A New Cluster-Based Hybrid Sampling Strategy for Probabilistic Roadmaps in Spatially Varying Wind Fields [1]

Kaushalya Madugalle, Zhe Xu and Salah Sukkarieh
Australian Centre for Field Robotics, The University of Sydney, Australia
[k.madugalle, z.xu, s.sukkarieh]@acfr.usyd.edu.au

Abstract

This paper proposes a workspace-guided and cluster-based sampling strategy to construct a probabilistic roadmap (PRM) for vehicles operating in regions with spatially varying conditions. While the majority of existing PRM planners focus on efficient roadmap construction in configuration spaces with obstacles and narrow passages, we focus on constructing a PRM based on clustering of the free space. Our method is evaluated on an aerial vehicle whose performance is mainly affected by external conditions such as wind strength and wind direction. Comparisons with the canonical uniform sampling strategy are conducted via Monte-Carlo simulations and potential path cost savings achievable using the proposed sampling strategy are demonstrated. These differences are found to be statistically significant at a confidence level of 99.9%.

1 Introduction

Vehicles such as autonomous underwater vehicles (AUVs), unmanned aerial vehicles (UAVs) and aircraft operate in workspaces that contain both obstacles and obstacle-free regions with varying degrees of favourability. The source of such variations may be winds or ocean currents that affect the performance of the vehicle; thus, these variations must be considered during the computation of a minimum cost route between two points. The state space generally contains both discrete and continuous variables, and there are also static and dynamic constraints that must be obeyed. Thus, optimal control methods are computationally infeasible for this class of complex nonlinear time-dependent problems. Methods such as approximate cell decomposition methods, potential field methods and roadmap methods have therefore been introduced for path planning. Of these three approaches, cell decomposition methods and potential field methods pose issues for problems containing more than five dimensions [Karaman and Frazzoli, 2011]. Potential field methods are also likely to get stuck in local minima [Karaman and Frazzoli, 2011]. Meanwhile, roadmap methods that are sampling-based are able to compute near-optimal paths in polynomial time [LaValle, 2006] for problems containing more than five dimensions [Karaman and Frazzoli, 2011].

Commonly used sampling-based techniques include probabilistic roadmap (PRM) planners and rapidly-exploring random tree (RRT) planners [Karaman and Frazzoli, 2011]. RRTs, being single-query and not involving preprocessing [Kuffner and LaValle, 2000] allow path planning in dynamic environments. Meanwhile, PRMs are preferred when information is available for preprocessing [Kuffner and LaValle, 2000]. This study uses useful workspace information known *a priori*; thus, PRMs form the foundation of this work.

Construction of a PRM requires a sampling strategy, a connection strategy and a collision checker. The sampling strategy is the focus of this study as it is the first step in building a PRM that represents the configuration space (C-space). Many existing path planning approaches that use PRMs split a C-space into either free space or occupied space, and mainly focus on C-spaces containing narrow passages. An acceptable minimum cost path is deemed as one which falls in the free space. Meanwhile, only a few studies have diverted from investigating narrow passages and looked at planning in free spaces with varying degrees of favourability [Belghith *et al.*, 2006; Chung *et al.*, 2015].

As opposed to the canonical sampling strategy which samples only based on a uniform distribution, the proposed sampling strategy is based on initial clustering of the workspace. Samples are then drawn from multivariate Gaussian distributions established using data obtained by clustering.

[1] This work was supported by Qantas Airways Limited and the Australian Centre for Field Robotics.

The novel strategy is tested on four time-invariant, two-dimensional (2-D) planar workspaces (wind fields). Comparisons with the canonical sampling strategy are made via Monte-Carlo simulations for an aerial vehicle. The differences between the two methods are shown to be statistically significant at a confidence level of 99.9% using an independent samples t-test for the cases considered.

The remainder of this paper is outlined as follows: Section 2 presents some related work on PRMs; Section 3 provides the problem formulation; Section 4 introduces the two-step algorithm outlining the new sampling strategy; Section 5 discusses the results obtained; and Section 6 concludes this work and lists future research directions.

2 Related Work

Motion planning is an active area of research with many challenges stemming from factors such as high degrees of freedom, time-dependency, uncertainty, static and dynamic constraints and both stationary and moving obstacles. A class of path planning methods known as sampling-based approaches were introduced for problems with high dimensional state spaces and they have seen widespread use for the computation of approximate but efficient solutions [LaValle, 2006; Lindemann and LaValle, 2005; van den Berg and Overmars, 2005; Karaman and Frazzoli, 2011].

A popular sampling-based method is the multi-query PRM planner that was introduced by Kavraki et al. [Kavraki *et al.*, 1996]. It initially constructs a roadmap and then a graph search algorithm is applied to determine the minimum cost path. PRMs display probabilistic completeness [Karaman and Frazzoli, 2011], which means that if a solution exists, the probability that it is found increases to one as the number of samples increases [Barraquand *et al.*, 1996]. The quality of a PRM is based on its connectedness and coverage of the free space [Hsu *et al.*, 1998], and is affected by the strategy initially used for sampling.

Existing PRM sampling techniques mainly attempt to efficiently provide coverage and connectivity in a C-space containing narrow passages and obstacles. Narrow passages pose a bottleneck for efficient PRM construction [van den Berg and Overmars, 2005; Kavraki *et al.*, 1996; Hsu *et al.*, 1997] and many different sampling strategies have been proposed as a consequence. Of these, non-uniform sampling methods have increased efficiency in dealing with narrow passages. For instance, workspace-guided strategies sample in such a way that information about the workspace is utilised [Kurniawati and Hsu, 2004; van den Berg and Overmars, 2005]. These strategies require the C-space to resemble the workspace [van den Berg and Overmars, 2005]. As a close mapping between the C-space and the workspace can be achieved for the vehicles addressed in this study, workspace-guided strategies are applicable. Other examples of workspace-guided methods include those that employ cell-decomposition techniques to infer knowledge of the workspace [van den Berg and Overmars, 2005; Siméon *et al.*, 2004; Kurniawati and Hsu, 2006]. Meanwhile, another non-uniform sampling method is Gaussian sampling that biases samples to be closer to obstacles in the C-space [Boor *et al.*, 1999]. This strategy can similarly be used to bias sampling towards favourable regions of a workspace.

The diversity in the strengths and weaknesses of different sampling strategies according to problem-specific factors has lead to the introduction of hybrid strategies [Morales *et al.*, 2005]. These seek to combine different sampling strategies based on the space being sampled to generate a roadmap [Hsu *et al.*, 2005]. For instance, the use of uniform sampling along with a bridge test, which increases the sampling density in narrow passages, was introduced to tackle the problem of poor connectivity in narrow passages [Hsu *et al.*, 2003]. A machine learning-based feature-sensitive motion planner was also introduced which partitions the C-space and then determines the best sampler to use for each partition to construct a roadmap. Afterwards, all the small roadmaps are combined together to obtain the complete roadmap [Morales *et al.*, 2005].

With planning in narrow passages being the focus of most of the PRM literature, only few studies have focused on planning in workspaces with a free space that varies in favourability. The problem investigated in this study was also addressed by Chung et al. [Chung *et al.*, 2015]; however, instead of using a clustering algorithm as proposed in this work, Gaussian processes were utilised to bias sampling and build a PRM. Furthermore, the presence of obstacles was ignored. Another study that focused on varying free space favourability proposed a sampling strategy targeted at replanning applications in dynamic environments [Belghith *et al.*, 2006]. This strategy randomly samples free regions and guides a robot to zones with higher desirability and minimises the distance to a given goal configuration. Due to the focus on replanning, sampling is done by first selecting a node based on a priority value. However, in the method proposed in this paper, node selection is not required as replanning is not the aim.

3 Problem Formulation

The problem of concern in this study is path planning for an aerial vehicle operating in a 2-D workspace represented by a time-invariant wind field. Obstacles and free regions are differentiated by the wind strength acting at each point in the workspace. An obstacle is

classified as a region composed of winds with strengths greater than what is deemed as safe for the chosen aerial vehicle, while free regions are those regions that remain. The tailwind and crosswind components acting in the free regions where the aerial vehicle will traverse through influence its flight time and energy consumption. Tailwinds are beneficial as they are the wind components that lie in the direction of travel. Meanwhile, crosswinds are the wind components that act perpendicular to the direction of travel and deviate the aerial vehicle from its path. Hence, a minimum cost path lies in regions with high tailwinds and low crosswinds. In this study, these regions are defined as 'favourable' regions. The minimum cost path p^* is then expressed as

$$p^* = \min_p \int_{t_i}^{t_f} E(t)\ dt \quad \text{s.t.} \quad \left\|\vec{V}\right\| < V_{max}, \tag{1}$$

where t_i and t_f denote the initial and final time, respectively, $\vec{V}$ represents each wind vector acting along the path and V_{max} denotes the maximum wind magnitude that is safe for the vehicle to encounter. Meanwhile, E represents energy and has the following dependence:

$$E = f(x, y, m, \vec{V}), \tag{2}$$

where x and y represent the 2-D position and m denotes the gross weight of the vehicle. The function f is vehicle-specific and outputs the amount of energy consumed by considering the weight and speed of the vehicle travelling in a given environment.

4 A New Cluster-Based Hybrid Sampling Strategy

The approach used to achieve the objectives of this study is outlined in a two-step algorithm (Algorithm 1). Firstly, the sampling strategy proposed identifies varying degrees of favourability in the workspace via clustering (Algorithm 2). This information is then used to bias sampling towards favourable regions using multivariate Gaussian distributions (Algorithm 3).

Algorithm 1: Minimum cost path computation

input : workspace data set $\mathcal{W}$, no. of clusters K, no. of nodes N_{max}, level of dependence on favourable clusters α, start location s, end location e
output: minimum cost path P, path costs P_{costs}
$[C, \tilde{C}] \leftarrow$ perform-clustering$(\mathcal{W}, K)$
//see Algorithm 2 ;
$A \leftarrow$ generate-PRM$(\mathcal{W}, N_{max}, C, \tilde{C}, \alpha, s, e)$
//see Algorithm 3 ;
$[P, P_{costs}] \leftarrow$ perform-graph-search(A, s, e);

Algorithm 2 performs clustering using the squared Euclidean distance measure and the widely used k-means clustering algorithm, which is suitable for problems with quantitative variables [Hastie *et al.*, 2001]. Note that while each workspace has an optimal number of clusters K_{opt}, finding this value was beyond the scope of this work.

A feature set $\mathcal{F}$ that is appropriate to the workspace (i.e., affects the performance of the vehicle and influences path costs) is used for clustering, and is defined as

$$\mathcal{F} = \{\vec{\mathcal{F}}_i\} \quad \text{for} \quad i = 1, 2, ..., F, \tag{3}$$

where $\vec{\mathcal{F}}$ represents each feature vector and F denotes the total number of features. $\mathcal{F}$ is dependent on the workspace data set $\mathcal{W}$ given as

$$\mathcal{W} = \{\vec{x}, \vec{y}, \vec{V_x}, \vec{V_y}\}, \tag{4}$$

where $\vec{x}$ and $\vec{y}$ represent vectors containing x and y position data, respectively, and $\vec{V_x}$ and $\vec{V_y}$ represent vectors containing wind speed data in the x and y directions, respectively.

The output of Algorithm 2 is information regarding the conditions of the clustered workspace, assigned as cluster data C and its normalised version $\tilde{C}$. C contains results gathered via the k-means clustering algorithm and is represented by the following matrix:

$$C = \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,F} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,F} \\ \vdots & \vdots & \ddots & \vdots \\ c_{K,1} & c_{K,2} & \cdots & c_{K,F} \end{pmatrix} \in \mathbb{R}^{K \times F}, \tag{5}$$

where K denotes the total number of clusters. Here, each element $c_{k,f}$ is the value of the cluster centroid w.r.t. each feature $1 \leq f \leq F$ in cluster $1 \leq k \leq K$.

Algorithm 2: Identification of workspace conditions via clustering

input : workspace data set $\mathcal{W}$, no. of clusters K
output: cluster data C, normalised cluster data $\tilde{C}$
$\mathcal{F} \leftarrow \emptyset$; $\tilde{C} \leftarrow \emptyset$;
foreach $w \in \mathcal{W}$ **do**
  $f_w \leftarrow$ find all features;
  $\mathcal{F}_w \leftarrow \{f_w\}$;
  $\mathcal{F} \leftarrow \mathcal{F} \cup \mathcal{F}_w$
end
$C \leftarrow$ perform-clustering$(\mathcal{F}, K)$;
foreach $k \leftarrow 1, K$ **do**
  $\tilde{C}_k \leftarrow$ normalise-data(C);
  $\tilde{C} \leftarrow \tilde{C} \cup \tilde{C}_k$;
end

To compute $\tilde{C}$, a min-max normalisation as described by [Han *et al.*, 2011] is done on each feature (i.e., column) of matrix C. Equal weights are assigned to the features and then summed to obtain $\tilde{C}$ which is given as

$$\tilde{C} = [\tilde{C}_1 \ ... \ \tilde{C}_k \ ... \ \tilde{C}_K], \tag{6}$$

where $\tilde{C}_k \in [0, 1]$ and corresponds to the normalised cluster data for each cluster k.

Algorithm 3 presents the new sampling strategy which uses both C and $\tilde{C}$. A parameter termed $\alpha \in [0, 1]$ is introduced to control sampling towards favourable clusters. Specifically,

$$\alpha \in [\min(\tilde{C}), \max(\tilde{C})]. \tag{7}$$

Once a suitable α is selected, it is used as a threshold to determine the number of samples N_k for each cluster k. This calculation also uses the user-defined number of nodes N_{max} and is described as follows:

$$N_k = \begin{cases} \tilde{C}_k N_{max}, & \text{if } \tilde{C}_k \geq \alpha \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

Note that the default value for α is 0, which refers to the least level of dependence on favourable clusters. As α increases, less favourable clusters are ignored and sampling is biased towards the more favourable clusters. The final number of samples N_{final} to be generated in the workspace based on α never exceeds N_{max} and is simply

$$N_{final} = \sum_{k=1}^{K} N_k. \tag{9}$$

A total of N_k samples q' are then generated within each cluster based on a multivariate Gaussian distribution. In this study, each distribution is a 2-D elliptical Gaussian with a diagonal covariance matrix Σ containing direction and distance as independent parameters. For each cluster k,

$$q' \leftarrow \mathcal{N}(\mu_k, \Sigma_k), \tag{10}$$

where

$$\mu_k = \begin{bmatrix} \mu_{k,direction} \\ \mu_{k,distance} \end{bmatrix} \tag{11}$$

and

$$\Sigma_k = \begin{bmatrix} \sigma^2_{k,direction} & 0 \\ 0 & \sigma^2_{k,distance} \end{bmatrix}, \tag{12}$$

where μ_k and σ^2_k denote the mean and variance, respectively, of the data points with regard to direction and distance within each cluster. This sampling process ensures good cluster coverage by generating samples all across a cluster. The strategy also increases the likelihood of samples being located in favourable regions instead of in completely random locations as per the canonical sampling strategy.

Algorithm 3: A new cluster-based hybrid PRM sampling strategy for path planning

input : workspace data set $\mathcal{W}$, no. of nodes N_{max}, cluster data C, normalised cluster data $\tilde{C}$, level of dependence on favourable clusters α, start location s, end location e
output: probabilistic roadmap A

```
M ← [s; e]                                   //milestone array;
if α ≠ 0 then
    α ← calculate-α(C̃);
end
foreach k ← 1, K do
    check if cluster centroid centroid_k is
    collision-free using W;
    if centroid_k is collision-free then
        M_q ← [centroid_k];
        M ← M ∪ M_q;
        N_k ← calculate-number-of-samples-
        in-cluster(C̃_k, N_max, α) ;
        foreach j ← 1, N_k do
            repeat
                q'_j ← generate-sample(C);
            until q'_j is collision-free;
            M_q' ← [q'_j];
            M ← M ∪ M_q';
        end
    else
        disregard k;
    end
end
A ← connect-milestones(M);
```

Collision-free samples termed 'milestones' are retained for connection via Delaunay triangulation (DT). The choice of DT arises due to the fundamental dependence of the new sampling strategy on clustering. As N_{max} increases, the concentration of samples to be generated within a cluster increases. A connection strategy such as the widely used k-nearest neighbour connection technique, which is dependent on a user-defined maximum distance to neighbours and maximum number of neighbours, leads to an increase in intra-cluster connections and a decrease in inter-cluster connections. Consequently, the overall coverage and connectivity of a PRM produced using the new sampling strategy deteriorates. Hence, both inter- and intra-cluster connections are vital in PRM construction.

Finally, a graph search algorithm is executed on the PRM generated by Algorithm 3 to determine the minimum cost path between two points in the workspace.

5 Results and Discussion

This section presents a comparative discussion of the canonical and new sampling strategies using results obtained via simulations conducted on four different wind fields. Note that wind vectors are denoted by blue arrows in each figure.

5.1 Canonical Sampling Strategy

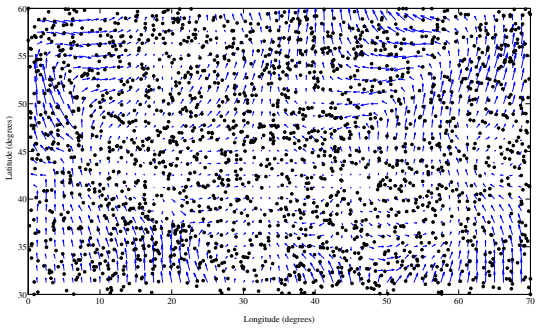

Figure 1: A workspace with 2000 nodes generated using the canonical sampling strategy. Uniform sampling and a high number of nodes allows good coverage of the workspace; however, workspace conditions such as wind strength and wind direction that affect the path cost are ignored.

Figure 1 illustrates milestones generated using the canonical sampling strategy for 2000 nodes. This strategy, which follows a uniform distribution, is only focused on collision avoidance instead of also traversing in *'favourable'* collision-free regions. It is evident how milestones can be located anywhere within the free space with equal probability, and with no consideration of aspects such as wind direction and wind speed.

5.2 Cluster-Based Hybrid Sampling Strategy

The squared Euclidean distance measure was used to compute the similarity between two data vectors of the feature data set. This distance measure presented better clustering results for the wind data available as compared to other distance measures such as L_1 distance (sum of absolute differences) and cosine similarity.

Figure 2 illustrates the clusters obtained on a particular region with spatially varying wind vectors using the k-means clustering algorithm, where the number of clusters K was 20. The choice of features for clustering is important to avoid clusters spanning over large areas and being located within other clusters. The clusters shown in Figure 2 were generated based on three features applicable for an aerial vehicle, namely 2-D location data, tailwind data and crosswind data. After these features were calculated, they were assigned the same weighting for use by the k-means clustering algorithm.

5.3 Level of Dependence on Favourable Clusters α

The influence of parameter α on sampling can be seen in Figures 3 and 4.

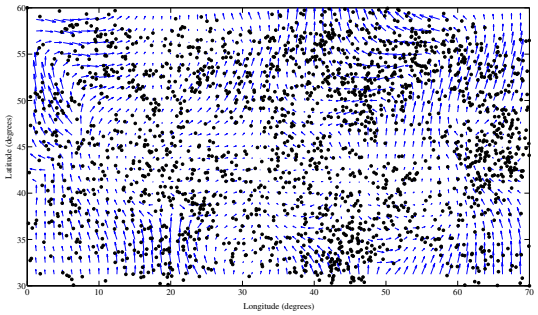

Figure 3: Milestone generation with $\alpha = 0$ and $N_{max} = 2000$. This is the least level of dependence on favourable clusters and milestones are generated across the workspace. Note that the locations of these milestones are biased to be inside the underlying clusters.

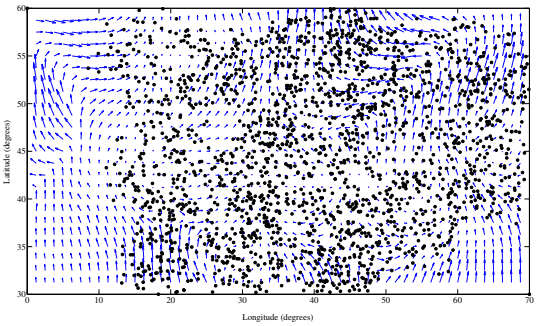

Figure 4: Milestone generation with $\alpha = Q(1)$ and $N_{max} = 2000$. Less favourable clusters are discarded and sampling is biased towards the more favourable regions of the workspace.

In Figure 3, α was fixed at the default value of 0 and milestones were generated across the workspace. The number of samples allocated to each cluster in the workspace was a function of the features acting within that cluster. Unlike the sampled workspace shown in Figure 4, milestones were located across more clusters overall. In Figure 4, α was set to be equal to the value at the 25^{th} percentile ($Q(1)$ or 1^{st} Quartile) of the normalised cluster data corresponding to the workspace. It is evident that as α increased, more of the weaker clusters were discarded; instead, samples were guided to more favourable regions of the workspace. This consequently allowed improved path costs to be determined as can be seen in Figure 7 for both $N_{max} = 500$ and $N_{max} = 2000$.

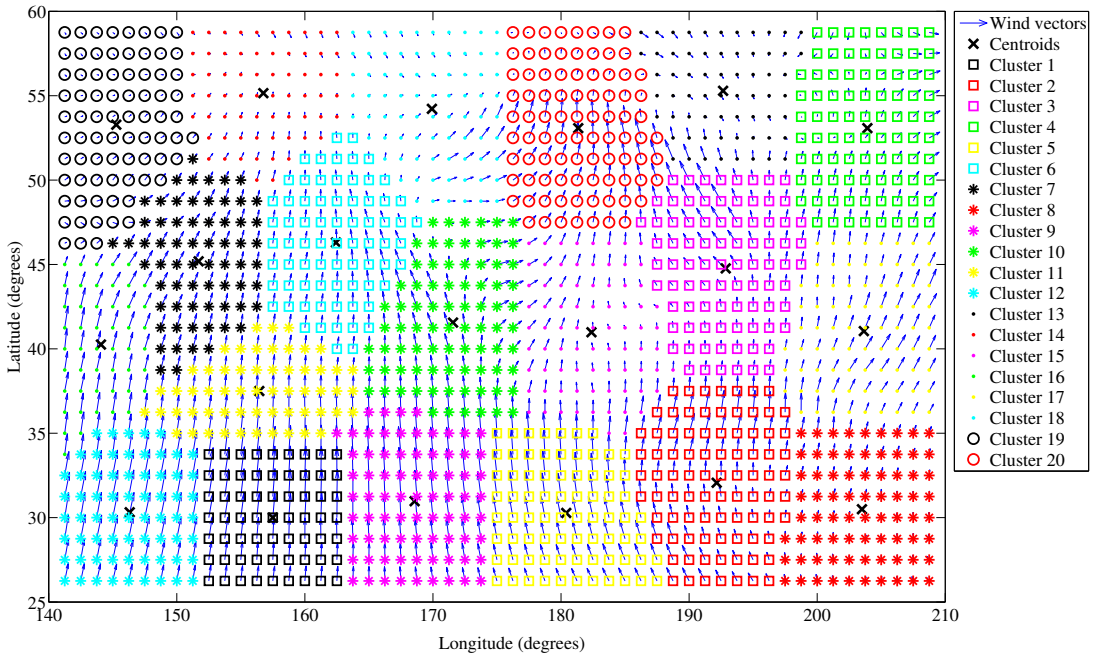

Figure 2: A clustered wind field with $K = 20$.

5.4 Comparisons

Figures 5 and 6 show the minimum energy cost path between two points on a PRM generated using the canonical sampling strategy and the new sampling strategy, respectively, in a given wind field. In both figures, N_{max} was 100 and the milestones were connected using DT. Note that while better connection strategies may exist, gauging their performance with the proposed sampling strategy was not within the scope of this research.

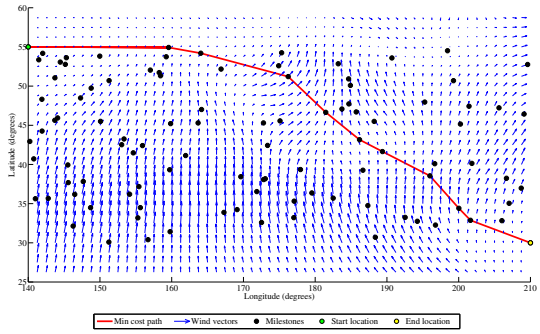

Figure 5: Minimum cost path determined using the canonical sampling strategy; normalised minimum cost = 0.5738; $N_{max} = 100$.

As evident, more milestones were generated in regions with high tailwinds and low crosswinds using the new method (Figure 6) as opposed to the uniformly random approach used by the canonical sampling strategy (Figure 5). Hence, there is higher probability that the minimum cost path determined by the new method is better than that found by the canonical method. This can be seen by the lower normalised minimum cost determined using the new method as compared to the canonical method. The resulting path cost improvement using the new method was 1.48%.

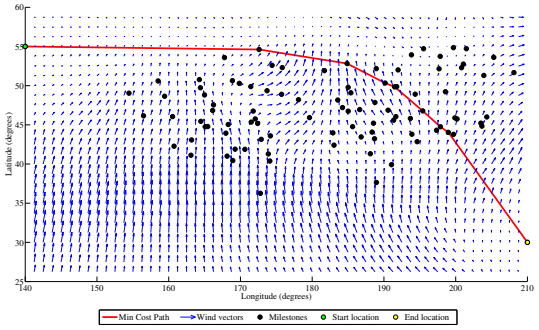

Figure 6: Minimum cost path determined using the new sampling strategy; normalised minimum cost = 0.5653; $N_{max} = 100$.

Case studies were conducted on four different wind fields (see Figures 8-11 in the Appendix) of approximately the same area. The same number of clusters ($K = 20$) was used for each workspace.

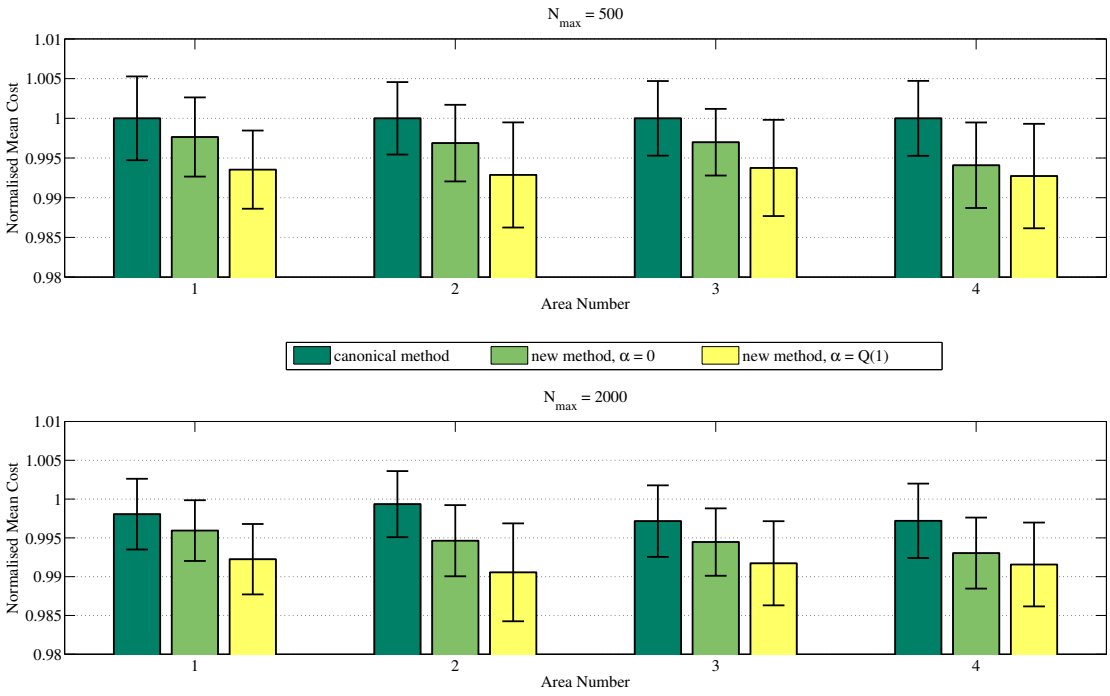

Figure 7: Comparison between normalised mean cost obtained via the canonical and new strategies with $N_{max} = 500$ and $N_{max} = 2000$ using 100 Monte-Carlo simulations.

Monte-Carlo simulations were carried out, where the minimum energy cost path in each wind field was simulated 100 times. The start and end points were kept constant for each workspace and Dijsktra's algorithm was used as the graph search algorithm. Figure 7 illustrates results of normalised mean energy consumption obtained with 500 and 2000 nodes on each of the four wind fields using both the canonical and new sampling strategies. Two levels of dependence (α) were considered. In the first case, α was set at zero. In the second case, α was set to equate the value at the 1^{st} Quartile of the normalised workspace cluster data.

As can be seen, the new sampling strategy (for both α values) produced lower mean costs as compared to the canonical method in all areas for both 500 and 2000 nodes. It is evident that when $\alpha = Q(1)$, further improvements in mean cost could be achieved as samples were now allocated to more of the favourable regions where the minimum cost path was likely to lie on. It must be noted that an optimal value for α exists for each region, which if exceeded will adversely affect the performance of the new method due to excessive pruning. In all of the areas tested, $\alpha = Q(1)$ produced better mean costs than $\alpha = 0$.

It can also be seen that the mean costs for $N_{max} = 2000$ were lower (i.e., better) than those for $N_{max} = 500$ for both sampling strategies. This indicates that as the number of nodes increased, the coverage and connectivity of the resulting PRMs increased in quality causing the minimum mean costs determined to improve. This provides evidence of the probabilistic completeness nature of PRMs.

The standard deviation error bars show the uncertainty inherent in the two PRM planners. The canonical method and the new method with $\alpha = 0$ had similar levels of uncertainties (as expressed by the length of the error bars) for each of the areas. However, the error bars corresponding to the new method with $\alpha = Q(1)$ were longer for seven of the eight scenarios than the canonical method and the new method with $\alpha = 0$. This behaviour was likely due to the choice of the number of clusters K. As α increases, the dependence of the new method on clustering increases. As in this study the value chosen for K was not optimal, the performance of the new method became more uncertain when further pruning was introduced (i.e., $\alpha > 0$). This indicates the importance of the proper selection of K and α.

In general, as the number of nodes increased the uncertainty decreased for both planners. Furthermore, each upper and lower error (i.e., maximum and minimum standard deviation from the mean, respectively) was smaller for the new method than the canonical method in all of the cases, meaning that the probability of obtaining improved costs using the new method is higher. This is true even for the case of $\alpha = Q(1)$, which showed higher uncertainty than the canonical method and the

new method with $\alpha = 0$.

Note that the data in Figure 7 were subjected to a 90% Winsorization (i.e., the bottom 5% and top 5% of the data were classed as outliers and were set to be equal to the value at the 5^{th} percentile and 95^{th} percentile, respectively). However, the general trends evident in these figures remain the same both with and without Winsorization. The only difference arises from the length of the error bars, which decreases with Winsorization due to the resulting reduction in uncertainty.

An independent samples t-test was performed on the data contained in Figure 7 to determine the significance of the differences between the canonical and new sampling strategies. The null (H_0) and alternative hypotheses (H_1) were set as follows:

- $H_0 : \mu_1 = \mu_2$
- $H_1 : \mu_1 \neq \mu_2$

where μ_1 and μ_2 represent the normalised mean of the canonical method and that of the new method, respectively. If H_1 is proven to be true, the results obtained can be regarded as being significant.

The test statistic (t) was calculated and a confidence level of 99.9% was used for hypothesis testing. The results are summarised as follows:

1. Canonical vs New ($\alpha = 0$; default)
   - $3.2416 \leq t \leq 8.2504$ for all areas
2. Canonical vs New ($\alpha = Q(1)$)
   - $8.1509 \leq t \leq 12.1304$ for all areas

These t values indicated H_1 as being true at a confidence level of 99.9% for both cases; thus, the results obtained were significant.

It must be noted that if the degree of wind variability in a given free-flight region is low, then the canonical sampling strategy will be more computationally efficient. This is because clustering the workspace does not provide much benefit in this scenario as the problem reverts back to the classical black and white problem composed of obstacles and free space with uniform favourability.

6 Conclusions and Future Work

This paper introduced a new cluster-based hybrid sampling strategy that is workspace-guided and utilises multivariate Gaussian distributions to build a probabilistic roadmap (PRM). This strategy is suitable for vehicles operating in workspaces with both obstacles and free-spaces composed of varying degrees of favourability. As opposed to typical robot motion planning problems that are generally concerned on obstacle avoidance and planning in narrow passages, problems involving vehicles such as manned aircraft, unmanned aerial vehicles and autonomous underwater vehicles that are mostly affected by environmental conditions should be formulated to consider such conditions during path planning.

Monte-Carlo simulations were conducted for an aerial vehicle operating in four workspaces with different wind fields using the sampling strategy proposed and the canonical sampling strategy. Statistically significant cost savings at a confidence level of 99.9% using the new method were made evident via an independent samples t-test. Further cost improvements were possible with the use of a non-zero variable level of dependence on favourable clusters (termed parameter α) as opposed to when it was set to zero. In conclusion, the new strategy showed much benefit for vehicles operating in workspaces with obstacles and free-spaces with spatially varying conditions.

Future work will focus on extending the current approach that only concerns a 2-D region to encompass different altitude levels to reflect the actual 3-D problem. This simplification made for the ease of algorithmic implementation indicates the potentially higher benefits that can be expected when the search is extended to 3-D space. The influence and selection of α will be investigated further, as well as its optimality that is specific to a given workspace. In addition, the sampling strategy will be modified to consider multiple wind fields known *a priori* to account for the temporal aspect of wind. Selection of the most appropriate number of clusters for a given workspace will also be studied to strengthen the new strategy. Finally, as the performance of a PRM planner and its probabilistic completeness guarantees are not only a function of the sampling strategy, a connection strategy that is tailored for the proposed cluster-based hybrid sampling strategy will be developed.

Appendix

The wind fields used in this study are presented here.

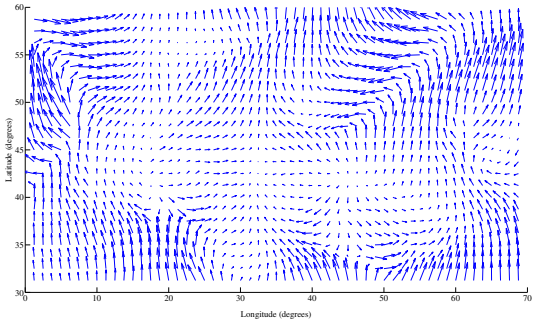

Figure 8: Wind field for Area 1.

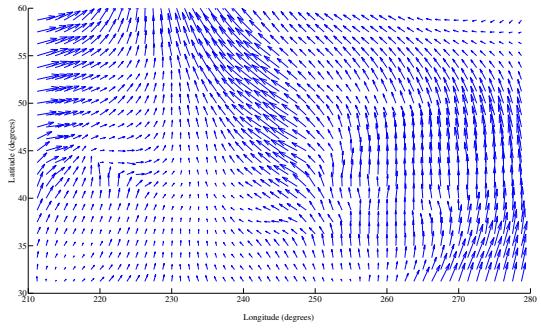

Figure 9: Wind field for Area 2.

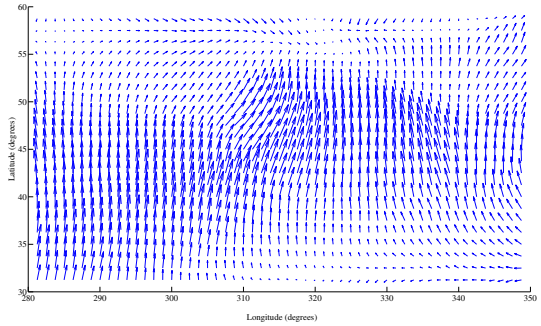

Figure 10: Wind field for Area 3.

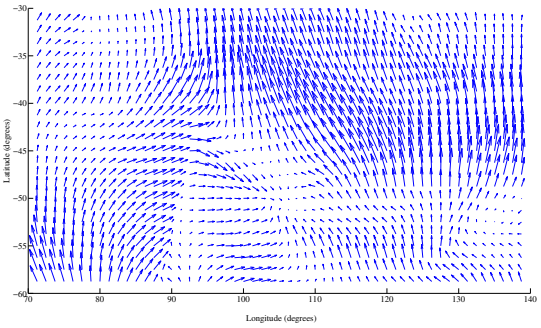

Figure 11: Wind field for Area 4.

References

[Barraquand *et al.*, 1996] Jérôme Barraquand, Lydia E Kavraki, Rajeev Motwani, Jean-Claude Latombe, Tsai-Yen Li, and Prabhakar Raghavan. A random sampling scheme for path planning. In *Proceedings of the 7th International Symposium on Robotics Research*, pages 249–264. Springer, 1996.

[Belghith *et al.*, 2006] Khaled Belghith, Froduald Kabanza, Leo Hartman, and Roger Nkambou. Anytime dynamic path-planning with flexible probabilistic roadmaps. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 2372–2377. IEEE, 2006.

[Boor *et al.*, 1999] Valdrie Boor, Mark H Overmars, and Others. The Gaussian sampling strategy for probabilistic roadmap planners. In *Proceedings of IEEE International Conference on Robotics and Automation*, volume 2, pages 1018–1023. IEEE, 1999.

[Chung *et al.*, 2015] Jen J Chung, Nicholas Lawrance, Seng K Gan, and Zhe Xu. Variable Density PRM Waypoint Generation and Connection Radii for Energy-Efficient Flight through Wind Fields. In *Proceedings of IEEE International Conference on Robotics and Automation*, number 3, pages 2–4, 2015.

[Han *et al.*, 2011] Jiawei Han, Micheline Kamber, and Jian Pei. *Data mining: concepts and techniques*. Elsevier, 2011.

[Hastie *et al.*, 2001] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*, volume 2nded. Springer-Verlag, Heidelberg, 2001.

[Hsu *et al.*, 1997] David Hsu, Jean-Claude Latombe, and Rajeev Motwani. Path planning in expansive configuration spaces. In *Proceedings of IEEE International Conference on Robotics and Automation*, volume 3, pages 2719–2726. IEEE, IEEE, 1997.

[Hsu *et al.*, 1998] David Hsu, Lydia E Kavraki, Jean-Claude Latombe, Rajeev Motwani, Stephen Sorkin, and Others. On finding narrow passages with probabilistic roadmap planners. In *Proceedings of the 1998 Workshop on the Algorithmic Foundations of Robotics*, pages 141–154, 1998.

[Hsu *et al.*, 2003] David Hsu, Tingting Jiang, John Reif, and Zheng Sun. The bridge test for sampling narrow passages with probabilistic roadmap planners. In *Proceedings of IEEE International Conference on Robotics and Automation*, volume 3, pages 4420–4426. IEEE, 2003.

[Hsu *et al.*, 2005] David Hsu, Gildardo Sánchez-Ante, and Zheng Sun. Hybrid PRM sampling with a cost-sensitive adaptive strategy. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 3874–3880. IEEE, IEEE, 2005.

[Karaman and Frazzoli, 2011] Sertac Karaman and Emilio Frazzoli. Sampling-based algorithms for optimal motion planning. *The International Journal of Robotics Research*, 30(7):846–894, 2011.

[Kavraki *et al.*, 1996] Lydia E Kavraki, Petr Švestka, Jean-Claude Latombe, and Mark H Overmars. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Transactions on Robotics and Automation*, 12(4):566–580, 1996.

[Kuffner and LaValle, 2000] James J Kuffner and Steven M LaValle. RRT-connect: An efficient approach to single-query path planning. In *Proceedings of IEEE International Conference on Robotics and Automation*, number April, pages 995–1001. IEEE, 2000.

[Kurniawati and Hsu, 2004] Hanna Kurniawati and David Hsu. Workspace importance sampling for probabilistic roadmap planning. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 2, pages 1618–1623. IEEE, IEEE, 2004.

[Kurniawati and Hsu, 2006] Hanna Kurniawati and David Hsu. Workspace-based connectivity oracle: An adaptive sampling strategy for PRM planning. In *Proceedings of the International Workshop on the Algorithmic Foundations of Robotics*. SpringerVerlag, 2006.

[LaValle, 2006] Steven M LaValle. *Planning algorithms*. Cambridge, 2006.

[Lindemann and LaValle, 2005] Stephen R Lindemann and Steven M LaValle. Current issues in sampling-based motion planning. In *Robotics Research. The Eleventh International Symposium*, pages 36–54. Springer, Springer Berlin Heidelberg, 2005.

[Morales *et al.*, 2005] Marco Morales, Lydia Tapia, Roger Pearce, Samuel Rodriguez, and Nancy M Amato. A machine learning approach for feature-sensitive motion planning. In *Algorithmic Foundations of Robotics VI*, pages 361–376. Springer Berlin Heidelberg, 2005.

[Siméon *et al.*, 2004] Thierry Siméon, Jean-Paul Laumond, Juan Cortés, and Anis Sahbani. Manipulation planning with probabilistic roadmaps. *The International Journal of Robotics Research*, 23(7-8):729–746, 2004.

[van den Berg and Overmars, 2005] Jur P van den Berg and Mark H Overmars. Using workspace information as a guide to non-uniform sampling in probabilistic roadmap planners. *The International Journal of Robotics Research*, 24(12):1055–1071, 2005.