Euro. Jnl of Applied Mathematics (2012), *vol.* 23, *pp.* 485–514. © Cambridge University Press 2012
doi:10.1017/S095679251200006X

A stochastic analysis of greedy routing in a spatially dependent sensor network

HOLGER P. KEELER

Department of Mathematics and Statistics, University of Melbourne, Victoria, 3010, Melbourne, Australia and INRIA, Paris, France
email: keeler@ms.unimelb.edu.au

(*Received 18 March 2010; revised 16 January 2012; accepted 16 January 2012; first published online 27 March 2012*)

For a sensor network, a tractable spatially dependent node deployment model is presented with the property that the density is inversely proportional to the sink distance. A stochastic model is formulated to examine message advancements under greedy routing in such a sensor network. The aim of this work is to demonstrate that an inhomogeneous Poisson process can be used to model a sensor network with spatially dependent node density. Symmetric elliptic integrals and asymptotic approximations are used to describe the random behaviour of hops. Types of dependence that affect hop advancements are examined. We observe that the dependence between successive jumps in a multi-hop path is captured by including only the previous forwarding node location. We include a simple uncoordinated sleep scheme, and observe that the complexity of the model is reduced when sufficiently many nodes are asleep. All expressions involving multi-dimensional integrals are derived and evaluated with quasi-Monte Carlo integration methods based on Halton sequences and recently developed lattice rules. An importance sampling function is derived to speed up the quasi-Monte Carlo methods. The ensuing results agree extremely well with simulations.

Key words: ad hoc network; elliptic integrals; non-stationary process

1 Introduction

Advancements in networking technologies are leading to sensor networks being a feasible and common technology. Sensor networks consist of electronic sensing devices known as *sensor nodes*. The nodes are deployed over a region known as a sensor field to gather environmental information. Each node has the ability to collect and process environmental data within its sensing range and to communicate with other nodes within its transmission range. The collected data are ultimately relayed, often via surrounding nodes, to a main station known as a *sink*. The applications of sensor networks are valuable and diverse, and they include security and traffic surveillance, environmental and animal monitoring, natural disaster warning and analysis, and building and structure assessment [2, 8, 30].

A significant issue for sensor networks is developing routing methods that can handle their dynamic topologies. A common approach in sensor networks and ad hoc networks in general is geometric or position-based routing [23, 28]. The operation of these algorithms is based on the assumption that each node knows its geographical location in relation to the sink and the location of neighbouring nodes within its transmission radius. Geometric routing is often considered attractive because of its localized nature and scalability. We

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

refer to the node from where a data message originates as the *source* node. A natural geometric routing approach is for this source node to forward a message to the node that is within the source node's transmission range and geographically the closest to the sink, and to repeat this step until the message finally reaches the target sink. This approach serves as a *greedy* routing method in itself or forms the basis for more intelligent routing methods in wireless ad hoc networks [5, 15, 19, 32].

Sensor nodes are often randomly scattered over the sensor field. To conserve power-consumption, a subset of sensor nodes may randomly fall into a low energy-consuming *sleep* state, in which they cannot relay messages. The inherent randomness in sensor networks motivates the need for a suitable stochastic model to determine the ability of a routing scheme to successfully deliver data. In recent years, stochastic models and methods are being increasingly employed in analyzing communication networks. In particular, there is a special issue [10] on stochastic geometry and related fields applied to communication networks, as well as a two-volume monograph by Baccelli and Blaszczyszyn [3, 4].

The majority of this work, however, is under the assumption that sensor nodes are deployed according to a homogeneous Poisson process. Although such a tractable model might not capture the underlying node density variation, it can serve as a first approximation for studying network characteristics. We wish to extend the standard model by examining inhomogeneous node deployment such that the node density is spatially dependent.

There are various reasons why inhomogeneous deployment models may be necessary. The deployment of sensor nodes depends on the environment and application of the sensor network. Consequently, sensor nodes may need to be deployed more densely in important sensing regions. The obstacles in the network surroundings may prevent nodes from being positioned in certain regions. Inconsistent system parameters (such as battery lifetime) and interference can reduce the effective node density in certain regions. Furthermore, the nature of the actual node deployment influences the node density. For example, an aerial dispersal of nodes could result in the nodes obeying some type of diffusion process. Also, there may be network protocols that require positioning the nodes in certain regions, which lead to performance advantages.

The design and deployment of sensor networks must address the problem of message collisions. The data messages in sensor networks converge towards the sink. Nodes closer to the sink need act as relay nodes more than nodes away from the sink. One possible solution to this problem is to deploy more nodes in these heavy traffic regions. Consequently, the node density would decay at some rate that is dependent on the distance to the sink.

It is in this last setting that we wish to examine greedy routing. We propose an approach using a tractable mathematical model similar to the one developed in previous work [17, 18]. Ideally, we want to offer a computationally quick and reliable way to obtain probabilistic descriptions of multi-hop paths in a sensor network with a simple stochastic sleep scheme. Moreover, we want to extend the model, analysis and calculations methods from the homogeneous case to the inhomogeneous case, and demonstrate that the techniques are still valid.

To achieve these goals, we analyze greedy routing in randomly deployed networks under the multi-hop situation. We propose a tractable spatially dependent node density function, and analyse the resulting stochastic characteristics. Furthermore, we examine the influence

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

of a simple stochastic sleep scheme. More specifically, we examine the effects of having a certain proportion of nodes awake at any given time. We derive probability distributions that involve multiple integrals and demonstrate their feasible evaluation via quasi-Monte Carlo integration methods based on Halton sequences and recently developed lattice rules.

The work presented here is focused on the stochastic behaviour of multi-hop paths, the calculation methods, and mathematically representing the effects of a sleep scheme. Overall, we show the application of this mathematical formulation in describing the stochastic behaviour of message delivery in sensor networks with inhomogeneous node deployment. Additionally, we demonstrate a calculation procedure based on quasi-Monte Carlo methods for evaluating multi-dimensional integrals.

2 Background work

The results presented here follow on from initial homogeneous Poisson model development and analysis [17], which was later extended by evaluating resulting integrals and analysing a simple sleep scheme [18]. Consequently, the majority of the formulation and solution techniques used here have been applied in the constant density setting.

Ishizuka and Aida [13] performed analysis on node placement approaches via simulations to gauge the fault tolerance of networks against random node failure and battery exhaustion. In particular, they assumed that individual nodes were scattered uniformly around the sink, and that the density decreased as the distance to the sink increased. In addition to the homogeneous model, Ishizuka and Aida examined two models where in one the density is inversely proportional to the sink distance, and in the other the density is a Gaussian function (its standard deviation was chosen such that 99% of nodes were found in the test region). They concluded that the simple inverse function resulted in the best fault tolerance overall. Ishizuka and Aida [14] later examined a more general power-law model and concluded again that the sensor networks are the most resilient to node failure when the density inversely proportional to sink distance.

A concept closely related to connectivity is the sensing coverage of a sensor network, which is the ability of a sensor network to successfully sense or cover the entire sensor field. Solutions to coverage problems have been based on coverage processes such as the Boolean model (see Hall [11] and Stoyan, Kendall and Mecke [29] for more details). In the sensor network setting, a more recent example is that of Pallavi *et al.* [25] who used coverage processes to examine the coverage of a sensor network with an exponentially decreasing node density.

The aforementioned work involving inhomogeneous node deployment cases did not examine the stochastic behaviour of any particular routing method. Furthermore, the work did not cover the effects that a sleep scheme has on stochastic dependencies in the routing model.

3 Mathematical model

We present a two-dimensional model that neglects the earth's curvature. We assume that nodes communicate data radially, and that a node's transmission radius is a constant that clearly cuts off at some distance, which implies that a node can relay data only to

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

another node when it is within the forwarding node's transmission radius. For numerical calculations and simulations, the transmission radius is set to one. However, we denote the transmission radius r in ensuing calculations and equations for clarity and future extensions.

We assume that nodes are scattered according to a two-dimensional Poisson process over a finite region, and that at any given time a random number of nodes are in sleep mode while the remaining are in their awake mode. Furthermore, we assume that the awake node density or the average number of awake nodes per unit area is a spatially dependent function; that is, the nodes are scattered according to an inhomogeneous Poisson process [29]. We assume the nodes are scattered uniformly around the sink, but decreases in some way as the distance to the sink increases. Hence, the awake node density is a radial function of the form

$$\lambda(u) = \lambda q(u) \qquad u \in (0, \infty), \tag{3.1}$$

where u is the distance to the sink, $q(u)$ is a non-negative shaping function and λ is a positive constant. The function $\lambda(u)$ can be interpreted as the mean number of awake nodes per infinitesimal area element. We refer to the constant λ as the *initial* node density, and use it to scale the density function.

Ideally, the $q(u)$ function should be amenable to analytical and asymptotic methods while reflecting a realistic node placement. We study a node density that is inversely proportional to the sink distance

$$q(u) = \frac{1}{u}, \tag{3.2}$$

which is the model proposed by Ishizuka and Aida [13]. Their simulation work showed that overall this node deployment model outperformed a Gaussian model in both tolerance against battery exhaustion and random node failure. We also believe that this positioning of nodes will better accommodate the convergence of messages near the sink. Consequently, we examine this model due its observed tolerance in simulations, while still being tractable to mathematical methods.

We introduce a Poisson process *mean* measure Λ, which for a bounded Borel set $B \subset \mathbb{R}^2$ with area A is the density function integrated over the region B, thus in polar coordinates

$$\Lambda(B) = \lambda \iint_B q(u) u \, du \, d\theta, \tag{3.3}$$

$$= \lambda Q(B), \tag{3.4}$$

where Q is referred to as the *re-scaled* mean measure, and is used such that the notation is analogous to the results based on the constant density case when Q reduces to the area A of the region [17, 18]. We emphasise that all the Q-type expressions we consider in this work are derived in a similar manner to the corresponding area expressions under the homogeneous model by integrating the density over specific regions [17, 18].

Under a sleep model, the initial node density parameter can be written as $\lambda = p\alpha$ where p is the probability of a node being awake and α is the underlying (that is, sleep and awake) node density parameter. The awake nodes form a thinned Poisson process. As in the homogeneous case, both the number of points kept and the number of points

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

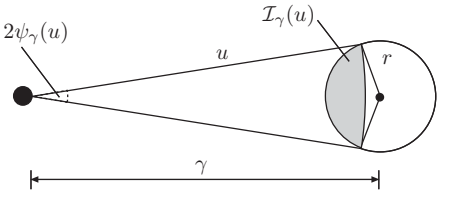


FIGURE 1. Integrate over the region $\mathscr{I}_\gamma(u)$ to derive $\Lambda(\mathscr{I}_\gamma(u))$ for the forwarding node.

removed form random variables that are independent of each other [29]. Hence, in our model the number of awake nodes is independent of the number of asleep nodes, which are examined further in Section 6.

The number of awake nodes N_B located within some region B is a inhomogeneous Poisson random variable with a probability mass function

$$\mathbb{P}(N_B = n) = \frac{(\lambda Q(B))^n}{n!} e^{-\lambda Q(B)}. \tag{3.5}$$

4 Single hop analysis

We introduce the parameter γ to represent the distance between a forwarding node and the sink (respectively located on the right and the left side in Figure 1). Often we shall present results such that the initial node density parameter λ is a product of the source-sink distance and some positive constant. For a forwarding node with a sink distance γ, let $\mathscr{I}_\gamma(u) \subset \mathbb{R}^2$ be its partial feasible region as a function of u. This region is formed by the intersection of two circles of radii r and u centred at the source node and the sink, respectively (as illustrated on the right in Figure 1). The area of $\mathscr{I}_\gamma(u)$ is denoted by $A_\gamma(u)$, and given by the integral

$$A_\gamma(u) = 2\int_{\gamma-r}^{u}\int_{0}^{\psi_\gamma(w)} w d\theta dw, \tag{4.1}$$

where the sink angle function $\psi_\gamma(u)$, by the law of cosines, is defined as

$$\psi_\gamma(u) = \arccos\left(\frac{u^2+\gamma^2-r^2}{2u\gamma}\right). \tag{4.2}$$

Furthermore, the mean measure of the region $\mathscr{I}_\gamma(u)$ is given by the integral

$$\Lambda(\mathscr{I}_\gamma(u)) = 2\int_{\gamma-r}^{u}\int_{0}^{\psi_\gamma(w)} \lambda(w) w d\theta dw, \tag{4.3}$$

which for the functions (3.1) and (3.2) reduces to

$$\Lambda(\mathscr{I}_\gamma(u)) = 2\lambda\int_{\gamma-r}^{u}\int_{0}^{\psi_\gamma(w)} d\theta dw \tag{4.4}$$

$$= 2\lambda\int_{\gamma-r}^{u} \psi_\gamma(w) dw. \tag{4.5}$$

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

Henceforth, we write

$$\Lambda_\gamma(u) = \Lambda(\mathscr{I}_\gamma(u)), \qquad Q_\gamma(u) = Q(\mathscr{I}_\gamma(u)),$$

to refer to the mean measure and the re-scaled mean measure, respectively, the motivation of which will become apparent by the analogous results that follow.

To calculate the mean measure $\Lambda_\gamma(u)$, we write integral (4.5) as

$$2\lambda \int_{\gamma-r}^{u} \psi_\gamma(w)dw = 2\lambda u\psi_\gamma(u) + 2\lambda \int_{\gamma-r}^{u} \frac{(w^2 + r^2 - \gamma^2)}{\left([w^2 - (\gamma - r)^2][(\gamma + r)^2 - w^2]\right)^{1/2}} dw, \tag{4.6}$$

whose solution is obtained with the reduction of general elliptic integrals by symmetric elliptic integrals (see Section 19.29 of [1] for further details and examples). The final solution of (4.5) is

$$\begin{aligned}\Lambda_\gamma(u) = \ & 2\lambda \Big[u\psi_\gamma(u) - \frac{1}{3}(\gamma - r)^2(\gamma + r)^2 R_D(v^2 + (\gamma - r)^2, v^2 + (\gamma + r)^2, v^2) + u(\gamma - r)/v \\ & + (r^2 - \gamma^2) R_F(v^2 + (\gamma - r)^2, v^2 + (\gamma + r)^2, v^2)\Big], \qquad u \in (\gamma - r, \gamma],\end{aligned} \tag{4.7}$$

where

$$v = \frac{\gamma - r}{u^2 - (\gamma - r)^2} \left([u^2 - (\gamma - r)^2][(\gamma + r)^2 - u^2)]\right)^{1/2}, \tag{4.8}$$

and R_F and R_D are symmetric elliptic integrals in Carlson form

$$R_F(x, y, z) = \frac{1}{2} \int_0^\infty \frac{dt}{[(t + x)(t + y)(t + z)]^{1/2}}, \tag{4.9}$$

$$R_D(x, y, z) = \frac{3}{2} \int_0^\infty \frac{dt}{(t + z)\,[(t + x)(t + y)(t + z)]^{1/2}}. \tag{4.10}$$

For calculation purposes, we note that naturally

$$\lim_{u \to \gamma - r} \Lambda_\gamma(u) = 0. \tag{4.11}$$

The results presented here are obtained via our purposely written elliptic integrals based on the papers by Carlson [6, 7], which give algorithms that are readily implementable and can handle both real and complex values under specified variable and parameter regimes. We found that the solutions are calculated quickly owing to the speedy convergence of the Carlson algorithms.

Take as our sample space Ω the set of two-dimensional point processes on $\mathbb{R}^2$, together with an appropriate σ-field $\mathscr{F}$ of subsets of Ω and the probability measure $\mathbb{P}$ induced on this space by our Poisson intensity measure (4.7). For a node at distance γ from the origin, let the random variable U with resepect to $(\Omega, \mathscr{F}, \mathbb{P})$ be the sink distance of the forwarding node after a single message hop. To derive the probability distribution of U, the nearest neighbour [29] approach is used in which $\mathbb{P}(U > u)$ is equated to the probability of no nodes existing in the feasible region at a distance less or equal to u. This

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

argument is analogous to that of the homogeneous case [17], hence the distribution

$$F_\gamma(u) = \begin{cases} 1 - e^{-\lambda Q_\gamma(u)} & \gamma - r \leqslant u < \gamma \\ 1 & u \geqslant \gamma \\ 0 & u < \gamma - r \end{cases} \tag{4.12}$$

follows. There is a jump discontinuity in the distribution at $u = \gamma$ owing to the positive probability that no nodes lie within the feasible region. On the support where the distribution (4.12) is absolutely continuous the probability density is defined by

$$f(u) = \lambda Q'_\gamma(u) e^{-\lambda Q_\gamma(u)}, \qquad \gamma - r \leqslant u < \gamma, \tag{4.13}$$

where the derivative of the re-scaled mean measure

$$Q'_\gamma(u) = 2\psi_\gamma(u).$$

Let $C = \gamma - U$ be the distance advanced towards the sink when the originating node is at a distance γ. Let $\bar{F}$ denote the distribution of C, which leads to

$$\bar{F}_\gamma(c) = \begin{cases} e^{-\lambda Q_\gamma(\gamma - c)} & 0 < c \leqslant r \\ 1 & c > r \\ 0 & c \leqslant 0, \end{cases} \tag{4.14}$$

and its probability density is given by

$$\bar{f}(c) = \lambda Q'_\gamma(\gamma - c) e^{-\lambda Q_\gamma(\gamma - c)}, \qquad 0 < c \leqslant r. \tag{4.15}$$

Given that C is non-negative, the mth moment expression

$$\mathbb{E}(C^m) = m \int_0^r c^{m-1} \mathbb{P}(C > c) dc, \tag{4.16}$$

$$= r^m - m \int_0^r c^{m-1} e^{-\lambda Q_\gamma(\gamma - c)} dc, \tag{4.17}$$

follows.

4.1 Asymptotic results

We derive an asymptotic approximation to the re-scaled mean measure $Q_\gamma(u)$, which gives a tractable and accurate expression. The feasible region of the forwarding node approaches zero at the point $u = \gamma - r$, and is the point closest to the sink. The angle function ψ_γ expanded at this point gives

$$\psi_\gamma(u) \approx b_0(u - \gamma + r)^{1/2} + b_1(u - \gamma + r)^{3/2} + b_2(u - \gamma + r)^{5/2}, \tag{4.18}$$

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

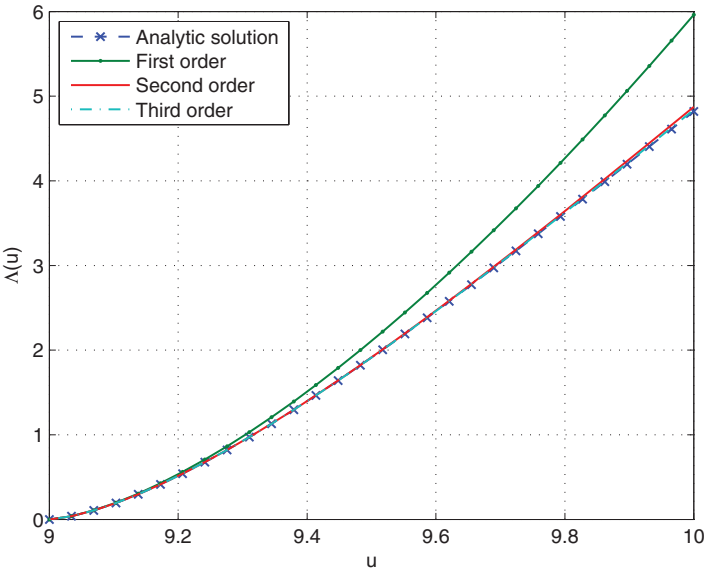


FIGURE 2. (Colour online) Comparison of $\Lambda_\gamma(u)$ asymptotic expressions ($\lambda = 3\gamma$, $r = 1$ and $\gamma = 10$).

where the expansion terms

$$b_0 = \left[\frac{2r}{\gamma(\gamma - r)}\right]^{1/2}, \tag{4.19}$$

$$b_1 = \left[\frac{2r}{\gamma(\gamma - r)}\right]^{1/2} \left[\frac{r^2 - 3r\gamma - 3\gamma^2}{12(\gamma^2 r - \gamma r^2)}\right], \tag{4.20}$$

$$b_2 = \left[\frac{2r}{\gamma(\gamma - r)}\right]^{1/2} \left[\frac{3r^4 + 25r^2\gamma^2 - 10r^3\gamma + 30\gamma^3 r - 5\gamma^4}{160\gamma^2(\gamma - r)^2 r^2}\right] \tag{4.21}$$

follow [17]. The asymptotic result for the re-scaled mean measure

$$Q_\gamma(u) \approx 4\left[\frac{b_0}{3}(u - \gamma + r)^{3/2} + \frac{b_1}{5}(u - \gamma + r)^{5/2} + \frac{b_2}{7}(u - \gamma + r)^{7/2}\right] \tag{4.22}$$

follows.

The results of the second-order approximation are accurate for a unit radius (see Figure 2). Adding the third term only improves the results slightly. However, it may be needed for larger transmission radius models (as Figure 3 suggests). Conversely, the two-term expansion gives accurate results (see Figure 4) when substituted into the sink distribution (4.12). In fact, it appears that the expansion (4.22) consisting of elementary functions can be used to give accurate results, which are clearly faster to evaluate than those based on elliptic integrals. However, we continue to use the exact solution of the integral (4.5), and later compare it to its approximation.

For large λ, we present asymptotic moment results for our spatially dependent node density model.

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

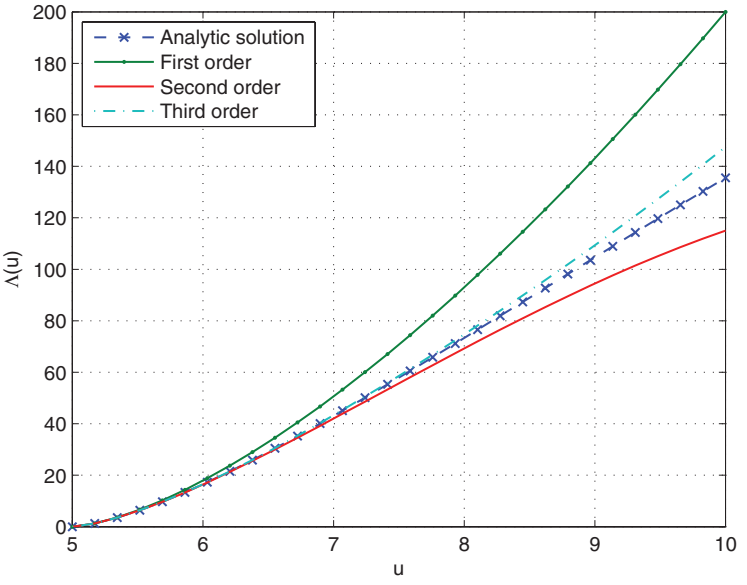


FIGURE 3. (Colour online) Comparison of $\Lambda_\gamma(u)$ asymptotic expressions ($\lambda = 3\gamma$, $r = 5$ and $\gamma = 10$).

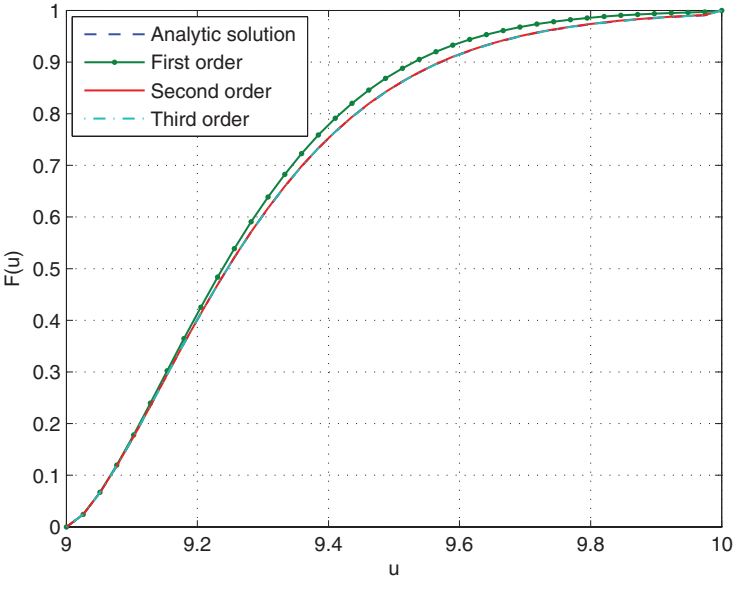


FIGURE 4. (Colour online) Comparison of $F_\gamma(u)$ asymptotic expressions ($\lambda = 3\gamma$, $r = 1$ and $\gamma = 10$).

Theorem 4.1 *For the mean measure (4.7), provided $\gamma > r$, under greedy routing the first hop moment*

$$\mathbb{E}(C) \sim r - \frac{\Gamma(5/3)}{(\lambda q_0)^{2/3}}, \tag{4.23}$$

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

and the second hop moment

$$\mathbb{E}(C^2) \sim r^2 - 2r\frac{\Gamma(5/3)}{(\lambda q_0)^{2/3}} + \frac{\Gamma(7/3)}{(\lambda q_0)^{4/3}}, \tag{4.24}$$

as the initial node density

$$\lambda \to \infty,$$

where Γ *is the gamma function, and*

$$q_0 = \frac{4}{3}\left[\frac{2r}{\gamma(\gamma - r)}\right]^{1/2}.$$

Proof Consider integrals of the form

$$I(\lambda) = \int_a^b \phi(t)e^{-\lambda Q(t)}dt. \tag{4.25}$$

Assume the real function $Q(t)$ has one minimum on the interval $[a, b]$, which occurs at $t = a$, and that

$$Q(t) \sim Q(a) + \sum_{s=0}^{\infty} q_s(t-a)^{s+\mu}, \tag{4.26}$$

and

$$\phi(t) \sim \sum_{s=0}^{\infty} \phi_s(t-a)^{s+\beta-1}, \tag{4.27}$$

as $t \to a^+$, and λ and μ are positive constants, and the constant β can be real or complex provided that the real part is positive (for more details see Laplace's method 31, page 58). Furthermore, assume that the first expansion (4.26) can be differentiated

$$Q'(t) \sim \sum_{s=0}^{\infty}(s+\mu)q_s(t-a)^{s+\mu-1}, \tag{4.28}$$

as $t \to a^+$. Provided that $Q(t) > Q(a)$ for all $t \in (a, b)$, then Laplace's method can be applied to integrals of the form

$$I(\lambda) = \int_a^b (t-a)^k e^{-\lambda Q(t)}dt, \tag{4.29}$$

where under our setting $k = 0$ or $k = 1$, and $a = 0$, thus giving

$$I(\lambda) \sim \frac{\Gamma(\tau)}{\mu(\lambda q_0)^{\tau}}, \quad \lambda \to \infty,$$

where

$$\tau = \frac{2(k+1)}{3}.$$

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

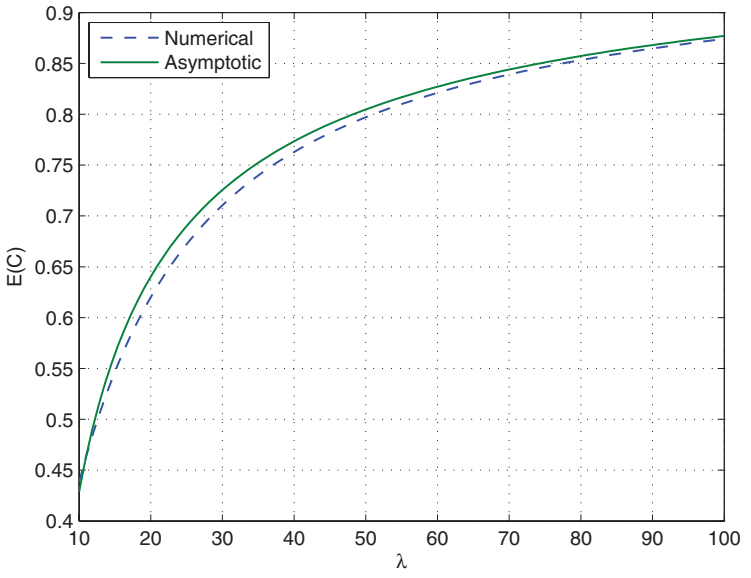


FIGURE 5. (Colour online) Numerical and asymptotic results of first moment $\mathbb{E}(C)$ ($r = 1$ and $\gamma = 10$).

Since $Q(a)$ needs to be the minimum on the integral interval, we use the change of variable $t = u - \gamma + r = r - c$ in the expansion of $Q(t)$, which, with a slight abuse of notation, leads to

$$Q_\gamma(t) \sim q_0 t^{3/2} + O(t^{5/2}) \quad \text{as} \quad t \to 0, \tag{4.30}$$

and

$$q_0 = \frac{4b_0}{3}, \qquad \mu = \frac{3}{2}.$$

The first moment result (4.23) follows by setting $k = 0$. The change of variable applied to the second moment equation gives

$$\mathbb{E}(C^2) = r^2 - 2\int_0^r (r-t)e^{-\lambda Q_\gamma(t)}dt,$$

which leads to the second result (4.24) by setting $k = 0$ and $k = 1$ accordingly. □

The asymptotic approximation (4.23) for $\mathbb{E}(C)$ demonstrates good results (in Figure 5), while similar results hold for the second moment.

4.2 Sink dependence

Since γ is the distance of an arbitrary node forwarding a message, we set $\gamma = \ell$ when the forwarding node is the source node. The node intensity function is clearly dependent on the source node sink distance. Comparing the hop distributions of messages from two different sources (in Figure 6) reveals that a message is relayed farther in a single hop if the forwarding node is closer to the sink. Intuitively, hops increase stochastically as the message approaches the sink as more potential forwarding nodes are available in the

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

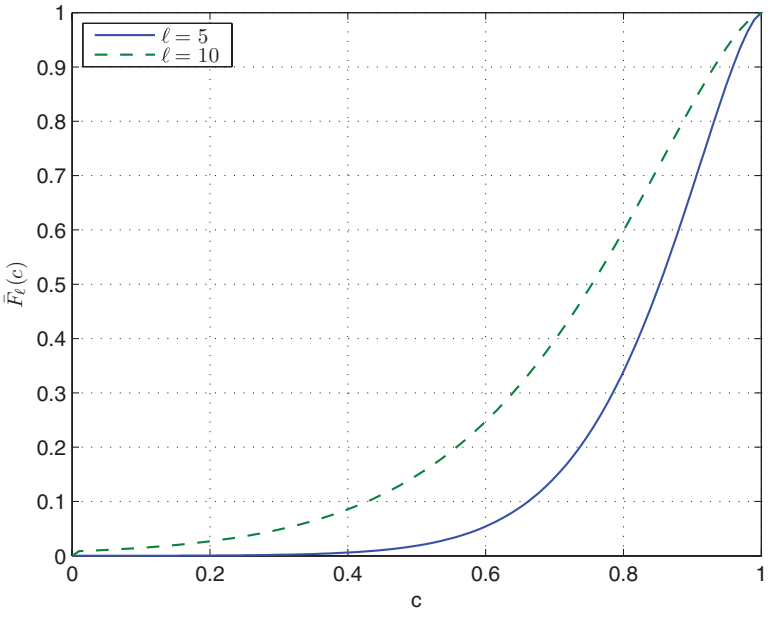


FIGURE 6. (Colour online) Comparison of hop distributions $\bar{F}_\ell(C)$ for different ℓ ($\lambda = 30$, $r = 1$, and $\ell = 5$ and $\ell = 10$).

forwarding region. Geometrically, we observe that the integral kernel in the mean measure equation (4.5) is the angle function ψ_γ, which decreases as γ increases; that is

$$\psi_{\gamma_1}(u) \leqslant \psi_{\gamma_2}(u), \qquad \gamma_1 \geqslant \gamma_2 \geqslant r. \tag{4.31}$$

Conversely, ψ_γ and, hence, the integral (4.7) increases as the sink distance decreases. This dependence on the sink distance of the forwarding node is simply referred to as the *sink* dependence.

The influence of the sink dependence can be observed by comparing the difference in two hop distributions with different sink distances γ_1 and γ_2. In previous work [17], the hop distribution dependence on the sink distance under the homogeneous model was examined by a Kullback-Leibler. The Kullback-Leibler divergence [20], also known as relative entropy, is an asymmetric measure of the difference between two probability distributions, and it applied to the mixed discrete-continuous hop distribution gives

$$D(\gamma_1, \gamma_2) = \int_0^r \bar{f}_{\gamma_2}(c) \log\left[\frac{\bar{f}_{\gamma_2}(c)}{\bar{f}_{\gamma_1}(c)}\right] dc + \bar{F}_{\gamma_2}(0^+) \log\left[\frac{\bar{F}_{\gamma_2}(0^+)}{\bar{F}_{\gamma_1}(0^+)}\right], \tag{4.32}$$

where the routing void (that is, no nodes in the feasible region) probability

$$\bar{F}_\gamma(0^+) = e^{-\lambda Q_\gamma(\gamma)}. \tag{4.33}$$

The Kullback-Leibler divergence is non-negative and is zero for identical distributions [20].

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

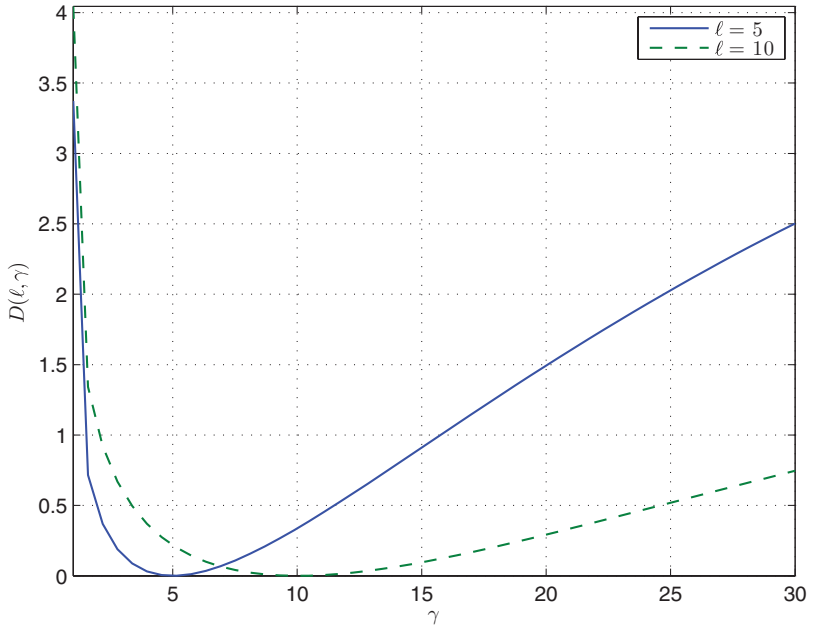


FIGURE 7. (Colour online) Kullback-Leibler analysis of hop distributions for different ℓ ($\lambda = 30$, $r = 1$, and $\ell = 5$ and $\ell = 10$).

We calculated the integral in equation (4.32) numerically to observe how the hop distribution is influenced when we set $\gamma_1 = \ell$ and vary $\gamma_2 = \gamma$. The comparison reveals that $D(\ell, \gamma)$ is high near the sink and decreases as γ increases (refer to Figure 7). This is the same expected behaviour as observed under the homogeneous model [17], however, $D(\ell, \gamma)$ decays relatively slowly under the inhomogeneous model. We observe after $D(\ell, \gamma) = 0$ at $\gamma = \ell$, it then increases at a rate that depends on the value of ℓ. This contrasts starkly with the homogeneous model where $D(\ell, \gamma)$ is mostly zero away from the sink, and only increases significantly near the sink [17]. Furthermore, $D(\ell, \gamma)$ differs significantly for two different ℓ; that is, the hop distribution varies with respect to the sink distance. This behaviour differs from the constant node density case where renewal processes can be used to model and bound message advancement over multi-hop routes due to the hop distribution varying only slightly [17, 32]; hence renewal processes cannot be used here.

5 Multi-hop analysis

For a multi-hop analysis, we introduce indexing for the random variables U and C by initially setting $U_0 = \ell$ and $U_1 = U$. For a node at distance ℓ from the origin, the random variables U_i that give the sink distance of the forwarding node after i hops are defined with respect to $(\Omega, \mathscr{F}, \mathbb{P})$. We define the ith hop advancement as $C_i = U_{i-1} - U_i$. Each C_i depends on the forwarding node's sink distance U_{i-1}; thus, the sink distance of the source node clearly affects the first hop and subsequent hops. As noted in the previous section, comparing the hop distributions demonstrates that each C_i is stochastically dominated by

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

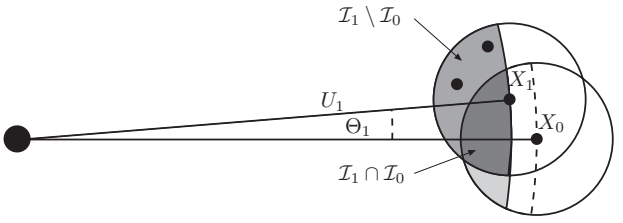


FIGURE 8. No awake nodes in the intersection region $\mathscr{I}_1 \setminus \mathscr{I}_0$ during the first message relay.

C_{i+1}. That is, for $i \geqslant 0$, we have the stochastic ordering

$$\mathbb{P}(C_{i+1} > c) \geqslant \mathbb{P}(C_i > c), \qquad c \in (0, r). \tag{5.1}$$

This inequality is opposite to the equivalent result under the homogeneous model as noted by Zorzi and Rao [32] and Keeler and Taylor [17]. Consequently, this stochastic ordering of C_i is dependent on the choice of $q(u)$, and does not hold in general. However, if given a decreasing shaping function such that $q(u) \leqslant 1/u$ for all u, then the inequality

$$uq(u)\psi_{\gamma_1}(u) \leqslant uq(u)\psi_{\gamma_2}(u), \qquad \gamma_1 \geqslant \gamma_2 \geqslant r, \tag{5.2}$$

holds, and the kernel in the mean measure integral (4.3) decreases as γ increases, and thus, under these conditions the stochastic ordering (5.1) holds.

5.1 Path dependence

Let the random variable Θ_i be the angle between the ith node and the previous node in relation to the sink. We assign the point $X_i = (U_i, \Theta_i)$ to the ith forwarding node. The source (or zeroth) node corresponds to the point $X_0 = (\ell, 0)$. A message travels i hops along a path that corresponds to a sequence of random points $\vec{X}_i = (X_0, X_1, \ldots, X_i)$.

Let $\mathscr{I}_i(u_{i+1}) \subset \mathbb{R}^2$ be the feasible region of the ith forwarding node as a function of u_{i+1} under the independent model. After the first hop the nature of greedy routing implies that another dependence arises in the distribution of U_2, which was observed under the homogeneous model [17, 18]. If a forwarding node is chosen, then there are no other nodes in the source feasible region closer to the sink. Hence, the intersection of the feasible regions of the source and the first node (that is, $\mathscr{I}_1 \cap \mathscr{I}_0$ in Figure 8) has no awake nodes. This implies that U_2 is dependent on both U_1 and Θ_1, the angle between the first node and the source node in relation to the sink. We call this dependence in both the sink distance and the sink angle the *path* dependence, and the hop model that includes both the path and the sink dependence simply the *dependent* model. Conversely, the *independent* model only includes the sink dependence.

In the dependent model with no sleep scheme, the randomness is all encapsulated in the inhomogeneous two-dimensional Poisson process that gives the node locations, and everything else is deterministically given. Conversely, the independent model has a different source of randomness. The two-dimensional Poisson process is re-sampled at each time step.

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

Thus we have to extend underlying sample space to the set Ω^∞ of sequences of realizations of two-dimensional point processes, together with the probability measure induced by the assumption that the realizations are independent and compatible with the Poisson intensity measure (4.4). The random variables U_i are now defined to give the sink distance of the forwarding node after i hops where, at each hop, we select the next node in the same manner as that described above for the dependent model, but according to the ith realization of the spatial Poisson process in the sequence.

The assumption that there are independent realizations of the underlying node distribution at each hop may be thought to be unusual, but it has been implicitly assumed before [32], and it does lead to tractable approximations and bounds. Furthermore, it becomes a good model if there is a sleep scheme in operation in which nodes are alternately available and unavailable to act as transits.

Under a sleep scheme, if nodes were asleep during the previous message relay, it is possible for recently awoken nodes to be present in this region during the current message relay. To perform the initial analysis we assume no sleep scheme exists (by setting $p = 1$). After our analysis, we include a simple sleep scheme and examine how varying p and α (while fixing λ) affects the path dependence.

We denote the probability measures derived under the dependent and independent models, respectively, by $\mathbb{P}_D$ and $\mathbb{P}_I$ (the subscript is dropped if a result applies to both models). The distribution of U_{i+1} under the independent model is dependent only on the sink distance of the current forwarding node, hence we write

$$F_i(u_{i+1}) = \mathbb{P}_I(U_{i+1} \leqslant u_{i+1} | U_i = u_i), \tag{5.3}$$

while under the dependent model the distribution is dependent on the message path, and so we write

$$G_i(u_{i+1}) = \mathbb{P}_D(U_{i+1} \leqslant u_{i+1} | \vec{X}_i = \vec{x}_i). \tag{5.4}$$

We denote the re-scaled mean measure of the feasible region under the independent and dependent models, respectively, as $Q_i(u_{i+1})$ and $\vec{Q}_i(u_{i+1})$.

The re-scaled mean measure under the independent model is always given by the original equation (4.7), and hence, the distribution and probability density of U_{i+1} are obtained by setting $\gamma = u_i$ in equations (4.12) and (4.13). Under the dependent model, if the re-scaled mean measure is given after i hops, the sink distribution

$$G_i(u_{i+1}) = \begin{cases} 1 - e^{-\lambda \vec{Q}_i(u_{i+1})} & u_i - r \leqslant u < u_i \\ 1 & u \geqslant u_i \\ 0 & u < u_i - r, \end{cases} \tag{5.5}$$

immediately follows, and where it is absolutely continuous its probability density

$$g_i(u_{i+1}) = \lambda \vec{Q}_i'(u_{i+1}) e^{-\lambda \vec{Q}_i(u_{i+1})}, \tag{5.6}$$

also follows.

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

The set representing the feasible region under the dependent model follows by excluding the intersections of previous feasible regions, namely

$$\mathscr{D}_i(u_{i+1}) = \mathscr{I}_i(u_{i+1}) \setminus \cup_{j=0}^{j=i-1} \mathscr{I}_j(u_{j+1}). \tag{5.7}$$

Under the dependent model, it is possible to calculate the re-scaled mean over the feasible region after one hop (see appendix for details). If $i \geqslant 2$, we approximate the feasible region under the dependent model

$$\mathscr{D}_i(u_{i+1}) \approx \mathscr{I}_i(u_{i+1}) \setminus \mathscr{I}_{i-1}(u_{j+1}), \tag{5.8}$$

where we refer to this approximation as the one-hop model. Results under the homogeneous model revealed that this approximation sufficiently captures the path dependence [18]. Consequently, only the location of the previous node is needed to calculate the re-scaled mean under the dependent model, and approximation (5.8) is used henceforth.

The path dependence depends on the precise locations of the previous nodes while the sink dependence which only depends on their distance from the sink. To capture this observation and use it to describe the random behaviour of message hops we need the joint density of U_i and Θ_i. Under our inhomogeneous Poisson model, the joint probability density is

$$g_i(u_{i+1}, \theta_{i+1}) = \lambda_{\mathscr{D}_i}(u_{i+1}, \theta_{i+1}) e^{-\lambda \hat{Q}_i(u_{i+1})}, \tag{5.9}$$

where the derivation is analogous to that of the homogeneous model [18] (see appendix for details). The spatially dependent initial density $\lambda_{\mathscr{D}_i}(u_{i+1}, \theta_{i+1}) = \lambda \mathbf{I}_{\mathscr{D}_i}(u_{i+1}, \theta_{i+1})$ and the indicator function of the dependent feasible region

$$\mathbf{I}_{\mathscr{D}_i}(u_{i+1}, \theta_{i+1}) = \begin{cases} 1, & (u_{i+1}, \theta_{0i}) \in \mathscr{D}_i, \\ 0, & \text{otherwise}, \end{cases} \tag{5.10}$$

where angular coordinate θ_{0i} is the angle between the source node and the ith forwarding node in relation to the sink. These expressions serve as the basis of the node density and the joint probability density under the sleep model in the next section.

6 Sleep model

We outline a simple sleep scheme that has been analysed under the homogeneous model [17]. We assume that the probability that a node is awake on each hop is p and the event that a node is awake during a transmission attempt is independent of the event that it is awake at other transmission attempts. Consequently, the intersection region ($\mathscr{I}_1 \cap \mathscr{I}_0$ in Figure 8) has a thinned initial node density $(1-p)\lambda$. The rest of the feasible region $\mathscr{I}_1 \setminus \mathscr{I}_0$ has an initial density λ. It follows that the initial node density function after one message hop is given by

$$\lambda_{\mathscr{D}_1}(u_2, \theta_2) = \lambda \left[\mathbf{I}_{\mathscr{I}_1 \setminus \mathscr{I}_0}(u_2, \theta_2) + (1-p) \mathbf{I}_{\mathscr{I}_1 \cap \mathscr{I}_0}(u_2, \theta_2) \right], \tag{6.1}$$

where the superscripts denote the indicator functions of the disjoint regions.

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

In the limit as p approaches zero and α approaches infinity with $\lambda = p\alpha$ held constant, the locations of the nodes after each hop is re-sampled, thus, completely removing the path-dependence and allowing each forwarding node to be treated like a source node. Its effects on the node density have been explored more thoroughly under the homogeneous model [18].

To calculate the Poisson mean measure over the region $\mathscr{D}_i$, the above node density function is integrated over the domain [29] leading to

$$\vec{\Lambda}_{\mathscr{D}_i}(u_{i+1}) = \int_{u_i - r}^{u_{i+1}} \int_{-\psi_{u_i}(w_{i+1})}^{\psi_{u_i}(w_{i+1})} \lambda_{\mathscr{D}_i}(w_{i+1}, \theta_{i+1}) q(w_{i+1}) w_{i+1} d\theta_{i+1} dw_{i+1}, \tag{6.2}$$

and we define

$$\vec{Q}_{\mathscr{D}_i}(u_{i+1}) = \frac{\vec{\Lambda}_{\mathscr{D}_i}(u_{i+1})}{\lambda}. \tag{6.3}$$

The joint probability density of U_i and Θ_i, that is

$$g_i(u_{i+1}, \theta_{i+1}) = \lambda_{\mathscr{D}_i}(u_{i+1}, \theta_{i+1}) e^{-\lambda \vec{Q}_{\mathscr{D}_i}(u_{i+1})}. \tag{6.4}$$

7 Multi-hop distribution

7.1 Hop advancements

We initially formulated this problem with the sink distance variable since greedy routing is naturally based on it. However, hop advancement is a more intuitive variable in describing message progress over a multi-hop path. We adopt similar notation used for the sink distance random variables, hence $\bar{F}$ and $\bar{G}$ denote the distributions under the two models. The complement of the sink distribution yields the hop distribution under both the independent and dependent models, the latter being

$$\bar{G}_i(c_{i+1}) = \begin{cases} e^{-\lambda \vec{Q}_i(u_i - c_{i+1})} & 0 < c_{i+1} \leqslant r \\ 1 & c_{i+1} > r \\ 0 & c_{i+1} \leqslant 0, \end{cases} \tag{7.1}$$

and its probability density defined on the absolutely continuous part of the support

$$\bar{g}_i(c_{i+1}) = \lambda \vec{Q}_i'(u_i - c_{i+1}) e^{-\lambda \vec{Q}_i(u_i - c_{i+1})}, \tag{7.2}$$

where a simple sum relates the hop and sink distance variables

$$u_i = \ell - \sum_{j=1}^{i} c_j. \tag{7.3}$$

7.2 Distribution of Z_n

Let the random variable Z_n represent the distance advanced by a message in n hops

$$Z_n = \sum_{i=1}^{n} C_i. \tag{7.4}$$

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

To calculate the distribution of Z_n we use the joint probability density of the random variables C_1 to C_n and Θ_1 to Θ_n, that is

$$\bar{g}_{(n-1)}(c_1,.,c_n,\theta_1,.,\theta_n)=\prod_{i=1}^{n}\lambda_{\mathscr{D}_{i-1}}(u_{i-1}-c_i,\theta_i)e^{-\lambda\vec{Q}_{i-1}(u_{i-1}-c_i)}, \tag{7.5}$$

which is defined for $c_i \in (0,r]$. The derivation of the joint probability density is similar to that under the homogeneous mode [18] which we have adapted and included in the appendix for completeness.

It follows that the distribution of message advancement after n hops is expressed by

$$\mathbb{P}_D(Z_n \leqslant z)=\int_{0^+}^{\min(z,r)} dc_1 \int_{-\psi_0(c_1)}^{\psi_0(c_1)} d\theta_1 \int_{0^+}^{\min(z-c_1,r)} dc_2 \int_{-\psi_1(c_2)}^{\psi_1(c_2)} \tag{7.6}$$

$$\ldots \int_{0^+}^{\min(z-\sum_{i=1}^{n-1}c_i,r)} dc_n \int_{-\psi_{n-1}(c_n)}^{\psi_{n-1}(c_n)} \bar{g}_{(n-1)}(c_1,.,c_n,\theta_1,.,\theta_n)d\theta_n \tag{7.7}$$

$$+\mathbb{P}_D(C_1=0)+\mathbb{P}_D((Z_1 \leqslant z)\cap(C_2=0))+\ldots \tag{7.8}$$

$$+\mathbb{P}_D((Z_{n-1} \leqslant z)\cap(C_n=0)), \tag{7.9}$$

where

$$\psi_i(c_{i+1})=\psi_{u_i}(u_i-c_{i+1}) \tag{7.10}$$

denotes the maximum angle for a sink distance given by the sink angle function (4.2). The distribution of Z_n under a sleep scheme is obtained by substituting the product of the joint probability densities (6.4).

The integral explicitly shown in the expression of $\mathbb{P}_D(Z_n \leqslant z)$ is the distribution of Z_n conditioned on the event that all hops C_i advance some positive distance. We will refer to this integral simply as the *conditional distribution* of Z_n, and denote it by

$$\mathbb{P}(Z_n \leqslant z|+)=\mathbb{P}(Z_n \leqslant z|C_1>0,\ldots,C_n>0). \tag{7.11}$$

Under the independent model the joint probability density is not a function of any of the variables θ_1 to θ_n. Hence, the equivalent integral can be analytically integrated over the sink angle domains [17], thus giving a simplified expression in the form of hop probability densities

$$\mathbb{P}_I(Z_n \leqslant z)=\int_{0^+}^{\min(z,r)} dc_1 \int_{0^+}^{\min(z-c_1,r)} dc_2 \tag{7.12}$$

$$\ldots \int_{0^+}^{\min(z-\sum_{i=1}^{n-1}c_i,r)} \bar{f}_0(c_1)\ldots\bar{f}_{n-1}(c_n)dc_n \tag{7.13}$$

$$+\mathbb{P}_I(C_1=0)+\mathbb{P}_I((Z_1 \leqslant z)\cap(C_2=0))+\ldots \tag{7.14}$$

$$+\mathbb{P}_I((Z_{n-1} \leqslant z)\cap(C_n=0)). \tag{7.15}$$

Under the dependent model, the probability of a message reaching a routing void after advancing i hops

$$\mathbb{P}_D(C_{i+1}=0|\vec{X}_i=\vec{x}_i)=e^{-\lambda\vec{Q}_i(u_i)} \tag{7.16}$$

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

follows. The routing void probability leads to the distribution of Z_n conditioned on the event that the message meets a routing void on the last hop

$$\mathbb{P}_D((Z_n \leqslant z) \cap (C_{n+1} = 0)) = \int_{0^+}^{\min(z,r)} dc_1 \int_{-\psi_0(c_1)}^{\psi_0(c_1)} d\theta_1 \cdots \int_{0^+}^{\min(z-\sum_{i=1}^{n-1} c_i, r)} dc_n \times \int_{-\psi_{n-1}(c_n)}^{\psi_{n-1}(c_n)} \bar{g}_{(n-1)}(c_1, ., c_n, \theta_1, ., \theta_n) \mathbb{P}_D(C_{n+1} = 0) d\theta_n. \quad (7.17)$$

Consider the event when a message does not advance, hence $X_i = X_{i+1}$. If there is a sleep scheme, a forwarding node can benefit by making multiple relay attempts. The number of different possible events soon results in the integral expressions needed to describe such a model growing to be intractable. We restrict the integrals by assuming that a message executes only one relay attempt.

On this note, we point out that under the homogeneous model [18], stochastic 'rules of thumb' have been proposed for how many re-attempts forwarding nodes should make before considering other options (such as message backtracking). Under the inhomogeneous model, the equivalent results can be obtained by appropriately replacing the area terms with the corresponding integrals from the re-scaled mean measures.

7.3 Number of hops

Let the random variable N represent the total number of hops required for a message to reach the sink. The distribution of N gives a different perspective of the performance of a routing method in a sensor network. The random variables N and Z_n are connected by a simple result [17], which in short says that for all $n \geqslant 1$ the relation $\mathbb{P}_D(N \leqslant n) = 1 - \mathbb{P}_D(Z_{n-1} < \ell - r)$ holds. This result has been used to calculate the distribution of N from the distribution of Z_n in the homogeneous setting [18]. We use this relation to calculate the equivalent results under our spatially dependent node density model (see results in Figures 10–12). Moreover, we note that the limiting value of $\mathbb{P}_D(N \leqslant n)$ as $n \to \infty$ is the probability that the message ever reaches the sink, which is an important performance measure of the system, and can be derived only when all the dependence is incorporated into the model.

8 Integration methods

To calculate the distributions of Z_n under the dependent model, a $2n$-fold integral (7.6) needs to be evaluated. For low n, traditional numerical integration schemes can be used. Unfortunately, integration by these methods is too slow at higher hop numbers, which motivates us to employ quasi-Monte Carlo methods.

8.1 Quasi-Monte Carlo

The integration description that follows is similar to the more detailed account 16, Chap. 4 where quasi-Monte Carlo methods were used to evaluate similar integrals under the

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

homogeneous model. In recent years, these integration methods have gained much interest owing to their speed and accuracy in evaluating high dimensional integrals.

Quasi-Monte Carlo methods are based on purely deterministic sequences of quasi-random numbers such as Halton [12] and Sobol [26] sequences. Mathematically, quasi-random sequences have low-discrepancy [24, 27]. Informally, such sequences appear 'less random' than sequences produced by regular pseudo-random number generators as they occur more evenly spaced apart.

Previous numerical work [18] has led us to use leaped Halton sequences to calculate integrals. Often integrals were also calculated via regular Monte Carlo methods to check the quasi-random approach. We also give some results based on so-called *lattice rules*, which lead to specific cases of quasi-random sequences. These rules can produce well-behaving quasi-random sequences and have been a research focus in recent years owing to their ability to counter the curse dimensionality [22]. The points arising from lattice rules are used in a similar manner to the quasi-Monte Carlo approach. There are many suggestions for lattice rules, but we lightly examine only one known as *rank-1 lattice rule*, which over the unit hyper-cube gives the integral estimate

$$\hat{I}_n f = \frac{1}{n} \sum_{k=1}^{n} f\left(\left\{k \frac{\mathbf{z}}{n}\right\}\right), \tag{8.1}$$

where the *generating vector* $\mathbf{z} \in \mathbb{Z}^s$, n is the number of function samples, and the braces give the fractional part in $[0, 1)$.

Given a generating vector, quasi-random sequences based on such a lattice rule can be clearly produced in an exceedingly fast manner. The way to quickly evaluate an integral is by choosing a suitable generating vector. However, the drawback is that lattice rules are based on input parameters known as weights which depend on the nature of the function. In particular, the weights depend on how the function varies with respect to all its variables. Furthermore, the choice of some lattice rules require the total number of function samples before the integral calculation starts. This differs from regular quasi-Monte Carlo methods in which the integral estimate can be calculated continually until a sufficient number of function samples has been taken.

Under lattice rules, the number of function samples also influences the choice of the generating vector. Consequently, using the most suitable generating vector may not be a simple task as it involves analysing the function of interest. However, a thorough analysis of which weights and quasi-random sequences are the most suitable in this setting is beyond the scope of this work. We simply give some complementary results based on the rank-1 lattice rule, and leave the analysis as a future task.

The research field of lattice rules has a relatively short history, and new lattice rules are being developed continually. For more information, we refer the reader to the introductory piece by Kuo and Soan [22], and an example of lattice rules used in a financial setting [9]. Suitable lattice rules may offer a substantially faster way of evaluating the high dimensional integrals presented here and in previous work [18]. However, we focus on reducing error regardless of the chosen sequences by employing importance sampling, which has been done under the homogeneous model [18].

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

8.2 Importance sampling

To reduce the variance of the integral estimate we employ importance sampling, that is, suitably generate C_i values to sample the function in key regions. Based on previous work [18], we derive an importance sampling function which is similar in form to that of the homogeneous case. We recall the expansion of the Q function (4.22) in which we use only the first term, hence

$$Q_\gamma(u) \sim q_0(u-\gamma+r)^{3/2} + O(u-\gamma+r)^{5/2} \quad \text{as} \quad u \to \gamma - r, \tag{8.2}$$

where

$$q_0 = \frac{4}{3}\left[\frac{2r}{\gamma(\gamma-r)}\right]^{1/2}.$$

A change of variable $c = \gamma - u$ leads to the function

$$\tilde{Q}_\gamma(c) = q_0(r-c)^{3/2},$$

which leads to an approximate solution for the hop distribution

$$\tilde{F}_\gamma(c) = e^{-\lambda \tilde{Q}_\gamma(c)}.$$

The quantity $\tilde{F}_\gamma(0)$ is subtracted from the above expression and the result is divided by $\Delta\tilde{F}_\gamma = \tilde{F}_\gamma(c_{\max}) - \tilde{F}_\gamma(0)$, to obtain an importance sampling function

$$\widehat{F}_\gamma(c) = \frac{1}{\Delta\tilde{F}_\gamma}\left[e^{-\lambda\tilde{Q}_\gamma(c)} - \tilde{F}_\gamma(0)\right], \qquad 0 \leqslant c \leqslant c_{\max}, \tag{8.3}$$

where $c_{\max}$ is the largest hop value. The subtracting of the routing void term has negligible effect for sufficiently large λ. The corresponding derivative needs to integrate to one, hence the re-scaling step. The derivative

$$\widehat{f}_\gamma(c) = \frac{3\lambda}{2\Delta\tilde{F}_\gamma}(r-c)^{1/2}e^{-\lambda\tilde{Q}_\gamma(c)}, \tag{8.4}$$

exists on the same domain as the sampling function. The inverse of the sampling function is given by

$$\widehat{F}_\gamma^{-1}(t) = r - \left(\frac{-1}{\lambda a_1}\ln\left[t\Delta\tilde{F}_\gamma + \tilde{F}_\gamma(0)\right]\right)^{3/2}, \qquad 0 \leqslant t \leqslant r. \tag{8.5}$$

Substituting a random variable from a uniform distribution, say $T \sim U(0,r)$, into the inverse of the importance sampling function gives a random variable C adhering to the importance sampling distribution (8.3).

After each hop is generated, the functions (8.4) and (8.5) should be updated by setting γ to the current sink distance for each sample. This step was not necessary under the homogeneous model. However, under this inhomogeneous model each hop distribution varies more with respect to the sink distance. Furthermore, for high λ, the importance sampling step should improve as it is based on the independent model, and at high node density the intersection regions grow stochastically smaller [17, 18].

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

9 Simulation

We compare results from routing simulations to those from our stochastic model and calculations to demonstrate that the one-hop approximation (5.8) sufficiently captures the dependence. Given a source node sink distance ℓ, message relaying is simulated in a circular sensor field $\mathscr{C}_\ell \subset \mathbb{R}^2$ of radius ℓ with the sink located at the origin. The fact that messages only advance towards the sink under greedy routings implies that sensor field edges do not influence the message routing. The total number of nodes per simulation is a Poisson random variable with the parameter

$$\begin{aligned}\Lambda(\mathscr{C}_\ell) &= \lambda \int_0^\ell \int_0^{2\pi} q(u)u\,du\,d\theta \\ &= 2\lambda\pi\ell.\end{aligned}$$

In simulation, a node is assigned two independent random variables Θ_S and R_S corresponding to their polar coordinates in relation to the sink. To simulate node deployment such that the nodes adhere to the spatially dependent node density (3.2), both random variables are uniformly distributed

$$\begin{aligned}\mathbb{P}(\Theta_S \leqslant \theta) &= \frac{\theta}{\pi}, \qquad \theta \in [0,\pi], \\ \mathbb{P}(R_S \leqslant r) &= \frac{r}{\ell}, \qquad r \in [0,\ell].\end{aligned}$$

10 Numerical results

All the numerical integration and routing simulations were performed in Matlab on a standard machine. The built-in Halton sequence generator was employed for the quasi-Monte Carlo integration. For a given value of z, it took between 10^2 and 10^4 points to obtain a quasi-Monte Carlo estimate of the conditional distribution $\mathbb{P}(Z_n \leqslant z|+)$ where n ranged from 2 to 20. The actual number of points depends on the number of hops n; more hops require more points. The exact relationship between the required number of points and n is not known as the number of points also depends on λ, but future analytic empirical work may shed light on the relationship. The importance sampling step improved the rate of convergence, particularly for high λ. All calculations took no longer than an hour to complete, and usually considerably less.

We compared the dependent model to routing simulations of various ensemble sizes with and without a blinking sleep scheme. Generally, between 10^3 and 10^5 routing simulations were required. Similar to the integration process, higher hop numbers required more simulations to give converged results. The routing simulations are based on the same assumptions made in the mathematical model.

We observed under the inhomogeneous model that more function samples and routing simulations are needed to give similarly converged results compared to those obtained under the homogeneous model. An extensive empirical investigation is needed to see how fast quasi-Monte Carlo methods are compared to routing simulations. Also, more empirical and theoretical evidence is needed to elucidate the advantages and disadvantages of calculating probabilistic behaviour of greedy routing via our model.

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

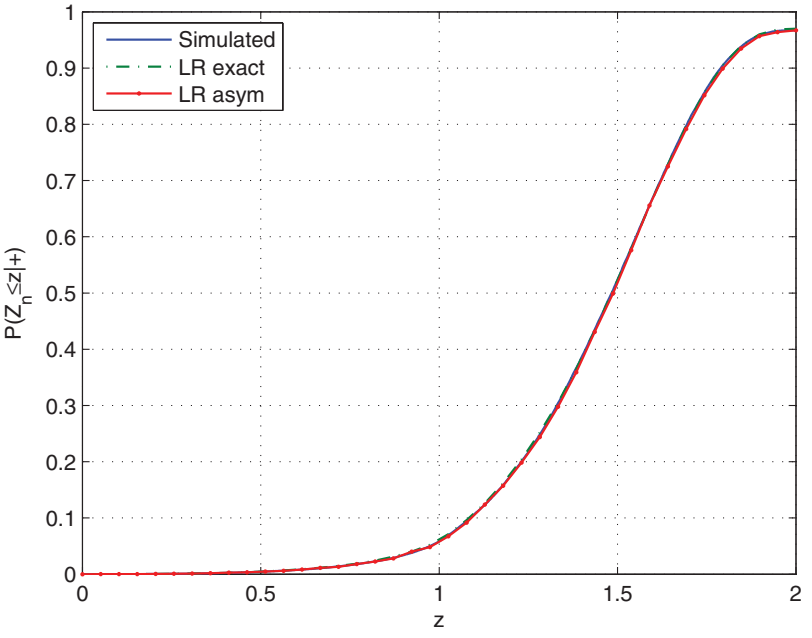


FIGURE 9. (Colour online) Results of $\mathbb{P}(Z_2 \leqslant z|+)$ via lattice rules integration (based on analytic and asymptotic expressions) and simulations ($p = 1$, $\lambda = 3\ell$, $r = 1$ and $\ell = 10$).

The lattice rule approach was only used to calculate one set of results (see Figure 9). We used generator vectors based on fixed lattice rules for 2^{10} function sample points and equal weights [21]. We applied ten random shifts to the lattice rules, (see Giles *et al.* [9] for an example). This approach generally performed well, however, under importance sampling sometimes erratic results arose. This is possibly due to an over bias in function sampling or an unknown numerical artefact. A more thorough examination is needed, but we believe that the preliminary results using lattice rules are promising.

We found that the results based on elliptic integral functions could be replaced with results that used the three-term approximations (4.18) and (4.22) with no discernible loss of accuracy (see the plots in Figure 9). Evaluating the approximations is faster as the expressions only involve elementary functions. Consequently, the remaining results are based on these approximations (Figures 10–12).

We calculated the distribution of N, and included a simple sleep scheme to see its influence on the path dependence. We compared the independent model to the dependent model by varying p (and accordingly α, hence holding λ constant). The difference between the two models has an accumulative effect on N, hence it serves as a good indicator of path dependence.

Under the independent model, the distribution $\mathbb{P}(N \leqslant n)$ gives greater values for each n than the equivalent result under the dependent model. Also, the path dependence clearly lessens as p approaches zero (compare Figures 10 and 11). These results are analogous to those under the homogeneous density model [18].

For large λ, the difference between the two models is less. As is the case for the homogeneous, a larger density results in the next forwarding node being closer to the sink,

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

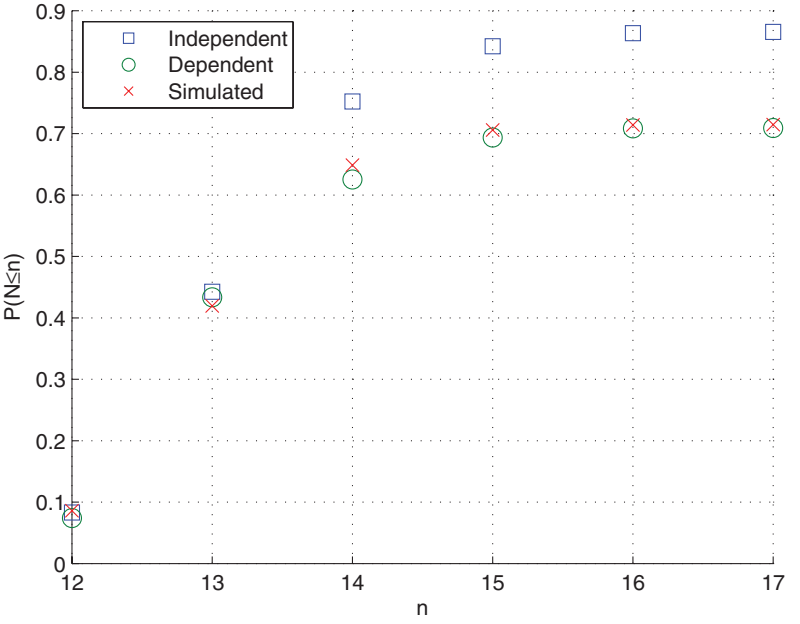


FIGURE 10. (Colour online) Independent and dependent model results of $\mathbb{P}(N \leqslant n)$ compared to simulations ($p = 1$, $\lambda = 2\ell$, $r = 1$ and $\ell = 10$).

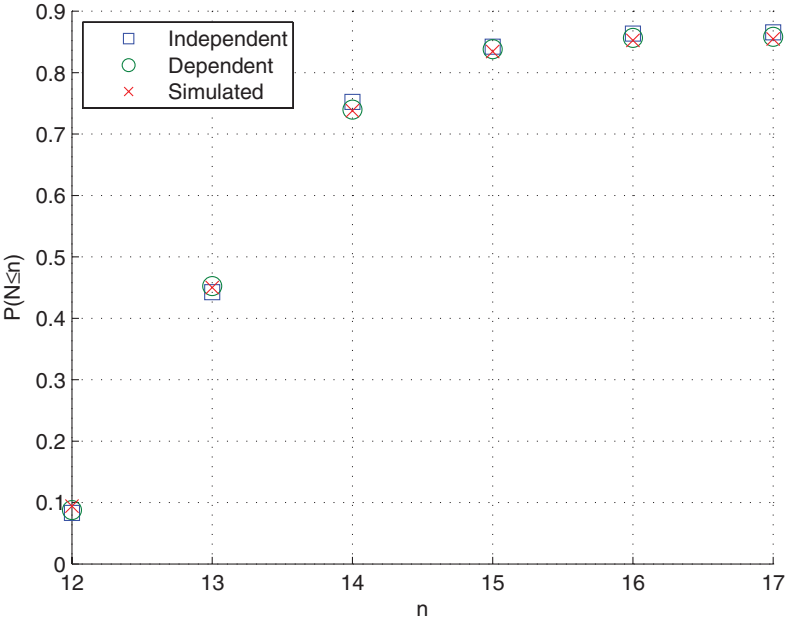


FIGURE 11. (Colour online) Independent and dependent model results of $\mathbb{P}(N \leqslant n)$ compared to simulations ($p = 0.1$, $\lambda = 2\ell$, $r = 1$ and $\ell = 10$).

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

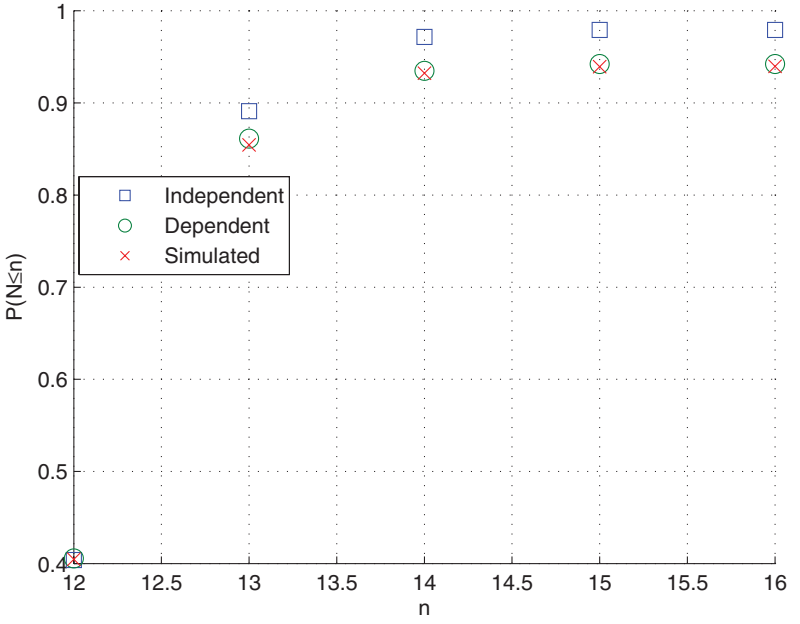


FIGURE 12. (Colour online) Independent and dependent model results of $\mathbb{P}(N \leqslant n)$ compared to simulations ($p = 1$, $\lambda = 3\ell$, $r = 1$ and $\ell = 10$).

thus reducing the intersection of the feasible regions and lessening the path dependence. Moreover, under the inhomogeneous model hops grow stochastically larger due to the increasing node density. Thus, we believe that the path dependence continues to decrease stochastically as the message approaches the sink.

In conclusion, the results reveal that the dependent model clearly captures the path dependence. The resulting expressions, both involving the elliptic integrals and the asymptotic expansions, give results that closely agree with simulations. We believe, more numerical investigation is needed to choose the most appropriate quasi-random sequences in this setting.

11 Future work

In realistic settings, the constant node density assumption may often not be appropriate. Under the spatially dependent model, the density function was chosen such that it was simple enough for obtaining analytic and asymptotic means, while still being a plausible node placement scenario. Other suggestions exist such as the node density decaying exponentially or according to some inverse power of the sink distance.

Furthermore, under our model the sink was located at the maximum of the node density. Placing the sink at an arbitrary point in the sensor field results in the node density being dependent on the sink angle. This is an additional increase in the complexity of the density function. Moreover, the angle of an individual node would not be a uniformly distributed random variable, thus importance sampling might be needed when integrating over the

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

angle domains. Consequently, these suggestions may result in analytic and asymptotic mean measures that can be used to model more realistic node deployment models.

Further investigation of the asymptotic approximations are needed. The approximation may break down when the radius is large compared to the sink distance. For a constant radius model, the lengths can always be re-scaled with respect to the transmission radius. However, this may not be possible for a randomly varying radius model. We stress that including random transmission radii into our model would be an interesting and realistic model extension in itself.

We presented some integration results based on lattice rules. Despite these methods giving mostly agreeable results, further investigation is needed to gauge which lattice rules and quasi-random sequences are the most suitable for the integrals that arise from our model. This investigation would be both analytic and numerical in nature. This work may lead to our mathematical model considerably outperforming regular routing simulations.

Finally, an attractive feature of regular Monte Carlo methods is that the error is obtained by estimating the variance of the integral. Conversely, quasi-Monte Carlo methods lack a practical way of estimating the error, despite them generally have a faster convergence rate. The idea of 'randomized' quasi-Monte Carlo methods seeks to combine the advantages of both approaches. Consequently, a future task lies in investigating these hybrid methods in evaluating hop integrals.

12 Conclusion

We presented a tractable inhomogeneous spatially dependent density model. Inspired by previous work, we developed and examined a greedy routing model that incorporates both sink and path dependence. Moreover, the spatially dependent density model verified that the formulation of the homogeneous model can be extended to an inhomogeneous case. This model is an alternative means of ascertaining the stochastic characteristics of greedy routing in sensor and ad hoc networks.

We used asymptotic methods to derive accurate approximations for hop length moments and the mean measure for our spatially dependent node density model. We used quasi-Monte Carlo methods and recently developed lattice rules coupled with importance sampling to estimate the resulting high dimensional integrals. For a sufficient number of function samples, all the results agreed admirably with those obtained by routing simulations.

Finally, we included a sleep scheme to demonstrate its effects on the local node density and the path dependence. For systems with a low p, the results imply that the independent model can be used, thus reducing computation time in calculating stochastic properties of the system.

Acknowledgements

The author thanks Peter G. Taylor at the University of Melbourne for valuable discussions and advice. A further thanks to Frances Kuo and Ian Sloan at the University of New South Wales for kindly explaining the subtleties of lattice rules. The author would also like to acknowledge the support of the Australian Research Council Centre of Excellence for Mathematics and Statistics of Complex Systems and the David Lachlan Hay Memorial Fund.

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

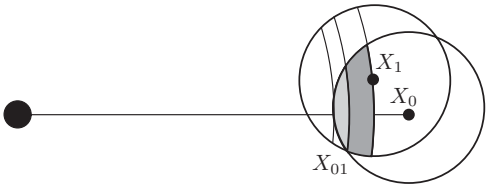


FIGURE A 1. The form of intersection region depends on the u_2 interval and the location of X_{01}.

Appendix A Derivation of $\vec{Q}_i(u_{i+1})$

We outline how to calculate the re-scaled mean measure for the feasible region under the dependent model. The method is akin to calculating the equivalent area function $\vec{A}_i(u_{i+1})$ in the homogeneous case [17]. In fact, the derivation of $\vec{Q}_i(u_{i+1})$ is included for completeness, and we refer the reader to previous work [17] for further details.

We use the function $\Delta\psi(u_2)$ again to describe the angular width of the intersection of the source and current feasible regions. Subsequently, the re-scaled mean measure on this intersection region is given by

$$Q_{1\setminus 0}(u_2) = \int_{\ell-r}^{u_2} \Delta\psi(w_2)dw_2,$$

which leads to the re-scaled mean under the dependent model

$$\vec{Q}_1(u_2) = Q_{u_1}(u_2) - Q_{1\setminus 0}(u_2).$$

It can be shown that the intersection angle expression is

$$\Delta\psi(u_2) = 2\psi_\ell(u_2)\mathbf{I}^-_{X_{01}}, \qquad \ell - r \leqslant u_2 \leqslant u_{01},$$

where $\mathbf{I}^-_{01}$ is an indicator function for when the intersection-point X_{01} is below the baseline that runs from X_0 to X_S (refer to Figure A 1), and u_{01} is the sink distance of X_{01}. On the second interval we obtain the intersection angle expression

$$\Delta\psi(u_2) = \psi_\ell(u_2) + \psi_{u_1}(u_2) - \theta_1, \qquad u_{01} \leqslant u_2 \leqslant u_1.$$

Recall under the independent model the re-scaled mean integral

$$Q_\gamma(u) = 2\int_{\gamma-r}^{u} \psi_\gamma(w)dw,$$

where the angle function

$$\psi_\gamma(u) = \arccos\left(\frac{u^2+\gamma^2-r^2}{2u\gamma}\right).$$

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

Thus, on the first interval, $[\ell - r, u_{01}]$, we have the re-scaled mean

$$\begin{aligned} Q_{1\setminus 0}(u_2) &= 2\int_{\ell-r}^{u_2} \psi_\ell(w_2)dw_2 \mathbf{I}^-_{X_{01}} \\ &= Q_\ell(u_2)\mathbf{I}^-_{X_{01}}. \end{aligned}$$

On the second interval, $[u_{01}, u_1]$, we have the slightly more complicated re-scaled mean expression

$$\begin{aligned} Q_{1\setminus 0}(u_2) &= \int_{\ell-r}^{u_2} [\psi_\ell(w_2) + \psi_{u_1}(w_2) - \theta_1]dw_2 + Q_\ell(u_{01})\mathbf{I}^-_{X_{01}}, \\ &= \frac{1}{2}\left(Q_\ell(u_2) + Q_{u_1}(u_2) + 2\theta_1\,[u_{01} - u_2]\right) \\ &\quad + \frac{1}{2}\left(Q_\ell(u_{01})[2\mathbf{I}^-_{X_{01}} - 1] - Q_{u_1}(u_{01})\right). \end{aligned}$$

This approach naturally extends to the re-scaled mean on the intersection of any two feasible regions. Consequently, for $i \geqslant 1$, under the one-hop dependent model the re-scaled mean measure on the feasible region is given by

$$\vec{Q}_i(u_{i+1}) = Q_{u_i}(u_{i+1}) - Q_{i\setminus i-1}(u_{i+1}).$$

Appendix B Joint probability density

We assume there is no sleep scheme, and note that to include a sleep scheme entails substituting the corresponding node density function (6.1) into the joint probability density.

We consider the probability density of Θ_i conditioned on the event $U_i = u_i$. Under our inhomogeneous Poisson model, the angle of any node is distributed uniformly around the sink (in the regions where nodes can exist). Hence, the conditional probability density under the dependent model

$$g_{u_1}(\theta_1|U_1 = u_1) = \frac{\mathbf{I}_{\mathscr{D}_0}(u_1, \theta_1)}{2\psi_{u_0}(u_1)}$$

follows, and this expression also applies to the independent model as it has the same feasible region. We introduce the function $\Psi_{\vec{x}_i}(u_{i+1})$ to denote the total angular width of the feasible region given $U_{i+1} = u_{i+1}$ and the path $\vec{X}_i = \vec{x}_i$. The conditional probability density

$$g_{u_{i+1}}(\theta_{i+1}|U_{i+1} = u_{i+1}) = \frac{\mathbf{I}_{\mathscr{D}_i}(u_{i+1}, \theta_{i+1})}{\Psi_{\vec{x}_i}(u_{i+1})}$$

follows. Under the independent model the angular width function simplifies to

$$\Psi_{\vec{x}_i}(u_{i+1}) = 2\psi_{u_i}(u_{i+1}).$$

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

The re-scaled mean of the feasible region written as an integral

$$\vec{Q}_i(u_{i+1}) = \int_{u_i - r}^{u_{i+1}} \Psi_{\vec{x}_i}(w_{i+1}) dw_{i+1}$$

gives the derivative of the re-scaled mean measure

$$\vec{Q}_i'(u_{i+1}) = \Psi_{\vec{x}_i}(u_{i+1}).$$

The probability density

$$g_{\vec{x}_i}(u_{i+1}) = \lambda \Psi_{\vec{x}_i}(u_{i+1}) e^{-\lambda \vec{Q}_i(u_{i+1})}$$

follows. Hence, the joint probability density of the two random variables U_{i+1} and Θ_{i+1} is given by

$$\begin{aligned} g_i(u_{i+1}, \theta_{i+1}) &= g_{\vec{x}_i}(u_{i+1}) g_{u_{i+1}}(\theta_{i+1} | U_{i+1} = u_{i+1}, \vec{X}_i = \vec{x}_i) \\ &= \lambda_{\mathscr{D}_i}(u_{i+1}, \theta_{i+1}) e^{-\lambda \vec{Q}_i(u_{i+1})}, \end{aligned}$$

where the spatially dependent node density function has been introduced

$$\lambda_{\mathscr{D}_i}(u_{i+1}, \theta_{i+1}) = \lambda \mathbf{I}_{\mathscr{D}_i}(u_{i+1}, \theta_{i+1}).$$

The joint probability density of the random variables U_1 to U_n and Θ_1 to Θ_n

$$g_{(n-1)}(u_1, ., u_n, \theta_1, ., \theta_n) = \prod_{i=1}^{n} \lambda_{\mathscr{D}_{i-1}}(u_i, \theta_i) e^{-\lambda \vec{Q}_{i-1}(u_i)} \tag{B 1}$$

follows, or in terms of hop advancements the equivalent expression

$$\bar{g}_{(n-1)}(c_1, ., c_n, \theta_1, ., \theta_n) = \prod_{i=1}^{n} \lambda_{\mathscr{D}_{i-1}}(u_{i-1} - c_i, \theta_i) e^{-\lambda \vec{Q}_{i-1}(u_{i-1} - c_i)}.$$

References

[1] *Digital Library of Mathematical Functions*, Accessed 15 January 2010. URL: http://dlmf.nist.gov/.

[2] Akyildiz, I. F., Weilian, S., Sankarasubramaniam, Y. & Cayirci, E. (2002) A survey on sensor networks. *IEEE Commun. Mag.* **40**, 102–114.

[3] Baccelli, F. & Blaszczyszyn, B. (2009) *Stochastic Geometry and Wireless Networks, Volume I - Theory, Vol. 1*, NOW Publishers, Delft, The Netherlands.

[4] Baccelli, F. & Blaszczyszyn, B. (2009) *Stochastic Geometry and Wireless Networks, Volume II - Theory, Vol. 1*, NOW Publishers, Delft, The Netherlands.

[5] Bose, P., Morin, P., Stojmenovic, I. & Urrutia, J. (1999) Routing with guaranteed delivery in ad hoc wireless networks. In: *3rd International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, ACM, pp. 48–55.

[6] Carlson, B. (1979) Computing elliptic integrals by duplication. *Numer. Math.* **33**, 1–16.

[7] Carlson, B. (1995) Numerical computation of real or complex elliptic integrals. *Numer. Algorithms*, **10**, 13–26.

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press

[8] Chong, C. & Kumar, S. P. (2003) Sensor networks: Evolution, opportunities and challenges. *Proc. IEEE* **91**, 1274–1256.
[9] Giles, M. B., Kuo, F. Y., Sloan, I. H. & Waterhouse, B. J. (Nov 2008) Quasi-monte carlo for finance applications. In: G. N. Mercer & A. J. Roberts (editors), *Proceedings of the 14th Biennial Computational Techniques and Applications Conference*, CTAC-2008, vol. 50 of ANZIAM Journal, pp. C308–C323.
[10] Haenggi, M., Andrews, J., Baccelli, F., Dousse, O. & Franceschetti, M. (2009) Stochastic geometry and random graphs for the analysis and design of wireless networks. *IEEE J. Sel. Areas Commun.* **27**, 1029–1046.
[11] Hall, P. (1988) *Intoduction to the Theory of Coverage Process*, 1st ed., John Wiley and Sons., Hoboken, NJ.
[12] Halton, J. (1960) On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numer. Math.* **2**, 84–90.
[13] Ishizuka, M. & Aida, M. (2004) Performance study of node placement in sensor networks. In: *Proceedings 24th International Conference on Distributed Computing Systems Workshops*, March 2004, pp. 598–603.
[14] Ishizuka, M. & Aida, M. (2007) Stochastic node placement improving fault tolerance in wireless sensor networks *Electron. Commun. Japan* **90**, 2181–2191.
[15] Karp, B. & Kung, H. T. (2000) Greedy perimeter stateless routing for wireless networks. In: *Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 2000)*, Los Angeles, California, pp. 243–254.
[16] Keeler, H. P. (2010) *Stochastic Routing Models in Sensor Networks*, PhD thesis, University of Melbourne.
[17] Keeler, H. P. & Taylor, P. G. (2010) A stochastic analysis of a greedy routing scheme in sensor networks. *SIAM J. Appl. Math.* **70**, 2214–2238.
[18] Keeler, H. P. & Taylor, P. G. (2011) A model framework for greedy routing in a sensor network with a stochastic power scheme. *ACM Trans. Sensor Netw.* **7**, 34:1–34:28.
[19] Kuhn, F., Wattenhofer, R., Zhang, Y. & Zollinger, A. (2003) Geometric ad-hoc routing of theory and practice. In: *22 nd Annual Symposium on Principles of Distributed Computing*, Boston, Massachusetts.
[20] Kullback, S. (1959) *Information Theory and Statistics*, 1st ed., Wiley, New York.
[21] Kuo, F. *Lattice rule generating vectors, http://web.maths.unsw.edu.au/∼fkuo/lattice/index.html.* Personal Website. Accessed 4 August 2009.
[22] Kuo, F. Y. & Sloan, I. H. (2005) Lifting the curse of dimensionality. *Not. AMS* **52**, 1320–1328.
[23] Mauve, M., Widmer, A. & Hartenstein, H. (2001) A survey on position-based routing in mobile ad hoc networks. *IEEE Netw.* **15**, 30–39.
[24] Niederreiter, H. (1992) *Ŕandom Number Generation and Quasi-Monte Carlo Methods*, SIAM, pp. 13–18.
[25] Pallavi, M., Ram, S. S. & Manjunath, D. (2009) Path coverage by a sensor field: The nonhomogeneous case. *ACM Trans. Sensor Netw.* **5** , 1–26.
[26] Sobol, I. (1967) The distribution of points in a cube and the approximate evaluation of integrals. *U.S.S.R. Comput. Math., and Math. Phys.* **7**, 86–112. (In Russian).
[27] Spanier, J. & Maize, E. H. (1994) Quasi-random methods for estimating integrals using relatively small samples. *SIAM Rev.* **36**, 18–44.
[28] Stojmenovic, I. (2002) Position-based routing in ad hoc networks. *IEEE Commun. Mag.* **40**, 128–134.
[29] Stoyan, D., Kendall, W. & Mecke, J. (1995) *Stochastic Geometry and Its Applications*, 2nd ed., Wiley, New York.
[30] Tubaishat, M. & Madria, S. (2003) Sensor networks: An overview. *IEEE Potentials* **22**, 20–23.
[31] Wong, R. (1989) *Asymptotic Approximations to Integrals*, Academic Press, New York.
[32] Zorzi, M. & Rao, R. R. (2003) Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Multihop performance. *IEEE Trans. Mobile Comput.* **2**, 349–365.

https://doi.org/10.1017/S095679251200006X Published online by Cambridge University Press