

Tractable Pathfinding for the Stochastic On-Time Arrival Problem

Mehrdad Niknami¹(✉) and Samitha Samaranyake²

¹ Electrical Engineering and Computer Science, UC Berkeley, Berkeley, USA
mniknami@berkeley.edu

² School of Civil and Environmental Engineering, Cornell University, Ithaca, USA

Abstract. We present a new and more efficient technique for computing the route that maximizes the probability of on-time arrival in stochastic networks, also known as the path-based stochastic on-time arrival (SOTA) problem. Our primary contribution is a pathfinding algorithm that uses the solution to the *policy*-based SOTA problem—which is of pseudo-polynomial-time complexity in the time budget of the journey—as a search heuristic for the optimal path. In particular, we show that this heuristic can be exceptionally efficient in practice, effectively making it possible to solve the path-based SOTA problem as quickly as the policy-based SOTA problem. Our secondary contribution is the extension of policy-based preprocessing to path-based preprocessing for the SOTA problem. In the process, we also introduce Arc-Potentials, a more efficient generalization of Stochastic Arc-Flags that can be used for *both* policy- and path-based SOTA. After developing the pathfinding and preprocessing algorithms, we evaluate their performance on two different real-world networks. To the best of our knowledge, these techniques provide the most efficient computation strategy for the path-based SOTA problem for general probability distributions, both with and without preprocessing.

1 Introduction

Modern advances in graph theory and empirical computational power have essentially rendered deterministic point-to-point routing a solved problem. While the ubiquity of routing and navigation tools in our everyday lives is a testament to the success and usefulness of deterministic routing technology, inaccurate predictions remain a fact of life, resulting in missed flights, late arrivals to meetings, and failure to meet delivery deadlines. Recent research in transportation engineering, therefore, has focused on the collection of traffic data and the incorporation of *uncertainty* into traffic models, allowing for the optimization of relevant *reliability* metrics desirable for the user.

The point-to-point *stochastic on-time arrival* problem [1], or SOTA for short, concerns itself with this reliability aspect of routing. In the SOTA problem, the network is assumed to have uncertainty in the travel time across each link, represented by a strictly positive random variable. The objective is then to

maximize the probability of on-time arrival when traveling between a given origin-destination pair with a fixed time budget.¹ It has been shown that the SOTA solution can appreciably increase the probability of on-time arrival compared to the classical *least expected travel time* (LET) path [3], motivating the search for efficient solutions to this problem.

1.1 Variants

There exist two primary variants of the SOTA problem. The path-based SOTA problem, which is also referred to as the *shortest-path problem with on-time arrival reliability* (SPOTAR) [4], consists of finding the a-priori most reliable path to the destination. The policy-based SOTA problem, on the other hand, consists of computing a *routing policy*—rather than a fixed path—such that, at every intersection, the choice of the next direction depends on the current state (i.e., the remaining time budget).² While a policy-based approach provides better reliability when online navigation is an option, in some situations it can be necessary to determine the entire path prior to departure.

The policy-based SOTA problem, which is generally solved in discrete-time, can be solved via a successive-approximation algorithm, as shown by Fan and Nie [5]. This approach was subsequently improved by Samaranayake et al. [3] to a pseudo-polynomial-time label-setting algorithm based on dynamic-programming with a sequence of speedup techniques and the use of *zero-delay convolution* [6, 7]. It was then demonstrated in Sabran et al. [8] that graph preprocessing techniques such as Reach [9] and Arc-Flags [10] can be used to further reduce query times for this problem.

In contrast with the policy-based problem, however, no polynomial-time solution is known for the general path-based SOTA problem [4]. In the special case of normally-distributed travel times, Nikolova et al. [11] present an $O(n^{O(\log n)})$ -algorithm for computing exact solutions, while Lim et al. [12] present a polylogarithmic-time algorithm for approximation solutions. To allow for more general probability distributions, Nie and Wu [4] develop a label-correcting algorithm that solves the problem by utilizing the first-order stochastic dominance property of paths. While providing a solution method for general distributions, the performance of this algorithm is still insufficient to be of practical use in many real-world scenarios; for example, while the stochastic dominance approach provides a reasonable computation time (on the order of half a minute per instance) for networks of a few hundred to a thousand vertices, it fails to perform well on metropolitan road networks, which easily exceed tens of thousands of vertices. In contrast, our algorithm easily handles networks of tens of thousands of edges in approximately the same amount of time *without* any kind of preprocessing.³

¹ The target objective can in fact be generalized to utility functions other than the probability of on-time arrival [2] with little effect on our algorithms, but for our purposes, we limit our discussion to this scenario.

² In this article, we only consider time-invariant travel-time distributions. The problem can be extended to incorporate time-varying distributions as discussed in [3].

³ Parmentier and Meunier [13] have concurrently also developed a similar approach concerning stochastic shortest paths with risk measures.

With preprocessing, our techniques further reduce the running time to less than half a second, making the problem tractable for larger networks.⁴

1.2 Contributions

Our primary contribution in this article is a practically efficient technique for solving the path-based SOTA problem, based on the observation that the solution to the policy-based SOTA problem is in practice itself an extremely efficient heuristic for solving the path-based problem.

Our secondary contribution is to demonstrate how graph preprocessing can be used to speed up the computation of the policy heuristic, and thus the optimal path, while maintaining correctness.⁵ Toward this goal, we present Arc-Potentials, a more efficient generalization of the existing preprocessing technique known as Stochastic Arc-Flags that can be used for both policy- and path-based preprocessing.

After presenting these techniques, we evaluate the performance of our algorithms on two real-world networks while comparing the trade-off between their scalability (in terms of memory and computation time) and the speedups achieved. Our techniques, to the best of our knowledge, provide the most efficient computation strategy for the path-based SOTA problem with general probability distributions, both with and without preprocessing.

2 Preliminaries

We are given a stochastic network in the form of a directed graph $G = (V, E)$ where each edge $(i, j) \in E$ has an associated probability distribution $p_{ij}(\cdot)$ representing the travel time across that edge.⁶ The source is denoted by $s \in V$, the destination by $d \in V$, and the travel time budget by $T \in \mathbb{R}^+$.

For notational simplicity, we present the SOTA problem in continuous-time throughout this article, with the understanding that the algorithms are applied after discretization with a time interval of Δt .

Definition 1 (SOTA Policy). *Let $u_{ij}(t)$ be the probability of arriving at the destination d with time budget t when first traversing edge $(i, j) \in E$ and subsequently following the optimal policy. Let $\delta_{ij} > 0$ be the minimum travel time along edge (i, j) , i.e. $\min\{\tau : p_{ij}(\tau) > 0\}$. Then, the on-time arrival probability $u_i(t)$ and the policy (optimal subsequent node) $w_i(t)$ at node i , can be defined via the dynamic programming equations below [1]. Note that Δt must satisfy $\Delta t \leq \delta_{ij} \forall (i, j) \in E$.*

⁴ It should be noted that the largest network we consider only has approximately 71,000 edges and is still much smaller than networks used to benchmark deterministic shortest path queries, which can have millions of edges [14].

⁵ As explained later, there is a potential pitfall that must be avoided when the pre-processed policy is to be used as a heuristic for the path.

⁶ We assume that at most one edge exists between any pair of nodes in each direction.

$$\begin{aligned}
u_{ij}(t) &= \int_{\delta_{ij}}^t u_j(t - \tau) p_{ij}(\tau) d\tau & u_i(t) &= \max_{j: (i,j) \in E} u_{ij}(t) \\
u_d(\cdot) &= 1 & w_i(t) &= \arg \max_{j: (i,j) \in E} u_{ij}(t)
\end{aligned}$$

The solution to the policy-based SOTA problem can be obtained by solving this system of equations using dynamic programming as detailed in [3]. This requires evaluating a set of convolution integrals. The computation of the policy $u_s(\cdot)$ for a source s and time budget T using direct convolution takes $O(|E|T^2)$ time, as computing $u_s(T)$ could in the worst case require convolutions of length $O(T)$ for each edge in the graph. However, by using an online convolution technique known as *zero-delay convolution* (ZDC) [6, 15], the time complexity can be reduced to $O(|E|T \log^2 T)$. Justifications for the results and time complexity, including details on how to apply ZDC to the SOTA problem, can be found in [3, 7].

Assumptions. Our work, as with other approaches to both the policy-based and path-based SOTA problems, makes a number of assumptions about the nature of the travel time distributions. The three major assumptions are that the travel time distributions are (1) *time invariant*, (2) *exogenous* (not impacted by individual routing choices), and (3) *independent*. The time-invariance assumption—which prevents accounting for traffic variations throughout the day—can be relaxed under certain conditions as described in [3]. Furthermore, the exogeneity assumption is made even in the case of deterministic shortest path problems. This leaves the independence assumption as a major concern for this problem.

It might, in fact, be possible to partially relax this assumption [3] to allow for conditional distributions at the cost of increasing the computation time by a factor linear in the number of states to be conditioned on. (If we assume the Markov property for road networks, the number of conditioning states becomes the in-degree of each vertex, a small enough constant that may make generalizations in this direction practical.) Nevertheless, we will only focus on the independent setting and make no claim to have solved the path-based SOTA problem in full generality, as the problem already lacks efficient solution methods even in this simplified setting. Our techniques should, however, provide a foundation that allows for relaxing these assumptions in the future.

3 Path-Based SOTA

In the deterministic setting, efficient solution strategies (from Dijkstra’s algorithm to state-of-the-art solutions) generally exploit the *sub-path optimality* property: namely, the fact that any optimal route to a destination node d that includes some intermediate node i necessarily includes the optimal path from i to d . Unfortunately, this does not hold in the stochastic setting.

Algorithm 1. Algorithm for computing the optimal SOTA path

Notation: $*$ is the convolution operator and \parallel is the concatenation operator
for all $i \in V, 0 \leq t \leq T$ **do**
 Compute the optimal policy's reliability⁸ $u_i(t)$
 $Q \leftarrow \text{PriorityQueue}()$
 $Q.\text{Push}(u_s(T), ([1.0], [s]))$ \triangleright Push (reliability, (cost dist., initial path))
while Q is not empty **do**
 $(r, (q, P)) \leftarrow Q.\text{PopMax}()$ \triangleright Extract most reliable path so far
 $i \leftarrow P[|P| - 1]$ \triangleright Get the last node in the path
 if $i = d$ **then**
 return P
 for all $j \in E.\text{Neighbors}(i)$ **do**
 $Q.\text{Push}((q * u_j)[T], (q * E.\text{Cost}(i, j), P \parallel [j]))$ \triangleright Append new edge.
return nil \triangleright No path found

Furthermore, blind enumeration of all possible paths in the graph is absolutely intractable for all but the most trivial networks, as the number of simple paths grows exponentially with the number of nodes in the graph. Naturally, this leads us to seek a heuristic to guide us toward the optimal path efficiently, while not compromising its optimality.

3.1 Algorithm

Consider a fixed path P from the source s to node i . Let $q_{si}^P(t)$ be the travel time distribution along P from node s to node i , i.e., the convolution of the travel time distributions of every edge in P . Upon arriving at node i at time t , let the user follow the optimal policy toward d , therefore reaching d from s with probability density $q_{si}^P(t)u_i(T - t)$. The reliability of following path P to node i and subsequently following the optimal policy toward d is⁷:

$$r_{si}^P(T) = \int_0^T q_{si}^P(t)u_i(T - t) dt$$

Note that the route from $s \rightarrow i$ is a fixed path while that from $i \rightarrow d$ is a policy.

The optimal path is found via the procedure in Algorithm 1. Briefly, starting at the source s , we add the hybrid (path + policy) solution $r_{si}^P(T)$ for each neighbor i of s to a priority queue. Each of these solutions gives an upper bound on the solution (success probability). We then dequeue the solution with the highest upper bound, repeating this process until a path to the destination is found.

Essentially, Algorithm 1 performs an A^* search for the destination, using the policy as a heuristic. While it is obvious that the algorithm would find the

⁷ The bounds of this integral can be slightly tightened through inclusion of the minimum travel times, but this has been omitted for simplicity.

⁸ Can be limited to those i and t reachable from s in time T , and can be further sped up through existing policy preprocessing techniques such as Arc-Flags.

optimal path *eventually* if the search were allowed to continue indefinitely, it is less obvious that the *first* path found will be optimal. We show this by showing that the policy is an admissible heuristic for the path, and consequently, by the optimality of A^* [16], the first returned path must be optimal.

Proposition 1 (Admissibility). *The solution to policy-based SOTA problem is an admissible heuristic for the optimal solution to the path-based SOTA problem using Algorithm 1.*

Proof. When finding a minimum cost path, an admissible heuristic is a heuristic that never overestimates the actual cost [17]. In our context, since the goal is to maximize the reliability (success probability), this corresponds to a heuristic that never underestimates the reliability of a routing strategy. The reliability of an optimal SOTA policy clearly provides an upper bound on the reliability of any fixed path with the same source, destination, and travel budget. (Otherwise, a better policy would be to simply follow the fixed path irrespective of the time remaining, contradicting the assumption that the policy is optimal.) Therefore, the SOTA policy is an admissible heuristic for the optimal SOTA path.

3.2 Analysis

The single dominant factor in this algorithm’s (in)efficiency is the length of the priority queue (i.e., the number of paths considered by the algorithm), which in turn depends on the travel time distribution along each road. As long as the number of paths considered is approximately linear in length of the optimal path, the path computation time is easily dominated by the policy computation time, and the algorithm finds the optimal path very quickly. In the worst-case scenario for the algorithm, the optimal path at a node corresponds to the direction for the worst policy at that node. Such a scenario, or even one in which the optimal policy frequently chooses a suboptimal path, could result in a large (even exponential) running time as well as space usage. However, it is difficult to imagine this happening in practice. As shown later, experimentally, we came across very few cases in which the path computation time dominated the policy computation time, and even in those cases, they were still quite reasonable and extremely far from such a worst-case scenario. We conjecture that such situations are extremely unlikely to occur in real-world road networks.

An interesting open problem is to define characteristics (network structure, shape of distributions, etc.) that guarantee pseudo-polynomial running times in stochastic networks, similar in nature to the *Highway Dimension* property [18] in deterministic networks, which guarantees logarithmic query times when networks have a low *Highway Dimension*.

4 Preprocessing

In deterministic pathfinding, preprocessing techniques such as *Arc-Flags* [10], *reach-based routing* [9, 19], *contraction hierarchies* [20], and *transit node routing*

[21] have been very successfully used to decrease query times by many orders of magnitude by exploiting structural properties of road networks. Some of these approaches allow for pruning the search space based solely on the destination node, while others also take the source node into account, allowing for better pruning at the cost of additional preprocessing. The structure of the SOTA problem, however, makes it more challenging to apply such techniques to it. Previously, Arc-Flags and Reach have been successfully adapted to the policy-based problem in [8], resulting in Stochastic Arc-Flags and Directed Reach. While at first glance one may be tempted to directly apply these algorithms to the computation of the policy heuristic for the path-based problem, a naive application of source-dependent pruning (such as Directed Reach or source-based Arc-Flags) can result in an incorrect solution, as the policy needs to be recomputed for source nodes that correspond to different source regions. This effectively limits any preprocessing of the policy heuristic to destination-based (i.e., source-independent) techniques such as Stochastic Arc-Flags, precluding the use of source-based approaches such as Directed Reach for the policy computation.

With sufficient preprocessing resources (as explained in Sect. 5.2), however, one can improve on this through the direct use of *path-based* preprocessing—that is, pruning the graph to include only those edges which may be part of the most reliable path. This method allows us to simultaneously account for both source and destination regions, and generally results in a substantial reduction of the search space on which the policy needs to be computed. However, as path-based approaches require computing paths between all $\approx |V|^2$ pairs of vertices in the graph, this approach may become computationally prohibitive for medium- to large-scale networks. In such cases, we would then need to either find alternate approaches (e.g. approximation techniques), or otherwise fall back to the less aggressive policy-based pruning techniques, which only require computing $|V|$ separate policies (one per destination).

4.1 Efficient Path-Based Preprocessing

Path-based preprocessing requires finding the optimal paths for each time budget up to the desired time budget T for all source-destination pairs. Naively, this can be done by placing Algorithm 1 in a loop, executing it for all time budgets from 1 to T . This requires T times the work of finding the path for a single time budget, which is clearly prohibitive for any reasonable value of T . However, we can do far better by observing that many of the computations in the algorithm are independent of the time budget and can be factored out when the path does not change with T .

To improve the efficiency of the naive approach in this manner, we make two observations. First, we observe that, in Algorithm 1, *only* the computation of the path’s reliability (priority) in the priority queue ever requires knowledge of the time budget. Crucially, the convolution $q * E.Cost(i, j)$ only depends on the *maximum* time budget T for truncation purposes, which is a fixed value. This means that the travel time distribution of any path under consideration can be computed once for the maximum time budget, and re-used for all lower time

budgets thereafter. Second, we observe that when a new edge is appended, the priority of the new path is the inner product of the vector q and (the reverse of) the vector u_j , shifted by T . As noted in the algorithm itself, this quantity in fact the convolution of the two aforementioned vectors evaluated at T . Thus, when a new edge is appended, instead of recomputing the inner product, we can simply convolve the two vectors once, and thereafter look up the results instantly for other time budgets.

Together, these two observations allow us to compute the optimal paths for all budgets far faster than would seem naively possible, making path-based preprocessing a practical option.

4.2 Arc-Potentials

As noted earlier, Arc-Flags, a popular method for graph preprocessing, has been adapted to the SOTA problem as Stochastic Arc-Flags [8]. Instead of applying it directly, however, we present *Arc-Potentials*, a more natural generalization of Arc-Flags to SOTA that can still be directly applied to the policy- and path-based SOTA problems alike, while allowing for more efficient preprocessing.

Consider partitioning the graph G into R regions (we choose $R = O(\log |E|)$, described below), where R is tuned to trade off storage space for pruning accuracy. In the deterministic setting, Arc-Flags allow us to preprocess and prune the search space as follows. For every arc (edge) $(i, j) \in E$, Arc-Flags defines a bit-vector of length R that denotes whether or not this arc belongs to an optimal path ending at some node in region R . We then pre-compute these Arc-Flags, and store them for pruning the graph at query time. (This approach has been extended to the dynamic setting [22] in which the flags are updated with low recomputation cost after the road network is changed.)

Sabran et al. [8] apply Arc-Flags to the policy-based SOTA problem as follows: each bit vector is defined to represent whether or not its associated arc is *realizable*, meaning that it *belongs to an optimal policy* to some destination in the target region associated with each bit. The problem with this approach, however, is that it requires computing arc-flags for *all* target budgets (or more, practically, some ranges of budgets), each of which takes a considerable amount of space. Instead, we propose a more efficient alternative Definition 2, which we call *Arc-Potentials*.

Definition 2 (Arc-Potentials). *For a given destination region D , we define the arc activation potential ϕ_{ij} of the edge from node i to node j to be the minimum time budget at which the arc becomes part of an optimal policy to some destination $d \in D$.*

The Arc-Potentials pruning algorithm only stores the “activation” potential of every edge. As expected, this implies that for large time budgets, every edge is potentially active. We could have further generalized the algorithm to allow for asymptotically *exact* pruning at relatively low cost by simply storing the actual potential *intervals* during which the arc is active, rather than merely their first activation potential. However, in our experiments this was deemed

unnecessary as Arc-Potentials were already sufficient for significant pruning in the time budgets of interest in our networks.

The computation of the set of realizable edges (and nodes) under a given policy is essentially equivalent to the computation of the policy itself, except that updates are performed in the reverse order (from the source to the destination). The activation potentials ϕ can then be obtained from this set. As with Arc-Flags, we limit the space complexity to $O(|E|R) = O(|E| \log |E|)$ by choosing R to be proportional $\log |E|$, tuning it as desired to increase the pruning accuracy. In our experiments, we simply used a rectangular grid of size $\sqrt{R} \times \sqrt{R}$. Note, however, that the preprocessing time does not depend on R , as the paths between all $\approx |V|^2$ pairs of nodes must be eventually computed.

5 Experimental Results

We evaluated the performance of our algorithms on two real-world test networks: a small San Francisco network with 2643 nodes and 6588 edges for which real-world travel-time data was available as a Gaussian mixture model [23], and a second (relatively larger) Luxembourg network with 30647 nodes and 71655 edges for which travel-time distributions were synthesized from road speed limits, as real-world data was unavailable. The algorithms were implemented in C++ (2003) and executed on a cluster of 1.9 GHz AMD Opteron™ 6168 CPUs. The SOTA queries were executed on a single CPU and the preprocessing was performed in parallel as explained below.

The SOTA policies were computed as explained in [3, 7] using zero-delay convolution with a discretization interval of $\Delta t = 1$ s.⁹ To generate random problem instances, we independently picked a source and a destination node uniformly at random from the graph and computed the least expected travel-time (LET) path between them. We then evaluated our pathfinding algorithm for budgets chosen uniformly at random from the 5th to 95th percentile of LET path travel times (those of practical interest) on 10,000 San Francisco and 1000 Luxembourg problems instances.

First, we discuss the speed of our pathfinding algorithm, and afterward, we evaluate the effectiveness and scalability of our preprocessing strategies.

5.1 Evaluation

We first evaluate the performance of our path-based SOTA algorithm without any graph preprocessing. Experimental results, as can be seen in Fig. 1, show that the run time of our solution is dominated by the time taken to obtain the solution to the policy-based SOTA problem, which functions as a search heuristic for the optimal path.

The stochastic-dominance (SD) approach [4], which to our knowledge is the fastest published solution for the path-based SOTA problem with general probability distributions, takes, on average, between 7 and 18 s (depending on the

⁹ Recall that we must have $\Delta t \leq \min_{(i,j) \in E} \delta_{ij}$, which is ≈ 1 s for our networks.

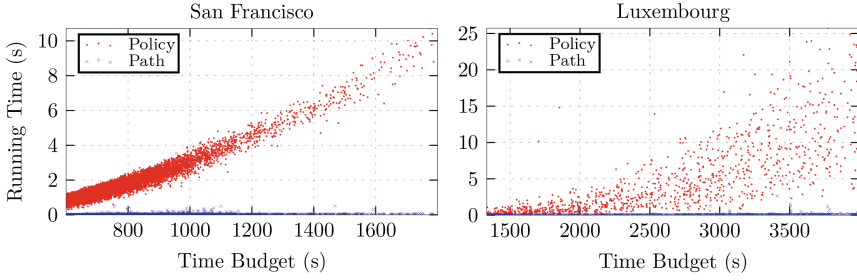


Fig. 1. Running time of the pathfinding algorithm as a function of the travel time budget for random unpruned (i.e., non-preprocessed) instantiations of each network. We can see that the path computation time is dominated by the policy computation time, effectively reducing the path-based SOTA problem to the policy-based SOTA problem in terms of computation time.

variance of the link travel time distributions) to compute the optimal path for 100 time-step budgets. For comparison, our algorithm solves for paths on the San Francisco network with budgets of up to 1400 s (= 1400 time-steps) in ≈ 7 s, even achieving query times below 1 s for budgets less than 550 s without any preprocessing at all. Furthermore, it also handles most queries on the 71655-edge Luxembourg network in ≈ 10 s (almost all of them in 20 s), where the network and time budgets are more than an order of magnitude larger than the 2950-edge network handled by the SD approach in the same amount of time.

Of course, this speedup—which increases more dramatically with the problem size—is hardly surprising or coincidental; indeed, it is quite fundamental to the nature of the algorithm: by drawing on the optimal policy as an upper bound (and quite often an accurate one) for the reliability of the final path, it has a very clear and fundamental informational advantage over any search algorithm that lacks any knowledge of the final solution. This allows the algorithm to direct itself toward the final path in an intelligent manner.

It is, however, less clear and more difficult to see how one might compare the performance of our generic discrete-time approach with Gaussian-restricted, continuous-time approaches [12, 24]. Such approaches operate under drastically different assumptions and, in the case of [12], use approximation techniques, which we have yet to employ for additional performance improvements. When the true travel times cannot be assumed to follow Gaussian distributions, however, our method, to the best of our knowledge, presents the most efficient means for solving the path-based SOTA problem.

As we show next, combining our algorithm with preprocessing techniques allows us to achieve even further reductions in query time, making it more tractable for industrial applications on appropriately sized networks.

Preprocessing. Figure 2 demonstrates policy-based and path-based preprocessing using Arc-Potentials for two random San Francisco and Luxembourg problem

instances. As can be seen in the figure, path-based preprocessing is in general much more effective than policy-based preprocessing.

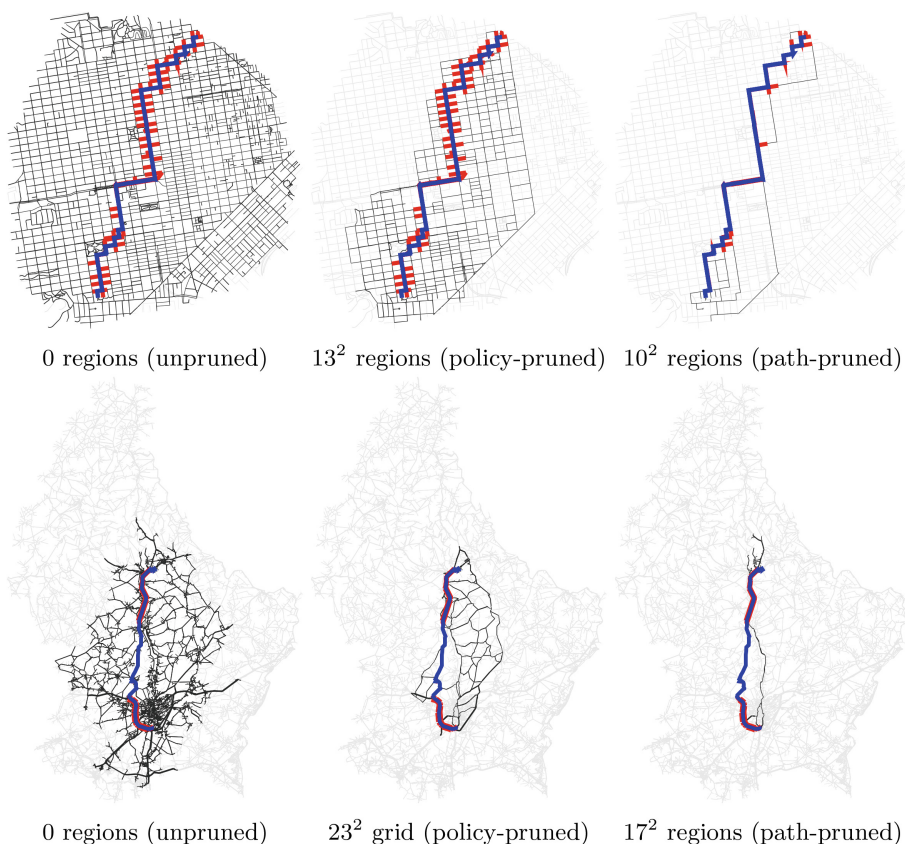


Fig. 2. Policy- vs. path-based pruning for random instances of San Francisco ($T = 837$ s, source at top) and Luxembourg ($T = 3165$ s, source at bottom). Light-gray edges are pruned from the graph and blue edges belong to the optimal path, whereas red edges belong to (sub-optimal) paths that were on the queue at the termination of the algorithm. (Color figure online)

Figure 3, summarized in Table 1, shows how the computation times scale with the preprocessing parameters. As expected, path-based preprocessing performs much better than purely policy-based preprocessing, and both become faster as we use more fine-grained regions. Nevertheless, we see that the majority of the speedup is achieved via a small number of regions, implying that preprocessing can be very effective even with low amounts of storage. (For example, for a 17×17 grid in Luxembourg, this amounts to $71655 \times 17^2 \approx 21$ M floats.)

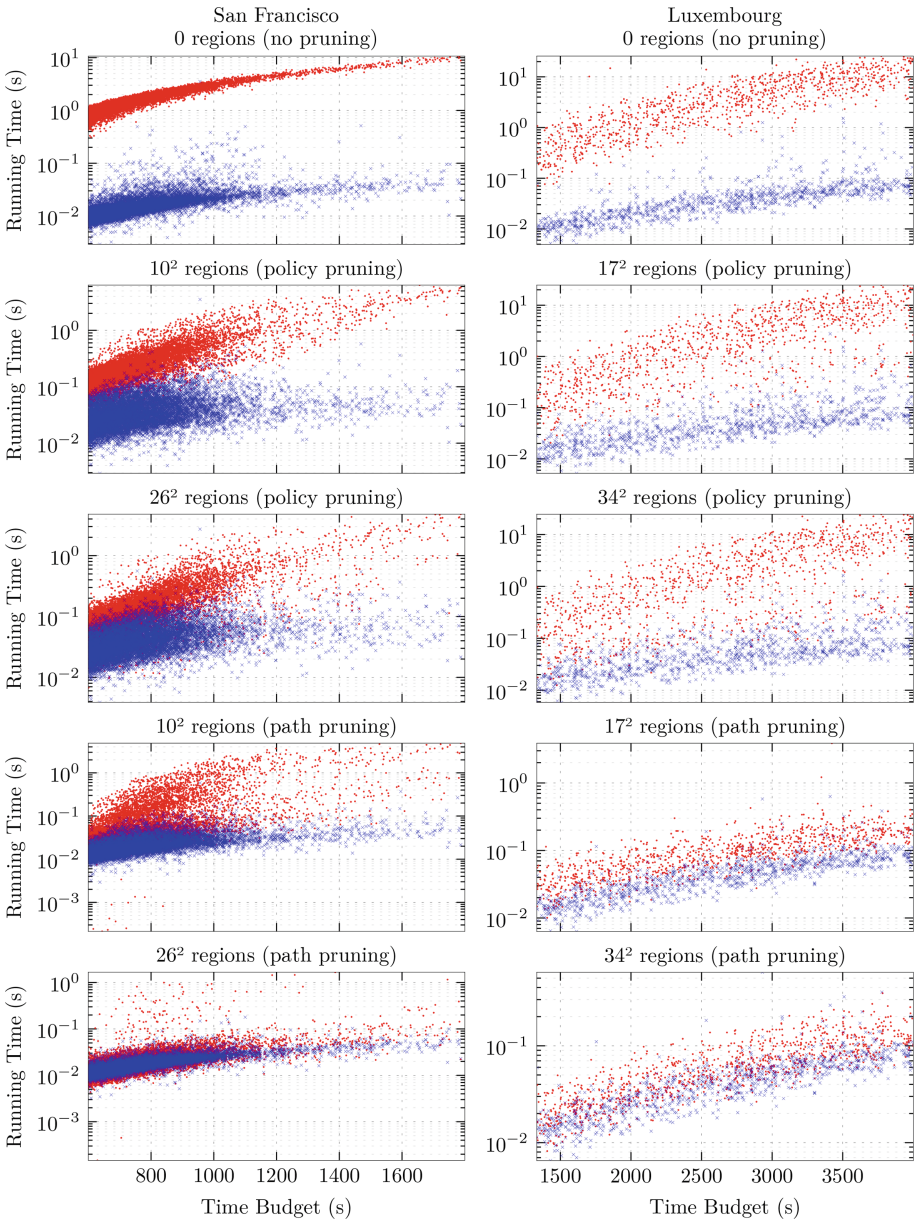


Fig. 3. Running time of pathfinding algorithm as a function of the time budget for each network. Red dots represent the computation time of the policy, and blue crosses represent the computation of the path using that policy. (Color figure online)

Table 1. The average query time with both policy-based and path-based pruning at various grid sizes and time budgets on the San Francisco network (left) and the Luxembourg network (right). We can see that in both cases, most of the speedup occurs at low granularity (and thus low space requirements).

Grid/ Pruning	Time budget (seconds)					Grid/ Pruning	Time budget (seconds)				
	800	1000	1200	1400	1600		1500	2000	2500	3000	3500
Unpruned	1.81	3.00	4.10	5.11	5.23	Unpruned	0.54	1.29	3.53	6.27	9.95
10 × 10, policy	0.30	0.69	1.11	1.66	1.72	17 × 17, policy	0.25	0.83	2.31	4.57	7.97
26 × 26, policy	0.17	0.40	0.63	0.93	0.97	34 × 34, policy	0.21	0.71	2.09	3.79	7.07
10 × 10, path	0.11	0.38	0.63	0.87	0.90	17 × 17, path	0.03	0.06	0.09	0.13	0.18
26 × 26, path	0.02	0.04	0.06	0.07	0.08	34 × 34, path	0.02	0.04	0.06	0.08	0.12

5.2 Scalability

Path-based preprocessing requires routing between all $\approx |V|^2$ pairs of vertices, which is quadratic in the size of the network and intractable for moderate size networks. In practice, this meant that we had to preprocess every region lazily (i.e. on-demand), which on our CPUs took 9000 CPU-hours. It is therefore obvious that this becomes intractable for large networks, leaving policy-based preprocessing as the only option. One possible approach for large-scale path-based preprocessing might be to consider the boundary of each region rather than its interior [8]. While currently uninvestigated, such techniques may prove to be extremely useful in practice, and are potentially fruitful topics for future exploration.

6 Conclusion and Future Work

We have presented an algorithm for solving the path-based SOTA problem by first solving the easier policy-based SOTA problem and then using its solution as a search heuristic. We have also presented two approaches for preprocessing the underlying network to speed up computation of the search heuristic and path query, including a generalization of the Arc-Flags preprocessing algorithm that we call Arc-Potentials. We have furthermore applied and implemented these algorithms on moderate-sized transportation networks and demonstrated their potential for high efficiency in real-world networks.

While unobserved in practice, there remains the possibility that our algorithm may perform poorly on stochastic networks in which the optimal policy is a poor heuristic for the path reliability. Proofs in this direction have remained elusive, and determining whether such scenarios can occur in realistic networks remains an important step for future research. In the absence of theoretical advances, however, our algorithm provides a more tractable alternative to the state-of-the-art techniques for solving the path-based SOTA problem.

While our approach is tractable for larger networks than were possible with previous solutions, it does not scale well enough to be used with regional or

continental sized networks that modern deterministic shortest path algorithms can handle with ease. In the future, we hope to investigate how our policy-based approach might be combined with other techniques such as the first-order stochastic dominance [4] and approximation methods such approximate Arc-Flags [8] for further speedup, and to also look into algorithms that allow for at least a partial relaxation of the independence assumption. We hope that our techniques will provide a strong basis for even better algorithms to tackle this problem for large-scale networks in the foreseeable future.

References

1. Fan, Y., Robert Kalaba, J.E., Moore, I.I.: Arriving on time. *J. Optim. Theor. Appl.* **127**(3), 497–513 (2005)
2. Flajolet, A., Blandin, S., Jaillet, P.: Robust adaptive routing under uncertainty (2014). [arXiv:1408.3374](https://arxiv.org/abs/1408.3374)
3. Samaranayake, S., Blandin, S., Bayen, A.: A tractable class of algorithms for reliable routing in stochastic networks. *Transp. Res. Part C* **20**(1), 199–217 (2012)
4. Nie, Y.M., Wu, X.: Shortest path problem considering on-time arrival probability. *Trans. Res. Part B Methodol.* **43**(6), 597–613 (2009)
5. Fan, Y., Nie, Y.: Optimal routing for maximizing travel time reliability. *Netw. Spat. Econ.* **6**(3–4), 333–344 (2006)
6. Dean, B.C.: Speeding up stochastic dynamic programming with zero-delay convolution. *Algorithmic Oper. Res.* **5**(2), 96 (2010)
7. Samaranayake, S., Blandin, S., Bayen, A.: Speedup techniques for the stochastic on-time arrival problem. In: *ATMOS*, pp. 83–96 (2012)
8. Sabran, G., Samaranayake, S., Bayen, A.: Precomputation techniques for the stochastic on-time arrival problem. In: *SIAM, ALENEX*, pp. 138–146 (2014)
9. Gutman, R.: Reach-based routing: a new approach to shortest path algorithms optimized for road networks. In: *ALENEX/ANALC*, pp. 100–111 (2004)
10. Hilger, M., Köhler, E., Möhring, R., Schilling, H.: Fast point-to-point shortest path computations with Arc-Flags. *Ninth DIMACS Implementation Challenge* **74**, 41–72 (2009)
11. Nikolova, E., Kelner, J.A., Brand, M., Mitzenmacher, M.: Stochastic shortest paths via quasi-convex maximization. In: Azar, Y., Erlebach, T. (eds.) *ESA 2006. LNCS*, vol. 4168, pp. 552–563. Springer, Heidelberg (2006)
12. Lim, S., Sommer, C., Nikolova, E., Rus, D.: Practical route planning under delay uncertainty: stochastic shortest path queries. *Robot. Sci. Syst.* **8**(32), 249–256 (2013)
13. Parmentier, A., Meunier, F.: Stochastic shortest paths and risk measures (2014). [arXiv:1408.0272](https://arxiv.org/abs/1408.0272)
14. Delling, D., Sanders, P., Schultes, D., Wagner, D.: Engineering route planning algorithms. In: Lerner, J., Wagner, D., Zweig, K.A. (eds.) *Algorithmics of Large and Complex Networks. LNCS*, vol. 5515, pp. 117–139. Springer, Heidelberg (2009)
15. Gardner, W.G.: Efficient convolution without input/output delay. In: *Audio engineering society convention 97*. Audio Engineering Society (1994)
16. Dechter, R., Pearl, J.: Generalized best-first search strategies and the optimality of A^* . *J. ACM (JACM)* **32**(3), 505–536 (1985)
17. Russell, S.J., Norvig, P.: *Artificial Intelligence: A Modern Approach*. Prentice-Hall Inc., London (1995). ISBN 0-13-103805-2

18. Abraham, I., Fiat, A., Goldberg, A., Werneck, R.: Highway dimension, shortest paths, and provably efficient algorithms. In: Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 782–793. Society for Industrial and Applied Mathematics (2010)
19. Goldberg, A., Kaplan, H., Werneck, R.: Reach for A^* : efficient point-to-point shortest path algorithms. In: ALENEX, vol. 6, pp. 129–143. SIAM (2006)
20. Geisberger, R., Sanders, P., Schultes, D., Delling, D.: Contraction hierarchies: faster and simpler hierarchical routing in road networks. In: McGeoch, C.C. (ed.) WEA 2008. LNCS, vol. 5038, pp. 319–333. Springer, Heidelberg (2008)
21. Bast, H., Funke, S., Matijevic, D.: Transit: ultrafast shortest-path queries with linear-time preprocessing. In: 9th DIMACS Implementation Challenge [1] (2006)
22. D’Angelo, G., Frigioni, D., Vitale, C.: Dynamic arc-flags in road networks. In: Pardalos, P.M., Rebennack, S. (eds.) SEA 2011. LNCS, vol. 6630, pp. 88–99. Springer, Heidelberg (2011)
23. Hunter, T., Abbeel, P., Bayen, A.M.: The path inference filter: model-based low-latency map matching of probe vehicle data. In: Frazzoli, E., Lozano-Perez, T., Roy, N., Rus, D. (eds.) Algorithmic Foundations of Robotics X. STAR, vol. 86, pp. 591–607. Springer, Heidelberg (2013)
24. Lim, S., Balakrishnan, H., Gifford, D., Madden, S., Rus, D.: Stochastic motion planning and applications to traffic. *Int. J. Robot. Res.* 3–13 (2010)