Near-Optimal Load Balancing in Dense Wireless Multi-Hop Networks

Esa Hyytiä†
Telecommunications Research Center Vienna (ftw.)
Donau-City Straße 1, 1220 Vienna, Austria

Jorma Virtamo‡
Helsinki University of Technology, TKK
P.O. Box 3000, FIN 02015 TKK, Finland

***Abstract*—We consider the load balancing problem in wireless multi-hop networks. In the limit of a dense network, there is a strong separation between the macroscopic and microscopic scales, and the load balancing problem can be formulated as finding continuous curves ("routes") between all source-destination pairs that minimize the maximum of the so-called scalar packet flux ("traffic load"). In this paper we re-formulate the problem by focusing entirely on the so-called d-flows (vector flow field of packets with a common destination $\mathbf{x}$) and by looking at the equation these flows have to satisfy. The general solution to this equation can be written in terms of a single unknown scalar function, $\psi(\mathbf{r}, \mathbf{x})$, related to the circulation density of the d-flow, for which function the optimization task can be presented as a problem of variational calculus. In this approach, we avoid completely dealing with systems of paths and calculating the load distribution resulting from the use of a given set of paths. Once the optimal solution for $\psi(\mathbf{r}, \mathbf{x})$ is found the corresponding paths are obtained as the flow lines of the d-flows. In the example of a unit disk with uniform traffic demands we are able to find a set of paths which is considerably better than any previously published results, yielding a low maximal scalar flux and an extraordinarily flat load distribution. We further illustrate the methodology for a unit square with comparable improvements achieved.**

I. Introduction

In a wireless multi-hop network a typical path consists of several hops and the intermediate nodes along a path act as relays. Thus, each node has two functions. First, it can act as a source or a destination for some flow, i.e., the nodes can communicate with each other. Secondly, when necessary, nodes have to relay packets belonging to the other flows.

Several types of wireless multi-hop networks have been proposed each with different unique characteristics. For example, wireless sensor networks are networks designed to collect some information from a given area and to deliver the information to one or more sinks [1]. Thus, the traffic distribution in sensor networks is typically highly asymmetric and the energy efficiency is often a critical design parameter. Another example of wireless multi-hop networks are the wireless mesh networks consisting of both mobile and fixed wireless nodes and one or more gateway nodes through which the users have access to the Internet [2], [3]. It has been envisioned that these types of networks, depending on the application, may consist of a large number of nodes such that a lot of relaying is needed to ensure end-to-end connectivity.

In this paper we study wireless multi-hop network in the limit when the number of nodes is large. In this limit the network is referred to as a dense network [4], [5]. In particular, we assume a strong separation in spatial scales between the macroscopic level, corresponding to typical distances between the source and destination nodes, and the microscopic level, corresponding to distances between the neighboring nodes. This justifies modeling the routes on the macroscopic scale as smooth geometric curves as if the underlying network fabric formed a homogeneous and isotropic[1] continuous medium.

The microscopic scale corresponds to a single node and its immediate neighbors. At this scale the above assumptions imply that only the direction in which a particular packet is traversing is significant. Considering one direction at a time there exists a certain maximum flow of packets a given MAC protocol can support (packets per unit time per unit length, 'density of progress'). This maximal sustainable directed packet flow depends on scheduling rules and possible coordination between the nodes. Determining the value of this maximum is not a topic of this paper but it is assumed to be given (known characteristic constant of the medium). By a simple time sharing mechanism this maximal value can be shared between flows propagating in different directions. As a result, the scalar flux (defined in Section III) of packets is bounded by the given maximum, and the load balancing task is to determine the paths in such a way that the maximum flux is minimized. For instance, using shortest paths (straight line segments in the current setting) typically concentrates significantly more traffic in the center of the network than elsewhere limiting the capacity of the network.

A. Our Contributions

In a series of papers we have dealt with the load balancing problem in dense wireless multi-hop networks with progressively improving results. In [6] we defined the basic concepts of angular and scalar flux (borrowed from the neutron transport theory [7]) and defined the load balancing problem as a minmax problem of finding the set of paths that minimizes the maximal scalar flux. Further we gave formulae for calculating the scalar flux at a given point for any curvilinear set of paths.

†For the part of E. Hyytiä, this work has been carried out in the Telecommunications Research Center Vienna (ftw.) in the framework of the Austrian Kplus Competence Centre programme.

‡For the part of J. Virtamo, this work was carried out in the project Fancy funded by the Academy of Finland (grant n:o 210275).

[1]Homogeneity and isotropicity are not crucial but are assumed here to simplify the discussion.

These results were then applied for the example of a unit disk with uniform traffic demands with the total packet stream Λ. Using shortest path routing, the maximal scalar flux occurring at the center of the unit disk is $0.637\,\Lambda$. By using so-called circular paths the maximal flux could be reduced to 2/3 of the shortest path value, i.e. to $0.424\,\Lambda$. An even better result was obtained by allowing multi-path routing and carrying given proportions of the traffic on paths chosen from three different sets. The minimum (over the mixture ratios) of the maximal flux was $0.379\,\Lambda$. Lower bounds were also established for the optimum. In the circular disk geometry a sure lower bound, so-called distance bound, is $0.288\,\Lambda$.

Publication [8] is an extended version of [6], and notably contains a new efficient formula for calculating the scalar flux distribution in the disk geometry for the routing where each path is a section of a curve in the so-called basic set of paths (a one-parameter family of paths) or a rotation thereof. By using certain type of modified circular paths the single-path routing optimum could be brought down to $0.384\,\Lambda$.

An important new insight was given in [9], where we introduced the concept of a d-flow $\mathbf{J}(\mathbf{r}, \mathbf{x})$, which means the vector flow field at $\mathbf{r}$ of the packets heading to a given destination $\mathbf{x}$. It was proven that no matter which routing (multi-path, single-path or whatever) was used, the scalar flux decreases (or certainly not increases) everywhere if the routing is replaced with the single-path destination-based forwarding (DBF), where the flow lines of the d-flow are used as the paths. This means that the optimum can always be obtained by single-path destination-based forwarding. By applying this principle to the multi-path routing of the previous paper the optimum decreased to $0.343\,\Lambda$, which is almost 50% improvement in comparison with the shortest path routes.

In the present paper, we make a major conceptual leap forward leading to a great simplification of the problem. Our starting point is the observation that the d-flow field $\mathbf{J}(\mathbf{r}, \mathbf{x})$ has a given source density (the density of $\mathbf{x}$-destined traffic emanating at $\mathbf{r}$ and negative δ-function sink at $\mathbf{x}$). Thus it satisfies a so-called source equation whose all solutions differ from each other by a pure circulation field defined by a single scalar function, $\psi(\mathbf{r}, \mathbf{x})$, related to the circulation density. Thus the whole optimization problem can be expressed in terms of optimizing with respect to this function $\psi(\mathbf{r}, \mathbf{x})$, with no reference to which routes are used. In particular, the difficult step of calculating the scalar flux for a given set of paths is completely avoided. When the optimal $\psi(\mathbf{r}, \mathbf{x})$ function is determined, then the paths are afterwards obtained as the flow lines of the resulting d-flow. Starting from a specific solution for the source equation and using a simple trial function $\psi(\mathbf{r}, \mathbf{x})$ with a single adjustable parameter, we obtain a result $0.329\,\Lambda$ which is lowest result so far reported. The resulting load distribution is remarkably flat, and we believe it is close to the real optimum. The proposed approach is not limited to any particular geometry and for a square area we find paths yielding a 43% reduction in the maximum scalar flux.

The rest of this paper is organized as follows. Section II presents related work. Then, in Section III we introduce the necessary definitions and review some earlier results. In Section IV we develop the new problem formulation, which is illustrated in Section V by finding near optimal routes for disk and square geometries. Section VI concludes the paper.

II. Related Work

The research community has recently shown considerable interest towards different routing problems in dense wireless multi-hop networks. The idea of studying routes as continuous curves at the macroscopic level was introduced in [10] and [4]. However, the trajectory based forwarding (TBF) scheme, proposed by Niculescu and Nath in [11], already provides the connection between the macroscopic and microscopic levels.

In TBF, the source specifies the route as a continuous curve to the destination (cf., dynamic source routing [12]). Additionally, TBF defines how the nodes at the microscopic level forward packets in order to approximately follow this path. Another closely related concept is the so-called geometric routing (or geographic routing) paradigm [13], [14], [15], where each packet carries the information (directly or indirectly) about the location of the destination node.

The shortest path routes, i.e., straight line segments, tend to concentrate traffic in the center of the network. This is particularly true with uniform traffic demands [16], [6], [17]. In [16], Pham and Perreau, and later, in [18], Ganjali and Keshavarzian have studied the load balancing using k-shortest paths instead of shortest path. The analysis assumes a disk area and a high node density so that the shortest paths correspond to straight line segments. In multipath situation the straight line segments are replaced by rectangular areas where the width of the rectangle is related to the number of multiple paths.

Considerable amount of recent work relies on strong analogy with physical systems. Also the irrotational solution for the d-flow in the present paper stems from physical systems. Similarly, Kalantari and Shayman, in [10] study dense wireless multi-hop networks by leaning to theory of electrostatics and consider a routing problem where a large number of nodes send data to a single destination. In this case the optimal paths are obtained by solving a set of partial differential equations similar to Maxwell's equations. Using a similar analogy with electrostatics, Toumpis and Tassiulas, in [19], consider a related problem of optimal placement of the nodes in a dense sensor network. It is assumed that at any point of the network the information flows exactly to one direction.

Furthermore, Catanuto et al., in [20], [21], have optimized the routes by exploiting the analogy with geometrical optics. The communication cost, which is related to the index of refraction in the analogy, is defined as a function of the node density. In [22] Popa et al. study the load balancing problem and by using linear programming approach are able to show that the optimal paths can be expressed as ray trajectories of geometrical optics. This implies two things. Firstly, the optimal paths correspond to shortest paths with a certain metric (index of refraction). Secondly, the optimal load balancing can be achieved by single path routes, a result obtained independently in [9] by using a different approach. The only obstacle with the approach seems to be the technical difficulties in actually computing the resulting traffic load and paths. Finally, a comprehensive survey on different modeling approaches and problem formulations is given by Toumpis [5].

III. Preliminaries

In this section we give basic definitions and review some earlier results. Let $\mathcal{A}$ denote the area of the network. In our setting the nodes form a continuum and therefore it is convenient to define the traffic demands as densities:

Definition 1 (traffic demand density) *The rate of flow of packets from a differential area element dA about $\mathbf{r}$ to a differential area element dA about $\mathbf{x}$ is $\lambda(\mathbf{r},\mathbf{x}) \cdot dA^2$, where $\lambda(\mathbf{r},\mathbf{x})$ is called the traffic demand density [pkts/s/m^4].*

The total offered traffic, denoted by Λ [pkts/s], is given by,

$$\Lambda = \int_{\mathcal{A}} \int_{\mathcal{A}} \lambda(\mathbf{r},\mathbf{x})\, d^2\mathbf{x}\, d^2\mathbf{r}.$$

The load balancing problem in dense multi-hop networks can be defined as a minmax problem for the scalar packet flux. Scalar packet flux in turn is defined in terms of so-called angular packet flux (see Fig. 1):

Definition 2 (angular flux) *Angular flux of packets at $\mathbf{r}$ in direction θ, denoted by $\varphi(\mathbf{r},\theta)$, is equal to the rate [1/s/m/rad] at which packets flow in the angle interval $(\theta, \theta+d\theta)$ across a small line segment of the length ds perpendicular to direction θ at point $\mathbf{r}$ divided by $ds \cdot d\theta$ in the limit when $ds, d\theta \to 0$.*

Definition 3 (scalar flux) *Scalar flux of packets [1/s/m] at $\mathbf{r}$ is given by*

$$\Phi(\mathbf{r}) = \int_0^{2\pi} \varphi(\mathbf{r},\theta)\, d\theta. \tag{1}$$

To elaborate this, we define *progress* as the advance a packet makes in a given time interval in the direction of its path.

Remark 1 (density of cumulative progress rate) *Scalar flux $\Phi(\mathbf{r})$ corresponds to the cumulative progress [m] of packets per unit time [s] per unit area [m^2] about point $\mathbf{r}$ (rendering 1/s/m as its dimension).*

Given the traffic demand density $\lambda(\mathbf{r},\mathbf{x})$, the scalar flux depends on the set of paths $\mathcal{P}$ used for routing. The load balancing problem in the context of dense multi-hop network can now be stated as follows [6], [8], [9], [22]:

Definition 4 (load balancing problem) *Find the set of paths $\mathcal{P}_{\text{opt}}$ which minimizes the maximum scalar flux,*

$$\mathcal{P}_{\text{opt}} = \arg\min_{\mathcal{P}} \max_{\mathbf{r}} \Phi(\mathcal{P},\mathbf{r}). \tag{2}$$

The corresponding maximum scalar flux is denoted by Φ_{opt}.

In [6], we derived two lower bounds for Φ_{opt}. Here we present the more useful of them.

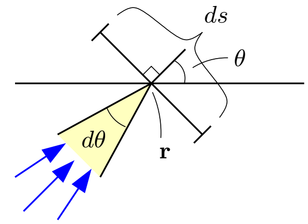

Fig. 1. Angular flux $\varphi(\mathbf{r},\theta)$ is the rate of packets crossing a small line segment ds in angle $(\theta, \theta+d\theta)$ divided by $d\theta \cdot ds$ at the limit $d\theta, ds \to 0$.

Proposition 1 (distance bound)

$$\max_{\mathbf{r}} \Phi(\mathcal{P},\mathbf{r}) \geq \frac{\Lambda \cdot \bar{\ell}}{A}, \tag{3}$$

where $\bar{\ell}$ denotes the mean path length of the set of paths $\mathcal{P}$.

The proposition says that the maximum of the local density of the cumulative progress rate is at least as great as the corresponding mean density. Obviously the shortest path routes have the lowest mean path length, denoted by $\bar{\ell}_{\text{sp}}$, which yields

$$\Phi_{\text{opt}} \geq \frac{\Lambda \cdot \bar{\ell}_{\text{sp}}}{A}. \tag{4}$$

In the analysis it is convenient to consider the packets having a given destination $\mathbf{x}$. To this end, we need some additional definitions.

Definition 5 (angular d-flux density) *Angular d-flux density, denoted by $\varphi(\mathbf{r},\theta;\mathbf{x})$ [1/s/m^3/rad], is equal to the angular flux $\varphi(\mathbf{r},\theta)$ resulting from the packets having their final destination in small area dA about $\mathbf{x}$ divided by dA in the limit when $dA \to 0$.*

Thus, by definition

$$\varphi(\mathbf{r},\theta) = \int_{\mathcal{A}} \varphi(\mathbf{r},\theta;\mathbf{x})\, d^2\mathbf{x}. \tag{5}$$

Definition 6 (d-flow intensity) *Destination flow intensity or for short d-flow intensity of packets at $\mathbf{r}$ having destination $\mathbf{x}$, denoted by $\mathbf{J}(\mathbf{r},\mathbf{x})$ [1/s/m^3], is equal to*

$$\mathbf{J}(\mathbf{r},\mathbf{x}) = \int_0^{2\pi} \varphi(\mathbf{r},\theta;\mathbf{x})\, \mathbf{e}_\theta\, d\theta, \tag{6}$$

where $\mathbf{e}_\theta$ is the unit vector in direction θ.

The angular d-flux density $\varphi(\mathbf{r},\theta;\mathbf{x})$ defines the specific intensity, in each direction θ, of the flow of packets having their final destination about $\mathbf{x}$. The d-flow intensity $\mathbf{J}(\mathbf{r},\mathbf{x})$ is a vector sum of these flows, i.e., it represents the net flow at $\mathbf{r}$ of packets destined to $\mathbf{x}$. For a given destination $\mathbf{x}$ the d-flow intensity $\mathbf{J}(\mathbf{r},\mathbf{x})$ defines a vector field in the variable $\mathbf{r}$. Note that while the net packet flow $\mathbf{J}(\mathbf{r}) = \int \mathbf{J}(\mathbf{r},\mathbf{x})\, d^2\mathbf{x}$ may be zero everywhere, e.g., in the case of symmetric traffic matrix $\lambda(\mathbf{r},\mathbf{x}) = \lambda(\mathbf{x},\mathbf{r})$ and bidirectional routes, the d-flow intensity $\mathbf{J}(\mathbf{r},\mathbf{x})$ can still be non-zero.

In [9] we have shown an important result that the optimal solution to the load balancing problem can be obtained using the destination-based forwarding. To be precise, we define:

Definition 7 (destination-based forwarding (DBF)) *Routing where the forwarding direction of a packet depends only on the current location* $\mathbf{r}$ *and the packet's destination* $\mathbf{x}$ *is referred to as the* destination-based forwarding. *The corresponding forwarding direction is denoted by* $\vartheta(\mathbf{r}, \mathbf{x})$.

Note that the DBF defines a single path between each source-destination pair. For example, the most elementary DBF scheme is the shortest path (SP) routes, for which we have $\vartheta_{\mathrm{SP}}(\mathbf{r}, \mathbf{x}) = \arg(\mathbf{x} - \mathbf{r})$. Because in DBF all packets located at $\mathbf{r}$ and destined to $\mathbf{x}$ are forwarded in the same direction $\vartheta(\mathbf{r}, \mathbf{x})$, it follows in view of (6) that

$$\varphi(\mathbf{r}, \theta; \mathbf{x}) = |\mathbf{J}(\mathbf{r}, \mathbf{x})| \cdot \delta(\theta - \vartheta(\mathbf{r}, \mathbf{x})),$$

where $\delta(\cdot)$ denotes the Dirac's δ-function, and $\vartheta(\mathbf{r}, \mathbf{x}) = \arg \mathbf{J}(\mathbf{r}, \mathbf{x})$. Then by (1) and (5) we have

$$\Phi(\mathcal{P}, \mathbf{r}) = \int\limits_{\mathcal{A}} |\mathbf{J}(\mathbf{r}, \mathbf{x})| \, d^2\mathbf{x}. \tag{7}$$

Thus, for DBF all relevant quantities (paths, traffic demands and the corresponding scalar packet flux) can be easily obtained once the d-flow $\mathbf{J}(\mathbf{r}, \mathbf{x})$ is known. Moreover, any routing system can be transformed in a straightforward manner to a corresponding DBF based routing by using the field lines of the d-flow field $\mathbf{J}(\mathbf{r}, \mathbf{x})$ of the original system as routes to $\mathbf{x}$. It can be shown that in this transformation the scalar flux $\Phi(\mathbf{r})$ at any points remains the same or decreases. This observation is at the core of the optimality of the DBF-paths [9].

IV. New Formulation of the Problem

As just discussed, the optimal routes for the load balancing problem are single-path, destination-based routes. No matter how a routing is originally done, let it be multi-path or non-destination-based, one can always obtain at least the same or better performance by using the flow lines of the d-flow field $\mathbf{J}(\mathbf{r}, \mathbf{x})$ as routes leading to a destination $\mathbf{x}$. Irrespective of the routing, the d-flow $\mathbf{J}(\mathbf{r}, \mathbf{x})$ must always satisfy the equation

$$\nabla \cdot \mathbf{J}(\mathbf{r}, \mathbf{x}) = \lambda(\mathbf{r}, \mathbf{x}) - \Lambda(\mathbf{x})\delta(\mathbf{r} - \mathbf{x}), \quad \forall\, \mathbf{r}, \mathbf{x}, \tag{8}$$

where the right hand side represents the source density of a d-flow with destination $\mathbf{x}$, i.e., a positive source density when $\mathbf{r} \neq \mathbf{x}$ and a singular sink at $\mathbf{x}$. Here $\Lambda(\mathbf{x}) = \int \lambda(\mathbf{r}, \mathbf{x}) d^2\mathbf{r}$ is the density of the total traffic destined to $\mathbf{x}$ per unit area about $\mathbf{x}$. In addition, $\mathbf{J}(\mathbf{r}, \mathbf{x})$ satisfies the boundary condition $\mathbf{J}(\mathbf{r}, \mathbf{x}) \cdot \hat{\mathbf{n}}(\mathbf{r}) = 0$ for all points $\mathbf{r}$ on the boundary, where $\hat{\mathbf{n}}(\mathbf{r})$ in unit normal vector of the boundary at $\mathbf{r}$, i.e., there is no flow across the boundary.

The solution to (8) is not unique; different routings obviously lead to different d-flow fields. If $\mathbf{J}_0(\mathbf{r}, \mathbf{x})$ is a specific solution to (8), then the general solution is of the form

$$\mathbf{J}(\mathbf{r}, \mathbf{x}) = \mathbf{J}_0(\mathbf{r}, \mathbf{x}) + \hat{\mathbf{k}} \times \nabla\psi(\mathbf{r}, \mathbf{x}), \tag{9}$$

where $\hat{\mathbf{k}}$ is the unit vector perpendicular to the plane. The proof is given in the Appendix, see also, e.g., [23]. The above expression satisfies (8) since $\nabla \cdot \hat{\mathbf{k}} \times \nabla\psi(\mathbf{r}, \mathbf{x}) = 0$. The boundary condition of $\mathbf{J}(\mathbf{r}, \mathbf{x})$ implies that the tangential component for $\nabla\psi(\mathbf{r}, \mathbf{x})$ at the boundary has to vanish, which means that $\psi(\mathbf{r}, \mathbf{x})$ *is constant on the boundary*. Otherwise the function $\psi(\mathbf{r}, \mathbf{x})$ is arbitrary. Note that (9) is the general solution, i.e. it exhausts the set of all possible d-flows to destination $\mathbf{x}$, including all routing possibilities. In passing, we note that

$$\nabla \times (\hat{\mathbf{k}} \times \nabla\psi(\mathbf{r}, \mathbf{x})) = \hat{\mathbf{k}}\, \nabla^2\psi(\mathbf{r}, \mathbf{x}),$$

i.e., $\nabla^2\psi(\mathbf{r}, \mathbf{x})$ gives the rotation or circulation density of the added vector field. Given the d-flow (9) in terms of $\psi(\mathbf{r}, \mathbf{x})$, the scalar flux can be calculated with the aid of (7). We are now ready to give a new formulation for the load balancing problem of minimizing the maximum of the flux:

Definition 8 (load balancing problem) *Find* $\psi(\mathbf{r}, \mathbf{x})$ *such that it is constant at the boundary* $\partial\mathcal{A}$, $\psi(\mathbf{r}, \mathbf{x}) = c \;\forall\, \mathbf{r} \in \partial\mathcal{A}$, *and minimizes the maximum scalar flux,*

$$\min_{\psi} \max_{\mathbf{r}} \int |\mathbf{J}_0(\mathbf{r}, \mathbf{x}) + \hat{\mathbf{k}} \times \nabla\psi(\mathbf{r}, \mathbf{x})| \, d^2\mathbf{x}. \tag{10}$$

Note that in contrast to the original formulation of the load balancing problem (2), where $\Phi(\mathcal{P}, \mathbf{r})$ is a complex functional of the used path system $\mathcal{P}$ (indeed finding this functional dependence itself is a difficult task, to the solution of which we have devoted considerable effort in [6] and [8]), the new equation (10) does not refer to the paths at all and is an explicit variational problem in terms of the unknown function $\psi(\mathbf{r}, \mathbf{x})$. This makes the new formulation much more easily amenable to systematic numerical approaches.

In summary, finding the optimal paths is accomplished as follows. First a specific solution $\mathbf{J}_0(\mathbf{r}, \mathbf{x})$ is obtained, e.g., by using the shortest paths for which the $\mathbf{J}_{\mathrm{SP}}(\mathbf{r}, \mathbf{x})$ is straightforward to calculate. Then, using (10) and appropriately chosen parameterized family of functions $\{\psi(\mathbf{r}, \mathbf{x})\}$, one numerically finds the parameter values that yield optimal $\psi(\mathbf{r}, \mathbf{x})$. Finally the corresponding paths are obtained as flow lines of the resulting $\mathbf{J}(\mathbf{r}, \mathbf{x})$ given by (9). It should be emphasized that in this new approach the actual paths are not explicitly part of the formulation (10), and are not needed in the optimization process. The paths are, however, straightforward to obtain afterwards. The advantage of this formulation is that finding a scalar function $\psi(\mathbf{r}, \mathbf{x})$ that realizes the minimum of (10) is conceptually a considerably easier task than working with all possible paths in a given area and computing the resulting scalar flux at each step (the approach taken in [6], [8], [9]).

In [22] the authors have studied the same problem by using optical paths. The optimization parameter in this case is the index of refraction and the paths are defined by the trajectory of a light path between the given points. However, determining a path or a local forwarding direction with a given index of refraction leads to a classical problem of variational calculus, which is not trivial to solve in general. Similarly, it is also

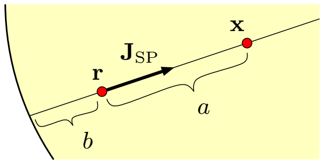

Fig. 2. Notation for the d-flow intensity $J_{\rm SP}(\mathbf{r},\mathbf{x})$ with the shortest paths.

tedious to use paths referred to as basic path set in [8], where one needs to determine the corresponding basic path from the given set by solving some equation before the forwarding direction of a packet can be deduced after appropriate rotations. In contrast, the presented new approach provides a convenient way to express the paths and forwarding directions by means of d-flow density $\mathbf{J}(\mathbf{r},\mathbf{x})$, i.e., $\vartheta(\mathbf{r},\mathbf{x}) = \arg\mathbf{J}(\mathbf{r},\mathbf{x})$. In some sense the paths are obtained in parallel with no extra effort when minimizing the maximal scalar flux. Assuming $\mathbf{J}(\mathbf{r},\mathbf{x})$ can be expressed in a compact form (e.g., by suitable Taylor series) one simply needs to provide this information to all nodes. This is demonstrated in the next Section V.

V. Examples

In this section we explore the use of the new formulation for the load balancing problem in disk and square geometries. We start by considering specific solutions to (8) and then proceed to finding an approximate solution $\psi(\mathbf{r},\mathbf{x})$ of the minmax problem (10). We assume uniform traffic demands, $\lambda(\mathbf{r},\mathbf{x}) = \Lambda/A^2$, for which the d-flow intensity of the shortest paths is given by

$$\mathbf{J}_{\rm SP}(\mathbf{r},\mathbf{x}) = \frac{\Lambda}{A^2}\left(d + \frac{d^2}{2}\right)(\mathbf{x}-\mathbf{r}), \tag{11}$$

where $d = b/a$ and for a and b we refer to Fig. 2.

A. Unit Disk

First we consider the standard example of a unit disk.

1) Specific solutions: In (9) any specific solution $\mathbf{J}_0(\mathbf{r},\mathbf{x})$ can be used; any other solution can be obtained by an appropriate choice of $\psi(\mathbf{r},\mathbf{x})$. As noted above, one solution is readily available, namely the d-flow field associated with the shortest path routing. For unit disk with uniform traffic demands, $\lambda(\mathbf{r},\mathbf{x}) = \Lambda/\pi^2$, we have from (11) for a sink on the x-axis at the point $(x,0)$,

$$\mathbf{J}_{\rm SP}(\mathbf{r},x) = \frac{\Lambda}{2\pi^2}\cdot \left[\frac{\left(x(x-r_x)+\sqrt{(x-r_x)^2+(1-x^2)r_y^2}\right)^2}{((x-r_x)^2+r_y^2)^2}-1\right](x-r_x,-r_y). \tag{12}$$

Alternatively, one can find a specific solution to (8) that is irrotational. Synonymously, an irrotational vector field is also called a potential flow, as it can expressed as the gradient of a scalar potential. Such a solution is provided by well-known physical systems such as heat or electrical conduction in a planar area with constant heat or electrical conductivity. In [9] we noted that a solution for the heat conduction between point source at $\mathbf{x}_1$ and sink at $\mathbf{x}_2$ in a unit disc, i.e., the irrotational, $\nabla\times\mathbf{K}(\mathbf{r};\mathbf{x}_1,\mathbf{x}_2) = \mathbf{0}$, solution for the equation

$$\nabla\cdot\mathbf{K}(\mathbf{r};\mathbf{x}_1,\mathbf{x}_2) = \delta(\mathbf{r}-\mathbf{x}_1) - \delta(\mathbf{r}-\mathbf{x}_2),$$

is given by (with $\mathbf{r} = (x,y)$)

$$\mathbf{K}(\mathbf{r};\mathbf{x}_1,\mathbf{x}_2) = -\frac{1}{4\pi}\nabla\log\left(\frac{(x-x_2)^2+(y-y_2)^2}{(x-x_1)^2+(y-y_1)^2}\cdot \frac{1-2xx_2-2yy_2+(x^2+y^2)(x_2^2+y_2^2)}{1-2xx_1-2yy_1+(x^2+y^2)(x_1^2+y_1^2)}\right).$$

From this, the solution to (8) can be obtained by integrating over the source points. To be precise, the d-flow intensity as defined in [9] is given by

$$\mathbf{J}(\mathbf{r},\mathbf{x}) = \int \lambda(\mathbf{x}',\mathbf{x})\,\mathbf{K}(\mathbf{r};\mathbf{x}',\mathbf{x})\,d^2\mathbf{x}'.$$

For a system with uniform traffic demands, however, this is unnecessarily complicated. A simpler way is provided by the observation that if $\mathbf{J}(\mathbf{r},\mathbf{x}_1)$ is the d-flow density to $\mathbf{x}_1$ at $\mathbf{r}$, then the d-flow density to $\mathbf{x}_2$ is

$$\mathbf{J}(\mathbf{r},\mathbf{x}_2) = \mathbf{J}(\mathbf{r},\mathbf{x}_1) + \frac{\Lambda}{\pi}\mathbf{K}(\mathbf{r};\mathbf{x}_1,\mathbf{x}_2),$$

together with the observation that for sink at the origin $\mathbf{x}_1 = \mathbf{0}$ we have complete symmetry and the irrotational flow field is trivially a pure inward radial flow

$$\mathbf{J}(\mathbf{r},\mathbf{0}) = -\frac{\Lambda}{\pi}\frac{1-r^2}{2\pi r}\,\hat{\mathbf{r}},$$

where $\hat{\mathbf{r}}$ is the radial unit vector. Thus the irrotational solution for the unit disk with uniform traffic demands is,

$$\mathbf{J}_{\rm NR}(\mathbf{r},\mathbf{x}) = \frac{\Lambda}{\pi}\left(-\frac{1-r^2}{2\pi r}\,\hat{\mathbf{r}} + \mathbf{K}(\mathbf{r};\mathbf{0},\mathbf{x})\right).$$

Explicitly we have in polar coordinates for a sink at $(x,0)$,

$$\begin{aligned} J_{\rm NR}(r,\theta,x)_r &= \frac{\Lambda}{2\pi^2}\,\frac{1-r^2}{r}\cdot \left(\frac{(1+r^2)\,x-(1+x^2)\,r\cos\theta}{(r^2-2rx\cos\theta+x^2)(r^2x^2-2rx\cos\theta+1)}\,x-1\right), \\ J_{\rm NR}(r,\theta,x)_\theta &= -\frac{\Lambda}{2\pi^2}\cdot \frac{(1+r^2)(1+x^2)-4\,rx\cos\theta}{(r^2-2rx\cos\theta+x^2)(r^2x^2-2rx\cos\theta+1)}\,x\sin\theta. \end{aligned} \tag{13}$$

The d-flows corresponding to this solution are depicted in Fig. 6 for different destinations on the positive x-axis.

2) Optimizing $\psi(\mathbf{r},\mathbf{x})$: From the principal point of view, it is immaterial which specific solution is used as the starting point in (9), since all solutions can be obtained from any one by adding $\hat{\mathbf{k}}\times\nabla\psi$. For numerical optimization methods $\mathbf{J}_{\rm NR}(\mathbf{r},\mathbf{x})$ is, however, a better starting point than $\mathbf{J}_{\rm SP}(\mathbf{r},\mathbf{x})$ as it alone gives a lower maximal scalar flux.

Without loss of generality assume that the destination $\mathbf{x}$ is on the positive x-axis at the point $(x,0)$. The specific solution $\mathbf{J}_0(\mathbf{r},\mathbf{x})$ has mirror symmetry with respect to the x-axis. We require that also the added term in (9) has the same

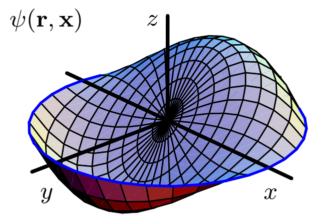

Fig. 3. Shape of the ψ-function of (14) for any destination on the positive x-axis. In the upper half disk there is a hill and the lower half disk a valley.

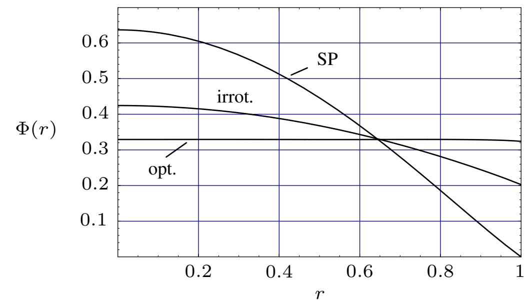

Fig. 4. Scalar flux distributions $\Phi(r)$ corresponding to shortest paths (SP), irrotational d-flows of (13) (irrot.), and optimized d-flows of (14) (opt.).

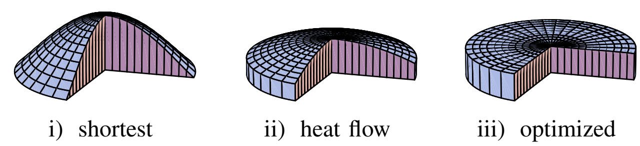

Fig. 5. Scalar flux distribution $\Phi(\mathbf{r})$ with shortest paths (12), paths according to heat conduction (13), and optimized paths of (14) with $c = 0.898$.

symmetry. One easily sees that this implies that $\psi(\mathbf{r},\mathbf{x})$ has to be *antisymmetric* with respect to the x-axis. This in turn implies that $\psi(\mathbf{r},\mathbf{x})$ is 0 on the x-axis, and further because $\psi(\mathbf{r},\mathbf{x})$ is constant on the perimeter, this constant is 0.

A function $\psi(\mathbf{r},\mathbf{x})$ that turns (eastward) flow away from the center is of the type that it has a hill in the upper half disk and a mirror symmetric valley in the lower one; in between and on the perimeter the ground level is zero. On the south side of a hill the gradient points to the north and $\hat{\mathbf{k}} \times \nabla\psi$ points to the west. Similarly, considering other sides of the hill, we find that the field $\hat{\mathbf{k}} \times \nabla\psi$ circulates around the hill clockwise. For a valley, the field circulates in counter clockwise direction. Adding a clockwise circulation in the upper half and a counter clockwise circulation in the lower half, turns the flow away from the center; the flow intensity becomes lower in the middle and larger closer to the perimeter. This demonstrates also the trade-offs. As the only degree of freedom one can play with is adding a divergence-free pure circulation field of type $\hat{\mathbf{k}} \times \nabla\psi$, then if one decreases the flow intensity somewhere then necessarily the intensity grows somewhere else; the flow is just pushed (routed) to another place.

A possible systematic expansion for the $\psi(\mathbf{r}, x)$ presented in polar coordinates, $\psi(r,\theta,x)$, with the required symmetries is provided by the Fourier series,

$$\psi(r,\theta,x) = \sum_{n=1}^{\infty}\sum_{m=0}^{\infty} c_{n,m}(x)\sin n\theta \sin m\pi r.$$

Then (10) becomes a variational problem for determining the system of functions of one variable $c_{n,m}(x)$. In our numerical examples we use the simple expression

$$\psi(r,\theta,x) = \frac{\Lambda}{2\pi^2}\, c\, x\, r(1-r^2)\sin\theta, \tag{14}$$

with a single adjustable parameter c, see Fig. 3. With this trial, by optimizing c we obtain extremely good results. The minimum of the maximal $\Phi(r)$ is $0.329\,\Lambda$ (obtained with $c = 0.898$) and is to our knowledge the best result reported so far (we presented the result $0.343\,\Lambda$ in [9]). We believe that this new result is indeed very close to the optimal. For comparison recall that the maximal flux with SP is $0.637\,\Lambda$ and with the routes defined by the irrotational d-flows of (13) it is $0.434\,\Lambda$. The distance bound (4) gives the lower bound $0.288\,\Lambda$.

The load distribution $\Phi(r)$ is very flat as shown in Figs. 4 and 5. The d-flows are depicted in Fig. 6 for heat flow (lower row) and optimized paths (upper row). The effect of the added circulation field is particularly visible in cases where the destination is close to the perimeter. The field lines are turned away from the central area assuming a slightly bell-shaped form. In addition, at the left part of the x-axis the field lines emanate from it approximately at right angle.

In general there may be several optimal solutions to a given load balancing problem. In [9] we defined the area where all optimal solutions obtain the maximum scalar flux Φ_{opt} as the *bottleneck area*. Outside the bottleneck area there obviously is some degree of freedom (slack) with regard to the choice of the paths. This is similar to the load balancing problem in wired networks, where the optimal solution (found by solving an LP problem) often is such that in the central part of the network the links are fully utilized while elsewhere some links remain underloaded. In [9] it was argued that in the bottleneck area the optimal paths are unambiguous and bidirectional, i.e., the same path is used in both directions, $\mathbf{r} \to \mathbf{x}$ and $\mathbf{x} \to \mathbf{r}$. A closer inspection of the optimized paths illustrated in Fig. 6 reveals that these paths are not entirely bidirectional, which suggests that some small further improvement is still possible.

B. Unit Square

Here we consider a multi-hop network in unit square with uniform traffic demands, $\lambda(\mathbf{r},\mathbf{x}) = \Lambda$.

1) Shortest paths and lower bound: Scalar flux for the shortest paths (SP) is straighforward to compute also in this case (cf., [24]), yielding the maximum at the origin,

$$\Phi_{\text{SP}}(0,0) = \frac{\Lambda}{2}\left(\sqrt{2} + \operatorname{arcsinh} 1\right) \approx 1.148\,\Lambda.$$

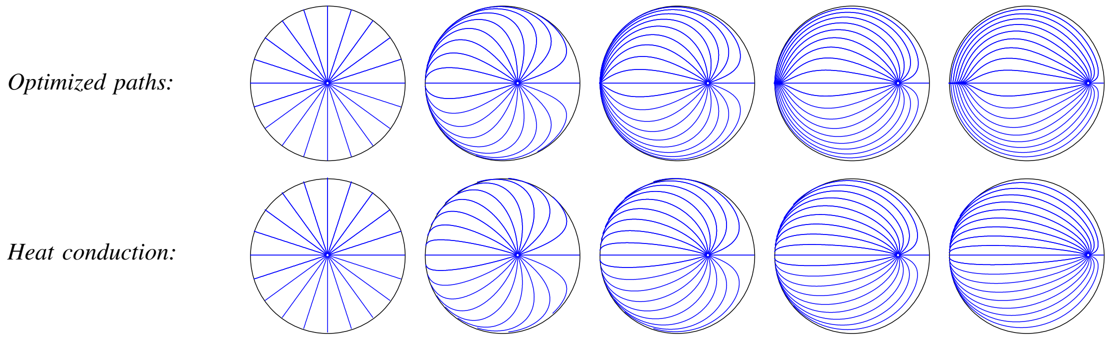

Fig. 6. Heat conduction paths and optimized paths in a disk for different locations of the destination.

The mean path length with SP routes, $\overline{\ell}_{\mathrm{SP}}$, is $(1/15) \cdot \big(2+\sqrt{2}+5\log(1+\sqrt{2})\big)$ and the distance bound (4) gives

$$\Phi_{\mathrm{opt}} \geq \frac{\Lambda \cdot \overline{\ell}_{\mathrm{SP}}}{A} = \Lambda \cdot \overline{\ell}_{\mathrm{SP}} \approx 0.521\,\Lambda.$$

2) Heat flow and optimizing $\psi(\mathbf{r},\mathbf{x})$: Similarly as in the case of a disk it is advantageous to start from a specific solution $\mathbf{J}_{\mathrm{NR}}(\mathbf{r},\mathbf{x})$ that is irrotational. This can again be found using the tools of complex analysis. Using the map $u(z) = \sin \pi z$ we get the complex flow field $J(z) = J_x(z) - iJ_y(z)$,

$$J(z, z_0) = \frac{u'(z)}{2\pi}\Big(\frac{1}{u(z)-u(z_0)} + \frac{1}{u(z)-u(z_0)^*}\Big)$$

for the flow in a semi-open tube $y > 0,\ -\frac{1}{2} \leq x \leq \frac{1}{2}$ with a sink at z_0. We can close the tube to form a square by the mirror image method leading to a rapidly converging series (excellent accuracy with six terms), add the flow $J_x = 0,\ J_y = y$ from the uniform source density, and finally shift the whole pattern downwards by $\frac{1}{2}$ to get the square in the position $-\frac{1}{2} \leq y \leq \frac{1}{2}$. Using the heat flow paths, the maximum of the scalar flux, attained at the origin, is $0.770\,\Lambda$.

For the ψ-function we use the trial

$$\psi(\mathbf{r},\mathbf{x}) = c\,\Lambda\,\hat{\mathbf{k}}\cdot\mathbf{x}\times\mathbf{r}\ \left(\frac{1}{4}-r_1^2\right)^d\left(\frac{1}{4}-r_2^2\right)^d,$$

where the factor $\left(\frac{1}{4}-r_1^2\right)^d\left(\frac{1}{4}-r_2^2\right)^d$ ensures that the boundary conditions are satisfied, and the vector product $\hat{\mathbf{k}}\cdot\mathbf{x}\times\mathbf{r}$ provides the desired circulation effect for paths to avoid the congested center; c and d are free optimization parameters. The optimum is obtained at $c = 5.99$ and $d = 1.06$ with

$$\max_{\mathbf{r}} \Phi(\mathbf{r}) \approx 0.650\,\Lambda,$$

corresponding to a 43 % improvement to the shortest path routing. The resulting scalar flux distribution, shown in Fig. 7, is flat except near the corners. Field lines of the heat flow and optimized paths are shown in Fig. 8. For comparison, the same trial ψ-function with fixed $d = 1$ when applied to SP routes yields a maximal flux of about $0.671\,\Lambda$ with $c = 21.1$.

To the best of our knowledge, the problem of finding paths that minimize the scalar flux in a rectangular area has not been studied explicitly in the literature. For example, Popa et al. in [22] propose a heuristic path set for unit disk referred

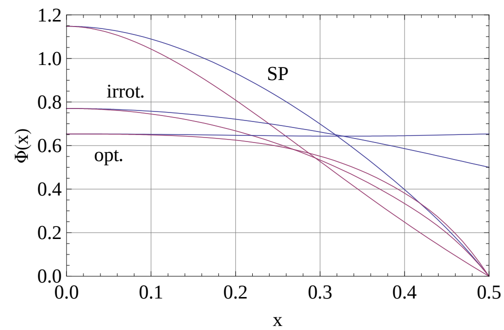

Fig. 7. Scalar flux in square for shortest paths, heat flow paths, and optimized paths. In each pair of curves the upper curve is along a cut from the origin to the center of a side and the lower curve along a diagonal cut.

to as curveball routes. They do evaluate these paths also for square area, for which they report a decrease of the maximal scalar flux by 18% to 24% in comparison to shortest paths depending on the parameter values. Consequently, we believe that the optimized paths presented in this paper yield the best reported performance in terms of load balancing.

VI. Conclusions

In this paper we have studied the routing problem in dense wireless multi-hop networks. In this limit the paths can be modeled as continuous curves between source-destination pairs. The objective in the load balancing problem is to find such paths that minimize the maximal scalar flux. Our main contribution is the new formulation of the problem entirely in terms of so-called d-flows. The most general solution for a d-flow can be expressed in terms of a scalar function $\psi(\mathbf{r},\mathbf{x})$ and the load balancing problem can be written as an optimization problem regarding this function. When the solution is available, the actual routes are obtained as the flow lines of the d-flows. In the given form, the problem is amenable to systematic numerical approaches. In particular, the new approach is notably more efficient than the previous one, where the problem formulation requires working explicitly with all possible paths for which the computation of the scalar flux is itself a complex and computationally heavy task.

The new methodology was illustrated with two example multi-hop networks with uniform traffic demands. By using simple trial functions, we were able to find paths that yield

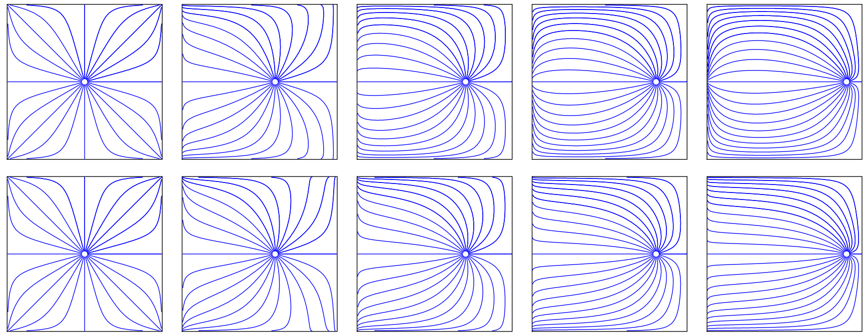

Fig. 8. Heat conduction paths and optimized paths in a square for different locations of the destination.

considerably better results than any of the previously published work to the best of our knowledge. The solution obtained for unit disk has a very flat load distribution, which is presumably very close to the real optimum. Also for unit square the resulting traffic load distribution is flat except near the corners.

The future work includes developing efficient algorithms to find optimal paths in distributed and scalable fashion. Another topic for future research is to investigate how well the optimal geometric routing works in real large scale networks.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, 2002.

[2] I. F. Akyildiz and W. Xudong, "A survey on wireless mesh networks," *IEEE Communications Magazine*, vol. 43, no. 9, Sep. 2005.

[3] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey," *Computer Networks*, vol. 47, no. 4, pp. 445–487, Mar. 2005.

[4] P. Jacquet, "Geometry of information propagation in massively dense ad hoc networks," in *ACM MobiHoc*, Roppongi Hills, Japan, 2004.

[5] S. Toumpis, "Mother nature knows best: A survey of recent results on wireless networks based on analogies with physics," *Computer Networks*, vol. 58, pp. 360–383, 2008.

[6] E. Hyytiä and J. Virtamo, "On load balancing in a dense wireless multihop network," in *NGI 2006, 2nd Conference on Next Generation Internet Design and Engineering*, València, Spain, Apr. 2006.

[7] G. Bell and S. Glasstone, *Nuclear Reactor Theory*. Reinhold, 1970.

[8] E. Hyytiä and J. Virtamo, "On traffic load distribution and load balancing in dense wireless multihop networks," *EURASIP Journal on Wireless Communications and Networking*, Jun. 2007.

[9] ——, "On optimality of single-path routes in massively dense wireless multi-hop networks," in *Proc. of MSWiM'07*, Crete Island, Greece, Oct. 2007, pp. 28–35.

[10] M. Kalantari and M. Shayman, "Routing in wireless ad hoc networks by analogy to electrostatic theory," in *Proc. of IEEE ICC'04*, Jun. 2004.

[11] D. Niculescu and B. Nath, "Trajectory based forwarding and its applications," in *Proc. of MobiCom'03*, San Diego, CA, USA, Sep. 2003.

[12] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*. Kluwer, 1996, vol. 353, pp. 153–181, chapter 5.

[13] F. Kuhn, R. Wattenhofer, and A. Zollinger, "Asymptotically optimal geometric mobile ad-hoc routing," in *Proc. of DIALM*, Sep. 2002.

[14] F. Kuhn, R. Wattenhofer, Y. Zhang, and A. Zollinger, "Geometric ad-hoc routing: Of theory and practice," in *Proc. of ACM PODC'03*, 2003.

[15] H. Frey, "Scalable geographic routing algorithms for wireless ad hoc networks," *IEEE Network Magazine*, vol. 18, no. 4, July-Aug. 2004.

[16] P. P. Pham and S. Perreau, "Performance analysis of reactive shortest path and multi-path routing mechanism with load balance," in *Proc. of IEEE Infocom '03*, vol. 1, San Francisco, USA, March-April 2003.

[17] S. Kwon and N. Shroff, "Paradox of shortest path routing for large multi-hop wireless networks," in *Infocom'07*, Anchorage, Alaska, 2007.

[18] Y. Ganjali and A. Keshavarzian, "Load balancing in ad hoc networks: Single-path routing vs. multi-path routing," in *Proc. of IEEE Infocom '04*, Hong Kong, Mar. 2004.

[19] S. Toumpis and L. Tassiulas, "Optimal deployment of large wireless sensor networks," *IEEE Trans. Information Theory*, vol. 52, no. 7, 2006.

[20] R. Catanuto and G. Morabito, "Optimal routing in dense wireless multihop networks as a geometrical optics solution to the problem of variations," in *Proc. of IEEE ICC'06*, Istanbul, Turkey, Jun. 2006.

[21] R. Catanuto, S. Toumpis, and G. Morabito, "Opti{c,m}al: Optical/optimal routing in massively dense wireless networks," in *Proc. of Infocom '07*. Anchorage, Alaska: IEEE, May 2007.

[22] L. Popa, A. Rostami, R. M. Karp, C. Papadimitriou, and I. Stoica, "Balancing the traffic load in wireless networks with curveball routing," in *Proc. the 8th ACM MobiHoc*, Montréal, Canada, Sep. 2007.

[23] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 6th ed. Elsevier Academic Press, 2005.

[24] E. Hyytiä, P. Lassila, and J. Virtamo, "Spatial node distribution of the random waypoint mobility model with applications," *IEEE Trans. Mobile Computing*, vol. 5, no. 6, pp. 680–694, Jun. 2006.

Appendix

A. *Proof of* (9)

We need to show that the general solution to partial differential equation,

$$\nabla \cdot \mathbf{J}(x, y) = s(x, y),$$

can be expressed as

$$\mathbf{J}(x, y) = \mathbf{J}_0(x, y) + \hat{\mathbf{k}} \times \nabla\psi(x, y),$$

where $\mathbf{J}_0(x, y)$ is a specific solution and $\psi(x, y)$ a function $\mathbb{R}^2 \to \mathbb{R}$.

Proof:

Let $\mathbf{J}$ denote an arbitrary solution and define $\mathbf{E} \equiv (E_1, E_2) = \mathbf{J} - \mathbf{J}_0$. The task is to show that $\mathbf{E}$ is of form $\hat{\mathbf{k}} \times \nabla\psi = \left(\frac{\partial\psi}{\partial y}, -\frac{\partial\psi}{\partial x}\right)$ for some $\psi = \psi(x, y)$. Clearly,

$$\nabla \cdot \mathbf{E} = \frac{\partial E_1}{\partial x} + \frac{\partial E_2}{\partial y} = 0.$$

One can always find $\psi^*(x, y)$ such that $E_1 = \frac{\partial}{\partial y}\psi^*(x, y)$. Substituting this into the above equation yields

$$\frac{\partial^2\psi^*}{\partial x\partial y} + \frac{\partial E_2}{\partial y} = 0 \quad \Rightarrow \quad E_2 = -\frac{\partial\psi^*}{\partial x} + f(x).$$

Defining $\psi(x, y) = \psi^*(x, y) - F(x)$, where $\partial F(x)/\partial x = f(x)$, then gives

$$E_1 = \frac{\partial\psi}{\partial y} \quad \text{and} \quad E_2 = -\frac{\partial\psi}{\partial x}.$$

■