Distributional Properties of Stochastic Shortest Paths for Smuggled Nuclear Material

Leticia Cuéllar, Feng Pan, Fred Roach, Kevin Saeger
Risk Analysis and Decision Support System D-6
Los Alamos National Laboratory
Los Alamos, NM 87545, USA

Abstract

The shortest path problem on a network with fixed weights is a well studied problem with applications to many diverse areas such as transportation and telecommunications. We are particularly interested in the scenario where a nuclear material smuggler tries to succesfully reach her/his target by identifying the most likely path to the target. The identification of the path relies on reliabilities (weights) associated with each link and node in a multi-modal transportation network. In order to account for the adversary's uncertainty and to perform sensitivity analysis we introduce random reliabilities. We perform some controlled experiments on the grid and present the distributional properties of the resulting stochastic shortest paths.

Keywords: Stochastic shortest path, network.

1 Introduction

There are many well studied optimizaton problems on transportation networks, including optimal routing [3] between selected nodes and interdiction [5]. Many of the existing models are based on the deterministic weights on both nodes and links. Examples include, Dijkstra's and Floyd Marshall's shortest paths algorithm [2]. We are particularly interested in the scenario where a nuclear material smuggler tries to successfully reach his/her target. The Pathway Analysis, Threat Response and Interdiction Options Tool (PATRIOT), a tool developed at the Los Alamos National Laboratory readily handles this problem using a multi-modal world transportation network by identifying the most likely threat pathway to the target. The identification of this path relies on a set of fixed reliabilities associated with each arc and terminal in the network that represent the adversary's perceived probability of successfully traversing the respective links or nodes. This path thus represents the adversary's most reliable path. In addition to the estimated reliabilities, to account for the adversary's uncertainty and to perform sensitivity analysis we relax the restriction of the deterministic link and node reliabilities and consider them as random variables. The main quantities of interest are the resulting stochastic most reliable paths or stochastic shortest paths for short, and their properties. These properties may depend on the network, and on the specific origin and destination. The most reliable path can be transformed to a shortest path problem (see, e.g., [2, exercise 4.39]). Different from the shortest path problem which has efficient algorithms, counting edges on shortest paths in a stochastic network is computationally hard [4]. One approach to solve this problem is based on Monte Carlo simulation [1]. To gain a better understanding on how paths are selected in a transportation grid with uncertainties, we perform some preliminary controlled experiments by performing Monte Carlo simulations on a grid. This paper presents the resulting findings.

2 Stochastic Reliabilities

The reliabilities in PATRIOT are the result of considering various stochastic search and detection processes on the multi-modal transportation network [6]. For the purpose of this study, it suffices to consider one detection process on the network. Letting this process be a Poisson process makes the reliability of each link be given by $e^{-\lambda d}$, where λ stands for the detection rate, and d for the link's length. To keep the distance dependency, we let the detection rate be a random variable rather than assigning a distribution directly to the reliability. In this manner, assuming that all the links' rates are independent, the problem of finding the most reliable path between an origin $\mathcal{O}$ and a destination $\mathcal{D}$, is equivalent to finding the shortest path between $\mathcal{O}$ and $\mathcal{D}$ using for the links' weights the corresponding link's detection rates, namely:

$$\max_{AllPaths(\mathcal{O},D)} \left\{ \prod_{i \in Path(\mathcal{O},D)} e^{-\Lambda_i d_i} \right\} \equiv \min_{AllPaths(\mathcal{O},D)} \left\{ \sum_{i \in Path(\mathcal{O},D)} \Lambda_i d_i \right\}. \quad (1)$$

Assuming the rates to be identically distributed, the prime question we would like to investigate is if the distribution of the random weights matters, and determine what distributional properties are important and how they affect the quantities of interest. We first consider Logormal random weights and then compare to Gamma distributed random weights. We first focus on the resulting lengths of the shortest paths (section 3.1), subsequently on the link usage (sections 3.2), and look for border effects (section 3.2), and conclude in section 4.

3 Stochastic Shortest Path on an $n \times m$ Grid

Consider the grid of size $n \times m$ with the origin $(0,0)$ as its lower left corner. This allows us for now to keep the length of all links equal to one. Let $(0,0)$ be the source (or origin) and let $(n,0)$, $(n,1) \cdots (n,m)$ be the destinations. Assign to the $(n+1) \times (m+1)$ links in the specified grid independent and identically distributed random weights, and for each realization find the shortest path or the path that minimizes the sum of the weights to each of the m possible destinations $(n,0)$, $(n,1) \cdots (n,m)$. Figure 1 shows the grid 10×10 with some sample paths. Note that for constant equal weights an *optimal* shortest path to destination (n,j) is of length $n+j$ (n steps to the right and j steps up), and that there are a total of $\binom{n+j}{n}$ such shortest paths! In contrast, using random weights can produce shortest paths that are of length greater than the *optimal* length as the yellow path in figure 1 illustrates. The questions is to determine under what conditions this will be the case, and figure out the behavior of the shortest path length's distribution.

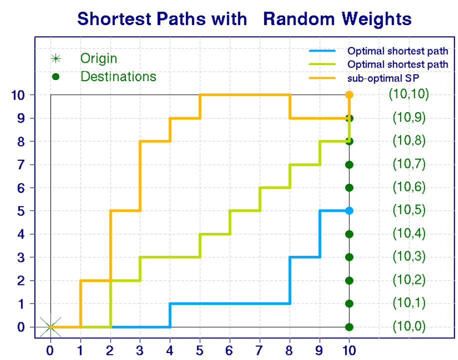


Figure 1: Stochastic shortest paths on a grid.

Weight's Distribution Since we are interested only in positive weights, we consider two distributions: the Lognormal and the Gamma distributions, both with the same means and variances. Recall that the Lognormal random variable is such that its logarithm is a Normal random variable. We let $\mu = 10$ and $\sigma = 1, 5, 10, 20$. Figure 2 shows the densities of Lognormal and Gamma random variables that are very similar for small variances, but differ as the variance increases.

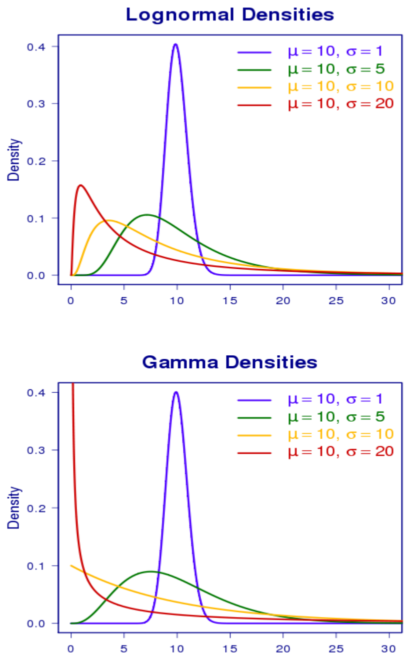


Figure 2: Logormal and Gamma Densities.

3.1 Shortest Path Length

We investigate how the path length changes depending on the distribution of the weights. For each set of parameters (mean equal to 10 and standard deviation equal to 1, 5, 10, and 20) we run 200 Monte Carlo simulations. For each realization of the weights we compute the shortest path between the origin and the m destinations (n,m), and for each of these shortest paths we compute its length, and plot for each destination the corresponding mean path length.

3.1.1 Lognormal Distribution

Figure 3 shows plots of the corresponding mean shortest path length for the grids 10×10 and 20×20. Some things to observe are:

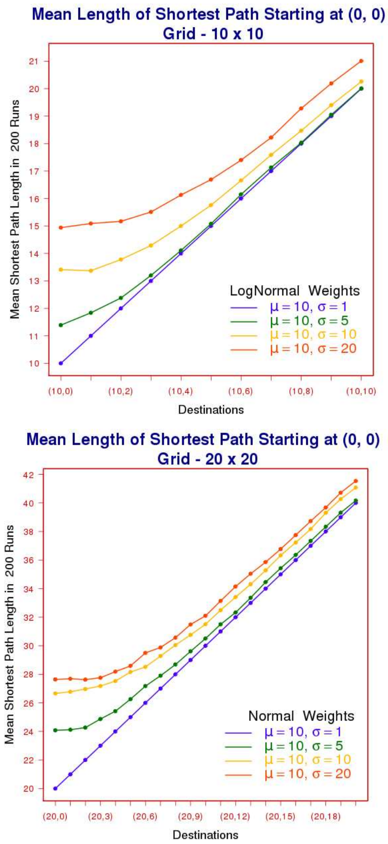


Figure 3: Simulations on a 10×10 and a 20×20 grids.

- For small variance the shortest path length is equal to the optimum length for each of the destinations. In fact, for $\sigma = 1.5$, there is small increase of 0.08 in the mean length of the shortest path only for destination $(10,0)$. For $\sigma = 3.5$ the mean length increase by one extra step for destination $(10,0)$ and produces a slight increase for destinations $(10,1)$ and $(10,2)$. It takes almost a value of $\sigma = 5$ to have an effect on all the destinations. As a consequence, for small variance there is no need to run any Monte Carlo simulations.

- The higher the variance the longer the shortest path. In fact, if we run the same simulations for say $\mu' = c\mu$ and $\sigma' = c\sigma$, where $c > 0$, we get the exact same plots. That's because a linear transformation of the weights ($W_i^{'} = cW_i$), yields the same shortest path. Thus, the quantity that really matters is the signal to noise-ratio, which is the reciprocal of the coefficient of variation

$$c_v = \frac{\sigma}{\mu}. \tag{2}$$

The general statement is then that the larger the coefficient of variation, the longer we expect the resulting shortest path to be.

- It seems that the greater the distance between the origin and the destination, the length of the stochastic shortest path gets closer to the optimum. This phenomena can also be observed when comparing the simulations on the 10×10 grid to the same set of simulations on the 20×20 grid (Figure 3).

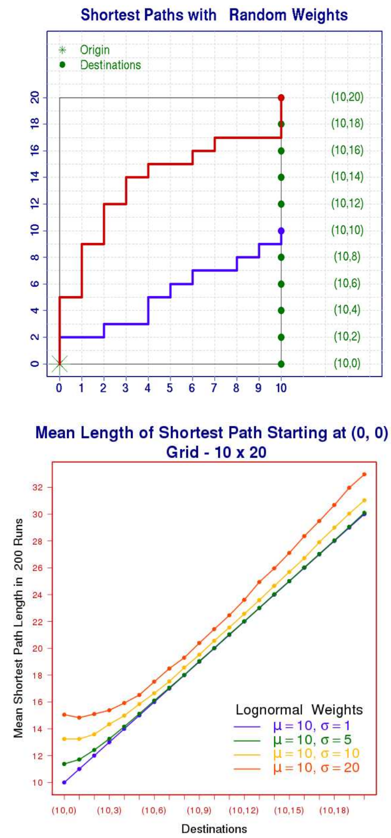


Figure 4: Simulations on a 10×20.

To verify the last statement we modify the simulation by taking a 10×20 grid as shown in Figure 4 (a). The resulting mean lengths shown in Figure 4 (b) disprove this statement: even though as the destinations get closer to (10, 10), or (n,n) in general, the number of shortest paths with optimal length increases making it feasible to find one of them, but as the destinations get beyond $(10,10)$, it seems that it is harder to find an optimal shortest path. Thus there is another factor, besides the distance between the origin and the destination, that affects the length of the path.

Next we explore if these results hold when we change the weigths distribution.

3.1.2 Gamma Distribution

We run the same Monte Carlo simulations using independent and identically distributed Gamma weights with mean $\mu = 10$ and standard deviations $\sigma = 1, 5, 10, 20$. These selections of moments correspond to the Gamma shape parameters $100, 4, 1$, and 0.25 and scale parameters $0.1, 2.5, 10$, and 40 respectively. Recall that the Gamma density has a similar shape as the Lognormal density for σ equal to 1 and 5 (see Figure 2) , thus we would expect to get similar results in these cases.

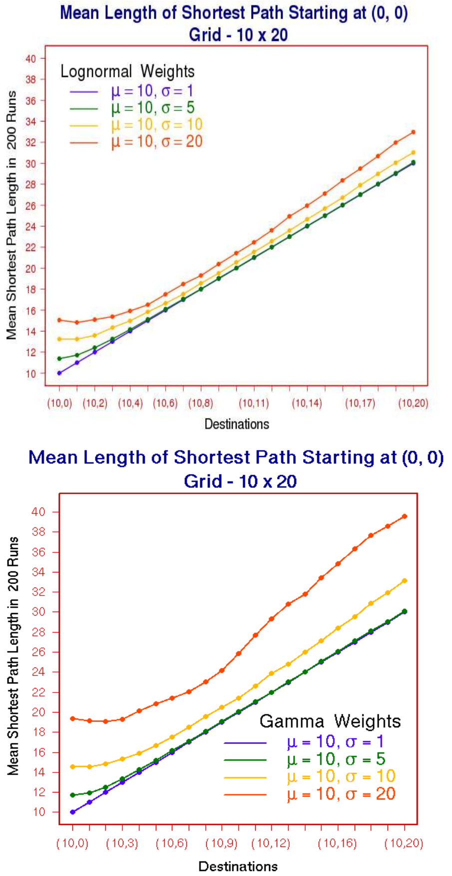


Figure 5: Lognormal vs. Gamma weights on 10×20 grid.

That is confirmed by Figure 5 that displays the mean shortest path lengths when using Lognormal and Gamma weights. For higher variance the stochastic shortest paths using Gamma weights become much longer than their Lognormal counterparts as can be seen in Figures 5 and 6. Longer paths are a result of encountering several heavy weights on the way that literally block the road. Thus the Gamma distribution must produce more heavy weights than the Lognormal one. Indeed, for $\sigma = 10$ and $\sigma = 20$, the Gamma density has a heavier tail than the Lognormal density. For $\sigma = 20$, and the Gamma produces many more lower weights, having a median less than two compared to the corresponding Lognormal median that is about two and a half times bigger. Similarly, as in the Lognormal case, mutiplying all

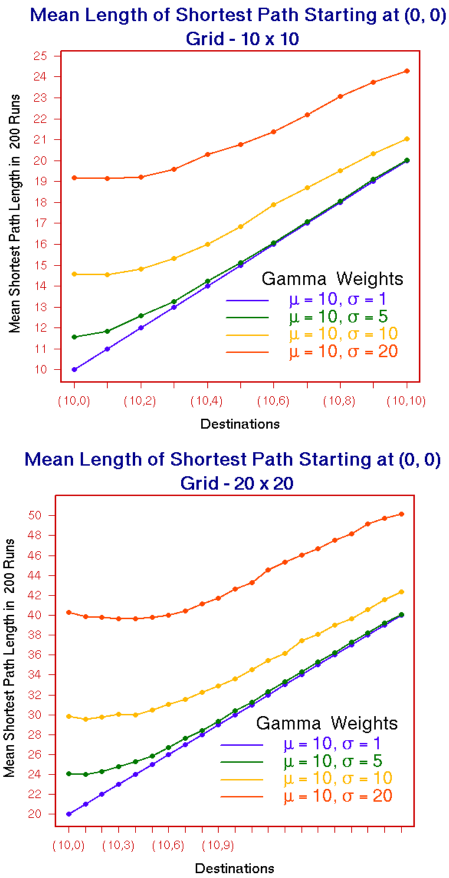


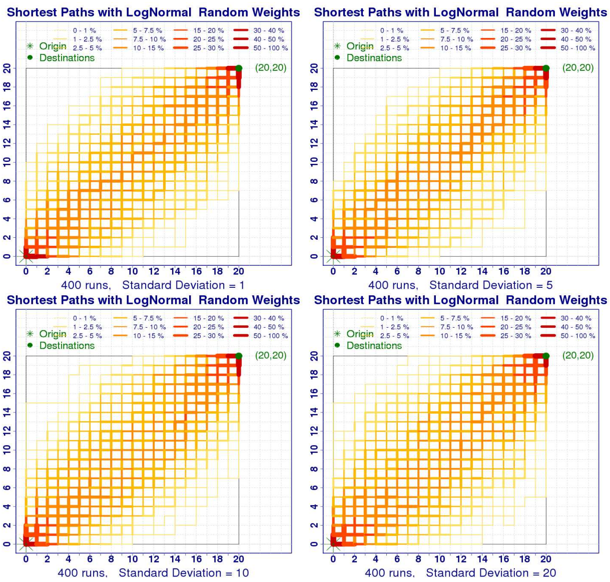
Figure 6: Simulations on 10×10 and 20×20 grids.

weights by a factor $c > 0$ does not affect the resulting shortest paths, thus the quantity that affects the random shortest path is the coefficient of variation.

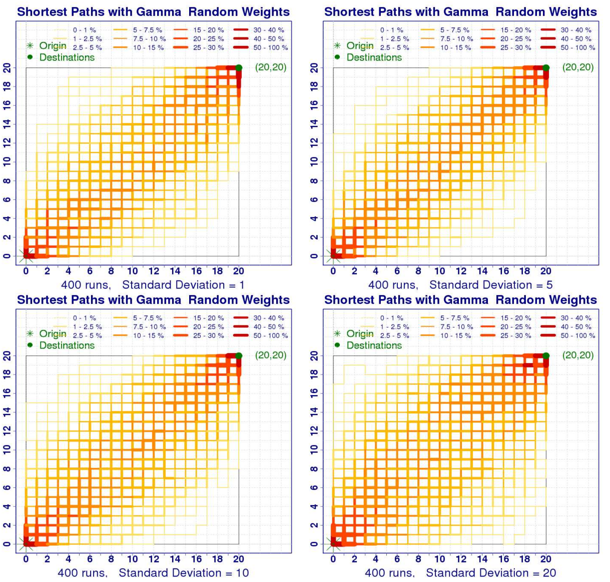
3.2 Edge Frequency

Next we investigate the edge frequency: for each scenario we run 400 Monte Carlo simulations and we calculate the percentage of times every edge was used. Figures 7 and 8 compare runs performed with Lognormal and Gamma weights on 20×20 and 10×20 grids respectively. Some straightforward observations:

- Edges in the *direct* path from the orgin to the destination are visited more frequently, this are the edges by the diagonal. This result was at first somewhat surprising since we expected equally edge usage given that we have independent and identically distributed random weigths on all the edges. But if we consider all the optimal length paths $\binom{40}{20}$ for the runs on the 20×20 grid and $\binom{30}{20}$ for the ones on the 10×20 grid, most of those shortest paths traverse the edges that are located on the *direct* path between the origin and the respective destination, and edges that are *better* connected. For example, the most frequently visited edges are the edges close to the origin and to the destination. To illustrate, the edges $(0,0)$ to $(0,1)$ and $(0,0)$ to $(1,0)$ are each visited about half the time since they are only two possible paths out from the origin (0, 0) and they are both equally likely (becuase they weights are independent and ientically distributed). Also, note that half of all the optimal length shortest paths go through the edge $(0,0)-(0,1)$ and the other half go through $(0,0)-(1,0)$. On the other extreme we have edges that are included in only a few optimal length shortest paths, and are very unlikely to be visited, e.g. the edge $(0,19)-(0,20)$ is included in only one optimal length shortest path, and it was actualy not visited in any of our 400 runs.

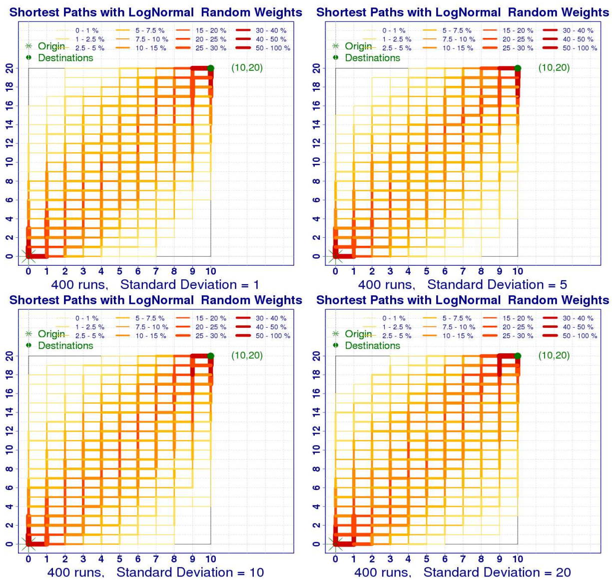


(a)

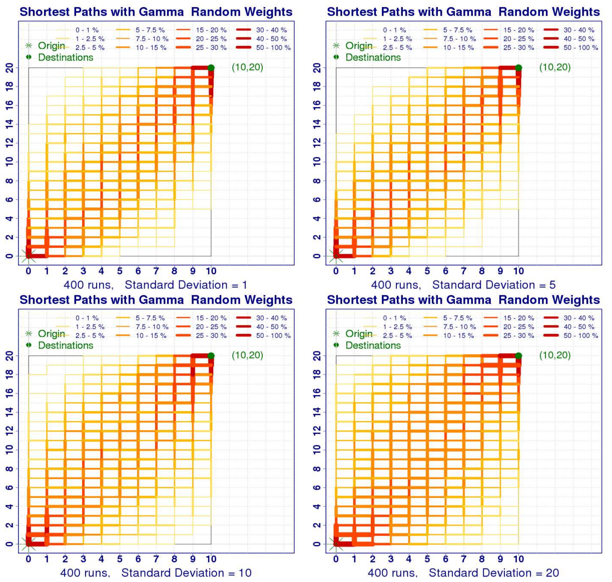


(b)

Figure 7: Edge frequencies for stochastic shortest paths from (0, 0) to (20, 20) using (a) Lognormal and (b) Gamma weights with $\mu = 10$ and $\sigma = 1, 5, 10, 20$.



(a)



(b)

Figure 8: Edge frequencies for stochastic shortest paths from (0, 0) to (10,20) using (a) Logormal and (b) Gamma weights on the 10×20 grid using $\mu = 10$ and $\sigma = 1, 5, 10, 20$.

- For both distributions it is clear that the higher the variance (the larger the coefficient of variation) the more edges are visited.

- For small variance, the plots produced using Lognormal and Gamma weights are very similar, but for higher variance, the plots with Gamma weights indicate that more edges are visited. This is to be expected, since from the previous section, Gamma weigths with higher variance produce longer paths.

3.3 Border Effects

To check if border effects alter our findings, we rerun the same scenarios on larger grids.

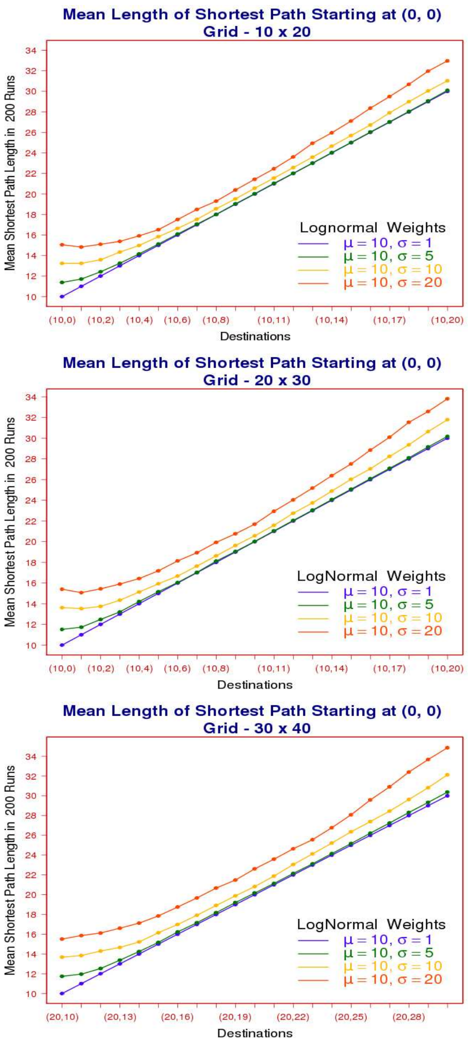


Figure 9: Comparing Mean SP Lengths for various grid sizes.

3.3.1 Shortest Path Length

The larger the grid the longer we expect the shortest path to be (i.e. the total number of hops needed from origin to destination), since there are more possible paths. This is confirmed by the first two plots in Figure 9 that show the mean length of the shortest path from the origin to each of the destinations $(10,0),\cdots,(10,20)$ for Monte Carlo runs on 10×20 and 20×30 grids; the third plot in figure 9 shows the comparable mean shortest path lengths from $(10,10)$ to $(20,10),\cdots,(20,30)$ on the 30×40 grid.

3.3.2 Edge Frequency

The length of the shortest path seems to increase with the size of the grid, but only up to a point. Using Lognormal random weights, we computed the shortest path from the origin $(10,10)$ to each of the destinations $(20,10),\cdots,(20,30)$ using a 30×40 grid, leaving plenty of space between the origin and the destinations and all the grid's borders. Figure 10 displays the edge frequency when traveling to $(20,20)$. The grid 30×40 is only partially displayed since the edges missing had frequency zero. These paths are equivalent to the shortest paths between $(0,0)$ and $(10,10)$ on the previous grids.

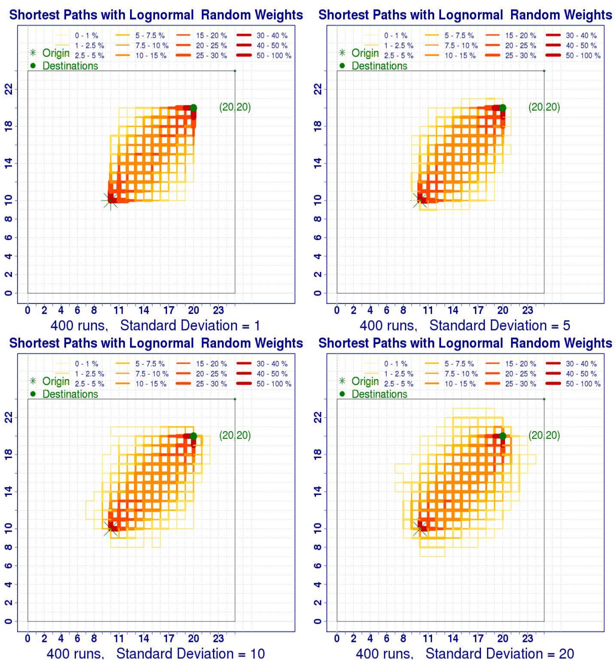


Figure 10: Edge frequency for all shortest paths produced with Lognormal random weights on a 30×40 grid for paths between $(10,10)$ and $(20,20)$.

Note that as the variance increases more edges are visited, and that now some of the shortest paths travel outside the grid defined by the four corners $(10,10),(20,10),(10,20)$ and $(20,20)$. The paths extended beyond this grid by no more than four units in any directions.

Next, Figure 11 shows the edge frequency when traveling from $(10,10)$ to $(20,30)$. We observe a similar behavior as before, the higher the variance the more edges are visited, and since the distance between the origin and the destination is larger, the excursions extend farther out beyond the grid defined by $(10,10),(20,10),(10,30)$ and $(20,30)$, namely by more than six and three untis in the horizontal and vertical directions respectively.

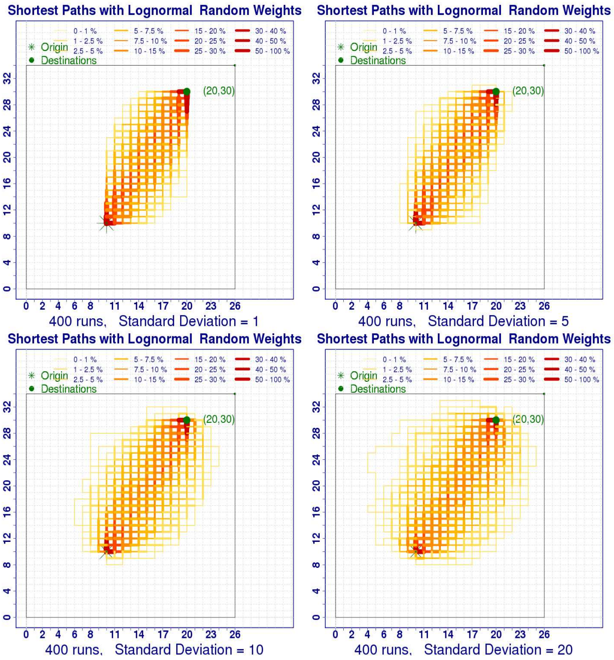


Figure 11: Edge Frequency for shortest paths produced with Lognormal random weights on a 30×40 grid for paths between (10, 10) and (20, 30).

4 Conclusion

To better understand the most reliable stochastic path for smuggling special nuclear material on a transportation network we perform some controlled experiments on a grid using positive random weights. We study the length and the link usage of shortest paths obtained by randomizing the links' weights on a regular grid of size $n \times m$. We found that the diversity and the length of the resulting shortest paths depend on the distribution of the weights. Since the weights are independent and identically distributed the quantity that regulates the bevahiour of the random shortest paths is the coefficient of variation c_v: the larger the coefficient of variation the longer their shortest path, and the more edges are visited.

The experiments were performed on relatively small grids (up to 30×40). Results from simulations on larger grids should be qualitatively similar, but would requiere more runs.

The two main phenomenas that we observed, but still need to be better understood are that the length of the path does not only depend on the distance between the origin and the destination, and the fact that the most frequently visited edges are those edges in the *direct* path between the origin and the destination that seem to have better accesibility.

References

[1] V. G. Adlakha, An improved conditional Monte Carlo technique for the stochastic shortest path problem, Management Science, 32(10), 1986.

[2] Ahuja, Ravindra K. and Magnanti, Thomas L. and Orlin, James B., Network Flows: Theory, Algorithms, and Applications, Prentice Hall, 1993.

[3] Goczylla, Krzysztof and Cielatkowski, Janusz, Optimal Routing in a Transportation Network, European Journal of Operational Research, 87(2), 1995.

[4] Hagstrom, J. N., Computing the probability distribution of project duration in a PERT network. Networks, 20, 1990.

[5] D. P. Morton and F. Pan and K. J. Saeger, Models for nuclear smuggling interdiction, IIE Transactions, 39(1), 2007.

[6] Saeger, K.J. and L. Cuéllar, Modeling Most Likely Pathways for Smuggling Radioactive and Special Nuclear Materials on a Worldwide Multimodal Transportation Network IEEE International Conference on Technologies for Homeland Security, November, 2010.