2013

Environment Aware Connectivity for Wireless Underground Sensor Networks

Xin Dong
University of Nebraska-Lincoln, xdong@cse.unl.edu

Mehmet C. Vuran
University of Nebraska-Lincoln, mcvuran@cse.unl.edu

Follow this and additional works at: http://digitalcommons.unl.edu/cseconfwork

Environment Aware Connectivity for Wireless Underground Sensor Networks

Xin Dong   Mehmet C. Vuran
Cyber-Physical Networking Laboratory
Department of Computer Science & Engineering
University of Nebraska-Lincoln, Lincoln, NE 68588
Email: {xdong, mcvuran}@cse.unl.edu

Abstract—**Wireless underground sensor networks (WUSNs) consist of sensors that are buried in and communicate through soil. The channel quality of WUSNs is strongly impacted by environmental parameters such soil moisture. Thus, the communication range of the nodes and the network connectivity vary over time. To address the challenges in underground communication, above ground nodes are deployed to maintain connectivity. In this paper, the connectivity of WUSNs under varying environmental conditions is captured by modeling the cluster size distribution under sub-critical conditions and through a novel aboveground communication coverage model for underground clusters. The resulting connectivity model is utilized to analyze two communication schemes: transmit power control and environment-aware routing, which maintain connectivity while reducing energy consumption. It is shown that transmit power control can maintain network connectivity under all soil moisture values at the cost of energy consumption. Utilizing relays based on soil moisture levels can decrease this energy consumption. A composite of both approaches is also considered to analyze the tradeoff between connectivity and energy consumption.**

I. Introduction

The potential applications of Wireless underground sensor networks (WUSNs) include intelligent irrigation, environment monitoring, infrastructure monitoring, localization, and border patrol [2], [7]. However, WUSNs face unique challenges since the sensor motes are buried in soil. Therefore, novel solutions are needed to address the connectivity and energy consumption issues brought by the unique communication medium.

In WUSNs, the quality of the communication channels varies over time due to the variations in soil moisture [15]. Therefore, for a given field with deployed sensor nodes, the network may not remain connected over time. It may be argued that the network can be designed and deployed according to the worst case to guarantee the connectivity. However, in practice, this is not a suitable solution because (1) the worst case may not be known *a priori* and (2) the deployment cost may be very high due to the large number of nodes required to keep the network connected in the worse case conditions. In fact, most applications, e.g., soil moisture monitoring in agriculture, can tolerate that only part of the network is connected. However, the nodes that are temporally not connected to the network should stop sending data in order to save energy. Therefore, the issue of balancing the number of nodes in the network and energy consumption of the network, considering the dynamics of the environmental parameters, is of importance for WUSNs.

In this paper, two communication schemes: *transmit power control* (TPC) and *environment-aware routing* (EAR) are an-

alyzed. TPC adjusts the transmit power of the underground nodes while EAR turns off the underground nodes that are not directly connected to an aboveground node when the soil moisture is beyond a threshold. The two schemes make local communication decisions based on the locally detected soil moisture level. Only local information is utilized such that no overhead message exchange among nodes is needed. Therefore, energy is saved for the underground nodes. To understand the impact of the two schemes on the network connectivity and the energy consumption, we analyze the connectivity of WUSNs with underground nodes and aboveground nodes in the dynamics of soil moisture.

To this end, the network connectivity is modeled in three steps. First, the distribution of the number of given size clusters is modeled. Second, the aboveground communication coverage, which indicates the probability that a cluster is connected to aboveground nodes, is estimated. Finally, the network connectivity is modeled. This model captures the impact of the variations of the environmental parameters, especially soil moisture, on the network connectivity.

The rest of the paper is organized as follows: the related work is discussed in Section II. Background in underground communication is provided in Section III. In Section IV, the problem is formally stated and in Section V the network connectivity is modeled, including a cluster size distribution model and an aboveground communication coverage model. In addition, these models are verified through extensive simulations in Section VI and the employment of the models to analyze the communication schemes is described in Section VII. Finally, the paper is concluded in Section VIII.

II. Related Work

The connectivity of ad hoc wireless networks has been analyzed by different methods. In [4], the relationship between k-connectivity and node density is investigated. The upper bound of the connectivity is expressed as a function of the density. In this paper, we define connectivity in a different manner, such that the portion of the nodes that are connected in the network since (1) in WUSNs, we focus on routing packets to base stations thus communication among different underground nodes far away is not a main concern; (2) the density of WUSNs is usually low and most nodes have only one path to a base station. Thus, k-connectivity does not apply.

In [8], [9], [19], percolation theory is employed to model wireless network connectivity. The concept of supercritical

state, where an infinite component is formed in the random graph, is utilized to model the case that the wireless network is well connected. The effects of base stations are also investigated in [9], [19], where it is shown that if the communication range of the base station is larger than the communication range of the nodes, a giant component exists even the density of the nodes is below critical density. In WUSNs, the underground motes can be modeled as nodes deployed in a 2-dimensional random graph [8] and the aboveground motes, which have a better communication performance, can be modeled as base stations as in [19]. The same model is also used in [17], where both fixed aboveground motes and mobile aboveground motes are considered.

These studies cannot be directly used in WUSNs since in [8], [9], [19], only the conditions, under which a well connected network is formed, are given. This is not sufficient when considering WUSNs. First, in [9], the network connectivity is considered in an infinite network. However, the conditions for an infinite network to be percolated may still result in nodes that are not in the giant component for a finite network. For example, for a density of $\lambda = 1.5$, which is above critical density λ_c, and a communication range of $r = 1$, simulations show that approximately 20% of nodes are not in the giant component. For these nodes, any communication attempts waste energy. Second, existing work focuses on the supercritical state of the network but the properties of the sub-critical state of networks are not well studied. Sub-critical state is of importance in WUSNs since the underground communication quality is affected by the environment and dynamically changes over time. Thus, it cannot be assumed that the network is percolated at all times.

Wireless underground communication can be thought of as similar to underwater communication [5]. However, due to even higher attenuation caused by sea water, underwater RF communication is not deemed practical. Instead, acoustic [1] and magneto-inductive [11] communications are explored in underwater communications. In [11], the connectivity of a magneto-inductive underwater network is modeled. This work cannot be applied in wireless underground communication since the magneto-inductive channel does not change over time and no above-water nodes are deployed to increase connectivity.

The connectivity of wireless underground sensor networks is analyzed in [17], where both fixed aboveground motes and mobile aboveground motes are considered. An underground node is considered as connected to the network if it can connect to a fixed or mobile aboveground node in a multi-hop fashion at least once within an interval. However, the focus is on the dynamic change in connectivity caused by the duty cycle and mobility of the aboveground nodes. The impacts of the dynamic changes in soil moisture on the connectivity is not considered. In this work, we focus on the impacts of channel variations on the connectivity.

III. UNDERGROUND COMMUNICATION

Unlike over-the-air communications, underground communications are significantly prone to the changes in the environment, especially the variation in soil moisture. For example, in aboveground-to-underground communication, an increase in
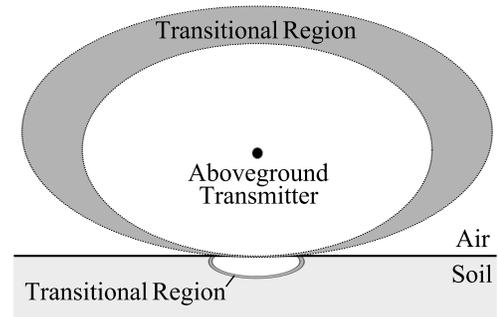

Fig. 1: Communication ranges in air and soil.

the volumetric water content of the soil from 9.5% to 37.3% results in a $9\,\text{dB}$ decrease in the received signal strength at $4\,\text{m}$ (from $89\,\text{dBm}$ to $98\,\text{dBm}$) [15]. Thus, the network topology and connectivity dynamically change over time according to the soil moisture level.

In this section, the unique channel models in WUSNs are introduced in Section III-A and the impact of the soil moisture on network connectivity is described in Section III-B.

A. Channel Models

A WUSN topology consists of two kinds of nodes: underground nodes that are buried at the subsurface of the soil and aboveground nodes that are deployed to gather data from underground nodes [18]. Consequently, there are three different cases of communication channels: underground-to-underground channel (U2U Channel), aboveground-to-underground channel (A2U Channel) and underground-to-aboveground channel (U2A Channel) based on the locations of the transceiver and the receiver.

In [6], [7], we have investigated the communication channels in WUSNs. The received signal strength for the U2U channel, $P_r^{uu}(r)$, U2A channel, $P_r^{ua}(r)$ and A2U channel, $P_r^{au}(r)$ are modeled as functions of soil type, soil moisture and distance. Even though the U2A channel is inherently better than the A2U channel, an aboveground node can increase its transmit power to compensate for the disadvantage of the A2U channel. Thus, in this paper, we consider the channel is symmetric and the calculations for the U2A channel is used.

The communication range of the U2U channel, r_{uu}, is defined as the maximum distance, at which the received signal strength is higher than a threshold, i.e., $r_{uu} = \sup\{r|P_r^{uu}(r) \geq P_{th}\}$.

Likewise, the communication range of the U2A channel, r_{ua}, is defined as the maximum distance at which the received signal strength is above a threshold, i.e., $r_{ua} = \sup\{r|P_r^{ua} \geq P_{th}\}$.

Due to the fact that part of the U2A/A2U channel is in air, the U2A channel has lower attenuation than the U2U channel [15]. For the same node, the communication range of the U2A channel, r_{ua}, can be more than 5 times than that of the U2U channel, r_{uu}, as shown in Fig. 1. Thus, even with a few aboveground nodes, the network connectivity can be significantly improved [17], [18].

B. Impacts of Soil Moisture on Connectivity

Water in soil increases its conductivity, and hence the attenuation of electromagnetic waves. Thus, the attenuation of

soil is a function of soil moisture. Therefore, a network well connected when soil moisture is low can become disconnected when the soil moisture increases (e.g., it rains). Our simulations show that a well connected network (98% of the nodes are in the same giant component) when volumetric water content (VWC) is 10% can be disconnected (less than 5% of the nodes are in the same component) when VWC increases to 25%. In other words, the underground nodes form small disconnected clusters. Therefore, in WUSNs, aboveground nodes should be deployed to improve the connectivity [18]. Since U2A channel has a much longer communication distance, the aboveground nodes work as bridges to connect the underground network components together. In our simulation, where 9 aboveground nodes are randomly deployed in the field, for VWC=20%, the percentage of the connected underground nodes increases from 15% to 72%. However, the network connectivity still varies over different soil moisture values, and hence strategies to improve connectivity while reducing energy consumption are needed.

IV. PROBLEM STATEMENT

Consider a WUSN, which consists of underground sensor nodes and aboveground nodes deployed in a square field, as shown in Fig. 2. The underground nodes are deployed according to a Poisson point process with density λ_u (black dots in Fig. 2). Poisson point process is employed because in agriculture applications, soil moisture sensors are deployed according to the soil types, elevations and slopes in the field. Due to the randomness in soil, a random deployment of underground nodes is required. In the same field, aboveground nodes (white squares in Fig. 2) are also deployed as another Poisson point process with density λ_a to gather the data from the underground nodes.

The set of the underground nodes is denoted as $\mathcal{U} = \{u_1, u_2, \ldots, u_n\}$ and the set of the aboveground nodes is denoted as $\mathcal{A} = \{a_1, a_2, \ldots, a_m\}$. The transmit power of each of the underground node is denoted as P_u. The communication ranges of each underground node, $r_{uu,i}$ (for the U2U channel) and $r_{ua,i}$ (for the U2A channel), are functions of P_u and volumetric water content ω. For a given ω, the nodes have the same U2U communication range, r_{uu}, and the same U2A communication range, r_{ua}. In Fig. 2, the topology of the underground nodes (black dots), aboveground nodes (white squares) and the communication ranges (r_{ua} and r_{uu}) are shown. The A2U channel is not considered our model since the aboveground nodes can always use the maximum transmit power.

A random geometric graph $G_u = G(\mathcal{U}, r_{uu})$ is used to represent the network of the underground nodes while graph $G_h = G(\{\mathcal{U} \bigcup \mathcal{A}\}, r_{uu}, r_{ua})$ is used to represent the underground-aboveground hybrid network. Note G_u and G_h are functions of ω since r_{uu} r_{ua} are functions of ω. In G_u, nodes form components (or clusters) where there exists a path between two nodes u_i and u_j if and only if they are in the same component. According to [12], if the density of the network, λ, is above the critical density λ_c, there exists a giant component $\mathcal{C}_g$ in $G(\mathcal{U}, 1)$ such that most of the nodes are in this component.

Here, it is important to note that in [12], [9], [19], the communication ranges of nodes do not fluctuate and are generally normalized to unit distance. Hence, the density, λ, can
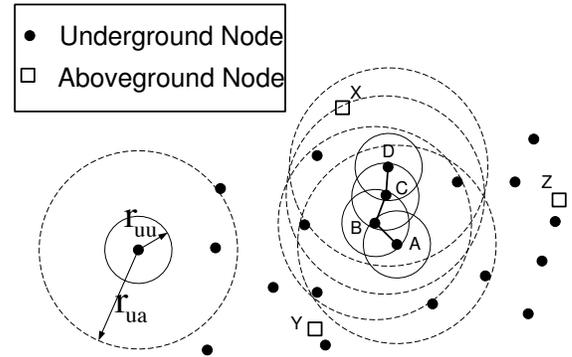

Fig. 2: The hybrid WUSN architecture.

be considered fixed. In WUSNs, however, the communication range fluctuates due to changes in soil moisture. Consequently, the density of the WUSN, λ_u, which is defined based on a unit communication range, *fluctuates* with the changes in soil moisture. Therefore, unless G_u is designed based on the worst case, the network cannot be guaranteed to be well connected at all times by only deployment.

For the hybrid network G_h, the network connectivity is defined as follows.

Definition 1: An underground node u_i is called a *connected node* if and only if it has a path to any aboveground node, i.e., u_i is connected $\equiv \exists a_k \in \mathcal{A}, \exists u_j \in \mathcal{C}_i$ such that $\|u_j - a_k\| \leq r_{ua}$, where $\mathcal{C}_i$ is the cluster that contains u_i.

In Fig. 2, underground nodes A, B, C, and D form a cluster. Since the aboveground node X is in the communication range of node D, these four underground nodes are connected to the network.

Definition 2: The *network connectivity*, φ, is the ratio of connected underground nodes to the total number of underground nodes.

Given a WUSN, G_h, we are interested in finding the best communication scheme that maintains underground network connectivity while saving energy under different soil moisture levels. To this end, a model for analyzing the connectivity of G_h is developed in the following.

V. CONNECTIVITY ANALYSIS

We model the ratio of underground nodes that have paths to aboveground nodes. This is done in three steps. (1) The cluster size distribution of an underground network G_u is derived. Accordingly, if any one node in the cluster is connected to an aboveground node, the whole cluster is connected to the network. (2) We define the concept of the aboveground communication coverage of a cluster, which is the area in which the cluster can communicate to an aboveground node. (3) By applying the first two models, the network connectivity, G_h, is modeled.

A. Cluster Size Distribution

To the best of our knowledge, cluster size distribution for random networks has not been analyzed so far. In this section, we derive an approximation of this distribution. For a random geometric graph $G_u = G(\mathcal{U}, r_{uu})$ with density λ_u, the cluster size distribution can be approximated as

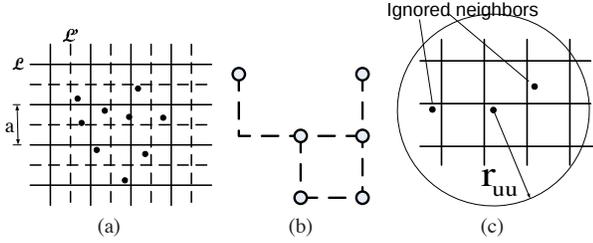$$p_{cs}(n) \approx \alpha n^{-\tau} e^{-n/n_\xi} , \qquad (1)$$

Fig. 3: Mapping the random graph to a lattice: (a) Lattice over the graph; (b) the resulting cluster in the lattice; (c) the ignored neighbors as the result of the lattice mapping

where α and τ are two empirical parameters and n_ξ is the crossover size, i.e., the mean size of the clusters in G_u, which is a function of node density, λ_u, and transmission range, r_{uu}.

To show this approximation, the cluster size distribution of a lattice topology, which has been considered in [3], [16], is employed. For a lattice topology, each site (node) has a site occupation probability. When two adjacent sites are occupied, they are connected. Thus, the size of a cluster is the number of sites in that cluster.

In [16], the cluster size distribution for a lattice topology is found as

$$n_s(p) = s^{-\tau} f[(p - p_c)s^\sigma] \qquad (p \to p_c, s \to \infty) , \quad (2)$$

where s is the size of the cluster, p is the site occupation probability, p_c is the critical probability and τ is an empirical parameter. The cluster size distribution $n_s(p)$ is defined as the number of clusters of size s per site. In this form, $f(z)$ is a function to be determined by numerical methods. An approximation of (2) is also given in [16] as $n_s \propto s^{-\tau} e^{-cs}$, where c can be further approximated as $c \propto |p - p_c|^{1/\sigma}$. In [3], this cluster size distribution for lattice topology is further approximated as

$$n_s \propto s^{-\tau} e^{-s/s_\xi} , \quad (3)$$

where s_ξ is called *crossover size*, which is the mean diameter of the clusters on the lattice.

As shown in Fig. 3(a), given a random graph of the underground network $G_u(r_{uu})$, a lattice $\mathcal{L}$ is built on top of it, with edge length $a = \frac{1}{\sqrt{5}} r_{uu}$. Another lattice, $\mathcal{L}'$, with the same edge length as $\mathcal{L}$, is shifted by $\frac{1}{a}$ on both axises. For each cell c_i of the lattice $\mathcal{L}$, a site s_i of $\mathcal{L}'$ is at the centroid. In other words, each site s_i of $\mathcal{L}'$ has a corresponding cell c_i in $\mathcal{L}$. A site s_i of $\mathcal{L}'$ is occupied if and only if $\exists u_k \in \mathcal{U}$ such that u_k is in cell c_i.

We call lattice $\mathcal{L}'$ as the mapping of the random network $G_u(r^{uu})$, and a cluster in $\mathcal{L}'$ as the result of the mapping is shown in Fig. 3(b). This mapping is imperfect since neighbors in $G_u(r^{uu})$ may not be mapped into adjacent cells in $\mathcal{L}$ as shown in Fig. 3(c). As pointed out in [10], it is impossible to perfectly map a random boolean model to a lattice model. However, when the density of the underground node is low, the number of unconsidered neighbors is negligible and the approximation becomes close. For example, if the density of the underground nodes is 1.5 per unit square, the average number of nodes in one cell is 0.3. Thus, the approximation error is small.
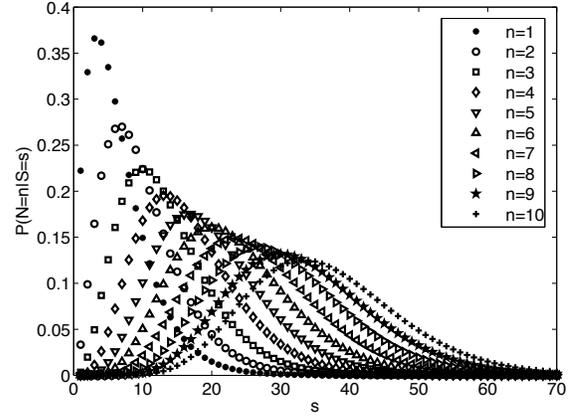

Fig. 4: The distribution of the number of random nodes in a lattice of size S.

For lattice $\mathcal{L}'$, (3) is applied to analyze the cluster size distribution of $\mathcal{L}'$. Denote S as the cluster size in $\mathcal{L}'$. For $S = s$, the total area of the cells represented by this cluster is $a^2 s$. Thus, the distribution of random nodes of G in this area is

$$P(N = n | S = s) = \frac{(\lambda_u a^2 s)^n e^{-\lambda_u a^2 s}}{n!} , \quad (4)$$

where n is the cluster size in G. Therefore, the cluster size distribution of G is

$$p_{cs}(n) = P(N = n) = \sum_{s=1}^{\infty} P(N = n | S = s)P(S = s) , \quad (5)$$

where $P(S = s)$ is given in (3).

The cluster size distribution $p_{cs}(n)$ is not in closed-form. Thus, an approximation is desired. This can be achieved by simplifying $P(N = n | S = s)$. For different n, $P(N = n | S = s)$ is plotted as a function of s in Fig. 4.

As shown in Fig. 4, for fixed n, the standard deviation of $P(N = n | S = s)$ is small, especially when s is small. Thus, $P(N = n | S = s)$ is approximated by a delta function, i.e.,

$$P(N = n | S = s) = \begin{cases} 1 & \text{if } S = s_n, \\ 0 & \text{Otherwise,} \end{cases} \quad (6)$$

where s_n is the size of the lattice cluster such that $N = n$ is most likely to happen. Therefore, the number of random nodes in a cluster can be approximated by $\lambda_u a^2 s$. Thus, the cluster size distribution of G_u has the same form as (3), i.e., it is approximated as (1).

According to the scaling property of random geometric graphs [12], [13], r_{uu} and λ_u can be considered together as $r_{uu}\sqrt{\lambda_u}$ and thus, n_ξ and α in (1) are functions of $r_{uu}\sqrt{\lambda_u}$. The relations are empirically found in Section VI.

B. Aboveground Communication Coverage

As explained in Definition 1, for a WUSN, the connectivity of an underground node u_i is the probability that the cluster, within which it resides, is connected to any aboveground node. In Section V-A, we captured the cluster size $|\mathcal{C}_i|$ of a given node u_i, and now, we model the area of the union of the U2A communication range of cluster $\mathcal{C}_i$. First, define the following.
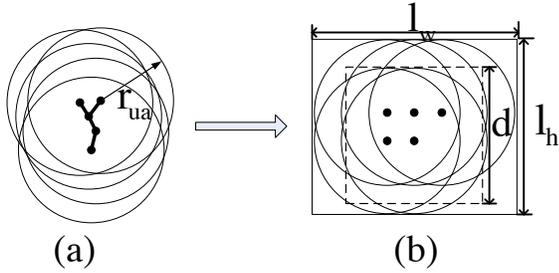
Fig. 5: Redeployment of the nodes to estimate aboveground coverage.

Definition 3: For a node u_i in G_u, the *aboveground communication coverage* of u_i, $\mathcal{S}_i$, is the union of the U2A communication ranges of all the nodes in $\mathcal{C}_i$.

An example of the aboveground communication coverage is shown in Fig. 5(a), where a cluster of size 5 is depicted. Each node in the cluster has a U2A communication range with a radius r_{ua}. The union of all the communication ranges is the *aboveground communication coverage* of the cluster.

Let A_i be the number of above ground nodes in the above ground coverage of u_i, $A_i = \lambda_a S_i$, where λ_a is the density of the above ground nodes and $S_i = |\mathcal{S}_i|$ is the size of the above ground coverage. Since all underground nodes in the same cluster are connected to each other, $\mathcal{S}_{k_1} = \mathcal{S}_{k_2} = \cdots = \mathcal{S}_{k_n}$ for all $u_{k_1}, u_{k_2}, \ldots, u_{k_n} \in \mathcal{C}$. We denote $\mathcal{S}_\mathcal{C}$ as the above ground coverage of the component $\mathcal{C}$.

Since the underground nodes are randomly deployed, even given the number of nodes in a component $\mathcal{C}$, the above ground coverage is still undetermined. This problem is similar to the union volume problem of the penetrable spheres in 2-dimensional space [14]. However, the solution in [14] cannot be directly employed in our case. In [14], the radius of the sphere is the same metric that determines if two particles are in the same cluster. However, in WUSNs, the cluster formation is determined by the U2U communication radius, r_{uu}, whereas the aboveground communication coverage area is determined by the U2A communication radius, r_{ua}. Therefore, in this section, we find an approximation of this union volume in 2-dimensional space.

The methodology is illustrated in Fig. 5. The aboveground communication coverage size S of cluster $\mathcal{C}$ in Fig. 5(a) is estimated by an equivalent cluster $\mathcal{C}'$ in Fig. 5(b). In Fig. 5(a), since the underground nodes are uniformly deployed, the average occupied area of cluster $\mathcal{C}$ in the underground network G_u is C/λ_u, where $C = |\mathcal{C}|$. In Fig. 5(b), the equivalent cluster $\mathcal{C}'$ is deployed in a grid in the square $\mathcal{F} = d \times d$, where $d = \lfloor \sqrt{C/\lambda_u} \rfloor$ and $|\mathcal{C}'| = |\mathcal{C}| = C$. The number of rows of the grid is M and the number of columns of the grid is N, which are given by $M = \lfloor \sqrt{C} \rfloor$, $N = \lceil \sqrt{C} \rceil$.

As $\mathcal{C}$ is redeployed in Fig. 5(b), the rectangle envelop tangential to the U2A communications ranges of all the nodes in $\mathcal{C}$ is formed. The area of the rectangle provides an approximation to the expected aboveground communication coverage of components with size C. The process is shown in Fig. 5(b).

The width of the rectangle is

$$
\begin{aligned}
l_w &= \frac{d}{2} + r_{ua} - \frac{d}{N+1} \\
&= \frac{\lfloor \sqrt{C/\lambda_u} \rfloor}{2} + r_{ua} - \frac{\lfloor \sqrt{C/\lambda_u} \rfloor}{\lceil \sqrt{C} \rceil + 1} ,
\end{aligned}
\tag{7}
$$

and the height of the rectangle, l_h, is found according to (7) by replacing N with M. Thus, the approximation of the expectation of the size of the above ground coverage is

$$
\tilde{\mathbf{E}}[S_C] = l_w \times l_h ,
\tag{8}
$$

where S_C is the aboveground communication coverage of a underground component with C nodes. This estimation is evaluated in Section VI-B.

C. Network Connectivity

Employing the cluster size distribution and the aboveground communication coverage developed in Sections V-A and V-B, the network connectivity is modeled in this section. In the network, G_h, the aboveground nodes are considered always connected to the network and an underground node is considered connected to the network if and only if it satisfies Definition 1.

Assuming the aboveground nodes are randomly deployed with density λ_a, for an underground cluster with size C, the size of the aboveground communication coverage of this cluster, S_C, is given by (8). Thus, the average number of aboveground nodes that the cluster is connected with is $\lambda_a S_C$. If $\lambda_a S_C \geq 1$, all the nodes in this cluster is connected to the network, and when $\lambda_a S_C < 1$, the cluster is connected to the network with probability $\lambda_a S_C$. The number of clusters with size C is given in (1). Therefore, the network connectivity is found as

$$
\varphi(G_h) = \frac{\sum_n \{|\mathcal{U}| p_{cs}(n) \Upsilon\}}{|\mathcal{U}|} = \sum_n p_{cs}(n) \Upsilon
\tag{9}
$$

$$
\Upsilon = \begin{cases} \lambda_a |\mathcal{S}| & \text{if } \lambda_a |\mathcal{S}| < 1, \\ 1 & \text{otherwise} \end{cases}
$$

As shown in Sections V-A and V-B, cluster size distribution is impacted by the underground node density, λ_u, and U2U communication range, r_{uu}, while aboveground coverage depends on the underground node density, λ_u, and U2A communication range, r_{ua}. Furthermore, soil moisture, ω, and transmit power of the underground nodes, P_u, impact the communication ranges of the two channels. Therefore, network connectivity, φ, is a function of λ_u, λ_a, ω, and P_u.

This connectivity model is verified by simulations in Section VI.

VI. Numerical Analysis

Simulations are conducted to analyze the model described in Section V. Four different values of soil moisture, ω, (10%, 20%, 30% and 40%) are considered in the simulations, which cover the range of values in a practical agricultural field. In the simulations, the underground communication range, r_{uu}, is normalized according to the range at $\omega = 10\%$ and r_{uu} changes as a function of ω. Accordingly, the size of the
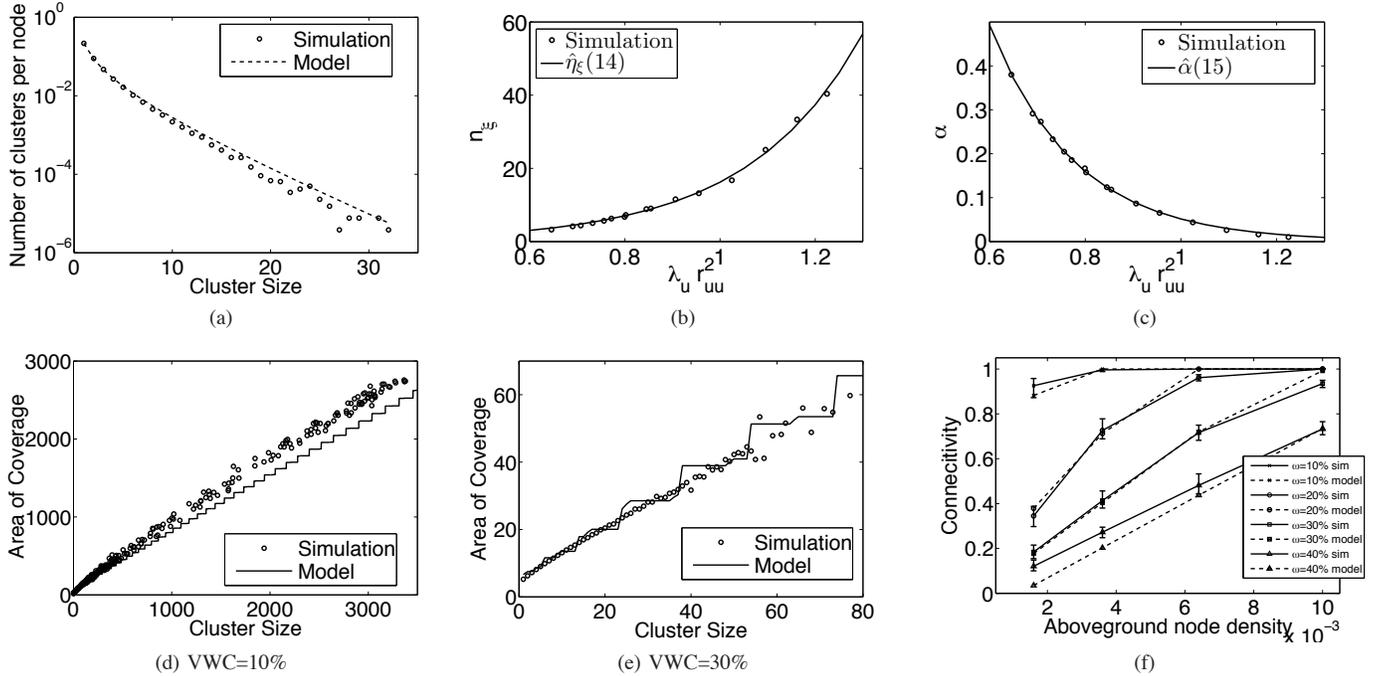
Fig. 6: The model and simulation results: (a) the cluster size distribution model, empirical fitting for (b) n_{xi} and (c) α, (d) aboveground coverage (VWC=10%), (e) aboveground coverage model (VWC=30%), (f) network connectivity.

field is 50×50 unit squares, the density of the underground network, λ_u, is in the range of $\{1.05, 1.2, 1.35, 1.5\}$, and the density of the aboveground network, λ_u, is in the range of $\{0.0016, 0.0036, 0.0064, 0.01\}$. The underground node densities are chosen such that when $\omega = 10\%$, all underground nodes are connected. The aboveground densities are chosen such that they correspond to 4, 9, 16, and 25 nodes in the field. For each pair of ω and λ_u, 100 topologies are generated. In each topology, underground nodes and aboveground nodes are generated according to two Poison point processes. The sizes and numbers of clusters in the topology are compared with our developed model.

The cluster size distribution and aboveground communicaiton coverage models are analyzed in Section VI-A and Section VI-B, respectively. Finally, the model for network connectivity is verified and discussed in Section VI-C.

A. Cluster Size Distribution

The cluster size distribution, (1), is a function of three parameters: α and τ, which are empirical parameters, and n_{ξ}, which is the mean size of the clusters. Using the simulation results, these parameters are estimated empirically to minimize the mean square error (MSE). Accordingly, the values shown in Table I are used in the evaluations, where it can be observed from the RMSE values that the estimated parameters fit well. It is observed that for all the cases, τ is closed to 1 and it is not listed in the table.

The result for $\lambda_u = 1.05$ and $\omega = 30\%$ is shown in Fig. 6(a). It is shown that the model matches the simulation results, especially for small cluster sizes. For cluster size less than 10, the error is less than 20%. The other combinations of λ_u and ω have similar results.

TABLE I: Simulation results for the parameter values.

λ_u	ω	$r_{uu}\sqrt{\lambda_u}$	n_{ξ}	α	RMSE
1.05	10%	1.0247	16.76	0.044	5.43×10^{-5}
1.05	20%	0.7993	6.74	0.17	3.90×10^{-4}
1.05	30%	0.7070	4.39	0.27	7.34×10^{-4}
1.05	40%	0.6456	3.26	0.38	9.88×10^{-4}
1.2	10%	1.0954	25.07	0.026	7.08×10^{-5}
1.2	20%	0.8544	9.03	0.12	1.55×10^{-4}
1.2	30%	0.7559	5.67	0.20	4.44×10^{-4}
1.2	40%	0.6901	4.15	0.29	8.43×10^{-4}
1.35	10%	1.1619	33.33	0.016	4.64×10^{-5}
1.35	20%	0.9063	11.55	0.087	2.22×10^{-4}
1.35	30%	0.8017	7.25	0.16	4.64×10^{-4}
1.35	40%	0.7320	5.06	0.23	5.72×10^{-4}
1.5	10%	1.2247	40.37	0.010	3.11×10^{-5}
1.5	20%	0.9553	13.19	0.07	6.75×10^{-5}
1.5	30%	0.8451	8.86	0.12	2.61×10^{-4}
1.5	40%	0.7716	6.27	0.19	4.87×10^{-4}

The parameters α and n_{ξ} in (1) are functions of the underground node density, λ_u, and the U2U communication range, r_{uu}. More specifically, the scaling property of random graphs make them monotonic functions of $r_{uu}\sqrt{\lambda_u}$. Based on the results in Fig. 6(b) and Fig. 6(c), exponential functions are utilized to describe the relations between α and $r_{uu}\sqrt{\lambda_u}$ as well as n_{ξ} and $r_{uu}\sqrt{\lambda_u}$. The exponential regression curves of n_{ξ} (Fig. 6(b)) and α (Fig. 6(c)), are found as follows: $\hat{n}_{\xi} = 0.2497e^{4.174x}$, $\hat{\alpha} = 14.61e^{-5.65x}$, where $x = r_{uu}\sqrt{\lambda_u}$, with RMSEs of of 0.6967 for n_{ξ} and 3.225×10^{-3} for α.

B. Aboveground Coverage Model

The aboveground communication coverage model in (8) is compared with the simulation results of different U2U communication radii, r_{uu}, and U2A communication radii, r_{ua}, which correspond to different soil moisture values. The results for

volumetric water content of $\omega = 10\%$ and 30% are shown in Fig. 6(d) and Fig. 6(e), respectively. It is shown that the estimation matches closely with the simulation results. When $\omega = 10\%$, the largest error is 10.3% while when $\omega = 30\%$, the largest error is 12.4%.

C. WUSN Connectivity

The network connectivity in (9) is evaluated with simulation results, where the underground density is chosen as $\lambda_u = 1.5$ and four different aboveground node density values are used. The number of aboveground nodes should be as small as possible and in practice and it is chosen based on the application requirement, deployment cost, and geography. Four volumetric water content values (ω) are investigated. The mean and the confidence intervals of 100 simulations are shown with the model results in Fig. 6(f). In most cases, the model matches well with the simulation results with an error less than 10%. For the cases where the soil moisture is high ($\omega = 40\%$) and the aboveground node density is very low (0.0016), higher deviation from the model is observed. Since this case corresponds to the network connectivity of less than 0.4, it should be avoided in practical deployments. It can be observed that for connectivity values of interest (0.8–1) the model agrees well with simulation results.

VII. ENVIRONMENT-AWARE CONNECTIVITY

In WSN systems, overhead traffic should be minimized to save energy. In WUSNs, this is even more important since it is not easy to change or recharge the battery of a buried sensor mote. Thus, it is desired that the buried motes can make communication decisions based on their own local information.

In this section, two communication schemes are developed using the model in (9):

Transmit Power Control (TPC): This scheme adjusts the transmit power of the nodes according to the local soil moisture readings.

Environment-Aware Routing (EAR): The second scheme divides the nodes into two categories: *direct node*, which has direct connection to an aboveground node and *indirect node*, otherwise. The scheme terminates the communication of the indirect nodes when soil moisture value is higher than a threshold.

The trade-off of connectivity and energy consumption is analyzed for each of the two schemes. In Section VII-C, we consider a solution that combines the two schemes for each soil moisture level.

To compare the impacts of each scheme on connectivity and energy consumption, a *total cost*, η, is defined as follows:

$$\eta = (1 - \gamma)(1 - \varphi) + \gamma\varepsilon, \tag{10}$$

where φ is the network connectivity in (9), $\varepsilon = \delta P_u / P_{max}$ is the normalized power consumption of the network, P_u is the transmit power of the underground nodes (in dBm), δ is the portion of active underground nodes in the network, which is determined by the specific scheme, and P_{max} is the maximum transmit power of a node (in dBm). In (10), γ adjusts the trade-off between network connectivity and energy consumption. The total cost, η, is in the range of $[0, 1]$ and a

lower value corresponds to more connected nodes and less the power consumption. The same function is employed in [20] to analyze the trade-off between detection accuracy and energy consumption.

The weight factor γ must be chosen according to the requirements of specific applications. A larger γ emphasizes energy conservation more. In the following, we consider $\gamma = 0.3$ as it prioritizes high connectivity. The impact of γ is discussed in Section VII-C. In the analysis, the results for the underground node density of $\lambda_u = 1.05$ are shown, which corresponds to the worst case density.

A. Transmit Power Control

As shown in Section III-A, the communication ranges of both the U2U and U2A channels change significantly with soil moisture. One impact of this change is the decrease in the network connectivity with soil moisture, especially when the aboveground node density is low (0.0016 as shown in Fig. 6(f)). To compensate for this impact, we employ transmit power control according to the soil moisture values. However, increasing transmit power also increases energy consumption. In this section, we analyze this trade-off.

The transmit power is set in the range of $1\,\text{dBm}$ to $20\,\text{dBm}$, which is the practical range for energy limited sensor nodes. The change in connectivity for different transmit power values is shown in Fig. 7(a), where $\lambda_a = 0.0036$ in four soil moisture levels as well as $\lambda_a = 0.0064$ and 0.01 in the worst case are shown. For a network connectivity requirement of 1, it is shown that by adjusting the transmit power, in all soil moisture conditions, the requirement can be achieved. However, the transmit power varies from $9\,\text{dBm}$ to $14\,\text{dBm}$. Thus, a static transmission power is definitely not optimal. Also notice that by increasing the aboveground node density, the negative impact of the soil moisture can be compensated. It is shown that when the $\omega = 40\%$, the network can achieve the same connectivity as $\omega = 30\%$ with an increase in λ_a from 0.0036 to 0.0064. In other words, deployment of more aboveground nodes increases connectivity without the need for higher transmit power.

In addition, the total cost η is shown in Fig. 7(b). It is shown that although increasing transmit power can increase network connectivity, it is achieved at the cost of energy consumption. The total energy consumption of the network is proportional to the transmit power adopted. Therefore, for different soil moisture levels, there is a best transmit power level to minimize η and balance connectivity and energy consumption. This transmit power level is also impacted by the density of the aboveground nodes as shown in Table II.

In the analysis, it is assumed that the soil moisture value is the same at all the underground nodes, thus the transmit power is set the same. In practice, however, due to the different soil moisture levels at different locations, the transmit power of the each node is set differently. In this case, the same look-up table can still be used.

B. Environment-Aware Routing

When the soil moisture is low and the network is highly connected, the underground nodes can exploit their neighbors to build a multi-hop path to an aboveground node instead of
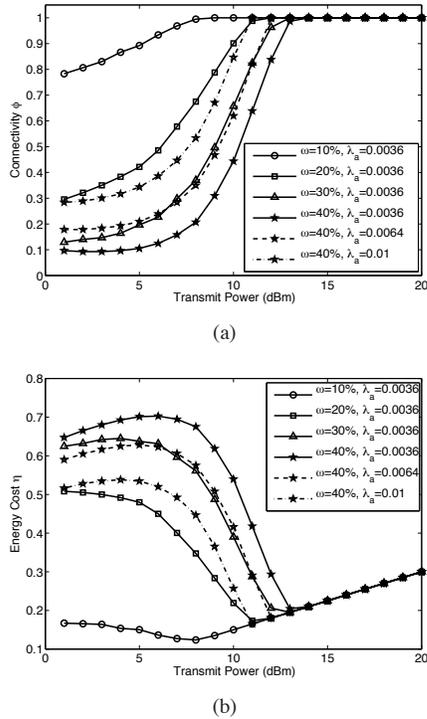
(a)

(b)

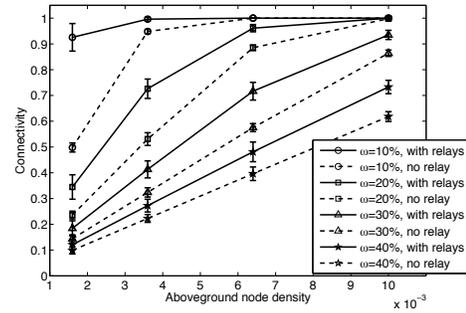Fig. 7: (a) Connectivity and (b) cost function for different transmit power levels.

Fig. 8: Network connectivity with or without underground relays.
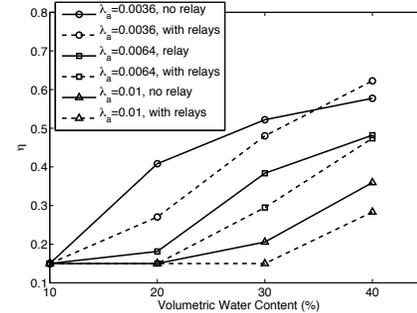
Fig. 9: The impact of the underground relays on the total cost.

direct communication. In this manner, more underground nodes are included in the network, which improves the connectivity. On the other hand, if the soil moisture is high, the U2U channel quality decreases such that the size of the underground clusters is limited. In this situation, if an underground node cannot directly communicate with an aboveground note, the probability that it is connected to the network through its neighbors is low. Thus, all communication attempts will be wasted since the packets will not reach the aboveground nodes. In this case, only underground nodes that are directly connected to the aboveground nodes should send packets.

The impact of this scheme on connectivity is shown in Fig. 8, where the results when underground relays are utilized are calculated using the model developed. When underground relays are not utilized, the results are calculated considering the number of underground nodes in the transmission ranges of the aboveground nodes.

In Fig. 8, it is shown that when the soil moisture value is low (VWC=10%) and the aboveground node density is low ($\lambda_a = 0.0016$), the underground relays have a major impact on the network connectivity by improving it from 0.5 to 0.9. However, if the soil moisture increases, the impact of the relays decreases, especially when the aboveground node density is low. In this case, most of the underground nodes are connected

TABLE II: Optimal transmit power, P_u^* (dBm), for different λ_a and VWC ω.

	P_u^* (dBm)		
ω	$\lambda_a = 0.0036$	$\lambda_a = 0.0064$	$\lambda_a = 0.01$
10%	8	1	1
20%	11	9	7
30%	13	11	10
40%	13	12	11

to the network through one hop connection to the aboveground node.

The impact on the total cost function, η, of utilizing underground relays is depicted in Fig. 9. It is shown that when the soil moisture is low, utilizing underground relays can improve energy performance (lower η). However, when soil moisture increases, the system without relay is more energy efficient when the aboveground node density is low ($\omega = 40\%$ and $\lambda_a = 0.0036$). For higher aboveground node density, in all soil moisture levels, including indirect nodes in the network achieves lower total cost.

C. The Composite Scheme

The two schemes described in Sections VII-A and VII-B have their own merits. In this section, we consider a composite scheme for the underground nodes in different soil moisture values. In other words, we minimize (10) by adjusting underground node transmit power and exploiting the choice of turning off indirect nodes at the same time. This results in a mapping between soil moisture values and communication parameters. Each node detects the soil moisture locally and selects the corresponding communication scheme individually.

Three aboveground node densities are considered, $\lambda_a = \{0.0036, 0.0064, 0.01\}$. In addition, the transmit power of the nodes is in the range of 1 dBm and 10 dBm. The resulting parameters are shown in Table III, where the transmit power values and whether relay is used are shown. The network connectivity is shown in Fig. 10(a). It is shown that when soil moisture is low (10%–15%), the network connectivity can be maintained high ($> 95\%$) even when the aboveground node density is low (0.0036). On the other hand, for higher soil moisture (35%–40%) and low aboveground node density (0.0036), maintaining connectivity has a higher cost for underground
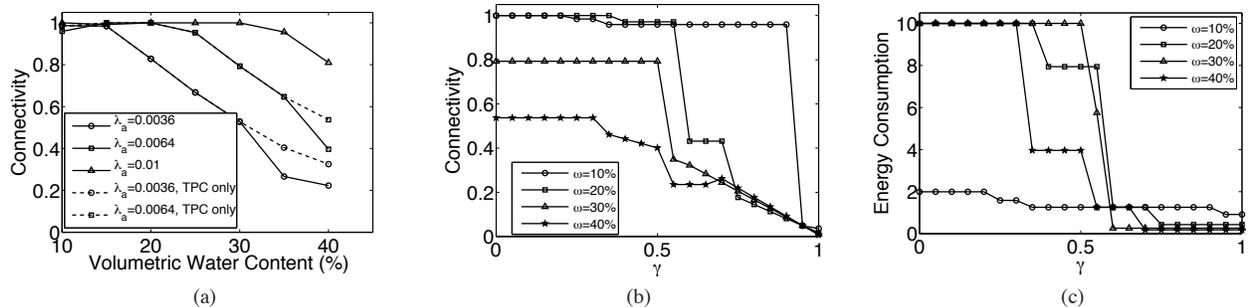
Fig. 10: The composite scheme: (a) the connectivity for the best strategy for different node densities, (b) the effect of γ on connectivity and (c) energy consumption.

TABLE III: Transmission Strategies ($\gamma = 0.3$).

ω	$\lambda_a = 0.0036$		$\lambda_a = 0.0064$		$\lambda_a = 0.01$	
	Relay	P_u (dBm)	Relay	P_u (dBm)	Relay	P_u (dBm)
10%	On	8	On	1	On	1
15%	On	10	On	8	On	2
20%	On	10	On	10	On	7
25%	On	10	On	10	On	9
30%	On	10	On	10	On	10
35%	**Off**	10	On	10	On	10
40%	**Off**	10	**Off**	10	On	10

nodes. Thus, routing is turned off to save energy.

In the high soil moisture conditions, to maintain connectivity, more aboveground nodes need to be deployed, as shown in Fig. 10(a). For an increase in the density of the aboveground nodes from 0.0036 to 0.01, the connectivity increases from 0.2 to 0.81 for $\omega = 40\%$. It is important to note that the energy consumption of aboveground nodes are not considered as they can be easily recharged using solar panels in agricultural fields.

The effects of the weight, γ, in (10) on the connectivity and the energy consumption are shown in Fig. 10(b) and Fig. 10(c), respectively. The energy consumption is normalized to the case where the transmit power is 1 dBm and all nodes are active. It is shown that by choosing a lower γ (< 0.5), the system maintains higher connectivity. However, connectivity is also limited by physical constrains. When $\omega = 40\%$, connectivity cannot be increased above 0.54 regardless of γ. When a higher weight ($\gamma > 0.5$) is used, energy consumption is weighted higher and the system decreases the energy consumption significantly at the cost of lower connectivity.

VIII. CONCLUSIONS

In this work, the connectivity of WUSNs under varying environment conditions is analyzed. By mapping the random network to a lattice topology, the cluster size distribution of a random network is estimated. Then, the aboveground communication coverage of the underground clusters is approximated. Building upon these two models, the underground network connectivity is captured. This model considers the variation of soil moisture, which significantly changes the communication ranges in underground channels.

Applying the developed connectivity model, we also analyze two communication schemes to balance connectivity and energy consumption. It is shown that adjusting transmit power and turning off indirect nodes in high soil moisture are effective in saving energy whiling maintaining connectivity.

REFERENCES

[1] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: research challenges," *Ad hoc networks*, vol. 3, no. 3, pp. 257–279, 2005.

[2] I. Akyildiz and E. Stuntebeck, "Wireless underground sensor networks: Research challenges," *Ad Hoc Networks*, vol. 4, no. 6, pp. 669–686, Nov. 2006.

[3] M. Z. Bazant, "Largest cluster in subcritical percolation," *Physical Review E*, vol. 62, no. 2, pp. 1660–1669, Aug. 2000.

[4] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proc. ACM MobiHoc '02*, Lausanne, Switzerland, Jun. 2002.

[5] J. Cui, J. Kong, M. Gerla, and S. Zhou, "The challenges of building mobile underwater wireless networks for aquatic applications," *IEEE Network*, vol. 20, no. 3, pp. 12–18, 2006.

[6] X. Dong and M. C. Vuran, "A channel model for wireless underground sensor networks using lateral waves," in *Proc. IEEE Globecom '11*, Houston, Tx, December 2011.

[7] X. Dong, M. C. Vuran, and S. Irmak, "Autonomous precision agriculture through integration of wireless underground sensor networks with center pivot irrigation systems," *Ad Hoc Networks*, 2012, accepted.

[8] O. Dousse, F. Baccelli, and P. Thiran, "Impact of interferences on connectivity in ad hoc networks," *IEEE/ACM Trans. Networking*, vol. 13, no. 2, pp. 425 – 436, April 2005.

[9] O. Dousse, P. Thiran, and M. Hasler, "Connectivity in ad-hoc and hybrid networks," in *Proc. IEEE INFOCOM '02*, Jun. 2002.

[10] E. Gilbert, "Random plane networks," *Journal of the Society for Industrial and Applied Mathematics*, vol. 9, no. 4, pp. 533–543, Dec. 1961.

[11] B. Gulbahar and O. Akan, "A communication theoretical modeling and analysis of underwater magneto-inductive wireless channels," *IEEE Trans. Wireless Communications*, vol. 11, no. 9, pp. 3326–3334, Sep. 2012.

[12] R. Meester and R. Roy, *Continuum percolation*. Cambridge Univ. Press, 1996.

[13] M. Penrose, *Random Geometric Graphs*. Oxford University Press, 2003.

[14] J. Quintanilla and S. Torquato, "Clustering properties of *d*-dimensional overlapping spheres," *Physical Review E*, vol. 54, no. 5, pp. 5331–5339, Nov. 1996.

[15] A. R. Silva and M. C. Vuran, "Communication with aboveground devices in wireless underground sensor networks: An empirical study," in *Proc. IEEE ICC '10*, Cape Town, South Africa, May 2010.

[16] D. Stauffer and A. Aharony, *Introduction to percolation theory*, 2nd ed. Taylor & Francis, 1994.

[17] Z. Sun and I. Akyildiz, "Dynamic connectivity in wireless underground sensor networks," *IEEE Trans. Wireless Communicatoins*, vol. 10, no. 12, pp. 4334–4344, Dec. 2011.

[18] J. Tooker and M. C. Vuran, "Mobile data harvesting in wireless underground sensor networks," in *Proc. IEEE SECON '12*, Seoul, Korea, June 2012.

[19] C. Yi and W. Wang, "On the connectivity of large-scale hybrid wireless networks," in *Proc. IEEE ICC '10*, Cape Town, South Africa, May 2010.

[20] D. Zois, M. Levorato, and U. Mitra, "A POMDP framework for heterogeneous sensor selection in wireless body area networks," in *Proc. IEEE INFOCOM '12*, Mar. 2012.