# Design Optimization of Multi-Sink Sensor Networks by Analogy to Electrostatic Theory

Mehdi Kalantari and Mark Shayman
Department of Electrical and Computer Engineering, University of Maryland
{mehkalan, shayman}@eng.umd.edu

*Abstract*— **In this work we introduce a new mathematical tool for optimization of routes, and topology design in wireless sensor networks. We introduce a vector field formulation that models communication in the network, and routing is performed in the direction of this vector field at every location of the network. The magnitude of the vector field at every location represents the density of amount of data that is being transited through that location. We define the total communication cost in the network as the integral of a quadratic form of the vector field over the network area. Our mathematical machinery is based on partial differential equations analogous to the Maxwell's equations in electrostatic theory. We use our vector field model to solve the optimization problem for the case in which there are multiple destinations (sinks) in the network. In order to optimally determine the destination for each sensor, we partition the network into areas, each corresponding to one of the destinations. We define a vector field, which is conservative, and hence it can be written as the gradient of a scalar function (also known as a potential function). Then we show that in the optimal assignment of the communication load of the network to the destinations, the value of that potential function should be equal at the locations of all the destinations. Also, we show that such an optimal partitioning of the network load among the destination is unique, and we give iterations to find the optimal solution.**

## I. Introduction

Wireless sensor networks have been studied extensively in recent years. Such networks are made up of several hundred to several thousand sensors distributed in a geographical area. There are many different applications for such networks including military, environment monitoring, agriculture and home applications. Sensors are very simple identical electronic devices equipped with a processor and small memory, a transmitter, and a receiver. Generally, sensors use radio frequency channels for the purpose of communication.

Most of the time it is desired to collect the data acquired by all sensors in a specific destination for processing, archiving and other purposes. This station is a data sink, and it has enough processing power, storage space, and capability of communicating with the sensors. For the purpose of communication to this destination, the sensors relay the packets of each other in a multi-hop way.

In order to solve the routing problem in a wireless network, we introduce an optimization problem based on a vector field, which represents the communication load at every place of the network, and it is conservative under certain assumptions. By using this conservative vector field, we define a tool for routing by writing the conservative vector field as the gradient of a scalar potential function. Routing of packets to each destination is done based on the value of this potential function on each node and its value on the neighboring nodes. In [1] and [2] we used a above vector field model to solve the routing problem in a *single-sink* scenario in which many sources are intended to send their data to a single destination. In this paper, we extend the approach to the case in which we have several destinations (sinks) in the network. The optimization problem in single destination scenario is finding efficient routes to that destination; however, in the multiple destination scenario, the problem is finding efficient routes from sensors to the destinations, as well as finding the optimal partitioning of the network area into regions corresponding to the destinations. Each destination collects all messages generated in its region.

We introduce a mathematical machinery based on partial differential equations very similar to the Maxwell's equations in electrostatic theory. In our formulation, sources of information are similar to the positive charges in electrostatics, the destinations are similar to negative charges, and the network is similar to a non-homogeneous dielectric media with variable dielectric constant (or permittivity coefficient).

One application of our methodology is to provide an optimal network partitioning algorithm into several regions. Each region has a destination inside it, which the wireless nodes inside that region communicate with. Given the geographic information of the communication demand in an area of the network, we give an optimal method for assigning the load of the network to the different destinations. We mathematically prove that in order to minimize a quadratic cost function of communication load, the value of potential function should be equal in the locations of all destinations.

The routing problem in sensor networks has been studied by many researchers. Sequential Assignment Routing (SAR) is proposed in [3], and it takes into account the energy constraints by making a tree rooted in the destination. Minimum Cost Forwarding Algorithm for Large Sensor Networks is proposed in [4]. Similar routing schemes can be found in [5] and [6]. Some approaches have been offered in which partial information about the node location is assumed to be known [7],[8],[9]. The routing approach proposed in [8] assumes some of the nodes in the network can serve as location proxies, and these nodes collect and forward the packets of their neighboring nodes that do not have their location information.

The idea of using a routing methodology in sensor networks inspired by the way the electrostatic field propagates in a

This work was partially supported by AFOSR under grant F496200210217 and and NSF under grant CNS-0519554.

dielectric medium was first introduced in our works [2], and [1]. Our work was followed by related work by Toumpis and Tassiulas [10], where they have shown that the approach minimizes the number of sensors required to handle the total communication burden of the network. A different approach that uses multipath routing based on electrostatic force is described in [11]. Surveys on the routing schemes of sensor networks have been given in [12] and [13]

The main contributions of our work are:

- Formulating of the communication load in a sensor network as a vector field.
- Introducing an optimization problem for minimizing the communication cost in a sensor network. This cost function leads to the solution in form of partial differential equations analogous to Maxwell's equations in electrostatic theory.
- Optimally solving the problem of assigning the network load among destinations.

The remainder of this paper is organized as follows: in Section II we give the background and introduce the basic formulation of communication load as a vector field for the case with one destination. In Section III we generalize the approach to the case with multiple destinations. Some clarifying examples and experiments are given in Section IV, and paper concludes in Section V.

## II. Background: Vector Field Formulation for the Single Destination Case

In this section, we show the main framework of routing problem in sensor networks as an optimization problem. Consider a network of $N$ wireless sensors that can communicate with each other through radio links. The nodes are densely located in a region $A$ in the plane, and they are intended to collect information about the events in the area of the network. Each sensor is responsible for events happening in its neighborhood. All messages are desired to be collected in a destination node (access point), and for now, we assume there is only one node of this type in the network. Later, we generalize our approach to the case in which we have multiple destinations. When an event is generated at some place in the network, the closest sensor to location of the event generates a message. All messages should be sent to the destination, which is assumed to have enough storage, energy and processing power. Furthermore, we make the following assumptions:

**A1**: The messages in the geographical area of the network happen with a known spatial density rate denoted by $r(x, y) \geq 0$ for the place $(x, y)$. $r(x, y)$ states how many messages are generated per unit of area per unit of time. We call this quantity the *load density*, which means that for the area $a \subseteq A$ the rate of messages generated inside $a$ is:

$$w(a) = \int_a r(x, y)\, dxdy \tag{1}$$

in which integration is over area $a$. Note that $r(x, y)$ does not include messages passing through $a$ on their path to the destination. The units of $r(x, y)$ are messages per square meter per second.

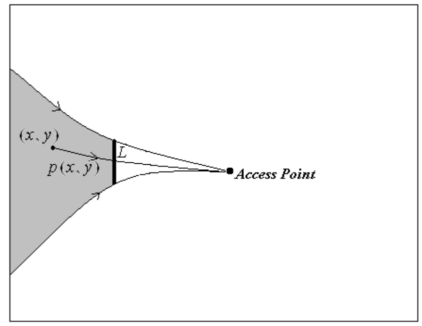


Fig. 1. The illustration of *Upstream Area* of a given set $L$. The Shaded area in this figure shows the upstream area of the rectangle shown as $L$

**A2**: For the purpose of routing, a single-valued direction is kept for every point $(x, y)$ of the network. At place $(x, y)$, this quantity represents the direction in which a message travels along the forwarding sensors from its source sensor to the destination; the message may be generated at place $(x, y)$, or received from an upstream place of the network. For place $(x, y)$ of the network, we denote this direction by the *unit* vector $\hat{\theta}(x, y)$. The value of $\hat{\theta}(x, y)$ is not defined for the location of the destination.

The above assumption implies that for every location of the network there exists a single path between that location and the destination. Mathematically, we define a path as a directed curved line starting at $(x, y)$ and ending at the destination. Let $p(x, y)$ denotes the set of points in $A$ that belong to this path. Note that based on assumption **A2** the paths have the *suffix property*, which means if the path of a point $(x, y)$ passes through a point $(x', y')$, then the path of $(x', y')$ coincides with the reminder (suffix) of the path from $(x, y)$.

It should be noted that the chosen paths are not constrained by the location of intermediate sensors. Instead, the paths are *abstract* paths in the plane that represent desired paths for the transmission of messages. For communication to occur, we need to define the routes in terms of the paths. Another important note is connection between the paths and routes. We define a *route* as a sequence of sensor nodes starting at a sensor and ending at the destination In order to find a route from a sensor to the destination, we approximate the path starting at the location of the sensor by a piecewise linear path with sensors in its vertices. This approximation is justified if we assume sensors are distributed dense enough in the area of the network. However, if sensors are not dense enough in certain areas, such areas will be considered as places with limited resources and we can adapt the framework by penalizing the traversal of such areas. (The weighting function $K(x, y)$ introduced later would take large values for these areas.)

Given a set of abstract paths for each location $(x, y)$, for a connected set $L$ in $A$, we define the notion of *Upstream Area*,

$\alpha(L)$ as:

$$\alpha(L) = \{(x, y) \in A \text{ s.t. } \exists\ (x', y') \in L \text{ and } (x', y') \in p(x, y)\} \quad (2)$$

The concept of upstream area is illustrated in Fig. 1. Next, we define a vector field on $A$ which we refer to as the *load density* vector field and denote by $\vec{D}$. Given a point $(x_1, y_1) \in A$, we choose a small line segment $L$ containing $(x_1, y_1)$, and perpendicular to $\hat{\theta}(x_1, y_1)$. The magnitude of the load density vector field is the density of messages passing through $L$, which can be written as the ratio of all load generated in the upstream area of $L$ divided by the length of $L$:

$$\vec{D}(x_1, y_1) = \lim_{|L| \to 0} \frac{\hat{\theta}(x_1, y_1)}{|L|} \int_{\alpha(L)} r(x, y)\, dxdy \quad (3)$$

It is important to note that $|\vec{D}(x_1, y_1)|$ is the sum of the communication load of all the paths that pass through $L$. So $|\vec{D}(x_1, y_1)|$ represents the actual amount of communication activity at point $(x_1, y_1)$. The direction of $\vec{D}(x_1, y_1)$ is $\hat{\theta}(x_1, y_1)$, which is the single valued direction on which the traffic at point $(x_1, y_1)$ is forwarded according to assumption **A2**.

Finally, we define a scalar function $\rho(x, y)$ on the network. This function represents the spatial density of rate at which the messages are generated in the network. This quantity is a function of location, and obviously $\rho(x, y) = r(x, y)$ for $(x, y) \neq (x_0, y_0)$, in which $(x_0, y_0)$ is the location of the destination. However, since all messages end at the destination, the density of the rate has a Dirac delta form at the location of the destination. Hence:

$$\rho(x, y) = r(x, y) - w_0 \delta(x - x_0)\delta(y - y_0) \quad (4)$$

in which $w_0$, the weight of delta function, is the integral of $r(x, y)$ in the network area (i.e., $w_0 = \int_A r(x, y) dxdy$). This definition implies that $\int_A \rho(x, y) dxdy = 0$.

The definition of $\vec{D}(x, y)$ given by equation (3) satisfies:

$$\oint_C \vec{D} \cdot \vec{dn} = \int_{S(C)} \rho(x, y) dxdy \quad (5)$$

in which the integral is over a closed contour $C$, $\vec{dn}$ is a differential vector normal to that contour at each point of its boundary and pointing to outside of the counter, dot represents the inner product of vectors in two-dimensional space, and $S(C)$ is the area surrounded by the closed contour $C$. Equation (5) is analogous to Gauss' law in the electrostatic theory, and it has a very simple explanation: the rate at which the messages exit a contour is the net amount of the sources inside that contour.

With the above definition of $\rho(x, y)$ and $\vec{D}(x, y)$ equation (5) can be expressed in partial differential equation form:

$$\vec{\nabla} \cdot \vec{D}(x, y) = \rho(x, y) \quad (6)$$

where $\vec{\nabla}$ is defined as:

$$\vec{\nabla} = \frac{\partial}{\partial x}\hat{i} + \frac{\partial}{\partial y}\hat{j} \quad (7)$$

in which $x$ and $y$ are the variables in the Cartesian coordinate frame, and $\hat{i}$ and $\hat{j}$ represent the unit vectors along $x$ and $y$ axes respectively ($\vec{\nabla}\cdot$ is also known as the divergence operator).

Depending on how we select the set of paths, the value of $\vec{D}(x, y)$ is different, but independent of path selection method, $\vec{D}(x, y)$ satisfies the following equations:

$$\begin{cases} \vec{\nabla} \cdot \vec{D}(x, y) = \rho(x, y) \\ D_n(x, y) = 0 \quad for\ (x, y) \in Boundary\ of\ A \end{cases} \quad (8)$$

in which $A$ denotes the geographical set that contains the network and $D_n(x, y)$ denotes the normal component of $\vec{D}(x, y)$ on the boundary of $A$. The first equation in (8) is the natural limitation imposed by the fact that all the traffic generated at the network should be delivered to the destination. The second equation comes from the fact that no load is desired to exit the geographical area of the network or enter into it through the boundary. It is important to notice that equations (8) do not give $\vec{D}(x, y)$ uniquely.

Conversely, if we have a $\vec{D}(x, y)$ that satisfies equations (8), we can find the path that can be used to send the traffic of each point $(x, y)$ to the destination. In order to define the routes based on the values of $\vec{D}(x, y)$, we need to define the concept of *load flow lines*. These lines are similar to the electric flux lines in electrostatic theory [14] [15][16]. The load flow lines are a family of curved lines that are always tangent to the direction of $\vec{D}(x, y)$ and their orientation is the same as the orientation of $\vec{D}(x, y)$. The load flow lines always end at the destination; this fact is because the value of divergence in equations (8) is nonnegative at every place of the network, except it is negative at the destination.

Based on the definition of the load flow lines, the path corresponding to each point $(x, y)$ can be easily found as the flux lines of the load density vector field, and finally $\hat{\theta}(x, y)$, can be written as:

$$\hat{\theta}(x, y) = \vec{D}(x, y) / |\vec{D}(x, y)|. \quad (9)$$

Since (8) does not specify $\vec{D}(x, y)$ uniquely, the remaining issue is to decide what additional condition(s) to place on $\vec{D}(x, y)$ so the resulting vector field generates a desirable set of routes. The intuition we follow is that by making $\vec{D}$ as uniform as possible, we obtain routes that cause the traffic to be highly dispersed throughout the network. In turn, this decreases both node congestion and collisions and leads to high throughput.

Spreading the network communication load can be formulated as minimizing the following quadratic cost function:

$$J(\vec{D}) = \int_A |\vec{D}(x, y)|^2 dxdy \quad (10)$$

The quadratic form of the cost function in equation (10) causes the load to be distributed as uniformly as possible. It prevents having high loads somewhere in the network while the resources are underutilized somewhere else. One interesting fact about this cost function is that it is similar to the definition of energy in electrostatic theory. The above optimization problem can be summarized as:

$$\begin{aligned}&\text{Minimize } J(\vec{D}) = \int_A |\vec{D}(x,y)|^2 dxdy\\&\text{Subject to:}\\&\vec{\nabla}\cdot\vec{D}(x,y) = \rho(x,y)\\&D_n(x,y) = 0,\ \forall(x,y)\in \text{Boundary of } A\end{aligned}$$

The following result is proved in [1], and it gives the key to finding the solution of the optimization problem of (11):

**Theorem 1:** *If $\vec{D^*}(x,y)$ denotes the optimal solution of equation (11), then it satisfies:*

$$\vec{\nabla}\times\vec{D^*}(x,y) = 0 \tag{11}$$

In the above equation $\vec{\nabla}\times$ is the two dimensional curl operator, and it is defined in the following way for a vector field $\vec{F} = [F_x\ F_y]$:

$$\vec{\nabla}\times\vec{F}(x,y) = (-\frac{\partial F_x(x,y)}{\partial y} + \frac{\partial F_y(x,y)}{\partial x})\hat{k} \tag{12}$$

in which $\hat{k}$ is a unit vector perpendicular to $\hat{i}$ and $\hat{j}$. More precisely, $\hat{k} = \hat{i}\times\hat{j}$.

Based on the result of Theorem 1, we can write a set of partial differential equations for the optimal $\vec{D^*}(x,y)$:

$$\begin{aligned}\vec{\nabla}\cdot\vec{D}^*(x,y) = \rho(x,y)\\\vec{\nabla}\times\vec{D}^*(x,y) = 0\end{aligned} \tag{13}$$

A more general case of stating the above optimization problem is to add a scalar function $K(x,y)$ as the coefficient of the integrand in cost defined by (11):

$$\begin{aligned}&\text{Minimize } J(\vec{D}) = \int_A K(x,y)|\vec{D}(x,y)|^2\,\mathrm{d}s\\&\text{Subject to:}\\&\vec{\nabla}\cdot\vec{D}(x,y) = \rho(x,y)\\&D_n(x,y) = 0\ (x,y)\in \text{Boundary of } A\end{aligned}$$

In this case $K(x,y)$ takes into consideration the cost of routing through point $(x,y)$ of the network. In [2] we have used $K(x,y)$ for energy efficient routing, and in that work, $K(x,y)$ takes high values in the regions of the network with low residual energy sensors. The following theorem provides the key to finding the solution of the optimization problem defined by (14).

**Theorem 2:** *If $\vec{D}^*(x,y)$ denotes the optimal solution of equation (14), then it satisfies:*

$$\vec{\nabla}\times\vec{E}^*(x,y) = 0 \tag{14}$$

*in which*

$$\vec{E}^*(x,y) = K(x,y)\vec{D}^*(x,y) \tag{15}$$

The proof of the above theorem can be found in [2]. Based on the result of Theorem 2, we can write a set of partial differential equations for the optimal $\vec{D}^*(x,y)$ and $\vec{E}^*(x,y)$:

$$\begin{aligned}\vec{\nabla}\cdot\vec{D}^*(x,y) = \rho(x,y)\\\vec{\nabla}\times\vec{E}^*(x,y) = 0\end{aligned} \tag{16}$$

The set of equations given by (16) is analogous to Maxwell's equations in the electrostatic theory. In this analogy, $\vec{E}^*(x,y)$ is analogous to the electric field density, $\vec{D}^*(x,y)$, is analogous to the electric displacement, and $K(x,y)$ is analogous to the inverse of permittivity factor in a non-homogeneous media. In the theory of partial differential equations it is proved that the above equations along with the boundary condition given by (8) give $\vec{D^*}(x,y)$ and $\vec{E^*}(x,y)$ uniquely.

Since $\vec{\nabla}\times\vec{E^*}(x,y) = 0$ then $\vec{E^*}(x,y)$ is conservative, and it can be expressed as the gradient of a scalar function:

$$\vec{E}^* = \vec{\nabla}U \tag{17}$$

Then the set of equations defined by (16) reduces to:

$$\nabla^2 U(x,y) - \frac{\vec{\nabla}K(x,y).\vec{\nabla}U(x,y)}{K(x,y)} = K(x,y)\rho(x,y) \tag{18}$$

in which the operator $\nabla^2$ is defined as:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \tag{19}$$

The boundary conditions for $\vec{D}(x,y)$ imply that the gradient of $U(x,y)$ is zero on the boundary along the direction that is normal to the boundary:

$$\vec{\nabla}U(x,y)\cdot\hat{n}(x,y) = 0,\ \forall(x,y)\in \text{Boundary of } A \tag{20}$$

in which $\hat{n}(x,y)$ is a unit vector normal to the boundary. The above boundary conditions are known as *Neumann* type boundary definition in the literature of partial differential equations. For the case in which $K(x,y)$ is constant, Equation (18) reduces to the well known *Poisson* equation:

$$\nabla^2 U(x,y) = K(x,y)\rho(x,y) \tag{21}$$

## III. Generalization To Multiple Destinations

So far we have introduced the vector field based routing method for the case where there is only one destination in the network, and all the traffic generated by the nodes is sent to that destination. While this is the case in many applications of sensor networks, a more general case is where we have more than one destination in the network.

In the case with a single destination, all the traffic generated by sensors is sent to that destination; however, a complexity of the case where we have several destinations is how to distribute the load of the network among them. In the case of multiple destinations, we write the optimization problem as:

$$\begin{aligned}&\text{Minimize } J(\vec{D}) = \int_A K(x,y)|\vec{D}(x,y)|^2 dxdy\\&\text{Subject to:}\\&\vec{\nabla}\cdot\vec{D}(x,y) = \rho(x,y)\\&\rho(x,y) = r(x,y) - \textstyle\sum_{i=1}^{M} w_i\delta(x-x_i)\delta(y-y_i)\\&D_n(x,y) = 0\ ,\forall(x,y)\in \text{Boundary of } A\end{aligned}$$

in which $M$ is the number of destinations, $(x_i, y_i)$ is the location of $i^{th}$ destination and $w_i$ is a nonnegative weight of the $i^{th}$ destination, which represents the amount of load that is sinked at the $i^{th}$ destination. Since in this case the load of the network is received by the $M$ destinations, we have:

$$\sum_{i=1}^{M} w_i = \int_A r(x,y)dxdy = w_0$$

This implies $\int_A \rho(x,y)dxdy = 0$. In the case with multiple destinations the optimization is both on $\vec{D}(x,y)$ and the values of $w_i$. In this case, the paths starting from each point of the network end at one of the destinations. Based on the paths, we partition the area of the network into $M$ disjoint sets corresponding to the $M$ destinations. Let $T_i$ denote the set for the $i^{th}$ destination. Then a point $(x,y)$ belongs to $T_i$ if the path corresponding to $(x,y)$ ends at the destination $i$.

We call $T_i$ the *region of attraction* of destination $i$. It is straightforward to verify that the region of attraction for each destination is a connected set. This is because $(x_i, y_i) \in T_i$, and $(x_i, y_i)$ belongs to every path that ends at the destination $i$. Then if two points such as $(x,y)$ and $(x', y')$ both belong to $T_i$, and if $p$ and $p'$ are their corresponding paths to the $i^{th}$ destination respectively, then based on the suffix property of paths the union of $p$ and $p'$ is a connected subset of $T_i$ containing both $(x,y)$ and $(x', y')$.

Based on the above definition of $T_i$, we can write the weight of $i^{th}$ destination as:

$$w_i = \int_{T_i} r(x,y)dxdy \tag{22}$$

The complexity of problem in the multiple destination case is based on the fact that optimization is both on $\vec{D}(x,y)$, and the weight values $w_1, w_2, \cdots, w_M$. If the weight values are fixed, the following lemma can be stated:

**Lemma 1:** *In the case of multiple destinations, if $w_1, w_2, \cdots, w_M$, are fixed, the necessary and sufficient condition for optimality of the cost function in (22) is:*

$$\vec{\nabla} \times \vec{E}^*(x,y) = 0 \tag{23}$$

*in which*

$$\vec{E}^*(x,y) = K(x,y)\vec{D}^*(x,y) \tag{24}$$

The proof of this lemma is similar to the proof of Theorem 1 and Theorem 2. If the weights are given, the value of $\rho(x,y)$ is known, and the optimal solution of the problem is found by solving the following PDE:

$$\begin{array}{l} \nabla^2 U(x,y) - \frac{\vec{\nabla} K(x,y).\vec{\nabla} U(x,y)}{K(x,y)} = K(x,y)\rho(x,y) \\ \vec{\nabla} U(x,y) \cdot \hat{n}(x,y) = 0, \forall (x,y) \in \text{Boundary of } A \end{array} \tag{25}$$

and ultimately $\vec{E}(x,y)$ is obtained from $\vec{E}(x,y) = \nabla U(x,y)$, and $\vec{D}(x,y)$ is found from $\vec{E}(x,y)$.

If the weight values of the destinations are added to the optimization variables of (22), the condition $\vec{\nabla} \times \vec{E}^*(x,y) = 0$ is a necessary but not a sufficient condition for optimality of the cost function. The following theorem gives the basic idea to solve the optimization problem of (22), where we have multiple destinations:

**Theorem 3:** *If the value of potential function at the locations of $M$ destinations is $U_1$, $U_2$, ..., $U_M$, then the necessary and sufficient conditions for the optimality of cost function in (22) are:*

$$\begin{array}{l} \vec{\nabla} \times \vec{E}(x,y) = 0 \\ U_i = U_j, \ \textit{for} \ \forall \ 1 \le i, j \le M \end{array}$$

*Proof:* The first condition in the theorem is $\vec{\nabla} \times \vec{E}(x,y) = 0$. The proof for this condition is similar to case of having a single destination. We assume this condition holds and we show that the second condition is necessary and sufficient for optimality.

In the forward proof, we show that for optimality of (22) the value of the potential function should be the same at all destinations. For this purpose, we use contradiction. Assume for some $i$ and $j$, we have $U_i < U_j$. Then we prove that the load distribution can be changed in the way of decreasing the cost function. We need the following two identities throughout the proof:

**Identity 1:** If $c$ is a scalar function and $\vec{F}$ is a vector field, then:

$$\vec{\nabla} \cdot (c\vec{F}) = c\vec{\nabla} \cdot \vec{F} + \vec{F} \cdot \vec{\nabla} c \tag{26}$$

**Identity 2:** If $A$ is a region in the plane with boundary $\partial A$, and $F$ is a vector field defined on $A$, then

$$\int_A \vec{\nabla} \cdot \vec{F} dxdy = \oint_{\partial A} \vec{F} \cdot \vec{dn} \tag{27}$$

in which $\vec{dn}$ is the differential vector normal to the boundary of $A$ pointing outward. The second identity is also known as the *Divergence Theorem* in the vector calculus literature.

To continue the proof, we make a small positive change in the weights of the $i^{th}$ and the $j^{th}$ destinations:

$$\begin{array}{l} w'_i = w_i + \epsilon \\ w'_j = w_j - \epsilon \end{array} \tag{28}$$

In other words, we increase the weight of the $i^{th}$ destination by $\epsilon$ and decrease the weight of the $j^{th}$ destination by the same amount. Assume $\vec{D}(x,y)$, $\vec{E}(x,y)$, $U(x,y)$ and $\rho(x,y)$ represent the values of vector fields, the potential function and and the density of sources before applying the above change and $\vec{D}'(x,y)$, $\vec{E}'(x,y)$, $U'(x,y)$, and $\rho'(x,y)$ represent the same quantities after applying the change. After this change, the density of sources is:

$$\rho'(x,y) = \rho(x,y) - \epsilon\delta(x - x_i)\delta(y - y_i) + \epsilon\delta(x - x_j)\delta(y - y_j) \tag{29}$$

in which $(x_i, y_i)$ is the location of the $i^{th}$ destination and $(x_j, y_j)$ is the location of the $j^{th}$ destination. Now we make the following definitions:

$$\begin{array}{l} \vec{D}'(x,y) = \vec{D}(x,y) + \vec{d}(x,y) \\ \vec{E}'(x,y) = \vec{E}(x,y) + \vec{e}(x,y) \\ U'(x,y) = U(x,y) + u(x,y) \end{array} \tag{30}$$

It is easy to verify that:

$$\begin{array}{l} \vec{\nabla} \cdot \vec{d}(x,y) = -\epsilon\delta(x - x_i)\delta(y - y_i) + \epsilon\delta(x - x_j)\delta(y - y_j) \\ \vec{\nabla} \times \vec{e}(x,y) = 0 \\ \vec{e}(x,y) = \vec{\nabla} u(x,y) \end{array} \tag{31}$$

The change of cost function after applying the above modification of weights is:

$$\Delta J = J(\vec{D}'(x,y)) - J(\vec{D}(x,y)) = \int_A K(x,y)|\vec{D}'(x,y)|^2 dxdy - \int_A K(x,y)|\vec{D}(x,y)|^2 dxdy \tag{32}$$

By substituting the value of $\vec{D}'(x,y)$ we have

$$\Delta J = 2\int_A K(x,y)\vec{d}(x,y)\cdot\vec{D}(x,y)dxdy + \int_A K(x,y)|\vec{d}(x,y)|^2 dxdy \tag{33}$$

Since we assume $\epsilon$ is a very small value, we can ignore the second term in the above equations, and we have:

$$\Delta J = 2\int_A K(x,y)\vec{d}(x,y)\cdot\vec{D}(x,y)dxdy = 2\int_A d(x,y)\cdot\vec{E}(x,y)dxdy \tag{34}$$

Now we use Identity 1 for $c = U(x,y)$ and $F = d(x,y)$:

$$\vec{\nabla}\cdot(U(x,y)\vec{d}(x,y)) = U(x,y)\vec{\nabla}\cdot\vec{d}(x,y) + \vec{d}(x,y)\cdot\vec{\nabla}U(x,y) \tag{35}$$

By using the fact that $\vec{E}(x,y) = \vec{\nabla}U(x,y)$, the above equation can be written as:

$$\vec{d}(x,y)\cdot\vec{E}(x,y) = \vec{\nabla}\cdot(U(x,y)\vec{d}(x,y)) - U(x,y)\vec{\nabla}\cdot\vec{d}(x,y) \tag{36}$$

By substituting this value for $\vec{d}(x,y)\cdot\vec{E}(x,y)$ in equation (34) we have:

$$\Delta J = 2\int_A \vec{\nabla}\cdot(U(x,y)\vec{d}(x,y))dxdy - 2\int_A U(x,y)\vec{\nabla}\cdot\vec{d}(x,y)dxdy \tag{37}$$

Now we use the Divergence Theorem given in Identity 2 for $\vec{F} = U(x,y)\vec{d}(x,y)$. We have:

$$\int_A \vec{\nabla}\cdot(U(x,y)\vec{d}(x,y))dxdy = \oint_{\partial A} U(x,y)\vec{d}(x,y)\cdot\vec{dn} \tag{38}$$

in which $\vec{dn}$ is a differential vector normal to $\partial A$ and pointing outward. Recall that the boundary conditions in the optimization problem of (22) implies both $\vec{D}(x,y)$ and $\vec{D}'(x,y)$ have zero components in the direction normal to the boundary of $A$. So $\vec{d}(x,y) = \vec{D}'(x,y) - \vec{D}(x,y)$ also has zero normal component at every point of the boundary of $A$. This causes the inner product in the integrand of equation (38) to be 0. Therefore

$$\int_A \vec{\nabla}\cdot(U(x,y)\vec{d}(x,y))dxdy = 0 \tag{39}$$

On the other hand from equation (31) we have $\vec{\nabla}\cdot\vec{d}(x,y) = -\epsilon\delta(x-x_i)\delta(y-y_i) + \epsilon\delta(x-x_j)\delta(y-y_j)$. Therefore:

$$\begin{aligned}&\int_A U(x,y)\vec{\nabla}\cdot\vec{d}(x,y)dxdy = \\ &-\epsilon\int_A U(x,y)\delta(x-x_i)\delta(y-y_i)dxdy \\ &+\epsilon\int_A U(x,y)\delta(x-x_j)\delta(y-y_j)dxdy = \\ &-\epsilon U(x_i,y_i) + \epsilon U(x_j,y_j) = -\epsilon(U_i - U_j)\end{aligned} \tag{40}$$

By substituting (40) and (39) in (37) we get:

$$\Delta J = 2\epsilon(U_i - U_j) < 0 \tag{41}$$

and this ends the forward part of the proof. Now we continue the proof of Theorem 3 in the backward part. From the forward part of the proof we know if $\vec{D}(x,y)$ is the optimal solution of (22), then its corresponding potential function takes the same value at the locations of all destinations. Also for this load vector field we have $\vec{\nabla}\times\vec{E}(x,y) = 0$.

Assume in addition to the optimal $\vec{D}(x,y)$, there exists a $\vec{D}'(x,y)$ that satisfies $\vec{\nabla}\times\vec{E}'(x,y) = 0$ as well as the constraints of the optimization problem given in (22). Also assume the corresponding potential function of $\vec{D}'(x,y)$ takes the same value at the locations of all destinations. Then we prove that $\vec{D}'(x,y) = \vec{D}(x,y)$.

Let $U(x,y), w_1, w_2, ..., w_M$, and $\rho(x,y)$ represent the quantities of the optimal solution, and $U'(x,y)$, $w'_1, w'_2, ..., w'_M$ and $\rho'(x,y)$ represent the similar quantities corresponding to $\vec{D}'(x,y)$. We define:

$$\begin{aligned}\vec{e}(x,y) &= \vec{E}(x,y) - \vec{E}'(x,y)\\ \vec{d}(x,y) &= \vec{D}(x,y) - \vec{D}'(x,y)\\ u(x,y) &= U(x,y) - U'(x,y)\\ \sigma(x,y) &= \rho(x,y) - \rho'(x,y)\end{aligned} \tag{42}$$

It is easy to verify that

$$\begin{aligned}&\vec{\nabla}\cdot\vec{d}(x,y) = \sigma(x,y)\\ &\vec{\nabla}\times\vec{e}(x,y) = 0\\ &\vec{e}(x,y) = \vec{\nabla}u(x,y)\end{aligned} \tag{43}$$

Recall that

$$\begin{aligned}\rho(x,y) &= r(x,y) - \textstyle\sum_{i=1}^M w_i\delta(x-x_i)\delta(y-y_i)\\ \rho'(x,y) &= r(x,y) - \textstyle\sum_{i=1}^M w'_i\delta(x-x_i)\delta(y-y_i)\end{aligned}$$

Hence:

$$\sigma(x,y) = -\sum_{i=1}^M (w_i - w'_i)\delta(x-x_i)\delta(x-y_i) \tag{44}$$

which implies that $\sigma(x,y)$ is zero everywhere in the network that is not a destination.

We use contradiction to prove $\vec{D}(x,y) = \vec{D}'(x,y)$. If $\vec{D}(x,y) \neq \vec{D}'(x,y)$, then there exist some destinations for which $w_i \neq w'_i$. This is because of the fact that if for all destinations we have $w_i = w'_i$, we have $\rho(x,y) = \rho'(x,y)$, and therefore $\vec{D}(x,y) = \vec{D}'(x,y)$. If we assume there exist some destinations for which $w_i \neq w'_i$, then we select the destination $j$ for which the corresponding value of $w_j - w'_j$ is minimum. Since $\sum_{i=1}^M (w_i - w'_i) = 0$, then $w_j - w'_j < 0$. Hence the flux lines of $\vec{d}(x,y)$ exit this destination; this is because the value of divergence of $\vec{d}(x,y)$ is positive at the location of destination $j$:

$$\vec{\nabla}\cdot\vec{d}(x_j,y_j) = \sigma(x_j,y_j) = -(w_j - w'_j)\delta(x-x_j)\delta(y-y_j) > 0 \tag{45}$$

The flux lines exiting the destination $j$ can end only at the destinations for which the value of $\sigma(x,y)$ is negative. This is because the value of divergence of $\vec{d}(x,y)$ should be negative at a location that flux lines end. Since $\sigma(x,y)$ can take nonzero values only at the locations of destinations, every flux line exiting destinations $j$ ends at a destination $k$ for which $w_k - w'_k > 0$, and hence $\sigma(x_k,y_k) < 0$. Let $L$ be one of such flux lines. Based on the definition of flux lines, $L$ is tangent to both of $\vec{d}(x,y)$ and $\vec{e}(x,y)$ at every point of it. Next we consider the following value of the potential difference of destinations $j$ and $k$:

$$u(x_j,y_j) - u(x_k,y_k) = \int_L \vec{e}(x,y)\cdot \vec{dl} \tag{46}$$

in which the integration is in the direction of the flux line $L$ (e.g., from the location of destination $j$ to the location of destination $k$). In equation (46) $\vec{dl}$ is a differential vector along $L$, and hence it has the same direction as $\vec{e}(x,y)$ at every point of $L$. Therefore, we have:

$$\vec{e}(x,y)\cdot\vec{dl} = |\vec{e}(x,y)||\vec{dl}| \tag{47}$$

The definition of flux lines implies that $\vec{e}(x,y)$ is nonzero at every point of $L$. Hence:

$$\vec{e}(x,y)\cdot\vec{dl} = |\vec{e}(x,y)||\vec{dl}| > 0 \tag{48}$$

By comparing (48) and (46) we have:

$$u(x_j,y_j) - u(x_k,y_k) > 0 \tag{49}$$

Recall the fact that $U(x,y)$ takes the same value in the locations of all destinations and the same fact is true for $U'(x,y)$. Hence $u(x,y) = U(x,y) - U'(x,y)$ takes the same value in the locations of all destinations. The statement of equation (49) contradicts this fact. Therefore we have $w_i = w'_i$ for all destinations and hence $\vec{D}(x,y) = \vec{D}'(x,y)$. **QED.**

The following is an intermediate result of Theorem 3:

**Corollary 1:** *If we increase the weight of destination $i$ by a small amount $\epsilon$, and subtract that amount from the weight of destination $j$, then the first order increment in the cost function in (22) is:*

$$\delta J = 2\epsilon(U_i - U_j) \tag{50}$$

We use Theorem 3 and Corollary 1 to introduce an iterative algorithm that gives the optimal assignment of load among the destinations. Assume we start with an arbitrary assignment of the weights to the destinations. Weight assignment is subject to the constraints that the sum of the weights should be the total amount of the load in the network:

$$\sum_{i=1}^{M} w_i = \int_A r(x,y)dxdy = w_0 \tag{51}$$

Given the initial weights, we solve the PDE given by (25) to find $U(x,y)$. Then if the value of $U(x,y)$ is the same at all destinations, we have found the optimal solution, otherwise we continue the iterations by reassigning the weights.

We use Corollary 1 to reassign the weights. We know that in order to improve the cost function we have to decrease the weight of destinations with a high value of the potential function and increase the weight of the destinations with small value of potential. For this purpose, we use the average value of the potential function at all destinations as a reference:

$$\bar{U} = \frac{1}{M}\sum_{i=1}^{M} U_i \tag{52}$$

If a destination has a smaller potential value than the above average, we increase its weight, otherwise, we decrease its weight. We use deviation from the above average to specify the amount of change for the weight of each destination.

$$\begin{aligned} \Delta w_i &= \gamma(\bar{U} - U_i) \\ w'_i &= w_i + \Delta w_i \end{aligned} \tag{53}$$

in which $\gamma$ is a small positive step size, and $w'_i$ is the weight of the $i^{th}$ destination after applying the change. The above change of weights preserves the property that $\sum_{i=1}^{M} w'_i = w_0$. We stop the iterations when the maximum of absolute value of $\bar{U} - U_i$ among all destinations is below a threshold:

$$\max_i |\bar{U} - U_i| < \xi \tag{54}$$

in which $\xi$ is a small positive value.

After finding the optimal weight values of all the destinations, the final issue will be to find the corresponding regions of attraction of the destinations. This is an easy problem to solve: for every point $(x,y)$ of the network, we follow the flux lines of $\vec{D}$, until that flux line ends at one of the destinations such as $d$. Then we declare point $(x,y)$ belongs to the region of attraction of destination $d$.

## IV. Numerical Examples

In this section we present the numerical examples with multiple destinations. The network is a $1\times 1$ area, and we assume the total load to be 100, which is uniformly distributed in the network (i.e., $r(x,y) = 100$). We place 4 destinations in the network. The destinations are located at: $(x_1,y_1) = (0.45, 0.45)$, $(x_2,y_2) = (0.75, 0.75)$, $(x_3,y_3) = (0.65, 0.25)$, and $(x_4,y_4) = (0.25, 0.75)$.

In the next step we divided the total load of 100 units evenly among the destinations and assumed $K(x,y) = 1$. This means that $w_1 = w_2 = w_3 = w_4 = 25$. Then we solved the PDE equation for the potential function $U(x,y)$, and from it we found $\vec{D}(x,y)$. The resulting potential values are shown in Fig. 2. The equipotential lines of the potential function are shown in Fig. 3. The value of $\vec{D}(x,y)$ is shown in Fig. 4 and the regions of attraction for the four destinations are shown in Fig. 5. The value of potential function at the destinations is: $U_1 = 0.0723$, $U_2 = 0.0601$, $U_3 = 0.0607$, and $U_4 = 0.0571$. The total value of the cost function in this case is 4.88. Since the value of potential function is not the same at all the locations of destinations, we know that we can update the weight of destinations to reduce the cost function. In the next step, we use the iterations of equation (53) to update the weights of destinations in order to reduce the cost function. The calculations show that the algorithm converges to the weight values within 1 percent of the optimal weights in 3 iterations. The values of optimal weights are: $w_1 = 22.07$, $w_2 = 25.50$, $w_3 = 25.94$, and $w_4 = 26.48$. With these weights, the cost function reduces to 4.07. We have used the gradient step size $\gamma = 200$, and for the stop criterion of iterations, we used $\xi = 0.001$. Also, the value of potential function at the destinations was calculated to be $U_1 = U_2 = U_3 = U_4 = 0.062$.

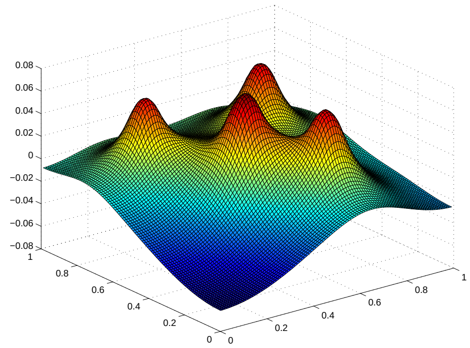


Fig. 2. The value of the potential function $U(x, y)$ for the case with four destinations

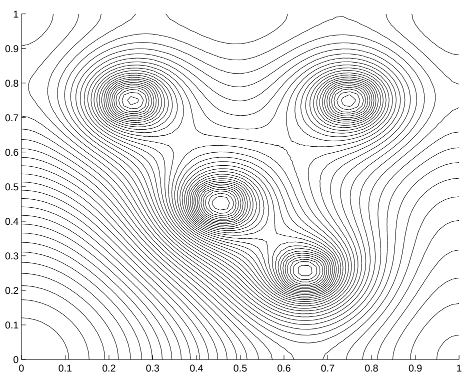


Fig. 3. The equipotential lines for the case with four destinations

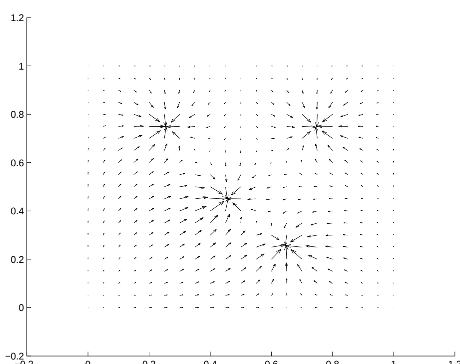


Fig. 4. The value of $\vec{D}(x, y)$ at different places of the network for the case with four destinations

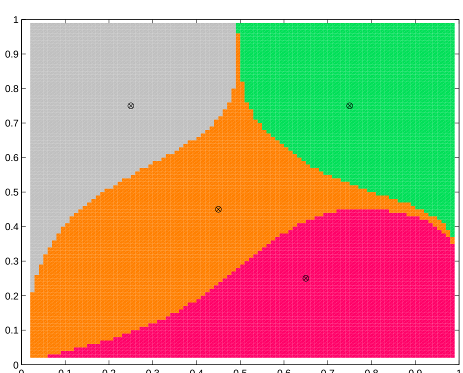


Fig. 5. The regions of attraction for different destinations in the case of four destinations. Destinations are shown by circles.

## V. Conclusion

In this paper, we presented a methodology for routing and optimally designing the topology of a wireless sensor network. We introduced a mathematical formulation based on vector fields and we showed that in order to optimize the communication cost in the network, we need to solve a set of partial differential equations similar to Maxwell's equations in electrostatic theory. Furthermore, we defined the conservative vector field related to the communication load in the network and expressed it as the gradient of a scalar function. In the case in which there are multiple destinations in the network, we have to partition the network into regions of attraction of the destinations. Each destination is responsible for collecting messages of sensors in its region of attraction. We showed in an optimal partitioning, the value of potential function at the location of all destinations should be equal. Also, we gave iterations that lead us to the optimal partitioning.

## References

[1] M. Kalantari and M. Shayman, "Routing in wireless ad hoc networks by analogy to electrostatic theory," in *Proceedings of IEEE International Communications Conference (ICC-04)*. Paris, France, June 2004.

[2] ——, "Energy efficient routing in sensor networks," in *Proceedings of Conference on Information Sciences and System*. Princeton University, March 2004.

[3] K. Sohrabi and J. Pottie, "Protocols for self-organization of a wireless sensor network," *IEEE Personal Communications*, vol. 7, pp. 16–27, 2000.

[4] F. Ye, A. Chen, S. Liu, and L. Zhang, "A scalable solution to minimum cost forwarding in large sensor networks," in *Proceedings of 10th international conference on computer communication and networks, 2001*.

[5] J. H. D. Estrin, R. Govindan and S. Kumar, "Next century challenges: Scalable coordination in sensor networks," in *Proceedings of 5th annual IEEE/ACM international conference on Mobile Computing and Networking, 1999*.

[6] W. H. J. Kulik and H. Balakrishnan, "Negotiation-based protocols for disseminating information in wireless sensor network," in *ACM Wireless Networks, vol. 8, no. 2-3, pp. 169-185, 2002*.

[7] A. P. R. Jain and R. Sengupta, "Geographical routing using partial information for wireless ad hoc networks," in *IEEE Personal Communications, pages 48-57, February 2001*.

[8] D. D. Couto and R. Morris, "Location proxies and intermediate node forwarding for practical geographic forwarding," in *Tech. Rep. MIT-LCSTR- 824, MIT Laboratory for Computer Science, June 2001*.

[9] C. P. S. S. A. Rao, S. Ratnasamy and I. Stocia, "Geographic routing without location information," in *Proceedings of IEEE MOBICOM, September 2003*.

[10] S. Toumpis and L. Tassiulas, "Efficient node deployment in massively dense sensor networks as an electrostatics problem," in *Proceedings of IEEE Infocom 2005*. available online at: http://userver.ftw.at/ toumpis/publications/infocom05.pdf.

[11] N. T. Nguyen, A. A. Wang, P. Reiher, and G. Kuenning, "Electric field routing: A reliable framework for routing in manets," in *Proceedings of ACM SIGMOBILE Mobile Computing and Communications Review, 2004*.

[12] S. G. P. Rentala, R. Musunnuri and U. Saxena, "Survey on sensor networks." available online at: http://www.mdpi.net/sensors.

[13] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *Computer Networks Journal (Elsevier)*, vol. 38, August 2002.

[14] B. R. McQuistan, "Scalar and vector fields: A physical interpretation," *John Wiley & Sons*, 1965.

[15] J. Jackson, "Classical electrodynamics," *John Wiley & Sons*, 1999.

[16] D. Cheng, "Field and wave electromagnetics," *Second Edition, Addison-Wesley*, 1989.