Optimal Deployment of Large Wireless Sensor Networks

Stavros Toumpis, *Member, IEEE*, and Leandros Tassiulas, *Member, IEEE*

***Abstract*—A spatially distributed set of sources is creating data that must be delivered to a spatially distributed set of sinks. A network of wireless nodes is responsible for sensing the data at the sources, transporting them over a wireless channel, and delivering them to the sinks. The problem is to find the optimal placement of nodes, so that a minimum number of them is needed. The critical assumption is made that the network is *massively dense*, i.e., there are so many sources, sinks, and wireless nodes, that it does not make sense to discuss in terms of microscopic parameters, such as their individual placements, but rather in terms of macroscopic parameters, such as their spatial densities. Assuming a particular interference-limited, capacity-achieving physical layer, and specifying that nodes only need to transport the data (and not to sense them at the sources, or deliver them at the sinks once their location is reached), the optimal node placement induces a traffic flow that is identical to the electrostatic field created if the sources and sinks are replaced by a corresponding distribution of positive and negative charges. Assuming a general model for the physical layer, and specifying that nodes must not only transport the data, but also sense them at the sources and deliver them at the sinks, the optimal placement of nodes is given by a scalar nonlinear partial differential equation found by calculus of variations techniques. The proposed formulation and derived equations can help in the design of large wireless sensor networks that are deployed in the most efficient manner, not only avoiding the formation of bottlenecks, but also striking the optimal balance between reducing congestion and having the data packets follow short routes.**

***Index Terms*—Capacity, electrostatics, node placement, physical layer, sensor networks, wireless ad hoc networks.**

I. Introduction

A. Wireless Sensor Networks

Wireless sensor networks are comprised of sensors that are equipped with wireless transceivers and so are able to form a wireless network [3]. The sensors use this network to coordinate their sensing activities, and so enhance their sensing capabilities, and also to relay the data they sense to specified data collection locations, typically referred to as data sinks.

Manuscript received May 22, 2005; revised March 21, 2006. This work was supported by **K***plus* funding for the ftw. project I0 "Signal and Information Processing," by the European Network of Excellence NEWCOM, and by the ARO under Grant W911NF-04-1-0306. This work was conducted while S. Toumpis was with the Telecommunications Research Center Vienna (ftw.). The material in this paper was presented in part at INFOCOM 2005, Miami, FL; and SECON 2005, Santa Clara, CA, March 2005.

S. Toumpis is with the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia 1678, Cyprus (e-mail: toumpis@ucy.ac.cy).

L. Tassiulas is with the Computer Engineering and Telecommunications Department, University of Thessaly, Volos 38221, Greece (e-mail: leandros@inf.uth.gr).

Communicated by E. Modiano, Associate Editor for Communication Networks.

Digital Object Identifier 10.1109/TIT.2006.876256

These networks differ from generic wireless ad hoc networks in that the traffic is not created by the nodes themselves, but rather by the environment in which the nodes exist. Therefore, for some applications, for example the sensing of temperatures in a planted area, the data that are sensed in neighboring sensors are correlated—if the nodes coordinate with their neighbors and compress the data in a distributed manner, the total amount of traffic that must be received by the data collectors will be significantly reduced [4], [5]. Another feature of wireless sensor networks is that, in most applications, it is expected that sensors will be totally immobile. This can significantly simplify the design of the routing protocols [6], [7]. On the other hand, it is expected that the sensors will typically be operating using a nonrenewable battery supply, therefore it is critical that they use their available energy as efficiently as possible [8]–[10].

B. Massively Dense Networks

It is envisioned that, in the future, wireless sensor networks may consist of a large number of nodes, potentially on the order of many thousands [3]. At first, it may seem that this can create insurmountable difficulties in their design and analysis. However, recent work has shown that the large number of nodes in a network could actually be a blessing in disguise, as it can allow researchers to make important simplifying assumptions.

For example, in [11] the authors investigate the asymptotic behavior of the capacity of a class of two-dimensional random networks as the number of nodes n approaches infinity, under a uniform traffic assumption. The authors present a scheme that achieves **with high probability (w.h.p.)**, i.e., with probability approaching 1 as n approaches infinity, a rate of communication equal to $c_1(n\log n)^{-\frac{1}{2}}$, where c_1 is a multiplicative constant, from each node to its randomly chosen destination. The authors also show that, with high probability, the n nodes cannot send data to their destinations with a per-node rate of communication equal to $c_2 n^{-\frac{1}{2}}$, where c_2 is a multiplicative constant. These results are based on a propagation model under which power decays polynomially with distance, and a given signal-to-interference-and-noise ratio (SINR) threshold is needed for the successful reception of a data packet. The results of [11] were recently tightened in [12], by dropping the logarithm in the expression of the lower bound, using percolation theory.

More recently, in [13] the author examines a *massively dense* network, that consists of so many nodes, that it does not make sense to describe the network in terms of microscopic quantities, such as the position of an individual node, or the rate with which it is receiving data, but rather in terms of macroscopic quantities, such as the spatial density of nodes $d(x,y)$ at a particular location (x,y), and the total traffic that goes through this location. In this setting, it is shown that the minimum-hop route

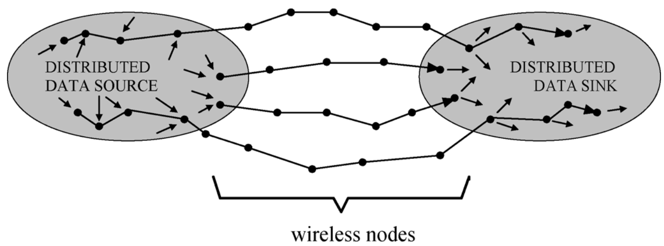


Fig. 1. A set of wireless nodes is deployed in an area to support the sensing, transport, and delivery of data from a distributed traffic source to a distributed traffic sink.

connecting two arbitrary points is identical to the path followed by a ray of light, traveling between the two points, if we assume that the network is substituted by an optically inhomogeneous medium whose index of refraction equals $\sqrt{d(x,y)}$. Therefore, an important problem in wireless networks can be shown to be a fundamental and well-understood problem in Optics.

C. Overview of Work

In this work we investigate a setting that, to the best of our knowledge, has not attracted significant research interest until now. As shown in Fig. 1, we consider an environment in which there are a spatially distributed set of data sources and a spatially distributed set of data sinks. We have available a large number of wireless nodes, to be used for *i)* the sensing of the data, *ii)* the transport of the data from the source locations to the sink locations, and *iii)* the delivery of the data to the sinks once their location is reached.

We are interested in calculating the minimum number of nodes needed to support the traffic, and the associated placement of nodes that achieves this minimum. In other words, we are given a task (the transfer of data from the sources to the sinks) and a set of resources (the wireless nodes), and we would like to determine what is the minimum of resources needed, and how this minimum of resources should be deployed to achieve the task.

Note that a fundamental tradeoff exists: On the one hand, the traffic must take relatively short routes, so that not many nodes are needed for each route. On the other hand, it is important that the traffic is sufficiently spread, to minimize the effects of interference. These two goals are competing, and the optimal placement of nodes should strike a balance in the most favorable way.

As in [13], the assumption is made that the network is massively dense, i.e., there are so many sources, sinks, and nodes available, that it is best to describe the network in terms of macroscopic parameters. In Section II, we define these parameters in detail.

We first study the problem under two specific assumptions.

A) Wireless nodes only have to transport the traffic from the locations of the sources to the locations of the sinks, and do not need to either sense the traffic at the sources, or deliver the traffic to the sinks, once it arrives at their physical locations.

B) We assume a specific relation between the traffic that can be transported through a location (x,y), and the node density $d(x,y)$ at that location. In particular, the amount of traffic that can cross a linear segment of incremental length ϵ, centered at (x,y), is at most $\epsilon|\mathbf{T}(x,y)|_{\max}$, and[1]:

$$|\mathbf{T}(x,y)|_{\max} = c\sqrt{d(x,y)} \quad (1)$$

where c is a constant. As discussed in detail in Section III, this assumption is justified for a physical layer in which *i)* the bandwidth available to the nodes is limited, therefore *ii)* adjacent transmissions interfere with each other, and *iii)* the nodes in each location share the bandwidth in the most efficient manner, so that the network locally operates at its capacity bound. Therefore, the network behaves locally as the networks studied in [11] and [12].

In Section IV, we show that, under these two assumptions, the optimal spatial density of nodes induces a traffic flow that is identical to the electrostatic field that will be induced if the distribution of sources is substituted by an identical distribution of positive charge, and the distribution of sinks is substituted by an identical distribution of negative charge. Many aspects of Electrostatics are shown to have a straightforward and illuminating interpretation in the context of wireless sensor networks, notably boundary conditions along the interfaces of different dielectric materials, Thomson's theorem on the placement of charges on conductors, and the potential function.

In Section V, we slightly modify the network model of Section II, and in Section VI we introduce a general physical layer model, that does not include Assumptions A) and B). In particular, we substitute (1) with

$$|\mathbf{T}(x,y)|_{\max} = F(x,y,d(x,y)) \quad (2)$$

where $F(\cdot)$ is some arbitrary function. Different choices of $F(\cdot)$ correspond to different physical layers models. We also include in our model the fact that nodes must not only transport the data, but also sense it at the sources and deliver it at the sinks, once their location is reached.

In Section VII we use calculus of variations to determine the optimal distribution of nodes, and the traffic flow it induces, under the general physical layer model of Section VI. The optimal distribution is given in terms of a scalar partial differential equation (PDE), which in general is nonlinear. The results of Section IV are contained in the results of this section as a special case.

In Section VIII, we briefly discuss extensions of our work that go beyond the scope of this paper. In particular, we discuss the case of networks with different types of traffic, the application of our formulation in alternative problems in wireless sensor networks (such as problems in delay and energy minimization), and the interpretation of our work as a more abstract problem in optimal commodity transportation. In Section IX we present some concluding remarks, and in particular discuss how our results could be used in a practical setting to optimize the performance of wireless sensor networks with a modest number of nodes.

[1]Formal definitions of $\mathbf{T}$ and $d(x,y)$ appear in Section II.

II. Network Model

In this section, we first introduce three macroscopic quantities: the information density function[2] $\rho(x, y)$, the node density function $d(x, y)$, and the traffic flow function $\mathbf{T}(x, y)$. We then derive an equation linking $\rho(x, y)$ and $\mathbf{T}(x, y)$.

A. Macroscopic Quantities

We consider the unbounded two-dimensional xy plane, on which are placed distributed sources and sinks of information. We model the sources and sinks jointly, by the continuous **information density function** $\rho(x, y)$, which is measured in $\mathrm{bps/m}^2$. At locations (x, y) where $\rho(x, y) > 0$, there is a distributed data source, such that the rate with which information is generated within a surface of infinitesimal area ϵ, centered at (x, y), is $\epsilon\rho(x, y)$. At locations where $\rho(x, y) < 0$, there is a distributed data sink, such that the required absorption rate within a surface of infinitesimal area ϵ, centered at (x, y), is $-\epsilon\rho(x, y)$.

We require that the total rate with which sinks must absorb data is the same as the total rate with which the data is created at the sources. This requirement translates into the equation

$$\int \rho(x, y)\, dS = 0 \tag{3}$$

where the surface integral is taken over the whole plane.

To facilitate the transfer of information from the sources to the sinks, we are given a large number of wireless nodes, that we are free to place anywhere on the plane. Because we assume the number of nodes to be very large, we will describe their placement not in terms of their individual positions which are *microscopic* quantities, but rather in terms of a *macroscopic* quantity, the **node density** $d(x, y)$, measured in $\mathrm{nodes/m}^2$, and assumed continuous. The total number of nodes, N, is given by

$$N = \int d(x, y)\, dS. \tag{4}$$

We stress that the assumption that the network is massively dense does not imply that the number of nodes N is infinite, but rather that N is very large. A similar situation occurs in Electromagnetism, where we model the electric charges that exist in a medium by a density function, without assuming that the total number of ions or their net electrical charge is infinite.

In networks, the flow of information is typically described in terms of the rate with which information arrives at individual nodes. However, in our setting, we have a massively dense network, in which the rate of arrival of information in a particular node is a microscopic quantity. In this setting, we can best model the flow of traffic in the network in terms of the **traffic flow function** $\mathbf{T}(x, y)$, which is a macroscopic quantity. $\mathbf{T}(x, y)$ is a continuous vector function whose magnitude is measured in bps/m. It is defined so that *i)* its direction coincides with the direction of the flow of information at point (x, y), and *ii)* $\epsilon|\mathbf{T}(x, y)|$ equals[3] the rate with which information crosses a linear segment of incremental length ϵ, that is centered on (x, y), and is perpendicular to $\mathbf{T}(x, y)$.

When viewing a specific location of the network, one may observe many distinct streams of traffic, possibly along different directions. However, the fact that the traffic streams all carry the same type of packets, allows us to combine them by performing vector addition, and thus abstract the movement of data at the microscopic level by a simple macroscopic quantity, the traffic flow function at that location. Our argument is identical to the argument used to justify the abstraction of the flow of a liquid in terms of a single vector function: Microscopically, different molecules of the liquid will be traveling along different directions and with different velocities. However, when the liquid is viewed from an adequate distance, a single dominant traffic direction and intensity emerge, that can be described jointly in terms of a flow vector.

B. The Divergence of the Traffic Flow Function

Let A_0 be a surface on the xy plane, of arbitrary shape. We will denote its boundary curve by ∂A_0 and its total area by $|A_0|$. For information to be conserved, it is necessary that the rate with which information is created in the area is equal to the rate with which information is leaving the area through its boundary ∂A_0. In other words, the following equality must hold:

$$\int_{A_0} \rho(x, y)\, dS = \oint_{\partial A_0} [\mathbf{T}(x, y) \cdot \hat{\mathbf{n}}]\, ds \tag{5}$$

where $\hat{\mathbf{n}}$ is the unit vector normal to the boundary curve ∂A_0 at the point (x, y) on ∂A_0, and pointing outside A_0, and the integral on the right hand side is the path integral of the function $[\mathbf{T}(x, y) \cdot \hat{\mathbf{n}}]$. This function represents the rate (measured in $\frac{\mathrm{bps}}{\mathrm{m}}$) with which information is leaving A at the point (x, y) of its boundary ∂A_0. Equation (5) must hold for any surface A_0. Therefore, it will also hold for a sequence of surfaces A_k that all include in their interior an arbitrary point (x_0, y_0), and are such that their areas $|A_k| \to 0$. Applying (5) for A_k we have

$$\int_{A_k} \rho(x, y)\, dS = \oint_{\partial A_k} [\mathbf{T}(x, y) \cdot \hat{\mathbf{n}}]\, ds. \tag{6}$$

Since $\rho(x, y)$ is assumed continuous, we have, by a known theorem of elementary calculus [14], that[4]

$$\int_{A_k} \rho(x, y)\, dS = \rho(x_0, y_0) \times |A_k| + o(|A_k|). \tag{7}$$

Combining (6) and (7), dividing by $|A_k|$, and taking the limit with respect to k, we arrive at

$$\rho(x_0, y_0) = \lim_{k \to \infty} \frac{1}{|A_k|} \oint_{\partial A_k} [\mathbf{T}(x, y) \cdot \hat{\mathbf{n}}]\, ds \triangleq \nabla \cdot \mathbf{T}(x_0, y_0). \tag{8}$$

The limit in (8) is defined as the **divergence** of the vector function $\mathbf{T}$, at the point (x_0, y_0) [15]. The divergence of the traffic flow function measures the degree with which the traffic

[2] We denote scalars by lower case letters, and vectors by bold capitals.

[3] If $\mathbf{x}$ is a vector, by $|\mathbf{x}|$ we denote its length.

[4] We write $f(x) = o(x)$ to denote that $\lim_{x \to 0} \frac{f(x)}{x} = 0$.

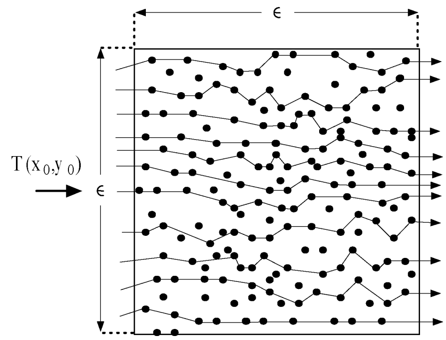


Fig. 2. A square of side ϵ, centered at the point (x_0, y_0). If the node density at that point is $d(x_0, y_0)$, then there are $\epsilon^2 d(x_0, y_0)$ nodes in the square and, with high probability, a 'highway system' can be constructed that consists of $\Theta(\epsilon\sqrt{d(x_0, y_0)})$ highways, each relaying $k_1 W$ bps from the left side to the right side.

flow increases (when information in injected in the network) or decreases (when information is removed from the network) at the particular point (x_0, y_0). In cartesian coordinates, the divergence is given by the formula $\nabla \cdot \mathbf{T} = \frac{\partial T_x}{\partial x} + \frac{\partial T_y}{\partial y}$, however its intuitive meaning can best be conveyed by the limit of (8), which is independent of the choice of the coordinate system.

To summarize, we have showed that

$$\nabla \cdot \mathbf{T} = \rho. \tag{9}$$

Note that there is nothing particular to wireless sensor networks, or to communications for that matter, in the derivations leading to (9). In fact, (9) is a well known equation of hydrodynamics. In that context, $\mathbf{T}(x, y)$ describes the flow of some liquid, and $\rho(x, y)$ models its sources and sinks.

III. Bandwidth-Limited Capacity-Achieving Physical Layer

In this section, we first specify a particular model for the capabilities of the nodes at the physical layer. We then proceed to show that, under this model, the traffic flow must be irrotational. Although the model will be significantly generalized in subsequent sections, it is worth to study it separately, as it leads to a particularly simple solution for the optimal placement of nodes.

A. Physical Layer

We make the assumption that no nodes are needed to sense the data at the sources, or deliver the data to the sinks once the data reach their physical locations. For now, the only task we assign to the nodes is the transport of the data from the physical locations of the sources to the physical locations of the sinks.

This assumption is reasonable if there is a secondary network of specialized nodes, exclusively dedicated to the sensing and delivery of the data. It is also reasonable if the most challenging task of the nodes is the transport of the data, so it is acceptable to optimize with respect to this task only, and then deploy a few extra nodes where needed to handle the sensing and the delivery tasks.

Let the wireless nodes be communicating over a common wireless channel of some finite bandwidth. Nodes can either transmit or receive at the same time. When receiving, a node is susceptible to thermal noise of power N, same for all nodes. When transmitting, a node X uses a power level P, same for all nodes, and a node Y at a distance $|X - Y|$ will receive the signal with power $KP|X - Y|^{-\alpha}$, where K is a normalizing constant and $\alpha > 2$. Let $\{X_k; k \in \mathcal{T}\}$ be the set of transmitting nodes at a given time instant. The transmission from node X_i to node X_j is received successfully iff

$$\frac{PK|X_i - X_j|^{-\alpha}}{N + \sum_{k \in \mathcal{T}, k \neq i} PK|X_k - X_j|^{-\alpha}} > \beta.$$

This means that a reception is successful iff the SINR is above a given threshold β. All successful transmissions happen with rate W bps, which we implicitly take to be a function of β and the available bandwidth.

Consider a location (x_0, y_0) of the network. We would like to determine how much traffic can pass through that location, given that the node density there is $d(x_0, y_0)$. Technically, we want to establish the maximum value for the norm $|\mathbf{T}(x_0, y_0)|$. For this, let us construct a small square of side ϵ, centered at (x_0, y_0), and oriented so that the traffic flow function $\mathbf{T}(x_0, y_0)$ is vertical to one of its sides, as shown in Fig. 2. The precise number and placement of nodes in the square will depend on how the network is constructed, but a reasonable assumption is that the nodes are thrown randomly in the square, according to a spatial Poisson process of intensity $d(x_0, y_0)$. Therefore, the expected number of nodes in the square will be $n = \epsilon^2 d(x_0, y_0)$. We need to calculate the maximum volume of traffic that can be carried into the left edge and out of the right edge of the square.

This problem was recently studied in [12], where it was shown (as an intermediate result) that the maximum possible traffic that can be carried from the left side to the right side of the square is[5] $\Theta(W\sqrt{n})$. This is achieved by the use of a "highway system" that consists of $\Theta(\sqrt{n})$ highways, each highway consisting of $\Theta(\sqrt{n})$ wireless nodes, and carrying $k_1 W$ bps. The constant k_1 captures the effects of a node having to locally share the channel with competing nodes. The highways are constructed in a manner that ensures that they can carry their traffic simultaneously, without being overwhelmed by the interference of each other. As discussed in the introduction, this result is proved by percolation theory, and is used in the tightening of a famous previous result by Gupta and Kumar [11]. It only holds with probability going to unity as $n \to \infty$.

To conclude, for the physical layer we are examining, the maximum amount of traffic that can cross a linear segment of length ϵ is on the order of $W\sqrt{n} = W\epsilon\sqrt{d(x_0, y_0)}$. Therefore, a reasonable assumption in this case is that

$$|\mathbf{T}(x, y)| \leq |\mathbf{T}(x, y)|_{\max} = c\sqrt{d(x, y)} \tag{10}$$

for some constant c that is proportional to the available bandwidth.

Clearly, our discussion does not constitute a proof of (10), but rather a justification for the use of (10) as a reasonable model of a more complicated reality.

[5] $f(n) = \Theta(g(n))$ iff $k_1 f(n) \leq g(n) \leq k_2 f(n)$, for all $n > n_0$, and for some n_0, k_1, k_2.

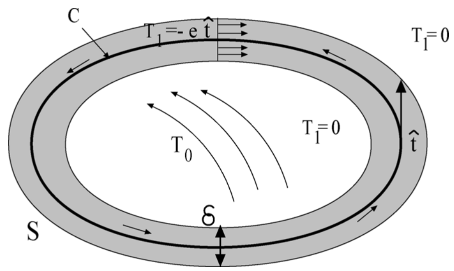


Fig. 3. The setup of the proof that the optimal traffic flow function is irrotational.

B. *The Curl of the Traffic Flow Function*

We now show that, under the model of Section III-A, among all traffic flow functions that satisfy (9), the one that needs the smallest number of nodes to be supported must also be **irrotational**, i.e., its curl must be zero everywhere:

$$\nabla \times \mathbf{T} = 0. \tag{11}$$

The curl $\nabla \times \mathbf{T}$ of a *two-dimensional* vector $\mathbf{T}$ at a point (x_0, y_0) is a scalar function defined as follows:

$$\nabla \times \mathbf{T} \triangleq \lim_{|A_k| \to 0} \frac{\oint_{\partial A_k} \mathbf{T} \cdot d\mathbf{s}}{|A_k|} \tag{12}$$

where $\{A_k\}$ is a sequence of surfaces of vanishing area, that contain (x_0, y_0) in their interior, and the integral of the right-hand side is the line integral of the function $\mathbf{T}$ over the curve ∂A_k (which is taken to have a counter-clockwise direction). Intuitively, the magnitude of the curl at a point (x_0, y_0) is a measure of how much circulation around the point (x_0, y_0) the function $\mathbf{T}$ has. The circulation is counterclockwise if the curl is positive, and clockwise if the circulation is negative. In cartesian coordinates, the curl of a function is given by $\nabla \times \mathbf{T} = (\frac{\partial \mathbf{T}_y}{\partial x} - \frac{\partial \mathbf{T}_x}{\partial y})$. A more detailed exposition on curl, with its generalization in the three dimensions (which is a vector function) can be found in [15].

We now prove (11), by assuming that it does not hold and arriving at a contradiction. In particular, suppose that the traffic flow $\mathbf{T}_0$ that needs the minimum number of nodes has a nonzero curl at some point in space. It follows from (12) that there is a curve $\mathcal{C}$, of length L, along which the line integral of $\mathbf{T}_0$ is nonzero. By choosing a proper direction for $\mathcal{C}$, we can assume that the line integral is positive

$$\oint_{\mathcal{C}} \mathbf{T}_0 \cdot d\mathbf{s} = p > 0. \tag{13}$$

As shown in Fig. 3, we form around $\mathcal{C}$ a strip S of infinitesimal and constant width δ. Because δ is infinitesimally small, the area of the strip can be taken to be $|S| = \delta \times L$.

We construct an auxiliary vector function $\mathbf{T}_1$ in the following manner: Outside the strip S, $\mathbf{T}_1 = \mathbf{0}$. Inside the strip, at a point (x, y), $\mathbf{T}_1 = -\epsilon\hat{\mathbf{t}}$, where $\hat{\mathbf{t}}$ is a unit vector tangential to $\mathcal{C}$, at the point where $\mathcal{C}$ is closest to the point (x, y), and with the same direction as $\mathcal{C}$. Therefore, we construct $\mathbf{T}_1$ to resemble the flow of a small quantity of liquid around a closed hose of impermeable boundaries, which goes *against* the average flow of $\mathbf{T}_0$ in S. By its physical interpretation, it is clear that $\mathbf{T}_1$ has a zero divergence everywhere. It is also straightforward to show this mathematically. Indeed, outside the strip $\mathbf{T}_1$ is identically zero, and inside the strip it can be shown that the divergence is zero by a direct application of the definition (8).

As the divergence operator is linear, we have

$$\nabla \cdot (\mathbf{T}_0 + \mathbf{T}_1) = \nabla \cdot \mathbf{T}_0 + \nabla \cdot \mathbf{T}_1 = \rho$$

and it suffices to show that the traffic flow function $(\mathbf{T}_0 + \mathbf{T}_1)$ can be supported by fewer nodes. Indeed, let N_0 be the total number of nodes needed to support $\mathbf{T}_0$ and N_{0+1} be the total number of nodes needed to support $\mathbf{T}_0 + \mathbf{T}_1$. We have

$$\begin{aligned} &c^2(N_0 - N_{0+1}) \\ &= \int_S (|\mathbf{T}_0|^2 - |\mathbf{T}_0 + \mathbf{T}_1|^2)\, dS \\ &= \int_S (|\mathbf{T}_0|^2 - |\mathbf{T}_0|^2 - |\mathbf{T}_1|^2 - 2\mathbf{T}_0 \cdot \mathbf{T}_1)\, dS \\ &= -\int_S \epsilon^2\, dS + \int_S 2\epsilon \mathbf{T}_0 \cdot \hat{\mathbf{t}}\, dS \\ &= -\epsilon^2|S| + 2\epsilon\delta \oint_{\mathcal{C}} \mathbf{T}_0 \cdot d\hat{\mathbf{t}} = -\epsilon^2|S| + 2\epsilon\delta p. \end{aligned}$$

The first equality comes from using (10) with the equality, and noting that the functions $\mathbf{T}_0$ and $\mathbf{T}_0 + \mathbf{T}_1$ differ only within the surface S. The last one comes from applying (13). It follows that, for a sufficiently small value of ϵ, the last expression is positive, and so the traffic flow $\mathbf{T}_0 + \mathbf{T}_1$ can be supported by a smaller number of nodes than the traffic flow $\mathbf{T}_0$. Therefore, we arrive at a contradiction, so (11) must hold.

IV. Analogy With Electrostatics

A. *Homogeneous Propagation Environments*

In the previous sections, we proved that the traffic flow function must satisfy (9) and (11). These equations jointly do not uniquely specify the traffic flow. Indeed, provided there is a solution $\mathbf{T}_0$ that satisfies both of them, then so does $\mathbf{T}_0 + \mathbf{c}$, where $\mathbf{c}$ is a constant vector. However, by Helmholtz's theorem [16], it follows that the solution to (9) and (11) exists and is uniquely specified if, in addition, we require that the traffic flow is zero at infinity: $\mathbf{T}|_\infty = \mathbf{0}$. Assuming that the traffic sources and sinks are constraint over a finite region, this is a reasonable boundary condition to take, as there is no need for the traffic flow to arrive at the sink by first going to infinity and back.[6] To summarize, it follows by Helmholtz's theorem that the equations

$$\nabla \cdot \mathbf{T} = \rho, \quad \nabla \times \mathbf{T} = 0, \quad \mathbf{T}|_\infty = \mathbf{0} \tag{14}$$

uniquely specify the optimal traffic flow function. Once the optimal traffic flow is known, by solving the system (14), the cor-

[6]Note that Helmholtz's theorem is typically mentioned in a three dimensional setting, however its two dimensional version follows as a special case.

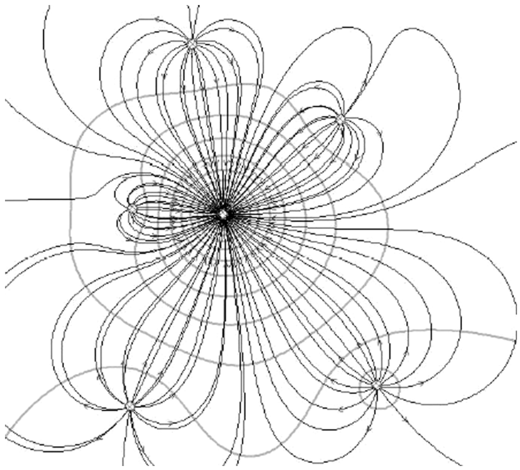


Fig. 4. Field lines (in thin black) and lines of constant potential (in thick gray), in a two-dimensional topology consisting of five positive singular charges of equal magnitude, and a single singular negative charge of five times that magnitude, placed in a homogeneous dielectric.

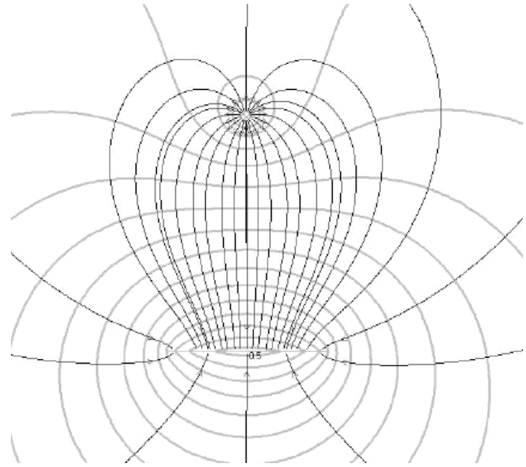


Fig. 5. Field lines (in thin black) and lines of constant potential (in thick gray), in a two dimensional topology consisting of a singular positive charge and a linear uniform distribution of negative charge, of equal total magnitude, placed in a uniform dielectric.

responding optimal node density can easily be derived, by applying (10) with equality.

However, it is a basic fact of electrostatic field theory [16], that (14) also uniquely specify the electric displacement $\mathbf{D}$ induced by a two-dimensional electric charge density $\rho(x, y)$ in a homogeneous dielectric (for example free space).[7] Note that, in homogeneous dielectrics, the electric displacement $\mathbf{D}$ is simply proportional to the electric field $\mathbf{E}$, i.e., $\mathbf{D} = \epsilon_r \epsilon_0 \mathbf{E}$, where ϵ_0 is the permittivity of free space and ϵ_r is the relative permittivity of the dielectric (equal to 1 for the case for free space). Note also that in the charge density ρ we only count free charges, and not charges that are induced by the polarization of the dielectric. Therefore, our optimal placement problem is identical to a standard problem of Electrostatics, namely the determination of the electric displacement (equivalently the electric field) in a homogeneous dielectric, in the presence of a distribution of free electric charge.

As an example, let us consider the topology of Fig. 4, in which we have placed in a homogeneous dielectric five positive singular charges of equal magnitude around a single singular negative charge of five times that magnitude. The induced electric displacement (and the associated electric field) can be calculated by using any of a large number of software tools that are available, either for solving arbitrary PDEs, or for solving Electrostatics problems in particular. In this, and the following examples of this section, we use the specialized software tool of [17]. As is standard in Electrostatics, in the figure we denote the electric field by **field lines**. These are defined in the following manner: The field line crossing a point (x, y) is parallel to $\mathbf{E}(x, y)$, and the density of field lines at that point is proportional to the magnitude $|\mathbf{E}(x, y)|$.

The above figure also has an interpretation in the context of wireless sensor networks. In particular, the field lines show the optimal packet trajectories in an environment in which there are five singular traffic sources of equal magnitude, and a single, singular traffic sink at the center, collecting all the created information. As there is a convergence of field lines toward the central traffic sink, more nodes will be placed around it, in order to support the large volumes of traffic, in accordance to condition (10) taken with equality.

As a second example, in Fig. 5 we plot the field lines in a topology consisting of a singular positive charge, and a distribution of negative charge, of equal total magnitude, along a horizontal linear segment. The field lines of the figure are also the trajectories of packets in a topology where the positive and the negative charges are substituted with a traffic source and a traffic sink, respectively.

As expected, the optimal node placement induces a traffic flow that is heaviest in the region between the source and the sink. The intuitive explanation is that the routes through this region are the shortest. On the other hand, quite a lot of traffic will actually travel along much longer routes, some of it actually arriving to the sink from below. The intuition behind this result is that, if *all* packets use short, more direct routes, then the congestion in the central region will be so high as to require a very large number of nodes to be supported, more than the number of nodes needed if some of the packets take a longer, but much less congested, route.

B. Nonhomogeneous Propagation Environments

Until now it has been assumed that all parts of the wireless network are equally efficient. This is a reasonable assumption when the network consists of a set of identical nodes, and in addition all parts of the environment present a similar challenge to the network. However, these assumptions may not always hold. For example, perhaps part of the network may be in an environment with heavy vegetation, which increases the attenuation of the signals with distance. As another example, parts of the bandwidth may not be available everywhere.

Such cases can be modeled by assuming that the coefficient c appearing in (10) is no longer a constant, but is a function

[7]A two-dimensional density is a density immersed in the two dimensional space. Equivalently, it could also mean a density immersed in the three dimensional space that is invariant with respect to the z coordinate. In the latter case, the electric displacement it induces will also be invariant with respect to the z coordinate, and with a zero z-component.

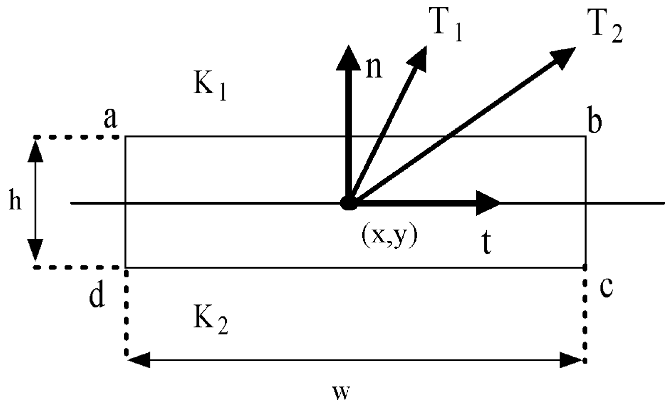


Fig. 6. The boundary between two regions with different node capabilities.

of the location. In particular, we assume that the xy plane is partitioned in a number of **propagation regions** P_i, where $i = 1, \ldots, p$, each associated with a coefficient c_i, such that within P_i condition (10) is substituted with

$$|\mathbf{T}(x,y)| \leq |\mathbf{T}(x,y)|_{\max} = c_i c\sqrt{d(x,y)}. \tag{15}$$

We now develop boundary conditions that connect the two traffic flow vectors across the two sides of the same boundary. Let us concentrate, with no loss of generality, at a point (x,y) on the boundary of regions P_1 and P_2. As shown in Fig. 6, let $\hat{\mathbf{n}}$ and $\hat{\mathbf{t}}$ be respectively the normal and tangential unit vectors of the boundary at (x,y). Also, let $\mathbf{T}_1$ and $\mathbf{T}_2$ be the traffic flows at point (x,y), at the two sides of the boundary, which we decompose as follows:

$$\mathbf{T}_1(x,y) = T_{n1}\hat{\mathbf{n}} + T_{t1}\hat{\mathbf{t}}, \quad \mathbf{T}_2(x,y) = T_{n2}\hat{\mathbf{n}} + T_{t2}\hat{\mathbf{t}}.$$

We first develop a boundary condition involving the normal components of the traffic flow. For this, let us apply (5) on the perpendicular region $A = \mathrm{abcd}$, shown in Fig. 6, centered at (x,y) and with height h and width w. The width w is taken to be so small, that the boundary appears locally as a straight line. By taking first $h \to 0$, (5) becomes

$$\oint_{\partial A} [\mathbf{T}(x,y) \cdot \hat{\mathbf{n}}]\, ds = 0.$$

As $\mathbf{T}$ must be continuous on each side of the boundary, we have that

$$\oint_{\partial A} [\mathbf{T}(x,y) \cdot \hat{\mathbf{n}}]\, ds = [T_{n1}(x,y) - T_{n2}(x,y)]w + o(w)$$

and by dividing by w and taking $w \to 0$, we have that

$$T_{n1} = T_{n2}. \tag{16}$$

Next, we develop a boundary condition on the tangential components of the traffic flow. For this, let us consider the two streams of traffic moving along either side of the boundary.

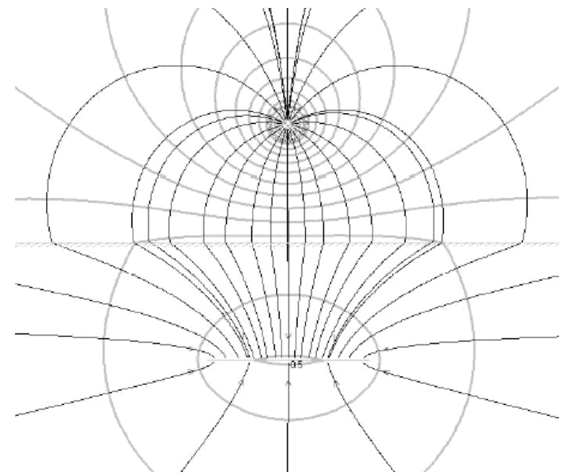


Fig. 7. The setting of Fig. 5, in which we have replaced the lower half of the plane with a dielectric with $\epsilon_r^1 = 10$, and the upper half of the plane is free space with $\epsilon_r^2 = 1$.

Since the traffic is optimally distributed, i.e., it uses the minimum number of wireless nodes, it follows that the moving of an incremental part of the tangential component of the traffic from one to the other side can not result to a net change of the number of nodes needed. Therefore, we must have

$$\frac{\partial T_{t1}}{\partial d_1} = \frac{\partial T_{t2}}{\partial d_2}$$

where d_1 and d_2 are the node densities on the two sides of the boundary. Noting that $T_{n1}^2 + T_{t1}^2 = c_1^2 c^2 d_1$ and $T_{n2}^2 + T_{t2}^2 = c_2^2 c^2 d_2$, and using (16), after some straightforward algebra, this equation becomes

$$(c_2^2)T_{t1} = (c_1^2)T_{t2}. \tag{17}$$

As is known from elementary Electromagnetics [16], [18], the boundary conditions (16) and (17) must also be satisfied by the electric displacement $\mathbf{D}$ if the regions P_i contain dielectrics characterized by relative permittivities $\epsilon_r^i = c_i^2$. In addition, these boundary conditions, together with the (14), which continue to hold at the interior of the regions P_i, uniquely specify the electric displacement.

As an example, in Fig. 7 we plot the field lines that are created by a distribution of free charge similar to that of Fig. 5, but where we now assume that the lower half of the plane is occupied by a dielectric with $\epsilon_r^1 = 10$, and the upper half of the plane is free space with $\epsilon_r^2 = 1$. In the context of wireless sensor networks, Fig. 7 shows how the optimal packet trajectories of the network of Fig. 5 would be modified if the lower half of the plane was modeled by (15) with a factor $c_1 = \sqrt{10}$ and the upper half of the plane was modeled by (15) with a factor $c_2 = 1$.

The introduction of different propagation regions allows us to handle the case where there are regions in which no nodes can be placed. As an example, we may have a situation in which both the traffic sources and the traffic sinks are placed in a large room, and we are not allowed to place any nodes outside the floor of the room.

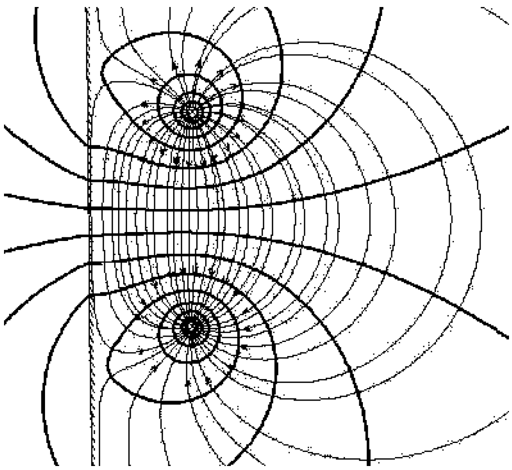


Fig. 8. Field lines (in thin black) and constant potential lines (in thick gray) created by a positive and a negative charge of equal magnitude, that are placed inside a dielectric of very large relative permittivity, which is adjacent to free space.

This case can be handled by assuming the existence of a special propagation region P_0 with a constant $c_0 \ll c_i$, for all $i = 1, \dots, p_i$. In this case, the cost of placing nodes in that propagation region is too high, because the physical layer in that region is very weak, as seen by (15). In the limit when $\frac{c_0}{c_i} \to 0$, for all $i = 1, \dots, p$, the optimal traffic pattern, as determined by solving (14), jointly with (16), (17), avoids routing any traffic through P_0, unless of course the topology absolutely requires that packets must pass through P_0. (That would be the case if, for example, there were sources or sinks of traffic within P_0.)

As an example, in Fig. 8 we have plotted the field lines created by a positive and a negative charge placed close to each other within a dielectric of very large relative permittivity, and close to a linear boundary surface with free space. The field lines also show the optimal traffic flow in the case when a point source and a point sink are placed close to the linear boundary with a region through which no traffic can flow.

C. Traffic Sources and Sinks With Limited Mobility

Until now, we have assumed that each location of infinitesimal size (x, y) is associated with a *fixed* rate of information creation (or absorption). However, there are situations in which the placement of data sources and sinks is also subject to optimization. As an example, let us consider a sensor network designed to monitor the levels of humidity and temperature of a large plantation, and forward the measurements to a large central building. If we assume that a large number of wireless receivers, connected over high capacity wired links with a central traffic sink, are placed along the circumference of the building, then the sensor network should be free to select which parts of the circumference of the building should receive how much traffic, in a way that minimizes the number of wireless nodes that must be deployed.

We model such scenarios by defining a set of t **traffic regions** $\{T_i\}$, where $i = 1, \dots, t$. The information density function ρ is only defined outside the traffic regions, in $\cup_{i=1}^{p} P_i = \mathbf{R}^2 - \cup_{i=1}^{t} T_i$. Each T_i is associated with an **information rate** Q_i, measured in bps, which represents the net amount of sources/sinks that must be placed in T_i. We assume that data can move with no cost inside T_i, and we require Q_i to be distributed only on the boundary of T_i, ∂T_i. Equation (1) is modified as follows:

$$\int_{\cup_{i=1}^{p} P_i} \rho(x, y)\, dS + \sum_{i=1}^{t} Q_i = 0.$$

For any distribution $\mathcal{D}$ of the information rates $\{Q_i\}$ on the boundaries of the $\{T_i\}$, there is an optimal node distribution, $d_{\mathcal{D}}(x, y)$ that minimizes the number of wireless nodes needed to support the traffic. A problem that arises naturally, is to find the optimal distribution $\mathcal{D}_{\text{opt}}$ of the rates $\{Q_i\}$ on the boundaries of the $\{T_i\}$, whose optimal node distribution $d_{\mathcal{D}_{\text{opt}}}(x, y)$ needs the minimum number N of sensor nodes. In other words, we have a problem that consists of two consecutive minimizations.

Let us consider the electrostatic equivalent of the total number of nodes N:

$$N = \int_{\cup_{i=1}^{p} P_i} d(x, y)\, dS = \int_{\cup_{i=1}^{p} P_i} \frac{|\mathbf{T}|^2}{c_i^2 c^2}\, dS$$
$$= \frac{\epsilon_0}{c^2} \int_{\cup_{i=1}^{p} P_i} \frac{|\mathbf{D}|^2}{\epsilon_0 \epsilon_r^i}\, dS = \frac{2\epsilon_0}{c^2} \mathcal{E}.$$

In the second equality, we use (15). In the third equality, we move from the networking quantities to their electrostatic equivalents. In the last equality, we substitute for the electrostatic field energy [16], [18]

$$\mathcal{E} = \frac{1}{2} \int_{\cup_{i=1}^{p} P_i} \mathbf{E} \cdot \mathbf{D}\, dS.$$

Therefore, the total number of nodes of the networking setting is mapped to the total energy of the Electrostatics setting, up to a constant coefficient.

We are now ready to consider the Electrostatics equivalent of our minimization problem: we have a setting with a fixed spatial electric charge density $\rho(x, y)$, and a set of regions $\{T_i\}$ on which we have placed a set of charges $\{Q_i\}$. Our assumption that the charges Q_i can move everywhere along their corresponding regions T_i means that these regions are conductors. Therefore, the equivalent problem becomes the calculation of the distribution of electric charge on a set of surfaces, such that the energy of the electric field, $\mathcal{E}$, is minimized.

Fortunately, this is exactly the same problem that nature solves when placing charges on conductors. In particular, Thomson's theorem [16], [18] states that charges placed on conductors distribute themselves so that the energy of the electric field is minimized. Furthermore, the electric displacement of the resulting field is uniquely specified by the boundary condition

$$\mathbf{D}(x, y) \cdot \hat{\mathbf{t}} = 0 \tag{18}$$

where (x, y) is any point on the boundary ∂T_i of a region T_i, and $\hat{\mathbf{t}}$ is the unit vector parallel to the boundary ∂T_i at point (x, y), together with the conditions (that follow from Gauss's law):

$$\oint_{\partial T_i^+} [\mathbf{D} \cdot \hat{\mathbf{n}}]\, dl = Q_i, \ \ \forall i = 1, \dots, m. \tag{19}$$

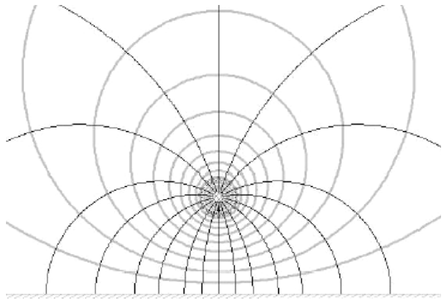


Fig. 9. Field lines (in thin black) and constant potential lines (in thick gray) when a singular positive charge is placed over a infinite conducting plane, infused with a negative charge of the same magnitude.

In the above conditions, $\hat{\mathbf{n}}$ is a unit vector normal at each point of the boundary and pointing outwards, and ∂T_i^+ is a closed curve that is continuously tangential to the boundary ∂T_i but lies outside of T_i. This complication of taking the integral along ∂T_i^+ and not along ∂T_i is due to the fact that the function $\mathbf{D}$ is in general discontinuous on ∂T_i.

Going back to sensor networks, it follows that the optimal distribution of traffic sources and sinks should create a traffic flow similar to the electric field induced by a placement of charges on a set of conductors, and this field $\mathbf{T}$ is uniquely determined by (18) and (19) (substituting $\mathbf{D}$ with $\mathbf{T}$, and taking ρ to be the information density function and the Q_i to be information rates) together with (14), and possibly (16) and (17).

Both (18) and (19) have simple meanings in the context of sensor networks, and should have been anticipated. Equation (18) requires that packets arriving at a traffic region hit the traffic region vertically. Indeed, if the traffic also has a tangential component, the packets arrive at the traffic source using a route that is longer than the absolute necessary, and a rearranging of the Q_i at the surface of the region will result to a traffic that needs fewer nodes to be supported. Equation (19) simply states that the total net rate with which information is leaving the traffic region T_i is equal to the net rate with which we stipulate that information is created inside the region.

As an example, in Fig. 9, we have placed a positive electric charge within a uniform dielectric and close to the linear boundary with a conductor infused with a negative electric charge of the opposite magnitude. In the figure, we have plotted the electric field, which can be calculated easily for this topology by the method of images [16]. The above figure has a dual interpretation in the context of wireless sensor networks. In particular, it shows the optimal routes that packets must follow when moving, through a uniform propagation environment, from a singular source of information to a linear traffic sink, when there is no particular restriction on how many packets should be received at each location on the boundary of the sink.

D. The Potential Function

In Electrostatics, the electric field $\mathbf{E}$ can be described in terms of the scalar **potential function** $U(\cdot)$. The two are related by the equations

$$\mathbf{E} = -\nabla U$$

where ∇U denotes the gradient of U, and

$$U(A) - U(B) = \int_A^B \mathbf{E} \cdot d\mathbf{s}$$

where the line integral is along any curve that starts at A and ends at B. The difference $U(A) - U(B)$ denotes the amount of energy that the field transfers to a positive unit charge as it moves from point A to point B. In Figs. 4, 5, 7, 8, and 9 we have plotted (in thick gray) lines of constant potential. As the potential is defined as the negative gradient of the electric field, the lines of constant potential intersect vertically the electric field lines.

A natural question to ask is the meaning of the potential in our sensor networks context. For this, we consider a curve $\mathcal{C}$ along the trajectory of a packet stream, starting at a point A and ending at a downstream point B, and possibly crossing different propagation regions. We have

$$\begin{aligned} U(A) - U(B) & \\ &= \int_A^B \mathbf{E} \cdot d\mathbf{s} = \int_A^B \frac{1}{\epsilon_r^i \epsilon_0} \mathbf{D} \cdot d\mathbf{s} = \frac{1}{\epsilon_0} \int_A^B \frac{1}{c_i^2} \mathbf{T} \cdot d\mathbf{s} \\ &= \frac{1}{\epsilon_0} \int_A^B \frac{1}{c_i^2} |\mathbf{T}| d\mathbf{s} = \frac{c}{\epsilon_0} \int_A^B \frac{1}{c_i} \sqrt{d(x,y)}\, ds. \end{aligned}$$

In the third equality, we moved from Electrostatics variables to networking variables. The fourth equality comes from noting that, by its construction, curve $\mathcal{C}$ is parallel to $\mathbf{T}$, and the inner product can be removed. The fifth equality comes from requiring that (15) holds with equality, so that at each point the network does not have more nodes than needed, and the network uses indeed the minimum number of nodes.

Note that $\sqrt{d(x,y)}\ ds$ is the approximate number of hops that a packet makes in order to traverse an incremental length ds at a point (x, y) where the node density is $d(x, y)$. Indeed, if $d(x,y)(ds)^2$ nodes are randomly placed in a square of size $(ds)^2$, there will be roughly $\sqrt{d(x,y)(ds)^2} = \sqrt{d(x,y)}ds$ nodes along each side of the square, and roughly as many hops will be needed by a packet to traverse the square from one edge to an opposite edge.

Therefore, if the curve $\mathcal{C}$ continuously stays in the same propagation region, then the difference $U(A) - U(B)$ is equal to the number of hops needed to go from A to B, multiplied by $\frac{c}{c_i \epsilon_0}$. In the general case, when $\mathcal{C}$ crosses multiple propagation regions, the difference $U(A) - U(B)$ is a weighted sum of the hops needed in each region, with the weight for region P_i equal to $\frac{c}{c_i \epsilon_0}$.

We stress that, although we are assuming a massively dense network, the total number of nodes in the network, and consequently the total number of hops to go from point A to point B, are both finite, albeit very large. The situation is very similar to that of Electrostatics, where we approximate the distribution of charges by a charge density, without assuming that the actual number of distinct charges is infinite.

We now summarize the findings of this section in the form of a theorem:

Theorem 1: Let the xy plane be partitioned in t traffic regions T_i, $i = 1, \ldots, t$, each associated with an information rate Q_i, and p propagation regions P_i, $i = 1, \ldots, p$, each characterized by a coefficient c_i such that the maximum norm for the traffic flow is given by

$$|\mathbf{T}(x,y)| \le |\mathbf{T}(x,y)|_{\max} = c_i c\sqrt{d(x,y)}.$$

Let $\rho(x,y)$ be an information density function, defined everywhere in $\cup P_i$.

The optimal distribution of nodes $d(x,y)$, that minimizes the number of nodes needed to transport all the traffic, induces a traffic flow $\mathbf{T}(x,y)$ that can be calculated by solving the differential equations

$$\nabla \cdot \mathbf{T} = \rho, \qquad \nabla \times \mathbf{T} = 0$$

together with the boundary conditions

$$\begin{aligned}
\mathbf{T}(x,y) &\to \mathbf{0}, && (x,y) \in \cup P_i, (x,y) \to \infty \\
\mathbf{T}(x,y)\cdot\hat{\mathbf{t}} &= 0, && (x,y) \in \partial T_i, \\
\oint_{\partial T_i^+} [\mathbf{T}\cdot\hat{\mathbf{n}}]\, ds &= Q_i \\
\mathbf{T}_i(x,y)\cdot\hat{\mathbf{n}} &= \mathbf{T}_j(x,y)\cdot\hat{\mathbf{n}}, && (x,y) \in \partial P_i \cap \partial P_j \\
c_j^2 \mathbf{T}_i(x,y)\cdot\hat{\mathbf{t}} &= c_i^2 \mathbf{T}_j(x,y)\cdot\hat{\mathbf{t}}, && (x,y) \in \partial P_i \cap \partial P_j
\end{aligned}$$

where $\mathbf{T}_i(x,y)$ is the limit of $\mathbf{T}$ at (x,y) from within P_i and $\partial P_i \cap \partial P_j$ is the boundary of the propagation regions P_i and P_j.

The optimal traffic flow is identical to the electric displacement vector $\mathbf{D}$ that exists if the T_i are conductors infused with charges Q_i and the P_i are dielectrics with relative permittivities $\epsilon_r^i = c_i^2$, infused with a free charge density ρ.

Let $U(\cdot)$ be the potential function of the electrostatic setting. Let A be a point in any of the propagation regions, and let B be a location downstream from A, in the same or another propagation region. In the network setting, $U(A) - U(B)$ equals the weighted sum of the hops taken by the packets in each region, with the weight for region P_i equal to $\frac{c}{c_i \epsilon_0}$:

$$U(A) - U(B) = \frac{c}{\epsilon_0}\int_A^B \frac{1}{c_i}\sqrt{d(x,y)}\, ds.$$

In the following sections we will extend Theorem 1 for the case of a physical layer model that is significantly more general than the model of (10). As we will see, the optimal traffic flow resembles an electrostatic field only in the special case when (10) holds. Before leaving the electrostatic case, however, it should be noted that a similar, but different, analogy between wireless networks and Electrostatics has also been explored, independently, in [6], [7]. In [7] the authors consider a wireless sensor network consisting of a central data sink and a large number of sensor nodes forming a wireless network. The placement of nodes is not subject to any optimization, The authors are interested in finding energy efficient routes to the central sink. They propose a quadratic optimization problem which also leads to the condition that the traffic must also be irrotational, and so must also resemble an electrostatic field. In this case, however, the quadratic optimization problem is not motivated by physical layer considerations, as in our case, but rather is adopted because of its intuitive appeal. Nevertheless, it leads to important energy savings. The authors also consider the network equivalents of dielectric materials and the potential function, but the analogies they use are different from our own. In [6] the same authors use a similar approach to suggest a heuristic, but very intuitive, optimization problem for optimizing the flow of traffic in a wireless ad hoc network with multiple types of traffic.

V. Network Model-Part II

We now slightly modify the model of Section II, to assume that all sources, sinks, and wireless nodes exist within a compact, i.e., closed and bounded, region A. We parametrize its boundary C by its arc length s, so that $C(s)$ traces out all the points in the curve as s goes from 0 to L. We denote by $\hat{\mathbf{n}}(s)$ the unitary vector normal to C at the point $C(s)$, and pointing outwards.

To model the sources and sinks that exist on the boundary of the region, we define the **boundary information density function** $\rho_b(s)$, measured in bps/m and assumed continuous. If $\rho_b(s) > 0$, then there is a distributed source of information located at the boundary point $C(s)$, such that the rate with which information is entering the network though an infinitesimal arc of length ds centered at that point is $\rho_b(s)\, ds$. If, however, $\rho_b(s) < 0$, then there is a sink of information located at the boundary point $C(s)$, such that the rate with which information is removed from the network through an infinitesimal arc of length ds of the boundary, centered at that point, is $-\rho_b(s)\, ds$. As the total volume of traffic entering the network must be equal to the total volume of traffic leaving the network, we substitute (3) with

$$\int_A \rho(x,y)\, dA + \oint_C \rho_b(s)\, ds = 0.$$

Clearly, an equation is needed that connects $\rho_b(s)$ with $\mathbf{T}(x,y)$, analogous to (9). To this end, let us consider a point $C(s) = (x_0, y_0)$ on the boundary of region A. As shown in Fig. 11, we construct a rectangle abcd of infinitesimal size, that is inside the region A and touches the boundary at the segment ab, which includes (x_0, y_0). The rectangle has width w and height h that are so small, that the boundary can be viewed locally as a straight line. Conservation of data requires that

$$\begin{aligned}
&\int_{\mathrm{a}}^{\mathrm{b}} \rho_b(s)\, ds + \int_{\mathrm{abcd}} \rho(x,y)\, dA \\
&\qquad = -\int_{\mathrm{a}}^{\mathrm{d}} [\mathbf{T}\cdot\hat{\mathbf{t}}]\, ds - \int_{\mathrm{d}}^{\mathrm{c}} [\mathbf{T}\cdot\hat{\mathbf{n}}]\, ds + \int_{\mathrm{c}}^{\mathrm{b}} [\mathbf{T}\cdot\hat{\mathbf{t}}]\, ds.
\end{aligned}$$

As the rectangle is very small, this equation can be approximated by:

$$\begin{aligned}
&w\rho_b(x_0,y_0) + wh\rho(x_0,y_0) \\
&\quad \simeq -h[\mathbf{T}(x_0,y_0)\cdot\hat{\mathbf{t}}] - w[\mathbf{T}(x_0,y_0)\cdot\hat{\mathbf{n}}] + h[\mathbf{T}(x_0,y_0)\cdot\hat{\mathbf{t}}].
\end{aligned}$$

Dividing by w, and then taking $h \to 0$, we arrive at the condition

$$\rho_b = -\mathbf{T}\cdot\hat{\mathbf{n}} \tag{20}$$

which must hold at any point $C(s)$ along C.

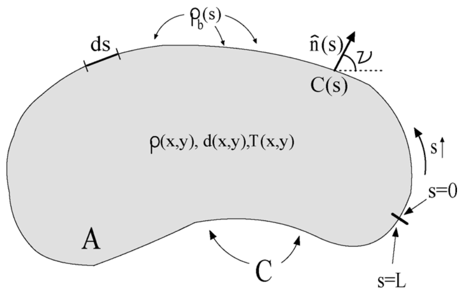


Fig. 10. The compact region A where the functions $\rho(x,y)$, $\rho_b(s)$, $d(x,y)$, and $\mathbf{T}(x,y)$ are specified.

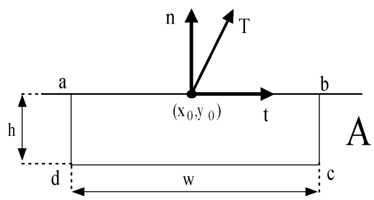


Fig. 11. A rectangle abcd of infinitesimal size placed inside A, and adjacent to its boundary around point (x_0, y_0).

VI. General Physical Layer

In this section, we first revisit and refine the physical layer model of Section III, and then consider an alternative, Ultra Wideband (UWB) model. We conclude by unifying the two models under a general model, which also accounts for the requirement that nodes must not only transport the data, but also sense them at the sources, and deliver them at the destinations once their physical location is reached.

A. Bandwidth Limited Physical Layers

The discussion that motivated (10) critically hinges on the assumption that a signal that is transmitted with power P will be received with power $PKd^{-\alpha}$, where d is the traveled distance. This implies that, as $d \to 0$, the received power becomes larger than the transmitted power, and in fact approaches infinity! A more realistic model would be to assume that the received power is given by a bounded function of the distance, for example the function $PK(\max\{d, d_{\min}\})^{-\alpha}$, for some appropriately chosen critical distance $d_{\min}$. In that case, however, it is intuitively clear that the traffic capabilities of the network will be dramatically altered. This problem was examined in [19]. There, it was shown that, as the number of nodes in a network n increases, at first the maximum achievable aggregate throughput that can be carried by the network increases like $\sqrt{n}$, however after some point the bounded nature of the power transfer function starts to have an effect, and for very large values of n the maximum aggregate throughput saturates at a constant value. Therefore, a more accurate macroscopic model for the capabilities of the network would be:

$$|\mathbf{T}(x,y)| \leq |\mathbf{T}(x,y)|_{\max} = k_1[d(x,y)]^{\beta(d(x,y))}, \qquad (21)$$

where now the exponent $\beta(d(x,y))$ is close to $\frac{1}{2}$ for small values of $d(x,y)$, but converges to 0 for $d(x,y) \to \infty$. The precise form of the exponent function $\beta(\cdot)$ will depend on the exact law by which power propagates over small distances, but an investigation toward this direction goes beyond the scope of this work.

The discussion until now assumed that the nodes can coordinate optimally to achieve the maximum possible transfer of traffic. In practice, this is not the case, as nodes have to operate under media-access control (MAC) protocols that operate suboptimally, particularly as the network becomes more and more dense. In such a case, the behavior of the physical layer could be approximated by (21), where now the exponent $\beta(\cdot)$ also models the imperfection of the MAC protocol used. Such a modeling also goes beyond the scope of our work.

B. Ultra Wideband Physical Layers

Until now it was assumed that all nodes transmit over a common wireless channel, interfering with each other's transmissions in the process. This is provably the best that nodes can do when the total bandwidth available for communications is limited [11]. Let us now consider the case where the available bandwidth is very large, ideally infinite. This is the case of ultrawideband (UWB) communications. As UWB transceivers are very inexpensive and simple to make, and still have excellent performance, it is expected that in the future many wireless sensor networks will be using UWB technology.

Since the available bandwidth is infinite, each transmission can occupy its own portion of the bandwidth, and will only be hampered by thermal noise. It would seem at first that, with infinite bandwidth, comes infinite capacity. This is not so, as, the more bandwidth we use, the greater becomes the power of the thermal noise. This can be seen by Shannon's formula for the additive white Gaussian (AWGN) channel

$$C = W\log_2\left(1 + \frac{P_r}{\eta W}\right) \underset{W\to\infty}{\longrightarrow} (\log_2 e)\frac{P_r}{\eta} \qquad (22)$$

where C is the capacity, W the available bandwidth, η is the noise spectral density, and P_r is the received power. Therefore, channels that are not bandwidth constrained are necessarily power constrained.

The capacity of wireless networks under an UWB physical layer was recently studied in [20]. There, the authors consider a setting in which n nodes are placed on a unit disk, each node randomly selecting as its destination another node in the network. Each node transmits with power P, and a receiver at a distance d will receive the signal with power $PKd^{-\alpha}$. Receivers are susceptible to thermal noise of spectral density η, and each transmission occupies its own bandwidth, which is so large that the limit of (22) is achieved, and all communication is with rate $\frac{(\log_2 e)PK}{\eta d^{\alpha}}$.

In this setting, the authors show that, with probability going to 1 as the number of nodes n goes to infinity, the maximum achievable aggregate throughput is[8] $\tilde{\Theta}(n^{\frac{\alpha+1}{2}})$. This result is in sharp contrast with the capacity result of Gupta and Kumar who found that, under limited bandwidth, but an identical topology and traffic pattern, the maximum achievable aggregate traffic is only $\tilde{\Theta}(n^{\frac{1}{2}})$ [11]. The gain comes from having infinite

[8]The notation $f(n) = \tilde{\Theta}(g(n))$ means that $k_1(n)f(n) \leq g(n) \leq k_2(n)f(n)$, for all $n > n_0$, and $k_1(n)$ and $k_2(n)$ are rational functions of $\log n$ [20].

bandwidth. Indeed, as more and more nodes are placed in the network, the average distance between nearest neighbors decreases, hence the received power increases. As there is no interference, this leads to an increase of the capacity of the links between neighbors, as specified in (22). Therefore, it is best for nodes to communicate by using multiple hops, exclusively between neighbors. Even after accounting for the effects of transmitting the same information multiple times, the end effect is that the maximum achievable aggregate throughput can increase very fast with the number of nodes, as $\tilde{\Theta}(n^{\frac{\alpha+1}{2}})$, and this is provably the maximum throughput that we can squeeze out of the network [20].

The calculations of [20] assume that nodes are placed on a disk, and the traffic starts and ends at nodes within the disk. However, the calculations can easily be modified to hold for the case where n nodes are placed in a square and the traffic must be carried from the left edge to the right edge, as in Fig. 2. Working as in the derivation of (10), we are motivated to consider the following macroscopic model:

$$|\mathbf{T}(x,y)| \leq |\mathbf{T}(x,y)|_{\max} = k_2[d(x,y)]^{\frac{\alpha+1}{2}}. \tag{23}$$

The derivation of the capacity bound of [20] critically hinges on the assumption that a signal transmitted with power P will be received by a node at a distance d with power $PKd^{-\alpha}$. As discussed in Section VI-A, this is unrealistic for very small distances d, as it implies that the received power can be arbitrarily large; a more realistic model would use a bounded power transfer function. In addition, we have assumed that nodes coordinate perfectly to achieve the capacity. In reality, they will be operating under MAC protocols that may be suboptimal, particularly as the node density increases. Therefore, a more realistic macroscopic model would be that

$$|\mathbf{T}(x,y)| \leq |\mathbf{T}(x,y)|_{\max} = k_3[d(x,y)]^{\beta(d(x,y))} \tag{24}$$

where $\beta(d(x,y))$ is approximately equal to $\frac{\alpha+1}{2}$ for small values of $d(x,y)$, but decreases as $d(x,y) \to \infty$. The determination of its precise shape goes beyond the scope of this work.

C. General Model

The aim of Sections III, VI-A and VI-B was not to *prove* (10), (21), (23), and (24). The aim was to only to *justify* them, as reasonable models of a much more complicated and intractable reality. A more diligent modeler could arrive at more accurate formulas, that incorporate effects of the physical layer that we ignored, such as the effects of fading, sleeping nodes, extraneous interferers, random node placement, or the effects of network coding, transmitter cooperation, MAC and routing protocols, etc. Instead of going through the messy details of more accurate models, and in order to widen the scope of our work, we will consider the following general model:

$$|\mathbf{T}(x,y)| \leq |\mathbf{T}(x,y)|_{\max} = F(x,y,d(x,y)) \tag{25}$$

where $F(\cdot)$ is a positive function, strictly increasing with respect to $d(x,y)$, but apart from that totally arbitrary. Equations (10), (21), (23), and (24) are special cases of (11). Note that (25) implicitly assumes that the maximum traffic will depend on the position (x,y). Therefore, it can model a nonhomogeneous medium. On the other hand, the maximum norm of $\mathbf{T}(x,y)$ will not depend on its direction, therefore the medium is implicitly assumed isotropic. An equivalent way of writing (25) is the following:

$$d(x,y) \geq G_T(x,y,|\mathbf{T}(x,y)|^2)$$

where $G_T(x,y,\cdot)$ is the inverse of $F^2(x,y,\cdot)$ with respect to the third argument, and so is also positive and strictly increasing.

Until now, we focused on the minimum node density required to transport the data. However, transporting the data is only one of the three tasks required of the nodes: the other two tasks is sensing the data, and also delivering the data to the sinks, once the data arrive at the physical location of the sinks. These two tasks are complementary: the first one is essentially the insertion of traffic in the wireless networks, and the second is the extraction of the traffic, once the destination has been reached. Researchers typically concentrate on the sensing and transport tasks and ignore the delivery task. However, by the explicit inclusion of the delivery task, we make our formulation a bit more general, and also more symmetric.

We assume that, in order to support the sensing (or delivery) at a location (x,y) where the information density function is $\rho(x,y)$, the node density $d(x,y)$ must satisfy the requirement

$$d(x,y) \geq G_{\mathrm{SD}}(x,y,\rho(x,y)).$$

Note that the information density function is a given, and not subject to any optimization. Therefore, the right hand side can be thought of as only a function of the location (x,y). The precise shape of $G_{\mathrm{SD}}(x,y,\rho(x,y))$ will depend on the sensing/delivering capabilities of the nodes, but it is intuitively clear that $G_{\mathrm{SD}}(x,y,\rho(x,y)) = 0$ when $\rho(x,y) = 0$.

Since the nodes must perform both the transport, and the sensing/delivery of the data, it follows that the density must satisfy the following:

$$\begin{aligned} d(x,y) &\geq f(G_T(x,y,|\mathbf{T}(x,y)|^2), G_{\mathrm{SD}}(x,y,\rho(x,y))) \\ &\triangleq G(x,y,|\mathbf{T}(x,y)|^2) \end{aligned} \tag{26}$$

where $f(\cdot)$ is some arbitrary function that captures the capability of the nodes to *jointly* perform sensing/delivery and transporting. Specifying this function goes beyond the scope of this work.

Let G' be the partial derivative of $G(x,y,|\mathbf{T}(x,y)|^2)$ with respect to the third argument $|\mathbf{T}(x,y)|^2$. For any reasonable physical layer model, the derivative cannot be negative, for any value of $|\mathbf{T}(x,y)|^2$, as this would imply that there is a situation in which we would need more nodes to carry less traffic. Furthermore, we make the mildly restricting assumption that G' must be strictly positive. Intuitively speaking, it is reasonable to assume that, the larger the traffic, the more nodes are needed to support it, even if locally most of the burden on the nodes comes from sensing/delivering the data. Mathematically speaking, if there is a range of $|\mathbf{T}(x,y)|^2$ for which G' is zero, then there will be some areas of the network in which the optimal traffic flow $\mathbf{T}(x,y)$ can not be uniquely defined, as we can increase the traffic with no extra cost. Technically speaking, at some point in the derivations of Section VII we will need to divide with G', and we need to ensure that it is always nonzero. The slightly more general case, where G' can be zero, can also be studied with our formulation but does not lead to a tidy solution, therefore we will ignore it.

VII. Optimal Node Placement

A. Problem Formulation

In this section we calculate the optimal node distribution $\hat{d}(x,y)$, that uses the minimum number of nodes and still is able to sense, transport, and deliver all the traffic, subject to the conditions (9) and (20), when the minimum required node density is given by (26). The problem can be written as

$$\begin{aligned} &\text{minimize} : I = \int_A G(x,y,|\mathbf{T}(x,y)|^2)\, dS \\ &\text{subject to} : \begin{cases} \nabla\cdot\mathbf{T}(x,y) = \rho(x,y), & (x,y)\in A, \\ [\mathbf{T}\cdot\hat{\mathbf{n}}(s)] = -\rho_b(s), & 0\le s\le \mathrm{L}. \end{cases} \end{aligned} \tag{27}$$

The minimization will be performed over all possible traffic flows $\mathbf{T}(x,y)$ that satisfy the constraints.

Note that we are not explicitly considering different propagation regions, as in Section IV-B, but implicitly, through the dependence of the $G(x,y,|\mathbf{T}(x,y)|^2)$ on the position (x,y). Also, we are not allowing any source/sink mobility, in the form of traffic regions. Our formulation will be extended to include source/sink mobility in Section VII-F.

B. Calculus of Variations

Let us write $\mathbf{T}(x,y) = u(x,y)\hat{\mathbf{x}} + v(x,y)\hat{\mathbf{y}}$, so that $|\mathbf{T}(x,y)|^2 = u^2(x,y) + v^2(x,y)$. Let us assume that the minimum of (27) is achieved by an optimal traffic flow $\hat{\mathbf{T}}(x,y) = \hat{u}(x,y)\hat{\mathbf{x}} + \hat{v}(x,y)\hat{\mathbf{y}}$. Therefore, the optimal value of the integral I of (27) is

$$\hat{I} = \int_A G(x,y,\hat{u}^2(x,y)+\hat{v}^2(x,y))\, dS.$$

Let the optimal functions $\hat{u}(x,y)$ and $\hat{v}(x,y)$ be perturbed by small variations δu and δv, such that the perturbed traffic $\hat{\mathbf{T}} + \delta\mathbf{T} = (\hat{u}+\delta u)\hat{\mathbf{x}} + (\hat{v}+\delta v)\hat{\mathbf{y}}$ continues to satisfy the constraints of (27). Note that the variations are not incrementally small *numbers*, but incrementally small *functions*. By standard calculus of variations arguments [21], [22], it follows that the first variation on the integral I should be zero

$$\delta I = \int_A \left(\frac{\partial G}{\partial u}\delta u + \frac{\partial G}{\partial v}\delta v\right) dS = 0 \tag{28}$$

where the partial derivatives are calculated at the points $(x,y,\hat{u}^2(x,y)+\hat{v}^2(x,y))$. This requirement is analogous to the requirement that the variation of a function $f(x)$, when we move an incremental distance dx from a stationary point $\hat{x}$ (which can be a maximum or a minimum), should be of the order $(dx)^2$, i.e., much smaller than the variation itself.

The perturbed functions $\hat{u}+\delta u$ and $\hat{v}+\delta v$ must satisfy the divergence constraint of (27), which for cartesian coordinates becomes:

$$\frac{\partial}{\partial x}(\hat{u}+\delta u) + \frac{\partial}{\partial y}(\hat{v}+\delta v) = \rho(x,y).$$

Since the optimal solution also satisfies the diversity constraint, we have that

$$\frac{\partial}{\partial x}\hat{u} + \frac{\partial}{\partial y}\hat{v} = \rho(x,y).$$

Combining the two equations, we arrive at

$$\frac{\partial}{\partial x}(\delta u) + \frac{\partial}{\partial y}(\delta v) = 0$$

which must hold at all points $(x,y)\in A$, and all small variations δu and δv for which the perturbed function satisfies the divergence constraint. Therefore, the following equation must hold, for any scalar, continuously differentiable function $\phi(x,y)$

$$\int_A \phi(x,y)\left(\frac{\partial}{\partial x}(\delta u) + \frac{\partial}{\partial y}(\delta v)\right) dS = 0. \tag{29}$$

Adding (28) and (29), we arrive at

$$\begin{aligned} &\int_A \left(\frac{\partial G}{\partial u}\delta u + \frac{\partial G}{\partial v}\delta v + \phi\frac{\partial}{\partial x}(\delta u) + \phi\frac{\partial}{\partial y}(\delta v)\right) dS \\ &= \int_A \left((\frac{\partial G}{\partial u} - \frac{\partial \phi}{\partial x})\delta u + (\frac{\partial G}{\partial v} - \frac{\partial \phi}{\partial y})\delta v \right) dS \\ &\quad + \int_A \left(\frac{\partial}{\partial x}(\phi\delta u) + \frac{\partial}{\partial y}(\phi\delta v)\right) dS = 0. \end{aligned} \tag{30}$$

However, by Green's Theorem [22], we have that:

$$\begin{aligned} \int_A \frac{\partial}{\partial x}(\phi\delta u)\, dS &= \oint_C \phi\delta u\cos\nu\, ds, \\ \int_A \frac{\partial}{\partial y}(\phi\delta v)\, dS &= \oint_C \phi\delta v\sin\nu\, ds \end{aligned}$$

where, as shown in Fig. 10, ν is the angle formed between the positive x-axis and the outward normal vector $\hat{\mathbf{n}}(s)$ at the point $C(s)$ of the boundary C of A, and s is the arc length along C. Adding the two together, we arrive at:

$$\int_A (\frac{\partial}{\partial x}(\phi\delta u) + \frac{\partial}{\partial y}(\phi\delta v))\, dS = \oint_C \phi(\delta u\cos\nu + \delta v\sin\nu)\, ds. \tag{31}$$

However, $(\delta u \cos\nu + \delta v \sin\nu)$ is the algebraic value of the projection of the perturbation traffic on the outward normal vector $\hat{\mathbf{n}}$ at point $C(s)$, i.e.

$$(\delta u \cos\nu + \delta v \sin\nu) = [(\delta \mathbf{T}) \cdot \hat{\mathbf{n}}](s).$$

On the other hand, both the optimal traffic flow, and its perturbation, must satisfy the boundary condition of (27):

$$\begin{aligned} \hat{\mathbf{T}} \cdot \hat{\mathbf{n}}(s) &= -\rho_b(s), \\ (\hat{\mathbf{T}} + \delta\mathbf{T} \cdot \hat{\mathbf{n}}(s)) &= -\rho_b(s). \end{aligned}$$

By subtracting the two, we have that $[(\delta \mathbf{T}) \cdot \hat{\mathbf{n}}](s) = 0$, everywhere on C. Therefore, the right hand side of (31) is 0, and using (30) if follows that:

$$\int_A (\frac{\partial G}{\partial u} - \frac{\partial \phi}{\partial x})\delta u \, dS + \int_A (\frac{\partial G}{\partial v} - \frac{\partial \phi}{\partial y})\delta v \, dS = 0. \quad (32)$$

The above equation must hold for any small perturbations δu and δv, and at the same time for any continuously differentiable $\phi(x,y)$. We now require that $\phi(x,y)$ satisfies $\frac{\partial G}{\partial v} - \frac{\partial \phi}{\partial y} = 0$, therefore (32) becomes

$$\int_A (\frac{\partial G}{\partial u} - \frac{\partial \phi}{\partial x})\delta u \, dS.$$

This equation must hold for *any arbitrary* variation δu (as long as we also define δv so that the constraints are satisfied). Clearly, this can only happen if the other factor of the integrand is identically zero, i.e., $(\frac{\partial G}{\partial u} - \frac{\partial \phi}{\partial x}) = 0$.

To conclude, if the integral of (27) is minimized by $\hat{u}$ and $\hat{v}$, then $\hat{u}$ and $\hat{v}$ must satisfy the system of equations:

$$\frac{\partial G}{\partial u} - \frac{\partial \phi}{\partial x} = 0, \; \frac{\partial G}{\partial v} - \frac{\partial \phi}{\partial y} = 0 \quad (33)$$

together with the constraints of (27). $\frac{\partial G}{\partial u}$ and $\frac{\partial G}{\partial v}$ are calculated at the points $(x, y, \hat{u}^2(x,y) + \hat{v}^2(x,y))$, and $\phi(x,y)$ is a continuously differentiable scalar function, which from now on we call the **potential**.

C. The Potential Equation

The system (33), together with the divergence constraint (9), consists of three partial differential equations, of which two are in general nonlinear. Note that we also have three unknowns, i.e., $\hat{u}$, $\hat{v}$, and ϕ. To simplify the notation, from now on we write u, v, and $\mathbf{T}$, instead of $\hat{u}$, $\hat{v}$, and $\hat{\mathbf{T}}$. Note that $|\mathbf{T}|^2 = u^2 + v^2$.

To simplify the problem, note that we have defined G' as the partial derivative of $G(x, y, u^2 + v^2)$ with respect to the third argument $u^2 + v^2$. It follows that $\frac{\partial G}{\partial u} = 2G'u$ and $\frac{\partial G}{\partial v} = 2G'v$, therefore the system (33) gives

$$2G'(x, y, u^2 + v^2)\mathbf{T}(x,y) = (\frac{\partial \phi}{\partial x}\hat{\mathbf{x}} + \frac{\partial \phi}{\partial y}\hat{\mathbf{y}}) \triangleq \nabla\phi. \quad (34)$$

Taking the square of the norm of both sides of (34), we arrive at

$$4(u^2 + v^2)G'^2(x, y, u^2 + v^2) = |\nabla\phi|^2. \quad (35)$$

In the above equation, u and v appear only through $u^2 + v^2$. We now make the mild assumption that (35) can be solved with respect to $u^2 + v^2$, and come to the form

$$u^2 + v^2 = H(x, y, |\nabla\phi|) \quad (36)$$

for some properly defined function $H(x, y, |\nabla\phi|)$. In this case, (34) can be written as

$$\mathbf{T}(x,y) = \frac{1}{2G'(x, y, H(x, y, |\nabla\phi|))}\nabla\phi. \quad (37)$$

Applying the divergence operator on both sides, and using (9), we arrive at:

$$\nabla \cdot \left(\frac{\nabla\phi}{2G'(x, y, H(x, y, |\nabla\phi|))} \right) = \rho. \quad (38)$$

The above equation, which we will refer to from now on as the **potential equation**, is a scalar partial differential equation (PDE) with a single unknown, the scalar potential function. Therefore, it is much easier to handle than the PDE system of (33) and (9).

Taking the inner product of each side of (37) with $\hat{\mathbf{n}}$, and using (20), gives the nonlinear boundary condition that accompanies (38):

$$\frac{1}{2G'(x, y, H(x, y, |\nabla\phi|))}[\nabla\phi \cdot \hat{\mathbf{n}}(s)] = -\rho_b(s). \quad (39)$$

In the special case where $\rho_b(s) = 0$, (39) becomes the Neumann condition $[\nabla\phi \cdot \hat{\mathbf{n}}(s)] = 0$, which is linear.

Using the PDE (38), together with the boundary condition (39), we can determine the potential function at all points in A. Then, we can determine $u^2 + v^2$ from (35), or equivalently (36), and finally $\mathbf{T}(x,y)$ can be determined using (37). Knowing $\mathbf{T}(x,y)$, we can find the optimal node distribution $\hat{d}(x,y)$ by using (26). The number of nodes needed will simply be its surface integral over A.

All the steps in this process are trivial, with the exception of solving the PDE (38), together with the boundary condition (39), which is highly nontrivial. Indeed, there is no general method for solving nonlinear PDEs analytically, and in the vast majority of settings the solution can only be calculated numerically.

Before moving to the study of a special case, we mention that our potential (38) can be thought of as a generalization of the linear scalar PDE (24) of [7]. As already mentioned, [7] studies a formulation related to our own. Due to its special form, (24) of [7] was shown using straightforward arguments, and without invoking calculus of variations.

D. Special Case: $F(x,y,d(x,y)) = Kd^{\beta}(x,y)$, $G_{\rm SD} \equiv 0$

As a special case, let us assume that $F(x,y,d(x,y)) = Kd^{\beta}(x,y)$. Therefore, (25) becomes

$$|\mathbf{T}(x,y)| \le |\mathbf{T}(x,y)|_{\max} = Kd^{\beta}(x,y). \tag{40}$$

As the maximum norm does not depend on the position (x,y), this model corresponds to a homogeneous medium. Note that (10) and (23) are special cases of (40). In addition, we set $G_{\rm SD} \equiv 0$, and we take $G(x,y,|\mathbf{T}(x,y)|^2)$ of (26) to be

$$G(x,y,|\mathbf{T}(x,y)|^2) = \frac{1}{K^{\frac{1}{\beta}}}\left[|\mathbf{T}(x,y)|^2\right]^{\frac{1}{2\beta}}.$$

Therefore,

$$G'(x,y,|\mathbf{T}(x,y)|^2) = \frac{1}{2\beta K^{\frac{1}{\beta}}}\left[|\mathbf{T}(x,y)|^2\right]^{\frac{1}{2\beta}-1}.$$

It then follows that

$$u^2+v^2 = H(x,y,|\nabla\phi|) \triangleq \left[K^{\frac{1}{\beta}}\beta|\nabla\phi|\right]^{\frac{2\beta}{1-\beta}}$$

and the potential (38) becomes

$$\Delta_p\phi \triangleq \nabla\cdot\left(|\nabla\phi|^{p-2}\nabla\phi\right) = f, \tag{41}$$

where $f = \left[K^{\frac{1}{\beta}}\beta\right]^{-\frac{\beta}{1-\beta}}\rho$, and $p = \frac{1}{1-\beta}$.

Equation (41) is known in the applied mathematics literature as p-Poisson's equation, and is the topic of much ongoing research [23], [24]. The operator Δ_p is called the p-Laplace operator. When $\rho = 0$, (41) becomes the homogeneous p-Laplace equation. Both equations appear often in variational problems such as our own. In the special case $p = 2$, which corresponds to $\beta = \frac{1}{2}$, the p-Poisson and p-Laplace equations become the well known, and much easier to solve, linear Poisson and Laplace equations.

The boundary condition (39) becomes

$$|\nabla\phi|^{p-2}[\nabla\phi\cdot\hat{\mathbf{n}}(s)] = -\left[K^{\frac{1}{\beta}}\beta\right]^{-\frac{\beta}{1-\beta}}\rho_b(s) \tag{42}$$

and is in general nonlinear, unless $p = 2$ or $\rho_b(x,y) = 0$.

E. Numerical Example

As a numerical example, let us consider a topology in which the area $A = \{|x| \le 1.5, |y| \le 1\}$. We place a distributed data source with $\rho = 100$ inside the rectangle $\{-0.5 \le x \le 0.5,\ 0.45 \le y \le 0.55\}$, and a symmetric distributed data sink with $\rho = -100$ inside the rectangle $\{-0.5 \le x \le 0.5, -0.55 \le y \le -0.45\}$. Regarding the physical layer, we assume that $G_{\rm SD} \equiv 0$ and $F(x,y,d(x,y)) = Kd^{\beta}(x,y)$ with $K = 1$, therefore to calculate the optimal traffic flow we need to solve (41) together with the boundary condition (42). We assume that there are no sources or sinks on the boundary, i.e., $\rho_s(s) = 0$, therefore (42) becomes the standard Neumann condition $[\nabla\phi\cdot\hat{\mathbf{n}}(s)] = 0$.

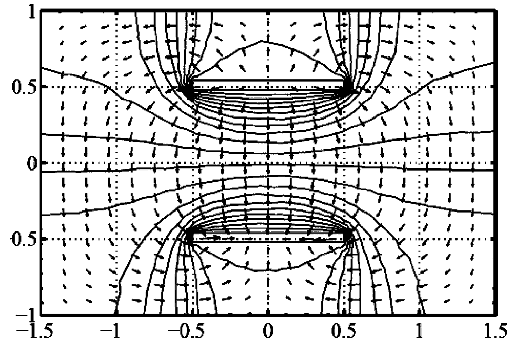


Fig. 12. Optimal traffic flow (denoted by arrows) and contours of constant potential, for the physical layer of Section VII-D, with $\beta = \frac{3}{8}$.

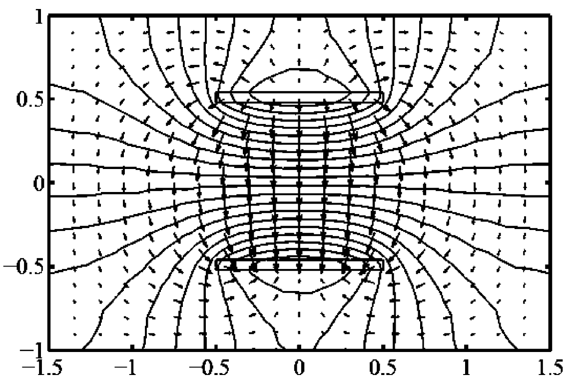


Fig. 13. Optimal traffic flow (denoted by arrows) and contours of constant potential, for the physical layer of Section VII-D, with $\beta = \frac{1}{2}$.

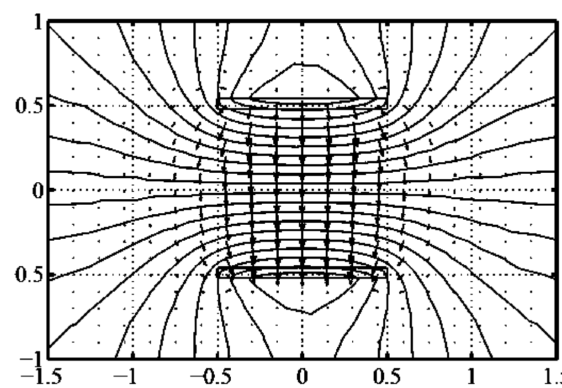


Fig. 14. Optimal traffic flow (denoted by arrows) and contours of constant potential, for the physical layer of Section VII-D, with $\beta = \frac{2}{3}$.

In Figs. 12–14 we plot the optimal traffic flow for the cases $\beta = \frac{3}{8}$ ($p = \frac{8}{5}$), $\beta = \frac{1}{2}$ ($p = 2$), and $\beta = \frac{2}{3}$ ($p = 3$) respectively. In Fig. 15 we plot the optimal traffic flow for the case where $\beta = \frac{2}{3}$ in the upper half $\{y \ge 0\}$ of A, and $\beta = \frac{3}{8}$ in the lower half $\{y < 0\}$ of A. In the figures, arrows denote the direction and size of $\mathbf{T}$ at the respective point, and contours denote loci of constant potential ϕ. Note that the loci of constant potential

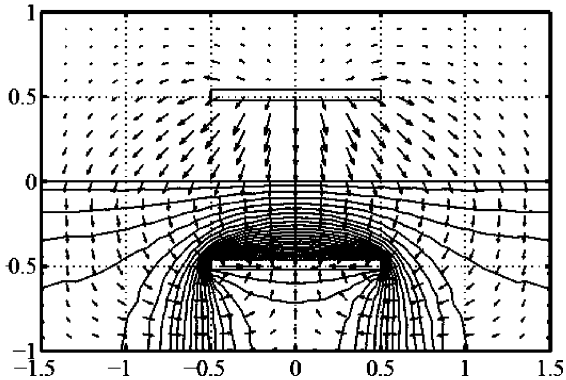


Fig. 15. Optimal traffic flow (denoted by arrows) and contours of constant potential, for the physical layer of Section VII-D, with $\beta = \frac{2}{3}$ on the upper half of the region, and $\beta = \frac{3}{8}$ on the lower half of the region.

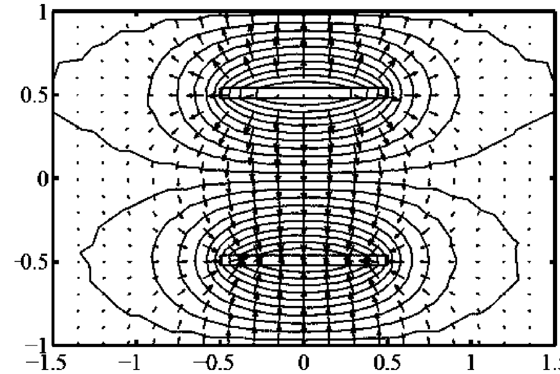


Fig. 16. Optimal traffic flow (denoted by arrows) and contours of constant potential, for the network of Fig. 13, when we substitute, on the upper and lower boundary, the Neumann condition (22) with the Dirichlet condition (23).

meet the boundaries always vertically, as we are requiring that $[\nabla\phi \cdot \hat{\mathbf{n}}](s) = 0$.

The plots are determined numerically, using the PDE toolbox of MATLAB. In the case $\beta = \frac{1}{2}$, (41) degenerates to the linear Poisson's equation of Electrostatics, and MATLAB calculates the solution using a standard Finite Element method. In the cases $\beta = \frac{2}{3}$ and $\beta = \frac{3}{8}$, however, (38) is nonlinear, and MATLAB solves it using an iterative method, based on damped Newton iterations with the Armijo-Goldstein line strategy.

As expected, in all three cases the optimal node placement induces a traffic flow that is heaviest in the region between the source and sink, however some of the traffic will travel along much longer routes. As already discussed, the intuition behind this result is that, if *all* packets use the short, direct routes, then the congestion in the central region will be so high as to require a very large number of nodes to be supported, more than the number of nodes needed if some of the packets take a longer, but much less congested, route.

Equally expected should be the fact that, the larger β becomes, the smaller the traffic that uses longer routes becomes. Indeed, as we see from (40), the larger the value of β, the easier it is for the network to support high levels of traffic, by increasing the node density, therefore short, congested routes do not come at a high cost. On the other hand, when β is small, as the traffic flow becomes larger, the number of nodes needed to support it increases very fast, so it is best for the routes to spread out as much as possible.

Perhaps not expected is how dramatically the traffic pattern changes with even moderate changes in the value of β. This suggests that network designers should carefully study the physical layer of their network, and deploy the nodes accordingly.

F. Dirichlet Boundary Conditions

Going back to the calculus of variations derivations of Section VII-B, note that we used the fact that all admissible traffic flows must satisfy (20), in order to show that the right hand side of (31) is zero. Then, using (30), (32) follows immediately, and after a few more derivations we arrive at (33). Note, however, that the right hand side of (31) will also be zero if we arbitrarily require that $\phi = 0$.

Formally, let us partition the boundary curve C of A into two parts, C_1 and C_2, and let us require that:

$$[\mathbf{T} \cdot \hat{\mathbf{n}}](s) = -\rho_b(s), \quad C(s) \in C_1 \tag{43}$$

$$\phi(s) = 0, \quad C(s) \in C_2. \tag{44}$$

Together, these assumptions guarantee that the right hand side of (31) is zero, and, replicating the derivations of Sections VII-B and VII-C, we have that $\mathbf{T}$ is given by (37), where the potential function ϕ can be determined by solving (38), together with the new mixed boundary conditions (43) and (44). Note that, in the special case when (40) holds with $\beta = \frac{1}{2}$, and (41) becomes Poisson's equation of Electrostatics, (44) means that C_2 is a grounded conductor.

The boundary condition (43) specifies, at the boundary section C_1, the value of the component of $\mathbf{T}$ that is vertical to the boundary. As we discussed in Section V, this component equals the rate with which data traffic enters (or leaves) C_1. On the other hand, the boundary condition (44) specifies that ϕ is constant on the boundary, hence its gradient $\nabla\phi$ must be vertical to the boundary. As the traffic flow $\mathbf{T}$ is parallel to $\nabla\phi$, it follows that the component of $\mathbf{T}$ that is parallel to the boundary is zero. In short, (43) places a constraint only on the component of $\mathbf{T}$ that is vertical to the boundary, and (44) places a constraint only on the component of $\mathbf{T}$ that is parallel to the boundary.

The physical interpretation of (44) is that there is a part of the boundary, i.e., C_2, through which packets are free to come and go, so long as they hit the boundary vertically. This constraint means that the packets should approach C_2 as fast as possible, without more nodes than necessary being spent on any transport parallel to the boundary. Therefore, the boundary C_2 models a fixed infrastructure network of very large capacity, that can move sources and sinks around in order to help the wireless network reduce its resources as much as possible.

In Fig. 16 we calculate the optimal traffic flow for the network of Fig. 13, where we substitute on the lower and upper boundaries the Neumann condition (43) with the Dirichlet condition (44). Clearly, some of the created packets are now diverted to the upper boundary, which acts as a sink, while some of the packets arriving at the distributed sink have actually started from the lower boundary, which acts as a source. The two boundaries may be thought of as belonging to the same infrastructure network, that can move packets around with a negligible cost.

We now summarize the findings of this section in the form of a theorem:

Theorem 2: Let A be a compact region, on the inside of which sources and sinks are placed, as described by the information density function $\rho(x,y)$. Let $\partial A = C_1 \cup C_2$, with $C_1 \cap C_2 = \emptyset$. Along C_1, information is entering the network with a prescribed net rate of $\rho_b(s)$. Along C_2 the net rate which which information is entering the network is subject to optimization. Let the minimum node density needed to support a traffic flow $|\mathbf{T}(x,y)|$ and the sensing/delivery of information at location (x,y) be equal to $G(x,y,|\mathbf{T}(x,y)|^2)$. Let G' be the partial derivative of G with respect to the third argument, and let $H(\cdot)$ be the function defined by writing $4aG'^2(x,y,a) = b^2$ as $a = H(x,y,b)$.

The optimal placement of nodes, that minimizes the number of nodes needed to transfer the data, induces the traffic flow

$$\mathbf{T}(x,y) = \frac{1}{2G'(x,y,H(x,y,|\nabla\phi|))}\nabla\phi$$

where the function ϕ satisfies the scalar nonlinear differential equation

$$\nabla \cdot \left(\frac{\nabla\phi}{2G'(x,y,H(x,y,|\nabla\phi|))} \right) = \rho$$

together with the boundary conditions

$$\begin{aligned} [\mathbf{T}\cdot\hat{\mathbf{n}}](s) &= -\rho_b(s), \quad C(s) \in C_1 \\ \phi(s) &= 0, \qquad\quad C(s) \in C_2. \end{aligned}$$

VIII. Extensions

A. Networks With Multiple Traffic Types

Until now, it was assumed that there is only one type of traffic in the network. Therefore, if more than one traffic stream flows through a location in the network, we are allowed to perform vector addition, and abstract the flow of traffic at that location by a single vector, the traffic flow function at that point.

If, however, there are $m > 1$ different types of traffic, each of them will have to be associated with its own traffic flow function $\mathbf{T}^i$, and its own information density function ρ^i, and boundary information density function ρ_b^i, for which we will have

$$\nabla \cdot \mathbf{T}^i = \rho_i,\ i = 1, \ldots, m \tag{45}$$

$$\rho_b^i = -\mathbf{T}^i \cdot \hat{\mathbf{n}},\ i = 1, \ldots, m. \tag{46}$$

A point in the network through which different types of traffic cross, will have to divide its resources to support all traffic types, therefore (26) will have to be extended to:

$$d(x,y) \geq G(x,y,|\mathbf{T}^1(x,y)|^2, \ldots, |\mathbf{T}^m(x,y)|^2).$$

Our new problem is the minimization of

$$\int G(x,y,|\mathbf{T}^1(x,y)|^2, \ldots, |\mathbf{T}^m(x,y)|^2)\, dS$$

subject to (45) and (46).

This optimization problem in principle can be studied using calculus of variations techniques, following the methodology of Section VII where now the optimization must be over $2m$ functions, i.e., the components u^i and v^i for each of the m traffic functions $\mathbf{T}^i = u^i\hat{\mathbf{x}} + v^i\hat{\mathbf{y}}$. Such an investigation goes beyond the scope of this work. We note, however, that this problem was investigated in [6].

B. Alternative Transport Optimization Formulations

Our formulation could be applied to study a variety of problems, other than the problem of minimizing the number of nodes in the network. Indeed, from a mathematical point of view, our aim was to solve the optimization problem (27), and the quantities appearing in (27) also admit alternative interpretations.

For example, we could have defined a slightly different setting in which the density of nodes is fixed, and not subject to optimization, and the cost $G(x,y,|\mathbf{T}(x,y)|^2)$ could be the power needed to support a level of traffic intensity $|\mathbf{T}(x,y)|$. This problem was first considered, with a similar but more restricted formulation, in [7]. Also, we could take the density of the nodes to be fixed, and the cost $G(x,y,|\mathbf{T}(x,y)|^2)$ to be the delay incured locally under a traffic intensity $|\mathbf{T}(x,y)|$.

Alternatively, we could use this formulation to determine if there is a traffic flow that can transport all the created traffic in a wireless sensor network in which the placement of resources (nodes, energy, etc.) is fixed and not subject to any optimization. In this case, the cost function $G(x,y,|\mathbf{T}(x,y)|^2)$ would be very small when $|\mathbf{T}(x,y)|^2)$ is below some threshold, which reflects that the resources at the location (x,y) can handle the traffic, and would grow very steeply as the threshold is reached, to reflect the fact that the traffic must not exceed the threshold.

Even more generally, our formulation could be viewed as an abstract problem in optimal transportation. In the most abstract setting, functions ρ and ρ_b specify the rate with which some indeterminate commodity is created or absorbed inside or on the boundary of an area A, $\mathbf{T}(x,y)$ specifies the rate and direction with which the commodity flows through point (x,y), and $G(x,y,|\mathbf{T}(x,y)|^2)$ is the cost of transporting this commodity through a point (x,y). The optimization problem is how to minimize the total cost needed for the transport of the commodity.

In contrast with other problems in optimal transportation, for example the Monge-Kantorovich formulation [25], [26], the transportation cost is not only a function of the distance covered by the transported commodity, but in general also depends on the competition among the transported commodities for the transportation resources along the way. This aspect of our formulation was forced by the nature of the wireless channel, but may be relevant in other transportation settings as well.

IX. Conclusion

We consider a setting in which a spatially distributed set of sources is creating data for a spatially distributed set of sinks. Our problem is how to optimally deploy a network of wireless nodes, so that all the data can be sensed at the sources, transported to the physical locations of the sinks, and delivered to the sinks, using the minimum number of nodes.

We make the critical simplifying assumption that the network is massively dense [13], i.e., there are so many sources, sinks, and nodes, that it is best to describe the network in terms of macroscopic parameters, such as their spatial distribution, rather than in terms of microscopic parameters, such as their individual placements.

We first focus on a particular physical layer model that is characterized by the following assumptions: *i)* the wireless nodes must only transport the data from the location of the sources to the location of the sinks, and do not need to sense the data at the sources, or deliver them at the sinks once the data arrive at their physical locations, and *ii)* the nodes have limited bandwidth available to them, but they use it optimally to locally achieve the network capacity. In this setting, the optimal distribution of nodes induces a traffic flow that resembles the electric displacement that will be created if we substitute the sources and sinks with positive and negative charges respectively. The analogy between the two settings is very tight, and many features of Electrostatics have a direct interpretation in wireless sensor networks.

Under a more general physical layer model, the optimal traffic flow no longer corresponds to an electrostatic field. Nonetheless, we can derive its form in terms of a scalar, nonlinear partial differential equation, by use of calculus of variations techniques.

Our work finds the most efficient deployment of networks, that strikes the most favorable balance between having short routes and keeping the levels of congestion down. Our numerical examples show that the optimal placement of nodes, and the traffic flow it induces, can heavily depend on the precise capabilities of the physical layer. Therefore, network designers need to carefully study the physical layer of the network to be created, before deciding on how the nodes should be deployed.

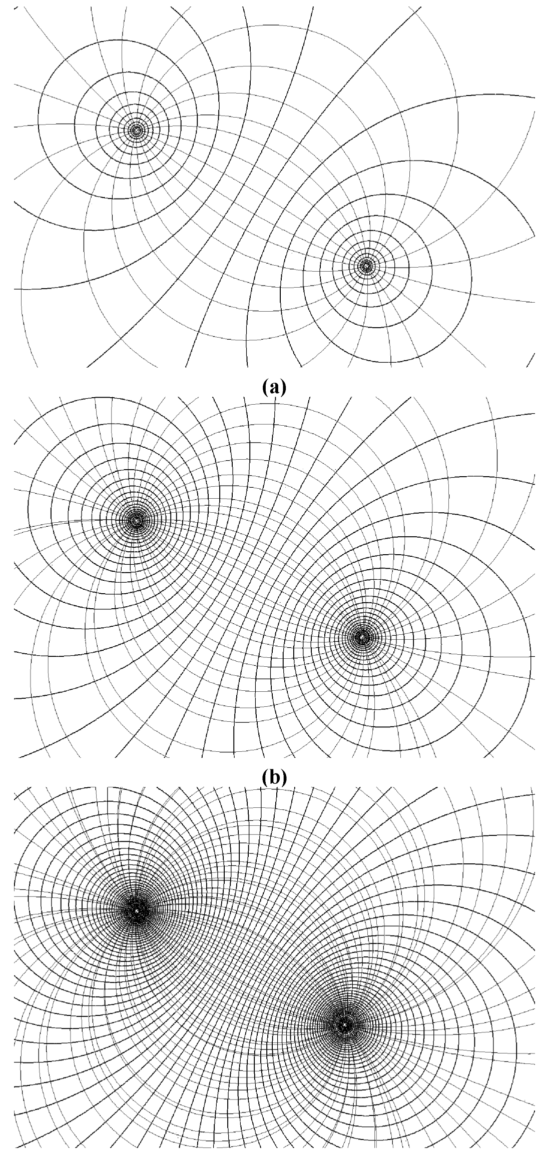
As discussed in Section VIII-B, the optimization problem we are studying readily admits alternative interpretations, therefore our work may also be of use in settings where we are interested in minimizing the energy per packet, or the delay per packet, and so on. Our work may also be viewed as an abstract problem in optimal transportation, and so may be of interest outside the field of wireless networks.

As in [11] and [12], our results only formally hold as the number of nodes goes to infinity. However, they are also relevant in networks with a finite (but relatively large) number of nodes. To benefit from our formulation, network designers should perform the following steps.

1) Approximate the distributed sources and sinks by an information density function.
2) Decide on the shape of the function $G(x, y, |\mathbf{T}(x, y)|^2)$ of (26), taking into account the various properties of the physical, MAC, routing, and sensing layers, etc.
3) Calculate the optimal spatial density of nodes, the total number of nodes N, and the induced traffic flow using our formulation.
4) Place the N nodes so that they form a node density function resembling as close as possible the optimal node density function. The accuracy of this approximation will depend on how large N is.

In the special case where (10) holds and the optimal traffic flow resembles an electrostatic field, there is a very intuitive and straightforward way of performing the last step, using electric field lines and constant potential loci. In this case, the optimal node density is proportional to the square of the Electrostatic field $|\mathbf{E}|^2 = |\mathbf{E}| \times |\nabla U|$, as specified by (10). Therefore, we can plot n electric field lines and m loci of constant potential, and then place one node in each intersection point between an electric field line and a constant potential locus. By construction, the spatial density of intersection points is proportional to $|\mathbf{E}| \times |\nabla U|$, as required. In this manner, a number of nodes that does not exceed $n \times m$ will be deployed. The parameters n and m must be chosen so that the resulting number of nodes is close to N, and the network looks *locally* like a square grid.

As an example of this process, in Fig. 17 we have plotted electric field lines and constant potential loci that are created by a single positive charge of magnitude q and a single negative charge of magnitude $-q$, and for the three cases $q = 1$ Cb, $q = 2$ Cb, $q = 4$ Cb. According to the discussion of the previous paragraph, these plots have an alternative interpretation:



(a)

(b)

(c)

Fig. 17. Electric field lines and constant potential loci created by a single positive charge of magnitude q and a single negative charge of magnitude $-q$. (a) $q = 1$ Cb. (b) $q = 2$ Cb. (c) $q = 4$ Cb.

The intersections of lines is where we need to place nodes so that all the information created at a source of q bps is transported to a sink of q bps, and the optimal number of nodes is needed. Note that, according to (10), in order to double the traffic, we need to increase the number of nodes by a factor of 4. In order to find places for four times more nodes, and still have the network locally resemble a square grid, we need to double the number of electric field lines and double the number of constant potential loci. Also note that, as the number of nodes increases, the network locally resembles more and more accurately a square grid. Each electric field line represents a route used by packets, and the number of hops needed for the packets to go from the source to the destination equals the number of times the line intersects the loci of constant potential.

The previous discussion is intuitive, but clearly imprecise, and so points to an issue of fundamental importance for this work, namely the calculation of the rate with which the performance of discrete networks converges to the performance of their corresponding massively dense networks as the number of nodes increases. This rate will depend, among other factors, on the physical layer parameters, the particular MAC and routing protocols, the use or not of network coding, the use or not of cooperation among the transmitters, how the nodes are placed, for example deterministically, or randomly with a prescribed spatial density, etc. To calculate the rate, we will need to be able to determine precisely the performance of wireless sensor networks of arbitrary size and topology. Related works have focused on calculating only scaling laws, and then on homogeneous topologies, which are more amenable to analysis than our own nonhomogeneous topologies [11], [12]. Therefore, calculating this rate of convergence for any setting of interest seems to be a major undertaking and, despite its obvious significance, is out of the scope of this work.

Acknowledgment

The authors would like to thank G. A. Gupta (IIT New Delhi) for suggesting the outline of the derivations in Section VII, C. F. Mecklenbräuker and J. Sayir (ftw.) for many useful discussions and suggestions, Claudio Casetti (PoliTo) for proposing the method of placing nodes at the intersections of electric field lines and constant potential loci discussed in the conclusions, Prof. M. Shayman (University of Maryland) for clarifying the connections between this work and [6], [7], and the anonymous reviewers for their numerous insightful remarks and suggestions.

References

[1] S. Toumpis and L. Tassiulas, "Packetostatics: Deployment of massively dense sensor networks as an electrostatics problem," in *Proc. IEEE INFOCOM*, Miami, FL, Mar. 2005.

[2] S. Toumpis and G. A. Gupta, "Optimal Placement of Nodes in Wireless Sensor Networks under a General Physical Layer Model," in *Proc. IEEE SECON*, Santa Clara, CA, Sep. 2005.

[3] I. F. Akyildiz, W. Su, Y. Sankarasubramanian, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[4] J. Barros and S. D. Servetto, "The sensor reachback problem," *IEEE Trans. Inf. Theory*, submitted for publication.

[5] S. D. Servetto, "Distributed signal processing algorithms for the sensor broadcast problem," in *Proc. Conf. Inf. Sci. Syst.*, Baltimore, MD, Mar. 2003.

[6] M. Kalantari and M. Shayman, "Routing in wireless ad hoc networks by analogy to electrostatic theory," in *Proc. IEEE ICC*, Paris, France, Jun. 2004.

[7] ——, "Energy efficient routing in wireless sensor networks," in *Proc. Conf. Inf. Sci. Syst.*, NJ, Mar. 2004.

[8] R. Verdone, "An energy-efficient communication protocol for a network of uniformly distributed sensors polled by a wireless transceiver," in *Proc. IEEE ICC*, Paris, France, Jun. 2004.

[9] J.-H. Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," in *Proc. IEEE INFOCOM*, Tel Aviv, Israel, Mar. 2000, vol. 1, pp. 22–31.

[10] A. Ephremides, "Energy concerns in wireless networks," *IEEE Wireless Commun. Mag.*, vol. 9, no. 4, pp. 48–59, Aug. 2002.

[11] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, pp. 388–404, Mar. 2000.

[12] M. Franceschetti, O. Dousse, D. N. C. Tse, and P. Thiran, "Closing the gap in the capacity of random wireless networks," in *Proc. IEEE ISIT*, Chicago, IL, Jun.-Jul. 2004.

[13] P. Jacquet, "Geometry of information propagation in massively dense ad hoc networks," in *Proc. ACM MobiHOC*, Roppongi Hills, Japan, May 2004, pp. 157–162.

[14] J. E. Marsden and M. J. Hoffman, *Elementary Classical Analysis*, 2nd ed. New York, NY: Freeman, 1993.

[15] H. M. Schey, *Div, Grad, Curl, and All That: An Informal Text on Vector Calculus*, 3rd ed. New York: Norton, 1996.

[16] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. New York: Wiley, 1998.

[17] N. Bigdely, Equipotential Surfaces and E Lines [Online]. Available: www.electrostatics3d.com

[18] W. K. H. Panofsky and M. N. Phillips, *Classical Electricity and Magnetism*. Reading, MA: Addison-Wesley, 1955.

[19] O. Dousse, M. Franceschetti, and P. Thiran, "Information theoretic bounds on the throughput scaling of wireless relay networks," in *Proc. IEEE INFOCOM*, Miami, FL, Mar. 2005.

[20] R. Negi and A. Rajeswaran, "Capacity of power constrained ad-hoc networks," in *Proc. IEEE INFOCOM*, Hong Kong, China, Mar.-Apr. 2004.

[21] F. Y. M. Wan, *Introduction to the Calculus of Variations and its Applications*. London, U.K.: Chapman & Hall Mathematics, 1995.

[22] F. B. Hildebrand, Ed., *Methods of Applied Mathematics*. New York: Dover, 1992.

[23] *The p-Laplace Equation, The Infinity-Laplace Equation, And Related Topics*, Conference Proceedings, Linköping, Sweden, Oct. 2004.

[24] L. C. Evans, *Partial Differential Equations*. New York: American Mathematical Society, 1998.

[25] ——, "Partial differential equations and Monge-Kantorovich mass transfer," *Curr. Dev. Math.*, 1997.

[26] G. Monge, "Memoire sur la theorie des déblais at des remblais," *Histoire de L'Acad. des Sciences de Paris*, 1781.