High-performance Heuristics for Optimization in Stochastic Traffic Engineering Problems

Evdokia Nikolova

Massachusetts Institute of Technology
Cambridge, MA, U.S.A.
enikolova @csail.mit.edu

Abstract. We consider a stochastic routing model in which the goal is to find the optimal route that incorporates a measure of risk. The problem arises in traffic engineering, transportation and even more abstract settings such as task planning (where the time to execute tasks is uncertain), etc. The stochasticity is specified in terms of arbitrary edge length distributions with given mean and variance values in a graph. The objective function is a positive linear combination of the mean and standard deviation of the route. Both the nonconvex objective and exponentially sized feasible set of available routes present a challenging optimization problem for which no efficient algorithms are known. In this paper we evaluate the practical performance of algorithms and heuristic approaches which show very promising results in terms of both running time and solution accuracy.

1 Introduction

Consider the problem one faces every day to go to work: Find the best route between two points in a complex network of exponentially many routes, filled with uncertainty. In the absence of uncertainty, there are known polynomial-time algorithms, as well as numerous metaheuristics developed to yield very fast and practical running times for shortest path computations. However, the case of uncertainty still leaves unresolved theoretical problems and calls for practical metaheuristic methods. Note that with uncertainty it is not even clear how to define the optimal route: is it the route that minimizes the *expected* travel time? Or its variance, or some other metric?

In this paper we consider the natural definition of an optimal route to minimize a convex combination of mean and standard deviation. As it turns out, the solution approaches for finding the optimal route under this metric also yield solutions to the related problem of arriving ontime (in which the optimal path maximizes the probability that the route travel time does not exceed a given deadline). The stochasticity is defined in terms of given mean and variance values for each edge in the network, under arbitrary distributions.

With this definition of an optimal route, traditional methods of solving shortest path problems fail (the problem no longer has the property that a subpath of an optimal path is optimal and thus one cannot use dynamic programming techniques to solve it). The structure of the problem allows us to reduce the number of local optima so that we do not need to examine all exponentially many routes. However in the worst case, the local optima are still superpolynomially many.

In the current paper we evaluate the performance of several algorithms and heuristic approaches. First, for a class of grid graphs, we provide experimental results that the number of local optima is sublinear and thus examining all of them can be done efficiently and yield the exact optimum. Second, for the purpose of practical implementation, it is preferable to examine only a small (constant or logarithmic) number of local optima. To this end, we examine heuristics that pick a small subset of the local optima, and provide bounds and experimental evaluation on the quality of the resulting solution.

2 Problem statement and preliminaries

We are given a graph G with n nodes and m edges and are interested in finding a route between a specified source node S and a destination node T. The edge lengths in the graph are stochastic and come from arbitrary independent[1] distributions with given mean μ_i and variance τ_i (which can be different for different edges i). Our goal is to find the optimal route that incorporates a measure of risk. As such we define the optimal route via a natural family of objectives, namely to minimize a convex combination of the route's mean and standard deviation:

$$\begin{aligned} \text{minimize} \quad & \alpha \sum_{i \in P} \mu_i + (1-\alpha)\sqrt{\sum_{i \in P} \tau_i} \\ \text{such that} \quad & P \text{ is an } ST\text{-path.} \end{aligned} \tag{1}$$

The summations above are over the edges i in a given ST-route and the minimization is over all valid ST-routes P. The parameter $\alpha \in [0,1]$ specifies the objective function (convex combination) of choice.

It will be helpful to consider a continuous formulation of the above discrete problem, by denoting an ST-route by its corresponding incidence vector $\mathbf{x} = (x_1, ..., x_m)$ where $x_i = 1$ if edge i is present in the route and $x_i = 0$ otherwise. Denote also the vector of means of all edges by $\boldsymbol{\mu} = (\mu_1, ..., \mu_m)$ and the vector of variances by $\boldsymbol{\tau} = (\tau_1, ..., \tau_m)$. The set of all feasible routes is represented by $\{01\}$-vectors $\mathbf{x} \in \mathbf{R}^m$, which are a subset of the vertices of the unit hypercube in m dimensions. The *convex hull* of these feasible vectors is called the *path polytope*. Thus, the continuous formulation of the routing problem is:

$$\begin{aligned} \text{minimize} \quad & \alpha \boldsymbol{\mu}^T \mathbf{x} + (1-\alpha)\sqrt{\boldsymbol{\tau}^T \mathbf{x}} \\ \text{such that} \quad & \mathbf{x} \in \text{path polytope.} \end{aligned} \tag{2}$$

[1] The results here generalize to dependent distributions, however we focus this presentation on the independent case for clarify and brevity of the mathematical exposition.

This objective function is concave for all values of the parameter $\alpha \in [0, 1]$, so it attains its minimum at an extreme point[2] of the feasible set [2]. This is a key property of the objective which establishes that the optimal solution of the continuous problem (which is a superset of the original feasible set) will be a valid path and thus will coincide with the optimal solution of the discrete problem. We emphasize this important observation in the following proposition.

Proposition 1. *The optimal solution of the continuous formulation (2) is the same as the optimal solution of the discrete problem (1).*

Furthermore the objective function is monotone increasing in the route mean and variance, thus its optimum also minimizes some convex combination of the mean and variance (as opposed to the mean and standard deviation!).

Proposition 2. *The optimal solution of the nonconvex problem (2) minimizes the* linear *objective* $\beta \boldsymbol{\mu}^T \mathbf{x} + (1-\beta)\boldsymbol{\tau}^T \mathbf{x}$ *for some* $\beta \in [0, 1]$.

This second observation is critical for yielding a *subexponential exact* algorithm for the stochastic routing problem. The exact algorithm enumerates the candidate set of extreme points or paths (which is a small subset of all extreme points) in time linear in the number of such paths, which is at most $n^{O(\log n)}$ in the worst case.

In this paper, we seek to reduce this superpolynomial complexity.

Related Work. Our work is most closely related to the work of Nikolova *et al.* [10] who propose and analyze the worst-case running time of the exact algorithm for a related objective (in the above notation, to maximize the objective $\frac{t-\boldsymbol{\mu}^T\mathbf{x}}{\sqrt{\boldsymbol{\tau}^T\mathbf{x}}}$, which arises in maximizing one's probability of arriving ontime, that is arriving within a specified time frame t, under normally distributed edge lengths).

Our problem is also related to the parametric shortest path problem [3], in which the edge lengths, instead of being stochastic, are deterministic linear functions of a variable λ and the question is to enumerate all shortest paths over all values of the variable $\lambda \in [0, \infty)$.

The literature on stochastic route planning and traffic engineering is vast, though most commonly, the different models minimize the expected route cost or length (*e.g.,* [12, 11, 8]. As such, the work in this domain is typically based on stochastic routing formulations of very different nature and solutions. A sample of work that is closest to the problem we consider is [9, 7, 5, 1]. In particular, Loui [7] considers monotone increasing cost functions, however, the algorithms proposed have exponential worst-case running time. On the other hand, Fan *et al.* [5] provide heuristics for different model of adaptive routing, of unknown approximation guarantee. Nikolova *et al.* [9] prove hardness results for a broad class of objective functions and provide pseudopolynomial algorithms. Lim *et al.* [6] provide empirical results showing that the independence assumption of edge distributions does not affect the accuracy of the answer by too much.

[2] An *extreme* point of a set $\mathcal{C}$ is a point that cannot be represented as a convex combination of two other points in the set $\mathcal{C}$.

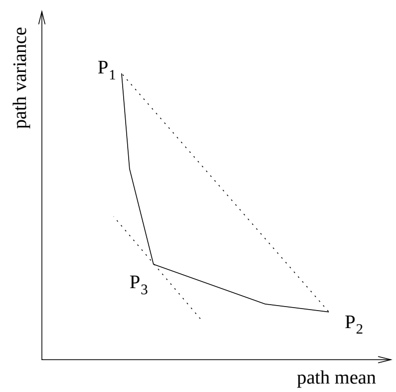

Fig. 1. Finding and enumerating the candidate solution-paths $P_1, P_2, ...$

3 Performance of exact algorithm

The exact algorithm for solving the problem can be thought of as a seach algorithm on the values of the parameter β from Proposition 2. For a given β and a path P with incidence vector $\mathbf{x}$, the *linear* objective $\beta\boldsymbol{\mu}^T\mathbf{x} + (1-\beta)\boldsymbol{\tau}^T\mathbf{x}$ can be separated as a sum of edge weights $\sum_{i\in P}(\beta\mu_i + (1-\beta)\tau_i)$ (note, the variance of the path P equals the sum of variances of the edges along the path, by the independence of the edge length distributions). Thus, finding the ST-route that minimizes the linear objective can be done with any deterministic shortest path algorithm such as Dijkstra, Bellman-Ford, etc. [4] with respect to edge weights $(\beta\mu_i + (1-\beta)\tau_i)$.

Both the exact and heuristic algorithms we consider will consist of a number of calls to a deterministic shortest path algorithm of the user's choice, for appropriately chosen values β: this makes our approach very flexible since different implementations of shortest paths algorithms are more efficient for different types of networks and one can thus take advantage of the most efficient implementations available. We thus characterize the running time performance in terms of the number of such calls or iterations to an underlying shortest path algorithm.

The exact stochastic routing algorithm first sets $\beta = 1$ and $\beta = 0$ and solves the resulting deterministic shortest paths problems (namely it finds the route with smallest mean and the route with smallest variance). Denote the mean and variance of the resulting routes P_1 and P_2 by (m_1, s_1) and (m_2, s_2) respectively. We next set β so that the slope of the linear objective is the same as the slope of the line connecting the points P_1 and P_2 (see Figure 1). Denote the resulting path, if any, by P_3. We continue similarly to find a path between P_1 and P_3 and between P_3 and P_2, etc, until no further paths are found. If there are k extreme points (paths) minimizing some positive linear objective of the mean and variance, then this algorithm finds all of them with $2k$ applications of a deterministic shortest path algorithm.

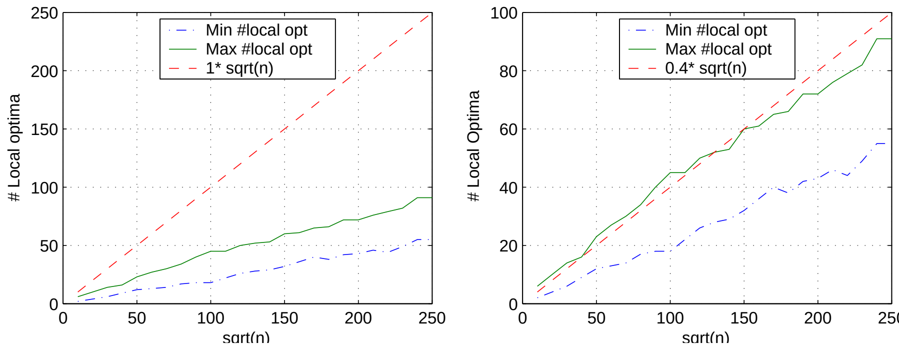

Fig. 2. Number of local optima (extreme point paths) in grids of size 10×10 to 250×250, with edge mean length $\mu_e \leftarrow Uniform[0, 1]$ and variance $\tau_e \leftarrow Uniform[0, \mu_e]$.

In the worst case the number of extreme points k can be $n^{1+\log n}$ [10]—that is superpolynomial (albeit subexponential) and too large to yield an efficient algorithm. However, in most networks and mean-variance values of interest, this number seems to be much lower, thus implying that the exact algorithm may have a good performance in practice.

We thus investigate the performance of the algorithm in a class of networks that are the standard first test set for indicating performance on realistic traffic networks. We consider grid networks of size ranging from 10×10 (100 nodes) to 250×250 ($62,500$ nodes) in additive increments of 10×10. For each network size type $10z \times 10z$ (where $z = 1, ..., 25$), we run 100 instances of the stochastic routing problem. In a single instance, we generate the mean values uniformly at random from $[0, 1]$, and the variance values uniformly at random from the interval $[0, mean]$ for a corresponding edge with an already generated *mean* value. (By scaling all edge means if necessary, we can assume without loss of generality that the maximum mean has value 1.) Out of these 100 simulations per network size, we record the minimum and the maximum number of extreme points and plot them against the square root of the network size (*i.e.,* the square root of the number of nodes in the network). The resulting plots are shown in Figure 2.

To put these empirical results in context: The maximum number of extreme points on a network with $10,000$ nodes found from the simulations, is $k = 45$ (meaning the exact algorithm consisted of only $2k = 90$ iterations of a deterministic shortest path algorithm to find the *optimal* stochastic route) as opposed to the predicted worst case value of $10,000^{1+\log 10,000} \approx 10^{57}$! Similarly, the highest number of extreme points found in graphs of $40,000$ and $62,500$ nodes is $k = 75$ and $k = 92$ respectively as opposed to the theoretical worst-case values of $40,000^{1+\log 40,000} \approx 10^{75}$ and $62,500^{1+\log 62,500} \approx 10^{81}$. In other words, despite the pessimistic theoretical worst-case bound of the exact stochastic routing algorithm, it has a good performance in practice that is orders of magnitude smaller.

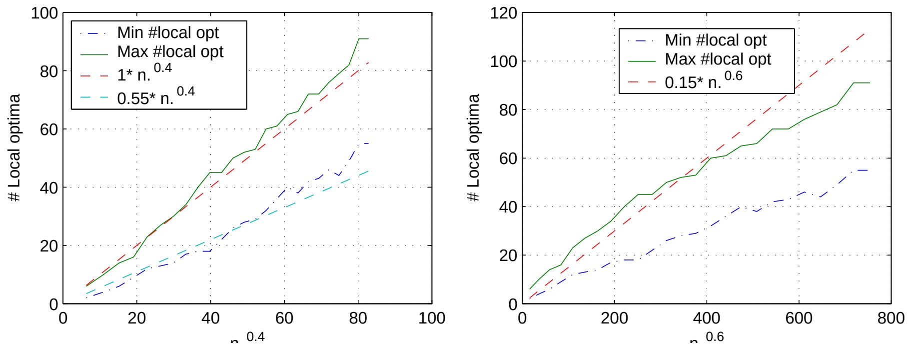

Fig. 3. Number of local optima in grids with n nodes vs $n^{0.4}$ (left) and $n^{0.6}$ (right).

In Figure 2, we have fitted the maximum number of extreme points to a linear function of $\sqrt{n}$: the extreme points for a graph with n nodes are always strictly less than $\sqrt{n}$ and asymptotically less than $0.4\sqrt{n}$. To ensure that $\sqrt{n}$ is the right function to bound the asymptotics, we have also plotted the number of extreme points with respect to $n^{0.4}$ and $n^{0.6}$, in Figure 3. The latter two plots confirm that the rate of growth of extreme points is faster than $\theta(n^{0.4})$ and slower than $\theta(n^{0.6})$.

On the basis of these empirical results, we conclude that a strict upper bound on the number of iterations of the exact algorithm for any grid graph with n nodes is $\sqrt{n}$. We leave as an intriguing open problem to give a theoretical proof of this performance bound.

4 High-performance heuristics

In this section we present a heuristic for finding the optimal stochastic route, which dramatically reduces the running time of the exact algorithm above. Instead of enumerating all extreme points, we select a very small subset of them, leading to high practical performance of the algorithm.

Remarkably, the big reduction of extreme points does not lead to a big sacrifice in the quality of the resulting solution. Our experiments show that even on the large networks of $40,000$ nodes, the heuristic examines only 3 to 6 extreme point paths (compare to the experimental 75 and the theoretical worst bound of $40,000^{1+\log 40,000} \approx 10^{75}$ points of the exact algorithm above), and in all our simulations the value of the solution is within a multiplicative factor of $0.0001 = 0.01\%$ of the optimum.

The heuristic again utilizes Proposition 2, but instead of searching all possible parameter values β that yield a different extreme point (path), it tests an appropriate geometric progression of values and selects the best of the resulting small set of paths.

We illustrate the details of the heuristic through Figure 4*(left)*. This figure plots all extreme point-paths for a $10,000$-node network (with mean and variance

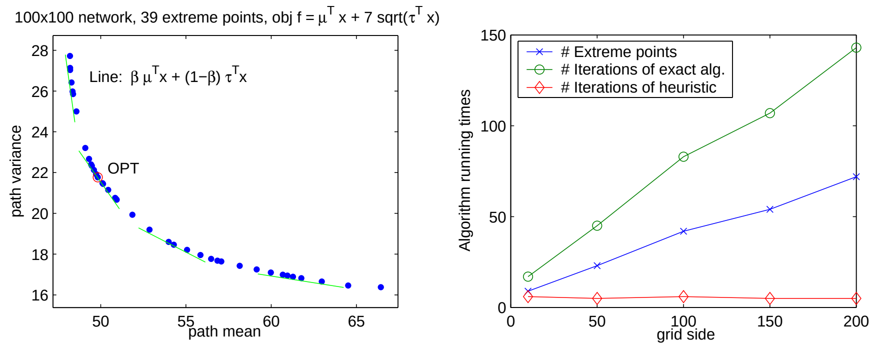

Fig. 4. *(left)* Plot of all extreme point paths of a 10,000-node network. The optimal path is marked with a circle and labelled 'OPT'. The heuristic algorithm minimizes a small number of linear functions as shown, each yielding one extreme point. It then outputs the extreme point (path) with smallest objective function value. *(right)* Number of iterations in the heuristic vs the exact algorithm, in 10x10 up to 200x200 grid networks. Recall that the number of iterations to find all extreme points in the exact algorithm is two times the number of extreme points, where each iteration is a call to a deterministic shortest path algorithm of the user's choice.

values of its edges generated as explained in the previous section). Each point in the figure corresponds to one path in the network with mean equal to the x-coordinate of the point, and variance equal to the y-coordinate of the point. In this mean-variance plot, the optimum happens to be a path with mean 49.83 and variance 21.77 (marked by a circle in the figure). By Proposition 2 and as depicted on the figure, the optimal path minimizes the linear objective $\beta\boldsymbol{\mu}^T\mathbf{x} + (1-\beta)\boldsymbol{\tau}^T\mathbf{x}$ (this corresponds to a line with slope $b = -\frac{\beta}{1-\beta}$) for some range of parameter values β or equivalently for some range of slopes.

If we could guess the optimal slope, then we would find the optimal route with a single iteration of a deterministic shortest path algorithm with respect to edge weights $\beta\mu_i + (1-\beta)\tau_i$. Instead, we test a geometric progression of slopes with a multiplicative step a. The smaller the step a, the more linear objectives we end up testing, which increases the accuracy but also increases the running time of the algorithm.

In our simulations, we experiment with different multiplicative steps. It turns out that using a multiplicative step of 1.01 results in very high accuracy of 99.99%, and also very few extreme point tests and iterations (up to 6 for all graphs with 2,500 to 40,000 nodes).

We compare the running time of the heuristic with that of the exact algorithm in Figure 4*(right)*. This plot gives the highest number of deterministic shortest path iterations that each algorithm has run over 100 simulations per network size. Recall again that the exact algorithm needs to run two times as many iterations as the total number of extreme points. The plot shows that the high performance and accuracy of the heuristic makes it a very practical and promising approach for the stochastic routing problem.

5 Conclusion

We investigated the practical performance of exact and heuristic algorithms for the stochastic routing problem in which the goal is to find the route minimizing a positive linear combination of the route mean and standard deviation. The latter is nonconvex integer optimization problem, for which no efficient algorithms are known. Our experimental results showed that the exact algorithm which is based on enumerating all paths that are potential local optima (extreme points of the feasible set), has surprisingly good running time performance $O(\sqrt{n})R$ on networks of practical interest compared to its predicted theoretical worst-case performance $n^{O(\log(n))}R$, where R is the running time of any deterministic shortest path algorithm of the user's choice. We also showed that a heuristic that appropriately selects a small subset of the potentially optimal paths, has very high performance, using a small constant number of deterministic shortest path iterations and returning a solution that has a 99.99% accuracy. Heuristics of this type are thus a very promising practical approach.

References

1. H. Ackermann, A. Newman, H. Röglin, and B. Vöcking. Decision making based on approximate and smoothed pareto curves. In *Proceedings of 16th ISAAC*, pages 675–684, 2005.
2. Dimitri Bertsekas, Angelia Nedić, and Asuman Ozdaglar. *Convex Analysis and Optimization*. Athena Scientific, Belmont, MA, 2003.
3. Patricia Carstensen. *The complexity of some problems in parametric linear and combinatorial programming*. Ph.D. Dissertation, Mathematics Dept., Univ. of Michigan, Ann Arbor, Mich., 1983.
4. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms, Second Edition*. The MIT Press, September 2001.
5. Y. Fan, R. Kalaba, and II J. E. Moore. Arriving on time. *Journal of Optimization Theory and Applications*, 127(3):497–513, 2005.
6. Sejoon Lim, Hari Balakrishnan, David Gifford, Samuel Madden, and Daniela Rus. Stochastic motion planning and applications to traffic. In *Proceedings of the Eighth International Workshop on the Algorithmic Foundations of Robotics (WAFR), Accepted*, Guanajuato, Mexico, December 2008.
7. R. P. Loui. Optimal paths in graphs with stochastic or multidimentional weights. *Communications of the ACM*, 26:670–676, 1983.
8. E. D. Miller-Hooks and H. S. Mahmassani. Least expected time paths in stochastic, time-varying transportation networks. *Transportation Science*, 34:198–215, 2000.
9. Evdokia Nikolova, Matthew Brand, and David R. Karger. Optimal route planning under uncertainty. In *Proceedings of the International Conference on Automated Planning and Scheduling*, 2006.
10. Evdokia Nikolova, Jonathan A. Kelner, Matthew Brand, and Michael Mitzenmacher. Stochastic shortest paths via quasi-convex maximization. In *Lecture Notes in Computer Science 4168 (ESA 2006)*, pages 552–563, Springer-Verlag, 2006.
11. C. Papadimitriou and M. Yannakakis. Shortest paths without a map. *Theoretical Computer Science*, 84:127–150, 1991.
12. G. Polychronopoulos and J. Tsitsiklis. Stochastic shortest path problems with recourse. *Networks*, 27(2):133–143, 1996.