urban air mobility," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10176–10183, 2022.

[14] S. G. Manyam, S. Rathinam, D. Casbeer, and E. Garcia, "Tightly bounding the shortest dubins paths through a sequence of points," *Journal of Intelligent & Robotic Systems*, vol. 88, pp. 495–511, 2017.

[15] S. G. Manyam, S. Rathinam, and S. Darbha, "Computation of lower bounds for a multiple depot, multiple vehicle routing problem with motion constraints," *Journal of Dynamic Systems, Measurement, and Control*, vol. 137, no. 9, 2015.

[16] P. Oberlin, S. Rathinam, and S. Darbha, "A transformation for a multiple depot, multiple traveling salesman problem," in *2009 American Control Conference*, pp. 2636–2641, IEEE, 2009.

[17] P. Oberlin, S. Rathinam, and S. Darbha, "Today's traveling salesman problem," *IEEE Robotics Automation Magazine*, vol. 17, no. 4, pp. 70–77, 2010.

[18] K. J. Obermeyer, P. Oberlin, and S. Darbha, "Sampling-based path planning for a visual reconnaissance unmanned air vehicle," *Journal of Guidance, Control, and Dynamics*, vol. 35, no. 2, pp. 619–631, 2012.

[19] S. Yadlapalli, S. Rathinam, and S. Darbha, "3-approximation algorithm for a two depot, heterogeneous traveling salesman problem," *Optimization Letters*, vol. 6, pp. 141–152, 2012.

[20] A. S. Bhadoriya, C. Montez, S. R. S. Darbha, D. W. Casbeer, and S. G. Manyam, "Assisted shortest path planning for a convoy through a repairable network," *arXiv preprint arXiv:2204.00697*, 2022.

[21] C.-Y. Lee, Z.-J. Lee, S.-W. Lin, and K.-C. Ying, "An enhanced ant colony optimization (eaco) applied to capacitated vehicle routing problem," *Applied Intelligence*, vol. 32, pp. 88–95, 2010.

[22] N. E. Du Toit and J. W. Burdick, "Robot motion planning in dynamic, uncertain environments," *IEEE Transactions on Robotics*, vol. 28, no. 1, pp. 101–115, 2011.

[23] E. Kaufman, T. Lee, and Z. 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)*, pp. 246–251, IEEE, 2016.

[24] T. Zaenker, J. Rückin, R. Menon, M. Popović, and M. Bennewitz, "Graph-based view motion planning for fruit detection," *arXiv preprint arXiv:2303.03048*, 2023.

[25] L. Paull, S. Saeedi, H. Li, and V. 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*, pp. 835–840, IEEE, 2010.

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

[27] C. Potthast and G. S. Sukhatme, "A probabilistic framework for next best view estimation in a cluttered environment," *Journal of Visual Communication and Image Representation*, vol. 25, no. 1, pp. 148–164, 2014.

[28] E. Palazzolo and C. Stachniss, "Effective exploration for mavs based on the expected information gain," *Drones*, vol. 2, no. 1, p. 9, 2018.

[29] C. Stachniss, G. Grisetti, and W. Burgard, "Information gain-based exploration using rao-blackwellized particle filters.," in *Robotics: Science and systems*, vol. 2, pp. 65–72, 2005.

[30] G. Paul, S. Webb, D. Liu, and G. Dissanayake, "Autonomous robot manipulator-based exploration and mapping system for bridge maintenance," *Robotics and Autonomous Systems*, vol. 59, no. 7-8, pp. 543–554, 2011.

[31] P. Quin, G. Paul, A. Alempijevic, D. Liu, and G. Dissanayake, "Efficient neighbourhood-based information gain approach for exploration of complex 3d environments," in *2013 IEEE International Conference on Robotics and Automation*, pp. 1343–1348, IEEE, 2013.

[32] G. Zhang, S. Ferrari, and M. Qian, "An information roadmap method for robotic sensor path planning," *Journal of Intelligent and Robotic Systems*, vol. 56, pp. 69–98, 2009.

[33] J. Denzler and C. M. Brown, "Information theoretic sensor data selection for active object recognition and state estimation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 2, pp. 145–157, 2002.

[34] J. Novakovic, "Using information gain attribute evaluation to classify sonar targets," in *17th Telecommunications forum TELFOR*, pp. 1351–1354, Citeseer, 2009.

[35] D. Deng, R. Duan, J. Liu, K. Sheng, and K. 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)*, pp. 1497–1503, IEEE, 2020.

[36] B. J. Julian, S. Karaman, and D. Rus, "On mutual information-based control of range sensing robots for mapping applications," *The International Journal of Robotics Research*, vol. 33, no. 10, pp. 1375–1392, 2014.

[37] S. Bai, J. Wang, F. Chen, and B. Englot, "Information-theoretic exploration with bayesian optimization," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1816–1822, IEEE, 2016.

[38] F. Amigoni and V. Caglioti, "An information-based exploration strategy for environment mapping with mobile robots," *Robotics and Autonomous Systems*, vol. 58, no. 5, pp. 684–699, 2010.

[39] N. Basilico and F. Amigoni, "Exploration strategies based on multi-criteria decision making for searching environments in rescue operations," *Autonomous Robots*, vol. 31, pp. 401–417, 2011.

[40] G. Laporte, "Fifty years of vehicle routing," *Transportation science*, vol. 43, no. 4, pp. 408–416, 2009.

[41] C. Montez, S. Darbha, C. Valicka, and A. Staid, "Routing of an unmanned vehicle for classification," in *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II* (T. Pham, L. Solomon, and K. Rainey, eds.), vol. 11413, p. 1141319, International Society for Optics and Photonics, SPIE, 2020.