Spatial Energy Balancing Through Proactive Multipath Routing in Wireless Multihop Networks

Seung Jun Baek, *Student Member, IEEE*, and Gustavo de Veciana, *Senior Member, IEEE*

***Abstract*—In this paper, we investigate the use of proactive multipath routing to achieve energy-efficient operation of ad hoc wireless networks. The focus is on optimizing tradeoffs between the energy cost of spreading traffic and the improved spatial balance of energy burdens. We propose a simple scheme for multipath routing based on spatial relationships among nodes. Then, combining stochastic geometric and queueing models, we develop a continuum model for such networks, permitting an evaluation of different types of scenarios, i.e., with and without energy replenishing and storage capabilities. We propose a parameterized family of energy balancing strategies and study the spatial distributions of energy burdens based on their associated second-order statistics. Our analysis and simulations show the fundamental importance of the tradeoff explored in this paper, and how its optimization depends on the relative values of the energy reserves/storage, replenishing rates, and network load characteristics. For example, one of our results shows that the degree of spreading should roughly scale as the square root of the bits · meters load offered by a session. Simulation results confirm that proactive multipath routing decreases the probability of energy depletion by orders of magnitude versus that of a shortest path routing scheme when the initial energy reserve is high.**

***Index Terms*—Gaussian random field, $M/GI/1$ queue, sensor networks, shot-noise process, stochastic geometry.**

I. Introduction

ENERGY-EFFICIENT design and operation of ad hoc multihop wireless networks is a key problem in the context of mobile and/or distributed sensing applications, where energy storage and availability may be quite limited. There are many levels at which one can address this problem. Advances in silicon technology can realize energy savings through power efficient circuitry, e.g., voltage scaling, while specialized architectures can be devised to allow components to enter "sleep" modes. At the same time power control and optimized MAC protocols which put nodes to sleep can bring substantial energy savings enabling networks with thousands of sensors. Particularly, in large ad hoc wireless networks the data originated from a source might need to be relayed a *long distance* to a destination or sink wireline node. Relaying through many hops may cause intermediate nodes to consume substantial amounts of energy and thus make energy-efficient routing a particularly critical task.

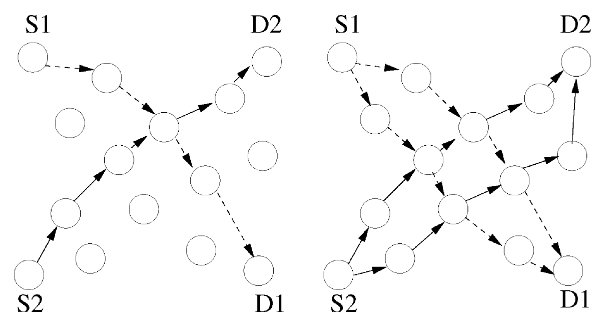


Fig. 1. Comparison of the shortest path routing scheme (on the left) and a typical load-balancing scheme (on the right). The dotted arrows represent flows for $S1$–$D1$ and the dashed arrows represent flows for $S2$–$D2$, respectively.

Consider the network shown in Fig. 1. Two sources $S1$, $S2$ send to destinations $D1$, $D2$ on opposite ends of the network, respectively. In the network on the left these sessions are supported along shortest hop routes. If one of these sessions were sustained for a long time, nodes along the route would eventually see depleted energy reserves, roughly "dividing" the network into two parts. Subsequently, if other nodes needed to communicate across this depleted zone they may result in exhaustion of energy along the diagonal, or require selection of routes around this area of the network, which in turn would incur additional energy burdens. This simple example shows how energy depletion along long routes combined with interactions with future overlapping and/or routing of additional traffic flows might exacerbate the energy problem. A natural solution to this problem is to spread out the energy burden of sustained sessions so as to obtain a spatially balanced energy burden. Specifically, one may split traffic across two disjoint routes as shown on the right in Fig. 1. Assuming energy consumption is roughly proportional to the load this leads to a more balanced energy burden across sets of intermediate nodes. At the same time this scheme may involve a larger number of nodes, e.g., a route with five versus four hops, and thus an increased overall energy burden.

In this paper, we consider overall system design aspects for such multipath routing strategies—we refer to this as *proactive* balancing of energy burdens over multiple routes. Our primary interest here is not to devise detailed multipath routing algorithms, but rather to investigate the design of, and possible improvements afforded by, such routing mechanisms. The key intuition is that the more we spread the traffic, the more the energy profile of the network will be balanced. However, spreading traffic requires that some packets take long "detours", which will incur extra energy cost. This tradeoff associated with the degree of spreading is the main topic investigated in this paper.

Manuscript received March 4, 2005; revised October 4, 2006; approved by IEEE/ACM Transactions on Networking Editor N. Shroff. This work was supported in part by the National Science Foundation under Grant ECS-0225448. This work was presented in part at IEEE INFOCOM 2005, Miami, FL.

The authors are with the Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712 USA (e-mail: sbaek@ece.utexas.edu; gustavo@ece.utexas.edu).

Digital Object Identifier 10.1109/TNET.2006.890102

To this end, we use a simple, idealized model to characterize and parameterize the spatial energy balancing aspects of proactive multipath routing schemes. Our model provides sharp insights on design choices under various scenarios. For example, one of the key issues studied in this paper is the degree to which a session's traffic should be spread, depending on the load and the distance it must travel. Not surprisingly, we show that traffic should be spread *more* as the load and the hop count *increase*, and provide a simple scaling rule to proactively adapt the degree of spreading.

The paper is organized as follows. In Section II, we discuss related work in this area. Section III introduces a concrete multipath routing and balancing strategy. In Section IV, we present a continuum model and characterize spatial energy burdens using a shot-noise process associated with the model. Section V uses a grid network to explicitly analyze a parameterized family of energy balancing strategies. In Sections VI and VII, we formulate and investigate the design and optimization of such spreading using second-order and asymptotic approximations. Section VIII includes simulation results and a discussion of various scenarios. Finally, Section IX presents our conclusions.

II. Related Work

There has been substantial research on the design and implementation of energy conserving routing protocols suitable for ad hoc networking applications. Let us review some of this work. In [1], a characterization and algorithm determining the most energy-efficient route between two nodes is proposed. However, it is not clear whether using such routes will extend network lifetime, nor how this would impact network capacity for nonhomogenous traffic loads. By contrast, [2] and [3] propose and evaluate routing mechanisms to maximize network lifetime based on nodes' current residual energy reserves, whereas scalability and the effectiveness of greedy routing to spread energy burdens are a concern. Among recent work, [4] and [5] show that, by properly defining cost metrics as exponential functions of the residual energy at each node, one can achieve competitive optimality for throughput under energy constraints. The work of [6] takes yet another tack—they propose packet-level randomized routing in order to proactively balance energy burdens across the network. A unifying principle emerges from this body of work: the tradeoff between minimizing the energy expended to carry an offered load versus the balancing of energy burdens across the network. To the best of our knowledge, the spatial character of this tradeoff has not been studied. The primary contribution of this paper is the use of a stochastic geometric framework to analyze, and then work towards realizing this tradeoff in an "optimal" manner.

III. Spatial Modelling

This section is divided into three parts. We start by stating our model assumptions. We then introduce a multipath routing scheme based on nodes' spatial relationships. Finally, we propose a continuum model where we regard the field of the wireless nodes as an infinitesimal "medium" that carries fluid, i.e., the traffic flow. This leads to a simple shot-noise process model for the spatial field of energy expenditures which is amenable to analysis.

A. Model Assumptions

We will use a simplified model for energy expenditures associated with data transmissions. Nodes are assumed to share a common transmit power level sufficient to guarantee the network is connected. They relay packets towards the destination via *neighboring* nodes in a hop-by-hop manner. We refer to a flow of traffic between a pair of source–destination nodes as a *session*, and refer to the source and destination nodes as a *session pair*.

A key assumption in this paper will be that the energy consumption at each node is roughly *proportional* to the traffic it is carrying, so we use the terms "traffic load" and "energy burden" interchangeably. This linear relationship is assumed to be satisfied with a common proportionality constant for each node.

In reality, there are many factors, in addition to the traffic load, that contribute to the energy burdens in a wireless network, e.g., interference and channel contention. Due to such factors our assumption may not hold in the following situations.

- The proportionality constant may vary, e.g., some nodes may experience larger energy burden per load if transmission ranges and the amount of local interference are higher.
- The relationship between traffic load and energy burden can be nonlinear, e.g., depending on protocols such as CSMA, loads beyond capacity may lead to throughput collapse.
- The number of paths used for multipath routing schemes may affect the interference level, e.g., when multiple paths are geographically clustered, the routes may interfere with each other versus "single" path routing.

Nevertheless, by properly defining the proportionality constant, this seems to be a reasonable model to capture the first-order relationship between energy expenditure and traffic for various types of wireless networks. We assume the network satisfies *homogeneity* conditions such that the average energy burden per unit load is roughly constant over time, and the traffic is not severely limited by interference. Moreover, we are interested in making a *comparative* analysis of proactive multipath routing schemes with different numbers of multipaths, for example one can consider a scheme which transmits packets alternating among multiple routes in a time-division manner so that inter-route interference among neighboring routes is limited. Finally, our assumption of a linear relationship between energy burden and carried load is intended to allow the study of the longer term burdens accrued under various routing approaches. As such the proportionality constant need not reflect the exact instantaneous relationship but rather some approximate longer term relationship between load and energy burden.

We assume traffic is relayed only via neighboring nodes which we shall define based on proximity as follows. We model the locations of the nodes as fixed and following a spatial point process in the $\mathbb{R}^2$ plane. A natural notion of proximity can be introduced via the Voronoi tessellation and Delaunay graphs induced by the locations of the nodes. These are discussed below.

Suppose the locations of the nodes constitute a point process Ψ on the $\mathbb{R}^2$ plane . Each point $x_i \in \Psi$ serves as a *seed* for a *cell* $V(x_i)$

$$V(x_i) = \{y \mid |x_i - y| \leq |x_j - y| \quad \forall\, x_j \in \Psi\}$$

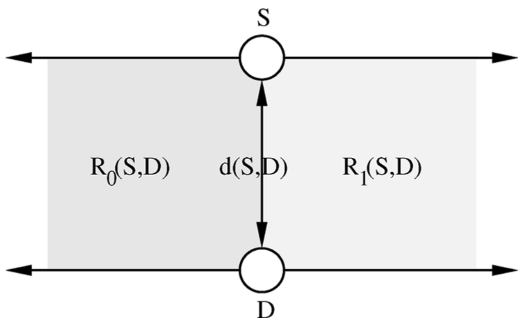


Fig. 2. Illustration of regions $R_0(S,D)$ and $R_1(S,D)$ for the source–destination pair (S,D).

in the Voronoi tessellation induced by Ψ. If $V(x_i) \cap V(x_j)$ is not an empty set, we refer to $V(x_i)$ and $V(x_j)$ as *neighboring cells* and we say that x_i and x_j are neighbors.

A *Delaunay graph* is a graph whose vertex set is Ψ and whose edges connect nodes that are neighbors. We denote the Delaunay graph by $G(\Psi, E)$ where E is the set of Delaunay edges. Routes considered in the discussion below will be based on the Delaunay graph. We shall assume that the spatial distribution of nodes is fairly uniform and sufficiently dense that each node can in fact reach its neighbors.

B. Proximity-Based Multipath Routing

Consider a route connecting two nodes $x_i, x_j \in \Psi$, which has a minimal length, i.e., sum of the Euclidean lengths of the edges it traverses. This path is referred to as the *Shortest Delaunay Route* (SDR) and has a length that is within a factor of 2.42 of the Euclidean distance between x_i and x_j, see, e.g., [7], [8]. Note that the SDR is based on the Euclidean norm, thus the SDR may not coincide with a route having a minimum number of hops. We will see in the sequel (Section VIII) that this subtle difference may impact the *spatial distribution* of energy expenditures significantly. Based on the SDR, we propose the following simple construction for a set of paths between two nodes, say $S, D \in \Psi$.

1) Draw a straight line segment $d(S,D)$ between S and D, and draw two additional lines, through S and D and orthogonal to $d(S,D)$. Let $R_0(S,D)$ and $R_1(S,D)$ denote the open planes with their boundaries being $d(S,D)$ and its orthogonal lines as shown in Fig. 2.
2) We let $N_1(S,D)$ denote the set of nodes included in the SDR from S to D, and refer to this route as the Level 1 route. S and D are referred to as the *Level-1 connectors*.
3) Find the set of nodes $N_2(S,D)$ which are neighbors of $N_1(S,D)$ and fall in $R_0(S,D)$. Among the nodes in $N_2(S,D)$ select two nodes each of which is closest[1] to S and D, respectively, and refer to them as *Level-2 connectors*. Create a route that connects the nodes in $N_2(S,D)$ with Delaunay edges where the endpoints of the route are Level-2 connectors.
4) Construct a SDR for a Level-2 connector to its closest Level-1 connector and repeat the same for the other connector. If this SDR crosses new nodes, update $N_2(S,D)$ by adding these new nodes in $N_2(S,D)$. Now the nodes in $N_2(S,D)$, S and D can be connected via a Delaunay route which is referred to as *Level-2 route*.
5) Next, determine the set of nodes $N_3(S,D)$ which are neighbors of $N_1(S,D)$ but falls in $R_1(S,D)$ this time. Following the similar process as above, find the Level-3 connectors, construct SDRs from these connectors to Level-1 connectors, update $N_3(S,D)$. This constructs the Level-3 route.
6) For $w \geq 4$, the Level-w route can be constructed in a similar manner. Determine the new set of nodes $N_w(S,D)$ that are neighbors of N_{w-2} and fall in $R_{(w \bmod 2)}(S,D)$. Find the Level-w connectors, construct SDRs from these connectors to Level-$(w-2)$ connectors and update $N_w(S,D)$.

The basic idea is to recursively construct higher level routes based on nodes which are neighbors of those included in previous levels but alternating between $R_0(S,D)$ and $R_1(S,D)$ in order to balance the spreading cost as the levels increase. We confine relaying nodes to the regions $R_0(S,D)$ and $R_1(S,D)$ so as to prevent routes from extending backward.[2] As will be clear from the construction, the role of connectors is to ensure connectivity via Delaunay routes among S, D and the routes at different levels. These routes can be constructed by each node if it has the information on the locations of its neighbors, the source and the destination.

An example of our route construction for a source–destination pair (S,D) is illustrated in Fig. 3. At each level, the level connectors are marked with arrows. Note that at level 3, the level connector associated with node D is not its neighbor (see dotted line), thus an SDR route is constructed between them. Also, note that this route includes a node, indicated by a star, which is in fact not contained in $R_1(S,D)$.

We shall refer to this construction as *proximity-based multipath* (PBM) routing which gives us a concrete set of paths over which to distribute traffic so as to spread out energy burdens.

As shown in Fig. 3, the set of PBM routes associated with a session is *spatially clustered* by construction. We shall refer to the set of nodes in such a cluster as the *spatial footprint* of a session. For example, the nodes within the shaded Voronoi cells in Fig. 3 correspond to the spatial footprints of the session as we increase the degree of spreading. Thus, the traffic pattern on a network can be viewed as a dynamic set of, possibly overlapping, spatial footprints. This is the basic idea behind the *continuum* model introduced in the next section.

IV. Characterization via Continuum Model

A. The Continuum Model and Shot-Noise Formulation

Consider the case where the density of the nodes in a network is large. We can think of an infinitesimal area in space as corresponding to a node with an initial energy reserve. Each footprint will correspond to a closed set in $\mathbb{R}^2$, and is assumed to have a well-defined, possibly random, shape. Since an infinitesimal area (node) serves as a "carrier" of a flow, a footprint can be regarded as a "vessel" which contains the flow from a source to a sink. Session pairs are located at the ends of footprints. We refer to the "length" of a footprint which corresponds to the distance

[1]We perform random tie-breaking if there are multiple nodes with the same closest distances.

[2]In step 4) of the construction, some nodes that are not contained in these regions may be included in a route.

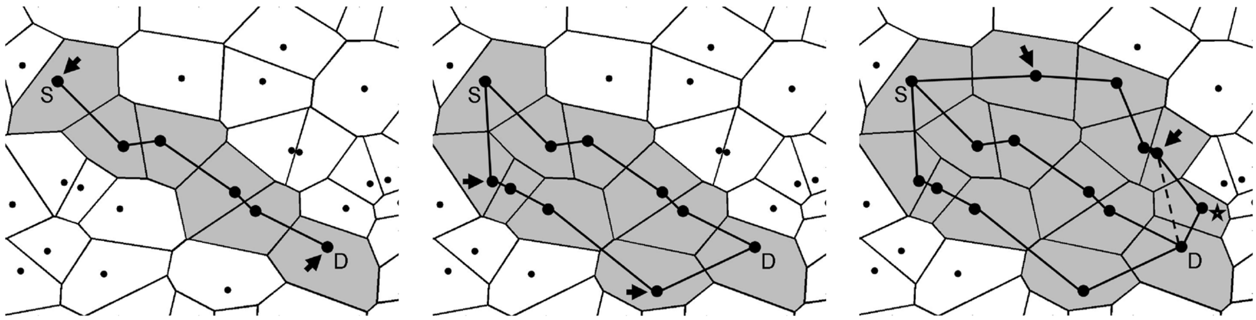


Fig. 3. Construction of Level 1, 2, and 3 routes for nodes S and D, from left to right. The route at each stage is shown with solid lines and the shaded regions are the cells for the nodes in routes. The nodes marked by an arrow are the *connectors* at each route.

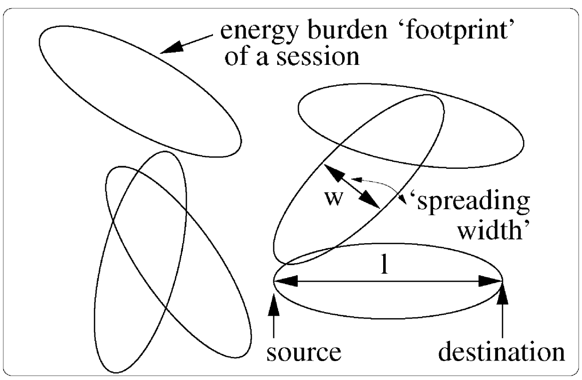


Fig. 4. A realization of the energy footprints for sessions in an ad hoc network. A footprint is assumed to have an elliptical shape for the purpose of illustration.

between a session pair as the *span*, and the maximum "width" of a footprint as the *spreading width*. We refer to the midpoint of the session pair as the *center* of the associated footprint. We assume that the centers of footprints constitute a point process in $\mathbb{R}^2$. Fig. 4 exhibits a realization of the process capturing the energy burdens incurred over a period of time—only sessions' footprints with span l and spreading width w are shown.

Each location within a footprint would, in general, experience a different load/energy burden. To model this, we define a *load distribution* function which is the load per unit area, i.e., the energy burden *density* at each location, in a footprint. The load distribution function depends on the "strategy" used to spread traffic within a footprint—this will be further quantified in the sequel.

The proposed continuum model can be mathematically formalized as follows. Let Φ_0 denote an isotropic, random closed subset of $\mathbb{R}^2$. We will assume Φ_0 has the distribution of a typical footprint in a given network. The load distribution function $h(\cdot, \Phi_0) : \mathbb{R}^2 \to \mathbb{R}_+$ gives the spatial density of energy burden at each location, for a session with unit load. Specifically, for $y \in \mathbb{R}^2$, if $y \notin \Phi_0$ then $h(y, \Phi_0) = 0$ and otherwise $h(y, \Phi_0)$ corresponds the relative energy burden per unit area at location y of the footprint Φ_0.

We assume that the centers of sessions constitute a homogeneous spatio-temporal Poisson point process Π with intensity λ per unit time per unit area. Let us denote by Π_t a *spatial* point process in $\mathbb{R}^2$ for the centers of the sessions/footprints that have been been offered to the network during time $[0, t]$. Thus, Π_t is a homogeneous spatial Poisson point process with intensity of λt. Each point $x_i \in \Pi_t$ has an associated footprint denoted by Φ_i. We assume $\{\Phi_i\}$ are i.i.d. copies of Φ_0. The contribution of the energy burden on location x from a session centered at x_i with footprint Φ_i, is given by $h(x - x_i, \Phi_i)$. Note since we equivocate load and energy burden, $h(\cdot, \Phi_i)$ depends on how a routing mechanism chooses to spread the flow of session i within its footprint Φ_i. For now, assume that the strategy $h(\cdot, \Phi_i)$ does not depend on the amount of load Φ_i is carrying, and denote the offered load, in bits, by U_i – these are assumed to be i.i.d. with the same distribution as U. Thus, $h_{U_i}(\cdot, \Phi_i) := U_i h(\cdot, \Phi_i)$ gives the spatial load density of session i.

The total energy burden per unit area accumulated at location $x \in \mathbb{R}^2$ during time $[0, t]$ can be represented as a *shot-noise process* as follows:[3]

$$G(x, t) = \sum_{x_i \in \Pi_t} h_{U_i}(x - x_i, \Phi_i). \tag{1}$$

Next, we state several known results from shot-noise theory. Since Π_t is stationary, we can consider a typical location at the origin based on the following result.

Lemma 1: (See [9].) Let us define $G_0(t) := G(O, t)$, the energy burden density at the origin. Also, let $\chi^{(n)}(t)$ be the nth order cumulant of $G_0(t)$. Since Π_t is a homogeneous Poisson process with intensity λt, we have that

$$\chi^{(n)}(t) = \lambda t \, E\left[\int_{\Phi_0} h_U(x, \Phi_0)^n dx\right].$$

Defining the normalized mean μ and variance σ^2 as

$$\mu := E\left[\int_{\Phi_0} h_U(x, \Phi_0) dx\right] \tag{2}$$

$$\sigma^2 := E\left[\int_{\Phi_0} h_U(x, \Phi_0)^2 dx\right] \tag{3}$$

we have that

$$E[G_0(t)] = \lambda t \mu, \text{ and } \mathrm{Var}[G_0(t)] = \lambda t \sigma^2.$$

We shall for now assume the load U equals 1 with probability 1, i.e., $h_{U_i}(\cdot, \Phi_i) = h(\cdot, \Phi_i)$, which we will revisit in Section VI.

[3]Precisely, if the session i has not ended by time t, the offered load by that session will be less than U_i. Thus, we assume that each load is offered in an *instantaneous* manner – we are interested in the *cumulative* spatial loads induced by footprints.

As mentioned earlier, the function $h(\cdot, \Phi_0)$ captures both the "shape" and how the flow is spread within a typical footprint—these are the design choices one can make to control the mean and variance of the spatial energy burdens. Although using only two moments to describe the statistical properties of $G_0(t)$ may not be sufficient, the following theorem suggests this might give a good approximation (see [9]).

Theorem 1: (Asymptotic normality of shot-noise process) *Consider $G_0(t)$ defined in Lemma 1. We have that*

$$\frac{G_0(t) - \lambda t \mu}{\sqrt{\lambda t}\sigma} \xrightarrow{d} N(0,1) \text{ as } \lambda t \to \infty$$

where $N(0,1)$ is the standard normal distribution.

From this theorem, we have that, for large λt, the probability that the energy burden per unit area exceeds a prescribed level b is given by

$$P(G_0(t) > b) \simeq \phi\left(\frac{b - \lambda t \mu}{\sqrt{\lambda t}\sigma}\right) \tag{4}$$

$$\phi(u) := \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-v^2/2} dv. \tag{5}$$

In order for this approximation to be useful, we assume a typical node in the network sees a large number of overlapping footprints on average.

B. Depletion Probability and Network Lifetime

A common criterion for the energy performance of a network is its lifetime, e.g., the time before some fraction of nodes (or any single node) drop(s) below a certain battery level. Our objective lies in the complementary question: given a *desired network operation time*, can we minimize the fraction of the depleted nodes? For example, if one wishes to operate a sensor network for a week, what fraction of nodes might survive the week and what is a good multipath routing strategy to achieve this? We believe this to be a practical objective in engineering such networks. To address this question, we shall use the approximation in Theorem 1.

Let τ be the desired network operation time, and assume $\lambda = 1$. Suppose the critical reserve level b per unit area is specified as a multiple k of $\tau\underline{\mu}$ where $\underline{\mu}$ is defined as the mean energy consumption of the *baseline* scheme, i.e., a scheme without multipath routing. Thus, b is specified in terms of a factor k times the mean energy consumption of baseline scheme during τ. Thus, by letting $b = k\tau\underline{\mu}$, and by defining $z_k(\tau)$ as

$$z_k(\tau) := \sqrt{\tau}\, \frac{k\underline{\mu} - \mu}{\sigma} \tag{6}$$

we can estimate the fraction of nodes that have not depleted the critical level b by time τ by $\phi(z_k(\tau))$. To reduce the likelihood of depletion we wish to maximize $z_k(\tau)$ for a given τ, i.e., minimize the probability of depletion $\phi(z_k(\tau))$ through admissible choices for μ and σ.

Equation (6) provides us with crucial insights. Certainly, we would like to minimize both μ and σ, however as we will see later, there is a *tradeoff* between these parameters, i.e., we can decrease μ at the cost of increasing σ, and vice versa. The optimal tradeoff will depend on k. If $k\underline{\mu}$ is small, one might try to decrease μ. Conversely, if k is relatively large, one might prefer strategies that give smaller σ. This captures the fundamental tradeoff addressed in this paper, i.e., that between the benefit of spatial energy balancing by traffic spreading (σ) and the cost associated with such spreading (μ). To this end, in the following section we discuss parameterized energy balancing strategies for which μ and σ can be estimated explicitly.

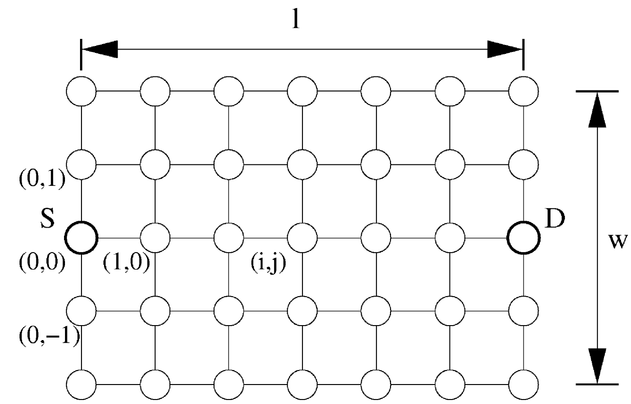


Fig. 5. Topology of a regular grid footprint. The coordinates of locations are shown for some nodes in their lower left corners. The source and the destination is marked by S and D, respectively, and the dimensions l and w of the grid are shown.

V. Proactive Energy Balancing Strategies

We will study strategies that give desirable footprints and flow distributions over the footprints, $h(\cdot, \Phi_0)$ so as to reduce $\phi(z_k(\tau))$. For simplicity, we assume the spans of footprints are *fixed* to l. Also, we let the loads offered by sessions be fixed to 1 and treat the maximum footprint width w as a design parameter. We will deal with the issues involving random spans and session load in later sections.

While it is possible to take an approach in continuum flow domain by making analogy with incompressible ideal fluid problems; see [10] for details, this approach requires numerical methods to estimate μ, σ, and $z_k(\tau)$. To obtain intuitive and closed-form results, we take another approach by considering parameterized spreading strategies over a *regular grid*.

Let us consider a session pair in a network. We assume that the intermediate relaying nodes form a regular grid topology,[4] as shown in Fig. 5. The hop count between the session pair is fixed to l. The source emits one unit of flow which is distributed among the intermediate nodes to reach the destination, and a flow can be relayed only among adjacent nodes. A *grid footprint* is defined to be the set of grid nodes which carry nonzero flow for a given session pair. Accordingly, a *spreading strategy* corresponds to assigning the flow rates at the intermediate nodes such that the total flow is conserved. These are analogous to Φ_0 with fixed span and spreading width, and $h(\cdot, \Phi_0)$, respectively.

We refer to the maximum degree, in number of hops, to which the flow is spread as the *spreading width* w. Our main goal is to determine "good" spreading strategy and grid footprint for given l and the constraint w. In particular, we will consider the *minimum variance* strategy specified as follows.

[4]This serves as a coarse approximation for a dense, uniform network. Also, note its similarity with the topology induced by PBM construction discussed in Section III, e.g., those shown in Fig. 3.

Let us label each node with an integer coordinate (i,j) placing the source node at the origin as shown in Fig. 5. The total *outgoing* flow rate at node (i,j) is denoted by $e_{i,j}$ which will also represents the energy consumed at node (i,j). Define the set $H := \{(i,j) \mid 0 \le i \le l,\ |j| \le (w-1)/2\}$. Under the above construction, we have that

$$\mu = E\left[\int_{\Phi_0} h(x,\Phi_0)dx\right] \approx \sum_{(i,j)\in H} e_{i,j} \tag{7}$$

$$\sigma^2 = E\left[\int_{\Phi_0} h(x,\Phi_0)^2 dx\right] \approx \sum_{(i,j)\in H} e_{i,j}^2 \tag{8}$$

where the continuum footprint Φ_0 corresponds to a grid footprint within H, i.e., $\{(i,j) \mid e_{i,j} > 0,\ (i,j) \in H\}$. Similarly, the load per unit area $h(\cdot,\Phi_0)$ corresponds to the flow rates $e_{i,j}$ associated with a grid footprint. We wish to minimize (8) by properly setting $e_{i,j}$ for given w and l.

The rationale behind the minimum variance strategy is that, in order to maximize $z_k(\tau)$, it is an efficient strategy to minimize σ assuming that k is moderately large, i.e., all the nodes have large energy reserves at the beginning so they would not suffer from depletion when only a few sessions overlap at a given location.

Before we formally define the flow distribution problem for the minimum variance, let us assume that l and w are even and odd integers, respectively, for symmetry purposes and $w \le l$, i.e., traffic is not spread beyond l. We also add an extra constraint that *the flow rates are equal along the middle abscissa (MA)*. The MA is defined to be the nodes that lie on the middle of the horizontal span of H, i.e., the nodes along $i = l/2$ axis in Fig. 5. We refer to this suboptimal approach as *equi-flow minimum variance* strategy. The implications of the additional constraint will be explained in the sequel.

1) Problem 1: Equi-flow minimum variance problem.

$$\begin{aligned}
&\text{Minimize:} \sum_{i,j\in H} e_{i,j}^2, && (9)\\
&\text{Constraints:} && \\
&\qquad \text{flow is conserved at each node,} && (10)\\
&\qquad e_{i,j} \ge 0, && \\
&\qquad e_{0,0} = 1, && (11)\\
&\qquad \max\{|j| \mid e_{i,j} > 0,\ (i,j) \in H\} && \\
&\qquad = (w-1)/2, && (12)\\
&\qquad e_{l/2,k} = \frac{1}{w},\ (|k| \le (w-1)/2), && \\
&\text{Variables:}\ e_{i,j}. && (13)
\end{aligned}$$

where (11) is the source emitting 1 unit of flow, (12) corresponds to the definition of the spreading width and (13) is the equal flow constraint at the MA.

The solution can be explicitly obtained by exploiting the convexity of the objective; see the Appendix .

Lemma 2: The solution to Problem 1 is, for all $(i,j) \in H$,

$$e_{i,j} = \begin{cases} \{2(|i|+|j|)+1\}^{-1}, & |i|+|j| \le \frac{w-1}{2} \\ \{2(|l-i|+|j|)+1\}^{-1} & |l-i|+|j| \le \frac{w-1}{2}, \\ w^{-1}, & \text{otherwise.} \end{cases}$$

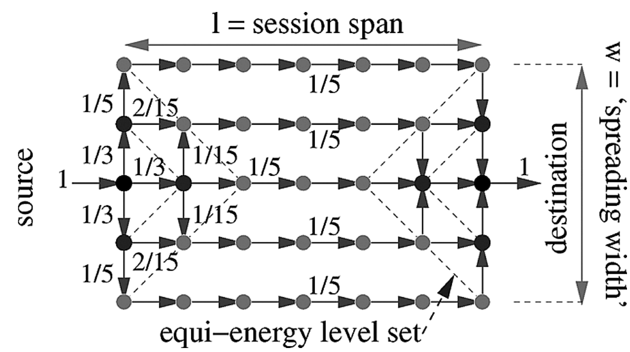


Fig. 6. Flow allocation under equi-flow minimum variance strategy when $w = 5$ and $l = 7$.

An example of such a flow assignment is illustrated in Fig. 6. The dotted lines are contours that represent level sets of nodes which have the same total flow. The value of the levels decreases harmonically, i.e., $(1, \frac{1}{3}, \frac{1}{5}, \ldots)$ as the contour expands outward from the source. Note $e_{i,j}$ is the sum of horizontal and vertical outgoing flows at location (i,j) in the figure.

We comment on the additional constraint. Without (13), the minimum variance flow allocation can be obtained using standard optimization techniques such as the projected gradient method [11]. However, with (13), the problem yields a closed-form, simple and intuitive solution. Moreover, it gives an excellent approximation especially for large values of l. In fact, one can show that the solution obtained for the suboptimal scheme converges to the optimal solution when l is large [10].

Given the flow assignment in Lemma 2, we have that

$$\mu \approx \sum_{(i,j)\in H} e_{i,j} = l + \frac{w}{2} - \frac{1}{2w} \tag{14}$$

$$\sigma^2 \approx \sum_{(i,j)\in H} e_{i,j}^2 = \frac{l}{w} - \frac{1}{2}\left(1+\frac{1}{w^2}\right) + \sum_{k=1}^{\frac{w+1}{2}} \frac{2}{2k-1}. \tag{15}$$

The following observations are in order. The grid footprint suggested by the solution is such that it corresponds to H, i.e., $e_{i,j} > 0$ for all $(i,j) \in H$, which in turn means that the grid footprint for the equi-flow minimum variance strategy has a rectangular shape. Recall that we imposed equal flow constraints on MA parameterized by its width w. However, our solution indicates that the flow should be spread as much as possible within H. For $1 \le w \le l$, we see that μ increases with w, i.e., the mean energy increases with the spreading width, indeed the flow will travel longer distances. However, σ decreases with w, i.e., the variance decreases with the degree of spreading. Thus, as mentioned earlier, one cannot minimize the mean and the variance simultaneously. The interesting case is where l is much larger than w. Then the mean energy is roughly invariant to small changes in w, but the variance is sensitive to w—the dominant term is l/w. Thus, for a load that traverses a long route, the benefit in terms of reducing variance from spreading is large. For a more detailed discussion, see [10].

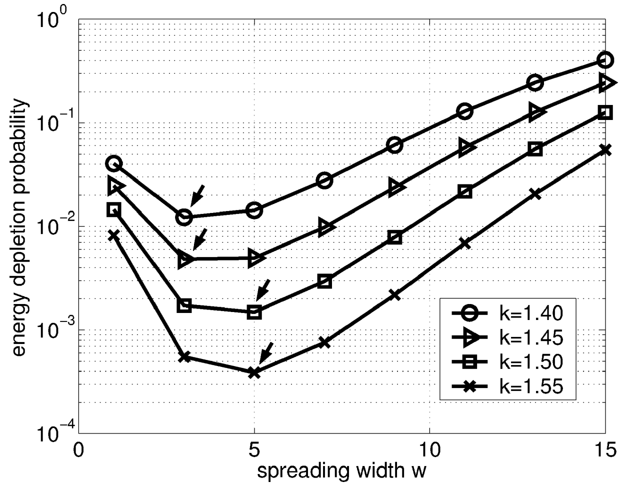


Fig. 7. A numerical evaluation of optimal design of spreading width under equi-flow minimum variance strategy. The session span l is fixed to 20 and the initial energy reserve parameter k is varied. Note the change in the tradeoff point indicated by arrows, i.e, the optimal w moves from 3 to 5 with increasing k.

VI. Design Tradeoffs: Networks Without Energy Replenishing Capability

A. Depletion Probability of the Typical Node

In this section, we numerically evaluate the depletion probability of a typical node, combining the estimates in (14) and (15) with $z_k(\tau)$. Here τ is assumed to be 1, l is set to 20 and $\underline{\mu} = l$ since it is the mean energy consumption without spreading – see (14) for the case where $w = 1$. Fig. 7 exhibits a plot of $\phi(z_k(\tau))$ for varying k, i.e., varying the initial energy reserves of network nodes versus the spreading width w. Clearly, there exists a w that minimizes the depletion probability for each k. As expected, for the case where nodes have high initial energy reserves, the optimal w gets larger. The intuition is that, whenever the nodes in the network have large residual reserves, they should cooperate to balance load on the network, i.e., the number of nodes participating in carrying a flow should increase, but up to a degree where the energy cost of load balancing does not overload the network.

B. Depletion Probability for a Network

One can also approximate the spatial energy burden pattern as a stationary, isotropic Gaussian random field in $\mathbb{R}^2$. Consider $G(x,t)$ in (1) for $x \in \mathbb{R}^2$. Let us model a network as occupying a 'nice' subset A in $\mathbb{R}^2$, e.g., a rectangle or circle [12]. Consider the probability that the node with the highest energy burden in A exceeds a prescribed level b by some time τ, i.e.,

$$P(\sup_{x \in A} G(x,\tau) > b). \tag{16}$$

We can estimate the asymptotic value of this probability as $b \to \infty$ via extreme value theory for homogeneous Gaussian fields, see [12] and [13]. Consider the normalized energy burden density $Z(x,\tau) := (G(x,\tau) - \tau\mu)/\sqrt{\tau}\sigma$ where $\lambda = 1$. Let us define the normalized spatial covariance function $r_\tau(y)$ at time τ by $r_\tau(y) = E[Z(x+y,\tau)Z(x,\tau)]$ for $y \in \mathbb{R}^2$. Since the field is isotropic, this function depends only on the norm of y, denoted by $|y|$. Suppose that the following holds for some positive constant a:

$$r_\tau(y) \approx 1 - a|y|^\alpha \quad \text{as } |y| \to 0.$$

Here α denotes the infinitesimal order of decay of the covariance with the magnitude $|y|$. Again, assume that b is given by $k\underline{\mu}\tau$ where k is large, then based on the Poisson clumping heuristic [12], we can rewrite and approximate (16) as

$$P(\sup_{x \in A} Z(x,\tau) > z_k(\tau)) \approx H_\alpha |A| a^{2/\alpha} \{z_k(\tau)\}^{4/\alpha} \phi(z_k(\tau))$$

where $|A|$ is the area of the region A and $H_\alpha > 0$ is the two-dimensional Pickand's constant which depends only on α. We see that the depletion probability is proportional to the physical area of the network, and is related to the covariance structure of the footprints.[5] Comparing this with the result for the typical node, we note that they share the term $\phi(z_k(\tau))$, but there is the extra term $\{z_k(\tau)\}^{4/\alpha}$ which may increase the depletion probability for large values of $z_k(\tau)$.

Thus, we expect that, by increasing w, we will observe a similar tradeoff curve as obtained for the typical node (Fig. 7), but the curve will be *flatter* due to the term $\{z_k(\tau)\}^{4/\alpha}$, which is shown in [10].

C. Optimal Choice of w When Sessions Have Different Spans and Varying Loads

A natural question arises: is it beneficial to change the spreading width w if the distance l between session pairs varies? An intuitive answer is that, if the span of a session is high, it helps to spread more, i.e., choose larger w. The rationale is that sessions spanning longer distances have an increased chance of overlapping with other sessions. Thus, it is more important to reduce hotspots induced by such sessions.

Further, we can also consider the impact of the *variability in traffic loads* associated with sessions should have on the spreading width. If a session has higher traffic loads, one can expect that we should spread *more*. In this section, we investigate how the optimal spreading width should scale with the span and traffic loads in a network/traffic model where these are indeed variable.

Let L be a random variable whose distribution is that of the span for the typical session pair and the distribution of variable loads be that of a random variable U. We have mild assumptions for L and U such that they have finite support and are independent of each other. Let us capture a strategy for choosing the spreading width depending on L and U by some function $g(L,U)$. From the previous arguments one would expect the optimal spreading width to be a *nondecreasing* function of L and U. Using our model we shall determine $g(\cdot,\cdot)$ in an approximate manner.

Considering only the dominant terms in (14) and (15), we will use the following approximations:

$$\mu \simeq l + \frac{w}{2}, \quad \sigma^2 \simeq \frac{l}{w} \tag{17}$$

[5]For the estimation results for a, α versus varying w, see [10].

for the case where l is typically much greater than w. With the above approximations and setting $\tau = 1$, we have that

$$z_k(1) \simeq \frac{kE[L]E[U] - E\left[U\left(L + g(L,U)/2\right)\right]}{\sqrt{E\left[U^2 L/g(L,U)\right]}}.$$

The goal is to find an optimal choice for $g(\cdot,\cdot)$, i.e., one that maximizes $z_k(1)$ and thus minimizes the depletion probability of a typical node.

Assume that L is continuously[6] distributed over a finite support $A := [l_{\min}, l_{\max}]$ and the probability density is nonzero and smooth on that support. Let us partition A into n subintervals $A_i := [l_i, l_{i+1}]$, $0 \le i \le n-1$, where $l_{\min} = l_0 < l_1 < l_2 < \cdots < l_n = l_{\max}$. Similarly, suppose U has a finite support $B = [u_{\min}, u_{\max}]$, define $B_j := [u_j, u_{j+1}]$, $0 \le j \le m-1$ for some integer m. Suppose $g(\cdot,\cdot)$ is stepwise constant over each box $A_i \times B_j$, i.e., for some l and u, $g(l,u) = w_{ij}$ for if $l \in A_i$ and $u \in B_j$. Thus, we would like to find choices for $\{w_{ij}\}$ that maximize $z_k(1)$.

Theorem 2: Under the above setup, the optimal spreading strategy $g(\cdot,\cdot)$ is a nondecreasing function of L and U.

To prove the theorem, we introduce the following notations: $E[L|A_i]$ (resp. $E[U|B_j]$) is the expectation of L (resp. U) conditioned on that $L \in A_i$ (resp. $U \in B_j$). We also need the following lemma.

Lemma 3: Let U have an arbitrary positive probability density function over the partitions B_j. Then $\frac{E[U^2|B_j]}{E[U|B_j]}$ is a nondecreasing function of j.

One can see this result intuitively by considering refining the partition of B, so that for some $u_j \in B_j$ we have that $E[U^2|B_j] \approx u_j^2$ and $E[U|B_j] \approx u_j$, thus the ratio is u_j which is nondecreasing in j. For a formal proof of the lemma, we refer the reader to [10].

Proof of Theorem 2: One can rewrite $z_k(1)$ as follows:

$$z_k(1) = \frac{kE[L]E[U] - \sum_{i,j} p_i q_j E[U|B_j]\left\{E[L|A_i] + \frac{1}{2}w_{ij}\right\}}{\sqrt{\sum_{i,j} p_i q_j E[L|A_i]E[U|B_j]w_{ij}^{-1}}}$$

where $p_i := P(L \in A_i)$ and $q_j := P(U \in B_j)$. It is easily seen that $z_k(1)$ admits unique set of maximizers denoted by $\{w_{ij}^*\}$ such that

$$w_{ij}^* = C\sqrt{\frac{E[U^2|B_j]E[L|A_i]}{E[U|B_j]}} \tag{18}$$

where C is a constant given by

$$C = \frac{\frac{2}{3}E[L]E[U](k-1)}{\sum_{i,j} p_i q_j \sqrt{E[U^2|B_j]E[U|B_j]E[L|A_i]}}$$

using the first-order necessary condition for optimality. Firstly, by inspecting (18), since $E[L|A_i] \in A_i$, we see $E[L|A_i]$ is an nondecreasing function of i, i.e., $w_{ij}^* \ge w_{kj}^*$ if $i > k$, from the definition of A_i. Secondly, by Lemma 3, $E[U^2|B_j]E[U|B_j]^{-1}$ is also nondecreasing in j. As one can see from the above derivations, such properties hold irrespective of the distributions of L and U. ■

[6] In fact, L takes integer values. However, our arguments yield more intuitive results and can be trivially extended to such discrete random variable cases.

Now we have the following observation regarding how the optimal spreading width scales with the span and traffic loads of a session. If we decrease the interval lengths of each A_i and B_j into infinitesimal ones, we have that $E[L|A_i] \simeq l_i$ and $E[U^2|B_j]E[U|B_j]^{-1} \simeq u_j$. Thus, from (18) we have the scaling rule of

$$w_{ij}^* \sim \sqrt{l_i u_j}$$

i.e., *the optimal spreading width for a given session approximately follows the square root of the bits · meters of its offered load.* We will verify by simulation that indeed such dynamic spreading schemes outperform those with fixed w.

VII. Design Tradeoffs: Networks With Energy Replenishing Capability

Next, we consider the case where nodes have the capability to replenish their energy at constant rate of c units per unit time and their energy storage capacity is b. We model the energy level of a node by a queue where arrivals correspond to new energy burdens to be served, i.e., replenished at rate c. Note that the dynamics of the queue and their physical interpretation are reversed: filling the queue with energy burden corresponds to consuming its energy reserves. Thus, we are interested in the likelihood that the queue length exceeds the level b.

For a typical node which is covered by multiple session footprints over time, the energy load burden for each footprint would be buffered in the node's energy queue which is replenished at rate c. In reality, an energy request fills the queue, i.e., consumes energy at a roughly constant rate, which can be modelled by using a continuous load model, e.g., fluid queues (see [14]). For simplicity we will assume that energy burdens are imposed *instantaneously* on nodes and the offered load at a typical node depends only on its location within the footprints that hit the location. We will again assume the footprint arrival process is a homogeneous Poisson process in time and space. With these assumptions we will use a discrete-time queueing model that approximates the $M/GI/1$ queue corresponding to these dynamics of energy burden at a typical node.

In this regime, we study the *asymptotic decay rate* of the queue content as an indicator of the probability that the energy burdens exceed a large initial energy reserve of b. For a stable, single-server queue, we denote the steady-state workload by W. If the following condition is satisfied for some $\theta^* > 0$:

$$b^{-1}\log P(W > b) \overset{b\to\infty}{\to} -\theta^*$$

then we refer to θ^* as its asymptotic decay rate [15]. We will use the results in [15] and [16] to describe the behavior of the tail probabilities.

Let us define the problem. The energy burdens of each footprint are assumed to be i.i.d. with a distribution that is not heavy-

TABLE I
DECAY RATES WITH VARYING SPREADING WIDTHS

Spreading width w	Decay rate θ^*	
	$\beta = 1.2$	$\beta = 2.0$
1	0.8673	1.7125
3	**1.2506**	2.7080
5	1.0965	**2.7593**
7	0.7965	2.6831

tailed. We denote the virtual workload for the energy queue associated with a typical location in time slot $(i, i+1]$ for $i \in \mathbb{Z}$ by W_i. Then we have that

$$W_{i+1} = \max[W_i + X_{i+1}, 0] = [W_i + X_{i+1}]^+$$

where $X_i = S_i - c$ and S_i is the total energy burden per unit time slot, and c is the replenished energy per time slot. These dynamics correspond to a Lindley process, and since $\{X_i\}$ are i.i.d., we can readily apply the following results on the decay rate function.

Theorem 3: [15] *Let us assume $\{W_i\}$ is stationary and thus stable under condition $E[X_i] < 0$, i.e., $E[S_i] - c < 0$. If $\{X_i\}$ are i.i.d., then θ^* satisfies*

$$\rho(\theta^*) = 0, \quad \frac{d}{d\theta}\rho(\theta^*) > 0$$

where $\rho(\theta) = \log E[e^{\theta X_i}] = \log E[e^{\theta S_i}] - c\theta$.

We can readily obtain the required cumulant generating function of energy burden per time slot as follows.

Theorem 4: The cumulant generating function $C(\theta)$ of S_i is given by

$$C(\theta) = \log E[e^{\theta S_i}] = \lambda E\left[\int_{\Phi_0} \{e^{\theta h(x,\Phi_0)} - 1\} dx\right].$$

Hence, we have that the rate decay function $\rho(\theta)$ is given by

$$\rho(\theta) = \lambda E\left[\int_{\Phi_0} \{e^{\theta h(x,\Phi_0)} - 1\} dx\right] - c\theta$$

under the stability condition

$$\lambda E\left[\int_{\Phi_0} h(x, \Phi_0) dx\right] < c. \quad (19)$$

Proof: See the Appendix. ■

The stability condition relates the replenishing rate c, and the rate of new energy burden requests per unit area, λ, times the average total energy per footprint. The root θ^* of $\rho(\theta) = 0$ may be found numerically. Using (14) and (15), several decay rates with varying spreading widths are given in Table I. Here $l = 8$ and $\lambda = 1$, and let us denote the critical replenishing rate to satisfy the stability condition when $w = 7$ as c^*. The replenishing rate c is set to βc^* where $\beta = 1.2$ and 2.0.

Again, we observe tradeoffs associated with different replenishing rates. When $\beta = 1.2$, the optimal spreading width is 3, but with a higher replenishing rate $\beta = 2.0$, the optimal w increases to 5. The intuition is that, with higher replenishing rates, one can spread traffic further to get more benefits from spatial balancing. However, if the replenishing rate decreases close to the critical value, the mean energy cost to spread is no longer negligible so that a smaller spreading width is preferred.

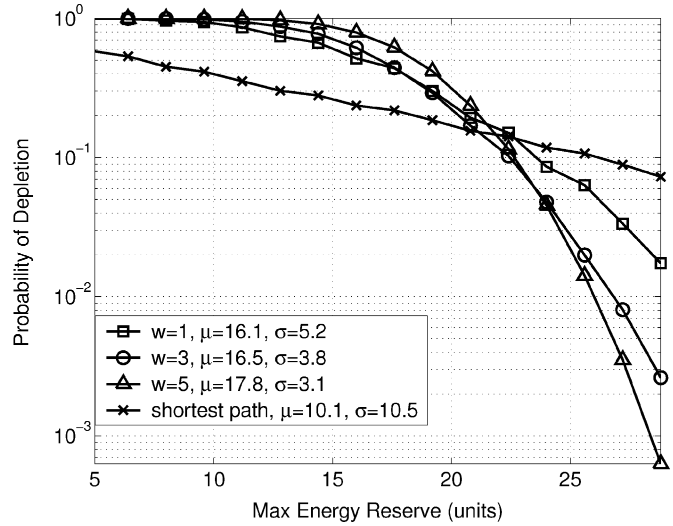


Fig. 8. Energy depletion probability for nodes without energy replenishing capability. μ and σ represents the mean and the standard variation of energy expenditure of each scheme.

VIII. SIMULATIONS

A. Basic Setup

In this section, we simulate several scenarios to further explore the benefits of proactive spreading. The performance metric will be the probability that a randomly selected node is depleted after a fixed time given a maximum energy reserve (MER). This metric is of fundamental interest from an engineering perspective, when given a network operation time and a MER, we wish to minimize the probability of the typical node is depleted, or equivalently, the fraction of depleted nodes in the network.

A total of 400 node locations were generated according to an uniform distribution on a 20×20 unit square area. Session arrivals are homogeneous in space, and a total of 200 sessions are generated at each simulation run. This is repeated 500 times to obtain an averaged energy profile. Unless otherwise specified, each session offers 1 unit of load per unit time with a holding time of 1 unit time. We simulate session arrivals by picking two nodes at random, which corresponds to a session pair and then setting up a unidirectional flow. We set up multipath routes based on the PBM route construction introduced in Section III, and the flow is equally divided on each path in order to approximate the scheme in Section V. In our simulations, the shortest path routing (SPR) is a routing that takes the minimum number of hops on the Delaunay graph of nodes. This must be distinguished from the shortest Delaunay routing (SDR) which is a PBM routing with a spreading width w of 1.

B. Scenarios

1) Nodes Without Replenishing Capability: Fig. 8 shows the average energy depletion probability for several values of the spreading width w and SPR. A point (x, y) in this plot should be interpreted as follows: "the probability that the energy expenditure of a typical node *will exceed* x is y". If x is the MER, then y is the probability that a typical node is depleted.

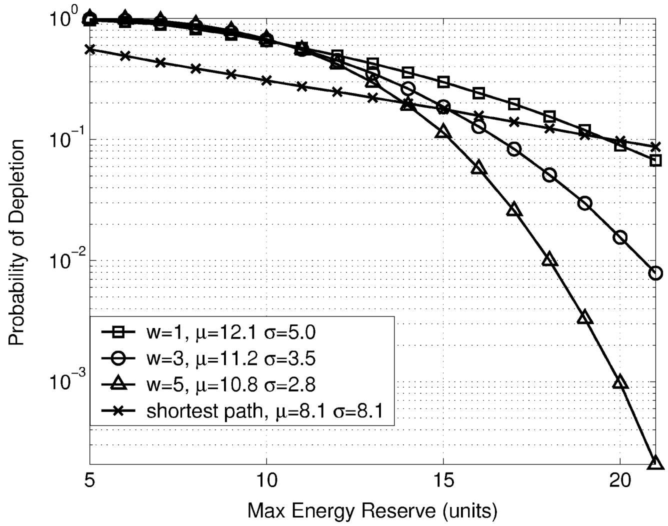


Fig. 9. Energy depletion probability for nodes with energy replenishing capability.

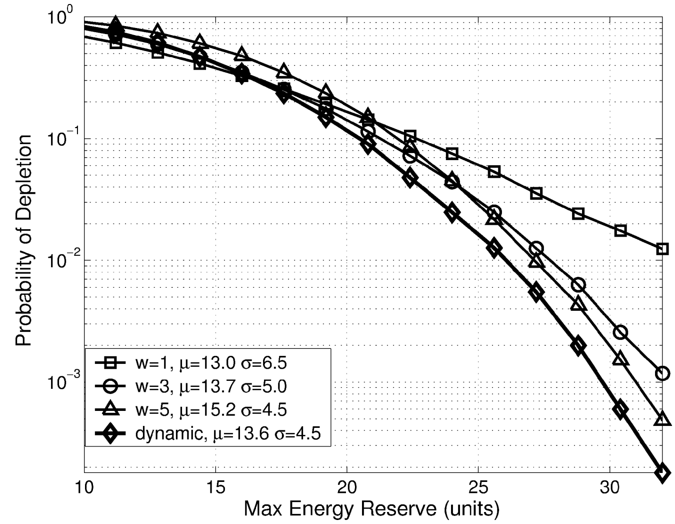


Fig. 10. Energy depletion probability for the dynamic spreading scheme adjusted to session load and hop length.

Let us consider only proactive routing first. When the MER is less than 20 units, routing with a minimal spreading width $(w = 1)$ performs best. However, as the MER increases to more than 25 units, proactive multipath routing with the largest spreading width $(w = 5)$ outperforms the others. These results are consistent with previous discussions, since with a high MER, a scheme with a lower variance in the energy expenditure $(w = 5)$ is preferable at the cost of higher mean energy expenditure. These tradeoffs occur when the maximum reserve is between 20 and 25 units in our simulations. SPR has a lowest mean energy expenditure but the highest variance, and suffers from the worst performance in tail behavior, i.e., the lowest slope in the decay for the probability of depletion with the MER. Also, note that it has different performance as compared to the SDR $(w = 1)$ case: the SDR performs better due to its steeper slope in the tail probability. We verified that, for SDR, the shape of the empirical histogram of energy burden indeed resembles the Gaussian probability density function (p.d.f.), while that for the SPR is monotonically decreasing with a heavy tail [10].

2) Nodes With Replenishing Capability: Fig. 9 shows the energy depletion probabilities when the nodes have the capability of replenishing their energy reserves. At each simulation run a total of 200 sessions arrive uniformly on the time interval [0, 200]. Nodes have replenishing rate of 0.125 energy units per unit time. The benefit from proactive spreading is greater than that seen for the non-replenishing case. The intuition here is that, for larger w, the average number of nodes that participate in a session is greater than that of a scheme with less w. Thus, more nodes have a chance to replenish their energy reserves, which results in a reduced mean *and* less variance in the energy expenditure (see μ and σ in Fig. 9) with the largest spreading width, $w = 5$.

3) Dynamic Spreading: Using the results from Section VI, we have simulated a scheme with dynamic spreading widths depending on the random load U and random session span L according to our scaling rule $\sqrt{LU}$ where nodes do not have energy replenishing capability. Fig. 10 shows the simulation results of such an dynamic spreading scheme. Here each session carries i.i.d. exponentially distributed load of mean 1. As shown in Fig. 10, for small MER region (< 15), the dynamic scheme performs reasonably well but not best perhaps due to the error in rounding w off to an integer. However, it is superior to other schemes with fixed spreading widths as the MER increases.

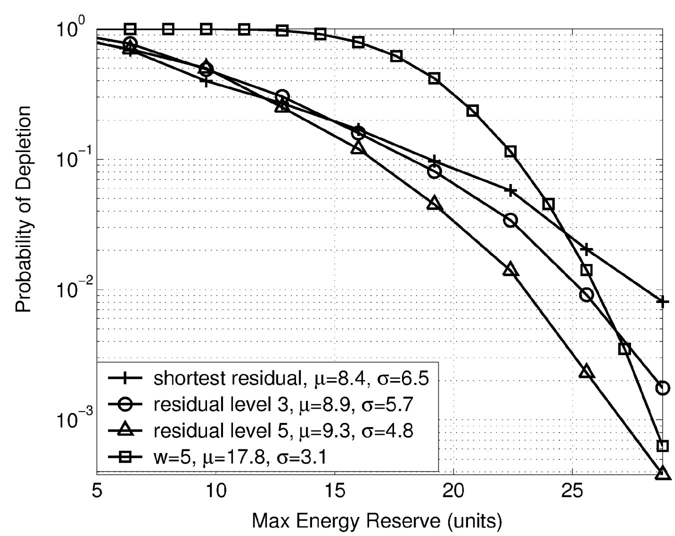


Fig. 11. Comparison of the shortest residual routing, the level-3 and the level-5 residual routing.

4) Routing Based on Residual Energy Reserves: Next, we consider a class of dynamic routing schemes and study how it benefits from proactive load balancing. Specifically, we consider a routing scheme which exploits knowledge of the residual energy reserve at each node, i.e., routing with state information. We use Bellman–Ford algorithm for minimum cost routing, where the cost is a decreasing function of the fraction of the residual energy to its full reserve. In this way, a route through nodes having relatively high energy reserves might be preferred even if that route involves a higher number of hops. Specifically, if the residual energy of ith node is $b_i(t)$ at time t, the cost of routing traffic to that node is $(b_i(t)/b_{\max})^{-\gamma}$ where γ is some positive constant and $b_{\max}$ is the maximum energy reserve at a node. Concerning the choice of γ, a related study [17] shows that a value within the range of $0.5 \sim 2.5$ is preferable, so we chose $\gamma = 1$. Also, we have assumed that the routes do not change once created since such changes can incur a severe scalability problem due to the large number of nodes.

For these simulations, we define a level-w residual routing to be such that, the w best disjoint routes are chosen. Fig. 11 shows one of such comparison. We see that proactive spreading reduces the tail probability although the performance of state-dependent routing schemes is sensitive to the variability of traffic. In addition, Fig. 11 exhibits the performance of proactive multipath routing without residual energy information when $w = 5$. The results show that, although it does not use dynamic state information, it may be adequate for a network whose nodes have high energy reserves.

IX. Conclusion and Future Work

In this paper, we propose a simple model for the spatial distribution of energy burdens in a multihop ad hoc wireless network. Our primary contribution is to use these models to investigate the design and potential benefits of proactive energy balancing multipath routing schemes. To do so, we develop a simple second-order approximation permitting one to investigate tradeoffs of several types, e.g., for ad hoc networks with or without replenishing and with energy storage capabilities. The essential tradeoff is between the mean and variance of a spatial energy (flow) balancing scheme. For our proposed models, one might attempt to identify Pareto optimal energy balancing strategies, e.g., one minimizing the variance subject to a mean energy constraint, or conversely one minimizing the mean energy burden subject to a variance constraint. To simplify matters, we consider flow/energy balancing on regular grid model for a simple parameterized family of spreading schemes. This permits us to concretely evaluate how this tradeoff should be optimized for the various network types and possible design criteria. The results are insightful but perhaps not unexpected. For networks with increased energy storage and/or replenishing capabilities it pays to be more aggressive in spreading traffic so as to reduce the variance in the energy burden since the additional energy burden can be smoothed by energy reserves or new energy sources—one must, however, ensure that the energy burden does not exceed the replenishing capability. For the most part our simulations confirm our analytical results and permitted us to evaluate more general regimes of interest.

We note, however, that the traffic patterns and network geometry used in our simulations are fairly benign in that they are fairly homogeneous in time and space. In practice, one would expect to see irregular topologies and imbalances and variability in traffic loads. These in turn would lead to additional variability in the energy burdens on the network. We expect, that the benefits of proactive load balancing to be be more prominent and sensitive to design in the presence of the aforementioned fluctuations. The degree of spreading, e.g., w, might advantageously be exploited to adaptively smooth out such spatial variabilities and achieve improved balancing of energy burdens coupled with improved performance on network lifetime. Indeed, one of the key results in this paper shows that w scales roughly as the square root of load times distance, which is well expected by the intuition that the more one should spread traffic when the spatial burden imposed on the network increases.

Finally, we note that our focus here has been on a preliminary analysis of proactive energy balancing. As such we have used a simplified energy model, appropriate to study a routing scheme. Yet overheads associated with setting up multipath routes, or other sources of energy expenditure or savings, e.g., putting nodes to sleep, will play a role. For example, in our model we have for the most part ignored MAC layer. In practice the temporal granularity on which load balancing is performed might be critical. For example, fine grain spreading of traffic might cause contention for transmission among neighboring paths lessening the benefits from an energy perspective. Such interactions need to be studied carefully, and might be lessened by increasing the granularity of spreading. These and additional aspects of the proposed routing strategies are part of our ongoing work.

Appendix

Proof of Lemma 2: We consider only the left half part of the geometry using the symmetric property of the problem. Let V be the grid set of the left half part, i.e., $V := \{(i,j) \mid 0 \leq i \leq l/2,\ |j| \leq ((w-1)/2)\}$. We have that

$$\sum_{i,j \in V} e_{i,j}^2 = \left\{ \sum_{k=0}^{(w-1)/2} \sum_{|i|+|j|=k} e_{i,j}^2 \right\} \tag{20}$$

$$+ \sum_{|i|+|j|>(w-1)/2} \{e_{i,j}^2\}. \tag{21}$$

In (20), since $\sum_{|i|+|j|=k} e_{i,j} = 1$ regardless of k by the flow conservation, we have that, using the Cauchy–Schwarz inequality

$$\sum_{|i|+|j|=k} e_{i,j}^2 \geq \frac{1}{2k+1} \left\{ \sum_{|i|+|j|=k} e_{i,j} \right\}^2$$

and the equality is achieved when $e_{i,j} = (1/(2k+1))$.

For (21), it can be shown that the minimization is achieved by setting all $e_{i,j}$ equal to $1/w$ due to the constraint (13); see [10] for the proof. ■

Proof of Theorem 4: We have that the energy request arrival process is Poisson with rate λ per unit time per unit space. Since S_i is defined as the energy request in unit time interval, S_i is stochastically equivalent to the shot-noise process in $\mathbb{R}^2$ with intensity λ. From Lemma 1, the nth order cumulant of S_i is $\chi^{(n)}(1)$, thus we have that

$$\begin{aligned} C(\theta) &= \sum_{n=1}^{\infty} \chi^{(n)} \frac{\theta^n}{n!} = \lambda \sum_{n=1}^{\infty} E\left[\int_{\Phi_0} h(x,\Phi_0)^n dx\right] \frac{\theta^n}{n!} \\ &= \lambda E\left[\int_{\Phi_0} \sum_{n=1}^{\infty} h(x,\Phi_0)^n \frac{\theta^n}{n!} dx\right] \\ &= \lambda E\left[\int_{\Phi_0} \{e^{\theta h(x,\Phi_0)} - 1\} dx\right]. \end{aligned}$$

■

REFERENCES

[1] V. Rodoplu and T. Meng, "Minimum energy mobile wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 8, pp. 1333–1344, Aug. 1999.

[2] J. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks," *IEEE Trans. Networking*, vol. 12, no. 4, pp. 609–619, Aug. 2004.

[3] C. Toh, H. Cobb, and D. Scott, "Performance evaluation of battery-life-aware routing schemes for wireless ad hoc networks," in *Proc. ICC*, 2001, vol. 1, pp. 2824–2829.

[4] K. Kar, M. Kodialam, T. Lakshman, and L. Tassiulas, "Routing for network capacity maximization in energy-constrained ad hoc networks," in *Proc. IEEE INFOCOM*, 2003, vol. 1, pp. 673–681.

[5] L. Lin, N. Shroff, and R. Srikant, "Asymptotically optimal power-aware routing for multihop wireless networks with renewable energy sources," in *Proc. IEEE INFOCOM*, 2005, vol. 2, pp. 1262–1272.

[6] S. Servetto and G. Barrenechea, "Constrained random walks on random graphs: Routing algorithms for large scale wireless sensor networks," in *Proc. 1st ACM Int. Workshop on Wireless Sensor Networks and Applications*, 2002, pp. 12–21.

[7] M. Keil and C. Gutwin, "Classes of graphs which approximate the complete Euclidean graph," *Discrete Comput. Geom.*, vol. 7, pp. 13–28, 1992.

[8] F. Baccelli, K. Tchoumatchenko, and S. Zuyev, "Markov paths on the Poisson-Delaunay graph with applications to routing in mobile networks," *Adv. Appl. Probabil.*, vol. 32, pp. 1–18, 2000.

[9] L. Heinrich and V. Schmidt, "Normal convergence of multidimensional shot noise and rates of this convergence," *Adv. Appl. Probabil.*, vol. 17, pp. 709–730, 1985.

[10] S. Baek and G. de Veciana, "Spatial energy balancing through proactive multipath routing in multi-hop wireless networks," Wireless Networking and Communications Group (WNCG), Univ. Texas, Austin, TX, WNCG Tech. Rep., 2004 [Online]. Available: http://www.ece.utexas.edu/~gustavo

[11] S. Nash and A. Sofer, *Linear and Nonlinear Programming*. New York: McGraw-Hill, 1996.

[12] D. Aldous, *Probability Approximations Via the Poisson Clumping Heuristic*. New York: Springer-Verlag, 1989.

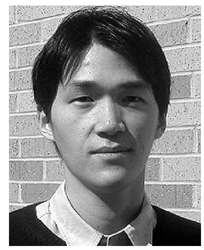
[13] R. Adler, *The Geometry of Random Fields*. New York: Wiley, 1981.

[14] J. Cohen, "The M/G/1 fluid model with heavy-tailed message length distributions," Centrum voor Wiskunde en Informatica, Amsterdam, The Netherlands, Tech. Rep. PNA-R9714, 1997.

[15] W. Whitt and P. Glynn, "Logarithmic asymptotics for steady-state tail probabilities in a single-server queue," in *Studies in Applied Probability, Papers in Honour of Lajos Takacs*, J. Galambos and J. Gani, Eds. Sheffield, U.K.: Applied Probability Trust, 1994, pp. 131–156.

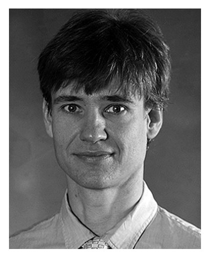
[16] G. de Veciana and J. Walrand, "Effective bandwidths: Call admission, traffic policing and filtering for ATM networks," *Queueing Syst.*, vol. 20, pp. 37–59, 1995.

[17] X. Su and G. de Veciana, "Predictive routing to enhance QoS for stream-based flows sharing excess bandwidth," *Comput. Netw.*, no. 42, pp. 65–82, May 2003.



Seung Jun Baek (S'04) received the B.S. degree from Seoul National University, Seoul, Korea, in 1998, and the M.S. degree from the University of Texas at Austin in 2002, both in electrical and computer engineering. He is currently working toward the Ph.D. degree at the University of Texas at Austin.

His research interests include wireless ad hoc networks and sensor networks.



Gustavo de Veciana (S'88–M'94–SM'01) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California at Berkeley in 1987, 1990, and 1993, respectively.

He is currently a Professor in the Department of Electrical and Computer Engineering, University of Texas at Austin, and the Director of the Wireless Networking and Communications Group (WNCG). His research focuses on the design, analysis and control of telecommunication networks. His current interests include measurement, modeling and performance evaluation, wireless and sensor networks, and architectures and algorithms to design reliable computing and network systems.

Dr. de Veciana has been an editor for the IEEE/ACM TRANSACTIONS ON NETWORKING. He is the recipient of General Motors Foundation Centennial Fellowship in Electrical Engineering and a 1996 National Science Foundation CAREER Award, co-recipient of the IEEE William McCalla Best ICCAD Paper Award for 2000, and co-recipient of the Best Paper in *ACM Transactions on Design Automation of Electronic Systems*, January 2002–2004.