Topology Control in Large-Scale Wireless Sensor Networks: Between Information Source and Sink$^{☆}$

Masoumeh Haghpanahi*, Mehdi Kalantari, Mark Shayman

Department of Electrical & Computer Engineering, University of Maryland-College Park, College Park, MD 20742, United States

Abstract

With non-uniform traffic patterns in wireless sensor networks due to the many-to-one nature of communications, the traditional definition of connectivity in graph theory does not seem to be sufficient to satisfy the requirements of sensor networks. In this work, a new notion of connectivity (called path-implementability) is defined which represents the ability of sensor nodes to relay traffic along a given direction field (referred to as information flow vector field) $\vec{D}$. The magnitude of information flow is proportional to the traffic flux (per unit length) passing through any point in the network, and its direction is toward the flow of traffic. The flow field may be obtained from engineering knowledge or as a solution to an optimization problem. In either case, information flow flux lines are abstract paths that are assumed to represent desired paths for flow of traffic. In a sensor network with a given flow field $\vec{D}(x, y)$, we show that a density of $n(x, y) = O(|\vec{D}(x, y)|^2)$ sensor nodes is not sufficient to implement the flow field. On the other hand, by increasing the density of wireless nodes to $n(x, y) = O(|\vec{D}(x, y)|^2 \log |\vec{D}(x, y)|)$, the flow field becomes path-implementable. Path-implementability requires more nodes than simple connectivity. However, it guarantees existence of enough paths connecting the information source to the sink so that all the traffic can be transmitted to the sink. We also propose a joint MAC and routing protocol to forward traffic along the flow field; the proposed tier-based scheme can be further exploited to build lightweight protocol stacks which meet the specific requirements of dense sensor networks.

Keywords: Dense sensor networks, Network connectivity, Optimal deployment, Scaling laws, Traffic modeling

*Corresponding author. Tel: +1 (650)796-6723, Address: 6 Landing view lane, Fairport, NY, 14450.

Email addresses: masoumeh@eng.umd.edu (Masoumeh Haghpanahi), mehkalan@eng.umd.edu (Mehdi Kalantari), shayman@eng.umd.edu (Mark Shayman)

1. Introduction

1.1. Path-implementability: a new notion of connectivity

A Wireless Sensor Network (WSN) is composed of a large number of sensor nodes densely deployed inside or very close to a phenomenon. The exact position of sensors need not be predetermined, and they can be randomly deployed in inaccessible terrains. Such networks are usually aimed at monitoring vast regions, much larger than the transmission radius of an individual node. Therefore, data transmission occurs in multihop communication. There has arisen a great demand for sensor networks in many fields including military, environment monitoring, transportation systems and agriculture.

WSN applications are classified based on their data-delivery requirements and their traffic characteristics. According to the work [1], most of the current WSN applications fall into one of the following broad classes of (i) monitoring and periodic reporting, (ii) event detection and reporting, (iii) sink-initiated reporting, (iv) object detection and tracking, and (v) hybrid applications with more than one of the above four characteristics.

Specific characteristics of WSNs distinguish these networks from ad hoc networks. Unlike in an ad hoc network, where any node can potentially communicate with any other node, a WSN exhibits the many-to-one communication paradigm. Also, sensor nodes have very limited energy, processing, and storage capacities. Thus, they require careful resource management. Moreover, design requirements of a sensor network change with its application. For example, while low latency is of utmost importance in applications such as tactical surveillance, extending the network lifetime is more important for applications such as periodic weather monitoring [2].

It is desirable to present a framework to study WSNs that can support more than one application and to provide a means to calculate required density of sensor nodes under different scenarios. To this end, authors in [3] have introduced a continuous space model called information flow vector field, with two components (magnitude and direction) at each point in the Euclidean space; the magnitude of information flow is proportional to the traffic flux (per unit length) passing through any point in the network, and its direction is toward the flow of traffic. This is very similar to visualization of fluid flows or electric fields by means of vector fields. In fluid flow visualization for example, the flow vector field represents the fluid velocity field and flow streamlines are trajectories of massless particles that are traversed by fluid particles in the flow field as a function of time. Similarly, information flow flux lines are a set of abstract paths used for data transmission to the destination. Note that flow flux lines are not constrained by the location of sensor nodes.

In a particular application, the flow field may be obtained from engineering knowledge. Alternatively, it may be obtained as a solution to an optimization problem. For example in [4], it is obtained as the solution to a *p-norm* optimization problem by minimizing the *p-norm* of the information flow vector field subject to the basic flow constraints, where $p > 1$ is a design parameter. The *p-norm* problem has load balancing property; for p close to 1, the information

2

flow tends to pass through geometric shortest paths. By increasing the value of p, the flow spreads in the network using resources more evenly, and it reaches maximum load balancing when $p \to \infty$.

In this work, we study the required density of nodes to forward traffic along information flow field. We show that following the optimal (desired) paths for flow of traffic requires more nodes than providing network connectivity. To perform this analysis, a new notion of connectivity called path-implementability is defined which represents the ability of sensor nodes to forward traffic along flux lines (also referred to as flow streamlines) of a known flow field as its magnitude scales linearly to infinity. The final results indicate that for a given flow field $\vec{D}(x, y)$, a density of $n(x, y) = O(|\vec{D}(x, y)|^2)$ sensor nodes is not sufficient to relay traffic along flow streamlines. This result contradicts the prior result of [5] where authors stated otherwise. On the other hand, we show that $\vec{D}(x, y)$ is path-implementable with a density of $n(x, y) = O(|\vec{D}(x, y)|^2 \log |\vec{D}(x, y)|)$ sensor nodes.

Note that previous results on throughput and connectivity in wireless networks were mainly derived based on symmetry and uniformity, either by assuming symmetric network geometries and/or by random matching of nodes into source-destination pairs. In most wireless sensor network applications, however, all the traffic generated in the network must be transmitted to a single destination; this results in a non-uniform traffic pattern. Therefore, a new notion of connectivity is required for such networks.

The notion of path-implementability not only provides a framework to study the density of sensor nodes required to follow a set of desired paths for flow of traffic, but also guarantees existence of enough paths connecting the information source[1] to the destination so that all the traffic can be transmitted to the destination.

We propose a joint MAC and routing protocol which requires no more than $n(x, y) = O(|\vec{D}(x, y)|^2 \log |\vec{D}(x, y)|)$ nodes to relay traffic along flow streamlines of $\vec{D}$. The proposed tier-based scheme also provides a framework to reduce the size of addressing fields (by requiring tier-level addressing instead of node-level addressing) and apply power-saving methods by using clustering techniques, and hence meeting the specific requirements of dense sensor networks. The details of addressing and clustering techniques should be specified based on network geometry and are out of the scope of this work.

The rest of this paper is organized as follows. After a review of notations and definitions used throughout the work in Section 2, information flow vector fields are defined in Section 3. Main results of the paper are summarized in Section 4. Next, a formal definition of path-implementability is presented in Section 5, and different steps involved in the implementability analysis are described in Sections 6-8 in a constructive manner. Finally, the paper is concluded in Section 9.

---

[1]Note that the term "information source", or as later defined "distributed source", refers to the region where all the traffic is generated and does not refer to a single sensor node.

1.2. Related work

Great efforts have been devoted in the literature to address different challenges such as network dynamics, energy efficiency, coverage and connectivity, and data aggregation and compression in wireless sensor networks. Some of the major works and results in these areas are summarized in surveys [6–9].

A vector field representation for flow of traffic in WSNs was first introduced in [3, 10] with inspiration from electrostatics. Approaches based on electrostatics have become popular thereafter. References [5], [11], [12], and [13] are among the works based on this analogy.

In [3], Kalantari and Shayman performed simulations to compare routing along information flow field with routing along geometric shortest paths to the destination. To this end, the optimal flow vector field for a special network set up was calculated and the paths from sensor nodes to the destination were found by following the flow field. The simulation results showed that the number of delivered messages to the destination increased significantly when following the routes generated by information flow field.

Focusing on routing algorithms, authors in [11] proposed a routing protocol by use of electrostatic potentials. The proposed solution introduces a potential field that covers the network such that each node is associated with a given potential value. The packets are then routed through nodes with a decreasing potential value. Similarly, authors in [12] present a heat-inspired model for link-diversity routing. The heat equation in steady state, similar to electrostatic fields, follows the Laplace equation. To exploit this property for routing, the destination in a network is modeled as a heat source. Then, the heat flow resulting from the source at all the nodes is evaluated according to the network connectivity. Once all the values are calculated, finding a path from the source to the destination is a gradient search problem. This way, the path a packet follows is analogous to the minimum-energy diffusion path of a particle in a real temperature field.

In [13], Nguyen et al. considered establishing a reliable framework for routing by using an analogy with electric fields. The authors proposed the use of several electric flux lines as routes generated by placing a positive charge in the source and a negative charge in the destination.

The spatial distribution of wireless nodes that can transport a given volume of traffic in a sensor network was first studied in [5]. Under a general assumption on the physical and medium access control (MAC) layers, the authors showed that the optimal distribution of nodes induces a traffic flow identical to the electrostatic field that would exist if the sources and sinks of traffic were substituted by an appropriate distribution of electric charge. This result was based on the assumption that a location (x, y) with node density $n(x, y)$ can support an information flow with magnitude less than or equal to $\lambda \sqrt{n(x, y)}$ for some constant λ. Although the quadratic relation between node and traffic density was shown to hold in a rather simplistic setting, in this work, we show that this assumption is not realistic under more accurate communication protocols.

The authors of [5] have also stated that the results of percolation theory inside square geometries [14] support existence of a quadratic relation between

4

node and traffic density. It is important to note that percolation theory does not impose any constraints on paths traversing a square geometry and connecting one side to the other side of the square. Hence, percolation results cannot be applied to wireless sensor networks where traffic is relayed from the information source (the region where all the traffic is generated) to the destination. This is because, as shown in Fig. 1, percolating paths inside individual square elements do not guarantee existence of connected paths from the information source to the sink. Therefore in such networks, it is necessary to not only study the node density that can support a given traffic density, but also to provide a means to guarantee existence of connected paths from the information source to the destination. In this work, we show that this occurs at the cost of increasing the node density.
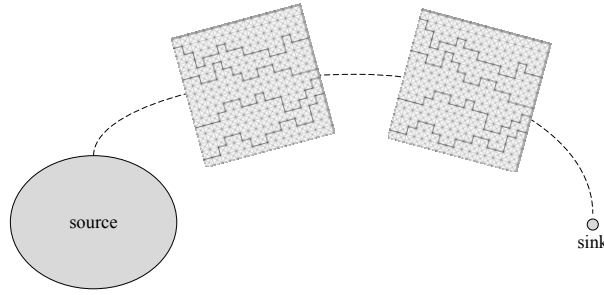

Figure 1: Percolating paths inside square elements cannot form connected paths from the information source to the sink.

2. Notations and Definitions

2.1. Notations

Throughout this work, X $= (x, y)$ represents the Cartesian coordinates of a point in Euclidean plane $\mathbb{R}^2$. The notation $\vec{v} = (v_x, v_y)$ is used to refer to a vector field, where v_x and v_y denote the field components along the x and y axes, respectively.

In partial differential equations, the nabla operator ∇ in two-dimensional Cartesian coordinates is defined as a vector differential operator $\nabla \equiv \frac{\partial}{\partial x} \vec{a}_x + \frac{\partial}{\partial y} \vec{a}_y$, where $\vec{a}_x$ and $\vec{a}_y$ are the unit vectors along x and y axes, respectively.

The divergence of a vector field $\vec{v}$ at a point is defined as the net outward flux of $\vec{v}$ per unit area as the area about the point tends to zero [15]. For an area with an enclosed contour there will be an excess of outward or inward flux through the contour only when the area contains a source or a sink, respectively. The net outward flux per unit area is therefore a measure of the strength of the enclosed source. Using the above definition, the divergence of a vector field $\vec{v}$ in two-dimensional Cartesian coordinates can be shown to be equal to

$$\nabla \cdot \vec{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}. \tag{1}$$

We refer to fields with zero divergence as incompressible fields.

On the other hand, the curl of vector field $\vec{v}$ is denoted by $\nabla \times \vec{v}$; its magnitude represents the maximum net circulation of $\vec{v}$ per unit area as the area tends to zero, and its direction is perpendicular to the area [15]. In two-dimensional Cartesian coordinates,

$$\nabla \times \vec{v} = (-\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x})\vec{a}_z, \tag{2}$$
$$\text{where, } \vec{a}_z = \vec{a}_x \times \vec{a}_y$$

We refer to fields with zero curl as irrotational fields.

The p-norm of a vector $Z = (z_1, z_2, \cdots, z_m) \in \mathbb{R}^m$ is defined as

$$|Z|_p = \Big(\sum_{i=1}^{m} |z_i|^p \Big)^{\frac{1}{p}}, \text{ for } p \geq 1$$

Similarly, the p-norm of a function $f : A \to \mathbb{R}^2$ defined over a subset A of $\mathbb{R}^2$ is defined as

$$|f|_p = \Big(\int_A |f(x,y)|^p \mathrm{d}x\mathrm{d}y \Big)^{\frac{1}{p}}.$$

Another measure for calculating the norm of a real or complex valued function $f : \mathcal{E} \to \mathbb{C}$ defined over a subset $\mathcal{E}$ of $\mathbb{R}^m$ is called the *supremum norm*[2] or *uniform norm* and is defined as

$$|f|_u = \sup\{|f(Z)| : Z \in \mathcal{E}\}$$

Note that the supremum norm is a special case of the p-norm where $p = \infty$. It can also be used to define a mode of convergence; a sequence f_n of functions converges *uniformly* to a function f on a set $\mathcal{E}$ provided that $\lim_{n \to \infty} |f_n - f|_u = 0$.

Throughout the work, when performing linear algebra operations, a vector is treated as a column vector and B^T refers to the transpose of matrix B.

In this paper we are mainly concerned with events that occur with high probability (w.h.p); that is, with probability tending to one as the argument of the function representing the event goes to infinity.

We use standard notational conventions for asymptotic analyses. Hence, for two non-negative functions $f(\cdot)$ and $g(\cdot)$, $f(n) = O(g(n))$ w.h.p, if there exists a positive constant k such that

$$\lim_{n \to \infty} P(f(n) \leq kg(n)) = 1.$$

---

[2]The maximum might not exist if f is not continuous on its domain of definition, so the supremum is used instead [16].

Using the above notation, $f(n) = \Omega(g(n))$, as $n \to \infty$, if $g(n) = O(f(n))$. Also, $f(n) = o(g(n))$, if $\lim_{n \to \infty} \frac{f(n)}{g(n)} = 0$.

2.2. Definitions

Consider a general network with predefined sources and sinks, where N nodes are distributed uniformly at random and independent from each other throughout the network. Data is sent from node to node in a multihop fashion from a source to an intended sink. For simplicity, we assume that time is slotted. Due to spatial diversity, at each time instant more than one node can transmit data to its next hop node, without facing destructive interference from other transmissions at the same time. Two models for multihop communication called the *Protocol Model* and *Physical Model* have been proposed and widely used in the literature [17]. In a homogeneous scenario, where sensor nodes have same transmission radius r_t and transmission power p_t, the aforementioned communication models can be described as below. Let X_i denote the location of a node as well as the node itself. Now, a transmission from node X_i to node X_j is successful

1. Under the Protocol Model if:
   (a)

$$|X_i - X_j| \le r_t, \tag{3}$$

   where $|\cdot|$ represents the Euclidean distance between two points in the Cartesian plane, and
   (b) for every other node X_k simultaneously transmitting over the same channel,

$$|X_k - X_j| \ge (1 + \Delta)r_t \tag{4}$$

   $\Delta > 0$ is a nonnegative parameter which models a guard zone around the receiving node to limit interference from other nodes. It imposes the constraint that no other neighboring nodes within the guard zone of the receiver can simultaneously send information in the same channel.

2. Under the Physical Model if: the *Signal to Interference Plus Noise Ratio (SINR)* is above a certain threshold Γ; i.e.,

$$SINR = \frac{\frac{p_t}{|X_i - X_j|^\alpha}}{N_0 + \sum_{\substack{k \in T \\ k \neq i}} \frac{p_t}{|X_k - X_j|^\alpha}} \ge \Gamma, \tag{5}$$

   where $\{X_k, k \in T\}$ is the set of nodes simultaneously transmitting over the same channel as X_i, and N_0 refers to the ambient noise power level. This models a situation where signal power decays with distance r as $\frac{1}{r^\alpha}$. The loss factor $\alpha > 2$ in practical situations.

While the Protocol Model describes the effects of interference based on a pairwise relation between nodes, the Physical Model takes the aggregate interference into account. Because of this, the Physical Model is generally considered to be more accurate than the Protocol Model. However, the simplicity of the Protocol Model has motivated its use in design and evaluation of communication protocols [18].

In order to reduce the complexity of studying successful transmissions under the Physical Model, approximate versions of this model have been proposed in the literature, especially for network simulators [19, 20]. A Bounded Physical Interference model (BPI model) is defined in [21] which neglects the effects of interference from nodes residing outside a region (referred to as interference region) enclosing the receiver of a particular transmission. Interference regions are considered to be circular regions which can be characterized by a single parameter, referred to as the interference range. Hence, in a more general scenario where sensor nodes have different transmission powers, a transmission from node X_i to node X_j is successful under the BPI model if

$$SINR = \frac{\frac{p_{t,i}}{|X_i - X_j|^\alpha}}{N_0 + \sum_{\substack{k \in T \cap IR(X_j) \\ k \neq i}} \frac{p_{t,k}}{|X_k - X_j|^\alpha}} \geq \Gamma, \tag{6}$$

where $IR(X_j)$ denotes the interference region of node X_j.

3. Background

In this section, we present an overview of the definition of information flow vector field, and briefly point out the constraints it must satisfy in order to model flow of traffic in a general network. Although the concept of information flow can be applied to networks with multiple information sinks (destinations), for the sake of simplicity, we consider networks with only one information sink in this work. Generalizing information flow modeling to networks with multiple sinks is fully described in [22].

Consider sensor networks with the application of monitoring and sending periodic reports to the destination. The destination is assumed to have unlimited energy, space and processing power, and is usually far away from the region to be monitored. Network geometry A is a subspace of the Cartesian plane including the monitoring region and destination where sensor nodes are deployed to carry the traffic to the destination. This is shown in Fig. 2.

Let $\rho : A \to \mathbb{R}$ denote the load density function which determines the total rate of traffic generated inside any closed contour within the network geometry. Hence, it takes positive values inside the monitoring region. In the rest of this work, we refer to the closure of the area with positive load density function as *distributed source*[3]. Let l denote the boundary of this area, and $S(l)$ be the area surrounded by l. Hence, the distributed source is represented by $S(l)$.

---

[3]Note that we assume that every event is picked up by only one sensor inside the distributed
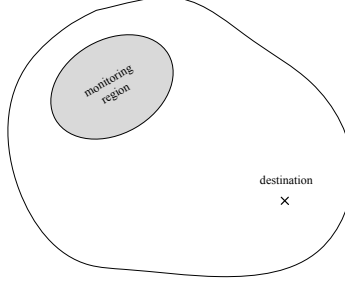
Figure 2: Network geometry A including the monitoring region and the destination.

The rate of traffic generated inside a closed contour $a \subset S(l)$ is denoted by $w(a)$ and is equal to

$$w(a) = \int_a \rho(x,y)\mathrm{d}x\mathrm{d}y$$

The value of load density function is defined to be zero for the region between the distributed source and the sink. Also, since all the traffic generated in the network must to be transmitted to the sink, the value of $\rho(x,y)$ at the sink can be defined as a Dirac delta function with weight $-w_0$, where $w_0 = \int_{S(l)} \rho(x,y)\mathrm{d}x\mathrm{d}y$.

The information flow field, denoted by $\vec{D}(x,y)$, is a vector field with two components (direction $\vec{Dir}(x,y)$ and magnitude $|\vec{D}(x,y)|$) at each point (x,y) in the network; $\vec{Dir}(x,y)$ represents the direction of flow of traffic at point (x,y), and $|\vec{D}(x,y)|$ represents the amount of traffic flux (per unit length) passing through (x,y). This implies that for any closed contour $c \subset A$,

$$\oint_c \vec{D}(x,y) \cdot \vec{\mathrm{d}n} = \int_{S(c)} \rho(x,y)\mathrm{d}x\mathrm{d}y, \tag{7}$$

where $\vec{\mathrm{d}n}$ is an outgoing differential normal vector to the boundary of c, and $S(c)$ is the area surrounded by closed contour c. Equation (7) is similar to Gauss' law in electrostatics, and can also be expressed in the form of the following partial differential equation:

$$\nabla \cdot \vec{D}(x,y) = \rho(x,y) \tag{8}$$

It must be noted that although we have assumed that the information sink is located at a point in the network, according to (7), non-zero dimensions must be chosen for the destination in order to avoid the magnitude of information flow to grow to infinity as it gets closer to the sink. Choosing non-zero dimensions

---

source. The challenges involved in data aggregation and correlation between messages are out of the scope of this work.

for data sink is also in accordance with networks where multiple antennas are located at a circle with non-zero radius around the sink.

In addition to equation (8), the information flow satisfies the Neumann boundary condition

$$\vec{D}_n(x,y) = 0, \quad \forall (x,y) \in \partial A \tag{9}$$

where $\vec{D}_n$ is the normal component of $\vec{D}$ on the boundary of A, denoted by ∂A. This constraint forces the information flow to stay within the network geometry without violating the boundaries. In reality, there are no sensors deployed outside of A; hence, no traffic passes through ∂A.

The above constraints on the flow field imply that for every location (x,y) inside A there exists a path from that point to the destination, which can be found by integrating the flow streamline from (x,y) all the way to the sink. Therefore, the information flow field represents a set of paths for transmitting traffic from the distributed source to the destination, which are not constrained by the location of sensor nodes.

Although the flow field satisfies equations (8) and (9), $\vec{D}(x,y)$ is not uniquely specified by these constraints. In a particular application, the flow field may be obtained from engineering knowledge. Alternatively, it may be obtained as a solution to an optimization problem. Additional constraints have been suggested to place on $\vec{D}$ such that the resulting vector field generates a desirable set of paths for traffic transmission to the sink.

For example, the quadratic cost function below was suggested in [10]:

$$J(\vec{D}) = \int_A |\vec{D}(x,y)|^2 \mathrm{d}x\mathrm{d}y$$

Minimizing the above cost function subject to the constraints on $\vec{D}$ can be summarized into the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & \int_A |\vec{D}(x,y)|^2 \mathrm{d}x\mathrm{d}y \tag{10}\\ \text{subject to} \quad & \nabla \cdot \vec{D}(x,y) = \rho(x,y)\\ & \vec{D}_n(x,y) = 0, (x,y) \in \partial A, \end{aligned}$$

which is a convex problem. It can be further proved that the optimal solution of (10) must have zero curl. Now the combination of the zero-curl property, together with the divergence property and the boundary constraint uniquely specify the optimal flow field $\vec{D}$.

The motivation behind using a quadratic cost function was to disperse traffic in the network. However, it does not achieve maximum load balancing in the network. Therefore, the following generalization of the quadratic optimization problem was suggested in [4]:
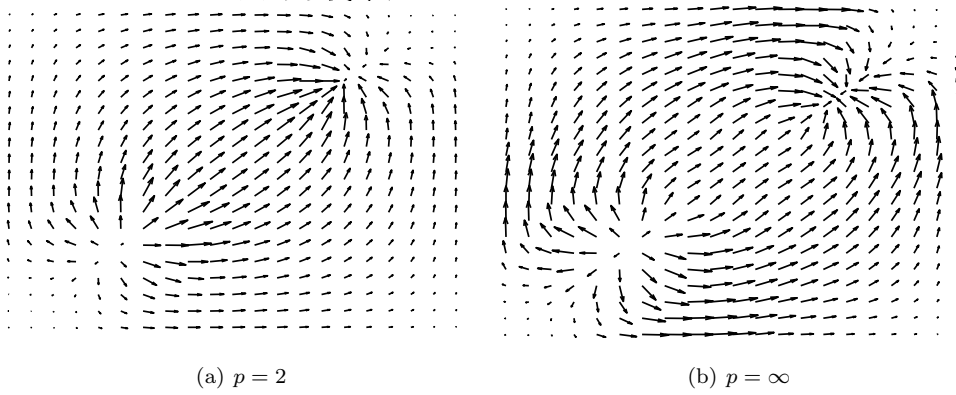
10

(a) $p = 2$ (b) $p = \infty$

Figure 3: The load balancing property of information flow increases as p changes from 2 to ∞ in the *p-norm* problem.

$$\text{minimize} \qquad \int_A |\vec{D}(x,y)|^p \mathrm{d}x\mathrm{d}y \tag{11}$$

$$\text{subject to} \qquad \nabla \cdot \vec{D}(x,y) = \rho(x,y)$$

$$\vec{D}_n(x,y) = 0, (x,y) \in \partial A,$$

in which $p > 1$.

Increasing p causes the optimization problem to increase spatial spreading in the network. To be more specific, when p is close to 1, most traffic tend to pass through geometric shortest paths from the source to the sink. As a result, while the average transport delay is less (compared to the case with higher values of p), sensor nodes along shortest paths will get depleted very fast which reduces the network lifetime. By increasing the value of p from 1, the optimization increasingly spreads the traffic in the network in order to use resources more evenly. This increases delay, however, sensor nodes are used more evenly and there is more load balancing compared to the case where p is smaller. The case $p \to \infty$ achieves maximum load balancing. The load balancing property of the *p-norm* problem is further shown in Fig. 3 in a square with a single source at the lower left and a single sink at the upper right side of the network.

The weighted version of the *p-norm* problem is also studied in [23], where it is shown that the weight function can be chosen cleverly to determine energy-efficient paths.

Given a time-invariant vector field $\vec{D}(x,y)$, flow flux lines (also called streamlines) of $\vec{D}$ from each location (x,y) in the network to the destination can be found by solving the following ordinary differential equation:

$$\frac{\mathrm{d}\mathrm{X}(t)}{\mathrm{d}t} = \vec{D}(\mathrm{X}(t)), \tag{12}$$

where t is the integration variable and is not to be confused with time. Therefore, each streamline can be represented parametrically from the source to the destination by $f(t) = (x(t), y(t); x_0, y_0)$ for $t_0 \le t \le t_1$, where $x_0 = x(t_0)$ and $y_0 = y(t_0)$ refer to the starting point of the streamline on the boundary of the distributed source and $x_1 = x(t_1)$ and $y_1 = y(t_1)$ refer to the corresponding landing point on the destination.

Later in this work, we consider network scenarios where the flow field scales linearly to infinity; i.e., we study networks with flow field $\vec{D}(x, y) = \theta \vec{D}_1(x, y), \forall (x, y) \in A$, where $\vec{D}_1$ is a given initial field and $\theta \in \mathbb{R}^+$. Note that by scaling the flow field as such the trajectory of flow streamlines does not change; however, the bounds of the parametrization variable t changes with the scaling factor θ. This can be mathematically represented as $f^\theta(t) = f^1(\theta t)$, where $f^\theta(t)$ is the parametric representation of a flow streamline of $\theta \vec{D}_1$. This is better understood by considering velocity fields; as the scaling factor θ increases, the trajectory of a flow streamline remains the same, however it is traversed faster.

4. Summary of Results

First we define a new notion of connectivity called path-implementability, which is an asymptotic approach where the load density function scales linearly to infinity.

Path-implementability provides a means to study the density of sensor nodes required to follow flow streamlines of a given flow field. The new notion of connectivity uses the magnitude and direction of information flow to form an increasing set of disjoint regions such that, as the load function scales linearly to infinity, traffic can be relayed to the destination in a multihop fashion at constant rate inside each region. For the flow field to be path-implementable, density of sensor nodes must be such that w.h.p piecewise linear routes can be formed inside each region. We show that in the limit, the set of piecewise linear routes converges uniformly to a set of flow streamlines, which by definition are the desired paths for flow of traffic.

To perform the implementability analysis, we first consider a simple example of a square with uniform information flow parallel to the sides of the square and study the required number of nodes to carry the traffic from one side to the other side of the square along a set of evenly-spaced flow streamlines. Next, we discuss the conditions based on which the results of the square example can be extended to more complicated network geometries. More specifically, we show that a general network geometry with a given flow field can be partitioned into smaller (approximately) square grids with uniform information flow inside each grid, if the flow field is irrotational. In other words, for irrotational flow fields, a flow-based coordinate system exists for the region between the source and destination and that the square network example can be interpreted as a curvilinear element of this coordinate system.

For the square example with uniform flow field $\vec{D}^{(l)}$, we first show that a quadratic relation between number of nodes and $|\vec{D}^{(l)}|$ is not sufficient for

12

path-implementability of the flow field. To show this, we prove that path-implementability requires more than simple connectivity by showing that a disconnected network of nodes cannot provide the requirements for following the flow streamlines; i.e., the probability that an isolated node prevents forming a piecewise linear route inside its associated region is strictly positive as $|\vec{D}^{(l)}| \to \infty$.

Next, we show that increasing the number of sensor nodes to $N = O(|\vec{D}^{(l)}|^2 \log |\vec{D}^{(l)}|)$ provides path-implementability. The sufficiency is shown by proposing a joint MAC and routing protocol which requires the same order of nodes to forward the traffic within each flow region using either the Protocol or Physical Model.

The main idea of the proposed scheduling scheme is to divide time and frequency domains into smaller slots and sub-channels and determine the transmission radius (power) of sensor nodes such that transmissions at each time slot and frequency sub-channel abide by either of the communication models. We tessellate the unit square to divide sensor nodes into different groups. The tessellation technique has become a rather common method in the literature to prove sufficiency bounds and has been previously used in prominent works such as [14, 24, 25].

5. Path-Implementability of Information Flow Field

The goal of this section is to provide a means to study the required density of nodes to communicate in a multihop fashion along flow streamlines.

Let C denote the transmission capacity of sensor nodes (measured in bps), and $0 < \gamma \leq 1$ be a parameter determined by the MAC protocol which represents the portion of time and/or frequency bandwidth assigned to each node for data transmission. Hence, with w_0 referring to the total amount of traffic generated at the distributed source, a total number of $n_s(A) = \lceil \frac{w_0}{\gamma C} \rceil$ routes is required to relay all the traffic from the source to the destination.

For a given information flow $\vec{D}(x, y)$, a set of $n_s(A)$ streamlines can be chosen to represent desired paths for flow of traffic. The goal of path-implementability is then to study the required density of nodes that can follow the chosen set of flow streamlines in a multihop fashion, such that traffic can be relayed along the resulting set of piecewise linear routes at constant rate simultaneously. In order to avoid intersection of the piecewise linear routes formed by nodes, the area between the distributed source and the destination must be partitioned into a set of disjoint regions (referred to as streamline regions), using the knowledge of flow streamlines.

The set of streamlines is chosen by placing flow seeds on the boundary of the distributed source and using relation (12) to integrate the flow lines all the way to the destination. Now, by bisecting the region between each pair of streamlines, network geometry A is partitioned into $n_s(A)$ streamline regions. Each region is the area between a pair of bisections and includes a streamline in its interior.

As an example, consider a rectangular area with a set of information flow streamlines passing through and parallel to the vertical sides of the rectangle,

similar to Fig. 4. This resembles streamlines passing through a small enough area element inside A. Using the same method described above, the rectangular area is partitioned into a set of streamline regions which in this case is a set of rectangular strips as shown in the figure.

Now consider a homogeneous scenario in which sensor nodes are deployed inside the rectangle uniformly at random. In order for sensor nodes to be able to form piecewise linear routes inside each streamline region, the transmission radius of nodes must be less than the width of the narrowest strip[4]. Otherwise, routes might violate the margins of their corresponding streamline regions[5].
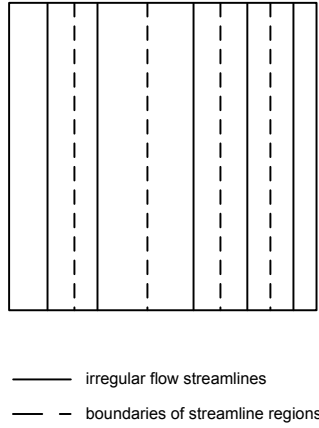

Figure 4: A set of irregular flow streamlines

As transmission radius of nodes decreases, more nodes must be deployed to form routes inside each strip. This is because density of nodes must be such that each node can find another node within its transmission radius. Therefore, required density of nodes is minimized if flow flux lines are evenly spaced within the rectangle.

Following a similar argument, the set of $n_s(A)$ streamlines chosen in A must be "locally regular" in order to minimize the required number of nodes inside each area element. A more precise definition of local regularity is stated in the sequel.

Implementability analysis is an asymptotic approach in which the probability of events are calculated as the load density function scales linearly to infinity. As it is stated in the following lemma, linear scale of $\rho(x, y)$ results in linear

---

[4]In order to ensure that the traffic inside each strip stays within itself, nodes inside adjacent strips must use different frequency sub-channels to carry traffic.

[5]More accurately, the transmission radius of nodes must be proportional to the width of the narrowest strip. The frequency channel must again be divided into sub-channels to conserve traffic within each streamline region.

scale of the traffic density (when derived from the *p-norm* problem), everywhere in the network.

Lemma 1. *Let $\vec{D}_1(x, y)$ denote the optimal solution of the p-norm problem in a region A with load density function $\rho_1(x, y)$. Then solving (11) for $\rho(x, y) = \theta \rho_1(x, y)$ results in $\vec{D}(x, y) = \theta \vec{D}_1(x, y)$.*

PROOF. See the Appendix.

From now on, we assume $\vec{D}(x, y) = \theta \vec{D}_1(x, y)$ for a given flow field $\vec{D}_1$ and scaling factor $\theta \in \mathbb{R}^+$. With the above introduction, a formal definition of implementability can be stated as follows.

Definition 1. *Let $\theta \in \mathbb{R}^+$ denote a positive scaling factor of $\vec{D}_1(x, y)$. Also, let $n(x, y; \theta)$ denote the density of nodes within the network, where $n(x, y; \theta) \to \infty$ as $\theta \to \infty$, $\forall (x, y) \in A$.*

Information flow $\vec{D}(x, y) = \theta \vec{D}_1(x, y)$ is said to be path-implementable by a density of $n(x, y; \theta)$ nodes if there exists:

(a) *transmission radius function $r_t(x, y; \theta)$ that goes to zero as $\theta \to \infty$, and*

(b) *increasing family of flow streamlines $\{f_i^\theta, i = 1, \cdots, n_s^\theta(A)\}$ of $\vec{D}$, and increasing family of streamline regions $\{h_i^\theta, i = 1, \cdots, n_s^\theta(A)\}$ associated with $\{f_i^\theta\}$,*

such that w.h.p as $\theta \to \infty$

(i) *$(n(x, y; \theta), r_t(x, y; \theta))$ topologically implement the flow streamlines $\{f_i^\theta\}$; i.e., without considering the possible interference due to simultaneous transmissions of nodes, connected paths can be formed to follow the flow streamlines $\{f_i^\theta\}$ within the margins of their associated regions $\{h_i^\theta\}$ by clusters of overlapping discs of radius $r_t(x, y; \theta)$ placed at nodes of density $n(x, y; \theta)$.*

(ii) *Multiple access schemes exist such that the topological implementation of the flow streamlines are sufficient to carry all of the traffic of $|\vec{D}|$.*

Remark 1: As mentioned earlier, if the streamline seeding process which constructs the set of flow streamlines $\{f_i^\theta\}$ is such that it provides local regularity, then the required density of nodes is minimized. A seeding process provides local regularity if as $\theta \to \infty$, streamlines passing through any area element (with approximately uniform $|\vec{D}|$) eventually get evenly spaced for θ large enough.

5.1. Regular streamline seeding process

The following seeding process can be performed in order to choose streamline seeds along the boundary l of the distributed source such that the corresponding set of streamlines are locally regular as $\theta \to \infty$. For each value of the scaling factor θ, the flow seeds can be chosen by starting from an arbitrary point (x_1, y_1) on l and choosing the rest of the seeds on the boundary such that the line integral

of $|\vec{D}|$ between consecutive seeds is equal to the traffic to be relayed within each streamline region; i.e.[6],

$$\int_{\substack{\text{consecutive} \\ \text{flow seeds}}} |\vec{D}| \mathrm{d}l = \gamma C. \tag{13}$$

Due to smoothness of $|\vec{D}(x,y)|$, as θ increases, flow seeds get evenly spaced along each small segment of l. This is because when $|\vec{D}|$ changes smoothly along the boundary of the distributed source, it can be assumed to be approximately constant along small segments of l, hence (using (13)) as $\theta \to \infty$, $\Delta l = \frac{\gamma C}{|\vec{D}|}$, where Δl is the distance between flow seeds (for each small segment) on the boundary of the source with uniform $|\vec{D}|$. Hence, streamlines integrated from these flow seeds will be locally regular.

Remark 2: Let $\{f_i^\theta\}_r$ denote a set of streamlines selected using the regular seeding process, and $\{h_i^\theta\}_r$ be its associated set of streamline regions constructed by bisecting the area between consecutive streamlines, as described earlier. Then as $\theta \to \infty$, the width of streamline regions decreases proportional to θ^{-1} across any cross section of the network perpendicular to the flow field.

Remark 3: Consider a set of locally regular streamlines $\{f_i^\theta\}_r$ and its associated set of regions $\{h_i^\theta\}_r$, and for each i, let g_i^θ denote the stochastic piecewise linear route inside streamline region h_i^θ. Then as $\theta \to \infty$, the set of piecewise linear routes $\{g_i^\theta\}_r$ converges to the set of flow streamlines $\{f_i^\theta\}_r$ uniformly; i.e., $\lim_{\theta \to \infty} \max_i |f_i^\theta - g_i^\theta|_u \to 0$. As defined in Section 2, $|f_i^\theta - g_i^\theta|_u$ is the maximum distance (greatest possible vertical distance) between the two functions from their initial starting points on the distributed source to the destination [16]. Using Remark 2, the width of streamline regions decreases proportional to θ^{-1} at any cross section of the network; hence, $\max_i |h_i^\theta|_u$ shrinks to zero with the same proportion.

It is important to mention that although information flow is defined everywhere inside the network (including inside the distributed source), the implementability analysis (in the sense defined in Definition 1) must be confined to the region between the distributed source and destination where the flow field has zero divergence. This is because in incompressible fields, assuming every flow streamline is drawn continuously with no lines terminating at any point between the source and destination, a one-to-one relationship can be established between $|\vec{D}(x,y)|$ and the number of flow streamlines crossing unit perpendicular line at that point [26]. This property will be found useful in the implementability analysis.

After describing the seeding process which constructs an increasing set of streamlines $\{f_i^\theta\}_r$, "suitable" transmission radius function $r_t(x,y;\theta)$ must be determined such that the rest of the conditions in Definition 1 are satisfied.

---

[6]More accurately, $\int_{\substack{\text{consecutive} \\ \text{flow seeds}}} |\vec{D}| \mathrm{d}l \leq \gamma C$, where the inequality is to take care of situations where $\frac{w_0}{\gamma C}$ is not an integer; hence, $n_s^\theta(A) = \lceil \frac{w_0}{\gamma C} \rceil = \frac{w_0}{\gamma C - \epsilon_\theta}$, where $\epsilon_\theta > 0$ is the smallest real value that makes the total number of streamlines be an integer value. Note that $\epsilon_\theta = o(1)$.

In order to specify $r_t(x, y; \theta)$ and determine its relation with $|\vec{D}(x, y)|$, the following steps are performed which will be described in detail subsequently:

- Perform a boundary-fitted flow-based gridding which results in a curvilinear coordinate system and divides the network region into curvilinear grids with approximately uniform information flow inside each grid.

- Study the implementability problem inside each curvilinear grid.

- Extend the results to the network geometry.

Each of these steps will be described in detail in the subsequent sections.

6. Boundary-Fitted Flow-Based Gridding

Node density is a quantity which must be specified locally. Hence, the network region must be partitioned into smaller area grids over which the magnitude of information flow is approximately uniform.

The gridding step must be such that the implementability analysis becomes easier along the grid boundaries. Hence, grids should adapt to the features of information flow, i.e., the curvature of flow streamlines. This can be viewed as a curvilinear coordinate system for the network. Furthermore, since the information flow vector field satisfies the Neumann boundary condition (9), flow-based gridding of the network results in generation of boundary-fitted coordinates.

In the rest of the paper, we rely on the results from complex analysis which prove existence of a flow-based coordinate system for ideal flow fields (fields that are incompressible and irrotational). A summary of these results are presented in Appendix A.

As mentioned earlier, the implementability analysis is for the region between the source and destination where the flow field has zero divergence. Therefore, in order to use the results from complex analysis and partition the region between the source and destination into curvilinear elements with desired properties, the flow field must be irrotational. Hence, we confine the implementability analysis to flow fields which are irrotational.

Based on Theorem A.1 and Theorem A.2 in Appendix A, for an irrotational flow field $\vec{D}_1$, the equipotential lines $\{\phi = a\}, a \in \mathbb{R}$ and flow streamlines $\{\psi = d\}, d \in \mathbb{R}$ of $\vec{D}_1$ form mutually orthogonal systems of level curves. Furthermore, the set of potential and stream functions of $\vec{D}_1 = (D_{1,x}(x, y), D_{1,y}(x, y))$ can be represented by the following nonlinear transformation:

$$
\begin{aligned}
\mathbf{g} : A &\rightarrow \Phi \\
(x, y) &\mapsto (\phi(x, y), \psi(x, y)),
\end{aligned}
\tag{14}
$$

with Jacobian matrix

$$
\begin{aligned}
\mathcal{J} &= \begin{pmatrix} D_{1,x} & D_{1,y} \\ -D_{1,y} & D_{1,x} \end{pmatrix} \\
&= |\vec{D}_1| \begin{pmatrix} \frac{D_{1,x}}{|\vec{D}_1|} & \frac{D_{1,y}}{|\vec{D}_1|} \\ \frac{-D_{1,y}}{|\vec{D}_1|} & \frac{D_{1,x}}{|\vec{D}_1|} \end{pmatrix}
\end{aligned}
\tag{15}
$$

17

The Jacobian matrix $\mathcal{J}$ represents local linear approximation to the change of coordinates between the two domains, and its determinant (referred to as the Jacobian J) is an area-change factor from each element in the A-domain to the corresponding element in Φ-domain.

Using the inverse function theorem, the Jacobian matrix of the transformation from the Φ-domain to A is equal to

$$\mathcal{J}^{-1} = |\vec{D}_1|^{-1} \begin{pmatrix} \frac{D_{1,x}}{|\vec{D}_1|} & -\frac{D_{1,y}}{|\vec{D}_1|} \\ \frac{D_{1,y}}{|\vec{D}_1|} & \frac{D_{1,x}}{|\vec{D}_1|} \end{pmatrix} \tag{16}$$

The inverse transformation $(\mathbf{g}^{-1})$ describes how the network region A (which is a subspace of the Cartesian plane) is partitioned into curvilinear grids corresponding to equal partitioning of the ϕ and ψ axes in the Φ-domain. Note that the matrix in the right hand side of (16) is a rotation matrix and the transformation is conformal; i.e., the Jacobian matrix is a product of a rotation and scaling at all points. Therefore, the grids of this curvilinear coordinate system can be approximated by square grids at each point in the region between the distributed source and destination.

The flow-based gridding process in shown in Fig. 5, where solid curves represent flow streamlines of a given ideal flow within a subregion of the network and the curvilinear grid specified by dashed lines is the result of flow-based gridding. As shown in the figure, curvilinear lines derived from constant ψ in the Φ-domain are tangent to flow streamlines, and curvilinear lines derived from constant ϕ are perpendicular to the direction of flow of information.
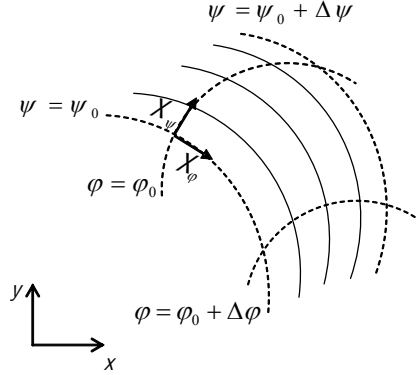

Figure 5: Flow-based gridding and curvilinear coordinates.

At the end, it is worth to mention that various numerical techniques for structured flow-based gridding have been proposed both within Cartesian and curvilinear grid frameworks [27–29]. These techniques are based on the use of streamlines computed from a single-phase flow problem. Also, a Jacobian-based elliptic grid generation technique is presented in [30]. This method, referred

to as vector-field adaptive method or VFA, uses the idea of alignment with a given vector field indirectly via control of the inverse Jacobian matrix of the transformation.

7. Implementability Analysis Inside a Square Grid

In the previous section, we concluded that the network region between the source and destination can be partitioned into smaller approximately square grids which are tangent to the equipotential lines and flow streamlines of $\vec{D}_1$. Note that the size of these square grids varies throughout the network proportional to $|\vec{D}_1|^{-1}$. This is because, based on (16), the Jacobian of the mapping $\mathbf{g}^{-1}$ (from the Φ-domain to A) is equal to $|\vec{D}_1|^{-1}$ at every point in the network.

In this section, we study path-implementability of the information flow inside a square grid as the scaling factor θ goes to infinity.

Recall that as the scaling factor θ increases, $|\vec{D}| = \theta|\vec{D}_1|$ scales linearly throughout the network and more flow streamlines are required to forward the traffic to the destination. Note that the level curves of the curvilinear coordinate system are tangent to the equipotential lines and flow streamlines of $\vec{D}_1$; hence, flow streamlines passing through each grid are parallel to the sides of the grid.

In order to study the required number of sensor nodes inside each (approximately) square grid, consider a unit square U_2 with uniform information flow $\vec{D}^{(l)} = \vec{D}(\phi, \psi)$ and flow streamlines parallel to the sides of the square, as shown in Fig. 6. N sensors are distributed uniformly at random throughout the square and relay traffic to their downstream with equal transmission capacity of C bps, over a common wireless channel; hence, the total number of streamlines required to carry all the traffic within the square is $n_s(U_2) = O(\frac{|\vec{D}^{(l)}|}{C}) = O(|\vec{D}^{(l)}|)$.
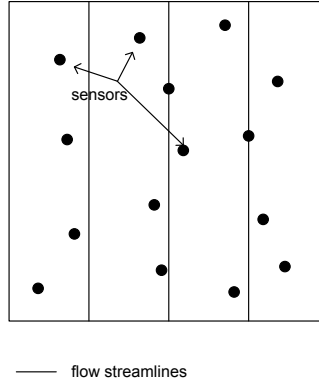

Figure 6: Unit square with uniform information flow and parallel streamlines.

Due to local regularity of the seeding process, streamlines passing through the unit square are evenly spaced. Hence, performing the bisection method,

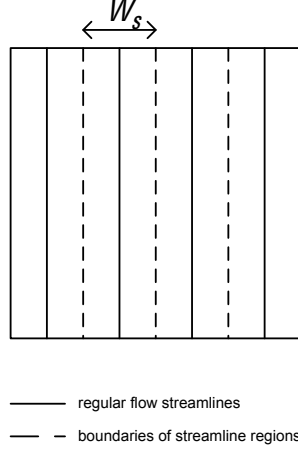streamline regions are a set of $n_s(U_2)$ strips of equal width W_s, as shown in Fig. 7.

Figure 7: Unit square divided into $n_s(U_2)$ strips of width W_s.

In what follows, we first consider a Boolean model for communication which does not account for interference. Using this model, we show that a quadratic relation between magnitude of information flow and number of nodes is not sufficient to follow flow streamlines inside the unit square; i.e., flow streamlines are not topologically implementable. Hence, we conclude that a quadratic relation cannot guarantee path-implementability of information flow under any realistic communication model which accounts for interference from simultaneous transmissions[7].

Next, a sufficient bound for number of nodes N is suggested and its sufficiency is shown under both communication models.

7.1. Topological implementability under a quadratic relation

Consider a bidirectional Boolean model for communication where nodes are connected to each other only if their distance is less than a specified transmission radius r_t. This link model represents wireless communication with signals diffusing isotropically with a certain signal attenuation. Let $p_t = p(r = 0)$ denote the signal power at the transmitter node and p_r denote the received power at a distance r from the transmitter. The received power falls off as $p_r \propto r^{-\alpha} p_t$, where α is the path loss exponent. The wireless transmission radius r_t can

---

[7]This is proved for the case where the set of streamlines are locally regular (evenly spaced). However, as explained in Remark 1, the locally regular case requires minimum node density for implementability. Thus, a quadratic relation cannot guarantee path-implementability regardless of whether or not the set of streamlines are locally regular.

then be mapped to the equivalent transmission power p_t using a threshold Γ for receiver sensitivity; i.e., assuming that a node can receive properly if $p_r \geq \Gamma$. Obviously, this model does not account for the interference caused by possible simultaneous transmissions.

Let $G(N, r_t(N))$ be the communication graph formed by N nodes distributed uniformly at random in the unit square of Fig. 6, where $r_t(N)$ denotes the transmission radius of each node. Also, let $X^{(l)}$ represent the location of nodes inside the unit square. Using the Boolean model, nodes $X_i^{(l)}$ and $X_j^{(l)}$ are connected if $|X_i^{(l)} - X_j^{(l)}| \leq r_t(N)$. Assuming all nodes use the same transmission radius for communication within the unit square, a well-studied problem is identifying the minimum radius (also called critical transmission radius r_c) such that the resulting communication graph is connected.

A communication graph G is connected if and only if there exists at least one path connecting any pair of nodes in the network, and otherwise it is disconnected. As shown in [31], disconnectedness manifests itself by presence of isolated nodes. Isolated nodes cannot find any other nodes within their transmission radius. Lemma 2 investigates the minimum value of r_t such that $G(N, r_t(N))$ is connected w.h.p as $N \to \infty$.

Based on Definition 1, the flow field is topologically implementable if, without considering possible interference due to simultaneous transmissions of nodes, connected paths can be formed to follow streamlines within the margins of their associated regions. Therefore, topological implementability of flow field inside the unit square is equivalent to existence of piecewise linear routes connecting the top to the bottom of the square within each region.

As we prove in Lemma 3, connectivity of the communication graph $G(N, r_t(N))$ is a necessary (yet not sufficient) condition for topological implementability of the flow field $\vec{D}^{(l)}$.

Lemma 2. *In order to ensure connectivity of $G(N, r_t(N))$ w.h.p, r_t must be chosen to be equal to or greater than the critical transmission radius*

$$r_c = \sqrt{\frac{\log N + \eta(N)}{\pi N}}, \tag{17}$$

where $\eta(N)$ is an arbitrary function such that $\eta(N) \to \infty$ as $N \to \infty$.

PROOF. See [32], Corollary 4.1.2.

Substituting N by $O(|\vec{D}^{(l)}|^2)$ in equation (17) results in $r_c(|\vec{D}^{(l)}|) = \Omega(|\vec{D}^{(l)}|^{-1}\sqrt{\log|\vec{D}^{(l)}|})$. On the other hand, since implemented paths must not intersect with each other, the transmission radius r_t cannot be chosen to be greater than $O(W_s)$[8]. Furthermore, when adding interference

---

[8]Having a non-intersecting set of implemented paths can be guaranteed by assigning different frequency sub-channels to nodes residing in neighboring streamline regions. The required number of frequency sub-channels grow to infinity if $r_t > O(W_s)$ which significantly reduces the throughput.

to the model, if the order of transmission radius r_t is greater than W_s, transmission along a path will introduce interference to infinitely many other paths as $|\vec{D}^{(l)}| \to \infty$ and will significantly reduce the throughput of the network. Therefore, r_t must be chosen to be $O(W_s) = O(|\vec{D}^{(l)}|^{-1})$. Hence, $r_t < r_c$ as $|\vec{D}^{(l)}| \to \infty$. Considering the relation between transmission power and transmission radius, and using the same argument as above, transmission power of sensor nodes must also be $O(|\vec{D}^{(l)}|^{-\alpha})$ to reduce the amount of interference on simultaneous transmissions in neighboring streamline regions.

Choosing $r_t < r_c$ results in a disconnected network; although we are not concerned with having a connected network in this work, but as stated earlier, disconnectedness manifests itself by presence of isolated nodes. The following lemma shows how the presence of isolated nodes affects topological implementability of flow streamlines.

Lemma 3. *The set of flow streamlines inside the unit square cannot be topologically implementable if the pair (N, r_t) is chosen such that $G(N, r_t)$ has isolated nodes w.h.p.*

PROOF. Throughout this argument, we assume that N nodes are distributed uniformly at random within the unit square. By the definition of topological implementability, if a set of flow streamlines is not implementable with (N, r_{t_1}), it cannot be implementable with any pair (N, r_{t_2}) such that $r_{t_2} < r_{t_1}$. Hence, without loss of generality, we will concentrate on cases where $r_t \geq W_s$.

Suppose $X_k^{(l)}$ is an isolated node in the network, and let d_{kf} denote its distance to the closest streamline. As mentioned before, the unit square is divided into $n_s(U_2)$ strips of width W_s; therefore, $0 \leq d_{kf} \leq \frac{W_s}{2}$.

Since $X_k^{(l)}$ is an isolated node, no other nodes exist in a region of distance r_t from $X_k^{(l)}$. This affects implementability of streamlines close to $X_k^{(l)}$, specifically the one within the same strip as $X_k^{(l)}$. For this streamline to get topologically implemented, there must exist at least one node inside the shaded area S_d shown in Fig. 8. Let $r_t = W_s$, and consider the largest S_d which corresponds to $d_{kf} = \frac{W_s}{2}$. Choosing a smaller d_{kf} can only decrease the probability of implementability.
$S_d|_{d_{kf} = \frac{W_s}{2}} = (\frac{\sqrt{3}}{4} - \frac{\pi}{12})W_s^2 = O(|\vec{D}^{(l)}|^{-2})$, and

$$
\begin{aligned}
P\{\text{the streamline not implementable}\} &\geq \\
P\{\text{there exists no node inside } S_d\} &= (1 - S_d)^{N-1} \\
&\geq \exp(-\frac{(N-1)S_d}{1 - S_d})
\end{aligned}
$$

The last inequality is derived using the relation $\exp(-\frac{x}{1-x}) \leq 1 - x$, for $|x| \leq 1$. Substituting N by $O(|\vec{D}^{(l)}|^2)$, and S_d by $O(|\vec{D}^{(l)}|^{-2})$, and taking the limit as

$|\vec{D}^{(l)}| \to \infty$,

$$\lim_{|\vec{D}^{(l)}| \to \infty} P\{\text{the streamline not implementable}\} \geq$$

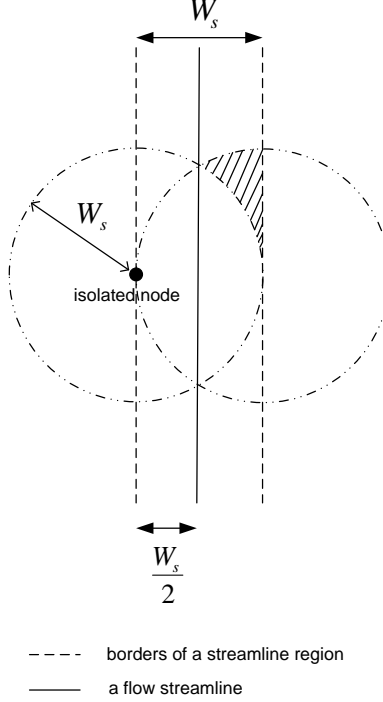$$\lim_{|\vec{D}^{(l)}| \to \infty} \exp(-NS_d) > 0$$

Figure 8: An isolated node with $r_t = W_s$, preventing topological implementability of a flow streamline.

The above inequality states that the probability that the closest streamline to $X_k^{(l)}$ is not topologically implementable is strictly positive. The same analysis can be done for $r_t > W_s$ resulting in a positive probability of the flow streamline not being implementable. Fig. 9 shows the scenario where $S_d|_{d_{kf} = \frac{W_s}{2}} = 0$ for $r_t = \frac{2}{\sqrt{3}} W_s$.

Note that this lemma is based on the assumption that the network has isolated nodes. Therefore, the transmission radius r_t cannot grow unconditionally and that it must be less than the critical transmission radius $r_c(N)$ for connectivity.

Theorem 1. *Information flow field $\vec{D}^{(l)}$ inside the unit square is not path-implementable with $N = O(|\vec{D}^{(l)}|^2)$ nodes distributed uniformly at random throughout the square.*
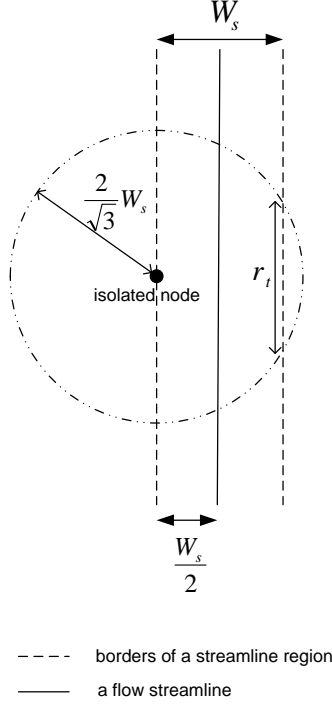
Figure 9: The area $S_d|_{d_{kf}=\frac{W_s}{2}} = 0$ for $r_t \geq \frac{2}{\sqrt{3}}W_s$.

PROOF. Lemma 3, together with the fact that $r_t < r_c(|\vec{D}^{(l)}|)$ for $N = O(|\vec{D}^{(l)}|^2)$ proves that a quadratic relation between the total number of nodes and the traffic to be relayed does not guarantee existence of piecewise linear routes inside each streamline region. Hence, $\vec{D}^{(l)}$ is not path-implementable with $N = O(|\vec{D}^{(l)}|^2)$ nodes distributed uniformly at random throughout the square.

7.2. A sufficient order of nodes for path-implementability

In this part we show that by increasing the number of nodes to $N = O(|\vec{D}^{(l)}|^2 \log |\vec{D}^{(l)}|)$, the flow field $\vec{D}^{(l)}$ becomes path-implementable inside the unit square. The sufficiency of this bound is shown by applying feasible scheduling schemes under both the Protocol and the Physical Model. The proposed scheme follows the lines of thought in [14] and [24], and is briefly described here. A more detailed description of the method is presented in [33].

The main idea of the scheduling scheme is to divide nodes into different groups using the tessellation technique, and divide time and frequency into smaller slots such that transmissions at each time slot and frequency sub-channel abide by either of the communication models.

Tessellate [9] the unit square into smaller square cells of length s similar to what is shown in Fig. 10. This divides the square into $\frac{1}{s^2}$ cells of equal size. Let $C_{a,b}$ denote a cell with coordinates $a, b = 1, \cdots, \frac{1}{s}$ starting from the bottom left, and let M be a fixed positive integer (to be determined later). Divide cells into M^2 groups $\{GC_i, i = 1, \cdots, M^2\}$ such that $C_{a,b}$ and $C_{a',b'}$ belong to the same group if and only if $a \equiv a' \pmod{M}$, and $b \equiv b' \pmod{M}$. Furthermore assume time is slotted, and each time frame has M^2 time slots. At each time slot $m = 1, \cdots, M^2$ only nodes within the mth group GC_m can communicate with their neighboring cells.
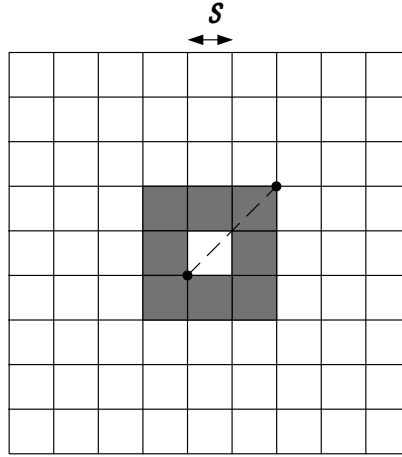

Figure 10: Tessellation of the unit square into cells of length s.

In order to guarantee existence of piecewise linear routes, N must be chosen such that w.h.p there exists at least one node inside each cell. Furthermore, the transmission power and correspondingly the transmission radius of nodes must be such that inter-cell communications are possible with finite M as $\theta \to \infty$.

To ensure that the implemented paths do not intersect with each other, the length of each cell s is chosen to be $O(W_s)$ and the frequency channel is divided into $\kappa \geq 2$ sub-channels of equal capacity $\{C_i = \frac{C}{\kappa}, i = 1, \cdots, \kappa\}$, such that nodes in adjacent implementation regions use different sub-channels to communicate.

Theorem 2. *Information flow $\vec{D}^{(l)}$ inside the unit square is path-implementable with $N = O(|\vec{D}^{(l)}|^2 \log |\vec{D}^{(l)}|)$ nodes distributed uniformly at random throughout the square using either the Protocol or the Physical Model.*

PROOF. See [33], Appendix B. The proof shows that a finite M (independent of θ) can be found under either communication model, and that w.h.p there exists

---

[9]Partition of an area into small units or subareas.

at least one node inside each cell of length $s = O(W_s) = O(|\vec{D}^{(l)}|^{-1})$ to relay traffic to its neighboring cell within its associated time slot.

8. Extending the Results to the Network Geometry

As stated earlier, the unit square example models the flow of traffic inside a small part of the network geometry. However, when studying the example, we only considered the interference effects of nodes communicating inside the square. While this is acceptable for the Protocol Model which describes the effects of interference locally and on a pairwise relation between nodes, it may not be acceptable for the Physical Model which takes into account the aggregate interference from all nodes in the network communicating at the same time slot.

The following lemma states that a bounded version of the Physical Model can be used for any arbitrary node X_i within the network and that the total interference due to simultaneous transmissions of nodes outside interference region $IR(X_i) = O(1)$ can be ignored in the implementability analysis. Hence, the results derived from the unit square can be extended to the rest of the network.

Lemma 4. *A bounded version of the Physical Model with interference region $IR = O(1)$ can be used in the implementability analysis when nodes are communicating based on a controlled access scheme with transmission power $p_t(x, y) \sim |\vec{D}(x, y)|^{-\alpha}$.*

PROOF. See Appendix B.

Hence, a sufficient bound on total number of nodes required to relay the traffic along a given irrotational flow field $\vec{D}(x, y)$ (scaling linearly to infinity) inside any subregion A_s of the network geometry between the distributed source and destination is found by calculating the following integral

$$\int_{A_s} |\vec{D}(x, y)|^2 \log |\vec{D}(x, y)| \mathrm{d}x\mathrm{d}y. \tag{18}$$

The following example further illustrates the different steps involved in the implementability analysis.

Example: Consider a circle centered at the origin with radius R, as shown in Fig. 11, and assume that uniform information flow enters the circular area toward the sink which is located at the center. In order to avoid numerical ambiguities, information sink is assumed to be a circle with non-zero radius r_d. However, for simplicity, it is plotted as a point sink in the figure.

Assuming that the distributed source is located outside the circular area,

$$\rho(x, y) = 0, \ \ r_d^2 < x^2 + y^2 < R^2, \tag{19}$$

and the flow field is incompressible and irrotational everywhere within the circle except at the sink.

Let θ denote the total amount of traffic flux entering the circle. Hence, $n_s^\theta(A) \sim \frac{\theta}{C}$ streamlines are required to carry all the traffic to the sink. Following

the streamline seeding process described in Section 5.1, any circle around the sink with radius $r_s > r_d$ can be used as a seeding path. Since the traffic flux passing through the seeding path is uniform, flow seeds are placed regularly on the seeding path and streamlines are integrated all the way to the boundary of the network.

Using basic properties of information flow, the flow field $\vec{D}$ can be expressed in polar coordinates as follows:

$$\vec{D}(r, \phi) = -\frac{\theta}{2\pi r} \vec{a}_r, \quad r_d < r < R$$
$$\phi \in [0, 2\pi),$$

where $\vec{a}_r = \vec{Dir}(r, \phi)$ is the unit vector along the r-coordinate.

It can be easily verified that the level curves of the polar coordinates (r, ϕ) are tangent to the equipotential lines and flow streamlines of $\vec{D}_1(r, \phi) = -\frac{1}{2\pi r} \vec{a}_r$. Hence, polar coordinates is used to partition the circular area into grids of size $r dr d\phi$ with approximately uniform information flow inside each grid.
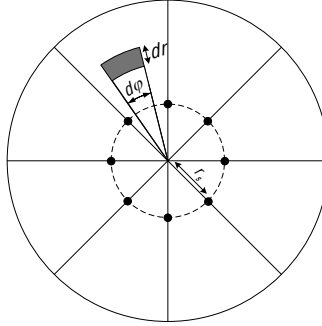

Figure 11: Streamline seeds are regularly placed on the circle with radius r_s from the sink.

Using the results of Section 7.2, the required density of nodes to follow flow streamlines inside each grid is $O(|\vec{D}(r, \phi)|^2 \log |\vec{D}(r, \phi)|)$. Therefore, as $\theta \to \infty$, the total number of nodes sufficient to carry all the traffic to the destination is proportional to

$$\int_0^{2\pi} \int_{r_d}^{R} \frac{\theta^2}{4\pi^2 r^2} \log \frac{\theta}{2\pi r} r dr d\phi = \frac{1}{2\pi} \log(\frac{R}{r_d}) \cdot \theta^2 \log(\frac{\theta}{2\pi\sqrt{r_d R}}).$$

9. Conclusions

In a network geometry with a given information flow vector field $\vec{D}$, the flow streamlines represent desired paths for flow of traffic. For such networks, we defined a new notion of connectivity (called path-implementability) which concerns the ability of sensor nodes to relay traffic along flow streamlines of $\vec{D}$.

27

We discussed that the notion of path-implementability not only provides a framework to study the density of sensor nodes required to follow a set of desired paths, but also guarantees existence of enough paths connecting the source to the destination so that all the traffic can be transmitted to the destination. Path-implementability requires more nodes than simple connectivity.

In a network with a given information flow $\vec{D}$, we showed that a density of $n(x,y) \sim |\vec{D}(x,y)|^2$ sensor nodes is not sufficient to implement the flow field, and that if the density increases to $n(x,y) \sim |\vec{D}(x,y)|^2 \log |\vec{D}(x,y)|$, the flow field becomes path-implementable. The sufficiency of the latter is shown by proposing a controlled access scheduling scheme.

The excess amount of nodes required for path-implementability can be exploited in cross-layer design and integration of protocol layers to make lightweight protocol stacks in order to meet specific requirements of dense sensor networks. This will be studied more in detail in our future work.

Appendix A. Existence of a Flow-Based Coordinate System

The following discussion, taken from [34], shows the existence of a flow-based coordinate system for ideal flow fields which, by definition, have zero divergence and zero curl. The following theorems (stated without proof from [34]) are used to introduce a curvilinear coordinate system using complex analytic functions, and are reported here for the sake of completeness.

Consider steady state flow field $\vec{v}(x,y) = (u(x,y), w(x,y))$ at the point $(x,y) \in \Omega$. Here $\Omega \subset \mathbb{R}^2$ is the domain in which the flow field is defined.

Theorem A.1. *The flow vector field $\vec{v} = (u(x,y), w(x,y))$ induces an ideal flow if and only if $f(z) = u(x,y) - iw(x,y)$ is a complex analytic function of $z = x + iy$.*

Hence, the components $u(x,y)$ and $-w(x,y)$ of an ideal flow field are harmonic conjugates. The corresponding complex function $f(z)$ is known as the *complex velocity* of the flow field.

Now, suppose that $f(z)$ admits a complex anti-derivative, i.e., a complex analytic function $\chi(z) = \phi(x,y) + i\psi(x,y)$ that satisfies $\frac{d\chi}{dz} = f(z)$. Then,

$$\frac{d\chi}{dz} = \frac{\partial \phi}{\partial x} - i\frac{\partial \phi}{\partial y} = u - iw \tag{A.1}$$

Therefore, $\nabla \phi = \vec{v}$. For this reason, the anti-derivative $\chi(z)$ is known as the complex potential function for the given flow field. Furthermore, $\phi(x,y)$ and $\psi(x,y)$ are referred to as the potential function and the stream function of the flow field, respectively. Following the Cauchy-Riemann equations:

$$\begin{aligned} \frac{\partial \phi}{\partial x} &= \frac{\partial \psi}{\partial y} = u \\ \frac{\partial \phi}{\partial y} &= -\frac{\partial \psi}{\partial x} = w \end{aligned} \tag{A.2}$$

28

The level curves of the potential function, $\{\phi(x,y)=a\}$, $a \in \mathbb{R}$, are known as equipotential lines. The flow vector $\vec{v}$ points in the normal direction to the equipotential lines. On the other hand, $\vec{v}$ is tangent to the level curves $\{\psi(x,y)=d\}$, $d \in \mathbb{R}$. But $\vec{v}$ is the flow field, and so tangent to the streamlines followed by the flow. Thus, these two systems of curves must coincide, and the level curves of the stream function are the streamlines of the flow.

Summarizing, for an ideal flow field, the equipotential lines $\{\phi = a\}$ and streamlines $\{\psi = d\}$ form mutually orthogonal systems of level curves. The flow field $\vec{v} = \nabla\phi$ is tangent to the streamlines and normal to the equipotential lines, whereas the gradient of the stream function, $\nabla\psi$, is tangent to the equipotential lines and normal to the streamlines.

The following theorem states the condition under which a unique anti-derivative function exists.

Theorem A.2. *Let* $f(z) = \frac{\mathrm{d}\chi(z)}{\mathrm{d}z}$, *where* $\chi(z)$ *is a single-valued complex function for* $z \in \Omega$. *If* $c \subset \Omega$ *is any closed curve, then* $\oint_c f(z)\mathrm{d}z = 0$. *Conversely, if this condition holds for all closed curves* $c \subset \Omega$ *contained in the domain of definition of* $f(z)$, *then* f *admits a single-valued complex anti-derivative* $\chi(z)$ *with* $\frac{\mathrm{d}\chi(z)}{\mathrm{d}z} = f(z)$.

Appendix B. Proofs

PROOF OF LEMMA 1. Since the *p-norm* problem (11) is a convex optimization problem with differentiable objective and constraint functions, the KKT[10] conditions are necessary and sufficient for the optimal solution to be met. let $L(\vec{D}, \nu_\theta)$ denote the Lagrangian function of the *p-norm* problem with load density function $\theta\rho_1(x,y)$, and ν_θ denote the dual variable associated with the problem. Now, if $(\vec{D}_1, \nu_1^*)$ be the optimal primal and dual solutions of the problem with $\theta = 1$, then based on the KKT conditions,

$$
\nabla_{\vec{D}} L(\vec{D}_1, \nu_1^*) = \nabla_{\vec{D}} \Big[\int_A |\vec{D}_1(x,y)|^p \mathrm{d}x\mathrm{d}y -
$$
$$
\int_A \nu_1^*(x,y)\big(\nabla \cdot \vec{D}_1(x,y) - \rho_1(x,y)\big)\mathrm{d}x\mathrm{d}y \Big] = 0,
$$
$$
\text{and}
$$
$$
\nabla \cdot \vec{D}_1(x,y) - \rho_1(x,y) = 0
$$

---

[10] *Karush-Kuhn-Tucker* conditions.

Note that $\nabla_{\vec{D}}$ is the gradient operator with partial derivatives with respect to D_x and D_y. Now, substituting ρ_1 by $\theta\rho_1$ and $\vec{D}_1$ by $\theta\vec{D}_1$,

$$\nabla_{\vec{D}}L(\theta\vec{D}_1, \nu_\theta^*)|_{\nu_\theta^* = \theta^{(p-1)}\nu_1^*} = \theta^p \nabla_{\vec{D}}L(\vec{D}_1, \nu_1^*)$$
$$= 0,$$

and

$$\nabla \cdot \big(\theta\vec{D}_1(x,y)\big) - \theta\rho_1(x,y) = \theta\big(\nabla \cdot \vec{D}_1(x,y) - \rho_1(x,y)\big)$$
$$= 0$$

This can be easily verified using the Lagrangian equation and the fact that, for scalar θ, $\nabla \cdot \big(\theta\vec{D}(x,y)\big) = \theta\nabla \cdot \vec{D}(x,y)$. Therefore, $\theta\vec{D}_1(x,y)$ is the optimal solution for the *p-norm* problem (11) with load density function $\theta\rho_1(x,y)$.

PROOF OF LEMMA 4. Note that the flow-based gridding of the network region A is performed using the initial flow field $\vec{D}_1(x,y)$ and independent of the scaling factor θ. Hence, as θ grows, the size of square grids remains unchanged. However, transmission power of nodes $p_t(x,y) \sim |\vec{D}(x,y)|^{-\alpha}$. Also, the width of streamline regions shrinks proportional to $|\vec{D}|^{-1}$ at any cross section of the network. Consider interference region of nodes to be circular regions with radius of $O(1)$ around each node. Let $d(x,y)$ denote the density of active cells at each time slot. With the proposed TDMA scheme, $d(x,y) = \frac{1}{s^2 M^2}$, where $s(x,y) = O(|\vec{D}(x,y)|^{-1})$ refers to the length of the tessellating cells inside each square grid, and M refers to the parameter of the controlled access scheme. Now, the total amount of interference at an arbitrary node X_u due to simultaneous transmission of nodes outside $IR(X_u)$ can be upper bounded as follows:

$$\begin{aligned}
I_{A \setminus IR(X_u)} &\leq \max_s\{\frac{1}{s^2 M^2}\} \cdot \max_{(x,y) \in A \setminus IR(X_u)}\{p_t\} \cdot O(1) \\
&= \frac{|\vec{D}_{\max}|^2}{|\vec{D}_{\min}|^\alpha} \cdot O(1) \\
&= \Big(\frac{|\vec{D}_{\max}|}{|\vec{D}_{\min}|}\Big)^2 \frac{1}{|\vec{D}_{\min}|^{\alpha-2}} \cdot O(1) \quad\quad\quad\text{(B.1)}
\end{aligned}$$

Note that as $\theta \to \infty$, $\frac{|\vec{D}_{\max}|}{|\vec{D}_{\min}|}$ remains constant due to linear scaling of $|\vec{D}|$, and bounded due to non-zero dimensions of the information sink. Also, based on Lemma 1, as $\theta \to \infty$, $|\vec{D}| \to \infty$ everywhere in the network with non-zero information flow. Hence, the second term in (B.1) goes to 0 as $\theta \to \infty$.

With a diminishing interference effect at an arbitrary node X_u due to transmissions outside $IR(X_u)$, a bounded version of the Physical Model (as stated in (6)) can be used for the implementability analysis.

References

[1] S. Kulkarni, A. Iyer, C. Rosenberg, An address-light, integrated MAC and routing protocol for wireless sensor networks, IEEE/ACM Transactions on Networking 14 (4) (2006) 793–806.

[2] J. N. Al-Karaki, A. E. Kamal, Routing techniques in wireless sensor networks: a survey, IEEE Wireless Communications 11 (6) (2004) 6–28.

[3] M. Kalantari, M. Shayman, Energy efficient routing in wireless sensor networks, in: Proceedings of Conference on Information Sciences and Systems (CISS'04), 2004.

[4] M. Kalantari, M. Haghpanahi, M. Shayman, A p-norm flow optimization problem in dense wireless sensor networks, in: The 27th Conference on Computer Communications (INFOCOM'08), 2008, pp. 341–345.

[5] S. Toumpis, L. Tassiulas, Packetostatics: deployment of massively dense sensor networks as an electrostatics problem, in: Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'05), vol. 4, 2005.

[6] J. Yick, B. Mukherjee, D. Ghosal, Wireless sensor network survey, Computer networks (Elsevier) 52 (12) (2008) 2292–2330.

[7] G. Anastasi, M. Conti, M. Di Francesco, A. Passarella, Energy conservation in wireless sensor networks: a survey, Ad Hoc Networks (Elsevier) 7 (3) (2009) 537–568.

[8] P. Santi, Topology control in wireless ad hoc and sensor networks, ACM Computing Surveys (CSUR) 37 (2) (2005) 164–194.

[9] E. Fasolo, M. Rossi, J. Widmer, M. Zorzi, In-network aggregation techniques for wireless sensor networks: a survey, IEEE Wireless Communications 14 (2) (2007) 70–87.

[10] M. Kalantari, M. Shayman, Routing in wireless ad hoc networks by analogy to electrostatic theory, in: IEEE International Conference on Communications (ICC'04), vol. 7, 2004.

[11] A. Basu, A. Lin, S. Ramanathan, Routing using potentials: a dynamic traffic-aware routing algorithm, in: Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM'03), Germany, 2003, pp. 37–48.

[12] V. Lenders, R. Baumann, Link-diversity routing: a robust routing paradigm for mobile ad hoc networks, in: IEEE Wireless Communications and Networking Conference (WCNC'08), 2008, pp. 2585–2590.

[13] N. T. Nguyen, A. I. A. Wang, P. Reiher, G. Kuenning, Electric-field-based routing: a reliable framework for routing in MANETs, ACM SIGMOBILE Mobile Computing and Communications Review 8 (2) (2004) 35–49.

[14] M. Franceschetti, O. Dousse, D. N. C. Tse, P. Thiran, Closing the gap in the capacity of wireless networks via percolation theory, IEEE Transactions on Information Theory 53 (3) (2007) 1009–1018.

[15] D. K. Cheng, Field and Wave Electromagnetics, Addison-Wesley, 1989.

[16] K. A. Ross, Elementary Analysis: The Theory of Calculus, Springer Verlag, 1980.

[17] P. Gupta, P. R. Kumar, The capacity of wireless networks, IEEE Transactions on Information Theory 46 (2) (2000) 388–404.

[18] P. Cardieri, Modeling interference in wireless ad hoc networks, IEEE Communications Surveys & Tutorials 12 (4) (2010) 551–572.

[19] http://isi.edu/nsnam/ns/, last access date: 09/08/2011.

[20] G. F. Riley, The Georgia Tech network simulator, in: Proceedings of the ACM SIGCOMM Workshop on Models, Methods and Tools for Reproducible Network Research, 2003, pp. 5–12.

[21] D. M. Blough, C. Canali, G. Resta, P. Santi, On the impact of far-away interference on evaluations of wireless multihop networks, in: Proceedings of the 12th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM'09), 2009, pp. 90–95.

[22] M. Kalantari, M. Shayman, Design optimization of multi-sink sensor networks by analogy to electrostatic theory, in: IEEE Wireless Communications and Networking Conference (WCNC'06), 2006.

[23] M. Kalantari, Design optimization and security for communication networks, Ph.D. thesis, University of Maryland, College Park, 2005.

[24] S. Toumpis, A. Goldsmith, Large wireless networks under fading, mobility, and delay constraints, in: IEEE INFOCOM 2004, 2004, pp. 609–619.

[25] A. El Gamal, J. Mammen, B. Prabhakar, D. Shah, Optimal throughput-delay scaling in wireless networks–part I: the fluid model, IEEE/ACM Transactions on Networking 14 (6) (2006) 2568–2592.

[26] N. Kemmer, Vector Analysis: A Physicist's Guide to the Mathematics of Fields in Three Dimensions, Cambridge University Press, 1977.

[27] S. Verma, K. Aziz, A control volume scheme for flexible grids in reservoir simulation, in: SPE Reservoir Simulation Symposium, 1997.

[28] M. Edwards, R. Elf, K. Aziz, Quasi K-orthogonal streamline grids: gridding and discretization, in: SPE Annual Technical Conference and Exhibition, 1998.

[29] A. Castellini, M. G. Edwards, L. J. Durlofsky, Flow based modules for grid generation in two and three dimensions, in: Proceedings of the 7th European Conference on the Mathematics of Oil Recovery (ECMOR VII), Italy, 2000.

[30] P. Knupp, Mesh generation using vector-fields, Journal of Computational Physics 119 (1) (1995) 142–148.

[31] P. Gupta, P. R. Kumar, Critical power for asymptotic connectivity, in: Proceedings of the 37th IEEE Conference on Decision and Control (CDC'98), vol. 1, 1998, pp. 1106–1110.

[32] P. Santi, Topology Control in Wireless Ad Hoc and Sensor Networks, Wiley, 2005.

[33] M. Haghpanahi, Design and implementation of information paths in dense wireless sensor networks, Ph.D. thesis, University of Maryland, College Park, 2011.

[34] P. J. Olver, Applied Mathematics Lecture Notes, http://www.math.umn.edu/~olver/appl.html, last access date: 09/08/2011.