Random Walk Models for Geographic Data Propagation in Wireless Sensor Networks

Silvija Kokalj-Filipović, Roy Yates, and Predrag Spasojević
WINLAB, ECE Department, Rutgers University, {*skokalj,ryates,spasojev*}*@winlab.rutgers.edu*

***Abstract*— For wireless sensor networks with many location-unaware nodes, we investigate a protocol, dubbed BeSpoken, that steers data transmissions along a straight path called a spoke. The protocol directs data transmissions by randomly selecting relays to retransmit data packets from crescent-shaped areas along the spoke axis. The packet retransmission by the selected relay constitutes a spoke hop. The current spoke direction at any given hop may deviate from the spoke axis. In this work, we model the current direction as a Markov Modulated Random Walk, and offer protocol design guidelines which ensure that the deviation is limited.**

Keywords: Wireless sensor networks, geographic information dissemination, Markov modulated random walk, Large Deviation Principle

I. Introduction

In wireless sensor networks, events are observed by data source nodes. A wireless dissemination protocol uses a sequence of wireless transmissions that propagate the information from the source to the sinks. Frequently, the data source is unaware of which sink nodes wish to learn of its observations. If the source is also unaware of its position, as the nodes in random sensor networks are, it is unable to send complete information to the sinks, in terms of both the nature of an event and its location. This problem is typical of sensor networks, due to their assumed simplicity and scarce energy resources. In particular, GPS, the most widely used positioning system, is a significant power consumer, and infeasible in environments with clear-sky impediments. Hence, data source localization would be an attractive complementary feature of a wireless dissemination protocol.

Flooding is the simplest dissemination strategy, nevertheless it typically yields excessive communication [1] and can lead to a "broadcast storm" of redundant transmissions [7], unless a mitigating technique is employed [11]. The idea of data source localization enabled by the dissemination protocol properties is absent from flooding-based dissemination techniques. Several papers consider spatial properties of the dissemination route. Different forms of spatially constrained random walk are discussed in [2], [4], [10], while the idea of a trajectory-based dissemination is presented in [5], [8]. None of these dissemination approaches enable unknown source localization.

Motivated by the radial symmetry of isotropic wireless transmission, we propose a dissemination protocol, dubbed BeSpoken, that conveys information about an event to the sinks in an energy efficient manner, and whose spatial properties can simplify the source-location estimation. This protocol generates relatively straight-line trajectories called spokes, without requiring the nodes to have any navigational information. We envision simple sources while sinks are likely to be more capable in terms of *direction-of-arrival (DoA)* estimation, and in being location-aware and application-cognizant. We examine networks in which the sinks are distributed uniformly along the perimeter of the network area. Since the exact positions of interested sinks are not known, we conjecture that the likelihood of successful data search would increase if both

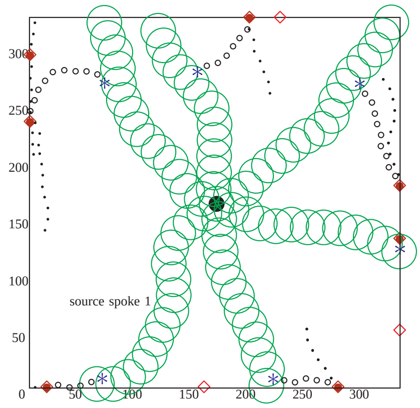


Fig. 1. The meaning of the name BeSpoken is twofold: the radial lines extending from the source form a pattern that resembles spokes of a wheel and, furthermore, spoke relays bespeak the source message. In this simulation snapshot, source spokes are shown as a sequence of relay transmission *ranges*, to illustrate the fact that each spoke is an ensemble of possible data routes. The sequence of wireless transmission *relays* forming a productive sink spoke is denoted by tiny circles(see the boxed spoke). Unproductive sink spokes are represented by dots. The search success is marked by a $*$.

the source and the sinks spawn several equally spaced radial spokes. As illustrated in Figure 1, a source disseminates data advertisements along the source spokes, and a sink sends a query along its spokes that may intersect the source spokes. Each intersection represents a successful search. The first sink spoke to reach one of the source spokes is called *productive.* Successful search is to be followed by the reinforcement of a route along the intersecting spokes and subsequent data dissemination. A GPS and DoA-enabled sink can determine the positions of the nodes along the productive sink spoke, and let them know of their positions. The other part of the enforced route can be learned based on the known position of the intersection nodes, provided that the direction of the source spoke is known. This can be achieved in a variety of ways, such as joint DoA estimation performed by the nodes surrounding the intersection, or by a mobile DoA-enabled sink which can move to the intersection and determine source spoke direction by polling the spoke nodes. We consider BeSpoken as both a dissemination protocol, and a tool to build an infrastructure of relatively straight paths (spokes) whose direction and length can be learned with moderate effort. A description on how the spokes would tessellate the sensor network space, and the implications and possible applications of such an infrastructure can be found in the companion paper [6]. The focus of this paper is the mathematical model of the BeSpoken, with an emphasis on the Markov-modulated random walk model of a spoke's direction.

II. System Model and Problem Statement

We consider a dense wireless network with a uniform spatial distribution of nodes. The *BeSpoken* protocol organizes a sequence of fixed-power *relay* transmissions that propagate the source message hop-by-hop, without positional or directional information. The hop relays form a *spoke* which may deviate from the radial *spoke axis*. Each spoke hop is organized using a sequence of two control message transmissions followed by the hop data transmission. We define the transmission range

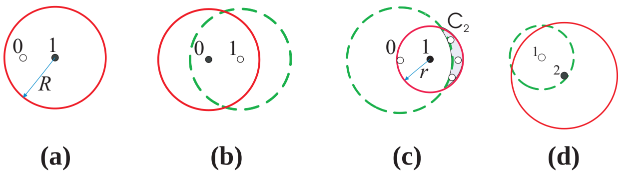


Fig. 2. BeSpoken Protocol: At each protocol stage, the current transmission range is denoted with the full circle while the previous range is denoted with a dashed circle.

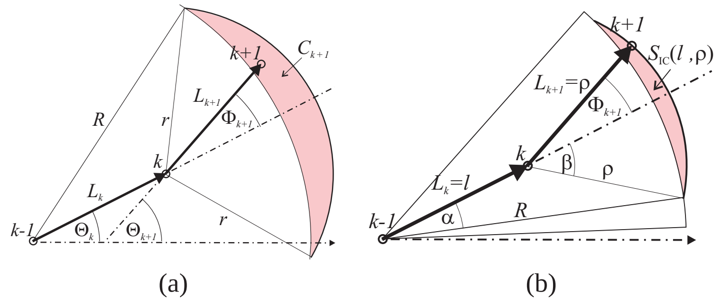


Fig. 3. **(a)** At hop $k+1$, node $k+1$ is distance L_{k+1} from node k and the current spoke direction is $\Theta_{k+1} = \Theta_k + \Phi_{k+1}$. **(b)** Given $L_k = l$ and $L_{k+1} = \rho$, the angular hop displacement Φ_{k+1} is constrained to the interval $-\beta \leq \Phi_{k+1} \leq \beta$ where the maximum angular displacement at hop $k+1$ is $\beta = \beta(l, \rho)$. The shaded area denotes the interior crescent of area $S_{\rm IC}(l, \rho)$.

as the maximum distance from the source at which nodes can reliably receive a packet. We assume that the physical layer modulation and coding are designed to compensate for short-scale fading effects and, thus, our transmit power requirements depend only on distance-dependent propagation path loss. Even though in a sensor network environment data rates are low relative to the available bandwidth and interference is not a primary issue, still, our protocol mitigates the interference as it always selects only one node to retransmit. Assuming radially symmetric attenuation (isotropic propagation), the area in which the transmitted packet is reliably received is a disk of a given radius. We use the same transmission power for both data and control packets, but different coding rate and/or modulation format, so that the communication rate for control messages is lower and translates to a longer range.

A. BeSpoken Protocol

The BeSpoken protocol implements a recursive process illustrated in Figure 2 in the following way:

- **(a)** The leading relay (node 1) sends an RTS (request to send) control packet with range $R = rq$ where $q = 2 - \epsilon$, for small ϵ.
- **(b)** The pivot (node 0) sends a BTS (block to send) control packet with range R.
- **(c)** The leading relay transmits the data packet with range r and becomes the new pivot. The region in which nodes receive this data packet but do not receive the preceding BTS packet forms the *1-st hop crescent* C_2.
- **(d)** A random node from the crescent C_2 becomes the new leading relay by transmitting a new RTS. The process returns to (a) with node 1 as the pivot and node 2 as the leading relay.

This recursive process is initialized by assigning the role of the pivot to the source node which transmits the data packet with a range r. The first node which receives the data packet and gets access to the medium becomes the first leading relay. The underlying ALOHA-type Carrier Sense Multiple Access protocol would resolve any collisions; hence, after a possible additional delay, only one random node from the crescent would transmit the RTS packet.

B. Problem Formulation

To describe the effects of the data and control ranges r and R, we evaluate the spoke behavior with respect to the constraints:

- **Outage:** the probability that a spoke dies before reaching a distance d is small,
- **Wobbliness:** the deviation of the instantaneous spoke direction with respect to the spoke axis is within defined limits.

The vector from node 0 to node 1 in Figure 2 defines the spoke axis. The crescent subtending angle determines how much the spoke may deviate from the spoke axis direction. The parameter $q = R/r$ determines the maximum crescent subtending angle. A large subtending angle fosters wobbleness, yet it implies a larger crescent, which increases chances that a relay will be found to retransmit data. Fixing q to a small value that limits wobbleness requires increasing r, to generate large enough crescent and decrease outage probability. Note that the energy per hop grows as r^α, where $\alpha \geq 2$ is the propagation loss coefficient, so that the total energy per spoke grows as $dr^{\alpha-1}$. Thus, minimizing the transmission range r corresponds to a minimum energy objective.

These contending influences illustrate the importance of the protocol parameters design. In this paper we develop closed-form expressions that serve as bounds for the values of q, ensuring that the wobbliness constraint is satisfied. In our companion paper [6], we show that outage and wobbliness constraints can be decoupled. We also show that satisfying the wobbliness constraint requires that one finds the minimum q so that the spoke direction is within the limits after η hops, where η is a sufficient number of hops to reach the target distance d.

III. Spoke Modeling

In this section we develop an analytical model of the spoke.

A. BeSpoken Geometry

Figure 3(a) depicts hops k and $k+1$. At the completion of hop k, the length L_k denotes the *current hop length* and the angle Θ_k denotes the *current spoke direction*.

From Figure 3(b) we observe that given $L_k = l$ and $L_{k+1} = \rho$ the control circle of radius R centered at node $k-1$ and the circle of radius ρ centered at node k specify a radius ρ arc for the possible positions of node $k+1$. The endpoints of this radius ρ arc constrain the *angular hop displacement* Φ_{k+1} to the interval $-\beta \leq \Phi_{k+1} \leq \beta$ where the maximum angular displacement is $\beta = \beta(l, \rho)$. Applying the law of cosines to the complementary angle $\pi - \beta(l, \rho)$ yields

$$\cos\beta(l, \rho) = \frac{R^2 - \rho^2 - l^2}{2l\rho}. \qquad (1)$$

We also observe that the region between the radius R control circle and the radius ρ arc defines an *interior crescent*, shown as the shaded area in Figure 3(b). From geometric arguments, it can be verified that the area of this interior crescent is

$$S_{\mathrm{IC}}(l,\rho) = 2\rho^2\beta(l,\rho) - 2R^2\alpha(l,\rho) + Rl\sin\alpha(l,\rho) \qquad (2)$$

where $\alpha(l,\rho)$ is found from the law of cosines to satisfy $\cos\alpha(l,\rho) = (R^2 - \rho^2 + l^2)/(2lR)$.

L_{k+1} can vary from a minimum value of $R - L_k$ to a maximum value of r. The induced interior crescent C_{k+1} in Figure 3(a) has an area $S_c(L_k) = S_{\mathrm{IC}}(L_k, r)$. We note that C_{k+1}, termed the *current crescent*, is the set of all possible positions of the node $k+1$.

B. Markov Process Model for Hop Length Evolution

For design purposes we assume that the spatial distribution of network nodes is a planar Poisson point process of intensity $\lambda = 1$. Thus, a current crescent yields a candidate set for node $k+1$ with cardinality Z_k that is, conditionally, a Poisson random variable with conditional expected value

$$E[Z_k|L_k = l_k] = S_c(l_k). \qquad (3)$$

The fact that the current hop length defines the expected size of the candidate set for the next relay illustrates the Markovian character of the hop length evolution. In particular, a small L_k will create a small crescent; this induces a support set $[R - L_k, r]$ for L_{k+1}, that excludes small hop lengths in the interval $[R - r, R - L_k)$.

The companion paper [6] shows that the analysis of the outage and wobbliness constraints can be decoupled by formally defining $\{L_k\}$ as a fictitious process that never encounters an empty crescent. Under the fictitious process model, the position of node $k+1$ will be uniformly distributed over the crescent C_{k+1}. From Figure 3 we see that, given the current hop length $L_k = l_k$, the arc of radius ρ has length $2\rho\beta(l_k,\rho)$. The conditional probability that we find node $k+1$ in the annular segment of width $d\rho$ along the arc of radius ρ is $2\rho\beta(l_k,\rho)d\rho/S_c(l_k)$. It follows that the conditional pdf of the next hop length L_{k+1} given $L_k = l_k$ is

$$f_{L_{k+1}|L_k}(\rho|l_k) = \frac{2\rho\beta(l_k,\rho)}{S_c(l_k)} \qquad R - l_k \le \rho \le r, \qquad (4)$$

and zero otherwise. We note that (4) provides a complete characterization of the fictitious process $\{L_k\}$, representing a Markov process model for the evolution of hop length.

C. Finite State Ergodic Markov Chain Model

Here, we develop a Markov Chain model that approximates the Markov process described above. We start by quantizing the L_k process, yielding the m-state Markov chain $\hat{L}_k$. We first select a chain state set that quantizes the process state space $[R-r, r]$, then describe a mapping from the process state space to the chain state set and, last, describe the resulting chain probability transition matrix. We define $\{h_1, \ldots, h_m\} \subseteq [R-r, r]$ to be the chain state set. Without loss of generality, we assume that $h_0 = R - r < h_1 < h_2 < \ldots < h_m = r$. As illustrated in Figure 4, whenever the kth hop Markov chain state is $\hat{L}_k = h_i$, the corresponding next process hop length is $L_{k+1} \in \mathcal{I}_i = [R - h_i, r]$, where $\mathcal{I}_i$ is the *next hop span* and its length $|\mathcal{I}_i|$ is also the width of the corresponding quantized crescent $\hat{C}_k$ of area $c_i = S_c(h_i)$. L_{k+1} is quantized to state h_j whenever $\hat{L}_{k+1} \in \mathcal{I}_{ij}$ where

$$\mathcal{I}_{ij} = \mathcal{I}_i \cap (h_{j-1}, h_j]. \qquad (5)$$

Note that the set $\{\mathcal{I}_{ij} : j = 1, \ldots, m\}$ partitions I_i and serves as a set of quantization intervals for L_{k+1} when $\hat{L}_k = h_i$. This quantization mapping is illustrated in Figure 4 where $L_{k+1} \in \mathcal{I}_{42}$ is extended to reach the quantized node position marked with a grey circle at $\hat{L}_{k+1} = h_2$. The chain proceeds by declaring a fictitious node at the quantized position as the new leading relay. As depicted in Figure 4, a quantization interval $\mathcal{I}_{ij}$ corresponds to the strip of area

$$d_{ij} = \begin{cases} \int_{R-h_i}^{h_j} 2\rho\beta(h_i,\rho)\,d\rho, & j = j^*(i), \\ \int_{h_j-\Delta}^{h_j} 2\rho\beta(h_i,\rho)\,d\rho, & j > j^*(i), \end{cases} \qquad (6)$$

(and zero otherwise), and of width $|\mathcal{I}_{ij}|$ within the crescent $\hat{C}_k$ of area $c_i = \sum_j d_{ij}$. Here $j^*(i) = \min\{j : h_j > R - h_i\}$ is the index of the leftmost non-empty quantization interval within I_i.

As shown in Figure 4, $c_{ij} = S_{\mathrm{IC}}(h_i, h_j)$ is the *quantized interior crescent area* formed by the control circle (of radius R) centered at the kth hop relay and a circle of radius h_j centered at node $k+1$ at distance $\hat{L}_k = h_i$. Note that $c_{ij} < c_{i(j+1)} \cdots < c_{im}$, where $c_{ij} = 0$ for $j < j^*(i)$, $c_{im} = c_i$, and $d_{ij} = c_{ij} - c_{i(j-1)}$. The hop-length transition probabilities

$$\begin{aligned} P_{ij} &= \Pr\{\hat{L}_{k+1} = h_j | \hat{L}_k = h_i\} \\ &= \Pr\{L_{k+1} \in \mathcal{I}_{ij} | L_k = h_i\} = d_{ij}/c_i \end{aligned} \qquad (7)$$

follow from the uniformity of Poisson spatial distribution of nodes and since the fictitious process assumes that the crescent $\hat{C}_k$ is not empty. Intuitively, when m is sufficiently large, the ergodic Markov chain will approximate well the ergodic Markov process. We consider Markov chain models with both uniform and non-uniform quantization of $[R-r, r]$. In this paper we focus on the uniform quantization model for the purpose of modeling the current spoke direction.

D. Spoke Direction Process

Figure 3 (a) indicates that the angular hop displacement Φ_{k+1} at hop $k+1$ changes the current spoke direction in that

$$\Theta_{k+1} = \Theta_k + \Phi_{k+1} = \sum_{i=1}^{k+1} \Phi_i. \qquad (8)$$

We observe that all points along the radius ρ arc in Figure 3 (b) are equiprobable locations for node $k+1$. Thus, given the sequence $\{L_k\}$, the angular hop displacements $\{\Phi_k\}$ form a sequence of conditionally independent uniform random variables with the conditional pdf

$$f_{\Phi_{k+1}|L_k,L_{k+1}}(\phi|l_k, l_{k+1}) = \frac{1}{2\beta(l_k, l_{k+1})}, \qquad (9)$$

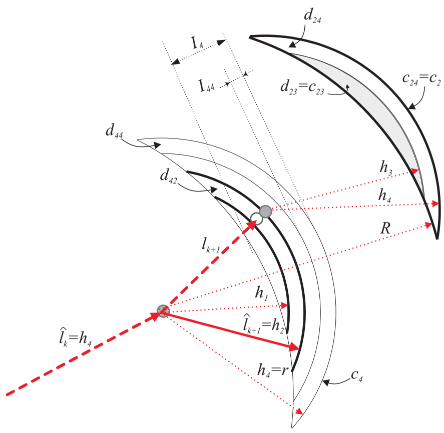


Fig. 4. Ergodic Finite State Markov Chain: quantization example for a four-state chain ($m = 4$): $\hat{L}_k = h_4 = r$ results in the first crescent $\hat{C}_k$ of area c_4 partitioned into four strips of total area $c_4 = d_{41} + d_{42} + d_{43} + d_{44}$; $L_{k+1} \in \mathcal{I}_{42}$, quantized to $\hat{L}_{k+1} = h_2$, is followed by a crescent $\hat{C}_{k+1}$ of area c_2 and a hop span $I_2 = [R - h_2, r]$ which is (uniformly) quantized into a crescent of area $d_{23} = c_{23}$ (shaded region) and a crescent strip $d_{24} = c_2 - c_{23}$ (the unshaded area).

for $|\phi| \leq \beta(l_k, l_{k+1})$, and zero otherwise. This probability distribution does not change when the conditioning sequence contains quantized values $\{\hat{L}_k\}$. The current angle sequence $\{\Theta_k\}$ is a random walk process modulated by the Markov chain $\{\hat{L}_k\}$, completely described by equations (7)-(9).

The transform domain analysis of Markov Modulated Random Walks (*MMRW*) [9] dictates that we first define the conditional moment generating functions of the incremental angular displacement Φ_{k+1} from (8)

$$g_{ij}(\omega) = E\left[\exp\left(\Phi_{k+1}\omega\right) | \hat{L}_k = h_i, \hat{L}_{k+1} = h_j\right] \quad (10)$$

$$= \frac{1}{2\varphi_{ij}} \int_{-\varphi_{i,j}}^{\varphi_{i,j}} \exp\left(\phi\omega\right) d\phi = \hbar\left(\varphi_{ij}\omega\right), \quad (11)$$

for ω in a convergence region $(\omega_{\text{-}}, \omega_+)$, where $\hbar(x) = \frac{\sinh x}{x}$ and $\varphi_{ij} = \beta(h_i, h_j)$. We create a matrix $\mathbf{\Gamma}(\omega)$ with elements

$$\Gamma_{ij}(\omega) = P_{ij} g_{ij}(\omega). \quad (12)$$

The Perron-Frobenius theorem (see e.g., [3]) dictates that its largest eigenvalue $\sigma(\omega)$ is real and positive. The elements of the corresponding right eigenvector $\nu(\omega) = [\nu_1(\omega) \cdots \nu_m(\omega)]^T$ are also real and positive. Next, we define the product martingale [9]

$$M_k(\omega) = \frac{\exp\left(\omega\, \Theta_k\right) \nu_{i(k)}(\omega)}{\sigma^k(\omega)\, \nu_{i(0)}(\omega)} \quad (13)$$

where $i(k)$ is the random state index of the chain at time k, and the random variable $\nu_{i(k)}(\omega)$ is the $i(k)$-th element of the right eigenvector.

IV. Wobbliness Constraint Model

The spoke goes off-course at hop k whenever the current angle Θ_k in (8) exceeds one of the following two thresholds ϕ_o and $-\phi_o$. To describe spoke wobbliness, we define

$$T_{\varphi_o} = \min\left\{k : |\Theta_k| \geq \varphi_o\right\}. \quad (14)$$

to be the first time that the spoke goes off-course. As we model the angle process evolution only up to that point, T_{φ_o} is the *stopping time* of the random walk Θ_k modulated by the ergodic Markov chain $\hat{L}_k$. Following [9, Chapter 7.7], T_{φ_o} is also a stopping rule for the martingale $M_k(\omega)$ relative to the joint process $\{M_k(\omega), L_k;\}$. Hence, following [9, Lemma 6] and the *optional sampling theorem* [9, Theorem 6] we have

$$E\left[M_{T_{\varphi_o}}(\omega)\right] = E\left[\frac{\exp\left(\omega\Theta_{T_{\varphi_o}}\right)\nu_{i(T_{\varphi_o})}(\omega)}{\sigma(\omega)^{T_{\varphi_o}}\nu_{i(0)}(\omega)}\right] = 1, \quad (15)$$

for $\omega \in (\omega_{\text{-}}, \omega_+)$. Since the stopping time T_{φ_o} is a random variable of unknown probability distribution, elaborate mathematical methods must be used to model it. Our methods utilize (15), which is an extension of the *Wald identity* to Markov modulated random walks. First, the wobbliness constraint is based on the first moment of T_{φ_o}, as

$$E\left[T_{\varphi_o}\right] \geq \eta. \quad (16)$$

Second wobbliness constraint is based on the *cumulative distribution function (CDF)* of T_{φ_o}, as follows

$$\Pr\left\{T_{\varphi_o} \leq \eta\right\} \leq p_t. \quad (17)$$

In subsection IV-A we demonstrate how to compute the mean $E\left[T_{\varphi_o}\right]$. Subsection IV-B describes a bound on the CDF of the stopping time. These two approaches together provide a good description of the stopping time, based on which a range of q values can be found for each φ_o.

A. Expected Threshold Crossing Time

The random variable $\Theta_{T_{\varphi_o}}$ is either $-\varphi_o$ or φ_o, assuming that there is no overshoot. We address the problem of overshoot later. By symmetry arguments, first and second moments of $\Theta_{T_{\varphi_o}}$ are

$$E\left[\Theta_{T_{\varphi_o}}\right] = 0, \quad \text{var}\left[\Theta_{T_{\varphi_o}}\right] = E\left[{\Theta_{T_{\varphi_o}}}^2\right] = \varphi_o^2. \quad (18)$$

We evaluate the second derivative of (15) with respect to ω at $\omega = 0$, and denote $\mu_i(\omega) = \nu_i''(\omega)/\nu_i(\omega)$, to obtain the expected number of hops until the hop angle hits the threshold as

$$E\left[T_{\varphi_o}\right] = \frac{\text{var}\left[\Theta_{T_{\varphi_o}}\right] + E\left[\mu_{i(T_{\varphi_o})}(\omega)\right]\Big|_{\omega=0} - \mu_{i(0)}(\omega)\Big|_{\omega=0}}{\frac{\sigma''(\omega)}{\sigma(\omega)}\Big|_{\omega=0}}. \quad (19)$$

One can show that, for $m = 2$, the denominator $\frac{\sigma''(\omega)}{\sigma(\omega)}\Big|_{\omega=0} =$

$$\frac{1}{3}\left(\pi_2 P_{22}\varphi_{22}^2 + \pi_1 P_{11}\varphi_{11}^2 + \pi_1 P_{12}\varphi_{12}^2 + \pi_2 P_{21}\varphi_{21}^2\right), \quad (20)$$

where $\pi_i,\ i = 1, 2$ are the elements of the vector of stationary state probabilities $\overline{\pi} = [\pi_i]_{(1\times m)}$. Note that terms $\varphi_{ij}^2/3$ are transition-specific variances. Direct generalization of (20) to an m state model has a form of a stationary average of transition-specific variances over m^2 transitions

$$\frac{\sigma''(\omega)}{\sigma(\omega)}\Bigg|_{\omega=0} = \text{var}\left[\theta\right]_p = \overline{\pi} P^{(v)} u^T,$$

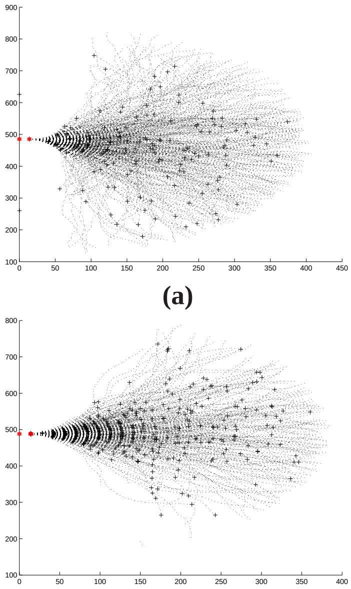
where, $P^{(v)} = \left[P_{ij}^{(v)}\right]_{(m\times m)}$ with elements $P_{ij}^{(v)} = \left(P_{ij}\varphi_{ij}^2\right)/3$, and $u = [1...1]_{(1\times m)}$. It can be shown that, for small crescent subtending angles relative to the threshold φ_o, we can ignore the terms $E\left[\mu_{i(T_{\varphi_o})}\right]$ and $\mu_{i(0)}$ in (19), thus

$$E\left[T_{\varphi_o}\right] = \frac{\text{var}\left[\Theta_{T_{\varphi_o}}\right]}{\text{var}\left[\theta\right]_p}. \tag{21}$$

Since (19) neglects the overshoot, we now seek to include the overshoot impact. We start with the overshoot analysis of the simple random walk $\Theta_n = \sum_{i=1}^{n} \Phi_i$, modulated by one-state Markov Chain, i.e. $\Phi_i \sim U\left(-\varphi_{11}, \varphi_{11}\right)$. Based on the derivation presented in the appendix, which assumes that undershoot and overshoot have the same uniform distribution, we obtain the overshoot-inclusive form of the numerator of (21) for a one-state MMRW

$$\text{var}\left[\Theta_{T_{\varphi_o}}\right] = \varphi_o^2 + (2/3)\varphi_o\varphi_{11} + \varphi_{11}^2/6. \tag{22}$$

Note that the second term in (22) contains the half-span φ_{11} of the uniform pdf. To extend the expression (22) to m-state Markov-modulated random walk, we replace φ_{11} with a weighted sum of transition-specific angle spans $\sum_{i,j=1}^{n} w_{ij}\varphi_{ij}$, where $w_{ij} = \pi_i P_{ij}$. Hence, the angle span associated with the trivial transition of the one-state MC is now replaced with a stationary average over angle-spans associated with m^2 transitions of the m-state MMRW. Note that the third term of (22) is one half of the angle variance for $\Phi_i \sim U\left(-\varphi_{11}, \varphi_{11}\right)$. For the m-state MMRW, we replace this term with another weighted sum where $(1/2)w_{ij}$-weighted terms are transition specific variances $\varphi_{ij}^2/3$. Hence, extended (22), in matrix notation, is



(a)

(b)

Fig. 5. Sample of spokes directed eastward - "constraint in the mean" vs."probability constraint" design.

$$\text{var}\left[\Theta_{T_{\varphi_o}}\right] = \varphi_o^2 + 2\varphi_o\left(\overline{\pi}P^{(a)}u^T\right) + 1/2\left(\overline{\pi}P^{(v)}u^T\right), \tag{23}$$

$P^{(a)} = \left[P_{ij}^{(a)}\right]_{(m\times m)}$ and $P_{ij}^{(a)} = P_{ij}\varphi_{ij}/3$. The overshoot-inclusive variant of (21) for a multi-state chain (23) yields values that match the simulation results closely and consistently. Note that (21) expresses the expected stopping time as a function of q only. Figure 5 (a) illustrates the achieved wobbliness in a sample of 500 spokes directed eastward, designed to propagate 160 length units with the wobble threshold of $\pi/4$. It is evident that a large number of spokes exceed the targeted propagation distance, while the straightness needs to be improved. Such a behavior is due to the fact that the outage constraint is a constraint in probability, while (16) is a constraint in the mean, where the pertinent pdf is long-tailed.

B. Probability of threshold crossing before time T_{φ_o}

Motivated by the observations illustrated by Figure 5 (a), we here analyze the wobbliness model, as defined in (17), from the point of view of *Large Deviation Theory (LDT)*. We determine a bound for $\Pr\{T_{\varphi_o} \le \eta\}$ based on the Gärtner-Ellis theorem [3, Thm 2.3.6] and its application to an empirical measure of finite Markov Chains, in particular [3, Exercise 3.1.4]. Let $\wp_{i(0)}^P$ denote the Markov probability measure associated with the transition probability matrix (7), and with the initial state $\hat{L}_0 = i(0)$. Precisely, $\wp_{i(0)}^P\left(\hat{L}_1 = y_1, \cdots, \hat{L}_n = y_n\right) = P_{i(0)y_1}\prod_{i=1}^{n-1} P_{y_iy_{i+1}}$ is the probability of a specific Markov chain path, starting at $i(0)$, and transitioning through the sequence of states $\{y_i\}_{i=1}^n$. Now, let us denote $\psi_{ij} = U\left(\varphi_{ij}, \varphi_{ij}\right)$, and thus, the conditional law of $\{\Phi_k\}$ for each realization $\{L_k = y_k\}_{k=1}^n$ is $\prod_{i=1}^n \psi_{y_{k-1}y_k}$. Denoting with $E_{i(0)}^P[.]$ the expected value with respect to $\wp_{i(0)}^P$ and the associated $\prod_{i=1}^n \psi_{y_{k-1}y_k}$, we further define $\Lambda_n(n\omega) = \log E_{i(0)}^P\left[e^{\omega\sum_{k=1}^n \Phi_k}\right]$. Following a derivation analogous to [3, Thm 3.1.2], we find that the logarithmic moment generating function of the current angle is related to the largest eigenvalue $\sigma(\omega)$ of (12) as $\Lambda(\omega) \triangleq \lim_{n\to\infty}\frac{1}{n}\Lambda_n(n\omega) = \log\sigma\left(\omega\right)$. According to [3, Thm 3.1.2], the empirical mean of the sum of angle deviations modulated by $\wp_{i(0)}^P$ has a rate function, which is a conjugate function of $\Lambda(\omega)$, i.e the Fenchel-Legendre transform $\Lambda^\star(x) = \sup_\omega\left\{\omega x - \Lambda(\omega)\right\}$. A geometric interpretation of $\Lambda^\star(x)$ is given in Figure 6. Using the fact that $\Lambda(\omega)$ is a convex function, and $\Lambda(\omega) \ge 0$ for $\omega \in (\omega_-, \omega_+)$, and applying the total probability formula over the event space $E_1 = \left\{T_{\varphi_o} \le \eta, \Theta_{T_{\varphi_o}} \ge \varphi_o\right\}$, $E_2 = \left\{T_{\varphi_o} \le \eta, \Theta_{T_{\varphi_o}} \le -\varphi_o\right\}$, $E_3 = \left\{T_{\varphi_o} > \eta, \Theta_{T_{\varphi_o}} \ge \varphi_o\right\}$, $E_4 = \left\{T_{\varphi_o} > \eta, \Theta_{T_{\varphi_o}} \le -\varphi_o\right\}$ to (15), assuming $\omega > 0$, we obtain:

$$\begin{aligned}
1 &= E\left[\frac{\exp\left(\omega\Theta_{T_{\varphi_o}}\right)\nu_{i(T_{\varphi_o})}\left(\omega\right)}{\sigma(\omega)^{T_{\varphi_o}}\nu_{i(0)}\left(\omega\right)}|E_1\right]\Pr\left\{E_1\right\} \\
&+ E\left[\frac{\exp\left(\omega\Theta_{T_{\varphi_o}}\right)\nu_{i(T_{\varphi_o})}\left(\omega\right)}{\sigma(\omega)^{T_{\varphi_o}}\nu_{i(0)}\left(\omega\right)}|E_3\right]\Pr\left\{E_3\right\} \\
&+ E\left[\frac{\exp\left(\omega\Theta_{T_{\varphi_o}}\right)\nu_{i(T_{\varphi_o})}\left(\omega\right)}{\sigma(\omega)^{T_{\varphi_o}}\nu_{i(0)}\left(\omega\right)}|E_2\right]\Pr\left\{E_2\right\} \\
&+ E\left[\frac{\exp\left(\omega\Theta_{T_{\varphi_o}}\right)\nu_{i(T_{\varphi_o})}\left(\omega\right)}{\sigma(\omega)^{T_{\varphi_o}}\nu_{i(0)}\left(\omega\right)}|E_4\right]\Pr\left\{E_4\right\} \\
&\ge E\left[\exp\left(\omega\Theta_{T_{\varphi_o}} - T_{\varphi_o}\log\sigma(\omega)\right)\frac{\nu_{i(T_{\varphi_o})}\left(\omega\right)}{\nu_{i(0)}\left(\omega\right)}|E_1\right]\Pr\left\{E_1\right\} \\
&\ge \exp\left(\omega\varphi_o - n\log\sigma(\omega)\right)\frac{\min_j\nu_j\left(\omega\right)}{\nu_{i(0)}\left(\omega\right)}\Pr\left\{E_1\right\},
\end{aligned} \tag{24}$$

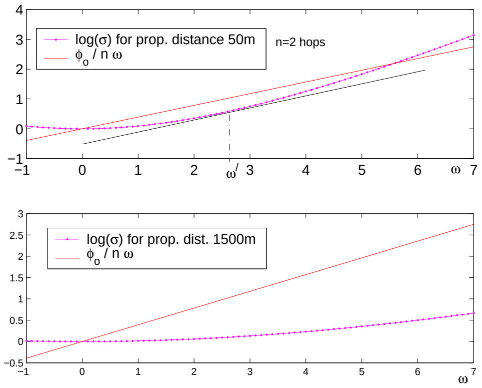


Fig. 6. Geometric Interpretation of the Bound for $\Pr\{T_{\varphi_o} \leq 2\}$: the upper subplot corresponds to a BeSpoken design where the targeted spoke length $d_s \approx 50$m. The lower plot corresponds to a design where the spoke length $d_l \approx 1500$m, corresponding to smaller crescent subtending angles than the d_s design - *the crossing probability after two hops is smaller for d_l; the probability bounds relate in the same way: $exp(-2\delta_l) < exp(-2\delta_s)$, $\delta_s = 0.5$ is the distance $\Lambda^*(\varphi_o/n) = \omega\varphi_o/n - \Lambda(\omega)$, evaluated for $n = 2$, $\omega = \omega^{'}$, for d_s design, and $\delta_l > 50$ is the equivalent for d_l design.*

then using $\omega < 0$ we get the similar result for event E_2, and, combining the two, we obtain a CDF bound

$$\Pr\{T_{\varphi_o} \leq \eta\} \leq \exp\left(-\eta(\omega\frac{\varphi_o}{\eta} - \log\sigma(\omega))\right)\frac{\nu_{i(0)}\left(\omega\right)}{\min_j \nu_j\left(\omega\right)}. \tag{25}$$

We base (25) on the largest eigenvalue $\sigma(\omega)$ of an m-state Markov Chain, for sufficiently large m. We apply numerical methods to obtain $\sigma(\omega)$ and observe that (25) (with $\nu_{i(0)}\left(\omega\right)/\min_j \nu_j\left(\omega\right) = 1$) tightly bounds the CDF obtained from the simulations, as shown by Figure 7.

The expression (25) evaluated for some desired $\Pr\{T_{\varphi_o} \leq \eta\} = p_t$ provides an upper bound for q, as opposed to the lower bound obtained through (21). Figure 5 (b) illustrates the achieved wobbliness in another sample of 500 spokes directed eastward, designed according to (25).

V. Results and Conclusion

We propose a protocol that generates spokes, relatively straight-line data dissemination trajectories, without requiring the nodes to have navigational information. Analysis of a Markov-modulated random walk model for the spoke results in a design for protocol parameters, necessary to produce sufficiently long and straight enough trajectories.

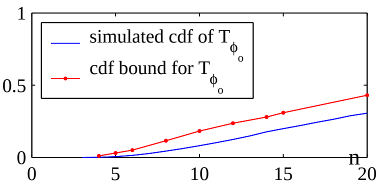


Fig. 7. Comparison of the LDT-based CDF bound and CDF obtained by "sampling" the underlying m-state Markov Chain.

We support our analysis with simulation results. We simulate a stationary network of unit-density, with uniformly distributed nodes deployed over a square region. Simulation statistics are generated by extending a large number of spokes to follow the same direction (as in Figure 5), over several network realizations. The collected averages for the spokes going off-course confirm the validity of the design with respect to the imposed constraint "in the mean". We also suggest a constraint based on the probability bound of the hop count before spoke goes off-course. For a given data range r, the statistics obtained from the simulations of the BeSpoken designed according to this constraint show a better control of the spoke direction at the expense of a slightly increased rate of prematurely stopped spokes due to outage (Figure 5 (b)).

References

[1] F. Bai and A. Helmy. Comparative analysis of algorithms for tree structure restoration in sensor networks. In *Proc. of 23rd IEEE ICPCC*, pages 385–391, 2004.

[2] D. Braginsky and D. Estrin. Rumor routing algorithm for sensor networks. In *Proc. 1st ACM Int'l. Wksp. Wireless Sensor Nets and Apps.*, 2002.

[3] A. Dembo and O. Zeitouni. *Large Deviation Techniques and Applications*. Springer, 1998.

[4] B. Hajek. Minimum mean hitting times of brownian motion with constrained drift. In *Proc. 27th Conf. on Stochastic Processes and Their Applications*, 2000.

[5] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk. Enhancing source-location privacy in sensor network routing. In *Proc. of the 25th Annual IEEE ICDCS*, pages 599–608, 2005.

[6] S. Kokalj-Filipovic, P. Spasojevic, and R. Yates. Bespoken protocol for data dissemination in wireless sensor networks. In *SpaSWIN 2007 workshop*. IEEE, April 2007.

[7] S. Ni, Y. Tseng, Y. Chen, and J. Sheu. The broadcast storm problem in a mobile ad hoc network. In *Proc. 5th ACM/IEEE MOBICOM*, 1999.

[8] D. Niculescu and B. Nath. Trajectory based forwarding and its applications. In *Proc. 9th ACM/IEEE MOBICOM*, pages 260–272, 2003.

[9] R.Gallager. *Discrete Stochastic Processes*. Kluwer Academic Publishers, 1995.

[10] S. Shakkottai. Asymptotics of query strategies over a sensor network. In *Proceedings 23rd IEEE Infocom, Hong Kong*, 2004.

[11] B. Willams and T. Camp. Comparison of broadcasting techniques for mobile ad hoc networks. In *Proc. of 3rd ACM MOBIHOC*, 2002.

VI. Appendix

MMRW Overshoot Analysis

The one-state Markov Chain modulated random walk $\Theta_n = \sum_{i=1}^{n} \Phi_i$ is in fact the IID random walk. This random walk stops if the condition in (14) is satisfied. We define the undershoot as $X = \varphi_o - \Theta_{T_{\varphi_o}-1}$, while the overshoot is defined as $Y = \Theta_{T_{\varphi_o}} - \varphi_o$. As the IID Φ_i is uniform over $\{-\varphi_{11}, \varphi_{11}\}$, and as at T_{φ_o} Φ_i assumes a positive value, we conjecture that random variables X and Y have the same pdfs $f_X(x) = f_Y(x)$ (or at least the first two moments), both uniform, with support set $\{0, \varphi_{11}\}$. We define random variable $Z = X + Y$ s.t. $Z|Y \sim U\left(Y, \varphi_{11}\right)$.

As $E\left[Z\right] = E\left[Y\right] + E\left[X\right] = 2E\left[Y\right] = 2m$ and $E\left[Z\right] = E_Y\left\{E\left[Z|Y\right]\right\} = E_Y\left\{(Y + \varphi_{11})/2\right\} = (m + \varphi_{11})/2$, we obtain the first moment of the overshoot as $E\left[Y\right] = m = \varphi_{11}/3$. Further, we establish

$$\begin{aligned} E\left[Y^2\right] &= m_2, E\left[Z^2\right] = 2m_2 + 2E\left[XY\right] = m_2 + m\varphi_{11} \\ E\left[Z^2\right] &= E_Y E\left[Z^2|Y\right] = (1/3)(m_2 + m\varphi_{11} + \varphi_{11}^2) \end{aligned} \tag{26}$$

Solving the system of equations (26) we obtain the second moment of the overshoot $E\left[Y^2\right] = \varphi_{11}^2/6$. For symmetry reasons the variance of the random walk at overshoot is equal at both φ_o and $-\varphi_o$. Thus, as both overshoot occurrences are equiprobable,

$$\begin{aligned} \text{var}\left[\Theta_{T_{\varphi_o}}\right] &= 0.5\left(2E\left[(\varphi_o + Y)^2\right]\right) \\ &= \varphi_o^2 + (2/3)\varphi_o\varphi_{11} + \varphi_{11}^2/6. \end{aligned} \tag{27}$$