Counting Geodesic Paths in 1D VANETs. IEEE Wireless Communications Letters.

Peer reviewed version
License (if available):
Unspecified

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) will be available online via IEEE. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

## General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms.html

# Counting Geodesic Paths in 1D VANETs 

Georgie Knight ${ }^{1}$, Alexander P. Kartun-Giles ${ }^{1}$, Orestis Georgiou ${ }^{1,2}$, and Carl P. Dettmann ${ }^{1}$<br>${ }^{1}$ School of Mathematics, University of Bristol, University Walk, Bristol, BS8 1TW, UK<br>${ }^{2}$ Toshiba Telecommunications Research Laboratory, 32 Queens Square, Bristol, BS1 4ND, UK


#### Abstract

In the IEEE 802.11p standard addressing vehicular communications, Basic Safety Messages (BSMs) can be bundled together and relayed as to increase the effective communication range of transmitting vehicles. This process forms a vehicular ad hoc network (VANET) for the dissemination of safety information. The number of "shortest multihop paths" (or geodesics) connecting two network nodes is an important statistic which can be used to enhance throughput, validate threat events, protect against collusion attacks, infer location information, and also limit redundant broadcasts thus reducing interference. To this end, we analytically calculate for the first time the mean and variance of the number of geodesics in 1D VANETs.


## I. Introduction

Vehicular ad-hoc networks (VANETs) are formed by vehicles, wirelessly connected in a communication network. VANETs dynamically and rapidly self-organise on roads and highways and can communicate with road side access point infrastructure. Primarily, vehicle to vehicle (V2V) and vehicle to infrastructure (V2I) communications involve safety related issues, such as collision warnings aimed at preventing imminent car accidents through broadcasting and relaying messages, thereby increasing local situation awareness [1]. V2V and V2I communications can also be exploited for applications such as intelligent cruise control or platooning, traffic information and management, as well as internet access and advertising. This paper is concerned with multi-hop communications in 1D VANETs, characterised by roads with no sharp bends and width much smaller than the wireless transmission range [2].

The IEEE 802.11 p standard defines a wireless area network (WLAN) for dedicated short range communication (DSRC) among vehicles. The standard defines protocols for the physical and MAC layers, has a 75 MHz bandwidth allocated at 5.9 GHz , and is the prime candidate currently being deployed in order to get IEEE 802.11p equipped cars on the roads [3]. Under the standard, it is possible to bundle together information on position, speed, direction, brake information, steering wheel angle, threat-events, etc., and append them to a basic safety message (BSM) which is then broadcasted [4], [5]. Vehicles within range can then actuate on this information, edit it, or append to the content message, and re-broadcast, thus locally flooding the network.

Flooding algorithms are commonplace in ad hoc networks, however here the algorithm is also spatially constrained to run along a one-dimensional road network. Such networks are typically modelled as random geometric graphs [6] formed by a 1D Poisson point process (PPP) and a communication


Fig. 1. A one-dimensional unit disk graph $\left(r_{0}=1\right)$ used to model a VANET. The geodesic length between the two extreme nodes is $k=3$.
range $r_{0}$ (see Fig.1) directly related to SNR [7] thus lending themselves to mathematical analysis and engineering. A major challenge in VANETs is the timeliness and latency in which information must arrive to be useful to a fast approaching vehicle. Hop-count statistics find application in a variety of other settings, e.g., in gas pipe sensor networks [8], nanowires [9], and map navigation problems in general. Therefore, hopcount statistics have been extensively studied in 1D [10] and 2D networks [11], [12]. They were first studied by Chandler [12], who looked at the probability that two wireless network nodes can communicate in $k$ hops. Such information can further assist the calculation of network centrality measures [13], or achieve range-free localisation [14].
In this paper we are concerned with the statistical properties of the shortest multihop paths, also referred to as geodesics, between nodes in 1D VANETs. To this end, we calculate for the first time the first few moments of the number of geodesics $\sigma$ between nodes in a 1D VANET, as a function of the Euclidean distance $L$ between them and the vehicle density $\lambda$. Clearly for $L \in\left((k-1) r_{0}, k r_{0}\right)$, the shortest possible path is of length $k$ hops, employing just $k-1$ relay nodes, thus defining a fundamental upper limit on the latency involved with such transmissions. On the other hand, due to the broadcast nature of wireless transmissions, multiple BSMs containing similar information may arrive via different $k$-hop paths almost simultaneously. It is therefore of interest to understand the statistical properties of the number of $k$-hop paths $\sigma_{k}$, as a function of $r_{0}, L$, and $\lambda$. Such statistics can be used to enhance throughput [15], validate threat events, protect against collusion attacks, infer location information, and also limit redundant broadcasts thus reducing interference. Moreover, calculating shortest paths is a computationally challenging task [16] which hop-count statistics can alleviate.

## II. System Model

Consider a source node $\mathcal{S}$ located at the origin, and a destination node $\mathcal{D}$ a distance $L>2 r_{0}$ to the right of $\mathcal{S}$ along the positive real line. Further, consider a 1D PPP of density $\lambda$ vehicles per unit length forming on the real line, with each point (node) representing a vehicle along an infinite stretch of road. Nodes are then connected via communication links


Fig. 2. System Model: All 3-hop paths between source (S) and destination (D) nodes separated by a distance $L \in\left(2 r_{0}, 3 r_{0}\right)$ must involve at least 1 relay node located in each of the shaded "lenses".


Fig. 3. Probability mass function of the shortest paths $\sigma$ for $r_{0}=1, \lambda=20$, and $L=2.5,3.5$, such that geodesics are of length 3 , and 4 respectively.
whenever their Euclidean distance is less than a predefined communication range $r_{0}$ (see Fig. 2), thus forming a 1 D network. The source and destination nodes are unable to communicate directly and must employ multihop communications in order to share information. We assume a separation in the time scales in that the messages are sufficiently fast relative to the nodes so that we assume a static network during a multihop transmission. Furthermore we assume a simplified MAC and PHY with no collision avoidance mechanism and a binary communication range mode i.e. the unit disk connection model. Depending on the density of vehicles $\lambda$, there may be none, one, or several multihop paths connecting $\mathcal{S}$ and $\mathcal{D}$. The length of these paths is the number of hops required for a message to pass between the two vehicles. It follows that the length of the shortest multihop paths is $k=\left\lceil\frac{L}{r_{0}}\right\rceil$. Therefore, paths of length $k$ are geodesic. Running a breadth-first search (BFS) algorithm can find all geodesic paths in linear-time since the underlying graph is neither directed, nor weighted. Let the set of all geodesics be described by $\Sigma\left(r_{0}, L, \lambda\right)$. Then the number of geodesic paths is

$$
\begin{equation*}
\sigma_{k}:=\operatorname{card}\left[\Sigma\left(r_{0}, L, \lambda\right)\right] \tag{1}
\end{equation*}
$$

Monte Carlo simulations of the pmf of $\sigma$ are shown in Fig. 3. We will first demonstrate the difficulties with obtaining the distribution of $\sigma_{k}$ for the case of $k=3$, and then calculate its first few moments for the general case of $k \geq 3$. The cases of $k=1$ and $k=2$ are trivial and therefore omitted.

## III. Distribution of Geodesic Paths

Let $L \in\left(2 r_{0}, 3 r_{0}\right)$ and $k=3$ as in Fig. 2 such that there are two sub-domains $L_{1}$ and $L_{2}$ within which relay nodes must be situated in order for a 3-hop path to exist. We call these sub-domains lenses, since in two dimensions they are formed by the intersection of two equal disks. This is because the first relay node located at a maximum distance of $r_{0}$ can form a 3-hop path by connecting with any node in $L_{2}=\left[L-r_{0}, 2 r_{0}\right]$.


Fig. 4. Schematic showing $N_{1}=3$ nodes in the left lens $L_{1}$, and the corresponding sub-domains $w_{i}$ in the right lens $L_{2}$. Note that $w_{3}$ is within range from all three nodes and therefore a fourth node located in $w_{3}$ will connect to all in $L_{1}$, to form three 3 -hop paths from $\mathcal{S}$ to $\mathcal{D}$. In contrast, $w_{2}$ is in range of nodes 1 and 2 (not 3 ), $w_{1}$ is only in range of node 1 (not 2 or 3 ), and $w_{0}$ is not in range from any of the nodes in $L_{1}$.

By symmetry $L_{1}=\left[L-2 r_{0}, r_{0}\right]$ such that the two lenses are of equal widths $\left|L_{1}\right|=\left|L_{2}\right|=3 r_{0}-L$. The number $N_{1}$ of relay nodes in $L_{1}$ is therefore a Poisson random variable with mean $\Lambda_{3}$, where we have defined $\Lambda_{k}=\lambda\left(k r_{0}-L\right)$. Moreover, for each relay node in $L_{1}$ there corresponds a subset of $L_{2}$ within which a second relay node must be located as to form a 3hop path from $\mathcal{S}$ to $\mathcal{D}$. Labelling the $N_{1}$ relays in descending distances $d_{i}$ from the source (i.e., $L-2 r_{0} \leq d_{N_{1}} \leq d_{N_{1}-1} \leq$ $\ldots \leq d_{1} \leq r_{0}$ ) we can identify subsets $\left[L-r_{0}, d_{i}+r_{0}\right] \subseteq L_{2}$ within which if located a second relay can successfully form a 3-hop path. Defining the sub-domains $w_{i}=\left[d_{i+1}+r_{0}, d_{i}+r_{0}\right]$, for $i=0,1, \ldots N_{1}$ with $d_{0}=2 r_{0}$ and $d_{N_{1}+1}=L-r_{0}$ it can be seen that a relay node in $w_{i}$ connects to $i$ relays in $L_{1}$. We therefore arrive at a simple expression for the number of shortest 3-hop paths

$$
\begin{equation*}
\sigma_{3}=\sum_{i=1}^{N_{1}} i n_{i} \tag{2}
\end{equation*}
$$

where the $n_{i}$ is the number of relays in $w_{i}$ and are thus Poisson random variables with mean $\lambda w_{i}$ (see Fig. 4). The widths $w_{i}$ are also random variables however must satisfy the constraint that $\sum_{i=0}^{N_{1}} w_{i}=3 r_{0}-L$, i.e., the $n_{i}$ are correlated.

The pmf of $\sigma$ can be expressed as follows:

$$
\begin{equation*}
\mathbb{P}\left[\sigma_{3}=x\right]=\mathbb{E}_{N_{1}, \mathbf{w}}\left[\mathbb{P}\left[\sigma_{3}=x \mid N_{1}, \mathbf{w}\right]\right] \tag{3}
\end{equation*}
$$

where $\mathbf{w}=\left\{w_{0}, w_{1}, \ldots w_{N_{1}}\right\}$ and any configuration of widths $\mathbf{w}$ is equally likely. We can attempt to obtain the pmf of $\sigma_{3}$ through the use of probability generating functions (PGFs). Namely, we have that the PGF of the random variable $Z_{i}=i n_{i}$ is given by $G_{Z_{i}}(z)=\mathbb{E}\left[z^{i n_{i}}\right]=G_{n_{i}}\left(z^{i}\right)=e^{\lambda w_{i}\left(z^{i}-1\right)}$ since $n_{i}$ is Poisson distributed with mean $\lambda w_{i}$. It follows that the sum of $N_{1}$ such random variables has a PGF given by

$$
\begin{equation*}
G_{\sigma_{3}}(z)=\prod_{i=1}^{N_{1}} e^{-\lambda w_{i}\left(1-z^{i}\right)} \tag{4}
\end{equation*}
$$

and the corresponding pmf given by

$$
\begin{align*}
& \mathbb{P}\left[\sigma_{3}=x\right]= \\
& \left.\sum_{k=0}^{\infty} \mathbb{P}\left[N_{i}=k\right] \int_{\left[0,3 r_{0}-L\right]^{N_{1}}} \mathbf{1}(\mathbf{w}) \frac{c}{k!} \frac{\mathrm{d}^{k} G_{\sigma_{3}}}{\mathrm{~d} z^{k}}\right|_{z=0} \mathrm{~d} w_{1} \ldots \mathrm{~d} w_{N_{1}} \tag{5}
\end{align*}
$$

where $\mathbb{P}\left[N_{i}=k\right]=\frac{\Lambda_{3}^{k}}{k!} e^{-\Lambda_{3}}$ and $\mathbf{1}(\mathbf{w})$ is the indicator function equal to 1 whenever $\|\mathbf{w}\|_{1}=3 r_{0}-L$ and zero otherwise such that $\int \mathbf{1}(\mathbf{w}) \mathrm{d} w_{1} \ldots \mathrm{~d} w_{N_{1}}=1 / c$, and $c>0$ is some normalisation constant. Geometrically, the indicator function defines a simplex polytope with $N_{1}+1$ vertices at $\left\{\mathbf{v}_{0}, \ldots \mathbf{v}_{N_{1}}\right\}^{T}=\left(3 r_{0}-L\right) \mathbf{I}_{N_{1}+1}$. The integral is therefore
over the surface of the $N_{1}$-simplex. Recall that the $N_{1}$-simplex is a triangle, a tetrahedron, a 5-cell, for $N_{1}=2,3$, and 4 respectively, and therefore is an evermore complex polytope embedded in the positive hyperoctant of $\mathbb{R}^{N_{1}+1}$ for which the integration of (5) becomes intractable. For this reason we next restrict our study to the mean and variance of $\sigma_{3}$.

## IV. 3-Hop Mean and Variance

We now describe a method which allows us to analytically derive the moments of $\sigma_{3}$. This involves dividing up the lenses into many small parts and making a simplifying approximation about the interactions. This allows us to treat the problem as one involving many independent random variables rather than trying to account for dependence. The final step is to take the limit of the number of divisions of the lenses to infinity, in which our approximation becomes exact.

We firstly split the lenses $L_{i}$ into a large number $l \gg 1$ of equally sized, disjoint domains $L_{i j}$ where $\left|L_{i j}\right|=\left(3 r_{0}-L\right) / l$ and $L_{i}=\bigcup_{j=1}^{l} L_{i j}$. The number of relay nodes in each $L_{i j}$ is then a Poisson distributed random variable $Y_{i j}$ with mean $\Lambda_{3} / l$. For finite $l$ we make the approximation that all relay nodes in $L_{11}$ connect with all those in $L_{21}$, all those in $L_{12}$ connect with all in $L_{21}$ and $L_{22}$ etc. The number of shortest 3-hop paths is then given by

$$
\begin{equation*}
\sigma_{3}=\lim _{l \rightarrow \infty} \sum_{q=1}^{l} \sum_{r=1}^{q} Y_{1 q} Y_{2 r} \tag{6}
\end{equation*}
$$

Using the independence of the $Y_{i j}$ we calculate the mean

$$
\begin{align*}
\mathbb{E}\left[\sigma_{3}\right] & =\lim _{l \rightarrow \infty} \sum_{q=1}^{l} \sum_{r=1}^{q} \mathbb{E}\left[Y_{1 q}\right] \mathbb{E}\left[Y_{2 r}\right]  \tag{7}\\
& =\lim _{l \rightarrow \infty}\left(\frac{\Lambda_{3}^{2}}{l^{2}}\right)\left(\frac{l^{2}+l}{2}\right)=\frac{\Lambda_{3}^{2}}{2}
\end{align*}
$$

To extract the variance we first define the random variable

$$
\begin{equation*}
T_{q}=Y_{1 q} \sum_{r=1}^{q} Y_{2 r} \tag{8}
\end{equation*}
$$

Given that the variance of a sum is equal to the sum of the variances plus the covariances we have

$$
\begin{equation*}
\operatorname{Var}\left(\sigma_{3}\right)=\lim _{l \rightarrow \infty} \sum_{q=1}^{l} \operatorname{Var}\left(T_{q}\right)+2 \sum_{t=2}^{l} \sum_{s=1}^{t-1} \operatorname{Cov}\left(T_{s}, T_{t}\right) \tag{9}
\end{equation*}
$$

We first evaluate the variance of $T_{q}$. We use the independence of $Y_{1 q}$ and $Y_{2 r}$ and note that $\sum_{r=1}^{q} Y_{2, r}$ is a Poisson random variable with mean $q \lambda\left(3 r_{0}-L\right) / l$. In addition we use the mean of the square of a Poisson random variable with mean $x$ is equal to $x^{2}+x$ and derive

$$
\begin{equation*}
\operatorname{Var}\left(T_{q}\right)=\frac{q^{2} \Lambda_{3}^{3}}{l^{3}}+q\left(\frac{\Lambda_{3}^{3}}{l^{3}}+\frac{\Lambda_{3}^{2}}{l^{2}}\right) \tag{10}
\end{equation*}
$$

Using (10) we evaluate the limit of the first sum in (9) as

$$
\begin{equation*}
\lim _{l \rightarrow \infty} \sum_{q=1}^{l} \operatorname{Var}\left(T_{q}\right)=\frac{\Lambda_{3}^{3}}{3}+\frac{\Lambda_{3}^{2}}{2} \tag{11}
\end{equation*}
$$

For the covariance terms in Eq.(9) we let $s<t \leq l$ and use the relation $\operatorname{Cov}\left(T_{s}, T_{t}\right)=\mathbb{E}\left[T_{s} T_{t}\right]-\mathbb{E}\left[T_{s}\right] \mathbb{E}\left[T_{t}\right]$. The expectation


Fig. 5. Mean (a) and variance (b) of the number of shortest shortest 3hop paths $\sigma$ as a function of $r_{0}$ calculated numerically from ensembles of $10^{6}$ realisations for a range of values of $\lambda$ (symbols). Also illustrated is the analytical results of Eq.(7) and Eq.(17) (lines).
of $T_{s}$ is given by

$$
\begin{equation*}
\mathbb{E}\left[T_{s}\right]=\frac{s \Lambda_{3}^{2}}{l^{2}} \tag{12}
\end{equation*}
$$

For the expectation of the product we have via (8)

$$
\begin{equation*}
\mathbb{E}\left[T_{s} T_{t}\right]=\mathbb{E}\left[\left(Y_{1 s} \sum_{r=1}^{s} Y_{2 r}\right) Y_{1 t}\left(\sum_{r=1}^{s} Y_{2 r}+\sum_{r=s+1}^{t} Y_{2 r}\right)\right] \tag{13}
\end{equation*}
$$

where by splitting the sum in $T_{t}$ we can factorise using the mutual independence of the terms as
$\mathbb{E}\left[T_{s} T_{t}\right]=\mathbb{E}\left[Y_{1 s}\right] \mathbb{E}\left[Y_{1 t}\right]\left(\mathbb{E}\left[\left(\sum_{r=1}^{s} Y_{2 r}\right)^{2}\right]+\mathbb{E}\left[\sum_{r=1}^{s} Y_{2 r}\right] \mathbb{E}\left[\sum_{r=s+1}^{t} Y_{2 r}\right]\right)$
evaluating the individual expectations and combining we have

$$
\begin{equation*}
\mathbb{E}\left[T_{s} T_{t}\right]=\frac{\Lambda_{3}^{4} s t}{l^{4}}+\frac{\Lambda_{3}^{3} s}{l^{3}} \tag{15}
\end{equation*}
$$

Combining (12) and (15) we have that $\operatorname{Cov}\left(T_{s}, T_{t}\right)=\frac{\Lambda_{3}^{3} s}{l^{3}}$. We can now evaluate the sum of covariances in (9)

$$
\begin{equation*}
\sum_{s \neq t} \operatorname{Cov}\left(T_{s}, T_{t}\right)=\frac{2 \Lambda_{3}^{3}}{l^{3}} \sum_{s=1}^{l-1} \sum_{t=s+1}^{l} s=\frac{2 \Lambda_{3}^{3}}{l^{3}}\left(\frac{l^{3}-l}{6}\right) \tag{16}
\end{equation*}
$$

Taking the limit $\lim _{l \rightarrow \infty} \sum_{s \neq t} \operatorname{Cov}\left(T_{s}, T_{t}\right)=\frac{\Lambda_{3}^{3}}{3}$ and combining it with (11) we may extract the variance

$$
\begin{equation*}
\operatorname{Var}\left(\sigma_{3}\right)=\frac{2 \Lambda_{3}^{3}}{3}+\frac{\Lambda_{3}^{2}}{2} \tag{17}
\end{equation*}
$$

Similarly we can extract higher order moments of the distribution using this technique. For example the third moment $\mathbb{E}\left[\left(\sigma_{3}-\mathbb{E}\left[\sigma_{3}\right]\right)^{3}\right]=-\frac{5 \Lambda_{3}^{5}}{6}-\frac{\Lambda_{3}^{4}}{5}$, which can be used to analyse the skewness of the distribution.

## V. Generalisation to K-Hop Shortest Paths

More generally for $L \in\left((k-1) r_{0}, k r_{0}\right)$ with integer $k$ there will be $k-1$ lenses of equal width $\left|L_{i}\right|=k r_{0}-L$. The method of (IV) can still be used. For general $k$ we have

$$
\begin{equation*}
\sigma_{k}=\lim _{l \rightarrow \infty} \sum_{q_{k-1}=1}^{l} \ldots \sum_{q_{2}=1}^{q_{3}} \sum_{q_{1}=1}^{q_{2}} Y_{k-1, q_{k-1}} \ldots Y_{2, q_{2}} Y_{1, q_{1}} \tag{18}
\end{equation*}
$$

where the $Y_{i, q_{i}}$ are Poisson with mean $\Lambda_{k} / l$. Using that $\sum_{k=1}^{n} k^{\theta}=n^{\theta+1} /(\theta+1)+o\left(n^{\theta+1}\right)$ we can derive the mean

$$
\begin{equation*}
\mathbb{E}\left[\sigma_{k}\right]=\frac{\Lambda_{k}^{k-1}}{(k-1)!} \tag{19}
\end{equation*}
$$

Now, letting $T_{q}^{(3)}=Y_{2, q} \sum_{r=1}^{q} Y_{1, r}$, we recursively define

$$
\begin{equation*}
T_{q}^{(n+1)}=Y_{n, q} \sum_{r=1}^{q} T_{r}^{(n)} \tag{20}
\end{equation*}
$$

By further defining $\tau_{l}^{(n)}=\sum_{r=1}^{l} T_{r}^{(n)}$, such that $\sigma_{k}=$ $\lim _{l \rightarrow \infty} \tau_{l}^{(n)}$ we can recursively define the expectation of $\tau_{l}^{(n)}$

$$
\begin{equation*}
\mathbb{E}\left[\tau_{l}^{(n+1)}\right]=\frac{\Lambda}{l} \sum_{r=1}^{l} \mathbb{E}\left(\tau_{l}^{(n)}\right) \tag{21}
\end{equation*}
$$

where $\Lambda / l$ is the mean of the Poisson variables $Y_{i, j}$. Similarly, for the variance of $\tau_{l}^{(n+1)}$ we calculate

$$
\begin{equation*}
\sum_{r=1}^{l} \operatorname{Var}\left(T_{r}^{(n+1)}\right)+2 \sum_{s=1}^{l-1} \sum_{t=s+1}^{l} \operatorname{Cov}\left(T_{s}^{(n+1)}, T_{t}^{(n+1)}\right) \tag{22}
\end{equation*}
$$

Using the recurrence relation we have

$$
\begin{equation*}
\operatorname{Var}\left(T_{r}^{(n+1)}\right)=\left(\frac{\Lambda^{2}}{l^{2}}+\frac{\Lambda}{l}\right) \operatorname{Var}\left(\tau_{r}^{(n)}\right)+\frac{\Lambda}{l} \mathbb{E}\left(\tau_{r}^{(n)}\right)^{2} \tag{23}
\end{equation*}
$$

For the covariance we have

$$
\begin{align*}
& \operatorname{Cov}\left(T_{s}^{(n+1)}, T_{t}^{(n+1)}\right)=\mathbb{E}\left[T_{s}^{(n+1)} T_{t}^{(n+1)}\right]-\mathbb{E}\left[T_{s}^{(n)}\right] \mathbb{E}\left[T_{t}^{(n)}\right] \\
& =\frac{\Lambda^{2}}{l^{2}}\left(\mathbb{E}\left[\sum_{p=1}^{s} T_{p}^{(n)}\left(\sum_{p=1}^{s} T_{p}^{(n)}+\sum_{p=s+1}^{t} T_{t}^{(n)}\right)\right]-\mathbb{E}\left[\tau_{s}^{(n)}\right] \mathbb{E}\left[\tau_{t}^{(n)}\right]\right) \\
& =\frac{\Lambda^{2}}{l^{2}}\left(\sum_{p=1}^{s} \sum_{r=s+1}^{t} \mathbb{E}\left[T_{p}^{(n)} T_{r}^{(n)}\right]+\mathbb{E}\left[\left(\tau_{s}^{(n)}\right)^{2}\right]-\mathbb{E}\left[\tau_{s}^{(n)}\right] \mathbb{E}\left[\tau_{t}^{(n)}\right]\right) \tag{24}
\end{align*}
$$

where we have used that $t>s$ and split the sum into parts. We now have

$$
\begin{align*}
& \operatorname{Cov}\left(T_{s}^{(n+1)}, T_{t}^{(n+1)}\right)=\frac{\Lambda^{2}}{l^{2}}\left[\operatorname{Var}\left(\tau_{s}^{(n)}\right)-\mathbb{E}\left[\tau_{s}^{(n)}\right] \mathbb{E}\left[\tau_{t}^{(n)}\right]\right. \\
& \left.+\mathbb{E}\left[\tau_{s}^{(n)}\right]^{2}+\sum_{p=1}^{s} \sum_{r=s+1}^{t} \operatorname{Cov}\left(T_{p}^{(n)}, T_{r}^{(n)}\right)+\frac{\Lambda^{2}}{l^{2}} \mathbb{E}\left[\tau_{p}^{(n)}\right] \mathbb{E}\left[\tau_{r}^{(n)}\right]\right] \tag{25}
\end{align*}
$$

Letting $n=k$ and $\Lambda / l=\Lambda_{k} / l$ we can combine (21), (22), (23) and (25) to obtain the variance for $\sigma_{k}$. For example, for $k=4$ we have $\operatorname{Var}\left(\sigma_{4}\right)=\frac{6 \Lambda_{4}^{5}+15 \Lambda_{4}^{4}+10 \Lambda_{4}^{3}}{60}$. This recursion relation allows us to derive the variance of $\sigma_{k}$, which involves evaluating a $(k-1)$-fold sum of products of random variables (see (18)) in terms of a simpler $(k-2)$-fold sum.

## VI. Conclusion

Motivated by the multihop diffusion of information in VANETs, realised through the periodic broadcasts of BSMs as mandated by the DSRC standard [3], we have studied the statistics of the number $\sigma_{k}$ of shortest $k$-hop paths in 1D random networks. Namely, we have derived simple closed form expressions for the mean and variance of $\sigma_{k}$ for $k=3,4$, provided a recursive formula for general $k$, and have confirmed them numerically using Monte Carlo simulations (see Fig. 5). We argue that knowledge of such statistics can be used to enhance throughput [15], validate threat events, protect against collusion attacks, infer location information, and also limit redundant broadcasts thus reducing interference. As an example, consider the realistic scenario where there are about
$\lambda=100$ vehicles per km , transmission range is $r_{0}=0.3$ km , and a vehicle detects an event and broadcasts a BSM containing relevant safety information which should reach at least a range of $L=1 \mathrm{~km}$ from the epicentre of the detected event. It follows that the length of the shortest multihop path is $k=\left\lceil\frac{L}{r_{0}}\right\rceil=4$, and that the expected number of shortest paths is $\mathbb{E}\left[\sigma_{4}\right]=1333.33$ (in either forward or backward direction). This is clearly unnecessary, and only a fraction of $\nu \in(0,1]$ vehicles should re-broadcast the original BSM. Thus inverting (19), we can calculate the re-broadcast probability $\nu=\frac{(\varsigma(k-1)!)^{1 /(k-1)}}{\lambda\left(k r_{0}-L\right)}$, where $\varsigma$ is the target number of shortest paths, e.g., setting $\varsigma=10$ we estimate that just $19.5 \%$ of vehicles should re-broadcast the original BSM.

## Acknowledgements

This work was supported by the EPSRC grant number EP/N002458/1 for the project Spatially Embedded Networks.

## REFERENCES

[1] K. A. Hafeez, L. Zhao, B. Ma, and J. Mark, "Performance analysis and enhancement of the DSRC for VANET's safety applications," Vehicular Technology, IEEE Transactions on, vol. 62, no. 7, pp. 3069-3083, 2013.
[2] Z. Yan, H. Jiang, Z. Shen, Y. Chang, and L. Huang, "k -connectivity analysis of one-dimensional linear vanets," IEEE Transactions on Vehicular Technology, vol. 61, pp. 426-433, Jan 2012.
[3] J. B. Kenney, "Dedicated short-range communications (dsrc) standards in the united states," Proceedings of the IEEE, vol. 99, no. 7, pp. 11621182, 2011.
[4] S. Najafzadeh, N. Ithnin, S. A. Razak, and R. Karimi, "Bsm: broadcasting of safety messages in vehicular ad hoc networks," Arabian Journal for Science and Engineering, vol. 39, no. 2, pp. 777-782, 2014.
[5] M. A. Javed, D. T. Ngo, and J. Y. Khan, "A multi-hop broadcast protocol design for emergency warning notification in highway vanets," EURASIP Journal on Wireless Communications and Networking, vol. 2014, no. 1, p. 179, 2014.
[6] M. Penrose, Random geometric graphs. Oxford University Press, 2003.
[7] O. Georgiou, C. P. Dettmann, and J. P. Coon, "Connectivity of confined 3D networks with anisotropically radiating nodes," Wireless Communications, IEEE Transactions on, vol. 13, pp. 4534-4546, Aug 2014.
[8] I. Stoianov, L. Nachman, S. Madden, and T. Tokmouline, "Pipenet: A wireless sensor network for pipeline monitoring," in 2007 6th International Symposium on Information Processing in Sensor Networks, pp. 264-273, April 2007.
[9] T. W. Larsen, K. D. Petersson, F. Kuemmeth, T. S. Jespersen, P. Krogstrup, J. Nygård, and C. M. Marcus, "Semiconductor-nanowirebased superconducting qubit," Phys.Rev.Lett., vol. 115, p. 127001, 2015.
[10] S. Vural and E. Ekici, "Probability distribution of multi-hop-distance in one-dimensional sensor networks," Computer Networks, vol. 51, no. 13, pp. $3727-3749,2007$.
[11] G. Mao, Z. Zhang, and B. Anderson, "Probability of k-hop connection under random connection model," Communications Letters, IEEE, vol. 14, no. 11, pp. 1023-1025, 2010.
[12] S. Chandler, "Calculation of number of relay hops required in randomly located radio network," Electronics Letters, vol. 25, no. 24, pp. 16691671, 1989.
[13] A. P. Giles, O. Georgiou, and C. P. Dettmann, "Betweenness centrality in dense random geometric networks," 2015 IEEE International Conference on Communications (ICC), London, UK, 2015.
[14] C. Nguyen, O. Georgiou, and Y. Doi, "Maximum likelihood based multihop localization in wireless sensor networks," in 2015 IEEE International Conference on Communications (ICC), London, UK, pp. 66636668, 2015.
[15] G. V. Rossi, K. K. Leung, and A. Gkelias, "Density-based optimal transmission for throughput enhancement in vehicular ad-hoc networks," in Communications (ICC), 2015 IEEE International Conference on, pp. 6571-6576, IEEE, 2015.
[16] C. Sommer, "Shortest-path queries in static networks," ACM Comput. Surv., vol. 46, pp. 45:1-45:31, Mar. 2014.

