Wave Addressing for Dense Sensor Networks

Serdar Vural and Eylem Ekici
Department of Electrical and Computer Engineering
Ohio State University, Columbus, OH 43210
{vurals,ekici}@ece.osu.edu

***Abstract*— Localization and coordinate assignment in sensor networks play an important role in relating the sensors with geographic locations. Existing localization schemes require use of specialized hardware or signal processing techniques in addition to infrastructure support, which may not be feasible for all deployment scenarios. In this work, we introduce *Wave Addressing for Dense Sensor Networks (WADS)* method to create a coordinate system in dense and large-scale sensor networks without using any special hardware, infrastructure support, or signal processing capability. We propose to use the hop count as distance metric to form a coordinate system. The resulting *Wave Mapping Coordinate (WMC)* system associates each sensor node in the network with a *region* that is at a particular distance to two randomly selected wave sources. We have shown through simulations that the WMC system is capable of segmenting the network into well-defined regions.**

I. Introduction

With the advances in hardware and communications technologies, it is now possible to produce simple sensing devices with wireless communication capabilities in large numbers and at low costs [1], [2]. These small and inexpensive devices can be deployed in large numbers forming wireless sensor networks to observe phenomena in the deployment area. Reconnaissance and close-range sensing in hostile environments can be accomplished efficiently using wireless sensor networks. Since accurate placement of sensors in preplanned positions is not always possible in hostile environments, random deployment of sensors such as scattering from moving (possibly airborne) vehicles is more plausible [1]. Furthermore, to achieve complete coverage of the sensor field, sensor networks must be composed of a high number of sensors as the sensing range of individual sensor nodes is limited. These two factors result in *random sensor network topologies with high sensor density*.

In sensor networks, the location of the observed phenomena and the location of the sensors are highly correlated. A number of existing solutions for data querying, task assignment, and data transfer rely on the availability of location information at the sensor nodes. As an example, assigning sensing tasks to nodes in a given geographical location has been proposed as one of the addressing methods in [3]. Furthermore, various geographic-based routing protocols for sensor [4] and ad hoc networks [5] assume the availability of accurate location information in all nodes. Yet, other protocols such as [6] use the location information to create information forwarding backbones in sensor networks.

In literature, various techniques have been proposed to determine the location of the sensor nodes. One of the most commonly used assumptions is the availability of a GPS [7] device in every sensor node. Assuming the availability of a GPS receiver in every sensor node not only increases the cost of the nodes (hence, limits the number of sensors in a sensor network) but also causes fast depletion of limited energy sources. There exist other schemes [8], [9], [10] to compute sensor locations without using GPS. While early examples like [11] utilize signal strength measurements to calculate distance between beacons and nodes, later examples use time of arrival, difference in time of arrival, and angle of arrival of beacons from known locations [12], [13], [14], [15]. These solutions can be implemented in sensor networks at the expense of increased complexity of the sensor nodes and provisioning of some reference infrastructure on which all other calculations are based.

Although a significant amount of work has been done to improve the accuracy of localization in sensor networks, two important questions require reconsideration:

- What level of location accuracy is actually needed for the application?
- Can hardware and infrastructure requirements necessary to implement the chosen localization method be satisfied for the given scale of deployment and properties of the terrain?

Many applications can tolerate higher level of uncertainty in location information. As an example, intrusion detection and interception applications can be realized with approximate location information of an event in a relatively coarse-grained coordinate system. Once the intrusion is detected, approximate location of the intruder may be sufficient to intercept the intruder. Similarly, a simple coordinate system that is consistent in itself can be used to route packet between two points in a sensor network. Therefore, it is not always necessary to implement costly, energy-inefficient, and infrastructure-reliant localization schemes that provide high fidelity localization.

In this paper, we introduce a simple and efficient mechanism called *Wave Addressing for Dense Sensor Networks (WADS)* to form a coordinate system called *Wave Mapping Coordinate (WMC)* without requiring specialized hardware in sensor nodes or infrastructure support in the network. Our new WADS mechanism creates the coordinate system at the network deployment time based on the hop distance of sensor nodes to two randomly selected sensor nodes called *wave sources*. Neither wave sources nor other sensor nodes require to have clock synchronization or any signal processing capability to implement WADS. As a result of WADS, every sensor can identify itself as belonging to a *region*. Using

the WMC system, it is possible to trace back or locate the source of an event at region level. If the locations of wave sources are known with respect to another (possibly finer-grained) coordinate system, it is also possible to map all regions of WMC to this new coordinate system. The rest of this paper is organized as follows: In Section II, our new WADS mechanism is introduced. Performance of the WADS mechanism is demonstrated with simulation results in Section III. Finally, Section IV concludes the paper.

II. Description of the WADS Mechanism

A. System Description and Assumptions

The sensor network in which we intend to deploy our new WADS system is composed of sensors of identical communication capabilities. In addition to their specific sensing hardware, sensor nodes are assumed to possess the capability of receiving information from and sending to other sensor nodes in their communication range, and performing simple calculations with the received information. An important property of these sensor nodes is that they do not require to have any time synchronization among themselves or any other node in the network. We assume that the sensing application requires densely and randomly deployed sensor nodes in the sensing field. It is also assumed that sensor nodes have very limited, if any, mobility. Our WADS mechanism does not require any specific nodes in the network that carry special properties. As such, synchronized beacons or nodes that can determine their absolute locations with GPS devices are not envisioned as parts of the sensor network. Consequently, WADS is well-suited for long-lived sensor deployment scenarios in large-scale and hostile environments. Since the node functionality is kept at a minimum, the cost of sensor nodes is reduced and the potential number of sensor nodes is increased to achieve more accurate sensing and for higher robustness.

B. WADS Operation

Our proposed WADS mechanism creates the Wave Mapping Coordinate (WMC) system using the hop distance of sensors from two randomly selected sensor nodes. These sensor nodes initialize messages that are broadcast in the network. In every hop, these messages propagate in the shape of waves, hence the name wave addressing. Every sensor node receives two wave IDs, one originating from each of the wave sources. Neighboring sensors that share the same ID pairs form a *region*, and this ID pair is used as the address of the region. Note that region IDs do not belong to a single sensor node but to a set of sensor nodes in the same locality.

An idealized wave generation and region formation example is depicted in Figure 1. In this example, we consider an extreme case where the sensor density is infinite, i.e., a packet can always travel in any direction precisely as far away as the communication range of the sender. Two wave sources, WS_1 and WS_2, generate wave packets that assign IDs to other sensors. The intersection areas of circular rings are the regions where sensor nodes sharing the same ID pairs are expected to reside. This grouping creates the points in our wave mapping coordinate system. In Figure 1, region A is assigned the coordinates (8,10) since it is 8 hops away from WS_1 and 9 hops away from WS_2. Similarly, region B is assigned the coordinates (12,5). Note that the resulting coordinate system is not linear and the regions are not of identical shapes and sizes. However, as will be shown in Section III, WADS results in localized groups that are consistently distributed with small variances around their mean.

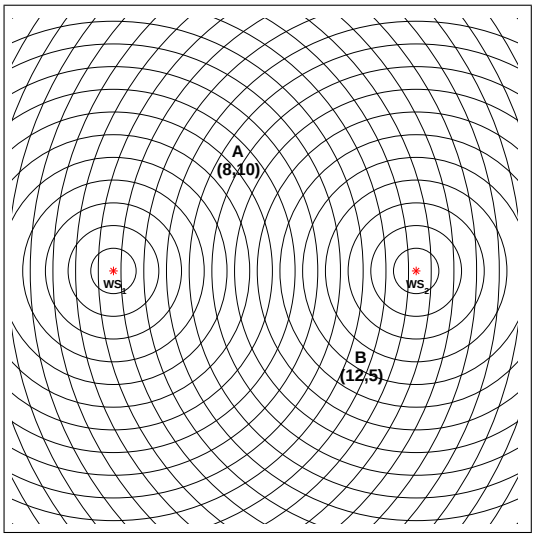


Fig. 1
Example Wave Mapping Coordinate System Generation

The creation of the WMC system is accomplished in three steps: (i) Wave Source Selection, (ii) Wave Propagation and Region Generation (iii) Address Ambiguity Elimination. After these steps, every sensor node associates itself with a region in WMC system. These regions can be used for various purposes ranging from addressing sensors in a locality to assisting routing decisions and estimating distance between event locations.

1) Wave Source Selection: The first step in WMC system generation is selecting the sensor nodes that will serve as the two reference points in the network. Although the entire WMC system will be based on these two points, they can be selected in a distributed manner. Alternatively, two sensor nodes can be pre-configured to serve as the wave sources. In this section, we will discuss two possible wave source deployment scenarios, namely, *pre-deployment* and *post-deployment selection*. Other deployment methods are definitely possible as long as two distinct wave sources can be selected.

In the pre-deployment wave source selection scenario, sensors that will act as wave sources are selected and programmed by the deployers. For this operation, two sources must be selected as primary wave sources. If it is possible to influence the approximate location of the wave sources, then these two wave sources must be placed sufficiently separated from each other[1]. As an example, if sensors are deployed from a moving aircraft, then the first wave source can be thrown into the sensing field shortly after the beginning and the second wave source right before the end of the deployment. This basic

[1]The effects of distance between wave sources are analyzed is Section III.

deployment strategy can be augmented with redundant standby wave sources to ensure operation in case the designated wave sources are damaged during deployment. The standby wave sources can activate themselves if no waves are received from one of the wave sources after a predetermined period of time and start acting as wave sources following the guidelines of the post-deployment selection scenario described below.

In the post-deployment wave source selection scenario, wave sources are not selected before deployment time. To determine the wave sources, it is possible to use leader election algorithms based on IDs sensors assign themselves randomly. Since we need two wave sources, sensors with highest and lowest random IDs can be selected as the two wave sources. However, since the sensor network in question is large, it is easier (and more efficient) to use simpler algorithms that incurs lower overhead. The outline of a possible simple distributed wave source selection operation is given as follows: Every sensor randomly determines if it is a potential candidate for the first or second wave source by generating a random number. Every sensor also generates a random number N_{wait} between 0 and N_{max}, and waits for $T = T_{turn} \times N_{wait}$ amount of time, where T_{turn} is the unit time period, which is in the order of the delay of a packet that would travel the maximum diameter of the network. N_{max} determines the maximum delay tolerance of the system for WMC system generation[2]. Once these two numbers are generated, the sensor starts waiting until either it receives an ID from another wave source with the same source ID (1 or 2) or its timer T expires. If a sensor receives a wave packet with a matching source ID as its own selected ID before its timer expires, this indicates that another node has assumed the role of a wave source. Hence, the sensor in question cancels its timer. Otherwise, the sensor generates the first wave packet as described in Section II-B.2. In case two waves with the same source ID is present in the network, which may happen in the case of two sensor nodes selecting the same N_{wait} value, ties may be broken by random numbers in the wave packets inserted by wave sources. The tie breaking details are presented in Section II-B.2, as well.

2) Wave Propagation and Region Generation: Our WADS mechanism creates the Wave Mapping Coordinate System (WMC) by measuring the distance to a given node from two randomly placed special nodes called *wave sources*. Once the wave sources are determined as described in Section II-B.1, these sources initiate wave mapping procedure by broadcasting *wave packets*. This packet contains a counter that is incremented every time the message is broadcast by a sensor. A node records and broadcasts a wave message only if it is smaller than its current ID. After the wave propagations from both sources is finalized, all sensor nodes will have an ID pair that designates which region they belong to. A wave packet carries the following information:

- *Source ID (SID):* Indicates the ID of the source. This field can take a value of 1 or 2.
- *Distance to Source (DS):* The number of times the message has been broadcast since its generation. Incremented by one every time the message is broadcast.

[2] N_{max} should scale with the number of sensors in the network.

0) (Initialization) $(\mathrm{ID}_1^i, \mathrm{ID}_2^i) := (\infty, \infty)$, $\mathrm{WS}_1 := \mathrm{WS}_2 := \infty$
   Received Wave Packet: (SID,DS,RSI)
1) IF (RSI $>$ RSI_1) OR ((RSI $==$ RSI_1) AND (DS+1$<$ ID_1^i)) THEN
   a) $\mathrm{RSI}_1 :=$ RSI
   b) $\mathrm{ID}_1^i :=$DS+1
   c) Update the wave packet with DS:$=$ ID_1^i and broadcast
2) ELSE discard the wave packet

Fig. 2
ALGORITHM TO PROCESS WAVE PACKETS

- *Random Source Identifier (RSI):* A random number generated by the wave source which is used to break ties.

Let every sensor node S_i be associated with an ID pair $(\mathrm{ID}_1^i, \mathrm{ID}_2^i)$ that indicates the region they belong to. Originally, all sensors have their ID pair initialized to $(\mathrm{ID}_1^i, \mathrm{ID}_2^i) = (\infty, \infty)$. Furthermore, every sensor also keeps track of the random source identifier RSI_1 and RSI_2 of both wave sources, which are also initialized to infinity. Let us consider a sensor node that identifies itself as wave source WS_1. The wave source WS_1 initiates the wave propagation by creating a wave packet in the format (SID=1, DS=0, RSI). A sensor node S_i receiving a wave packet processes the information only if the RSI in the wave packet is smaller than RSI_1 or RSIs match but the distance to the wave source is smaller than or equal to ID_1^i. If one of these conditions is satisfied, then the S_i sets ID_1^i to ID+1 and RSI_1 to RSI. Then it increments the distance to source field of the wave packet by one and rebroadcasts this information. These steps are formally shown in Figure 2.

The random source identifier is used only to break ties that can occur in the post-deployment wave source selection method. If two or more sources assume the role of the same wave source (1 or 2) and start a wave propagation before others' waves can reach them, then waves of all but one wave source must be suppressed. With the algorithm of Figure 2, the IDs of wave sources that have larger random source identifications are removed from the system. Clearly, there is a non-zero probability that more than one nodes select the same random IDs. However, it is possible to minimize this probability by selecting a number from a larger range. Consequently, the probability of having two sources that start waves at almost the same time and that have the same random source identification can be confined to a satisfactory level.

Once a sensor node S_i obtains an ID pair (ID_1^i,ID_2^i), it identifies itself as belonging to a region with the same ID pair. We define a *region* as a group of neighboring sensor nodes that share the same ID pair and designate it as R_{nm}, where $n = \mathrm{ID}_1$ and $m = \mathrm{ID}_2$. Note that not all nodes in the same region are within each other's direct communication range. However, we assume that all nodes of a region form a connected subgraph. IDs and connectivity of nodes in a region can be determined by creating a local broadcast message that is forwarded only among the nodes of a particular ID pair.

3) Address Ambiguity Elimination: The WADS mechanism generates approximately circular waves of nodes around two wave sources. The regions are formed in the intersection areas of these waves. When we consider two circles with two distinct center points and possibly different radii, these two circles can intersect in zero, one, or two points. Excluding the no intersection case, in the remaining two cases, regions are formed in the WMC system. Obviously, the case of a single intersection region does not create any problems in identifying the region and its members. However, it is likely to have two regions, each of which falls onto one side of the line connecting the two wave sources, that share the same ID pair. If the wave mapping coordinate system is used to identify locations of the sensors, then these two regions must be differentiated from each other. This is accomplished in the address ambiguity elimination phase. Note that this phase is not required if the application or other protocols do not need to differentiate between relative location of regions with same ID pairs.

In this section, we introduce a simple method to assign a third code to each region that will be used to tell two regions with the same ID pair apart. The basic idea behind this algorithm is to create an imaginary line that connect the two wave sources and assign a differentiating code to one side and another to the other side. Although we do not necessarily require a convex sensor field for this algorithm to work, we guarantee its correct operation in all convex sensor fields. This method works without any problems in many concave scenarios, as well. However, it is always possible to engineer special concave sensor fields and wave source placement scenarios under which the described method will not be able to differentiate two regions that share an ID pair.

The first step for address ambiguity elimination is to designate a sensor as the representative of a region. The designated sensor DS_{nm} of a region R_{nm} is selected among the members of a region. Although any leader selection algorithm can be used, we propose to select the sensor node that has the highest number of neighbors of the same ID pair as the designated sensor to reduce the communication overhead in the region. After the designated sensor selection, the coordination messages will be exchanged only between these designated nodes (possibly over multiple hops) rather than randomly between all group members. As a special case, each of the wave sources WS_1 and WS_2 act as the designated sensor of regions R_{1k} and R_{l1}, respectively. Once selected, each designated sensor DS_{nm} generates a *random designated sensor identifier* ($RDSI_{nm}$) to uniquely identify its region. Note that although RDSI can be used to identify the regions uniquely, it is not sufficient to create a neighborhood relationship. As an example, consider two regions R_{nm} with $RDSI_{nm}$ and $R_{(n+1)m}$ with $RDSI_{(n+1)m}$. Since $RDSI_{nm}$ and $RDSI_{(n+1)m}$ are generated independent of each other, one cannot tell if $R_{(n+1)m}$ is the neighbor of R_{nm} or the neighbor of the mirror image of R_{nm}. Hence, we will use RDSIs as temporary indicators that we utilize to associate neighboring regions with correlated codes.

The second step of the address ambiguity elimination is to identify the regions that lie on the line that connects two wave sources. Such regions that have two neighboring regions with the same ID pair but with different random designated sensor identifier are called *border regions*. The border regions of the example of Figure 1 are shown in Figure 3 In this Figure, all shaded areas are adjacent to two regions that have the same ID pair[3]. Note that these regions lie on the line that connects the two wave sources. However, since the location information of the wave sources are not known in a general case, this information cannot be used to determine the location of the border nodes. To discover the identities of neighboring regions, the designated sensor DS_{nm} of a region R_{nm} generates a HELLO packet that contains the ID pair (ID_1=n,ID_2=m) and the random designated sensor identifier $RDSI_{nm}$. The member sensors of region R_{nm} broadcast this message only once. Upon receiving HELLO messages from other regions, sensors of region R_{nm} relay those messages to DS_{nm}. Consequently, DS_{nm} will have a list of all its neighboring regions in the form (ID_1,ID_2, RDSI). A region will be considered *border region* if the neighbor region list, which also includes an entry for itself, contains at least two entries that have the same ID pair but different RDSI values.

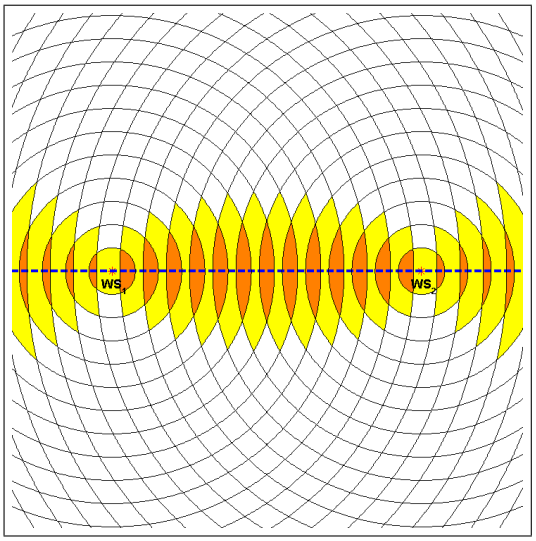


Fig. 3
BORDER REGIONS IN THE WAVE MAPPING COORDINATE SYSTEM

The third step involves dissemination of identifying codes to both sides of the borders. Using this code, one can distinguish if two regions that can potentially be neighbors according to their region ID pairs are in fact on the same side of the border, therefore actual neighbors, or not. These distinguishing codes are *partition identifiers (PID)*. PIDs are generated by messages sent by both wave sources. After these messages are disseminated, each region will have a four bit partition identifier. The first two bits of PID indicate whether or not a message has been received from wave sources WS_1 and WS_2. If a message is received from WS_1, then the value of the third

[3]In Figure 3, the darker border regions share vertexes with the regions with same ID pairs. We assume that messages can be transmitted between these regions and the border regions and, therefore, they are considered neighbors.

PID	Interpretation
00XX	No message received from WS_1 or WS_2
10AX	A message received from WS_1 with code A, but not from WS_2
01XB	A message received from WS_2 with code B, but not from WS_1
11AB	A message received from WS_1 and WS_2, with codes A and B respectively

TABLE I
PARTITION IDENTIFIER COMBINATIONS AND THEIR INTERPRETATIONS

bit contains the code that is assigned by WS_1 to the region. Similarly, if a message is received from WS_2, then the value of the fourth bit contains the code that is assigned by WS_2 to the region. Possible combinations are listed in Table I. In this Table, A and B are binary values and X represents the *"don't care"* condition.

The steps to establish PIDs start after the border regions identify themselves. We will describe the PID distribution generation with an example. Let us consider the first wave source WS_1. In Figure 4, a close-up of the neighborhood of WS_1 of Figure 3 is shown. WS_1, which acts as the designated sensor of region $R_{1,14}$, determines that two neighboring regions share the ID pair (2,14) and but have different RDSIs. Upon determining this, WS_1 sends a PID 1010 to the the designated sensor of one of the regions, and another PID 1000 to the designated sensor of the other region. These designated sensors are responsible for disseminating this PID to all their neighbors by sending a copy of the PID they receive to the designated sensors of their neighbors. The received PID is processed by a designated sensor as follows: All designated sensors keep a PID which is initialized to 0000. An incoming PID from a neighboring region is processed only if the receiving designated sensor does not belong to a border region or if the previous hop of the PID message is not a designated sensor with the same ID pair but different RDSI. Otherwise, the message is discarded. These two steps ensure that PID messages are retained on one side of the border region. Then, the designated sensor updates its PID by applying bitwise OR operation on the new copy received and the PID that it had. Then, it forwards the originally received PID to its neighboring designated sensors to be processed further.

After the PIDs are disseminated in the network, ideally, all nodes on one side of the region should have the same PID value. However, if the sensor field is not convex, then some of the regions may not receive the PID initiated by one of the wave sources. In such a case, neighborhood relationships can be concluded based on the PID received from one of the neighbors. Note that, despite this extra precaution, it is still possible to generate special topologies where a subset of regions do not receive PIDs although they are not border regions.

Two regions R_{ij} with PID=ABCD and R_{kl} with PID=EFGH is determined as neighbors if all three conditions below are satisfied:

1) $\mid i-k \mid \leq 1$

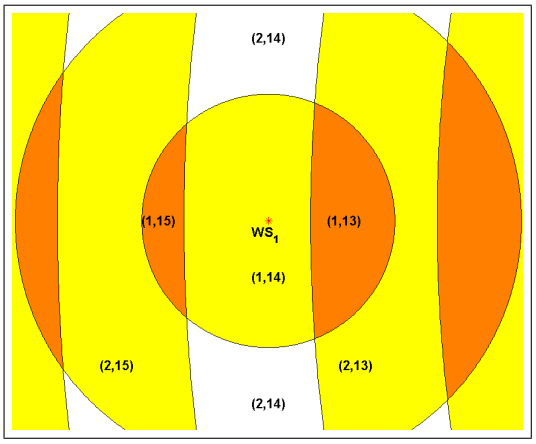


Fig. 4
INITIATION OF THE PID DISTRIBUTION

2) $\mid j-l \mid \leq 1$
3) (ABCD=0000) OR (EFGH=0000)
   OR (A=1 AND E=1 AND C=G)
   OR (B=1 AND F=1 AND D=H)

Using these criteria, the neighborhood of groups can be determined either locally or remotely.

At this point, it is important to note that a third source can also be used to assign a differentiating code to the regions. Although this method is definitely less complicated than the method described in this section, it is also more likely to fail: If the third wave source falls into a border region, then the differentiating codes would be distributed symmetrically, hence failing to actually differentiate regions. If, however, one can guarantee the three wave sources not to be, even roughly, on the same line, three codes generated by three wave sources would uniquely identify the regions.

III. PERFORMANCE EVALUATION

To evaluate the performance of our WADS protocol, we have conducted a set of experiments on random sensor network topologies. In all topologies, we have chosen a communication range of 100m for each sensor node. The sensor nodes are distributed over the sensor fields randomly following 2-D Poisson distribution:

$$P\{N(A)=n\}=\frac{(\rho A)^n}{n!}e^{-\rho A}, \quad (1)$$

where $N(A)$ is the number of nodes in area A, and ρ is the average number of nodes per unit area. In our experiments, we have considered four cases, where ρ takes values of 750, 1000, 1500, and 2000 nodes/km^2. Two types of sensor fields of dimensions 1000m x 1000m and 1000m x 200m are chosen for performance evaluations. These two fields are referred to as *square field* and *corridor field*, respectively. For all cases, two wave source placements are considered: In the distant wave source placement, the wave sources are 960m apart from each other. In the close wave source placement, this distance is reduced to 330m. In both cases, the sensors are placed equally distant from the top and bottom boundaries of the sensor fields.

Parameter	Value
Communication Range	100m
ρ	750, 1000, 1500, 2000 nodes/km^2
Field Dimensions	1000m x 1000m and 1000m x 200m
WS Separation	960m and 330m

TABLE II
SIMULATION PARAMETERS

These parameters are summarized in Table II. Note that these parameter selections result in $4\times2\times2 = 16$ different scenarios. However, only a subset of these results are presented due to space limitations. For each scenario, 100 independent random networks have been generated that share these parameters. The results in this section reflect the averages obtained from those 100 random networks for every combination of the simulation parameters.

A. Sample Wave Mapping Coordinate Systems

In this section, we present example Wave Mapping Coordinate (WMC) systems. The examples shown in Figures 6 and 7 depict the distribution of the regions over the sensor fields for eight cases composed of four different node densities for square and corridor fields. In these figures, the wave sources are shown as asterisks and indicated as WS_1 and WS_2. The regions are generated with our proposed WADS mechanism with wave packets initiated from the two wave sources shown in the figures. Each plot in these figures correspond to the average of 100 independent runs with the same density and shape parameters. Following attributes are calculated for the different scenarios:

Center of Gravity of Regions: For each generated region, the center of gravity is calculated by taking the average of the x and y coordinates of sensors of the same region. For a region R_{nm}, the coordinates of the center of gravity G_{nm} is calculated as

$$\begin{aligned} \overline{x_{nm}} &= \frac{1}{N}\sum_{S_i\in \text{R}_{nm}} x(i) \\ \overline{y_{nm}} &= \frac{1}{N}\sum_{S_i\in \text{R}_{nm}} y(i), \end{aligned} \tag{2}$$

where $\overline{x_{nm}}$ and $\overline{y_{nm}}$ are the x and y coordinates of the center of gravity G_{nm} of the region R_{nm}, N is the number of sensors in the region, and $x(i)$ and $y(i)$ are the coordinates of the location of sensor S_i. The resulting center of gravities for 100 independent runs are grouped according to their ID pairs and their averages are calculated. In Figures 6 and 7, the center of gravity averages are marked with plus signs.

The standard deviation of the distance of center of gravities from wave sources WS_1 and WS_2 are also calculated for all regions generated. If the average of the center of gravity for a region R_{nm} over 100 samples is denoted as $\mathcal{G}_{nm}$, then the standard deviation of center of gravity distances $\sigma_{C1}(nm)$ and $\sigma_{C2}(nm)$ from wave sources WS_1 and WS_2 are calculated as

$$\sigma_{C1}(nm) = \sqrt{\frac{1}{100}\sum_{i=1}^{100}(\mid G_{nm}(i) - \text{WS}_1 \mid - \mid \mathcal{G}_{nm} - \text{WS}_1 \mid)^2}$$

$$\sigma_{C2}(nm) = \sqrt{\frac{1}{100}\sum_{i=1}^{100}(\mid G_{nm}(i) - \text{WS}_2 \mid - \mid \mathcal{G}_{nm} - \text{WS}_2 \mid)^2}, \tag{3}$$

where $\mid G_{nm}(i) - \text{WS}_k \mid$ is the distance of the center of gravity of the region R_{nm} of run i from the wave source WS_k, $k = 1, 2$, and $\mid \mathcal{G}_{nm} - \text{WS}_2 \mid$ is the distance of the average center of gravity $\mathcal{G}_{nm}$ from the wave source WS_k, $k = 1, 2$. The standard deviations $\sigma_{C1}(nm)$ and $\sigma_{C2}(nm)$ are depicted as the semi-minor and semi-major axes of the dashed ellipses around the plus signs.

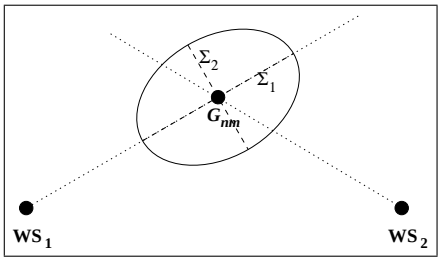


Fig. 5
EXAMPLE ELLIPSE ORIENTATION

Distribution of Sensors in Regions: To assess the distribution of sensors that belong to the same region, we have computed two standard deviations of the sensor distances from wave sources WS_1 and WS_2. For every region R_{nm} in each run, the standard deviations are calculated as

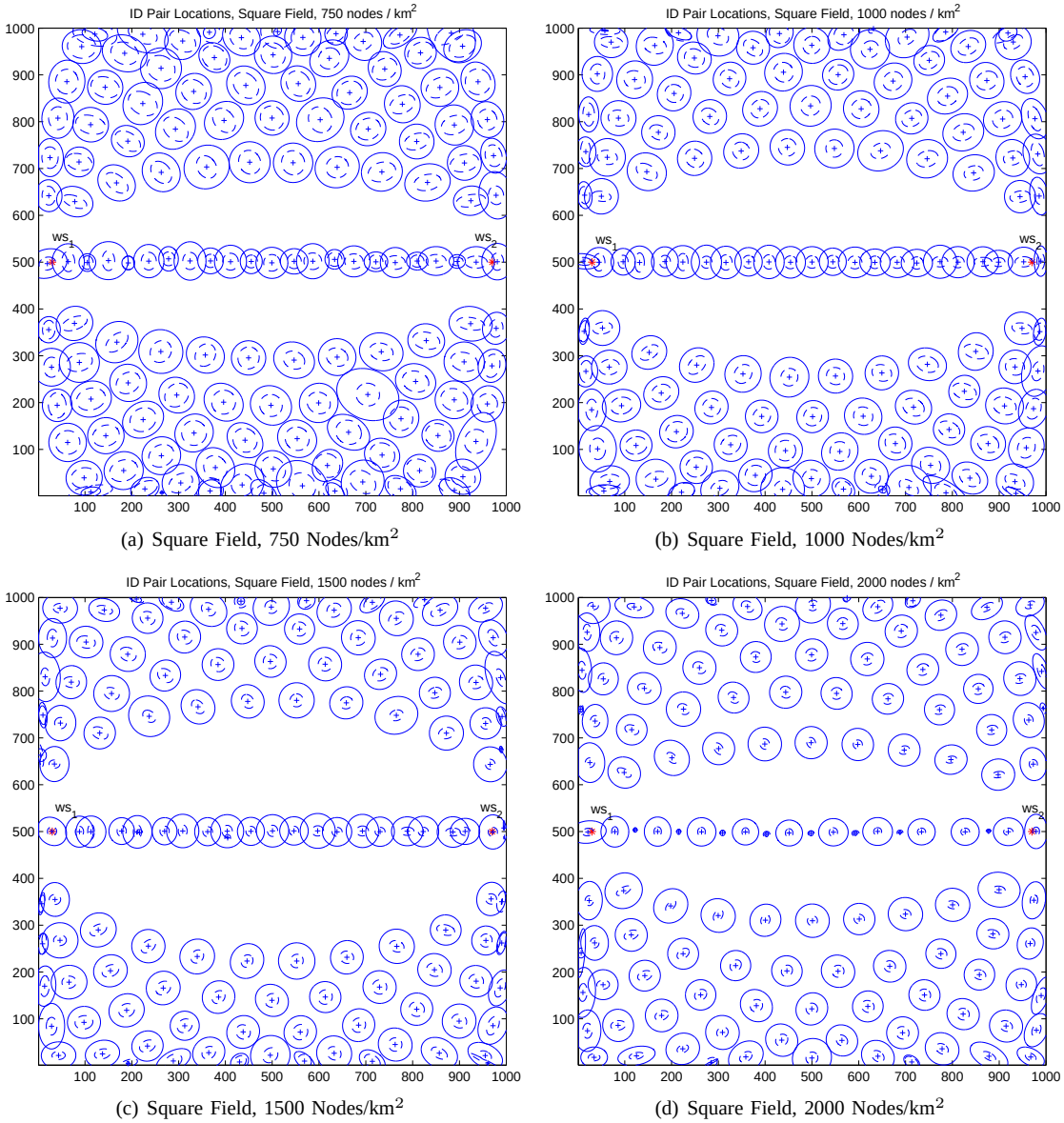
$$\sigma_1(nm) = \sqrt{\frac{1}{N}\sum_{S_i\in \text{R}_{nm}}(\mid S_i - \text{WS}_1 \mid - \mid G_{nm} - \text{WS}_1 \mid)^2}$$

$$\sigma_2(nm) = \sqrt{\frac{1}{N}\sum_{S_i\in \text{R}_{nm}}(\mid S_i - \text{WS}_2 \mid - \mid G_{nm} - \text{WS}_2 \mid)^2}, \tag{4}$$

where $\mid S_i - \text{WS}_k \mid$ is the distance of the sensor S_i from the wave source WS_k, $k = 1, 2$, and $\mid G_{nm} - \text{WS}_2 \mid$ is the distance of the center of gravity G_{nm} of the region R_{nm} from the wave source WS_k, $k = 1, 2$. Consequently, $\sigma_1(nm)$ ($\sigma_2(nm)$) is the standard deviation of the distance of the sensors of region R_{nm} from the wave source WS_1 (WS_2). Once these standard deviations are calculated for individual runs, we average them over 100 independent runs to obtain $\Sigma_1(nm)$ and $\Sigma_2(nm)$ that represent the average standard deviation:

$$\begin{aligned} \Sigma_1(nm) &= \frac{1}{100}\sum_{i=1}^{100}\sigma_1(nm) \\ \Sigma_2(nm) &= \frac{1}{100}\sum_{i=1}^{100}\sigma_2(nm). \end{aligned} \tag{5}$$

These standard deviations are depicted in Figures 6 and 7 as the semi-minor and semi-major axes of solid ellipses around the plus signs.



(a) Square Field, 750 Nodes/km^2

(b) Square Field, 1000 Nodes/km^2

(c) Square Field, 1500 Nodes/km^2

(d) Square Field, 2000 Nodes/km^2

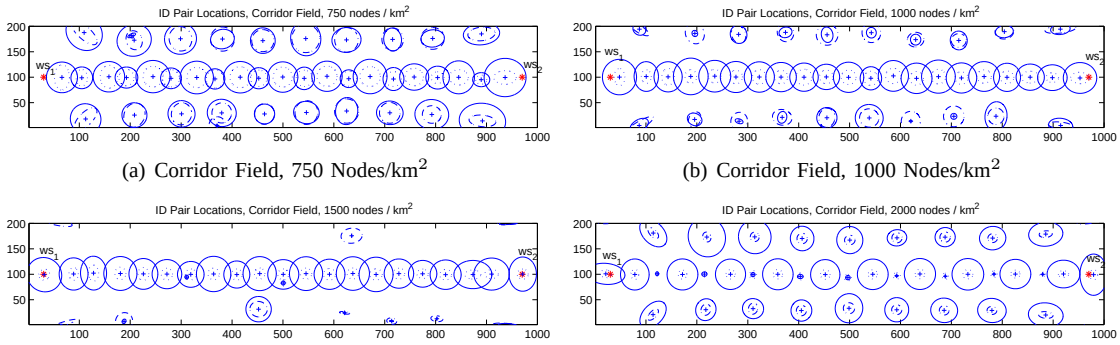
Fig. 6

EXAMPLE WAVE MAPPING COORDINATE SYSTEMS FOR SQUARE FIELD

Both dashed as well as solid ellipses are centered at the plus signs they encircle. The orientation of the ellipses is determined such that axis associated with the standard deviation with respect to the wave source WS_1 falls on the line connecting WS_1 and $\mathcal{G}_1(nm)$. As the other axis must be at a right angle, it does not fall on the line connecting WS_2 and $\mathcal{G}_1(nm)$. This orientation is depicted in Figure 5. The major axis, which is associated with the standard deviation Σ_1, is aligned with the line that connects WS_1 and the center of gravity G_{nm}, whereas the minor axis associated with Σ_2 is not aligned with the line connecting WS_2 and G_{nm}. Note that, if $\Sigma_1 < \Sigma_2$, then the names of the axes will change, i.e., the alignment is done based on WS_1 and not on the major axis.

In Figure 6, the distribution of the center of gravities of regions over a square field is shown for four different node densities. As the density of the sensors increases, the standard deviations of the center of gravities decrease, i.e., we expect to find the center of gravity of the regions more precisely and in a much more smaller area. Furthermore, the standard deviation of the sensor distances to the wave sources decrease with increasing density, which indicates that the regions become better defined and less overlap occurs across the regions. This can be attributed to the fact that as the density increases, the wave packets propagate almost radially away from the wave sources on straight lines. Hence, the formation of regions approaches to the ideal case depicted in Figure 1.

An important property of these figures is the shape of the boundary regions in the middle of the plots. As depicted in Figure 1, the boundary regions in the ideal case are much



(a) Corridor Field, 750 Nodes/km^2

(b) Corridor Field, 1000 Nodes/km^2

(c) Corridor Field, 1500 Nodes/km^2

(d) Corridor Field, 2000 Nodes/km^2

Fig. 7
EXAMPLE WAVE MAPPING COORDINATE SYSTEMS FOR CORRIDOR FIELD

longer in the y-direction than in x-direction. Since the ellipses show the standard deviation in the distance from wave sources (and not the standard deviations in x and y directions), the ellipses drawn have almost circular shapes. The boundary regions, in fact, are centered at the plus signs in the middle of the plots and stretch across the empty spaces along the y-direction. Hence, the unoccupied spaces in these figures do not mean that there are not sensors present in those areas.

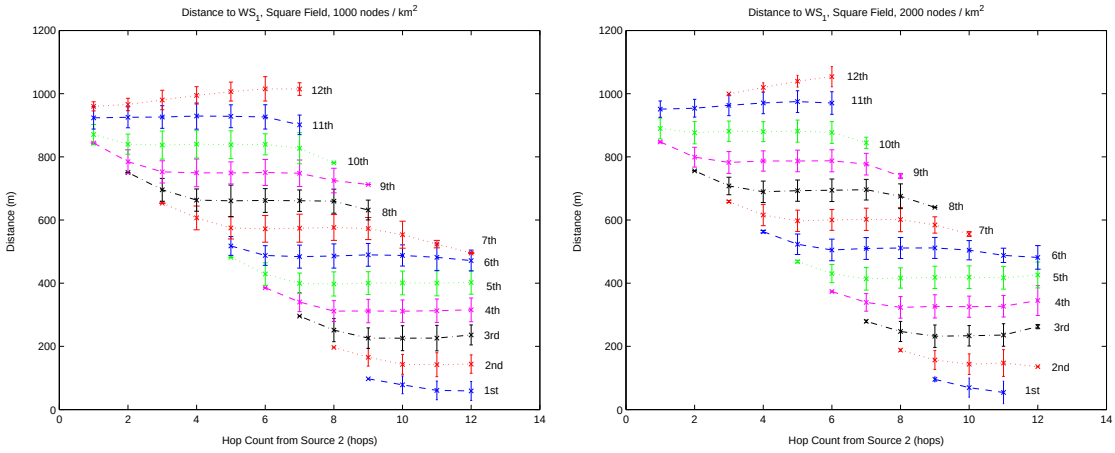
Another important observation can be made about the effect of the edges of the sensor field. In all plots of Figure 6, the regions closer to the sensor field boundaries are not spaced equally as other regions in the middle parts of the field. Overlaps with neighboring regions are more pronounced in these regions. In some cases, the standard deviation of the center of gravities and the sensor nodes of regions become closer, indicating that these regions contain very small number of sensor nodes if/when they exist. Hence, the shape of the regions and the accuracy of the WMC system deteriorate when the sensor field boundaries are approached. This effect is more visible in cases where the sensor density is smaller. In denser sensor networks, although the deviations are larger in regions closer to the field boundaries, the increased overall precision reduces the overlap rates.

In Figure 7, average locations of the centers of gravity, and the associated standard deviations are shown for the same node densities for the corridor field. As in the square field case, the dashed ellipses represent the standard deviation of center of gravity distances from the two wave sources and the solid ellipses represent the standard deviation of the sensor distances from the two wave sources. As in the case of the square field, all standard deviations diminish as the node density increases. Hence, the precision of the regions increases with increasing density. In the corridor field, the sensors are located in a much narrower region. The regions closer to the boundaries occupy a larger percentage of the sensor field. Hence, the uncertainty of the location near the borders have a larger effect. This effect is also observable across sample regions near borders where, in some cases, the standard deviation of the center of gravity is larger than the standard deviation of individual sensors in a region. This anomaly can be attributed to the lack of sufficient points that share that particular ID pair.

When Figures 6 and 7 are compared among themselves, an interesting pattern emerges: As the node density increases, the region layout changes in contrary to expectation that a more precise version of lesser densities would become visible. The change of the patterns is not surprising. As the density increases, average distance traveled by a wave packet increases as will be discussed in Section III-C. These changes directly affect the interference patterns of the topology waves, which in turn changes the locations and distribution of the regions. The variation of the distance traveled may result in a high number of smaller regions as shown in Figure 7(a) or into vertical slices as shown in Figure 7(c).

B. Distance of Region Centers of Gravity from the Wave Sources

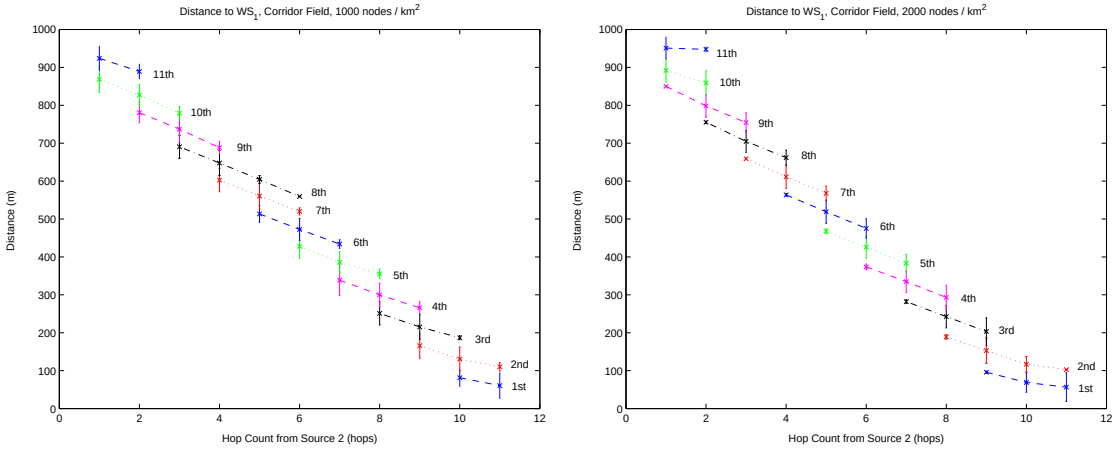
In this section, we analyze the distance of center of gravities of the regions formed by the proposed WADS mechanism from the wave sources. For this purpose, we focus on the distance of the center of gravities from the wave source WS_1 as a function of the second ID of the regions. In Figure 8, these distances are plotted for varying values of ID_2 for two different densities (1000 and 2000 nodes/km^2) in the square fields. The numbers next to the connected lines indicate the first ID (ID_1) of the regions. The x signs mark distance of the average center of gravity of regions from the first wave source. The error bars associated with the data points show the standard deviation of sensor distances from the source WS_1. In both plots, the average center of gravity locations that share the same ID_1 form almost a horizontal line as one might expect. The main reason for the deviations from the horizontal lines is the boundary effects that cause the left and right ends of the curves to curl up. For larger values of ID_1, this effect pushes the center of gravity locations together in a small area as was observed in Section III-A. Consequently, the left ends of the curves have higher values than the average values on the respective curves. The right ends of the curves for $7 \leq ID_1 \leq 11$ in Figure 8(a) have lower values than their respective averages for the same reason. Another important observation is that the standard deviations are much larger in



(a) Square Field, 1000 Nodes/km^2 (b) Square Field, 2000 Nodes/km^2

Fig. 8

DISTANCE OF REGION CENTERS OF GRAVITY FROM WAVE SOURCE WS_1 IN SQUARE FIELD



(a) Corridor Field, 1000 Nodes/km^2 (b) Corridor Field, 2000 Nodes/km^2

Fig. 9

DISTANCE OF REGION CENTERS OF GRAVITY FROM WAVE SOURCE WS_1 IN CORRIDOR FIELD

Figure 8(a) than in Figure 8(b), which can be attributed to the fact that when the sensor density is higher, it is more likely to form the same regions in the same geographic locations in a dense formation.

Figure 9 depicts the region center of gravity distances from WS_1 as a function of ID_2 values for corridor fields and two node densities. When compared with the plots of Figure 8, the corridor fields have much shorter connected lines since the field has limited width. On the other hand, the slopes of the curves are much steeper, and a flat region, which was present in the plots of Figure 8, do not occur for this narrow field. This is to be attributed to the dominance of the edge effects. Since the regions bordering the edges occupy a larger area percentage, the adverse effects of the borders are more clearly visible in these charts. The standard deviations of Figure 9 does not change visibly for these two densities.

C. *Average Wavelengths*

When the wave packets are disseminated in the sensor field, sensors forward only the packets that have traveled a smaller number of hops than seen so far. Hence, when a packet is forwarded in a direction other than radially moving away from the wave source, it is likely that it will be suppressed by another packet that may reach that point over another shorter path. However, since it is not always possible to locate a node very close to the edge of a sensor's transmission range, and because packets may be forwarded on paths that are not perfectly straight, the effective distance traveled by a wave is not equal to the transmission radius. We call this reduced average transmission range *Average Wavelength*. More specifically, the effective wavelength is the average thickness of the ring that contains all regions sharing the same ID_k, where $k = 1, 2$. This analysis is important to understand the

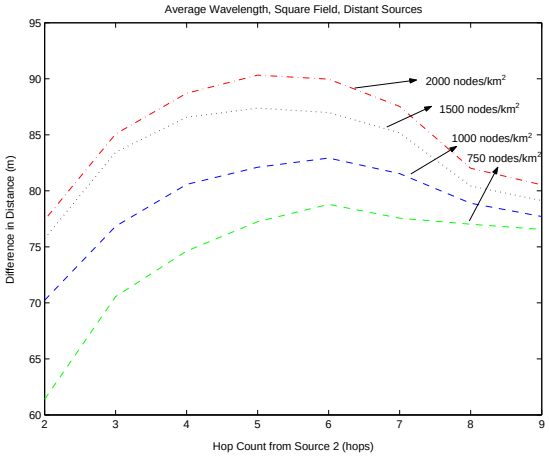
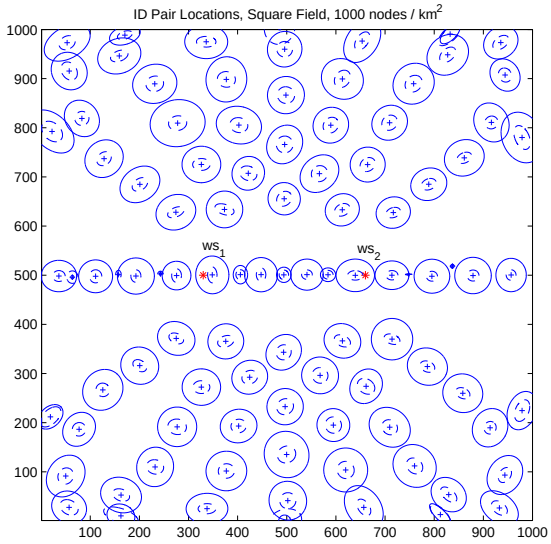
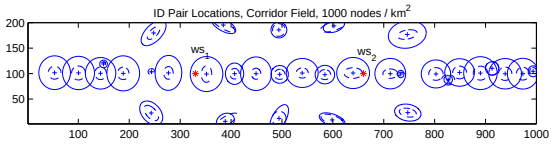


Fig. 10
EFFECTIVE WAVELENGTH OF THE WAVE PACKETS FROM WS_1



(a) Square Field, 1000 Nodes/km^2



(b) Corridor Field, 1000 Nodes/km^2

Fig. 11
EXAMPLE WAVE MAPPING COORDINATE SYSTEMS FOR CLOSE WAVE SOURCES

relationship between the region locations and the density of the nodes in the sensor field.

In Figure 10, the average wavelength of the wave source WS_1 are shown as functions of ID_2 values for square fields, 1000 nodes/km^2 density, and distant wave sources. The values plotted in this figure correspond to the distance a wave packet of WS_1 travels away from WS_1 in a single hop at different locations of the field. As the density increases, wave packets are forwarded on paths closer to a straight line away from the wave source since it is more likely to find next hops in a particular direction in denser topologies. This distance is a function of ID_2, as well: The average wavelength reaches its peak for medium values of the ID_2 range. These correspond to the areas away from the edges as well as from the close neighborhood of the wave source in question. When ID_2 values increase further, edge effects cause the wavelength to become smaller.

D. Effect of the Wave Source Locations

In Section II-B.1, we have claimed that wave sources must be selected far away from each other if at all possible. In this section, we present experimental results that show the effects of closely placed wave sources and compare them with the results presented in the previous sections. The experiments of this section are performed in square and corridor fields of node density 1000 nodes/km^2, where the wave sources WS_1 and WS_2 are only 330m apart. We present the WMC system maps that show the region centers of gravity and the two standard deviation types as in Section III-A.

In Figure 11, the region centers of gravity and associated standard deviations are depicted for square and corridor fields with 1000 nodes/km^2, where the wave sources are 330m away from each other. In order to assess the effect of source locations, we need to compare Figure 11(a) with Figure 6(b) and Figure 11(b) with Figure 7(b). In both comparisons, regions between the wave sources and to the north and south away from the line connecting the wave sources show similarities in shape and layout. However, in locations where the wave packets of WS_1 and WS_2 propagate in parallel, regions of larger sizes are formed. East of WS_2 and west of WS_1 are such locations. As larger-sized regions do not reflect a high level of precision, such occurrences are not very desirable. To avoid such formations, wave sources should be placed closer to the boundaries of the sensor fields and away from each other. This observation may also suggest concatenated use of limited range WMC systems if the deployment area is very large to avoid formation of very large regions. Alternatively, the wave sources can be placed or selected such that they are aligned on one edge of the sensor field and as far away from each other as possible. Placement of wave sources along an edge would also eliminate formation of larger border regions and eliminate the need for the address ambiguity elimination procedures.

IV. CONCLUSIONS

In this paper, we have introduced the Wave Addressing Mechanism for Dense Sensor Networks (WADS) to create a virtual coordinate system in dense sensor networks. The proposed WADS mechanism does not require the use of special equipment such as GPS devices or signal processing capabilities in the sensor nodes. Furthermore, no infrastructure support such as special location beacons are necessary, either. Our WADS mechanism utilizes two randomly selected nodes

called wave sources to disseminate messages that assign ID numbers to the sensors to form the Wave Mapping Coordinate (WMC) System. Sensors that share the same ID pair in their communication radius form the so-called regions that serves as the unit area in the WMC system. We have shown through simulations that WMC systems created with our proposed WADS mechanism effectively serves as consistent and simple coordinate systems without any infrastructure support or complex sensor architectures.

REFERENCES

[1] I. Akyildiz, W. Su, Y. Sankarasubramaninam, and E. Cayirci, "Wireless Sensor Networks: A Survey," *Computer Networks Journal (Elsevier)*, vol. 38, pp. 393–422, March 2002.

[2] G. Pottie and W. Kaiser, "Wireless Integrated Network Sensors," *Communications of the ACM*, vol. 43, pp. 551–558, May 2000.

[3] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed Diffusion for Wireless Sensor Networking," *IEEE/ACM Transactions on Networking*, vol. 11, pp. 2–16, February 2003.

[4] T. He, J. Stankovic, C. Lu, and T. Abdelzaher, "SPEED: A Stateless Protocol for Real-Time Communication in Sensor Networks," *Proceedings of IEEE International Conference on Distributed Computing Systems*, pp. 46–55, 2003.

[5] B. Karp and H. Kung, "Greedy Perimeter Stateless Routing for Wireless Networks," *Proceedings of IEEE/ACM Mobicom 2000*, pp. 243–254, 2000.

[6] F. Ye, H. Luo, J. Cheng, S. Lu, and L. Zhang, "A two-tier data dissemination model for large-scale wireless sensor networks," *Proceedings of ACM Mobicom 2002*, September 2002.

[7] B. Hoffman-Wellendorf, H. Lichtenegger, and J. Collins, *Global Positioning System: Theory and Practice*. Springer Verlag, 4^{th} edition ed., 1997.

[8] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less Low-Cost Outdoor Localization for Very Small Devices," *IEEE Personal Communications*, vol. 7, pp. 28–34, October 2000.

[9] S. Capkun, M. Hamdi, and J. Hubbaux, "GPS-free Positioning in Mobile Ad Hoc Networks," in *Proceedings of the 34th IEEE Hawaii International Conference on System Sciences*, pp. 3481–3490, 2001.

[10] X. Ji and H. Zha, "Robust Sensor Localization Algorithm in Wireless Ad Hoc Sensor Networks," in *Proceedings of IEEE International Conference on Computer and Communication Networks*, pp. 527–532, 2003.

[11] P. Bahl and V. Padmanabhan, "Radar: An In-Building RF-Based User Location and Tracking System," in *Proceedings of IEEE INFOCOM 2000*, vol. 2, pp. 775–784, 200.

[12] D. Niculescu and B. Nath, "Ad Hoc Positioning System (APS) using AoA," in *Proceedings of IEEE INFOCOM 2003*, vol. 3, pp. 1734–1743, 2003.

[13] A. Savvides, C.-C. Han, and M. Srivastava, "Dynamic Fine-Grained Localization in Ad Hoc Networks of Sensors," in *Proceedings of ACM MOBICOM 2001*, pp. 166–179, 2001.

[14] D. Niculescu and B. Nath, "Localized Positioning in Ad Hoc Networks," in *Proceedings of First IEEE International Workshop on Sensor Network Protocols and Applications*, pp. 42–50, 2003.

[15] N. Priyantha, A.Chakraborty, and H. Padmanabhan, "The Cricket Location Support System," in *Proceedings of ACM MOBICOM 2000*, pp. 32–43, 2000.