

Geographic Data Propagation in Location-Unaware Wireless Sensor Networks: A Two-Dimensional Random Walk Analysis

Silvija Kokalj-Filipović, Predrag Spasojević, and Roy Yates

Abstract—For wireless sensor networks with many location-unaware nodes, which can be modeled as a planar Poisson point process, we investigate a protocol, dubbed BeSpoken, which steers data transmissions along a straight path called a spoke. BeSpoken implements a simple, spatially recursive process, where a basic set of control packets and a data packet are exchanged repeatedly among daisy-chained relays that constitute the spoke. Hence, a data packet originated by the first relay makes a forward progress in the direction of the spoke. Despite the simplicity of the protocol engine, modeling the spoke process is a significant challenge. BeSpoken directs data transmissions by randomly selecting relays to retransmit data packets from crescent-shaped areas along the spoke axis. The resulting random walk of the spoke hop sequence may be modeled as a two dimensional Markov process. Based on this model, we propose design rules for protocol parameters that minimize energy consumption while ensuring that spokes propagate far enough and have a limited wobble with respect to the spoke axis. The energy efficiency is demonstrated through simulations of the BeSpoken-based data search, and a comparison with the energy consumption of a search based on directed diffusion.

Index Terms—Wireless sensor networks, Poisson point process, geographic data propagation, stochastic analysis, location-unaware, Markov-modulated random walk, energy efficient data search.

I. INTRODUCTION

IN RANDOMLY deployed *Wireless Sensor Networks* (WSNs) composed of position unaware nodes, data sources are frequently unaware of which data sinks have interest in their observations. An example is a network of cheap, battery-operated sensor nodes scattered over an area, used for environmental monitoring and expected to efficiently deliver gathered information to a data collector (sink) located at an arbitrary (and frequently random) position at the network boundary.

Given the scarce resources and the limited processing power of WSN nodes, the unknown position of a data sink makes the task of delivering data especially challenging. Several new communication paradigms, like geocasting, data dissemination, and data search, emerged from this problem [7]. In the geocasting problem [12] data needs to be routed to a geographic region instead of a destination node specified by an

address. Flooding the network is a trivial form of geocasting when sources are unaware of sink locations. Unrestricted flooding as a dissemination method leads to a “broadcast storm” of redundant transmissions [13] and consumes more resources than necessary [6]. Two dissemination techniques that use flooding selectively are briefly described next. In a *push* approach [5], a publishing process plants pointers in the network that can be used by the interested sinks to establish a path to the correct source. Publishing mechanisms are largely based on flooding and consequent path reinforcement. Alternatively, in [3] the authors introduce a data-centric *pull* mechanism called *directed diffusion* in which interest requests (queries) are flooded into the network leaving gradient paths back to the sink. With location-unaware nodes, a more efficient alternative to flooding is to use landmark-based routing protocols [4] to store state information in selected nodes (possibly along a path) to direct the search toward the correct source [14]. As another closely related example to our work, Rumor Routing [1] introduces the concept of agents, packets that advertise source data along a random walk path that resembles a fairly straight trajectory. The query packet follows a similar random walk path, and the success of the search is based on the high probability that the two sufficiently long lines in a bounded rectangle intersect.

In this work we propose BeSpoken, a wireless communication protocol that also steers data transmissions along fairly straight trajectories, and enables energy efficient data dissemination and search in networks with location unaware nodes. In addition, we illustrate the energy efficiency of an example push-pull dissemination infrastructure, mechanized using the BeSpoken-enabled trajectories.

The BeSpoken protocol implements a simple, spatially recursive process controlling data transmissions, where a basic set of control packets and a data packet are exchanged repeatedly among daisy-chained relays that constitute a trajectory, also called a *spoke*. The barrier-crossing analysis of a Markov-modulated random walk model for the spoke evolution process results in design conditions which protocol parameters need to satisfy to produce sufficiently long and straight trajectories.

Figure 1 illustrates an example of BeSpoken-based push-pull data dissemination. Here a source disseminates data advertisements along the source spokes, and a data collector (sink) sends a query along its own spokes that may intersect the source spokes. Each intersection represents a successful search. Successful search is followed by the endorsement of a

Manuscript received 30 August 2008; revised 31 January 2009. This paper was supported in part by the NSF grant 0721888. This paper is in part based on [11] and [9].

The authors are with WINLAB, ECE Department, Rutgers University (e-mail: {skokalj,spasojev,ryates}@winlab.rutgers.edu).

Digital Object Identifier 10.1109/JSAC.2009.0909xx.

route along the intersecting spokes, and by the corresponding data dissemination (see the closeup in Figure 1). In this example, the sinks are distributed uniformly along the perimeter of the network area and their position is unknown in advance. Hence, we conjecture that the likelihood of successful data search will increase if both the source and the sinks spawn several equally spaced radial spokes. Hence the twofold meaning of the name BeSpoken: the radial lines extending from the source form a pattern that resembles spokes of a wheel and, furthermore, spoke relays *bespeak* the source message. The relative direction of spokes allowing for the wheel pattern is controlled by *an extension of the BeSpoken* not discussed here due to space considerations.

The taken approach is conceptually closest to geographic greedy forwarding schemes [8], [16] used for routing to known destinations, with an important distinction that, instead of greedily approaching the sink, in our approach the data is greedily directed away from the source. In the greedy geographic forwarding scheme a packet is forwarded to a one-hop neighbor which is closer to the destination than the current node. The similarity is only conceptual, since the assumptions are orthogonal to ours: a source node knows the location of the destination node, and network nodes are location-aware.

The focus of this paper is the mathematical model of the BeSpoken and its analysis for the purpose of spoke design, as well as demonstrating that the energy consumption of a BeSpoken based search is much smaller than with the search based on directed diffusion [3], and that it decreases with the increasing number of events in the network. The rest of the paper is organized in the following manner: Section II introduces BeSpoken - II-A describes the protocol engine, and II-B depicts the effects of the protocol parameters on the spoke metrics, expressed through wobblines and propagation outage constraints; Section III focuses on spoke modeling - III-A introduces BeSpoken geometry, while III-B and III-C describe Markov process and Markov chain models of the spoke; Section IV analyzes the propagation outage constraint, and Section V provides analysis of the wobbliness constrain; Section VI presents simulation results for both a single BeSpoken trajectory and an example dissemination infrastructure, and concludes.

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 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

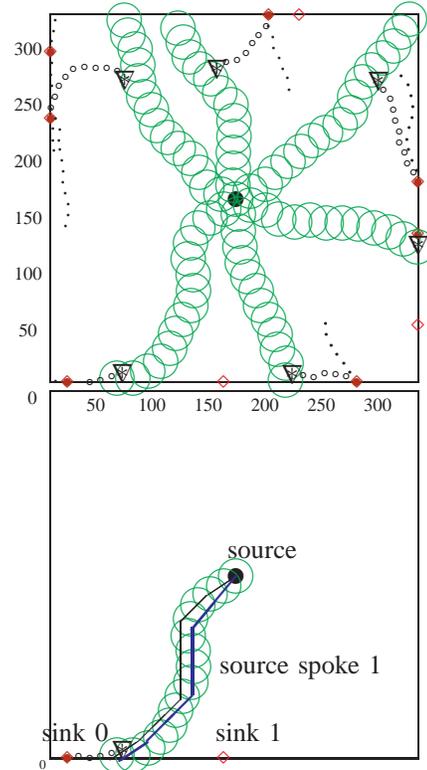


Fig. 1. In this simulation snapshot, sink 1 is not interested in advertised data while sink 0 has one productive sink spoke, and selects a data dissemination route from the ensemble of routes within that spoke, and within the intersecting source spoke 1. The sequence of wireless transmission *relays* forming the productive sink spoke (the one that hits a source spoke first) is denoted by tiny circles (see the sink spoke in the closeup). Unproductive sink spokes are represented by dots. The search success is marked by a ∇ with an inscribed $*$.

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.

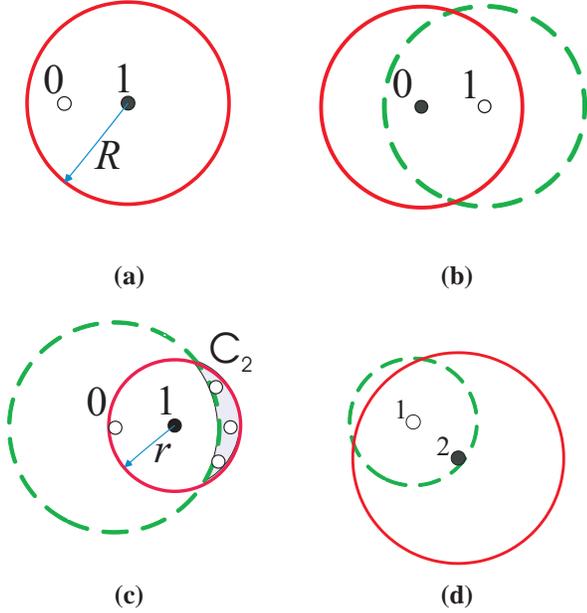


Fig. 2. BeSpoken Protocol: At each protocol stage, the current transmission range is denoted with the full circle while the previous range is denoted with a dashed circle.

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 wobbliness, 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 wobbliness requires increasing r to generate a large enough crescent and decrease the 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 $dr^{\alpha-1}$. Hence, minimizing the transmission range r corresponds to a minimum energy objective.

These competing tendencies illustrate the importance of the protocol parameters design. In this paper we show that outage

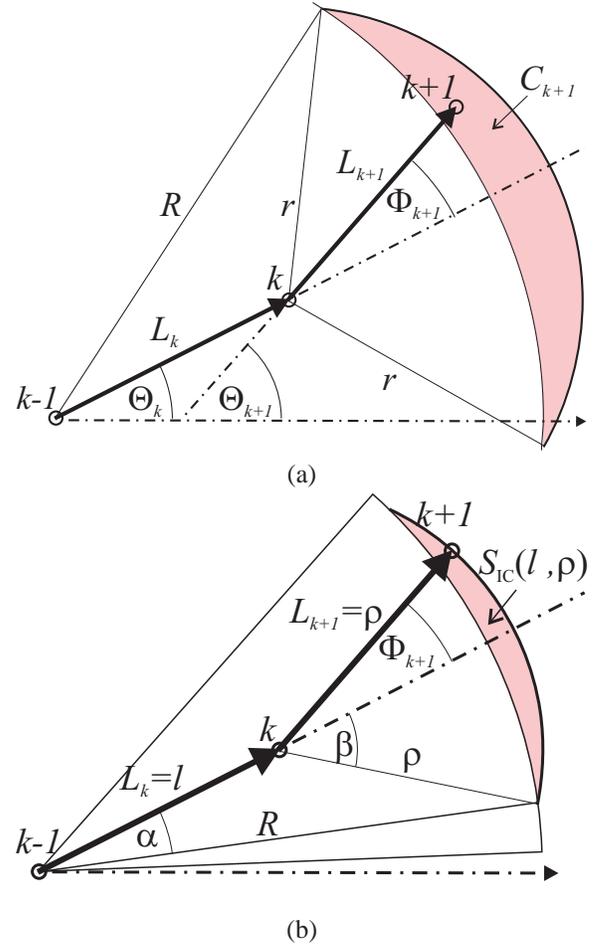


Fig. 3. (a) At hop $k + 1$, node $k + 1$ is distance L_{k+1} from node k and the current spoke direction is $\Theta_{k+1} = \Theta_k + \Phi_{k+1}$. (b) Given $L_k = l$ and $L_{k+1} = \rho$, the angular hop displacement Φ_{k+1} is constrained to the interval $-\beta \leq \Phi_{k+1} \leq \beta$ where the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

and wobbliness constraints can be decoupled. Consequently, as a result of the outage constraint analysis, we give the design guidelines for the parameter r . We demonstrate that satisfying the wobbliness constraint requires one to find 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 , given r . We develop closed-form expressions that serve as bounds for the values of q , ensuring that the wobbliness constraint is satisfied.

III. SPOKE MODELING

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

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

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

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

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

Note that 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 $S_c(L_k) = S_{IC}(L_k, r)$. 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 forms a candidate set for node $k + 1$ with cardinality Z_k that is, conditionally, a Poisson random variable with conditional expected value

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

A spoke stops at stage k when the crescent C_k is empty and thus spoke generation is a transient process. The outage constraint depends only on the crescent sizes $S_c(L_k)$ but not on the hop direction process Θ_k . On the other hand, the spoke wobbliness depends on the Θ_k but is meaningful only as long as each current crescent C_k is non-empty. Thus, we separate the analysis of the outage and wobbliness constraints 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 (b) 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

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

and zero otherwise. We note that (4) provides a complete characterization of the fictitious process $\{L_k\}$.

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 $\{h_1, \dots, h_m\} \subseteq [R - r, r]$ to be the chain state set. Without loss of generality, we assume that $h_0 = R - r < h_1 < h_2 < \dots < h_m = r$. As

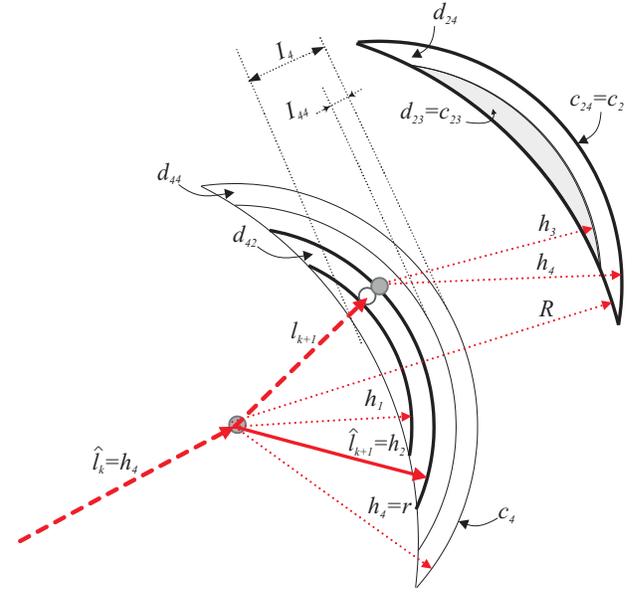


Fig. 4. Ergodic Finite State Markov Chain: quantization example for a four-state chain ($m = 4$): $\hat{L}_k = h_4 = r$ results in the first crescent \hat{C}_k of area c_4 partitioned into four strips of total area $c_4 = d_{41} + d_{42} + d_{43} + d_{44}$; $L_{k+1} \in \mathcal{I}_{42}$, quantized to $\hat{L}_{k+1} = h_2$, 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).

illustrated in Figure 4, whenever the k th 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

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

Note that the set $\{\mathcal{I}_{ij} : j = 1, \dots, m\}$ partitions \mathcal{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

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

(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 \mathcal{I}_i .

As shown in Figure 4, $c_{ij} = S_{IC}(h_i, h_j)$ is the *quantized interior crescent area* formed by the control circle (of radius R) centered at the k th hop relay and a circle of radius h_j centered at node $k + 1$ at distance $\hat{L}_k = h_i$. Note that $c_{ij} < c_{i(j+1)} \dots < c_{im}$, 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

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

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. Driven by the modeling criteria of simplicity and efficiency, we consider Markov chain models with both uniform and non-uniform quantization of $[R - r, r]$. With only $m = 2$ levels, the uniform quantization lacks accuracy. However, a carefully chosen two-state chain provides a useful non-uniform quantization model. The transition matrix for both two-state systems is

$$\mathbf{P} = \begin{bmatrix} 0 & 1 \\ c_{21}/c_2 & (c_2 - c_{21})/c_2 \end{bmatrix}, \quad (8)$$

since $c_{12} = c_1$ and $c_{22} = c_2$.

1) *Uniform Quantization Model*: In this model, the hop-length states $\{h_i\}$ uniformly quantize the process state space $[R-r, r]$ so that $h_i = R-r+i\Delta$, where $\Delta = (2r-R)/m$ is the quantization interval. Furthermore, $j^*(i) = m - i$ so that the next-hop quantization intervals \mathcal{I}_{ij} satisfy $\mathcal{I}_{ij} = (h_j - \Delta, h_j]$ for $j > m - i$ and are empty for $j \leq m - i$. The transition probabilities are now

$$P_{ij} = \frac{c_{ij} - c_{i(j-1)}}{c_i}, \quad i + j > m, \quad (9)$$

and $P_{ij} = 0$ whenever $i + j \leq m$ follows since, in that case, $(h_{j-1}, h_j]$ and $\mathcal{I}_i = [R - h_i, r]$ intersect in at most one point. For example, the uniformly quantized $m = 2$ Markov chain has $\Delta = r - R/2$, $h_1 = R - r + \Delta = R/2$ and $h_2 = r$, and, accordingly, $c_1 = S_c(R/2)$ and $c_2 = S_c(r)$.

2) *Non-Uniform Quantization Model for the Outage Constraint*: Non-uniform quantization, being inherently more complex than uniform, qualifies only if its application renders a simple two-state model possible. The proposed non-uniform quantization, two-state Markov chain model has a simpler definition with $c_1 = 1$, and $c_2 = S_c(r)$. The corresponding set of hop length states includes $h_2 = r$ and h_1 , which is a solution of $c_1 = 1 = S_c(h_1)$. Hence, the next hop partition mapping satisfies $d_{21} = c_{21}$, and $d_{22} = c_2 - c_{21}$. Let $R/2 > h_1 = S_c^{-1}(1) > R - r$, and, in this case, we have that $j^*(1) = 2$, $d_{11} = 0$, and $d_{12} = c_1 = 1$. The non-uniform partitioning differs from the uniform in that $c_2 \gg c_1$ and $c_{22} \gg c_{21}$ for large enough r . The rationale behind such a design follows in the next section.

D. The 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

$$\Theta_{k+1} = \Theta_k + \Phi_{k+1} = \sum_{i=1}^{k+1} \Phi_i. \quad (10)$$

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

$$f_{\Phi_{k+1}|L_k, L_{k+1}}(\phi|l_k, l_{k+1}) = \frac{1}{2\beta(l_k, l_{k+1})}, \quad (11)$$

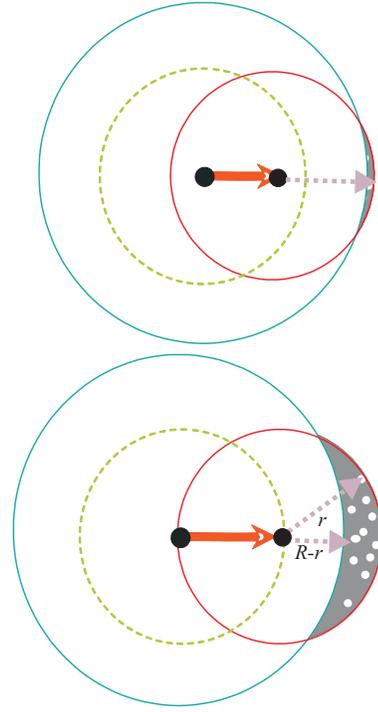


Fig. 5. Spring-coil analogy

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), (11).

The transform domain analysis of a Markov Modulated Random Walk (MMRW) [15] dictates that we first define the conditional moment generating functions of the incremental angular displacement Φ_{k+1} from (10)

$$\begin{aligned} g_{ij}(\omega) &= E \left[\exp(\Phi_{k+1}\omega) | \hat{L}_k = h_i, \hat{L}_{k+1} = h_j \right] \quad (12) \\ &= \frac{1}{2\varphi_{ij}} \int_{-\varphi_{ij}}^{\varphi_{ij}} \exp(\phi\omega) d\phi = \hbar(\varphi_{ij}\omega), \quad (13) \end{aligned}$$

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

$$\Gamma_{ij}(\omega) = P_{ij}g_{ij}(\omega). \quad (14)$$

The Perron-Frobenius theorem (see e.g., [2]) 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 [15]

$$M_k(\omega) = \frac{\exp(\omega \Theta_k) \nu_{i(k)}(\omega)}{\sigma^k(\omega) \nu_{i(0)}(\omega)} \quad (15)$$

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. MODEL FOR THE OUTAGE CONSTRAINT

Here we evaluate the outage probability for given $q = R/r$, in order to evaluate the associated outage constraint (18). With

respect to outage, a spoke stops at hop k when the crescent C_k is empty, i.e., $Z_k = 0$. Since the nodes obey a planar Poisson process, it follows from (3) that the conditional probability the crescent C_k is empty is

$$\Pr\{Z_k = 0 | L_k = l_k\} = e^{-S_c(l_k)}. \quad (16)$$

We define

$$D = \min\{n : Z_n = 0\} \quad (17)$$

as the first time the process encounters an empty crescent.

For analytical tractability, instead of requiring the spoke to travel distance d with high probability, we require it to travel η hops with high probability. In particular, we define $\eta = \lceil d/r \rceil$ as the number of hops corresponding to an idealized straight-line spoke extending to the distance d . The design outage constraint can be formalized as

$$\Pr\{D \leq \eta\} \leq p. \quad (18)$$

However, the analysis of (18) is challenging due to the complex way in which the hop length process $\{L_k\}$ evolves with time. 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 $[R - r, R - L_k]$. As illustrated in Figure 5, an imaginary coil is attached between a fixed pivot and a moving leading relay: when contracted, it pulls the leading relay's data circle inside the blocking control circle, exposing only a tiny area with possible relays. Note that the next hop length has to be long (close to r), if the relay is found in this tiny area. At the other extreme, when the coil is completely relaxed to length r , it exposes the largest possible area. This reduces the likelihood of an empty crescent yet it increases the likelihood of the next hop length being small. This oscillatory effect illustrates the importance of the Markov property for the hop length evolution model (4).

For the m -state Markov chain, let us denote the event that the first η crescents \hat{C}_k , $k = 1, \dots, \eta$, are not empty as $A_\eta = \{\min_{k \leq \eta} Z_k > 0\}$. The probability that the crescents $\hat{C}_1, \dots, \hat{C}_\eta$ are not empty, and that the system is in state j at time η is denoted $\kappa_j^{(\eta)} = \Pr\{\hat{L}_\eta = h_j, A_\eta\}$. Using Markovity of \hat{L}_k and conditional independence of Z_k given \hat{L}_k , it is straightforward to show that

$$\kappa_j^{(\eta)} = \sum_{i=1}^m e_j P_{ij} \kappa_i^{(\eta-1)} \quad (19)$$

where $e_j = 1 - \exp(-\lambda c_j)$ is the probability of a non-empty crescent while in state j . Let us define the $m \times m$ matrix $\check{\mathbf{P}}$ where $\check{P}_{ij} = P_{ij} e_j$ is the conditional probability to transition from state i to state j , and that the resulting crescent of area c_j is not empty. Note that (8) implies $P_{11} = \check{P}_{11} = 0$. In addition, by defining the vector $\kappa^{(\eta)} = [\kappa_1^{(\eta)}, \dots, \kappa_m^{(\eta)}]$, (19) becomes $\kappa^{(\eta)} = \kappa^{(\eta-1)} \check{\mathbf{P}}$. Recursively, we obtain $\kappa^{(\eta)} = \kappa^{(1)} \check{\mathbf{P}}^{\eta-1}$. Given the initial state m , we see that $\kappa_i^{(1)} = 0$ for $i < m$ and $\kappa_m^{(1)} = e_m$. Thus, $\kappa^{(\eta)} = [0 \dots e_m] \check{\mathbf{P}}^{\eta-1}$. As $\Pr\{A_\eta\} = \sum_{i=1, \dots, m} \kappa_i^{(\eta)} = \kappa^{(\eta)} [1 \dots 1]^T$, the probability that the spoke will stop at or before hop η (assuming that the chain always starts in state h_m) becomes

$$\begin{aligned} \Pr\{D \leq \eta\} &= 1 - \Pr\{A_\eta\} \\ &= 1 - [0 \dots e_m] \check{\mathbf{P}}^{\eta-1} [1 \dots 1]^T. \end{aligned} \quad (20)$$

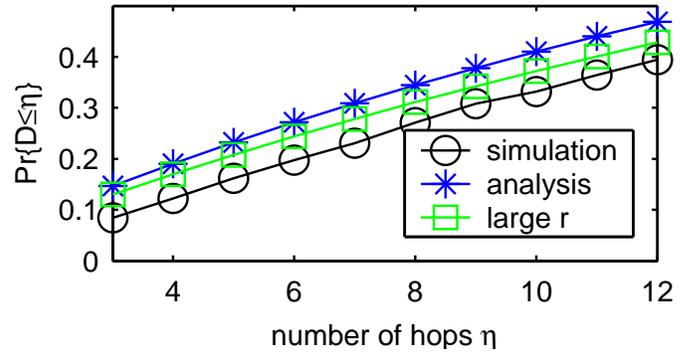


Fig. 6. Outage probability curves

The following asymptotic (large r) analysis of the outage probability (20) is based on the two state non-uniform quantization model (8). Let $\lambda_1 > \lambda_2$ be the two eigenvalues of $\check{\mathbf{P}}$ in (20) based on (8). The eigenvalue λ_1 describes the rate at which the outage probability increases with the number of hops, while the negative eigenvalue λ_2 describes the oscillatory, self-recovery mechanism depicted in Figure 5. Let $r \gg 1$ and q be close to two. Then $c_{22} = c_2 \gg c_1 = 1$ in (8). Now, λ_1 is close to one, while λ_2 one is close to zero. Furthermore, by combining (20) and (18), while expressing the two-state $\check{\mathbf{P}}$ through its singular value decomposition, and truncating the Taylor expansions of λ_1 and λ_2 to their significant terms, we show that, for a spoke to reach η hops with probability p , given q , the range is required to be

$$r \geq 1/\sqrt{\exp(1) \left(1 - (1-p)^{\frac{1}{\eta-1}}\right) f(q)}, \quad (21)$$

where $f(q) = S_c(r)/r^2$. Note that (2) implies that $f(q)$ is the crescent area for unit r (the ratio $S_c(r)/r^2$ does not depend on r). Figure 6 illustrates how well (21) matches the simulation results for large r .

V. WOBBLINESS CONSTRAINT MODEL

The spoke goes off-course at hop k whenever the current angle Θ_k in (10) exceeds one of the following two thresholds ϕ_o and $-\phi_o$. To describe spoke wobbliness, we define

$$T_{\phi_o} = \min\{k : |\Theta_k| \geq \phi_o\}. \quad (22)$$

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 [15, 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 [15, Lemma 6] and the *optional sampling theorem* [15, Theorem 6] we have

$$E[M_{T_{\phi_o}}(\omega)] = E\left[\frac{\exp(\omega \Theta_{T_{\phi_o}}) \nu_{i(T_{\phi_o})}(\omega)}{\sigma(\omega)^{T_{\phi_o}} \nu_{i(0)}(\omega)}\right] = 1, \quad (23)$$

for $\omega \in (\omega_-, \omega_+)$. 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 (23), which is an extension of the *Wald identity* to Markov

modulated random walks. The first wobbliness constraint is based on the first moment of T_{φ_o} , as

$$E[T_{\varphi_o}] \geq \eta. \quad (24)$$

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

$$\Pr\{T_{\varphi_o} \leq \eta\} \leq p_t. \quad (25)$$

In subsection V-A we demonstrate how to compute the mean $E[T_{\varphi_o}]$. Subsection V-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

$$E[\Theta_{T_{\varphi_o}}] = 0, \quad \text{var}[\Theta_{T_{\varphi_o}}] = E[\Theta_{T_{\varphi_o}}^2] = \varphi_o^2. \quad (26)$$

We evaluate the second derivative of (23) 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

$$E[T_{\varphi_o}] = \frac{\text{var}[\Theta_{T_{\varphi_o}}] + E[\mu_i(T_{\varphi_o})(\omega)]|_{\omega=0} - \mu_{i(0)}(\omega)|_{\omega=0}}{\frac{\sigma''(\omega)}{\sigma(\omega)}|_{\omega=0}}. \quad (27)$$

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

$$\frac{1}{3} (\pi_2 P_{22} \varphi_{22}^2 + \pi_1 P_{11} \varphi_{11}^2 + \pi_1 P_{12} \varphi_{12}^2 + \pi_2 P_{21} \varphi_{21}^2), \quad (28)$$

where π_i , $i = 1, 2$ are the elements of the vector of stationary state probabilities $\bar{\pi} = [\pi_i]_{(1 \times m)}$. Note that terms $\varphi_{ij}^2/3$ are transition-specific variances (for uniform angular displacement). Direct generalization of (28) to an m state model has a form of a stationary average of transition-specific variances over m^2 transitions

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

where, $P^{(v)} = [P_{ij}^{(v)}]_{(m \times m)}$ with elements $P_{ij}^{(v)} = (P_{ij} \varphi_{ij}^2)/3$, and $u = [1 \dots 1]_{(1 \times m)}$. Now, it can be shown that, for small crescent subtending angles relative to the threshold φ_o , we can ignore the terms $E[\mu_i(T_{\varphi_o})]$ and $\mu_{i(0)}$ in (27), thus

$$E[T_{\varphi_o}] = \frac{\text{var}[\Theta_{T_{\varphi_o}}]}{\text{var}[\theta]_p}. \quad (29)$$

Since (27) 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. $\Phi_i \sim U(-\varphi_{11}, \varphi_{11})$. 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 (29) for a one-state MMRW

$$\text{var}[\Theta_{T_{\varphi_o}}] = \varphi_o^2 + (2/3)\varphi_o\varphi_{11} + \varphi_{11}^2/6. \quad (30)$$

Note that the second term in (30) contains the half-span φ_{11} of the uniform pdf. To extend the expression (30) to m -state Markov-modulated random walk, we replace φ_{11} with a weighted sum of transition-specific angle spans $\sum_{i,j=1}^m w_{ij} \varphi_{ij}$, 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 (30) is one half of the angle variance for $\Phi_i \sim U(-\varphi_{11}, \varphi_{11})$. For the m -state MMRW, we replace this term with another weighted sum where $(w_{ij}/2)$ -weighted terms are transition specific variances $\varphi_{ij}^2/3$. Hence, extended (30), in matrix notation, is

$$\text{var}[\Theta_{T_{\varphi_o}}] = \varphi_o^2 + 2\varphi_o \left(\bar{\pi} P^{(a)} u^T \right) + \frac{\bar{\pi} P^{(v)} u^T}{2}, \quad (31)$$

$P^{(a)} = [P_{ij}^{(a)}]_{(m \times m)}$ and $P_{ij}^{(a)} = P_{ij} \varphi_{ij}/3$. The overshoot-inclusive variant of (29) for a multi-state chain (31) yields values that match the simulation results closely and consistently. Figure 7 (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 (24) is a constraint in the mean, where the pertinent pdf is long-tailed.

B. Probability of threshold crossing before certain time

Motivated by the observations illustrated by Figure 7 (a), we here analyze the wobbliness model, as defined in (25), from the point of view of *Large Deviation Theory (LDT)*. We determine a bound for $\Pr\{T_{\varphi_o} \leq \eta\}$ based on the Gärtner-Ellis theorem [2, Thm 2.3.6] and its application to an empirical measure of finite Markov Chains, in particular [2, Exercise 3.1.4]. Let $\varphi_{i(0)}^P$ denote the Markov probability measure associated with the transition probability matrix (7), and with the initial state $\hat{L}_0 = i(0)$. Precisely,

$$\varphi_{i(0)}^P \left(\hat{L}_1 = y_1, \dots, \hat{L}_n = y_n \right) = P_{i(0)y_1} \prod_{i=1}^{n-1} P_{y_i y_{i+1}}$$

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 $\psi_{ij} = U(\varphi_{ij}, \varphi_{ij})$, 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 $E_{i(0)}^P[\cdot]$ the expected value with respect to $\varphi_{i(0)}^P$ and the associated $\prod_{i=1}^n \psi_{y_{k-1} y_k}$, we further define $\Lambda_n(n\omega) = \log E_{i(0)}^P [e^{\omega \sum_{k=1}^n \Phi_k}]$. Following a derivation analogous to [2, 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 (14) as $\Lambda(\omega) \triangleq \lim_{n \rightarrow \infty} \frac{1}{n} \Lambda_n(n\omega) = \log \sigma(\omega)$. According to [2, Thm 3.1.2],

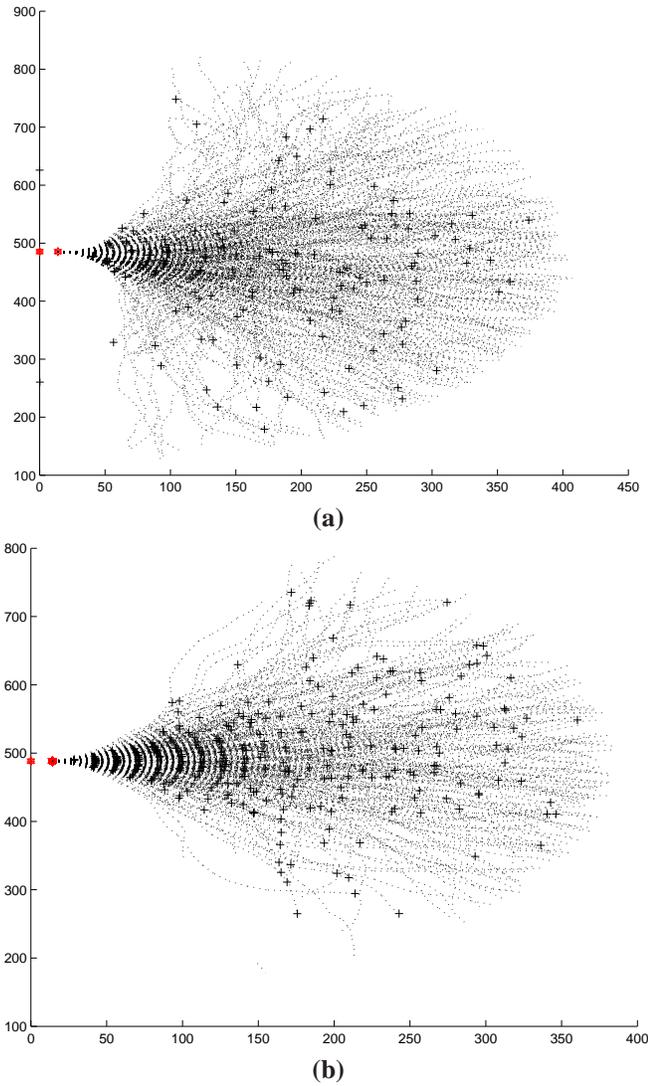


Fig. 7. Sample of spokes directed eastward shows the statistics of spoke angle deviation - design based on the "constraint in the mean" vs. "probability constraint" design.

the empirical mean of the sum of angle deviations modulated by $\varphi_{i(0)}^P$ has a rate function, which is a conjugate function of $\Lambda(\omega)$, i.e the Fenchel-Legendre transform

$$\Lambda^*(x) = \sup_{\omega} \{ \omega x - \Lambda(\omega) \}. \quad (32)$$

A geometric interpretation of $\Lambda^*(x)$ is given in Figure 8. Using the fact that $\Lambda(\omega)$ is a convex function, and $\Lambda(\omega) \geq 0$ for $\omega \in (\omega_-, \omega_+)$, and applying the total probability formula over the event space $E_1 = \{T_{\varphi_o} \leq \eta, \Theta_{T_{\varphi_o}} \geq \varphi_o\}$, $E_2 = \{T_{\varphi_o} \leq \eta, \Theta_{T_{\varphi_o}} \leq -\varphi_o\}$, $E_3 = \{T_{\varphi_o} > \eta, \Theta_{T_{\varphi_o}} \geq \varphi_o\}$, $E_4 = \{T_{\varphi_o} > \eta, \Theta_{T_{\varphi_o}} \leq -\varphi_o\}$ to (23), assuming $\omega > 0$, we obtain:

$$\begin{aligned} 1 &= \sum_{k=1}^4 E \left[\frac{\exp(\omega \Theta_{T_{\varphi_o}}) \nu_{i(T_{\varphi_o})}(\omega)}{\sigma^{T_{\varphi_o}}(\omega) \nu_{i(0)}(\omega)} | E_k \right] \Pr \{ E_k \} \\ &\geq E \left[e^{\omega \Theta_{T_{\varphi_o}} - T_{\varphi_o} \log \sigma(\omega)} \frac{\nu_{i(T_{\varphi_o})}(\omega)}{\nu_{i(0)}(\omega)} | E_1 \right] \Pr \{ E_1 \} \\ &\geq \exp(\omega \varphi_o - \eta \log \sigma(\omega)) \frac{\min_j \nu_j(\omega)}{\nu_{i(0)}(\omega)} \Pr \{ E_1 \}. \end{aligned} \quad (33)$$

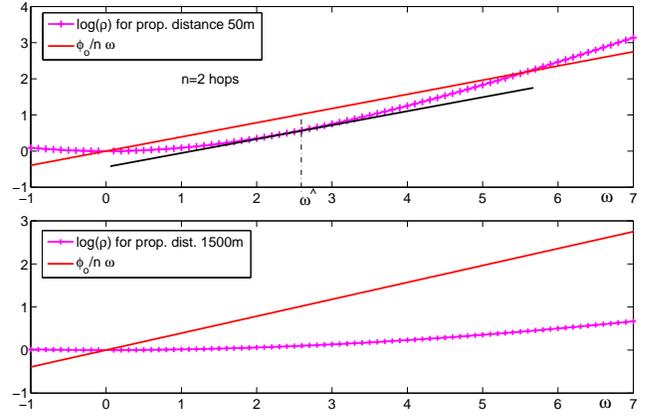


Fig. 8. Geometric Interpretation of the bound for $\Pr \{ T_{\varphi_o} \leq 2 \}$ (probability to go off-course in two or less steps) The **upper** subplot corresponds to a BeSpoken design where the targeted spoke length $d_s \approx 50\text{m}$; this design implies a certain maximum subtending angle of the spoke crescents. The **lower** plot corresponds to a design where the desired spoke length $d_l \approx 1500\text{m}$, corresponding to smaller maximum subtending angle than the design for d_s - Consequently, note that the off-course probability after two hops should be smaller for d_l design; the **probability bounds** relate in the same way: $\exp(-2\delta_l) < \exp(-2\delta_s)$, where $\delta_s = 0.5$ is the maximum distance (at $\omega = \omega'$ as shown in the plot) between the line $\omega \varphi_o/n$ and $\Lambda(\omega) = \log \sigma(\omega)$, which is exactly (32) with $x = \varphi_o/n$, evaluated for d_s design and for $n = 2$. Similarly, $\delta_l > 5$ is the equivalent for d_l design.

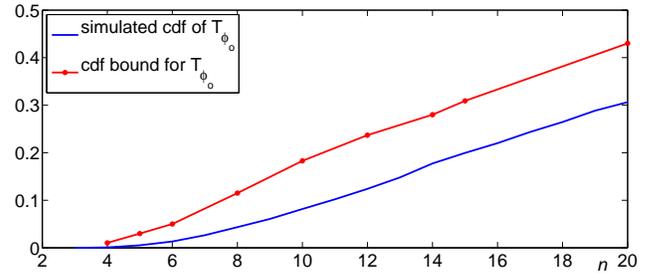


Fig. 9. Comparison of the LDT-based CDF bound and CDF obtained by "sampling" the underlying m-state Markov Chain.

Note that $\Pr \{ E_1 \} = \Pr \{ E_2 \}$, due to the random walk and the threshold symmetries. Hence, by combining the two bounds we have

$$\begin{aligned} \Pr \{ T_{\varphi_o} \leq \eta \} &= \Pr \{ E_1 \} + \Pr \{ E_2 \} = 2\Pr \{ E_1 \} \\ &\leq 2 \exp \left(-\eta \left(\omega \frac{\varphi_o}{\eta} - \log \sigma(\omega) \right) \right) \frac{\nu_{i(0)}(\omega)}{\min_j \nu_j(\omega)}, \end{aligned} \quad (34)$$

where $\omega \in (\omega_-, \omega_+)$. We base (35) 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 (35) (with $\nu_{i(0)}(\omega) / \min_j \nu_j(\omega) = 1$) tightly bounds the CDF obtained from the simulations for the relevant range of values, as shown by Figure 9. The expression (35) 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 (29).

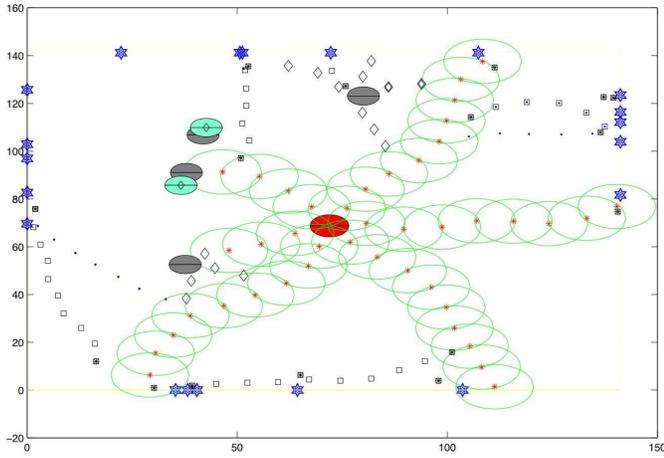


Fig. 10. Simulation snapshot of the BeSpoken search with multiple sources, randomly positioned across the network, and multiple sinks, distributed randomly along the network perimeter. The central source spokes are drawn as a sequence of circles representing transmission ranges of spoke relays. The spoke relays of other sources are denoted by diamonds, while the relays of productive sink spokes are denoted by squares, as in Figure 1.

Given the desired distance d , and the angle threshold φ_o

$n = 1, \dots, \infty$

- (a) Calculate q^* assuming $n = E[T\varphi_o]$ from (29)
- (b) Given $q = q^*$ from (a), calculate r^* from (21)
- (c) If $d/r^* < n$, goto (a) else BREAK.

VI. RESULTS AND CONCLUSION

We propose a protocol that generates spokes, relatively straight-line data dissemination trajectories, without requiring the nodes to have navigational information. The analysis of a Markov-modulated random walk model for the spoke process results in design conditions which protocol parameters need to satisfy to produce sufficiently long and straight trajectories. We here summarize the iterative design algorithm for BeSpoken parameters.

We support our analysis with simulation results. We simulate a stationary WSN of unit-density, with uniformly distributed nodes deployed over a square region.

A. Evaluation of BeSpoken Design for Straight-line Data Propagation

To empirically present the statistics of the spoke direction process, we extend a large number of spokes to follow the same direction (by fixing the first relay to have the same y coordinate as the source, like in Figure 7), over several realizations of a random field of points representing the WSN. The protocol parameters are designed for a network of $20K$ nodes, but the spokes are created over a ten times larger network to better observe the statistics. Figure 7 (b) presents statistics obtained when the protocol design is based on the angular constraint “in probability” (35), and demonstrates a better control of the spoke direction at the expense of a slightly increased rate of prematurely stopped spokes due to outage.

B. Evaluation of BeSpoken Infrastructure for Efficient Data Search

We consider BeSpoken as both a dissemination protocol, and a tool to build an infrastructure of relatively straight

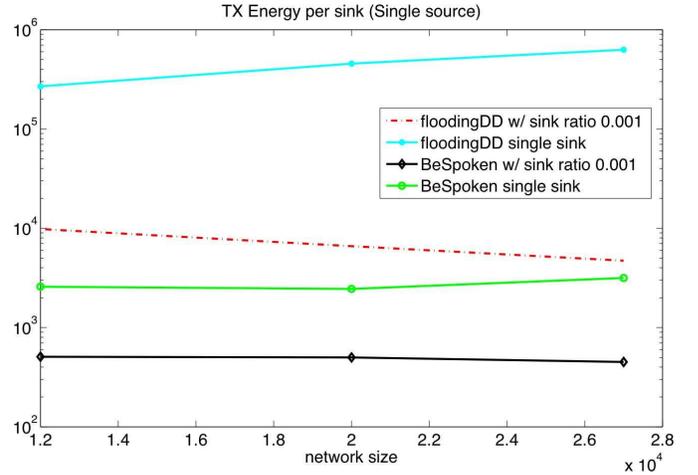


Fig. 11. Comparison of the consumed total network transmission energy per sink as a function of network size for a simulated search based on the BeSpoken, and another one, dubbed FloodingDD, based on directed diffusion.

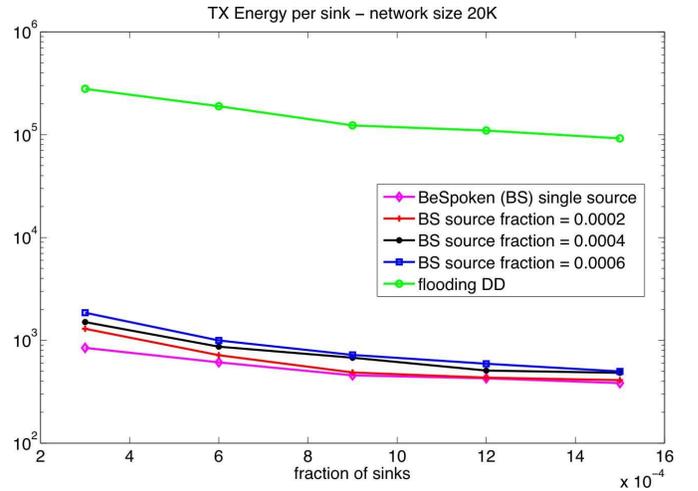


Fig. 12. Comparison of the consumed total network transmission energy per sink as a function of the ratio of sink nodes vs. sensor nodes, for a simulated search based on the BeSpoken, and another one, dubbed FloodingDD, based on the directed diffusion.

paths (spokes) whose direction and length can be learned with moderate effort (see [10]), and that serves as an overlay network for easier, more efficient search. Both source and sink spokes, utilized to convey the information of an event from the source to the sink, may remain active for a relatively long time period. The spokes tessellate the sensor network space, providing a way to map subsequent events to areas between the known paths, and to aid efficient navigation toward the associated sources. It is not hard to envision how this framework may bootstrap a number of mechanisms for energy-efficient dissemination, load balancing and rapid data propagation in desired directions. Of course, additional protocols (outside the scope of this paper) are needed to manage and control such applications.

For energy efficiency statistics, we simulate a data search where several data collectors, randomly distributed along the network perimeter, use BeSpoken to create search trajectories for specific events (characterized by some descriptors), while a number of sources, distributed randomly across the network

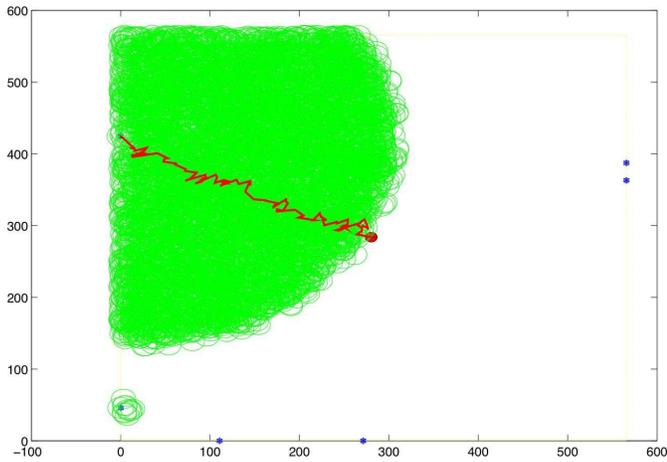


Fig. 13. Search based on directed diffusion: a simulation snapshot showing the area covered by the flooding process as the first random sink searched for data, and also the reinforced (most-efficient) path of the forwarding relays that lead to the source. Observe the beginning of another flooding process in the lower left corner.

area, advertise their events along equally spaced source spokes extended radially away from the source, as shown in Figure 10. These simulations simultaneously evaluate BeSpoken in its capacity of a spoke-infrastructure building tool. The nodes located at the intersections of source spokes hold data descriptors propagated along all the incoming spokes and practically serve as hash hubs that allow for easy discovery of any advertised event. The successful search is when the sink spoke hits any spoke of the interconnected infrastructure. Hence, with increasing number of data sources (producers), the energy efficiency of data search per source improves. In the simulation snapshot in Figure 10, central source spokes are shown as a sequence of relay transmission *ranges*, to illustrate the fact that the spoke is not only a sequence of forwarded relays, but includes all the overhearing nodes that may be used to replace the current relays, which makes each spoke an *ensemble of possible data routes*. Note that new events (marked by two disks, Gray and Cyan) that occur on other source's spokes do not need to extend their own spokes as they are immediately connected to the advertising infrastructure. Even for a single source, this turns out to be a better approach than a reference search strategy based on directed diffusion [3], as shown in Figures 11 and 12. In both approaches, we calculate per hop cost as free-space energy consumption Cr^2 , where coefficient C is normalized to one, and r is the designed data transmission range. For BeSpoken strategy, we account for every hop performed by both source and sink spokes, while for directed diffusion we count every flooding transmission but leave out the reinforcing packets, and for both we normalize the energy consumption per sink. Also, the simulations are designed in favor of the directed diffusion search, as the search order corresponds to geographic proximity, i.e. the sinks on other sides of the network do not start searching for data until all sinks on one side are done, which makes flooding processes much shorter due to the proximity of already reinforced paths. For the visualization of the directed diffusion based search, we present two consecutive snapshots of a simulation in Figures 13 and 14. The efficiency of this approach also

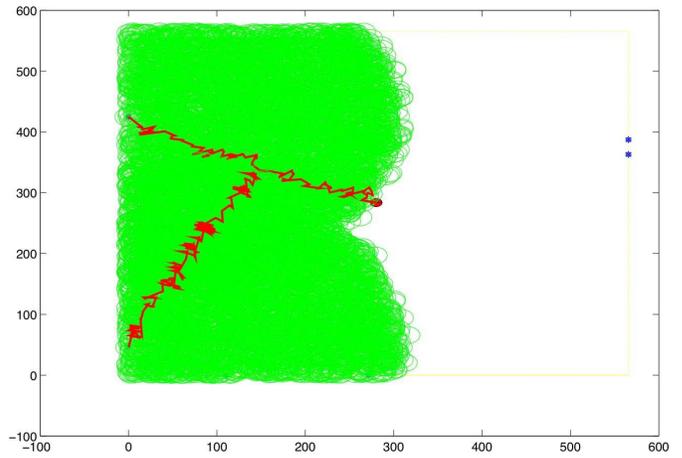


Fig. 14. Search based on directed diffusion: a simulation snapshot showing the area covered by the flooding processes of 10 random sinks, and the reinforced (most-efficient) paths leading from those sinks to the source pointers.

increases with the number of searches. However, if the sinks are searching for different sources, the efficiency does not improve since the reinforced paths are not interconnected as in the case of BeSpoken infrastructure. In general, the losses are large even though each node forwards only once, since the innovative coverage area of each retransmission is small due to high overlap of transmission ranges (note how the flooded area in Figure 13 is almost uniformly green, even though each transmission is represented by one green circle, visible on the edges of the covered area).

Overall, the simulation results confirm the validity of our design, and prove that data search strategies based on the BeSpoken are a scalable and efficient alternative to the existing approaches to search in randomly deployed lightweight WSNs.

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 (22) 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. $Z|Y \sim U(Y, \varphi_{11})$.

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

$$\begin{aligned} E[Y^2] &= m_2, E[Z^2] = 2m_2 + 2E[XY] = m_2 + m\varphi_{11} \\ E[Z^2] &= E_Y E[Z^2|Y] = (1/3)(m_2 + m\varphi_{11} + \varphi_{11}^2) \end{aligned} \quad (35)$$

Solving the system of equations (35) we obtain the second moment of the overshoot $E[Y^2] = \varphi_{11}^2/6$. 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,

$$\begin{aligned} \text{var} [\Theta_{T\varphi_o}] &= 0.5 (2E [(\varphi_o + Y)^2]) \\ &= \varphi_o^2 + (2/3)\varphi_o\varphi_{11} + \varphi_{11}^2/6. \end{aligned} \quad (36)$$

REFERENCES

- [1] Karp B. and Kung H.T. Greedy perimeter stateless routing for wireless networks. In *ACM/IEEE MobiCom 2000*, Boston, Massachusetts, Aug 2000.
- [2] D. Braginsky and D. Estrin. Rumor routing algorithm for sensor networks. In *1st ACM Workshop on WSNA*, Atlanta, Georgia, Sep 2002.
- [3] A. Dembo and O. Zeitouni. *Large Deviation Techniques and Applications*. Springer, 1998.
- [4] C. Intanagonwivat et al. Directed diffusion for wireless sensor networking. *IEEE/ACM Transactions on Networking (TON)*, 11(1):2–16, 2003.
- [5] Q.Fang et al. Glider: Gradient landmark-based distributed routing for sensor networks. In *IEEE INFOCOM 2005*, Miami, Florida, Mar 2005.
- [6] J. Heidemann, F. Silva, and D. Estrin. Matching data dissemination algorithms to application requirements. In *ACM SenSys 2003*, Los Angeles, California, Nov 2003.
- [7] W. R. Heinzelman, J. Kulik, and H. Balakrishnan. Adaptive protocols for data dissemination in wireless sensor network. In *ACM MobiCom 1999*, Seattle, Washington, Aug 1999.
- [8] K. Holger and A. Willig. *Protocols and Architectures for Wireless Sensor Networks*. J. Wiley, 2005.
- [9] S. Kokalj-Filipovic, P. Spasojevic, and R. Yates. Bespoken protocol for data dissemination in wireless sensor networks. In *SpaSWIN 2007 Workshop*, Limassol, Cyprus, Apr 2007.
- [10] S. Kokalj-Filipovic, R. Yates, and P. Spasojevic. Infrastructures for data dissemination and in-network storage in location-unaware wireless sensor networks. Technical report, WINLAB, Rutgers University, 2007.
- [11] S. Kokalj-Filipovic, R. Yates, and P. Spasojevic. Random walk models for geographic data propagation in wireless sensor networks. In *CISS 2007*, Baltimore, Maryland, Mar 2007.
- [12] J. C. Navas and T. Imielinski. Geocast - geographic addressing and routing. In *ACM/IEEE MobiCom 1997*, Budapest, Hungary, Sep 1997.
- [13] S. Ni, Y. Tseng, Y. Chen, and J. Sheu. The broadcast storm problem in a mobile ad hoc network. In *ACM/IEEE MobiCom 1999*, Seattle, Washington, Aug 1999.
- [14] Q.Fang, J. Gao, and L. Guibas. Landmark-based information storage and retrieval in sensor networks. In *IEEE INFOCOM 2006*, Barcelona, Spain, Apr 2006.
- [15] R.Gallager. *Discrete Stochastic Processes*. Kluwer Academic Publishers, 1995.
- [16] H. Takagi and L. Kleinrock. Optimal transmission ranges for randomly distributed packet radio terminals. *IEEE Trans. on Comms.*, 1984.



Silvoja Kokalj-Filipović received the Diploma of Engineering degree in 1989 and M.S. degree in 1995, both from the School of Electrical Engineering, University of Belgrade, Serbia, and Ph.D. degree in electrical and computer engineering from Rutgers University in 2008. Currently she is a Visiting Researcher at Princeton University.

From 1989 to 2002 she worked in the industry as a developer of cutting-edge communication and networking products. Immediately prior to starting her PhD work, her industrial engagements involved both 802.11 and 3G/UMTS RLC and MAC protocol development.

Her research interests include wireless networks, coding techniques, data dissemination and stochastic modeling in communication networks.



Predrag Spasojević received the Diploma of Engineering degree from the School of Electrical Engineering, University of Sarajevo, in 1990; and M.S. and Ph.D. degrees in electrical engineering from Texas A&M University, College Station, Texas, in 1992 and 1999, respectively.

Since 2000, he has been with the Wireless Information Networks Laboratory (WINLAB) and the Electrical and Computer Engineering Department at Rutgers University. Presently, he is an Associate Professor in the ECE Dept.

Dr. Spasojevic was an Associate Editor of the IEEE Communications Letters from 2002 to 2004 and served as a co-chair of the DIMACS Princeton-Rutgers Seminar Series in Information Sciences and Systems 2003-2004. He is currently Publications Editor of IEEE Transactions of Information Theory.



Roy Yates received the B.S.E. degree in 1983 from Princeton and the S.M. and Ph.D. degrees in 1986 and 1990 from MIT, all in Electrical Engineering. Since 1990, he has been with the Wireless Information Networks Laboratory (WINLAB) and the ECE department at Rutgers University. Presently, he is an Associate Director of WINLAB and a Professor in the ECE Dept. He also serves as an associate editor of the IEEE Transactions on Information Theory. He is a co-author (with David Goodman) of the text *Probability and Stochastic Processes: A Friendly*

Introduction for Electrical and Computer Engineers published by John Wiley and Sons. He is a co-recipient of the 2003 IEEE Marconi Prize Paper Award in Wireless Communications and the best paper award for the ICC 2006 Wireless Communications Symposium. His research in wireless networks includes data dissemination, interference mitigation, secret communication, and spectrum regulation.