Choosing Landmarks for Risky Planning

Liz Murphy[1], Peter Corke[1] and Paul Newman[2]

Abstract— **This work examines the effect of landmark placement on the efficiency and accuracy of risk-bounded searches over probabilistic costmaps for mobile robot path planning. In previous work, risk-bounded searches were shown to offer in excess of 70% efficiency increases over normal heuristic search methods. The technique relies on precomputing distance estimates to *landmarks* which are then used to produce probability distributions over exact heuristics for use in heuristic searches such as A* and D*. The location and number of these landmarks therefore influence greatly the efficiency of the search and the quality of the risk bounds. Here four new methods of selecting landmarks for risk based search are evaluated. Results are shown which demonstrate that landmark selection needs to take into account the centrality of the landmark, and that diminishing rewards are obtained from using large numbers of landmarks.**

I. INTRODUCTION

Through judicious use of a-priori available information large performance gains in path planning for mobile robots can be obtained. In previous work [1] it was shown how probabilistic costmaps can be generated from overhead imagery; these costmaps can be used to precompute heuristics for search algorithms such as A* and D*, wherein the probabilistic nature of the heuristic can be exploited to tradeoff the chance of the shortest path being returned against the speed of the search [2]. This tradeoff is done in a controlled fashion, the user specifies the risk they wish to accept that the algorithm will not return the shortest path.

These heuristics are obtained using a probabilistic variation of A* search, Landmark and Triangle inequality (ALT) technique [3], which precomputes distances between every grid cell in the costmap to a select number of *landmarks* scattered throughout the map. The precomputed distances to the landmarks are used in conjunction with the triangle inequality to provide tight lower bounds on the heuristic estimate on the distance between any node and the goal node during an A* search. The result is a probability distribution over an exact heuristic suitable for use in the A*/D* family of heuristic searches. The R^*_δ algorithm [4] operates within the framework of A* search to transform this probability distribution over the exact distance to a scalar value which is used to determine the order of node expansion during the search; this transformation takes into account the user specified risk governing the likelihood of returning the shortest path. In this work we propose techniques to optimally site the landmarks to best satisfy the separate but coupled demands of limiting node expansions for individual searches (efficiency) and ensuring the probability distribution we calculate over the exact heuristic is accurate (accuracy), in the domain of probabilistic search.

In [2], we followed the recommendation of [3] and used *planar landmark selection* to site landmarks for the ALT search. As in [3], who were constrained by memory capabilities, we used 16 landmarks for all costmaps. This naive technique places the landmarks at regular intervals around the border of the map. Yet it works well in the traditional case (when grid cell values are scalar), because the reverse triangle inequality produces tighter lower bounds when the landmark is located behind the goal (see Figure **??**). However, in risky planning with probabilistic costmaps we have an additional consideration that must be taken into account — the variance of the precomputed path. *Planar landmark selection* can give us a good approximation to the mean of the probability distribution over the exact heuristic but the probability guarantees of the R^*_δ search are only as reliable as the probability density function we use to model the exact distance and that must take into account the variance. In previous work we found that planar landmark selection lead to an overestimation of the variance, and thus impaired the efficiency of the more risk-averse path planning strategies we examined.

The contribution of this this work lies in a thorough investigation into the effect of both landmark placement and the number of landmarks used on the risk guarantees and efficiency of risky planning with probabilistic costmaps. We aim to extend the work in [2] where planar landmark selection was shown to work well for planning traverses between cells located in the outer 30% of a large-scale grid map, so that heuristic estimates are accurate regardless of where in the map the start and goal nodes for the search lie. Here, four techniques of landmark selection are compared with the baseline planar landmark selection technique. For the first time, we challenge the assumption that using more landmarks leads to superior performance. We also show that good landmarks have a physical analog in overhead imagery, thus choosing landmarks with the human eye could serve to bypass expensive precomputation.

II. RELATED WORK

The most prominent path planning algorithms for mobile robots [5] [6] [7] are built on the foundations of the A* search [8]. A* is a heuristic search technique, for each node it computes an estimate of the path length from start to goal of the best path that passes through that node; this estimate determines the order in which nodes of the graph being searched are visited. This heuristic itself has two parts, one is

[1] School of Engineering Systems, Queensland University of Technology, Brisbane, Australia liz.murphy@qut.edu.au
[2] Mobile Robotics Group, Oxford University, Oxford, UK

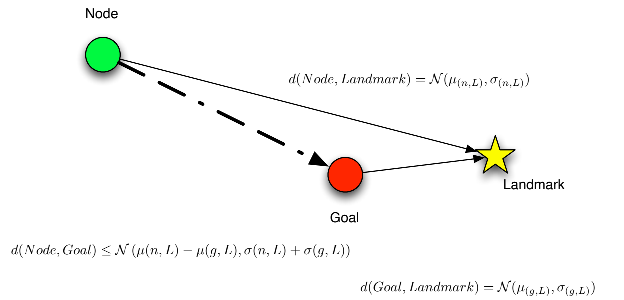


Fig. 1: Landmarks located behind the Goal Node produce good estimates of the mean of a probability distribution over the node-goal distance (dotted line). Conversely, they tend to grossly overestimate the variance as under the additive property of the Gaussian distribution the variances of two much longer paths than that we desire are added together to produce the variance estimate.

the cost of travelling from the start node to the current node n under examination, usually denoted $g(n)$. The second part, denoted $h(n)$ is a heuristic estimate of the distance from n to the goal. The sum $f(n) = g(n) + h(n)$ is used to order the priority queue governing the expansion of the search. The speed and accuracy of the search hinges on the accuracy of the estimate $h(n)$ provided to the algorithm. The straight-line (Euclidean) distance estimate is commonly used as the heuristic, but it fails to take into account the presence of obstacles and dangerous terrain for the robot and is much less efficient than using an exact heuristic over the same search area.

Typically under the A* paradigm the heuristic must be both consistent (monotonic) and admissible, meaning it never overestimates the cost of reaching the goal. It is the admissibility condition that guarantees that A* will find an optimal solution path if it exists. Without setting the heuristic to zero and degenerating to an inefficient Dijkstra shortest-path search, it is impossible to postulate an admissible heuristic for a continuous probability distribution.

An unfortunate by-product of an admissible heuristic is that it can frequently cause the search to spend large amounts of time deliberating between roughly equal solution paths and does not give us the option of terminating the search with an acceptable but not optimal solution path. In *Bandwidth Search* [9] Harris showed that if the (scalar) heuristic overestimation is bounded by ε then the resulting cost cannot exceed $(1+\varepsilon)C^*$, where C^* is the optimal path cost. The concept of ε*-admissibility* has more recently been used in [10] to produce bounded suboptimal solutions which produce increasingly better solutions to robotic path planning problems in a given time window.

Pearl [4] extended this approach to probabilistic heuristics; seeking to invoke likelihood considerations into the admissibility guarantee and allowing the user to control the *risk* of not obtaining the shortest path from the search. The resulting R^*_δ algorithm is a variant of A* that relaxes Harris' $\varepsilon - admissibility$ condition even further, to allow the use of a precise estimator that may occasionally overestimate the exact heuristic distance h^* by more than ε, such as is the case when the arc costs of a graph are known to be drawn from a probability distribution. This technique was demonstrated on an instance of the Travelling Salesman Problem with statistics generated from multiple previous searches.

In [2] we applied R^*_δ search to the problem of path planning over probabilistic costmaps for mobile robots. We showed that up to 70% efficiency increases over normal heuristic search methods can be obtained. In order to precompute the probabilistic heuristic required by R^*_δ we used the ALT algorithm [3] [11]. It is one of many algorithms [12] [13] [14] recently derived which focus on preprocessing the map to obtain better heuristics and in so doing restrict the point to point search to examining only relevant areas of the input graph.

Potomias et al. [15] studied the problem of selecting landmarks for approximate shortest path searches and proved that the problem of optimal landmark selection is NP hard. They validated experimentally that choosing landmarks with high centrality [16] [17] best served the approximate shortest path problem. In terms of the exact shortest path problem (of which A* and D* are instances), in [18] the problem of landmark selection is framed as case of the unconstrained facility location problem. Here n landmarks are chosen from a larger candidate set such that the aggregate cost of serving search queries from a large set of start-goal pairs over the map is minimized and solutions are obtained using a k-median approach.

III. RISKY PLANNING

In this section we provide a brief recap of the Risky Planning process used in [2]. We begin with a probabilistic costmap, computed from a-priori information about the environment we wish to plan over. We assume the cost of traversing each grid cell is represented as a probability distribution over cost, this relates directly to the uncertainty in our knowledge of the underlying terrain represented in the grid cell. If we take the grid cell values to be Gaussian probability distributions, we can use the additivity property of the Gaussian distribution together with the ALT technique to precompute an estimate of an exact heuristic between any cell-goal pair in the map. Briefly, the ALT method involves choosing k landmarks throughout the map and precomputing, using a Dijkstra search, the distance between each cell in the map and those k landmarks. Then, when an A* search is run, every time the heuristic is evaluated we invoke the reverse triangle inequality to compute a lower bound on the distance between that cell (n) and the goal (g).

$$|d(n, L_k) - d(g, L_k)| \leq d(n, g) = \hat{h}_{L_k}(n, g) \quad (1)$$

To achieve the best lower bound possible (and thus the fastest search), the maximum of the lower bounds proposed by all k landmarks is chosen.

$$\hat{h}_{ALT}(n) = \max_{k \in K}\left(\hat{h}_{L_k}(n, g)\right) \quad (2)$$

In order to use this estimate in conjunction with R^*_δ search, further steps need to be taken. This is because the heuristic obtained through ALT is always an overestimate of the exact distance, and for the risk guarantees of R^*_δ to apply we require this heuristic to be an accurate probability distribution over the exact distance. To better approximate this exact distance, small scaling parameters for both the mean (τ) and the variance (φ) are used.

$$\hat{h} = \mathcal{N}(\mu_{\hat{h}_{ALT}} - \tau, \varphi(\sigma^2_{\hat{h}_{ALT}})) \tag{3}$$

The values of these parameters are learnt by conducting multiple searches and comparing the results to optimal paths found by A* search using an admissible (Euclidean) heuristic.

In the course of carrying out an A^* search we observe $\hat{g}(n)$, which is the best known approximation to $g(n)$ — the minimum cost of navigating from the start s to node n. Note that this is a scalar value. Knowledge of $\hat{g}$ induces a conditional density function on $f^\dagger(n)$, the cost of a path from the start to goal via node n.

This observation, coupled with Equation 3 leads to a probability distribution over $f^\dagger$:

$$p(f^\dagger \mid \hat{g}, \hat{h}) = \hat{g} + p(h \mid \hat{h}) \tag{4}$$

Once we have this probability distribution over $f^\dagger$, we need to transform it to a scalar value for compatibility with A* and its priority queue ordering mechanism. R^*_δ performs this translation through the mechanism of *risk functionals*. An R^*_δ search imposes the requirement that the underlying A* search will continue until no node on the OPEN list has a risk associated with it that is greater than some level δ. This is called *δ-risk admissibility*, and guarantees that the search always terminates at a cost C such that risk $R(C)$ that the node could lead to a better path than the terminating path cost is less than δ for all nodes left on the OPEN priority queue.

So instead of using f-values to order the OPEN list, under R^*_δ we use a *threshold cost function* $C_\delta(n)$ which is given by the solution to the equation

$$R(C) = \delta. \tag{5}$$

The *Probability of Suboptimal Termination* risk functional is one method of mapping risk bounds and probability distributions over f to the A* paradigm. Here, we choose δ to be the probability of obtaining a suboptimal solution which we are prepared to accept. For instance, if we were prepared to risk obtaining a suboptimal solution 5 out of every 100 iterations of a particular problem instance we would set δ to 0.05.

For the purposes of this paper, where we are interested in measuring the quality of the heuristic approximation offered by different landmark selection techniques, it is convenient to use the $R_{ST}(C)$ risk functional together with the properties of the Gaussian distribution and write δ as an expression of the distance from the mean.

$$C_\delta(n) = \begin{cases} \mu & \delta = 0.5 \\ \mu - \sigma & \delta = 0.159 \\ \mu - 2 * \sigma & \delta = 0.023 \end{cases} \tag{6}$$

In effect, if we wish to have the search return the optimal path 50% of the time, when mapping the probability distribution over $f^\dagger$ to a scalar we would simply use the mean of the distribution. If we wanted to return the optimal path 84.1% of the time, we use the mean minus one standard deviation. These three delta values in Equation 6 are sufficient to provide a measure of the quality of our estimate of $p(f^\dagger \mid \hat{g}, \hat{h})$.

IV. CHOOSING LANDMARKS

Four methods of choosing landmarks are proposed. Potamias [15] proved choosing landmarks with high centrality to be the best technique for selecting landmarks for scalar costmaps for the approximate shortest path problem [15]. We wish to evaluate whether choosing landmarks with high centrality can lead to better results than planar selection for our probabilistic costmaps; when used in conjunction with exact shortest path techniques such as A* and D* prevalent in robotics. In addition we postulate that unlike scalar costmaps, landmarks are required throughout the centre of the map in order to estimate the variance of the probability distribution over the exact heuristic well. Furthermore, we wish to investigate the effect the number of landmarks has on the quality of the heuristic and the efficiency of the search.

The four methods we propose are:

- **Betweenness Centrality on a Grid** (BCGrid) The centrality of grid cells in the map is measured using Betweenness Centrality [16] [17]. The Betweenness Centrality of a node in a graph is a measure of the number of shortest paths that pass through the node, aggregated over the all-pairs shortest-paths of the graph. Concretely:

  $$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{7}$$

  where σ_{st} is the number of shortest paths between nodes s and t, and σ_{st} is the number of those paths that pass through v.

  In the BCG method we divide the map up into a regular grid and choose the landmark with the highest betweenness centrality for that portion of the grid.
- **Betweenness Centrality with K-Medoids** (BCKMed) In this selection method we frame the problem as an instance of the unconstrained facility location problem. The challenge is to site the landmarks so that we obtain the lowest 'cost' to service all possible queries that might originate from a given map. We view this as a k-medoid problem and apply a variation of local search to realise an approximate solution.

  To begin, we start with a large candidate pool C of landmarks, between 8 and 10 times the amount of landmarks, k, we desire in the final set. This pool contains the b nodes with the highest betweenness centrality in each coarse grid cell. We have a large pool of sample queries (1% of the number of cells in the map) which are deemed to be representative of the types of searches the landmarks will be required to 'serve' if chosen. An arbitrary selection of k landmarks from C is made as the initial landmark set. Then, i iterations of a local search are carried out, whereby one landmark from the set is swapped with a new landmark (the proposed landmark) from the candidate pool.

To arbitrate between the landmark sets we introduce a cost measure, which is evaluated over the set of landmarks — not the landmarks themselves. For each sample query, the lower bound given by the $h = \mu - \sigma$ (ST1) heuristic is evaluated using each landmark in the set. This risk functional is used as it encompasses the need for the landmarks to estimate both the mean and the variance of the exact distance well.

We find the optimal landmark, L_q^*, which offers the tightest lower bound on the distance between the two cells which form the query q.

$$L_q^* = \arg\max_L \hat{h}_q. \tag{8}$$

The cost of this optimal landmark is zero. To penalize 'bad' landmarks that overestimate the distance between the nodes in the query, the log of the difference in heuristic estimate between the optimal landmark estimate and the 'bad' landmark estimate is used.

$$C(L_k, q) = \log\left(\hat{h}_{L_q^*} - \hat{h}_{L_k}\right) \tag{9}$$

The cost of the set of landmarks under evaluation is

$$C_{set} = \sum_{q \in Q} \sum_k C(L_k, q). \tag{10}$$

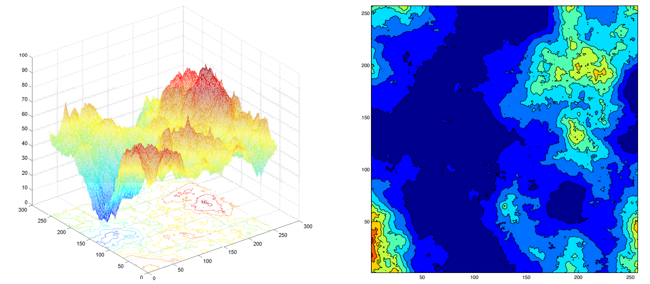
Once the swap is made, the cost of the previous set is compared to the cost of the new set. If the cost has decreased then the proposed landmark is kept, otherwise it is discarded and the initial set retained. We conduct several hundred iterations of the swapping step, terminating once the cost of the sets has converged and negligible improvements are noted between swaps.

- **Simple Landmark selection with K-Medoids** (SIMP-KMed) The same K-Medoid approach to selecting the set of landmarks is used, but the candidate pool of landmarks are drawn from an ordered grid spacing over the map rather than using the nodes with the highest betweenness centrality.
- **Simple Landmark selection on a Grid** (SIMPGrid) The most basic of the proposed methods, here we place landmarks on the map in a regular grid.

The different landmark configurations obtained using these techniques are shown on a sample overhead image in Fig 3. Note how pixels (grid cells) with high betweenness centrality correspond to roadways in the map. It is possible to pick good values for the BCGrid heuristic using the human eye.

V. RESULTS

The performance of the four landmark selection techniques were compared with the baseline technique of planar landmark selection (PlanarLM). In the absence of real world probabilistic costmaps, we tested these methods on a ten



(a) Fractal Generated Terrain Map - Mean (b) Fractal Generated Terrain Map - Variance

Fig. 2: An example from the 10 member fractally generated costmap dataset.

costmap set of synthetic fractal terrain maps, where both the mean and the variance of the map was generated using the diamond-square algorithm [19] (see Figure 2). This was done to ensure that there would be distinctive areas of high and low variance in the map — rather than randomly scattered throughout — and that this would reflect real world situations where certain regions would be better known than others.

The performance of the landmark selection technique is judged against two criteria — the efficiency of the results obtained when compared with a basic A* search with euclidean distance heuristic, measured in terms of the percentage of nodes expanded compared with the Euclidean heuristic search; and the accuracy in approximating the underlying exact heuristic. The accuracy is measured using statistics obtained from the $h = \mu$, $h = \mu - \sigma$ and $h = \mu - 2\sigma$ heuristic searches. We use the cumulative densities provided by these results to construct a Gaussian distribution which is scaled and offset from a $\mathcal{N}(0, 1)$ Gaussian representing the 'correct' distribution we seek, and compare the two using the Kullback-Leibler divergence. For brevity, we refer to searches employing these heuristics as ST0, ST1 and ST2 from here on, meaning that they use the suboptimal termination risk functional with the mean minus 0, 1 and 2 standard deviations respectively. Results were generated by running 50 separate start-goal searches, and for each search generated 50 different samples from the probabilistic costmap in order to test the probabilistic bounds.

In terms of efficiency, the results in Table I show that the risk planning searches become more efficient (expanding less nodes/grid cells) the larger the number of landmarks used. Comparing across results for the ST0 - ST2 risk functionals, we see that the SIMPGrid landmark selection method consistently results in the most efficient risk based searches. However, Table II shows the accuracy of heuristic approximations given this method to be poor. The landmarks are neither central enough or sufficiently close to the border of the map to produce good heuristic approximations. PlanarLM and BCGrid exhibit roughly equal efficiency, and significantly outperform both K-Medoids based selection methods.

The results shown in Table II are less clear. We do not see any obvious trend towards greater accuracy with increasing number of landmarks. The performance of each selection type has been ranked within the number of landmarks (ie

Method	No. LMs	EFF ST0	EFF ST1	EFF ST2
BCKMed	1	51.41	87.99	95.42
	4	46.67	89.64	96.84
	9	35.29	92.78	99.04
	16	24.14	93.73	99.69
SIMPKMed	1	47.54	82.47	91.25
	4	38.01	92.84	98.59
	9	32.22	90.90	98.29
	16	21.02	90.92	99.34
PlanarLM	1	48.24	93.65	98.97
	4	15.97	85.76	97.70
	9	11.79	92.95	99.79
	16	11.45	93.27	99.75
	32	10.40	93.33	99.85
	64	9.23	81.43	88.43
BCGrid	1	50.72	89.26	96.23
	4	49.66	88.54	95.49
	9	47.00	86.37	94.10
	16	12.79	95.34	99.68
	36	9.57	96.63	99.97
	64	9.23	94.75	99.32
SIMPGrid	1	62.69	96.44	99.41
	4	18.98	74.29	90.49
	9	11.56	66.30	84.65
	16	9.57	58.03	80.41
	36	7.30	50.35	73.23
	64	6.30	50.71	74.25

TABLE I: Efficiency, measured in terms of nodes expanded during the course of an A* search using ALT with landmarks selected using the various landmark selection techniques as a percentage of those expanded using a standard A* search with euclidean distance heuristic. The best result from within each landmark selection technique is highlighted.

Method	No. LMs	ST0	ST1	ST2	KL Div	Rank
BCKMed	1	48.71	63.25	81.52	0.3535	5
	4	48.84	73.96	91.66	0.0462	2
	9	55.70	79.63	97.35	0.0557	2
	16	55.52	76.80	97.66	0.0979	3
SIMPKMed	1	44.09	59.79	79.71	0.3261	4
	4	49.22	78.16	93.76	0.0141	1
	9	48.00	70.30	86.59	0.0899	4
	16	52.42	73.72	96.65	0.0968	4
PlanarLM	1	52.41	87.62	99.40	0.0037	1
	4	43.37	61.40	81.87	0.2335	4
	9	49.64	77.05	94.21	0.0230	1
	16	57.21	79.35	98.17	0.0855	2
	36	53.20	70.09	87.96	0.2234	2
	64	56.50	71.14	89.12	0.3376	2
BCGrid	1	47.64	69.32	86.43	0.1033	2
	4	49.28	74.21	87.39	0.0471	3
	9	52.02	70.72	84.34	0.1614	5
	16	51.96	78.91	99.00	0.0235	1
	36	48.46	84.53	99.57	0.0014	1
	64	44.99	70.23	89.10	0.0630	1
SIMPGrid	1	45.83	65.97	85.89	0.1485	3
	4	39.02	39.25	66.95	NaN	5
	9	35.37	31.59	61.03	0.0766	3
	16	38.74	25.64	55.16	NaN	5
	36	34.81	16.79	40.42	0.4312	3
	64	31.85	17.75	42.62	0.3526	3
	Optimal	**50.00**	**84.2**	**97.8**	**0.0000**	**1**

TABLE II: Optimality of the Approximations to the Exact Heuristic using various landmark selection techniques and number of landmarks. Optimality is measured using the percentage of the paths returned by the ST0, ST1 and ST2 risk functionals which are equal to the optimal path that is returned by an A* search with Euclidean distance heuristic. A perferct heuristic that models a gaussian distribution over the exact heuristic distances between nodes and goals in the A* search would return 50%, 84.2% and 97.8% of solutions equal to the optimal solution for ST0, ST1 and ST2 respectively. The KL divergence models the discrepancy between the results achieved with the various landmark selection techniques and this optimal heuristic. The best result from within each landmark selection technique is highlighted.

best performance with k=1 landmarks is PlanarLMBorder). From this ranking we see that PlanarLM and BCGrid consistently outperform the other 3 selection techniques. Of these two, BCGrid exhibits significantly less variance in the KL divergence of its results, leading to the conclusion that spacing landmarks with high centrality regularly over the grid does outperform the planar border selection technique for probabilistic costmaps. The accuracy obtained on the ten costmap data set using 36 landmarks in conjunction with the BCGrid selection technique is almost ideal, and with 16 landmarks the performance is almost as good.

Also evident from Table I is that beyond 16 landmarks the efficiency gains from adding more landmarks starts to decrease rapidly. This challenges the notion that the size of the landmark set should be the maximum that can be supported in memory by the machine running the search [20]. The minor performance gains (measured in terms of nodes scanned by an A* search) that are obtained by adding extra landmarks starts to be outweighed by the large overhead of needing to evaluate the lower bound over a larger landmark set at *every heuristic evaluation* during the course of the search. Table II shows that reasonable accuracy can be obtained using only 1 landmark, and that notable efficiency gains (of the order of 50%) are registered when using the ST0 heuristic and only 1 landmark. In short, risky planning does not require extensive precomputation; these results show intelligent placement of a few (≤ 16) landmarks with high centrality spread evenly over the map performs comparably to using 4 times as many landmarks.

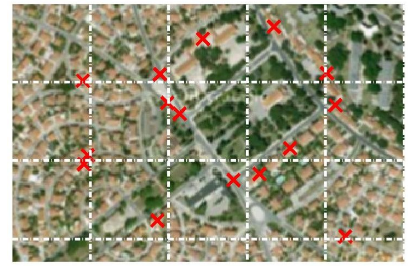
VI. CONCLUSIONS AND FUTURE WORKS

In this work we sought to augment the risky planning framework introduced in [2] with a landmark selection technique that leads to both efficient searches and accurate risk heuristics. It was found that the most robust way to place landmarks on the map such that the demands of efficiency and accuracy of the approximation are met, is to impose a coarse grid over the map and choose the node(s) with highest betweenness centrality in each broad cell as landmarks. We note that betweenness centrality is often easy to sense with the human eye, meaning that good results may be obtained by choosing likely landmarks with the human eye and thus avoiding costly precomputation. In addition, the results here show that a small number of landmarks is sufficient to obtain good results and therefore that landmark precomputation need not test the limits of available memory space.

The betweenness centrality grid landmark selection method we found to work best was also shown to produce near perfect approximations to Gaussian distributions over exact heuristics. This guarantees that the risk bounds of our Risky Planning framework will hold regardless of where the search lies in the map, and extends and improves upon our prior work in which over 70% efficiency increases over standard heuristic planning techniques were obtained. In future we seek to validate this work on real world datasets obtained from the construction of probabilistic costmaps from fielded robotics platforms.

(a) Planar Landmark Selection: Border. The 16 landmarks are spaced at regular intervals around the map border.

(b) Simple Grid. The 16 landmarks are spaced on a regular grid, covering the map.



(c) Betweenness Centrality: Grid. The map is divided into a rectangular grid and the cell with the highest betweenness centrality is chosen as the landmark for that cell.

(d) Simple K-Medoids. Yellow crosses denote members of the candidate set of landmarks passed to the K-Medoids function, the candidate set contains both landmarks spaced around the border of the map and a large number randomly scattered in the internal section of the map.

(e) Betweenness Centrality K-Medoids. Yellow crosses again denote the candidate set, drawn from the n locations with the highest betweenness centrality in each of the cells of a coarse grid imposed on the map.

Fig. 3: The 5 different landmark selection techniques used in the paper demonstrated on an overhead image. A probabilistic costmap has been constructed from the image following the framework introduced in [1], the landmarks have been precomputed using this costmap. Note how the landmarks for both Betweenness Centrality methods are located roadways and intersections: these locations are obvious to the human eye. Potentially, the need to precompute the Betweenness Centrality of grid cells could be bypassed in favour of choosing sparsely located landmarks corresponding to 'good' locations with the human eye.

References

[1] E. Murphy and P. Newman, "Planning paths from overhead imagery," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Anchorage, AK, May 2010.

[2] L. Murphy and P. Newman, "Risky planning: Path planning over costmaps with a probabilistically bounded speed-accuracy tradeoff," in *Proc. IEEE International Conference on Robotics and Automation (ICRA'11)*, Shanghai, China, May 2011, to Appear.

[3] A. V. Goldberg and C. Harrelson, "Computing the shortest path: A* search meets graph theory," in *16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '05)*, Vancouver, Canada, January 2005.

[4] J. Pearl and J. H. Kim, "Studies in semi-admissible heuristics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 4, no. 4, pp. 392–399, July 1982.

[5] A. Stentz, "Optimal and efficient path planning for partially-known environments," in *Proceedings of the International Conference on Robotics and Automation*, 1994, pp. 3310–3317.

[6] S. Koenig and M. Likhachev, "D*lite," in *Eighteenth national conference on Artificial intelligence*. Menlo Park, CA, USA: American Association for Artificial Intelligence, 2002, pp. 476–483.

[7] D. Ferguson and A. T. Stentz, "The field d* algorithm for improved path planning and replanning in uniform and non-uniform cost environments," Robotics Institute, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-05-19, July 2005.

[8] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions of Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, July 1968.

[9] L. R. Harris, "The bandwidth heuristic search," in *International Joint Conference on Artificial Intelligence*, Stanford, USA, 1973, pp. 23–29.

[10] M. Likhachev, D. Ferguson, G. Gordon, A. T. Stentz, and S. Thrun, "Anytime dynamic a*: An anytime, replanning algorithm," in *Proceedings of the International Conference on Automated Planning and Scheduling (ICAPS)*, June 2005.

[11] A. V. Goldberg, "Point-to-point shortest path algorithms with preprocessing,," in *Current Trends in Theory and Practice of Computer Science (SOFSEM)*, Harrachov, Czech Republic, 2007.

[12] A. Goldberg, H. Kaplan, and R. Werneck, "Reach for a: Efficient point-to-point shortest path algorithms," Microsoft Reasearch, Tech. Rep. MSR-TR-2005-132, 2005.

[13] K. Grant and D. Mould, "LPI: Approximating shortest paths using landmarks," in *ECAI 2008 - Workshop on AI and Games*, Patras, Greece, July 2008.

[14] D. Mould and M. C. Horsch, "An hierarchical terrain representation for approximately shortest paths," in *8th Pacific Rim International Conference on Artificial Intelligence*, Auckland, New Zealand,, August 2004.

[15] M. Potamias, F. Bonchi, C. Castillo, and A. Gionis, "Fast shortest path distance estimation in large networks," in *Proceeding of the 18th ACM conference on Information and knowledge management*, ser. CIKM '09. New York, NY, USA: ACM, 2009, pp. 867–876.

[16] L. C. Freeman, "A Set of Measures of Centrality Based on Betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, Mar. 1977.

[17] U. Brandes, "A faster algorithm for betweenness centrality," *Journal of Mathematical Sociology*, vol. 25, no. 2, pp. 163–177, 2001. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.11.2024

[18] A. V. Goldberg and R. F. Werneck. (2009, September) Selecting landmarks in shortest path computations. Patent Application Publication. Microsoft Corporation.

[19] A. Fournier, D. Fussel, and L. Carpenter, "Computer rendering of stochastic models," *Communications of the ACM*, vol. 25, pp. 371–384, 1982.

[20] A. V. Goldberg and R. F. Werneck, "Computing point-to-point shortest paths from external memory," in *Proceedings of the Seventh Workshop on Algorithm Engineering and Experiments (ALENEX'05)*, 2005.