On the distance between random events on a network

Article *in* Networks · November 2019

DOI: 10.1002/net.21919

CITATIONS
0

READS
44

3 authors:

Ningji Wei
University at Buffalo, The State University of New York
1 PUBLICATION **0** CITATIONS
SEE PROFILE

Jose L. Walteros
University at Buffalo, The State University of New York
19 PUBLICATIONS **116** CITATIONS
SEE PROFILE

Rajan Batta
University at Buffalo, The State University of New York
187 PUBLICATIONS **4,675** CITATIONS
SEE PROFILE

On the Distance Between Random Events on a Network

Ningji Wei, Jose L. Walteros*, Rajan Batta
Department of Industrial and Systems Engineering
University at Buffalo, The State University of New York
342 Bell Hall, Buffalo, NY, 14260
{ningjiwe, josewalt, batta}@buffalo.edu

Abstract

In this paper, we study several statistical properties regarding the distance between events that take place on random locations along the edges of a given network. We derive analytical expressions for the arbitrary moments of such a distance, its probability density function, its cumulative distribution function, as well as their conditional counterparts for the cases in which the position of one event is known in advance. As part of this study, we implement our developments as a callable library for the Python language, to provide potential users with a computational engine able to calculate and visualize these statistics for any given network. We test our implementation on several networks of different sizes and topologies, analyze some of interesting properties we observed in our experiments, and discuss several applications for our proposed methodology. In particular, we focus our discussion on applications aimed to help with the optimal design of emergency response systems on infrastructure networks.

Keywords: Random events, Distance distributions, Infrastructure networks, Emergency response systems, Spatial random processes.

1 Introduction and Motivation

A wide variety of problems arising from different scientific domains involve the careful study of random events that occur on scattered locations of a given network (3; 10; 15; 24; 44). Among the many different examples, perhaps the studies that have permeated the literature the most, are the ones that involve the analysis of events that take place on infrastructure networks (2). Of particular importance, given the major role they play in our society, are the studies aimed to optimally design emergency response systems (ERS), like police, fire, and medical services (11; 25; 29; 40). In this particular context, a detailed analysis of the location of criminal acts, car accidents, and medical emergencies, among others, can be enormously beneficial to maximize the coverage and efficiency of such systems.

Motivated by the fact that most ERS operations require one to rapidly dispatch specialized units in response to randomly occurring critical incidents, there has been a notable interest in designing methodological tools to help analyze several random properties regarding such events (4). Moreover, with the recent emergence of new technologies that facilitate the collection, storage, and processing of large amounts data, the need for new developments able to estimate and characterize locations, distances, and the nature of such random events has become even more prevalent in recent years (11).

*Corresponding author. Phone: 716-645-8876; Fax: 716-645-3302; Address: 413 Bell Hall, Buffalo, NY 14260; Email: josewalt@buffalo.edu

One of the key components pertaining to the study of random events in networks is the statistical characterization of the distances between them. Given that the events take place on random locations across a given network, studying the random variables representing these distances can bring valuable information regarding the dynamics behind these complex systems (27).

A few examples of key problems in the context of ERS, whose solution approaches can benefit from these methodologies are: 1) the design of districting policies to partition large operational areas into smaller districts to minimize the probability that two occurring events are separated by more than a predefined threshold; 2) the design of random patrolling operations to effectively cover areas with high rates of criminal events, without generating noticeable patterns that can be predicted by potential criminals; 3) the identification of location and dispatching policies for response units to minimize the expected traveling distance to attend incoming emergency services.

Interestingly, there are many other domains on different sides of the research spectrum that can also benefit from studying these distances. For instance, in the context of urban transportation systems, such as ride-sharing services, the random events can be assumed to be the arrival of new customers in need of a ride. Based on historical service data, these methodologies can help the operators of ride-sharing systems to preposition the drivers during inactive periods, to capture a larger number of customers, as well as to balance the transportation load (2). Similarly, studying the distances between service requests on wireless communication networks can improve the routing of information by predicting potential interference (19).

Following this line of research, this paper develops a general framework that allows for the characterization of several statistical properties of the distance between random events on a network. We aim to derive analytical expressions for the arbitrary moments, the probability density function, the cumulative distribution function, and their conditional counterparts (i.e., the resulting functions given that the position of one of the events is known). As part of this study, we also implement these developments into a callable library for the Python language, to provide potential users with a computational engine able to calculate and visualize these statistics for any given network[1].

We focus our attention on events that take place on random locations over the edges of a network and assume that the probability distribution of the location of such events is known in advance. Precise estimations of such distributions can be typically obtained by analyzing historical data. Interestingly, thanks to the GPS functionality that is prevalent in most electronic devices nowadays, collecting rich data about the location of any type of event has become a common task for most industries and services (28). In fact, it is not unreasonable to find rich data sets of this type in the public domain already. To name a few examples, comprehensive data sets with information about car accidents, as well as crime incidents in many areas of the US are available to the public in (1; 31). Similarly, databases with information about pick-up locations of taxi trips in several cities of the US, like New York City and Chicago, are also open for public use (13; 39).

2 Literature Review

The problem we study in this paper belongs to a larger class of problems whose main objective is to study the distance between random points on a given metric space. Some of the most well-studied variations are cases developed for compact path-connected subspaces of $\mathbb{R}^2$ or $\mathbb{R}^3$, with norms L_1 or L_2 acting as the metric functions (27; 33). There are three main aspects that define the different variations found on the literature: the geometric shape of the space, the metric function used to measure the distances in the space, and the distribution of the location of the points (12; 38).

[1]Our implementation is publicly available at `https://github.com/tidues/RanDist`

At the most general core, there are the studies that model the problem over a bounded plane. For example, the probability distributions of the distance between uniformly and Gaussian distributed points in a rectangular area were studied in (26); the expected distance between two points in grids, hypercubes, and shufflenets were studied in (34); the cumulative distribution functions of the distance between points inside a regular n-sided polygon were studied in (22); and the distances inside hyperspheres and $\mathbb{R}^k$ subspaces have been studied in (18).

Early applications of this approach in the context of facility location can be found in (6) and (14). Due to the huge density difference between the inner-city and highway road networks, these approaches approximate city networks as bounded connected subspaces of $\mathbb{R}^2$ and model them as vertices within a large network connected by edges representing the highways. While the results from these papers have been adopted to frame location problems in several contexts (9), the underlying metric space used therein poses major limitations in terms of the accuracy and modeling abilities when dealing with practical applications. More specifically, the assumption that the underlying space is a compact, dense subset of $\mathbb{R}^k$ implies that metric norms like L_1 or L_2 are used to represent the distances, which often results in notable imprecisions. This is particularly challenging when modeling infrastructure networks, as their underlying spaces are not two-dimensional manifolds, but their topology can be better described by a graph.

Surprisingly, the problem of studying the distance between random events on networks has received significantly less attention than the variation over subspaces of $\mathbb{R}^2$. A general explanation for this is that in many contexts the events are assumed to take place on the vertices of the networks and not on random locations along the edges (17). The characterization of the distance between random vertices is a rather simple exercise of conditional probability that is generally derived case-by-case based on the given context. There are however, a few studies that generalize some of these concepts and use them to model applications in communications and chemistry (15; 20; 35; 43).

A different line of research that is also worth mentioning focuses on studying distances between the vertices of random graphs (7; 19; 34; 37). There are two main areas covered by these studies. The first assumes that the topology of the graphs is not given in advance, but some of the vertices and their respective connections are added to or removed from the graph via a known random process. A typical example of this type of graphs are the networks that are created by the users of wireless communication networks at a given point in time. The second area mainly covers studies related to the distance between the members of social networks (8) and focuses on analyzing the degree of separation between members over graphs that have characteristics like being scale-free or having small-world properties (41). The distances in this context typically represent the number of hops between vertices instead of other more general metric values.

When studying events that occur on any place along the edges of a network, one possible modeling approach is to discretize each edge with a chain of auxiliary vertices, so that all the potential locations where the events can take place are captured by these additional vertices. The fundamental problem that exists behind this approach is the emergence of a trade-off between the accuracy and complexity of the resulting model. That is, to get a more precise representation of the event locations, one may need to increase the granularity of the discretization by introducing a larger set of auxiliary vertices, which may in turn pose some computational challenges when using the resulting network due to its size.

In this paper, we are interested in developing a general characterization of the random variables that represent the distance between random events taking place on any location along the edges of a given network (including the endpoints). The proposed framework is general enough to be used as a stepping stone to model a wide variety of applications.

This paper is organized as follows. In Section 3, we give a formal characterization of the problem at hand and describe some important notation. In Section 4, we study the kth order moments of

these random distances (for any arbitrary $k \in \mathbb{N}$); we develop analytical expressions for them and discuss some interesting properties. In Section 5, we derive the probability density functions and the cumulative distribution functions of the distances. In Section 6, we present a formal characterization of the domain for these probability functions and develop further insights on how to compute them more efficiently. In Section 7, we provide analytical formulas for all the conditional statistics. In Section 8, we provide some interesting properties of the probability density function when the distribution of the event locations are modeled as uniform random variables. In Section 9, we discuss the complexity of evaluating each of the expressions we developed and provide further details regarding our implementation. In Section 10, we discuss some interesting applications for optimizing several tactical decisions when designing emergency response systems on infrastructure networks. In Section 11, we present some computational experiments. Finally, in Section 12, we conclude and offer further insights.

3 Problem Definition and Notation

We consider a simple, connected, undirected graph $G = (V, E)$, where $V = \{1, \ldots, n\}$ is a set of n vertices and $E \subseteq \binom{V}{2}$ is a set of m edges. We will refer to each edge by its index e or by its unordered pair of endpoints $\{u, v\}$. For a given edge e, we also say that e_1 and e_2 are the pair of endpoints of e and assume by convention that $e_1 < e_2$, according to the corresponding index of the vertices in V. We associate the edge set E with a non-negative vector $[l_e]_{e \in E}$ that comprises the *lengths* of the edges in E. For a given pair of vertices $u, v \in V$, let d_{uv} be the *distance* between such vertices with respect to the edge length values $[l_e]_{e \in E}$; i.e., d_{uv} denotes the length of a shortest path between u and v in G. We assume that $d_{uu} = 0$, for all $u \in V$. Importantly, we consider here the more general case in which the distances are not necessarily restricted by the triangle inequality; therefore, it is possible that for some $e = \{u, v\}$, $d_{uv} \leq l_e$.

In this paper, we study random events that occur on the edges of graph G. The specific location of an event is given by a pair (e, p), where e indicates the edge where the event takes place, and p the *relative location* of the event on edge e; i.e., p is the length of the edge segment from e_1 to the location of the event, divided by l_e. Figure 1 depicts an example of two of such events on a small network.

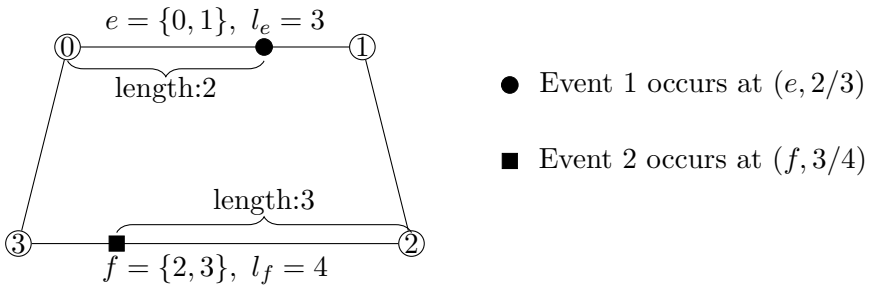


Figure 1: The location of two random events on a network

Given two random events that occur on G—referred to as *Event 1* and *Event 2*—we use the pairs of random variables $(X, P(X))$ and $(Y, Q(Y))$ to describe the location where such events take place. Here, X and Y are discrete random variables that represent the edges where Events 1 and 2 occur. Then, for any given pair of realizations $X = e$ and $Y = f$, $P(e)$ and $Q(f)$ are the continuous random variables that represent the relative location $p, q \in [0, 1)$ of the events on the corresponding edges e and f in E. We assume that the joint probability mass function (pmf) $g_{X,Y}$

for random variables X, Y, and the joint probability density function (pdf) $\phi_{P(e),Q(f)}$ for every pair of $(e, f) \in E^2$ are given, and that $\phi_{P(e),Q(f)}$ is Lebesgue integrable and bounded for all $e, f \in E$. Under these settings, we are interested in the random variable D that describes the (shortest) distance between the two events. We will particularly focus on deriving analytical expressions for the arbitrary order moments, the probability density functions (pdf), the cumulative distribution function (cdf), and the conditional counterparts of D given joint distributions $g_{X,Y}$ and $\phi_{P(e),Q(f)}$. In what follows, to reduce the complexity of the notation, we will refer to $P(X)$ and $Q(Y)$ simply as P and Q, still noting their dependence on random variables X and Y, respectively.

3.1 Notation and Definitions

3.1.1 Special sets and multi-index notation

We define I as the interval $[0, 1)$, $\mathbb{2}$ as the set $\{1, 2\}$, and

$$\mathbb{2}^2 = \{(1, 1), (1, 2), (2, 1), (2, 2)\},$$

as the Cartesian product of $\mathbb{2}$ with itself.

A multi-index $\alpha = (\alpha_i)_{i=1}^n$ is a n-tuple in $\mathbb{N}^n$ (we take the convention that $0 \in \mathbb{N}$) with a norm $|\alpha| = \sum_{i=1}^n \alpha_i$. When applying $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ as the power of $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we have

$$(x_1, x_2, \ldots, x_n)^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}.$$

Also, when α is used as a multinomial coefficient for any $k \in \mathbb{N}$ so that $|\alpha| \leq k$, we have

$$\binom{k}{\alpha} = \frac{k!}{\alpha_1! \alpha_2! \cdots \alpha_n! (k - |\alpha|)!}.$$

3.1.2 Special functions

We define the gate function $\eta_a^b : \mathbb{R} \to \{0, 1\}$ for any a, b in the extended real line $[-\infty, \infty]$ as

$$\eta_a^b(x) = \begin{cases} 0, & x < a \\ 1, & a \leq x \leq b \\ 0, & x > b. \end{cases},$$

and use it to control the nonzero domain of any function on $\mathbb{R}$ by multiplication. For a defined interval $K = [a, b]$, we will also denote the gate function $\eta_a^b(x)$ as $\eta_K(x)$.

We define the threshold function $\theta_a^b : \mathbb{R} \to \mathbb{R}$ as

$$\theta_a^b(x) = \begin{cases} a, & x < a \\ x, & x \in [a, b] \\ b, & x > b \end{cases},$$

and use it to limit the non-constant domain of any function on $\mathbb{R}$ by composition. We use θ_a and θ^b to represent $\theta_a^{+\infty}$ and $\theta_{-\infty}^b$ respectively. Similarly, η_a^∞ and $\eta_{-\infty}^b$ are denoted η_a and η^b.

An important property of the gate and threshold function is that:

$$(\theta_a^b)'(x) = \eta_a^b(x),$$

except at points a, b where the derivative is undefined. Here, $(\theta)'(x)$ denotes the first derivative of the threshold function with respect to x.

For any set $K \subseteq \mathbb{R}^k$ and a given pdf $\phi : K \to \mathbb{R}$, the mass operator $m_\phi(K)$ is defined as

$$m_\phi(K) = \int_K \phi = Prob\{x \in K\}.$$

If the pdf function ϕ is clear from the context, we will drop the subscript ϕ.

3.1.3 Partition (mod 0)

Given a measure space $(X, \mathcal{A}, \mu)$, where X is a set, $\mathcal{A}$ is a σ-algebra on X, and μ is a measure on $(X, \mathcal{A})$, a partition of X is defined as a family $\xi = \{C_k\}_{k \in K} \subset 2^X$ of non-empty disjoint subsets of X whose union is X. Often, there is a need for more general versions of the partition concept, like the case in which either the elements in ξ may have non-empty, but zero-measure intersections, or the case where ξ is a partition of $X' \subsetneq X$ with $\mu(X') = \mu(X)$. The latter case has been referred to as a partition (mod 0) in (36). We generalize this concept to incorporate both of the aforementioned cases.

Definition 1. Given a measure space $(X, \mathcal{A}, \mu)$, and an equivalence relation defined on $\mathcal{A}$ such that $A \sim B$ if $\mu(A \triangle B) = 0$, where $\triangle$ is the symmetric difference operator on sets, then a family $\{C_k\}_{k \in K}$ is a partition (mod 0) of X if there exists some partition $\{C'_k\}_{k \in K}$ of X such that $C_k \sim C'_k$ for all $k \in K$. In particular, a cover of X whose elements intersect only on sets of measure 0 is a partition (mod 0) of X.

Remark. If $(X, \mathcal{A}, \mu)$ is a probability space, and $\xi = \{C_k\}_{k \in K}$ is a partition (mod 0) of X, then the law of total probability and total expectation are valid for ξ. This is because we can pick $\xi' = \{C'_k\}_{k \in K}$ as the partition of X that corresponds to ξ. Since ξ' is a partition of the sample space X, both laws are valid conditioning on each C'_k in ξ'. But each C_k is equivalent to C'_k, so $\mu(C_k) = \mu(C'_k)$ for all $k \in K$; that is, the probability of each case is still same if we replace the partition ξ' by ξ, the partition (mod 0).

In the following sections, we use the laws of total probability and total expectation over partitions and partitions (mod 0) of several probability spaces. We will henceforth refer to both as partitions, unless further clarifications are needed.

4 The kth Moment of D

In this section, we will derive an analytical formula for $E[D^k]$—the kth order moment of D. As a consequence, we will also have an analytical moment-generating function for D. To this end, we first condition on the edges where the events take place, and then condition on their relative positions.

For a given pair of realizations e, f of random variables X and Y, we use D^{ef} to denote the shortest distance between the two events, given that they occur on edges e and f. Then, by total expectation, we have

$$\begin{aligned} E[D^k] &= \sum_{(e,f) \in E^2} g_{X,Y}(e,f) E[D^k \mid X = e, Y = f] \\ &= \sum_{(e,f) \in E^2} g_{X,Y}(e,f) E[(D^{ef})^k] \end{aligned} \quad . \tag{1}$$

where $E[D^k \mid X = e, Y = f]$ is the conditional kth moment of D.

We consider the cases where the events occur on the same or different edges. In each case, we explore the possible values of the pair (P, Q).

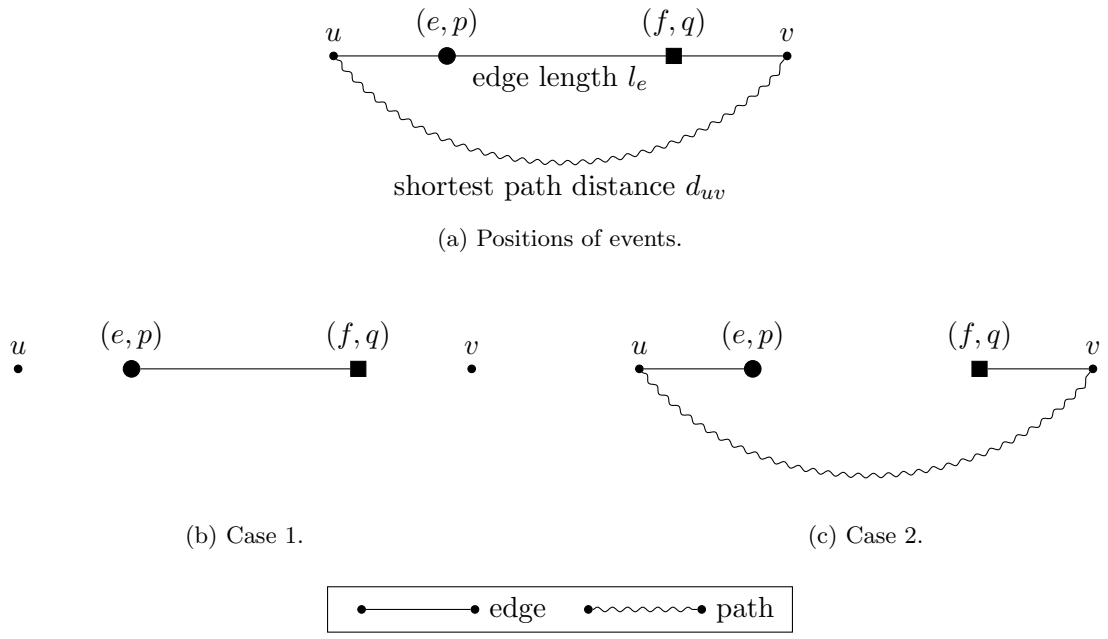
4.1 Scenario 1: $e = f$

In this section, we assume both events happen on the same edge, say $e = \{u, v\}$. For the sake of simplicity, we abuse notation and drop the superindices for all the edge pair dependent symbols, unless further clarifications are needed. For example, the distance D^{ee} will be written as D.

Under the assumption $e = f$ and given that $P = p, Q = q$, two possible cases may happen (see Figure 2):

Case 1: the shortest path between the events is along edge e, in which case the corresponding distance is $l_e|p - q|$,

Case 2: the shortest path passes through the two endpoints of edge e, in which case the corresponding distance is $d_{uv} + l_e(1 - |p - q|)$.



(a) Positions of events.

(b) Case 1.

(c) Case 2.

Figure 2: Two cases of the shortest path.

We can further split each of these cases into two additional subcases based on the relative locations p and q and refer to those as D_{11}, D_{12}, D_{21} and D_{22}, respectively. Then, the shortest distance under each case is

$$\begin{array}{lcll} D_{11}(p,q) & = & l_e(p-q) & \text{Case 1 subcase 1, } p \geq q \\ D_{12}(p,q) & = & -l_e(p-q) & \text{Case 1 subcase 2, } p < q \\ D_{21}(p,q) & = & d_{uv} + l_e - l_e(p-q) & \text{Case 2 subcase 1, } p \geq q \\ D_{22}(p,q) & = & d_{uv} + l_e + l_e(p-q) & \text{Case 2 subcase 2, } p < q \end{array} \tag{2}$$

or in general,

$$D_{ij}(p,q) = \hat{c}_{ij} + (-1)^{i+j} l_e p + (-1)^{i+j-1} l_e q \quad \forall (i,j) \in \mathbb{2}^2$$

where $\hat{c}_{ij} = (i-1)(d_{uv} + l_e)$ is constant. By the definition of the shortest path, we have

$$D = \begin{cases} \min\{D_{11}, D_{21}\}, & p \geq q \\ \min\{D_{12}, D_{22}\}, & p < q. \end{cases} \tag{3}$$

We now proceed by conditioning on the relative locations p and q. We will show that it is possible to partition the region I^2 into four subregions, referred to as $\{R_{ij}\}_{(i,j)\in 2^2}$, such that R_{ij} defines the domain where the minimum distance D is equal to D_{ij}. From such a partition, $E[D^k \mid X = Y = e]$ is then given by

$$\begin{aligned} E[D^k \mid X = Y = e] &= \sum_{(i,j)\in 2^2} Prob\{(p,q) \in R_{ij}\} E[D^k \mid X = Y = e, R_{ij}] \\ &= \sum_{(i,j)\in 2^2} m(R_{ij}) E[D^k_{ij} \mid X = Y = e, R_{ij}], \end{aligned} \tag{4}$$

where the mass operator m is defined over the joint pdf $\phi_{P,Q}$.

Proposition 1. *Let $R_{ij} = \{(p,q) \in I^2 \mid D = D_{ij}\}$, then $R = \{R_{ij}\}_{(i,j)\in 2^2}$ is a partition (mod 0) of I^2.*

Proof. We will prove R is a partition (mod 0) (as defined in Section 3.1.3) by showing it is a cover of I^2 whose elements only intersect on sets of measure zero. Clearly, $\bigcup_{(i,j)\in 2^2} R_{ij} = I^2$, since for any point $(p,q) \in I^2$, the shortest distance between the corresponding events (e,p) and (e,q) is always described by at least one of the functions $D_{ij}(p,q)$, so it will belong to some R_{ij}. It is left to show that the intersection between any two regions R_{ij} and R_{kl} has 2-dimensional measure zero. From (2), we notice that all D_{ij}s lie on different 2-dimensional hyperplanes embedded in $\mathbb{R}^3$. Therefore, the set $\{(p,q) \in \mathbb{R}^2 \mid D_{ij} = D_{kl}, (i,j) \neq (k,l)\}$ is the projection of the intersection of D_{ij} and D_{kl} onto the pq-plane, which is at most a line. Then, the intersection $R_{ij} \cap R_{kl} = \{(p,q) \in I^2 \mid D = D_{ij} = D_{kl}\}$ is a subset of such a line, thus it also has 2-dimensional measure zero. □

To visualize the partition R, we can find all the boundary lines by solving $D_{ij} = D_{kl}$ for the cases given in (2). We present such partition in Figure 3. For further reference, we name these boundary lines h_0, h_1, and h_2, respectively.

Equality	Line Name	Line
—	h_0	$q = p$
$D_{11} = D_{21}$	h_1	$q = p - (d_{uv} + l_e)/2l_e$
$D_{12} = D_{22}$	h_2	$q = p + (d_{uv} + l_e)/2l_e$

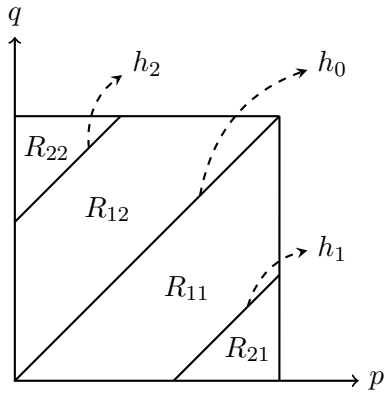


Figure 3: Boundary lines and partition R

Based on this partition, from (4), we have

$$\begin{aligned} E[D^k \mid X = Y = e] &= \sum_{(i,j)\in 2^2} m(R_{ij}) E[D_{ij}^k \mid X = Y = e, R_{ij}] \\ &= \sum_{(i,j)\in 2^2} m(R_{ij}) E[(\hat{c}_{ij} + (-1)^{i+j} l_e P + (-1)^{i+j-1} l_e Q)^k \mid R_{ij}]. \end{aligned}$$

Using the multi-index notation $\alpha = (\alpha_1, \alpha_2)$ of Section 3.1.1, we expand the term

$$\begin{aligned} & \left(\hat{c}_{ij} + (-1)^{i+j} l_e P + (-1)^{i+j-1} l_e Q\right)^k \\ = & \sum_{|\alpha|\le k} \binom{k}{\alpha} ((-1)^{i+j-1} l_e Q)^{\alpha_1} ((-1)^{i+j} l_e P)^{\alpha_2} \hat{c}_{ij}^{k-|\alpha|} \\ = & \sum_{|\alpha|\le k} \binom{k}{\alpha} (-1)^{(i+j-1)\alpha_1 + (i+j)\alpha_2} l_e^{|\alpha|} \hat{c}_{ij}^{k-|\alpha|} (Q, P)^{\alpha}, \end{aligned}$$

which yields

$$\begin{aligned} E[D^k \mid X = Y = e] &= \sum_{(i,j)\in 2^2} m(R_{ij}) \sum_{|\alpha|\le k} c_{ij\alpha} E[(Q, P)^{\alpha} \mid R_{ij}] \\ &= \sum_{(i,j)\in 2^2} \sum_{|\alpha|\le k} c_{ij\alpha} m(R_{ij}) E[(Q, P)^{\alpha} \mid R_{ij}] \\ &= \sum_{(i,j)\in 2^2} \sum_{|\alpha|\le k} c_{ij\alpha} m(R_{ij}) \frac{\int_{R_{ij}} (q, p)^{\alpha} \phi_{P,Q}(p, q)}{m(R_{ij})} \\ &= \sum_{(i,j)\in 2^2} \sum_{|\alpha|\le k} c_{ij\alpha} \int_{R_{ij}} (q, p)^{\alpha} \phi_{P,Q}(p, q), \end{aligned} \tag{5}$$

where

$$c_{ij\alpha} = \binom{k}{\alpha} (-1)^{(i+j-1)\alpha_1 + (i+j)\alpha_2} l_e^{|\alpha|} \hat{c}_{ij}^{k-|\alpha|}$$

is a constant.

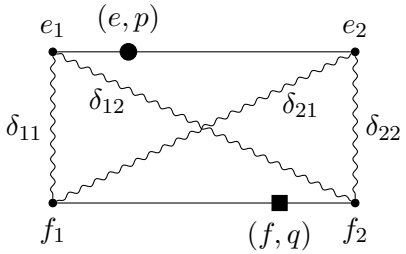
In what follows, we explore similar partitions for the case in which the events occur on different edges. Given the context, whenever there is the need of differentiating to which case the partition is referring to, we add the superindices ee or ef, respectively (i.e., R_{ij}^{ee} or R_{ij}^{ef})

4.2 Scenario 2: $e \neq f$

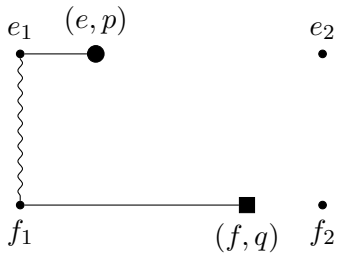
Now, we assume both events X and Y happen on different edges e and f. Thus, the shortest path between the two events could be one of four possible cases, each corresponding to the shortest path through nodes e_i and f_j for $i, j \in \mathbb{2}$ (see Figure 4). The distances of these cases are then $d_{e_i f_j}$ for $(i, j) \in \mathbb{2}^2$. For the sake of simplicity, we define $\delta_{ij} = d_{e_i f_j}$ and as before, we drop the superindex for D^{ef} and all other symbols that depend on ef.

As for the previous scenario, D_{ij} is the associated length of the path in each case,

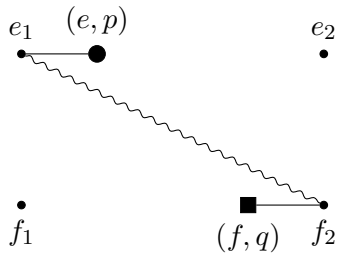
$$\begin{aligned} D_{11}(p, q) &= \delta_{11} + l_e p + l_f q \\ D_{12}(p, q) &= \delta_{12} + l_e p + l_f (1 - q) \\ D_{21}(p, q) &= \delta_{21} + l_e (1 - p) + l_f q \\ D_{22}(p, q) &= \delta_{22} + l_e (1 - p) + l_f (1 - q) \end{aligned} \tag{6}$$



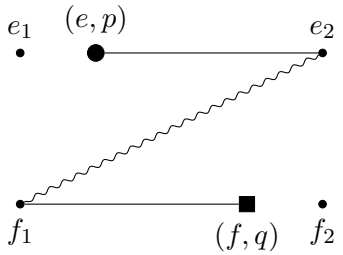
(a) Positions of the two events.



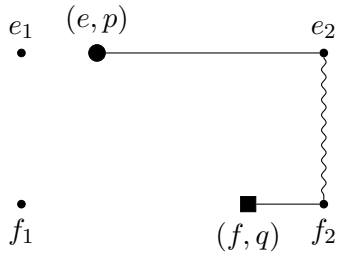
(b) Case $(1,1)$.



(c) Case $(1,2)$.



(d) Case $(2,1)$.



(e) Case $(2,2)$.

Figure 4: Four cases of the shortest path.

and in general,

$$D_{ij}(p,q) = \hat{c}_{ij} + (-1)^{i-1} l_e p + (-1)^{j-1} l_f q$$

where $\hat{c}_{ij} = \delta_{ij} + ((-1)^i)^+ l_e + ((-1)^j)^+ l_f$ is a constant. By definition, the shortest distance between the two events is,

$$D(p,q) = \min_{(i,j)\in\mathbb{2}^2} D_{ij}(p,q) \tag{7}$$

Again, we define $R_{ij} = \{(p,q) \in I^2 \mid D_{ij} = D\}$ as the region in which D_{ij} is minimum. Following a similar argument as in Proposition 1, it can be shown that $R = \{R_{ij}\}_{(i,j)\in\mathbb{2}^2}$ is a partition of I^2. We will provide a detailed characterization of R in Section 4.2.1. Given such a partition, we have

$$\begin{aligned} E[D^k \mid e,f] &= \sum_{(i,j)\in\mathbb{2}^2} m(R_{ij}) E[D_{ij}^k \mid R_{ij}] \\ &= \sum_{(i,j)\in\mathbb{2}^2} m(R_{ij}) E[\left(\hat{c}_{ij} + (-1)^{i-1} l_e P + (-1)^{j-1} l_f Q\right)^k \mid R_{ij}] \end{aligned}$$

Let $\alpha = (\alpha_1, \alpha_2)$ be a multi-index as defined in Section 3.1.1. We expand the term

$$\begin{aligned}
&\left(\hat{c}_{ij} + (-1)^{i-1} l_e P + (-1)^{j-1} l_f Q\right)^k \\
= &\sum_{|\alpha| \le k} \binom{k}{\alpha} ((-1)^{(j-1)} l_f Q)^{\alpha_1} ((-1)^{(i-1)} l_e P)^{\alpha_2} \hat{c}_{ij}^{k-|\alpha|} \\
= &\sum_{|\alpha| \le k} \binom{k}{\alpha} (-1)^{(j-1)\alpha_1 + (i-1)\alpha_2} (l_f, l_e)^{\alpha} \hat{c}_{ij}^{k-|\alpha|} (Q, P)^{\alpha},
\end{aligned}$$

which leads to

$$\begin{aligned}
E[D^k \mid e, f] &= \sum_{(i,j) \in 2^2} m(R_{ij}) \sum_{|\alpha| \le k} c_{ij\alpha} E[(Q, P)^{\alpha} \mid R_{ij}] \\
&= \sum_{(i,j) \in 2^2} \sum_{|\alpha| \le k} c_{ij\alpha} m(R_{ij}) E[(Q, P)^{\alpha} \mid R_{ij}] \\
&= \sum_{(i,j) \in 2^2} \sum_{|\alpha| \le k} c_{ij\alpha} \int_{R_{ij}} (q, p)^{\alpha} \phi_{P,Q}(p, q),
\end{aligned} \tag{8}$$

where

$$c_{ij\alpha} = \binom{k}{\alpha} (-1)^{(j-1)\alpha_1 + (i-1)\alpha_2} (l_f, l_e)^{\alpha} \hat{c}_{ij}^{k-|\alpha|}$$

is a constant.

Now, we proceed to characterize the partition R.

4.2.1 The partition of R when $e \neq f$

We first notice that the partition R is determined by six lines that correspond to six boundary cases. That is, $D_{ij} = D_{lk}$ for all $(i, j), (l, k) \in 2^2$. Solving this equality for all the cases in (6) gives us six lines in the space $\mathbb{R}^2$.

Equality	Line Name	Line	Constant Name
$D_{11} = D_{12}$	h_{21}	$q = (-\delta_{11} + \delta_{12} + l_f)/2l_f$	q_1
$D_{11} = D_{21}$	h_{11}	$p = (-\delta_{11} + \delta_{21} + l_e)/2l_e$	p_1
$D_{11} = D_{22}$	h_3	$l_e p + l_f q = (-\delta_{11} + \delta_{22} + l_e + l_f)/2$	
$D_{12} = D_{21}$	h_4	$l_e p - l_f q = (-\delta_{12} + \delta_{21} + l_e - l_f)/2$	
$D_{12} = D_{22}$	h_{12}	$p = (-\delta_{12} + \delta_{22} + l_e)/2l_e$	p_2
$D_{21} = D_{22}$	h_{22}	$q = (-\delta_{21} + \delta_{22} + l_f)/2l_f$	q_2

Table 1: Boundary lines

For instance, solving the equality $D_{11} = D_{12}$ yields the line $q = (-\delta_{11} + \delta_{12} + l_f)/2l_f$. We name this line h_{21}. Notice that this line is parallel to the p-axis since the *right hand side* of the line expression is a constant. We will show later that this constant dictates the shape of partition R. Given the importance of these constants, we give them shorter names for future reference (e.g., we use q_1 for this case). The full information of the six boundary cases is presented in Table 1.

Lemma 1. *The six boundary lines described in Table 1 have the following properties:*

1. *Given edges e and f with edge length l_e, l_f and the shortest path distances $\delta_{ij} = d_{e_i f_j}$ for all $(i,j) \in \mathbb{2}^2$, the values of p_1, p_2, q_1, q_2 lie within the range $[0,1]$.*
2. *h_3 passes through the points $h_{11} \cap h_{22}$ and $h_{12} \cap h_{21}$, and h_4 passes through $h_{11} \cap h_{21}$ and $h_{12} \cap h_{22}$.*
3. *Define $\Delta = \delta_{12} + \delta_{21} - (\delta_{11} + \delta_{22})$, then $\Delta \leq 0$ implies $p_1 \leq p_2$ and $q_1 \leq q_2$, the equality takes place if and only if $\Delta = 0$. $\Delta > 0$ implies $p_1 > p_2$ and $q_1 > q_2$.*

Proof. For Property 1, we only prove the case for q_1, as the other cases are similar. Assume q_1 is either less than 0 or greater than 1. After simplifying the inequalities $q_1 < 0$ and $q_1 > 1$ (q_1 was defined in Table 1), we get $\delta_{11} > \delta_{12} + l_f$ and $\delta_{12} > \delta_{11} + l_f$, but both cases imply the δ_{ij} on the left hand side is not a shortest path from vertex e_i to f_j, which contradicts the definition of δ_{ij}.

Property 2 can be proved by simply calculating the intersection points $h_{11} \cap h_{22}$ and $h_{12} \cap h_{21}$, then plugging the results back into h_3 and h_4.

For Property 3, after simplifying the inequalities $p_1 \leq p_2$ and $q_1 \leq q_2$, we obtain $\Delta \leq 0$, which is satisfied at equality if and only if $p_1 = p_2$ and $q_1 = q_2$. □

With this lemma, we can now determine the possible shapes of partition R.

Lemma 2. *The shape of partition $R = \{R_{ij}\}_{(i,j)\in\mathbb{2}^2}$ of domain I^2 is determined by the value*

$$\Delta = \delta_{12} + \delta_{21} - (\delta_{11} + \delta_{22}),$$

i.e., there are two possible shapes of R that correspond to $\Delta \leq 0$ and $\Delta > 0$ respectively.

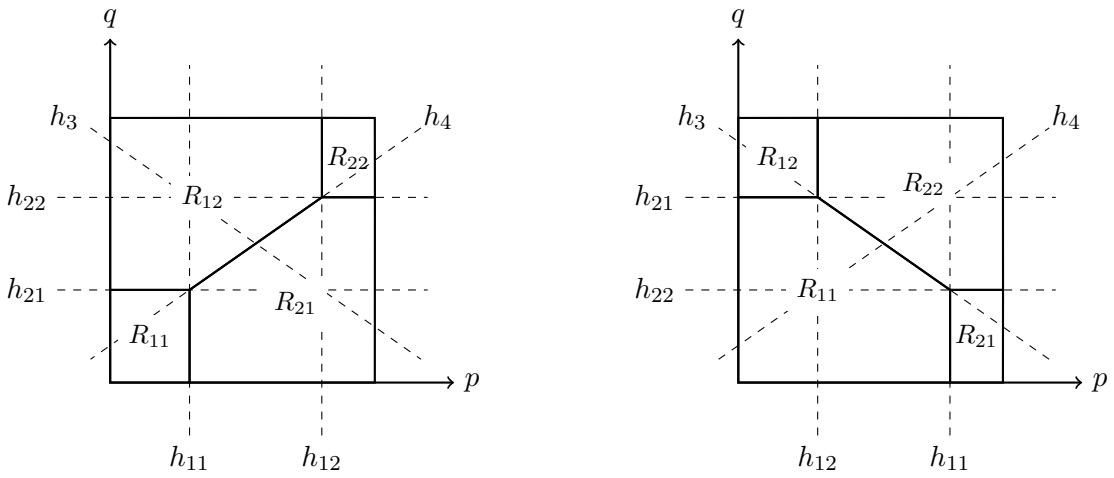
Proof. When $\Delta \leq 0$, from Properties 1 and 3 of Lemma 1, we know $0 \leq p_1 \leq p_2 \leq 1$ and $0 \leq q_1 \leq q_2 \leq 1$. From Table 1, we see that these values determine the position of lines $h_{11}, h_{12}, h_{21}, h_{22}$. Hence, we know that h_{11} lies to the left of h_{12} and h_{21} lies below h_{22}. Then, by Property 2 of Lemma 1, lines h_3 and h_4 are also determined, they pass through two intersection points with the fixed slope $\pm l_e/l_f$. With the relative positions of these six boundary lines determined, we can conclude that R looks like that shown in Figure 5a. Otherwise when $\Delta > 0$, R is illustrated in Figure 5b. □

It is interesting to see the implication of the value Δ. From Figure 4, we can see that Δ is actually the difference of the lengths between two cycles, one is the cycle that passes though the paths with lengths δ_{11} and δ_{22}, the other is the cycle that passes through the paths with lengths δ_{12} and δ_{21}. It is an intrinsic relation that the difference between the lengths of those two cycles determines the shape of the partition R.

Now, applying the law of total expectation, we can merge the results from formulas (5) and (8) to get an analytical formula for the moments of D. We summarize this result in the following theorem.

Theorem 1. *Given two random events (X,P) and (Y,Q) for which $g_{X,Y}$ is the joint pmf of X,Y and $\phi_{P,Q}$ is the joint pdf of P,Q, for each pair of realizations $(e,f) \in E^2$, then, assuming that all $\phi_{P,Q}$ are Lebesgue integrable and bounded, the kth moment of the distance D between such events is given by:*

$$E[D^k] = \sum_{(e,f)\in E^2} \sum_{(i,j)\in\mathbb{2}^2} \sum_{|\alpha|\leq k} c_{ij\alpha}^{ef} \int_{R_{ij}^{ef}} (q,p)^\alpha \phi_{P,Q}(p,q), \tag{9}$$



(a) Case 1: $\Delta \leq 0$ (b) Case 2: $\Delta > 0$

Figure 5: Two cases of the partition R

where

$$c_{ij\alpha}^{ef} = g_{X,Y}(e,f)\binom{k}{\alpha}(l_f, l_e)^{\alpha} \begin{cases} (-1)^{(i+j-1)\alpha_1+(i+j)\alpha_2}((i-1)(d_{uv}+l_e))^{k-|\alpha|}, & e=f=\{u,v\} \\ (-1)^{(j-1)\alpha_1+(i-1)\alpha_2}(d_{e_if_j}+(i-1)l_e+(j-1)l_f)^{k-|\alpha|}, & e \neq f. \end{cases}$$

Remark. Notice that the integrability of $\phi_{P,Q}(p,q)$ is sufficient to guarantee the integrability of $(q,p)^{\alpha}\phi_{P,Q}(p,q)$ because

$$\int |q^{\alpha_1}p^{\alpha_2}\phi_{P,Q}(p,q)| = \int |q^{\alpha_1}||p^{\alpha_2}||\phi_{P,Q}(p,q)| \leq \int |\phi_{P,Q}(p,q)| < +\infty,$$

where the first inequality comes from the fact that p and q are values from the interval $[0,1]$ and the second inequality is due to the integrability of $\phi_{P,Q}(p,q)$.

Here, all the edge pair (e,f) dependent symbols regain their superindices to specify the different cases described above. Notice, that expression (9) also gives us directly the coefficients of the moment-generating function expansion of D, which is defined as

$$M_D(t) = E[e^{tX}] = \sum_{k=0}^{\infty} \frac{t^k E[D^k]}{k!}.$$

5 Probability Distributions of D

Let $\Psi_D(x) = Prob(D \leq x)$ be the cdf of distance D, then the pdf is $\psi_D(x) = \dfrac{d}{dx}\Psi_D(x)$. First, we derive the cdf. To this end, recall the definition of random variable D (equations (3) and (7))

$$D = \begin{cases} D^{ee} = \begin{cases} \min\{D_{11}^{ee}, D_{21}^{ee}\}, & p \geq q \\ \min\{D_{12}^{ee}, D_{22}^{ee}\}, & p < q \end{cases}, & \text{if } X = Y = e \\ D^{ef} = \min\limits_{(i,j)\in 2^2} D_{ij}^{ef}, & \text{if } X = e, Y = f, e \neq f \end{cases}.$$

Similarly as before, we first condition on edges $X = e$ and $Y = f$.

$$\Psi_D(x) = \sum_{(e,f)\in E^2} g_{X,Y}(e,f)\Psi_{D|X,Y}(x \mid e,f),$$

where $\Psi_{D|X,Y}(x \mid e,f) = Prob(D \le x \mid X = e, Y = f)$ is the conditional cdf. Again, we consider two possible cases.

5.1 Scenario 1: $e = f$

In this section, we drop the superscript ee for all edge pair dependent symbols. That is, R_{ij}^{ee} and D_{ij}^{ee} are written as R_{ij} and D_{ij}, respectively.

5.1.1 Formulas for the cdf and pdf

Given the partition R and the distance functions D_{ij}, for $(i,j) \in \mathbb{2}^2$, defined in Section 4.1 (case $e = f$), we have

$$\begin{aligned}
\Psi_{D|e=f}(x) &= \sum_{(i,j)\in\mathbb{2}^2} m(R_{ij})\Psi_{D|e=f,R_{ij}}(x \mid e = f, R_{ij}) \\
&= \sum_{(i,j)\in\mathbb{2}^2} m(R_{ij}) Prob\{D \le x \mid e = f, R_{ij}\} \\
&= \sum_{(i,j)\in\mathbb{2}^2} m(R_{ij}) Prob\{D_{ij} \le x \mid e = f, R_{ij}\}.
\end{aligned} \tag{10}$$

Let region function $R_{ij}(x)$ be the region inside R_{ij} that has the property that $D_{ij}(p,q) \le x$, for a given x. That is,

$$R_{ij}(x) = \{(p,q) \in I^2 \mid D_{ij}(p,q) \le x\} \cap R_{ij}.$$

Then, we have

$$\begin{aligned}
Prob\{D_{ij} \le x \mid e = f, R_{ij}\} &= \frac{Prob\{D_{ij}(p,q) \le x \text{ and } (p,q) \in R_{ij}\}}{Prob\{(p,q) \in R_{ij}\}} \\
&= \frac{Prob\{(p,q) \in \{(p,q) \in I^2 \mid D_{ij}(p,q) \le x\} \cap R_{ij}\}}{Prob\{(p,q) \in R_{ij}\}} \\
&= \frac{Prob\{(p,q) \in R_{ij}(x)\}}{Prob\{(p,q) \in R_{ij}\}} \\
&= \frac{m(R_{ij}(x))}{m(R_{ij})}
\end{aligned}$$

plugging this back into (10), we have

$$\begin{aligned}
\Psi_{D|e=f}(x) &= \sum_{(i,j)\in\mathbb{2}^2} m(R_{ij})\frac{m(R_{ij}(x))}{m(R_{ij})} \\
&= \sum_{(i,j)\in\mathbb{2}^2} m(R_{ij}(x)) \\
&= \sum_{(i,j)\in\mathbb{2}^2} \int_{R_{ij}(x)} \phi_{P,Q}(p,q),
\end{aligned} \tag{11}$$

where the last equality is from the definition of the mass operator m.

Accordingly, the pdf is

$$\psi_{D|e=f}(x) = \sum_{(i,j)\in 2^2} \frac{d}{dx} \int_{R_{ij}(x)} \phi_{P,Q}(p,q) \tag{12}$$

5.1.2 The region functions $R_{ij}(x)$ when $e = f$

The above expression can only be computed if the regions $R_{ij}(x)$ are simple for every possible value x. Here, by simple we mean that the integration of ϕ over $R_{ij}(x)$ can be easily calculated.

To visualize the region function $R_{ij}(x)$, we first consider the domain of x; that is, the upper and lower bounds for the distance between the two events. The upper bound is $(d_{uv} + l_e)/2$ where $e = \{u, v\}$, which results directly from the definition of the edge length l_e and the shortest distance d_{uv} between two endpoints u and v of edge e. The lower bound depends on whether the shortest path is along the edge e or not. In the former case, the lowest possible distance is 0 when both events happen on the same position of the edge e; in the latter case, the lower bound is d_{uv}, for any shorter distance will force one of the events out of edge e, which contradicts the assumption that both events are on the same edge e. These bounds are presented in Table 2.

For any fixed x in its domain, the region $R_{ij}(x)$ is the subset of R_{ij} that is bounded by the inequality $D_{ij}(p, q) \leq x$. The results are illustrated in Figure 6. We will give the characterization of these regions functions later in Section 6.

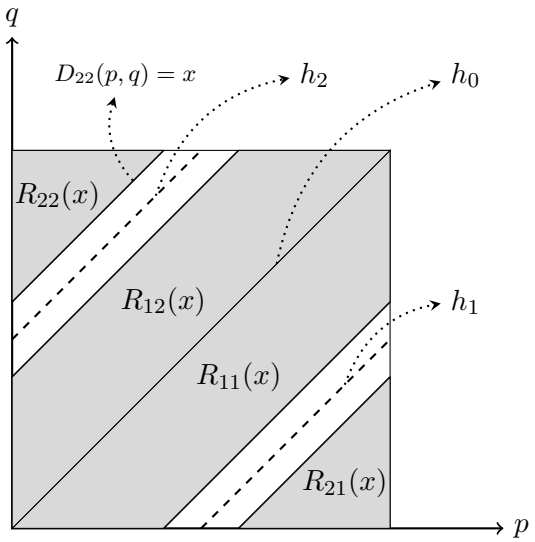


Figure 6: Region function $R_{ij}(x)$

5.2 Scenario 2: $e \neq f$

In this section, we drop the superindex ef for all edge pair dependent notations. That is, R_{ij}^{ef} and D_{ij}^{ef} are written as R_{ij} and D_{ij}, respectively.

5.2.1 Formulas for the cdf and pdf

As for Scenario 1, we define region functions $R_{ij}(x) = R_{ij} \cap \{(p,q) \in I^2 \mid D_{ij} \leq x\}$ for the current case, where the D_{ij} and R_{ij} are shorthands for D_{ij}^{ef} and R_{ij}^{ef} defined in Section 4.2. We have,

$$\begin{aligned}
\Psi_{D|e\neq f}(x) &= \sum_{(i,j)\in 2^2} m(R_{ij})\Psi_{D|e\neq f,R_{ij}}(x \mid e \neq f, R_{ij}) \\
&= \sum_{(i,j)\in 2^2} m(R_{ij}) Prob\{D \leq x \mid e \neq f, R_{ij}\} \\
&= \sum_{(i,j)\in 2^2} m(R_{ij}) Prob\{D_{ij} \leq x \mid e \neq f, R_{ij}\} \\
&= \sum_{(i,j)\in 2^2} m(R_{ij}) \frac{m(R_{ij}(x))}{m(R_{ij})} \qquad (13) \\
&= \sum_{(i,j)\in 2^2} m(R_{ij}(x)) \\
&= \sum_{(i,j)\in 2^2} \int_{R_{ij}(x)} \phi_{P,Q}(p,q).
\end{aligned}$$

Therefore, the pdf can be written as

$$\psi_{D|e\neq f}(x) = \sum_{(i,j)\in 2^2} \frac{d}{dx} \int_{R_{ij}(x)} \phi_{P,Q}(p,q). \qquad (14)$$

5.2.2 The region functions $R_{ij}(x)$ when $e \neq f$

To visualize the region functions $R_{ij}(x)$ for this case, we follow the same approach. First, we consider the domain of x. As shown in Figure 7, in which we use region R_{11} (recall region R_{11} from Figure 5b) as an example, we can determine the domain of x by solving two simple optimization problems $\min_{(p,q)\in R_{ij}} D_{ij}(p,q)$ and $\max_{(p,q)\in R_{ij}} D_{ij}(p,q)$. Graphically, these two values correspond to the cases where the contour of the objective function D_{ij} is binding at the boundaries of the region R_{ij}. That is, when x is at its lowest level, the line $D_{11}(p,q) = x$ passes through the origin, thus $R_{11}(x)$ contains only the single point $(0,0)$; when x is at its upper bound, the line will overlap with the boundary line h_3, therefore the region $R_{11}(x)$ is equal to the region R_{11}; when x is in between, the line $D_{11}(p,q) = x$ is also between these two extreme cases with constant slope.

To solve these two optimization problems, we refer back to the formula

$$D_{11}(p,q) = \delta_{11} + l_e p + l_f q,$$

which was defined in (6). The minimum is obtained when p and q are both 0, so the lower bound is δ_{ij}; the maximum is obtained when then line $D_{11}(p,q) = x$ overlaps with h_3, which yields $x = (l_e + l_f + \min\{\delta_{11} + \delta_{22}, \delta_{12} + \delta_{21}\})/2$. Notice that for a given pair of edges e and f, this upper bound is a constant and does not depend on (i,j). As a reference for later use, we provide the domain information for all the regions in Table 2.

We merge expressions (11), (12), (13) and (14) by the total probability law to obtain the analytical formulas for cdf and pdf of D. We summarize these results in the following theorem.

Theorem 2. *Given two random events (X,P) and (Y,Q) for which $g_{X,Y}$ is the joint pmf of X,Y and $\phi_{P,Q}$ is the joint pdf of P,Q, for each pair of realizations $(e,f) \in E^2$, then, assuming that all*

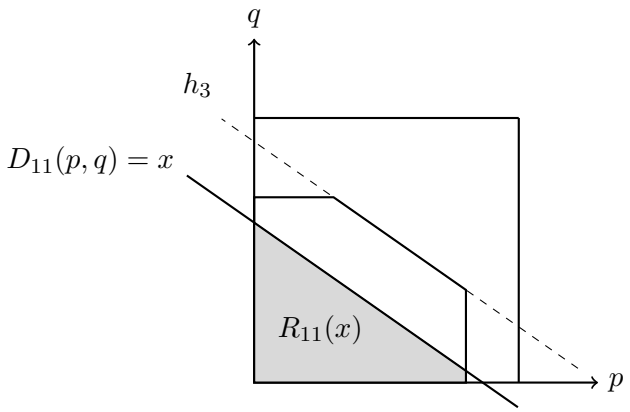


Figure 7: $R_{11}(x)$ inside region R_{11}

Region	Lower Bound a_{ij}^{ef}	Upper Bound b_{ij}^{ef}
R_{11}^{ee}	0	$(d_{uv}+l_e)/2$
R_{12}^{ee}	0	$(d_{uv}+l_e)/2$
R_{21}^{ee}	d_{uv}	$(d_{uv}+l_e)/2$
R_{22}^{ee}	d_{uv}	$(d_{uv}+l_e)/2$
$R_{ij}^{ef},\ e\neq f$	δ_{ij}	$(l_e+l_f+\min\{\delta_{11}+\delta_{22},\delta_{12}+\delta_{21}\})/2$

Table 2: The domain of x

$\phi_{P,Q}$ are Lebesgue integrable and bounded, the cdf and pdf of the distance D between such events follow:

$$\begin{aligned}\Psi_D(x) &= \sum_{(e,f)\in E^2} g_{X,Y}(e,f) \sum_{(i,j)\in\mathbb{2}^2} \int_{R_{ij}^{ef}(x)} \phi_{P,Q}(p,q) \\ \psi_D(x) &= \sum_{(e,f)\in E^2} g_{X,Y}(e,f) \sum_{(i,j)\in\mathbb{2}^2} \frac{d}{dx} \int_{R_{ij}^{ef}(x)} \phi_{P,Q}(p,q)\end{aligned} \tag{15}$$

Remark. Notice that $\phi_{P,Q}$ is Lebesgue integrable, therefore, its integration over the compact regions $R_{ij}^{ef}(x)$ is absolutely continuous. This is also true for $\Psi_D(x)$, because finite summation preserves absolute continuity. Moreover, since any absolutely continuous function has a derivative that is well defined almost everywhere; then, it follows that the pdf $\psi_D(x)$ is well defined almost everywhere. Later in Theorem 3, we will show that there are actually only a finite number of non-differentiable points and it is possible to identify their location.

In Sections 4 and 5, we derived analytical expressions for the arbitrary order moments (9), the cdf, and pdf (15) of D. However, to apply these formulas in practice, we need to provide a detailed characterization of regions R_{ij} and region functions $R_{ij}(x)$. Notice that all formulas require the integration of different functions (e.g., distribution functions $\phi_{P,Q}(p,q)$) over these regions and even though we have characterized them in terms of their boundary lines, a more suitable representation is necessary for the computation of such integrals.

Furthermore, for a given function $\phi_{P,Q}$, to calculate the pdf, we need to first solve an integral followed by its differentiation. This could be numerically unstable because the result of the integration

may be non-smooth and have many non-differentiable points for common input joint distributions $\phi_{P,Q}$ (see Section 8 for the case of uniform distributions). We will address these challenges in the following section.

6 Description of the Regions

All the formulas we derived above have shown the importance of the regions R_{ij}^{ef} and region functions $R_{ij}^{ef}(x)$. Their description will reveal further insights regarding the complexity of the calculation of those formulas. From the definition of these regions and region functions, we notice that $R_{ij}^{ef} = R_{ij}^{ef}(b_{ij}^{ef})$, where b_{ij}^{ef} is the upper bound (see Table 2) of the distance variable x. Therefore, it suffices to focus on the description of region functions $R_{ij}^{ef}(x)$.

For all the possible values of x, we have shown that sets $R_{ij}^{ef}(x)$ are simple polytopes inside I^2. Therefore, we can describe these sets by the tuple

$$(p_l, p_u, q_l(p), q_u(p)),$$

where p_l, p_u describe the lower and upper bounds of p, and $q_l(p), q_u(p)$ are the lower and upper bounds of q for a given value of p. We choose this representation for it allows us to reformulate the integrals over the regions $R_{ij}^{ef}(x)$ as double integrals over the domains of p and q. That is, if the region $R_{ij}^{ef}(x)$ can be described by the form $(p_l(x), p_u(x), q_l(x,p), q_u(x,p))$, then for any given function $\phi(p,q)$ we have the following equivalency,

$$\int_{R_{ij}^{ef}(x)} \phi = \int_{p_l(x)}^{p_u(x)} \int_{q_l(x,p)}^{q_u(x,p)} \phi(p,t)\ dtdp.$$

The tuple representations for all different regions are given in Table 3. We will use the region $R_{11}^{ee}(x)$ (see Figure 8a) as an example to show how to derive this representation. First, we can pick $p_l(x) = 0$ and $p_u(x) = 1$ since this is the domain of p. Then, the lower and upper boundaries of region R_{11}^{ee} are depicted by the thick dashed and solid lines in Figure 8a. The expressions of these lines are given in the table in Figure 3, from where we know the upper bound is the line $q = p$ and the nonzero part of the lower bound is the line $q = p - x/l_e$. Then, the lower bound can be represented by $q = \max\{0, p - x/l_e\}$. Thus, the tuple representation for R_{11}^{ee} is

$$(0,\ 1,\ \max\{0, p - x/l_e\},\ p), \quad x \in [a, b],$$

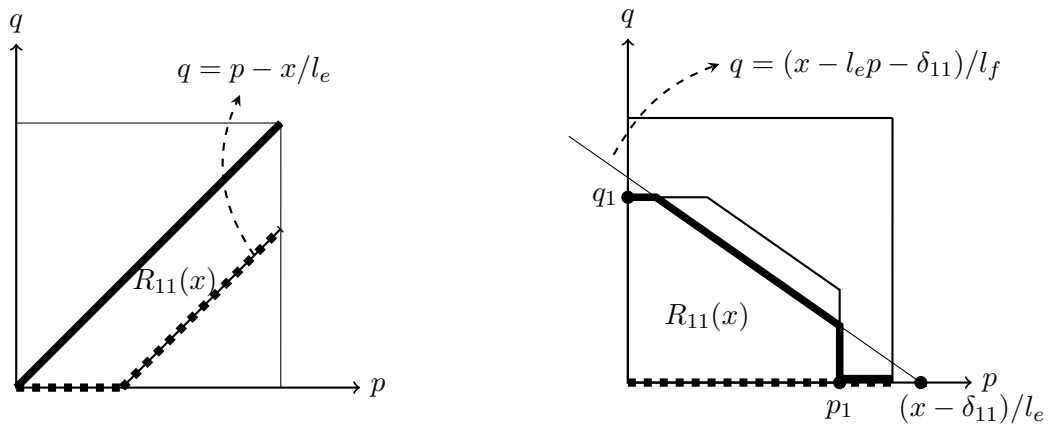
where a and b are the bounds of the domain of x for the case $X = Y = e$ and $(i, j) = (1, 1)$ given in Table 2. To incorporate this domain restriction inside the tuple representation, we replace all the occurrences of x by its adjusted value

$$\hat{x} = \theta_a^b(x),$$

where θ is the threshold function introduced in Section 3.1.2. The property of the threshold function ensures once x is outside its bounds, it will retain the same value. So, the final representation of R_{11}^{ee} is

$$(0,\ 1,\ \max\{0, p - \hat{x}/l_e\},\ p).$$

When $e \neq f$, we do the same. Take $R_{11}^{ef}(x)$ (Figure 8b) as an example. Here, we can still let p_l and p_u be 0 and 1, respectively. Then the lower and upper boundaries are shown in Figure 8b. The lower bound $q_l(x,p)$ (red line) is simply 0. For the upper bound (blue line), we need to first restrict



(a) Region description of $R_{11}^{ee}(x)$

(b) Region description of $R_{11}^{ef}(x)$

Figure 8: Determine the region representation

the nonzero domain of q inside the interval from 0 to $\min\{p_1, (x - \delta_{11})/l_e\}$, where $(x - \delta_{11})/l_e$ is the intersection between the line $q = (x - l_e p - \delta_{11})/l_f$ and the p-axis. We can achieve this by multiplying a corresponding gate function introduced in Section 3.1.2. Once this domain is fixed, the upper boundary is simply the minimum between the line $q = q_1$ and $q = (x - l_e p - \delta_{11})/l_f$. Again, in all the occurrences of x, we need to replace it with its adjusted $\hat{x}$. The representation is then

$$(0,\ 1,\ 0,\ \eta_0^{\min\{p_1,(\hat{x}-\delta_{11})/l_e\}}(p) \cdot \min\{q_1, (\hat{x} - l_e p - \delta_{11})/l_f\}).$$

Region	Boundary	Value
$R_{11}^{ee}(x)$	$q_l(p)$	$\max\{0, p - \hat{x}/l_e\}$
	$q_u(p)$	p
$R_{12}^{ee}(x)$	$q_l(p)$	p
	$q_u(p)$	$\min\{1, p + \hat{x}/l_e\}$
$R_{21}^{ee}(x)$	$q_l(p)$	0
	$q_u(p)$	$\max\{0, p - 1 + (\hat{x} - d_{uv})/l_e\}$
$R_{22}^{ee}(x)$	$q_l(p)$	$\min\{1, p + 1 - (\hat{x} - d_{uv})/l_e\}$
	$q_u(p)$	1
$R_{11}^{ef}(x)$	$q_l(p)$	0
	$q_u(p)$	$\eta_0^{\min\{p_1,(\hat{x}-\delta_{11})/l_e\}}(p) \cdot \min\{q_1, (\hat{x} - l_e p - \delta_{11})/l_f\}$
$R_{12}^{ef}(x)$	$q_l(p)$	$\eta_0^{\min\{p_2,(\hat{x}-\delta_{12})/l_e\}}(p) \cdot \max\{q_1, (-\hat{x} + l_e p + \delta_{12} + l_f)/l_f\}$
	$q_u(p)$	1
$R_{21}^{ef}(x)$	$q_l(p)$	0
	$q_u(p)$	$\eta^1_{\max\{p_1,(-\hat{x}+\delta_{21}+l_e)/l_e\}}(p) \cdot \min\{q_2, (\hat{x} + l_e p - \delta_{21} - l_e)/l_f\}$
$R_{22}^{ef}(x)$	$q_l(p)$	$\eta^1_{\max\{p_2,(-\hat{x}+\delta_{22}+l_e)/l_e\}}(p) \cdot \max\{q_2, (-\hat{x} - l_e p + \delta_{22} + l_e + l_f)/l_f\}$
	$q_u(p)$	1

Table 3: Region description

An interesting observation is that, for any fixed value of p, the region descriptions in Table 3 are step functions in x, except in a small region where a linear non-constant function takes place.

Take $q_l(p)$ in $R_{11}^{ee}(x)$ as an example: the non-constant function is $p - x/l_e$, which is a linear function of x for the given p. The other values are either 0, $p - a/l_e$, or $p - b/l_e$ (a, b are lower and upper bounds), which are all constants. This observation leads us to an alternative expression for the pdf, in which no differentiation operation is needed.

6.1 The pdf Expression Revisited

The pdf formula (15), as stated before, can be numerically unstable when applied directly, due to the potential non-differentiable points that may appear after the integration. Notice that the distance variable x in (15) is not part of the integrand; thus, we may reformulate this expression via the fundamental theorem of calculus. Indeed, according to the region descriptions in the last section, we take advantage of some special properties of these regions to devise an alternative formula for the pdf.

As mentioned in the last section, if a region $R_{ij}^{ef}(x)$ can be described by the tuple

$$(p_l(x), p_u(x), q_l(x,p), q_u(x,p)),$$

then we have

$$\int_{R_{ij}^{ef}(x)} \phi = \int_{p_l(x)}^{p_u(x)} \int_{q_l(x,p)}^{q_u(x,p)} \phi(p,t)\ dtdp.$$

After calculating these region function representations explicitly in Table 3, we observe the following properties:

1. $p_l(x)$ and $p_u(x)$ are constants, so we can simply refer them as p_l and p_u.
2. For all the cases, only one boundary of the inner integration region $[q_l(x,p), q_u(x,p)]$ depends on x. Without loss of generality, we can always assume the upper boundary $q_u(x,p)$ is the x-dependent one and refer to the other simply as $q_l(p)$. This is because we can always exchange the upper and lower bounds of any integration with the adjustment of a negative sign.
3. For a fixed p, the domain of x in the x-dependent function $q_u(x,p)$ can be split into $M(p)$ where $q_u(x,p)$ is equal to some non-constant linear function $q(x,p)$ and the rest where $q_u(x,p)$ is a step function with finite jumps. Taking the lower bound $\max\{0, p - \hat{x}/l_e\}$ of the region $R_{11}^{ee}(x)$ (Table 3) as an example, given p, this is a function in x. Clearly, the main component of this function is the non-constant linear part $p - x/l_e$ and it is easy to verify that it is equal to $p - x/l_e$ when $x \in [0, \min\{d_e p, b_{11}^{ee}\}]$ where b_{11}^{ee} is the upper bound of x (see Table 2); but when x is out of this interval, it is either equal to 0, or the lower bound a_{11}^{ee}, or the upper bound b_{11}^{ee}, which is a step function with finite jumps.

From these properties, we derive the following theorem.

Theorem 3. *If for every $p \in [p_l, p_u]$, a subset $M(p)$ can be found in the domain of x such that*

$$q_u(x,p) = \begin{cases} q(x,p), & x \in M(p) \\ \textit{some step function with finite jumps}, & x \notin M(p) \end{cases}$$

then for any given function ϕ that is Lebesgue integrable and bounded in the region $R(x)$ for all x, the derivative of the function

$$F(x) = \int_{R(x)} \phi = \int_{p_l}^{p_u} \int_{q_l(p)}^{q_u(x,p)} \phi(p,t)\ dtdp$$

is equal to (almost everywhere)

$$\int_{L(x)} \phi(p, q(x,p)) \cdot \partial_x q(x,p)\ dp$$

if $L(x) = \{p \in [p_l, p_u] \mid x \in M(p)\}$ *is measurable.*

This theorem basically states that when we calculate the derivative of such expressions, we are only required to compute the following three things:

1. the non-constant region $L(x)$
2. the non-constant function $q(x,p)$
3. the partial derivative $\partial_x q(x,p)$

and most importantly, the differentiation process does *not* depend on ϕ.

Proof. Since p_l and p_u are not x-dependent, we have

$$\frac{d}{dx}F(x) = \int_{p_l}^{p_u} \left(\frac{d}{dx} \int_{q_l(p)}^{q_u(x,p)} \phi(p,t)\ dt \right) dp.$$

We first notice that p is a fixed value for the inner integral, so we can apply a change of variables. Let $y = q_u(x,p)$ and let $g(y) = \int_{q_l(p)}^{y} \phi(p,t)\ dt$, then by the chain rule we have

$$\frac{d}{dx}g(y(x)) = \frac{d}{dy}g(y) \cdot \frac{dy}{dx}.$$

Given that $y = q_u$, the second term is

$$\frac{dy}{dx} = \begin{cases} \partial_x q(x,p), & x \in M(p) \\ 0 \text{ (almost everywhere)}, & x \notin M(p) \end{cases}.$$

For the first term, we have

$$\frac{d}{dy}g(y) = \frac{d}{dy} \int_{q_l(p)}^{y} \phi(p,t)\ dt.$$

Notice that p is a constant fixed by the outer integration, so we can directly apply the fundamental theorem of calculus, and get

$$\frac{d}{dy}g(y) = \phi(p,y).$$

Multiplying these terms assuming that ϕ is bounded in I^2, we have

$$\frac{d}{dx}g(y(x)) = \begin{cases} \phi(p, q_u(x,p)) \cdot \partial_x q(x,p), & x \in M(p) \\ 0, & x \notin M(p) \end{cases}$$

Now we can do the outer integration by splitting the domain $[p_l, p_u]$ into $L(x)$ and $[p_l, p_u] \setminus L(x)$, that is,

$$\frac{d}{dx}F(x) = \int_{L(x)} \frac{d}{dx}g(y(x,p))\ dp + \int_{[p_l,p_u]\setminus L(x)} \frac{d}{dx}g(y(x,p))\ dp.$$

By the definition of $L(x)$, the integrand of the second intergal is 0, so what is left is the first integral

$$\frac{d}{dx}F(x) = \int_{L(x)} \phi(p, q_u(x,p)) \cdot \partial_x q(x,p)\ dp.$$

Notice, in the region $L(x)$, x is in $M(p)$, thus $q_u = q$, so we are done.

□

Remark. Note that this proof not only validates the fact that the derivative of $F(x)$ is well defined almost everywhere, but also describes where those non-differentiable points may exist. These are the non-differentiable points of $q(x,p)$ for each given p and all the points where q_u has jumps. In particular, there are only a finite number of non-differentiable points, given the definition of $q(x,p)$ (see Table 4).

As mentioned before, one of the benefits from this theorem is that we can calculate the pdf by simply computing three things, the non-constant region, the non-constant function $q(x,p)$, and its partial derivative with respect to x. We use R_{11}^{ee} as an example to show how to obtain the region $L(x)$ as well as the function $q(x,p)$. From Table 3, we see the function $q_l(p) = \max\{0, p - \hat{x}/l_e\}$ is the x-dependent function in this case. When x is within certain domain, the dominant function $q(x,p)$ is $p - \hat{x}/l_e$. The original function equates to this dominant function only if $p - \hat{x}/l_e \geq 0$, which implies the region $L(x) = [\hat{x}/l_e, 1]$. Then, the value $\partial_x q(x,p)$ is simply $-\eta_a^b(x)/l_f$ (in this case $e = f$). In this case, the x-dependent function $q_l(x,p)$ is the lower bound of the inner integration, so we need to adjust the result with a negative sign. The results for all the regions are presented in Table 4.

It is easy to verify that in all the cases, the value $\partial_x q(x,p)$ after adjustment (determined by whether the x-dependent bound is the lower or upper bound of the inner integral) is equal to $\eta_a^b(x)/l_f$. Here a, b are the corresponding lower and upper bounds of x (Table 2). For that reason, we do not include this information in Table 4.

Theorem 3 and Table 4 allow us to express the pdf formula as the following form.

Theorem 4. *Given two random events (X,P) and (Y,Q) for which $g_{X,Y}$ is the joint pmf of X,Y and $\phi_{P,Q}$ is the joint pdf of P,Q, for each pair of realizations $(e,f) \in E^2$, then, assuming that all $\phi_{P,Q}$ are Lebesgue integrable and bounded, the pdf of the distance D between such events is given by:*

$$\psi_D(x) = \sum_{(e,f)\in E^2} g_{X,Y}(e,f)/l_f \sum_{(i,j)\in 2^2} \eta_{K_{ij}^{ef}}(x) \cdot \int_{L_{ij}^{ef}(x)} \phi_{P,Q}(p, q_{ij}^{ef}(x,p))\ dp, \tag{16}$$

almost everywhere, where K_{ij}^{ef} is the interval $[a_{ij}^{ef}, b_{ij}^{ef}]$, and a_{ij}^{ef}, b_{ij}^{ef} are the bounds of x in the corresponding case (see Table 2).

7 Conditional Statistics of D

If the location of Event 1 is fixed and given as (e,p), we can easily derive the conditional analytical formulas.

Conditional moments:

$$E[D^k \mid (e,p)] = \sum_{f\in E} \sum_{(i,j)\in 2^2} \sum_{|\alpha|\leq k} \hat{c}_{ij\alpha}^{ef} p^{\alpha_2} \int_{R_{ij}^{ef}|p} q^{\alpha_1} \phi_{Q|P}(q \mid p)\ dq$$

$$\hat{c}_{ij\alpha}^{ef} = c_{ij\alpha}^{ef}/g_X(e)$$

Region	Term	Value
R_{11}^{ee}	$L(x)$ lb	$\hat{x}/l_e$
	$L(x)$ ub	1
	$q(x,p)$	$p-\hat{x}/l_e$
R_{12}^{ee}	$L(x)$ lb	0
	$L(x)$ ub	$1-\hat{x}/l_e$
	$q(x,p)$	$p+\hat{x}/l_e$
R_{21}^{ee}	$L(x)$ lb	$1-(\hat{x}-d_{uv})/l_e$
	$L(x)$ ub	1
	$q(x,p)$	$p-1+(\hat{x}-d_{uv})/l_e$
R_{22}^{ee}	$L(x)$ lb	0
	$L(x)$ ub	$(\hat{x}-d_{uv})/l_e$
	$q(x,p)$	$p+1-(\hat{x}-d_{uv})/l_e$
R_{11}^{ef}	$L(x)$ lb	$\max\{0,(2\hat{x}-\delta_{11}-\delta_{12}-l_f)/(2l_e)\}$
	$L(x)$ ub	$\min\{p_1,(\hat{x}-\delta_{11})/l_e\}$
	$q(x,p)$	$(\hat{x}-l_ep-\delta_{11})/l_f$
R_{12}^{ef}	$L(x)$ lb	$\max\{0,(2\hat{x}-\delta_{11}-\delta_{12}-l_f)/(2l_e)\}$
	$L(x)$ ub	$\min\{p_2,(\hat{x}-\delta_{12})/l_e\}$
	$q(x,p)$	$(-\hat{x}+l_ep+\delta_{12}+l_f)/l_f$
R_{21}^{ef}	$L(x)$ lb	$\max\{p_1,(-\hat{x}+\delta_{21}+l_e)/l_e\}$
	$L(x)$ ub	$\min\{1,(-2\hat{x}+\delta_{21}+\delta_{22}+2l_e+l_f)/(2l_e)\}$
	$q(x,p)$	$(\hat{x}+l_ep-\delta_{21}-l_e)/l_f$
R_{22}^{ef}	$L(x)$ lb	$\max\{p_2,(-\hat{x}+\delta_{22}+l_e)/l_e\}$
	$L(x)$ ub	$\min\{1,(-2\hat{x}+\delta_{21}+\delta_{22}+2l_e+l_f)/(2l_e)\}$
	$q(x,p)$	$(-\hat{x}-l_ep+\delta_{22}+l_e+l_f)/l_f$

Table 4: Components for the pdf formula

Conditional cdf:

$$\Psi_{D|(e,p)}(x) \quad = \quad \sum_{f\in E} g_Y(f) \sum_{(i,j)\in 2^2} \int_{(R_{ij}^{ef}|p)(x)} \phi_{Q|P}(q \mid p)\ dq$$

These two formulas are derived by fixing values related to e and p, then divided by the term $g_X(e)\phi_P(p)$. The conditional region $(R_{ij}^{ef} \mid p)(x)$ is defined as the region of q when p is fixed. That is, if $R_{ij}^{ef}(x)$ is described by

$$(p_l, p_u, q_l(x,p), q_u(x,p));$$

then, for a fixed value p_0, $(R_{ij}^{ef} \mid p_0)(x) = (q_l(x,p_0), q_u(x,p_0))$. Furthermore, $R_{ij}^{ef} \mid p$ is defined as $(R_{ij}^{ef} \mid p)(b_{ij}^{ef})$, where b_{ij}^{ef} is the upper bound for the distance x in the corresponding region (Table 2). It is easy to check that under our choice of region representation, for any $p \in I$, the family $\{R_{ij}^{ef} \mid p\}_{(i,j)\in 2^2}$ consists of a partition of I. Then, $\phi_{q|p}(q \mid p)$ in the integrand is the conditional distribution of Q given P. In the case where P and Q are independent, it is simply equal to $\phi_q(q)$.

Conditional pdf:

From the conditional cdf, we have

$$\begin{aligned}\psi_{D|(e,p)} &= \frac{d}{dx}\Psi_{D|(e,p)} \\ &= \sum_{f\in E} g_Y(f) \sum_{(i,j)\in\mathbb{2}^2} \frac{d}{dx}\int_{(R_{ij}^{ef}|p)(x)} \phi_{q|p}(q \mid p)\ dq \\ &= \sum_{f\in E} g_Y(f) \sum_{(i,j)\in\mathbb{2}^2} \frac{d}{dx}\int_{q_l(p)}^{q_u(x,p)} \phi_{q|p}(q \mid p)\ dq\end{aligned} \tag{17}$$

In the derivative part, we can follow the same procedure as in the proof of Theorem 3. First, we change the variable $y(x) = q_u(x,p)$ and let $g(y) = \int_{q_l(p)}^{y} \phi_{Q|P}(q \mid p)\ dq$. Then, applying the chain rule and the fundamental theorem of calculus, we get

$$\frac{dg}{dy} = \phi_{Q|P}(q_u(x,p) \mid p), \qquad \frac{dy}{dx} = \begin{cases}\partial_x q(x,p) = \eta_K(x)/d_f, & x \in M(p) \iff p \in L(x) \\ 0 \text{ almost everywhere}, & x \notin M(p) \iff p \notin L(x).\end{cases}$$

Again, we can use the gate function to restrict the nonzero domain of p. Then, plugging this back into (17), we get

$$\psi_{D|(e,p)}(x) = \sum_{f\in E} g_Y(f)/l_f \sum_{(i,j)\in\mathbb{2}^2} \eta_{L_{ij}^{ef}(x)}(p) \cdot \eta_{K_{ij}^{ef}}(x) \cdot \phi_{Q|P}(q(x,p) \mid p).$$

8 Pdf when $\phi_{P,Q}$ is Uniformly Distributed

In this section, we prove a interesting property of the pdf of D when the relative positions of the events are uniformly distributed. We show that in such case, the pdf of D is piecewise linear.

Theorem 5. *If the $\phi_{P,Q}$ is uniformly distributed in its domain I^2, then the pdf of D is piecewise linear.*

Proof. The pdf from (16) can be written as,

$$\psi_D(x) = \sum_{(e,f)\in E^2} \sum_{(i,j)\in\mathbb{2}^2} g_{X,Y}(e,f)/l_f \cdot \eta_{K_{ij}^{ef}}(x) \cdot \int_{L_{ij}^{ef}(x)} \phi_{P,Q}(p, q_{ij}^{ef}(x,p))\ dp.$$

If $\phi_{P,Q}$ is uniform, then the composed function is

$$\phi_{P,Q}(p, q_{ij}^{ef}(x,p)) = \begin{cases}1, & \text{when } p \in I \text{ and } q_{ij}^{ef}(x,p) \in I \\ 0, & \text{otherwise.}\end{cases}$$

Since all $q_{ij}^{ef}(x,p)$ are linear functions both in x and p, we can transform the requirement $q_{ij}^{ef}(x,p) \in I$ into the domain of p as $[h_0(x), h_1(x)]$, where both h_0 and h_1 are linear functions in x, giving us

$$\phi_{P,Q}(p, q_{ij}^{ef}(x,p)) = \begin{cases}1, & \text{when } p \in [\max\{0, h_0(x)\}, \min\{1, h_1(x)\}] \\ 0, & \text{otherwise.}\end{cases}$$

Therefore, integrating this function over the region $L_{ij}^{ef}(x)$, whose bounds are both linear functions in x, gives us a piecewise linear function, say $h_2(x)$. Then, for each given $(e, f) \in E^2$ and $(i, j) \in \mathbb{2}^2$, the term inside the double summation is

$$g_{X,Y}(e, f)/l_f \cdot \eta_{K_{ij}^{ef}}(x) \cdot h_2(x)$$

where $g_{X,Y}(e, f)/l_f$ is a scalar, $\eta_{K_{ij}^{ef}}(x)$ is a gate function that preserves the piecewise linearity. As a consequence, $\psi_D(x)$ is also piecewise linear since finite summations preserve piecewise linearity. □

9 Implementation Details

In the previous sections we derived analytical expressions for the arbitrary order moments, cdf, pdf, and their conditional counterparts. In this section, we will discuss the complexity of evaluating these formulas for a given input and provide some details regarding our implementation.

9.1 Complexity of Computation

The first thing to notice is that all the calculations require as an input the distances between every pair of vertices $u, v \in V$. This can be done, for example, by either applying the Floyd-Warshall algorithm in $O(n^3)$ time or by applying n consecutive runs of Dijkstra's algorithm, rooted at every vertex for a total of $O(nm + n^2 \log n)$ time (assuming that Fibonacci heaps are used), depending on the graph's density.

From all the above expressions, it is easy to see that the complexity depends on the size of the graph, as well as the complexity of the integration. We use $\mathcal{I}_i$ to denote the complexity of integrating a function over a i-dimensional compact and path-connected manifolds embedded in $\mathbb{R}^i$. Then, all the complexity information is given in Table 5. Here, we make the specific distinction about the complexity of the integration component because it can be done either symbolically or numerically, depending on the expected results as well as the precision.

Formula	Complexity
$E[D^k]$	$O(m^2\mathcal{I}_2)$
Ψ_D	$O(m^2\mathcal{I}_2)$
ψ_D	$O(m^2\mathcal{I}_1)$
$E[D^k \mid F = (e, p)]$	$O(m\mathcal{I}_1)$
$\Psi_{D\|F=(e,p)}(x)$	$O(m\mathcal{I}_1)$
$\psi_{D\|F=(e,p)}(x)$	$O(m)$

Table 5: Formula evaluation complexity

If performing the integration numerically, the corresponding complexity is mainly determined by the dimension of the domain of the integrand and the required error bound (30). If the integration regarding input function $\phi_{P,Q}$ has a simple closed-form formula, then evaluation of the integrated function can usually be done in constant time, so the corresponding complexities in Table 5 are either linear or quadratic in m. Moreover, the process of calculating these formulas can be easily processed in parallel, since all of them are simply summations of individual cases.

9.2 The `RanDist` Library

All the statistics and plots provided in the following sections were generated using a callable library called `RanDist` that we implemented in Python 3.7.2. With a connected simple graph G along with the joint pmf $g_{X,Y}$ and joint pdf $\phi_{P,Q}$ as the input, this library provides a direct calculation of the formulas we derived in this paper, giving the option of calculating those either numerically (via the open-source `SciPy` Python library) or symbolically (via the open-source `SymPy` Python library). The numerical implementation is rather fast for evaluating and plotting the resulting formulas, especially for large networks. The symbolic calculation is in general slower, but offers extra functionalities, such as printing and saving the analytical expressions for all the formulas.

This library encapsulates all the proposed tools in a user-friendly interface that can be readily incorporated within more complex implementations. All the details about the library, including installation guides and user manuals can be found in the library's homepage (42).

10 Applications

In this section, we discuss several applications of the proposed methodology. We focus the discussion on models that optimize some tactical decisions for emergency response systems, but note that these examples can be easily extended for modeling similar applications in other contexts as well.

Designing Districting Policies for Emergency Response Systems

One of the most important goals for many emergency response systems is to ensure that the time required by the service units to reach the location of the emergency calls is below a certain threshold with a high probability (5; 16; 23). In the presence of multiple service units, it is a common practice to partition the service networks into several smaller subregions (i.e., districts) so that each unit attends exclusively the services that come from its assigned district, without the need of traversing long distances (24; 32).

Since our models have no assumption on the triangle inequality regarding the distances between the events, the traveling time, as well as other distance-based metrics can be considered as the edge lengths of the infrastructure network. Under this settings, we can apply the proposed formulas to determine an optimal partition (assuming such a partition exists) that guarantees that a desired service confidence level is achieved in each of the districts.

In what follows, we provide some of the building blocks of a valid formulation, but note that several alternatives can be constructed using the proposed tools. Let K be the set of expected districts comprising the partition, and r be the expected confidence level that the distance between any two potential services coming from the same district is shorter than a given value d. We can define the set of binary decision variables $\{z_{ke}\}_{k\in K, e\in E}$ that indicates whether or not the edge e belongs to the partition k. Then, the following sets of constraints define a feasible solution space that satisfies these requirements.

$$\sum_{k \in K} z_{ke} = 1 \quad \forall e \in E, \tag{18}$$

$$\sum_{(e,f) \in E^2} g_{X,Y}(e,f) \Psi_{D|X=e,Y=f}(d) z_{ke} z_{kf} \geq r \sum_{(e,f) \in E^2} g_{X,Y}(e,f) z_{ke} z_{kf} \quad \forall k \in K, \tag{19}$$

$$\mathbf{z} \in \mathcal{P}, \tag{20}$$

where $\Psi_{D|X=e,Y=f}(d)$ is the conditional cdf evaluated at $x = d$, which is equal to the probability $Prob\{D \leq d \mid X = e, Y = f\}$. Constraint set (18) ensures each edge belongs to exactly one district, constraint set (19) forces the conditional probability of $D \leq d$ to be greater than the confidence level requirement r, given that both the random service requests come from the same district, and constraints (20) enforce the z variables to satisfy any further requirements regarding the districts configuration (e.g., further limitations on the size and connectivity of the districts). We note here that the bi-linear terms in (19) can be easily linearized with traditional methods.

There are several variations that can be constructed based on the given model. We can minimize the number of district by introducing a set of additional indicator variables $\{y_k\}_{k \in K}$ that describe whether or not the district k is non-empty. Furthermore, we can also maximize the minimum confidence level across all districts by letting r be a variable. We can also introduce additional constraints regarding the expectation and variance of these distances using variations of the constraints in (19), given the moment generating functions of Section 4.

All these models provide optimal partitions of the network, ensuring that a service quality with controllable confidence level is satisfied. Notice that our framework helps determining the coefficients of the constraints in (19). Given the generality of our developments, these coefficients accurately describe such conditions provided that a good characterization of the distributions that model the event locations is given.

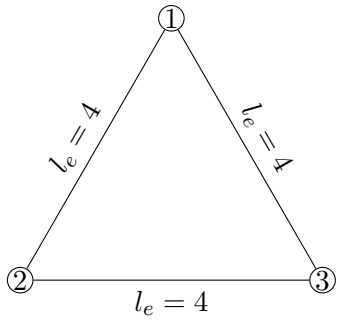
11 Computational Results

In this section, we apply our framework to some special types of networks and discuss some properties of the given statistics. The computational experiments reported in this section were conducted on a Dell OptiPlex 7060 desktop computer equipped with a 8-core Intel i7-6700 3.4GHz cpu, 32 GB of RAM, and running Linux x86-64, Arch 5.1. All the algorithms were implemented in Python 3. All the integrations needed in this section were performed numerically with an error bound $\epsilon = 10^{-3}$.

11.1 Triangle Graph

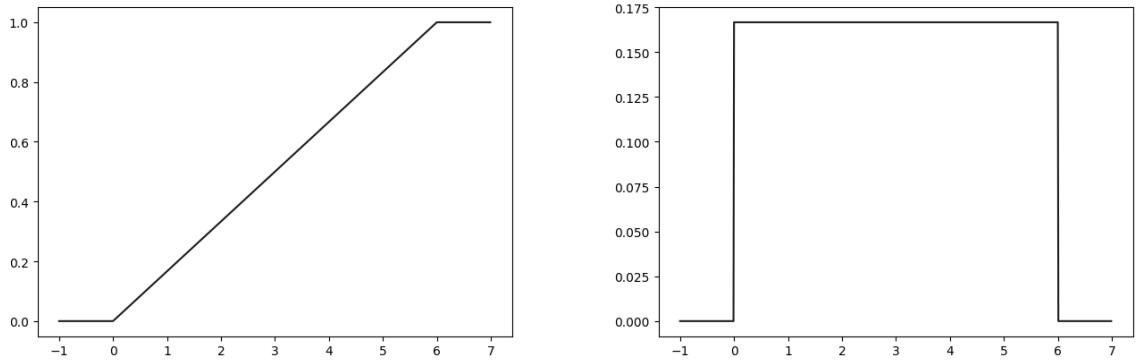
We begin our study with a very simple example. We ran our formulas on the triangle graph given in Figure 9, assuming that the location of both events is uniformly distributed on any point of the three edges of the graph. The information of the moments regarding the distance between such events is presented the table of Figure 9 and the corresponding cdf and pdf are depicted in Figure 10.

Notice that the pdf of D is essentially a uniform distribution from 0 to 6. This is in fact expected because this instance is equivalent to the case where the location of both events is uniformly distributed over a circle with a circumference of 12. The events can be separated by a maximum



Statistics	Value
$E[D]$	3
$E[D^2]$	12
$E[D^3]$	54
$Var(D)$	3

Figure 9: Triangle network and its moments



(a) Cdf

(b) Pdf

Figure 10: Pdf and cdf of D

distance of 6 (when the events take place at diametrically-opposed points) and since the locations are uniformly distributed, the distances are also uniform.

Now, consider the case in which the distribution of the relative locations ϕ_P and ϕ_Q for both events is beta with parameters $\alpha = \beta = 2$ (see Figure 11).

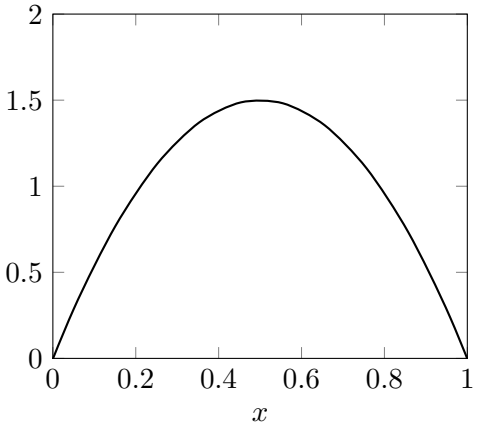
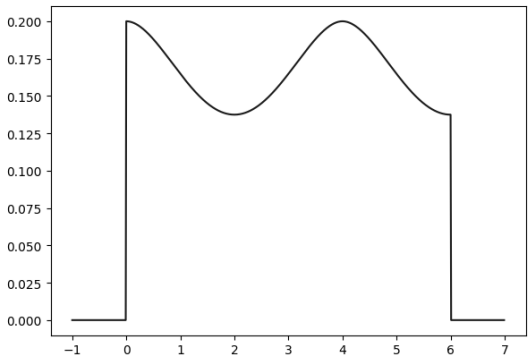


Figure 11: Pdf of Beta distribution with $\alpha = \beta = 2$

Then, assuming that ϕ_P and ϕ_Q are independent random variables, the joint probability dis-

tribution is $\phi_{P,Q}(p,q) = 36p(1-p)q(1-q)$. The corresponding moments as well as the pdf are presented in Figure 12. It is interesting to see how changing the distribution from uniform to beta has a major impact on the shape of the resulting pdf. We notice that in this pdf, any distance $d \in [0,2]$ has the exactly same density as $d+4$. This is because there is a bijection between any pair of points (x,y) separated by a distance d and the pair of points (x,y') separated by a distance $d+4$, since for the fixed location x, the locations y and y' have the exactly same density.

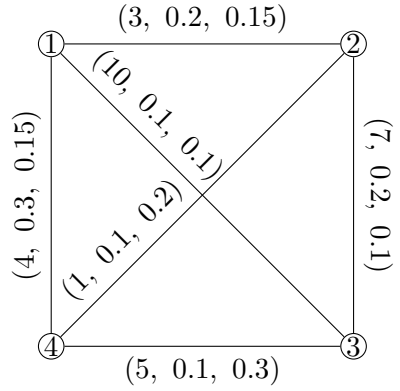


Statistics	Value
$E[D]$	2.975
$E[D^2]$	11.852
$E[D^3]$	52.732
$Var(D)$	3.002

Figure 12: Pdf and moments of D

11.2 Clique of Size Four

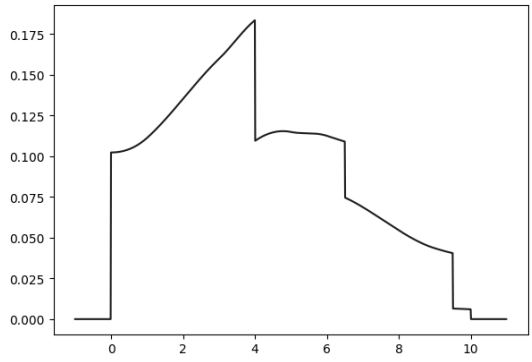
We also ran our formulas for a clique of size four. The given graph is depicted in Figure 13, where the tuple on top of each edge e contains the information of l_e, $g_X(e)$ and $g_Y(e)$, in such an order. We assume that the relative location ϕ_P of Event 1 is uniform, and assume that ϕ_Q is beta with parameters $\alpha = \beta = 2$. Some statistics regarding the distance D for this graph can also be found on the table provided in Figure 11.



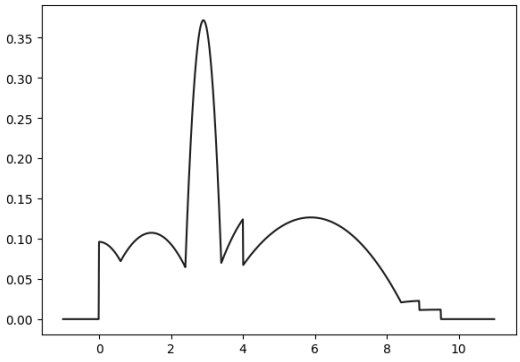
Statistics	Value
$E[D]$	4.035
$E[D^2]$	22.000
$E[D^3]$	139.713
$Var(D)$	5.713

Figure 13: Clique of size four and its moments

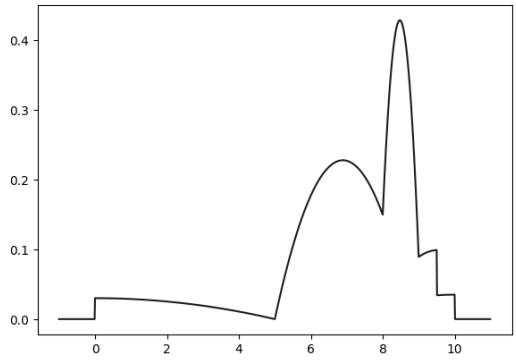
We also present the pdf, and several conditional pdf's, given that Event 1 occur at the locations $(\{1,2\},0.2)$, $(\{1,3\},0.5)$ and $(\{3,4\},0)$, respectively.



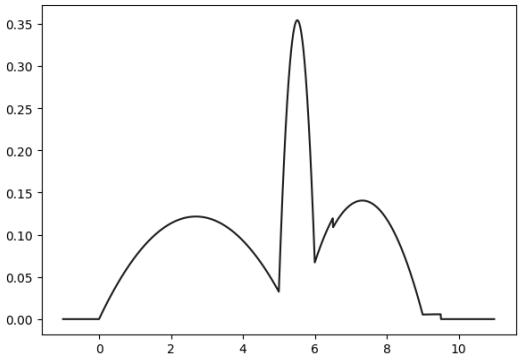
(a) Pdf of D



(b) Conditional pdf for $(\{1,2\}, 0.2)$



(c) Conditional pdf for $(\{1,3\}, 0.5)$



(d) Conditional pdf for $(\{3,4\}, 0)$

Figure 14: The pdf and conditional pdfs for the clique graph

11.3 Grid Network

In this subsection, we study a 10×10 square grid having a unit length on each edge. We are interested in this type of graph due to its similarities to the configurations of many road networks. With this relatively large network, we aim to analyze the evaluation times of various statistics corresponding to different input joint distributions.

Figure 15 depicts this network, where the solid lines represent edges between vertices while dashed lines indicate the fact that there are some ignored vertices and edges in between. We assume that the input pmf $g_{X,Y}$ are uniformly distributed. For a given edge pair (e, f), we consider three different types of input distributions: in Case 1, the events are independent and uniformly distributed; in Case 2, the events are independent, where Event 1 is uniformly distributed and Event 2 is distributed as beta with $\alpha = \beta = 2$ (see Figure 11); in Case 3, the events are jointly distributed with a mixture Dirichlet distribution (Figure 16) having the following pdf

$$\phi_{P,Q}(p,q) = \begin{cases} 180pq(p+q-1)^2, & p>0,\ q>0,\ p+q<1 \\ 180(p-1)(q-1)(p+q-1)^2, & p<1,\ q<1,\ p+q>1 \\ 0, & \text{otherwise.} \end{cases}$$

Over the region $p>0,\ q>0,\ p+q<1$, this pdf defines a Dirichlet distribution with parameters

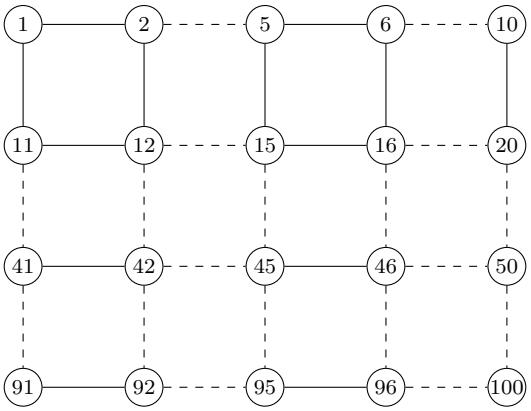


Figure 15: 10×10 grid graph

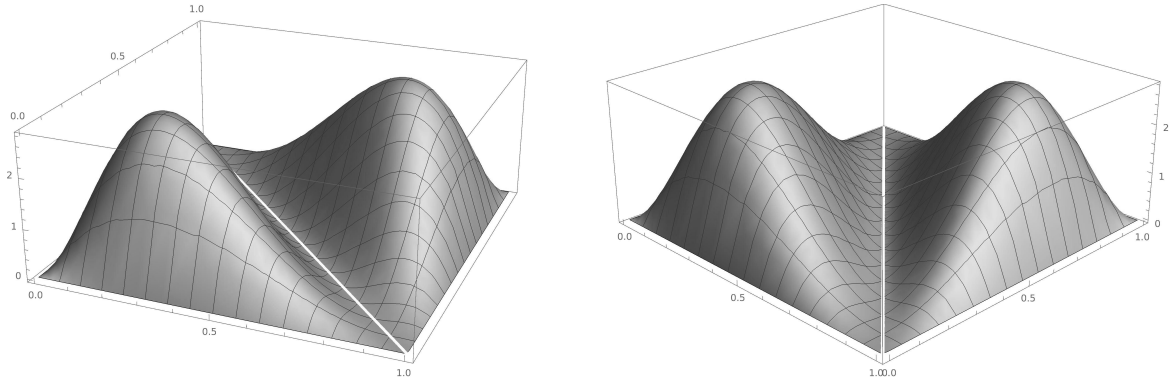


Figure 16: The mixture Dirichlet distribution from two angles

$\alpha = (2, 2, 3)$ weighted by 0.5, and over the other half region, it defines a same weighted Dirichlet distribution reflected over the line $p + q = 1$. We use this joint distribution for its non-trivial properties such as dependence, multimodality, and the fact that it has an uncountable set of non-differentiable points. To calculate the conditional distributions, we fixed the location of Event 1 to be $\mathrm{E}_1 = (\{5, 6\}, 0.5)$.

The pdfs and conditional pdfs associated with each of the three cases are presented in Figures 17, 18, and 19, respectively. As before, the overall shapes of the three cases are similar since the input pmf $g_{X,Y}$ is the same; but, they differ in the local patterns since these are dictated by the input pdf $\phi_{P,Q}(p, q)$. In Case 3, there are more zero points in the conditional pdf than in the other two cases. This is simply because the mixture Dirichlet distribution has non-differentiable points while the other two input joint distributions do not.

Various statistics and their corresponding evaluation times are presented in Table 6. First, notice that the evaluation times reflect the complexities presented in Table 5 of Section 9. All the conditional statistics are evaluated much faster than their unconditional counterparts, which is expected as evaluating the former takes time linear in the number of edges m while the latter takes time quadratic in m. Second, after comparing the evaluation times between all three cases, a general pattern that emerges is that Case 1 has in general the shortest evaluation time, Case 3 has the longest. This pattern can be attributed to the additional calculations required to evaluate the more complex integrals of Cases 2 and 3. As discussed in Section 9.1, the complexity of the numerical integration is determined by the dimension of the domain, and the error bound. Case 3

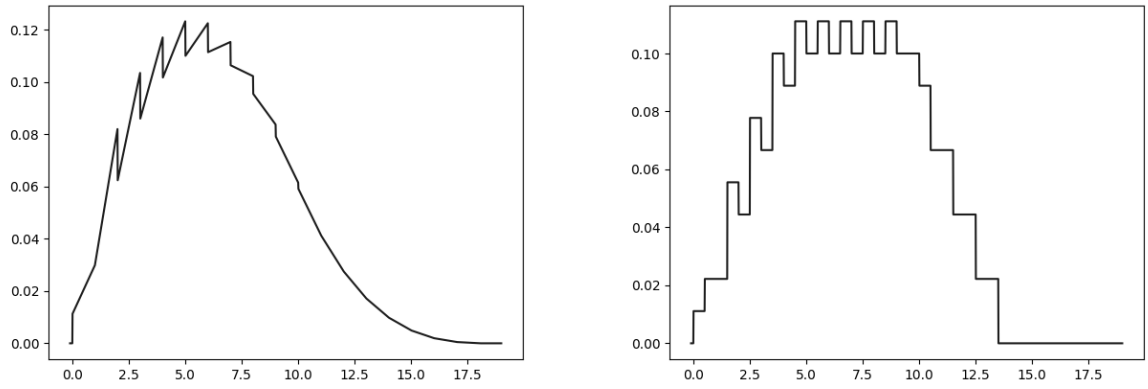


Figure 17: Case 1: Pdf and conditional pdf given event 1 at $(\{5,6\}, 0.5)$

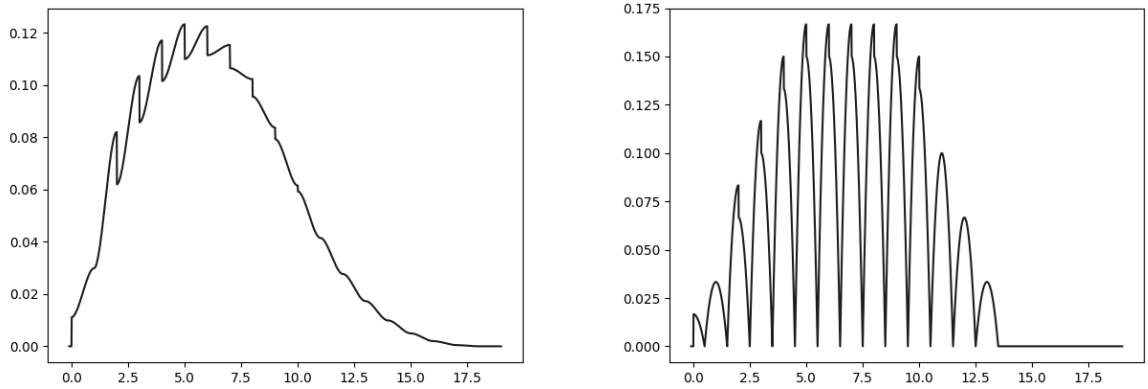


Figure 18: Case 2: Pdf and conditional pdf given event 1 at $(\{5,6\}, 0.5)$

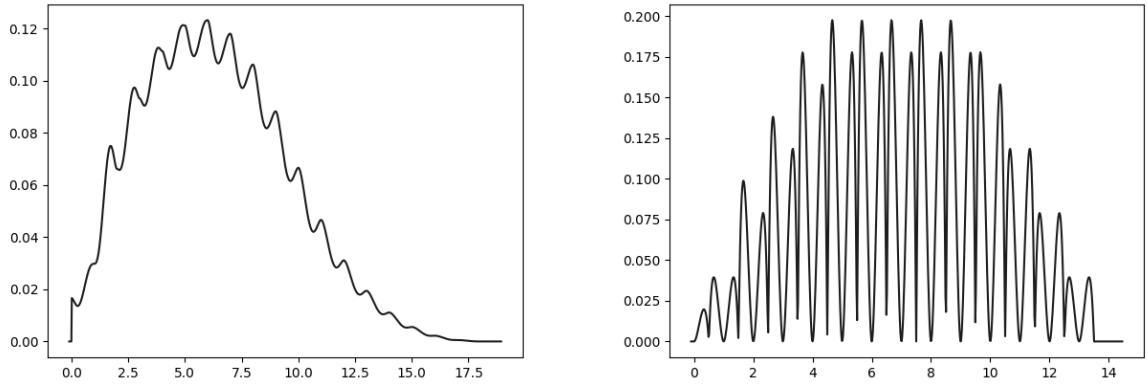


Figure 19: Case 3: Pdf and conditional pdf given event 1 at $(\{5,6\}, 0.5)$

has the most intricate distribution among all three cases; thus, to reach the required error bound, the numerical integration procedure needs to sample more points, which causes an increment of the evaluation time. Interestingly, from Table 6, we observe that this increment is quite small, especially for the conditional statistics.

Statistics	Case 1		Case 2		Case 3	
	Value	Time (sec)	Value	Time (sec)	Value	Time (sec)
$E[D]$	6.333	157.399	6.335	167.013	6.328	229.026
$E[D^2]$	50.067	184.148	50.049	195.744	49.970	248.895
$E[D^3]$	455.493	258.665	454.993	273.606	454.076	337.714
$\Psi_D(3)$	0.159	33.750	0.159	37.359	0.159	46.033
$\Psi_D(6)$	0.489	31.934	0.489	36.239	0.489	46.901
$\Psi_D(9)$	0.796	30.287	0.796	33.532	0.796	48.619
$\psi_D(3)$	0.103	10.864	0.103	11.249	0.093	12.406
$\psi_D(6)$	0.122	10.488	0.122	10.943	0.123	11.886
$\psi_D(9)$	0.084	10.123	0.084	10.637	0.088	11.402
$E[D \mid \mathrm{E}_1]$	6.900	0.229	6.903	0.228	6.898	0.258
$E[D^2 \mid \mathrm{E}_1]$	56.900	0.422	56.904	0.420	56.887	0.487
$E[D^3 \mid \mathrm{E}_1]$	520.538	0.686	520.247	0.678	520.563	0.789
$\Psi_{D\mid\mathrm{E}_1}(3)$	0.117	0.014	0.117	0.013	0.117	0.015
$\Psi_{D\mid\mathrm{E}_1}(6)$	0.406	0.015	0.406	0.014	0.406	0.018
$\Psi_{D\mid\mathrm{E}_1}(9)$	0.722	0.016	0.722	0.015	0.722	0.022
$\psi_{D\mid\mathrm{E}_1}(3)$	0.077	0.032	0.117	0.030	0.000	0.031
$\psi_{D\mid\mathrm{E}_1}(6)$	0.111	0.030	0.167	0.028	0.000	0.029
$\psi_{D\mid\mathrm{E}_1}(9)$	0.111	0.028	0.167	0.027	0.000	0.027

$$\mathrm{E}_1 = (e, p) = (\{5, 6\},\ 0.5)$$

Table 6: Grid network statistics and their evaluation times

11.4 Manhattan Road Network

We conclude this section by testing the proposed framework on a large real-life graph—the road network of the island of Manhattan in New York City (see Figure 20)—to test the scalability of our implementation. We extracted this instance from the data set provided by (21), which comprises the entire road network of the city of New York. In order to guarantee that the input set is connected, we only used the largest connected component of the graph and assumed that all the edges are undirected. The resulting graph has $10,566$ vertices and $16,450$ edges. We selected five key landmarks and three emergency response facilities: the Empire State building, the Charging Bull on Wall Street, Time Square, the Broadway Theater, the Metropolitan Museum of Art, the New York Presbyterian Hospital, the New York Police Department, and the New York City Fire Department. We collected the corresponding conditional statistics, assuming that Event 1 is located at each of these eight places.

Table 7 contains the conditional first and second moments, the conditional cdf and pdf evaluated at the point $x = 2,000$ meters, and their corresponding evaluation times in seconds, assuming that the other event is uniformly distributed. The same information is also presented in Table 8, with the assumption that the location of Event 2 has the marginal distribution of the mixture Dirichlet

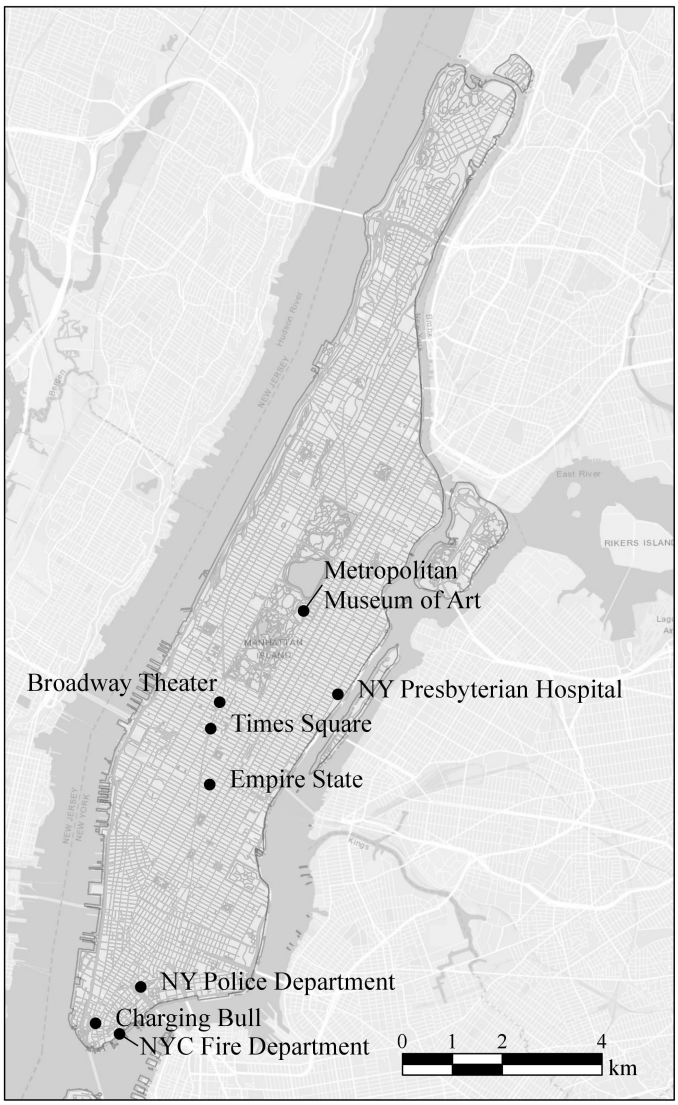
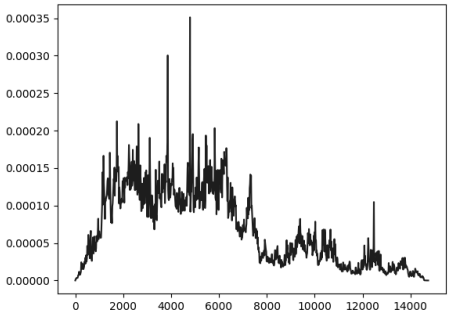


Figure 20: Manhattan road network

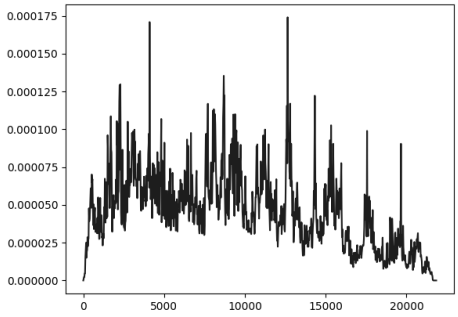
distribution (see Figure 16). The data shows that the evaluations of these statistics in both cases are quite fast. This demonstrates that the implementation of the conditional statistics scales well even for complicated input distributions.

The conditional pdfs of the distance D with respect to each of the locations are presented in Figure 21, with the assumption that the other event is uniformly distributed. From these figures, we see that the Charging Bull, the NYC Fire Department, and the NY Police Department are the locations with the longest average distance among all the landmarks and facilities; this is because they are close to an extreme position in this stripe-like network. Also, notice that the positions of these landmarks and facilities in the network naturally affect the shape of their corresponding distributions. For instance, Time Square and the Broadway Theater are very close to each other (see Figure 20), hence their distributions (Figures 21a and 21h) also share similar shapes.

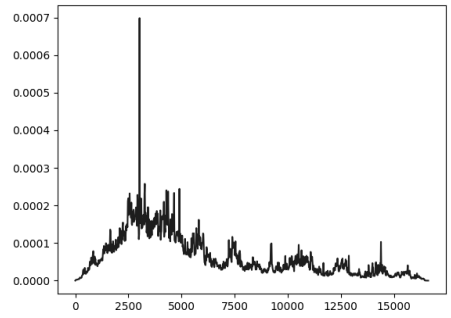
Notice that the distributions that we used in this example are for illustrative purposes. If instead, we use distributions that are estimated from the real data, it is possible to obtain valuable



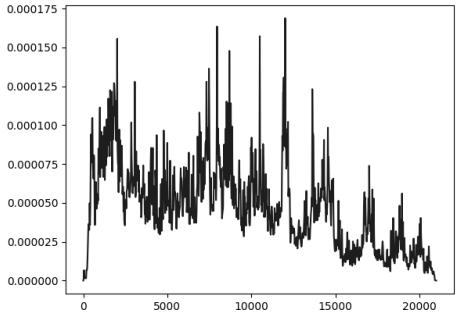
(a) Pdf given event 1 at Broadway Theater



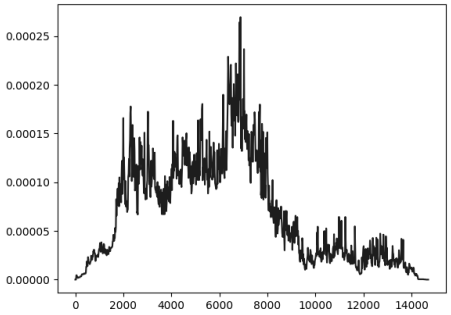
(b) Pdf given event 1 at Charging Bull



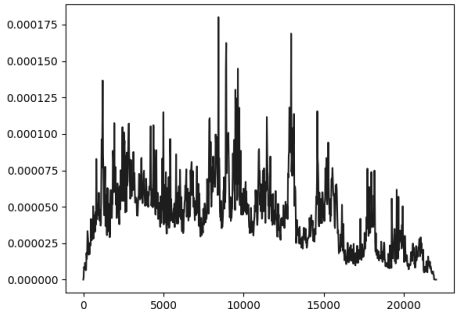
(c) Pdf given event 1 at Empire State



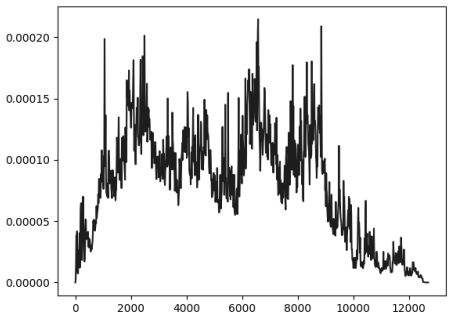
(d) Pdf given event 1 at NY Police Dept



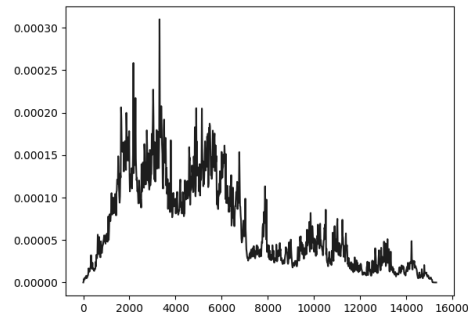
(e) Pdf given event 1 at Presbyterian Hosp



(f) Pdf given event 1 at NYC Fire Dept



(g) Pdf given event 1 at Met Museum



(h) Pdf given event 1 at Time Square

Figure 21: Conditional pdfs for the Manhattan road network

Event 1 Location	$E[D \mid \mathrm{E}_1]$		$E[D^2 \mid \mathrm{E}_1]$		$\Phi_{D\mid\mathrm{E}_1}(2,000)$		$\phi_{D\mid\mathrm{E}_1}(2,000)$	
	Value	Time	Value	Time	Value	Time	Value	Time
Broadway	5,279	23.38	37,711,334	39.81	0.15	1.29	0.00	2.94
Charging Bull	9,034	23.25	110,536,159	41.02	0.09	1.28	0.00	2.97
Empire State	5,732	23.42	46,709,993	40.22	0.11	1.28	0.00	2.96
NY Police Dept	8,341	23.48	97,501,396	40.51	0.14	1.29	0.00	2.93
NY Presbyterian Hosp	5,947	23.35	43,637,573	39.73	0.07	1.27	0.00	2.98
NYC Fire Dept	9,184	23.12	113,658,611	41.11	0.09	1.29	0.00	2.96
The Met	5,312	24.73	36,377,132	39.90	0.15	1.29	0.00	2.94
Times Square	5,386	23.32	39,937,980	40.69	0.14	1.29	0.00	2.94

$$\mathrm{E}_1 = (e, p) = (\{5, 6\},\ 0.5)$$

Table 7: Statistics and evaluation times (seconds) with uniform distribution

Event 1 Location	$E[D \mid \mathrm{E}_1]$		$E[D^2 \mid \mathrm{E}_1]$		$\Phi_{D\mid\mathrm{E}_1}(2,000)$		$\phi_{D\mid\mathrm{E}_1}(2,000)$	
	Value	Time	Value	Time	Value	Time	Value	Time
Broadway	5,280	27.38	37,737,370	47.42	0.15	1.54	0.00	3.15
Charging Bull	9,040	27.22	110,663,929	47.66	0.09	1.41	0.00	3.14
Empire State	5,730	26.49	46,661,779	46.60	0.11	1.43	0.00	3.14
NY Police Dept	8,344	26.64	97,562,346	46.30	0.14	1.45	0.00	3.07
NY Presbyterian Hosp	5,945	26.49	43,605,559	46.59	0.07	1.35	0.00	3.14
NYC Fire Dept	9,177	26.27	113,526,184	46.54	0.09	1.37	0.00	3.11
The Met	5,312	28.02	36,366,504	49.05	0.15	1.51	0.00	3.08
Times Square	5,388	25.93	39,976,357	44.91	0.14	1.47	0.00	3.16

$$\mathrm{E}_1 = (e, p) = (\{5, 6\},\ 0.5)$$

Table 8: Statistics and evaluation times (seconds) with marginal mixture Dirichlet distribution

statistics related to the coverage and efficiency of these emergency response facilities. Furthermore, in reality, there are multiple facilities of this kind in Manhattan, so calculating the distance statistics over the entire road network is not very useful, since the emergencies are typically served by the nearest facility. To obtain a more realistic analysis, we can first partition the network into connected sub-networks according to the distance, so that each sub-network has only one facility. Then, the proposed framework can be applied to each of the sub-networks.

12 Conclusion

In this paper, we aim to develop a collection of formulas that can be used to characterize several statistical properties of the distance between events that take place on random locations along the edges of a given network. We assume that the probability distribution of the location of such events is known in advance. We derived analytical expressions for the arbitrary moments of such distance, its probability density function, its cumulative distribution function, as well as their conditional counterparts, given that the position of one of the events is known in advance.

The proposed framework was implemented as a callable library for the Python language and was used to calculate and visualize some of these statistics for several networks of different sizes and topologies. We included a brief study of the road network of Manhattan island in New York City (USA), composed of more than 10 thousand vertices and 15 thousand edges. We analyzed some interesting properties of the resulting expressions and discussed several applications for developing models aimed at optimally designing emergency response systems on infrastructure networks. Our implementation is publicly available for others to improve upon, extend, or compare with it.

Acknowledgements

We thank María Nariné Torres Cajiao from *Centro de Investigaciones en Ingeniera Ambiental* at *Universidad de Los Andes* for her assistance with the generation of the Manhattan road network data set and Demetrios Papazaharias for his comments and suggestions that helped us improve this paper.

References

[1] ACJIC, 2019. Alabama Criminal Justice Information Center Safety Portal (ADVANCE). `https://safety.aladata.com` (Accessed Feb 14, 2019).

[2] N. Agatz, A. Erera, M. Savelsbergh, and X. Wang, *Optimization for dynamic ride-sharing: A review*, Eur. J. Oper. Res. **223** (2012), 295–303.

[3] S. Alpern, A. Morton, and K. Papadaki, *Patrolling games*, Oper. Res. **59** (2011), 1246–1257.

[4] R. Aringhieri, M.E. Bruni, S. Khodaparasti, and J.T. van Essen, *Emergency medical services and beyond: Addressing new challenges through a wide literature review*, Comput. Oper. Res. **78** (2017), 349–368.

[5] M.A. Badri, *Combining the analytic hierarchy process and goal programming for global facility location-allocation problem*, Int. J. Production Econ. **62** (1999), 237–248.

[6] R. Batta and U.S. Palekar, *Mixed planar/network facility location problems*, Comput. Oper. Res. **15** (1988), 61–67.

[7] C. Bettstetter and J. Eberspacher, *Hop distances in homogeneous ad hoc networks*, The 57th IEEE Semiannual Vehicular Technology Conference, 2003. VTC 2003-Spring., Vol. 4, 2003, pp. 2286–2290 vol.4.

[8] S.P. Borgatti, *Identifying sets of key players in a social network*, Comput. Math. Organization Theory **12** (2006), 21–34.

[9] J. Brimberg, H.T. Kakhki, and G.O. Wesolowsky, *Locating a single facility in the plane in the presence of a bounded region and different norms*, J. Oper. Res. Soc. Jpn. **48** (2005), 135–147.

[10] L. Brotcorne, G. Laporte, and F. Semet, *Ambulance location and relocation models*, Eur. J. Oper. Res. **147** (2003), 451–463.

[11] V. Bélanger, A. Ruiz, and P. Soriano, *Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles*, Eur. J. Oper. Res. **272** (2019), 1–23.

[12] S.N. Chiu, D. Stoyan, W.S. Kendall, and J. Mecke, *Stochastic geometry and its applications*, John Wiley & Sons, 2013.

[13] City of Chicago, 2019. Chicago data portal. `https://data.cityofchicago.org/` (Accessed Feb 14, 2019).

[14] Z. Drezner, C.H. Scott, and J. Turner, *Mixed planar and network single-facility location problems*, Networks **68** (2016), 271–282.

[15] B. Elenbogen and J.F. Fink, *Distance distributions for graphs modeling computer networks*, Discr. Appl. Math. **155** (2007), 2612–2624.

[16] R.Z. Farahani and M. Hekmatfar, *Facility location: Concepts, models, algorithms and case studies*, Springer, 2009.

[17] W. Goddard and O.R. Oellermann, *Structural Analysis of Complex Networks* chapter Distance in Graphs, pp. 49–72, Birkhäuser Boston 2011.

[18] M. Haenggi, *On distances in uniformly random networks*, IEEE Trans. Informat. Theory **51** (2005), 3584–3586.

[19] M. Haenggi and D. Puccinelli, *Routing in ad hoc networks: A case for long hops*, IEEE Commun. Magazine **43** (2005), 93–101.

[20] H. Hosoya, *On some counting polynomials in chemistry*, Discr. Appl. Math. **19** (1988), 239–257.

[21] A. Karduni, A. Kermanshah, and S. Derrible, *A protocol to convert spatial polyline data to network formats and applications to world urban road networks*, Sci. Data **3** (2016), 160046.

[22] Z. Khalid and S. Durrani, *Distance distributions in regular polygons*, IEEE Trans. Vehicular Technology **62** (2013), 2363–2368.

[23] A. Klose and A. Drexl, *Facility location models for distribution system design*, Eur. J. Oper. Res. **162** (2005), 4–29.

[24] M.E. Mayorga, D. Bandara, and L.A. McLay, *Districting and dispatching policies for emergency medical service systems to improve patient survival*, IIE Trans. Healthcare Syst. Eng. **3** (2013), 39–56.

[25] L.A. McLay and M.E. Mayorga, *Evaluating the impact of performance goals on dispatching decisions in emergency medical service*, IIE Trans. Healthcare Syst. Eng. **1** (2011), 185–196.

[26] L.E. Miller, *Distribution of link distances in a wireless network*, J. Res. National I. Standards Technology **106** (2001), 401–412.

[27] D. Moltchanov, *Distance distributions in random networks*, Ad Hoc Networks **10** (2012), 1146–1166.

[28] L. Moreira-Matias, J. Gama, M. Ferreira, J. Mendes-Moreira, and L. Damas, *Predicting taxi–passenger demand using streaming data*, IEEE Trans. Intelligent Transp. Syst. **14** (2013), 1393–1402.

[29] M. Nourinejad and M.J. Roorda, *Agent based model for dynamic ridesharing*, Transp. Res. Part C: Emerging Technologies **64** (2016), 117–132.

[30] E. Novak, *Some results on the complexity of numerical integration*, Monte Carlo and Quasi-Monte Carlo Methods, Springer International Publishing, Cham, 2016, pp. 161–183.

[31] (OHSP), 2019. Office of Highway Planning, State of Michigan. MTCF Michigan traffic crash facts. `https://www.michigantrafficcrashfacts.org` (Accessed Feb 14, 2019).

[32] M. Pavone, A. Arsie, E. Frazzoli, and F. Bullo, *Equitable partitioning policies for robotic networks*, Robotics and Automation, 2009. ICRA'09. IEEE International Conference on, 2009, pp. 2356–2361.

[33] R. Pure and S. Durrani, *Computing exact closed-form distance distributions in arbitrarily-shaped polygons with arbitrary reference point*, Mathematica J. **17** (2015).

[34] C. Rose, *Mean internodal distance in regular and random multihop networks*, IEEE Trans. Commun. **40** (1992), 1310–1318.

[35] B.E. Sagan, Y.N. Yeh, and P. Zhang, *The Wiener polynomial of a graph*, Int. J. Quantum Chemistry **60** (1996), 959–969.

[36] J. Segert, Hyperbolic dynamical systems and the noncommutative integration theory of connes, Ph.D. thesis, Princeton University, 1987.

[37] S. Srinivasa and M. Haenggi, *Distance distributions in finite uniformly random networks: Theory and applications*, IEEE Trans. Vehicular Technology **59** (2010), 940–949.

[38] D. Stoyan and H. Stoyan, *Fractals, Random Shapes, and Point Fields: Methods of Geometrical Statistics* Vol. 302, John Wiley & Sons Inc, 1994.

[39] (TLC), 2019. Taxi and Limousine Commision, New York City. trip record data. `http://www.nyc.gov` (Accessed Feb 14, 2019).

[40] H. Toro-Díaz, M.E. Mayorga, S. Chanta, and L.A. McLay, *Joint location and dispatching decisions for emergency medical services*, Comput. Indus. Eng. **64** (2013), 917–928.

[41] D.J. Watts and S.H. Strogatz, *Collective dynamics of 'small-world' networks*, Nature **393** (1998), 440.

[42] N. Wei, 2019. `RanDist` home page. `https://github.com/tidues/RanDist` (Accessed Feb 14, 2019).

[43] H. Wiener, *Structural determination of paraffin boiling points*, J. Am. Chemical Soc. **69** (1947), 17–20.

[44] P. Zhang, I. Nevat, G.W. Peters, G. Xiao, and H. Tan, *Event detection in wireless sensor networks in random spatial sensors deployments*, IEEE Trans. Signal Process. **63** (2015), 6122–6135.