IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 2, 2018, accepted February 2, 2018, date of publication February 9, 2018, date of current version March 12, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2804381

A Grid-Based Localization Algorithm for Wireless Sensor Networks Using Connectivity and RSS Rank

ZENGFENG WANG[1,2], HAO ZHANG[2,3], (Senior Member, IEEE), TINGTING LU[2], AND T. AARON GULLIVER[3], (Senior Member, IEEE)

[1]School of Information and Electrical Engineering, Ludong University, Yantai 264025, China
[2]Department of Electrical Engineering, Ocean University of China, Qingdao 266100, China
[3]Department of Electrical and Computer Engineering, University of Victoria, Victoria, V8W 2Y2, Canada

Corresponding author: Zengfeng Wang (wzf_ldu@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grants 41527901 and 61602229, in part by the Natural Science Foundation of Shandong Province under Grant ZR2016FM13, in part by the Major Program of China's Second Generation Satellite Navigation System under Grant GF**********03, and in part by Fundamental Research Funds for the Central Universities under Grant 201713018.

ABSTRACT To improve localization accuracy and reduce system costs, a new range-free localization algorithm for wireless sensor networks using connectivity and received signal strength (RSS) rank is proposed. An unknown node is first localized in an initial residence area according to the connectivity constraints. Then, the residence area is refined using the RSS rank vector. The RSS rank vector of an unknown node is formed from the ranks of the RSS values obtained from neighboring anchors. The estimated location of the unknown node is the centroid of the refined residence area. The proposed localization algorithm uses a grid scan approach to avoid the complex geometric computations. This algorithm is improved by using an adaptive strategy to determine the grid size. The performance of these methods is examined via analysis and simulation. The results obtained verify that the proposed approaches provide better localization accuracy than competing algorithms in the literature.

INDEX TERMS Wireless sensor networks, localization, received signal strength, connectivity.

I. INTRODUCTION

Localization is a fundamental and critical issue for wireless sensor networks (WSNs) which has received increasing attention in recent years [1]–[3]. Most WSN applications require the location of the sensor nodes, for example, battlefield monitoring, environment surveillance and object tracking [4]–[6]. Many WSN routing protocols and network management techniques are based on the assumption that the location of each sensor node is available [7]. However, only a small percentage of the sensor nodes, called anchors, obtain their location information via GPS or other sophisticated technologies. The remaining nodes, called unknown nodes, estimate their locations using location information from the anchors.

Many localization algorithms have been proposed for WSNs. These algorithms can be divided into two categories, range-based and range-free. Range-based techniques first measure the distance between two nodes using range information such as the time of arrival (TOA), time difference of arrival (TDOA), angle of arrival (AOA), or received signal strength (RSS) [1]. Then the location estimate is obtained using this information. The estimation accuracy is affected by multipath fading and measurement noise. Further, range-based techniques are complex, for example, AOA measurements require information from multiple antennas. In contrast, range-free approaches localize nodes based on simple sensing information such as wireless connectivity, anchor proximity, and/or event detection. Thus, the system requirements are much lower, and therefore range-free localization is better suited to WSNs due to the hardware limitations of the nodes.

RSS is a popular and widely used neighborhood sensing technique. It is available in many sensor platforms such as Mica2, MicaZ, and TelosB. In range-based techniques, the RSS can be used to estimate the distance between nodes according to the radio propagation model. However, multipath fading, signal propagation variations, and noise can affect the estimation accuracy. Although it is not a good choice for precise ranging measurements, the RSS has been

shown to be an effective metric for range-free localization, especially in outdoor environments [8]–[11]. For instance, the experimental results in [8] indicate that a network-wide monotonic relationship between the RSS and distance does not hold, but there is a per-node monotonic RSS-distance relationship. Thus unknown node RSS sensing results for neighboring nodes can be used as an indication of the distance between them. In addition, it was demonstrated in [12] that the RSS can be used to indicate the distance relationships among nodes in an environment with obstacles. These studies indicate that the rank order of the RSS values consistently reflects the distance relationships between nodes.

In this paper, a new range-free localization algorithm using connectivity and the RSS rank vector (CRRV) is proposed. The RSS rank vector for an unknown node is formed from the ranks of the RSS values sensed from neighboring anchors. This vector can easily be obtained by sorting the RSS values in descending order. An unknown node is first localized in an initial residence area based on the connectivity constraints. Then this area is refined using the RSS rank vector. Last, the centroid of the refined residence area is calculated as the estimated location of the node. A grid scan algorithm is also used to avoid complex geometric calculations. Further, an improved CRRV algorithm called adaptive-CRRV is presented which uses an adaptive strategy to determine the grid size. In adaptive-CRRV, a large grid size is employed if an unknown node has few anchors while a small grid size is used if it has many anchors. Analytic and simulation results are presented which verify that the proposed algorithms can achieve better localization accuracy compared to competing algorithms.

The main contributions of this paper are as follows.

(1) Different from previous RSS-based range-free localization approaches, CRRV constructs an RSS rank vector for each unknown node using all the RSS values sensed from neighboring anchors. An unknown node is then localized using the corresponding RSS rank vector.

(2) To examine the performance of the proposed algorithms, an analytic expression for the mean size of the residence area of an unknown node is obtained.

(3) An adaptive technique is developed to determine the grid size in the grid scan phase. With grid-based localization, there is a tradeoff between localization accuracy and computational cost. Thus, employing an adaptive grid size allows for good localization accuracy at a reasonable cost.

The rest of this paper is organized as follows. Section II gives an overview of related work in the literature. The proposed CRRV algorithm is presented in Section III. In Section IV, the mean size of the residence area is derived and the adaptive-CRRV method is presented. Simulation results are given in Section V to illustrate the localization performance. Finally, some concluding remarks are given in Section VI.

II. RELATED WORK

Range-based localization algorithms determine the node locations using distances calculated between the unknown nodes and anchors [13]–[19]. These solutions can provide excellent accuracy but are not cost-effective for typical WSN deployments.

Because of the severe resource constraints with small, low-cost sensor nodes, many range-free localization algorithms have been proposed [20]–[29]. With the centroid algorithm [20], unknown nodes collect location information from neighboring anchors and determine their locations using the centroid of this data. This algorithm is simple and can provide accurate location estimates when the number of anchors is large and the anchor distribution is uniform. However, when the anchor ratio is low or the distribution is uneven, the estimated locations can be inaccurate.

A distributed range-free localization scheme (DRLS) was proposed in [25]. With DRLS, the location of an unknown node is estimated using both positive and negative connectivity constraints. Each node first calculates an estimation rectangle and then the initial estimated location is obtained using a grid scan algorithm. This location is refined using a vector-based technique with the negative connectivity constraints. DRLS can provide good accuracy when the signal behavior is regular. However, for an irregular model (such as in multipath fading), this negative information can lead to poor results. Consider the example shown in Fig. 1 with anchors A_1 and A_2, and unknown node N. The area marked with solid lines represents the actual communication region of the anchor, while the dotted circle indicates the ideal region. Thus, node N can hear from anchor A_1, and its initial residence area is determined as a circle C_1 centered at A_1 with radius given by the ideal communication range. Node N discovers the two-hop anchor A_2, and uses this negative information to refine its residence area by discarding the intersection region of circles C_1 and C_2, where C_2 is a circle centered at A_2 with the same radius as C_1. It is clear that N is outside the refined residence area, which is undesirable. Further, DRLS has a high communication overhead as each node has to gather information via two-hop flooding to obtain the negative constraints.

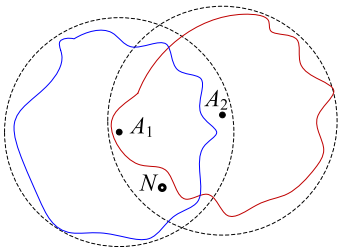

FIGURE 1. The actual and ideal communication range regions with anchors A_1 and A_2, and an unknown node N.

A range-free localization algorithm based on a half-symmetric lens (HSL) was proposed in [28]. A half-symmetric lens can be obtained using the location

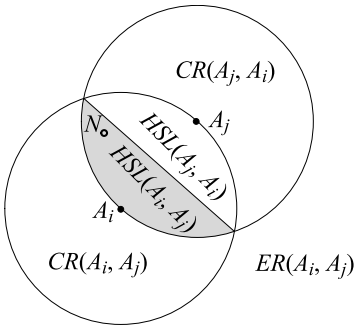

FIGURE 2. Sub-regions of half-symmetric lens presence test.

information of two anchors. For any two neighboring anchors A_i and A_j, HSL draws two circles C_{Ai} and C_{Aj} centered at A_i and A_j, respectively, with the same radius equal to the distance between A_i and A_j. The intersection area of these circles has a symmetric lens (SL) shape. The perpendicular bisector of line segment $\overline{A_iA_j}$ divides the SL into two half symmetric lens, $HSL(A_i, A_j)$ and $HSL(A_j, A_i)$, as shown in Fig. 2. Let d_{AB} denotes the distance between nodes A and B. If $d_{A_iN} < d_{A_iA_j}, d_{A_jN} < d_{A_iA_j}$ and $d_{A_iN} < d_{A_jN}$, then N must be in $HSL(A_i, A_j)$. The HSL algorithm uses RSS values to indicate the distances between nodes. If a node N lies within $HSL(A_i, A_j)$, then $RSS_{A_iN} > RSS_{A_iA_j}, RSS_{A_jN} > RSS_{A_iA_j}$, and $RSS_{A_iN} > RSS_{A_jN}$. Therefore, $HSL(A_i, A_j)$ is the estimated residence area of N from these RSS relationships. This process is referred as a half-symmetric lens presence test of neighboring anchors A_i and A_j for the unknown node N. Different RSS relationships result in different estimated sub-regions. The other possible sub-regions are $CR(A_i, A_j)$, $CR(A_j, A_i)$ and $ER(A_i, A_j)$, as shown in Fig. 2. Pseudo code for the HSL presence test is given in Algorithm I. The inputs are the locations of two neighboring anchors and the measured RSS values. The output is the residence sub-region of the unknown node.

With the HSL algorithm, the network is first divided into cells based on a Voronoi diagram. The initial residence area of an unknown node is the Voronoi cell of the nearest anchor, and this area is refined using half-symmetric lens presence tests with all combinations of pairs of neighboring anchors. While the HSL algorithm can achieve good localization accuracy, it involves many complex geometric computations and so is not well suited to resource constrained nodes. A grid scan algorithm was proposed in [28] to provide an approximate residence area, but determining the Voronoi diagram requires significant computations.

In summary, existing range-free localization schemes for WSNs are prone to errors and have high computational complexity. In the next section, a new range-free localization algorithm is proposed to overcome these limitations.

III. LOCALIZATION USING CONNECTIVITY AND THE RSS RANK VECTOR

In this section, a new localization algorithm is presented which is based on connectivity and RSS rank vectors. It is designed for WSNs with sensor nodes that are randomly deployed in a region. Every sensor node is assumed to have a unique ID. To illustrate the principle of localization using RSS rank vectors, the distance rank vectors are defined in the next section.

Algorithm 1 HSL Presence Test

```
init: A_i and A_j, two neighboring anchors of unknown
  node N, the RSS values are given.
if RSS_{A_iN} > RSS_{A_iA_j} and RSS_{A_jN} > RSS_{A_iA_j} then
    if RSS_{A_iN} > RSS_{A_jN} then
        N is in HSL(A_i, A_j)
    else
        N is in HSL(A_j, A_i)
    end if
else
    if RSS_{A_iN} > RSS_{A_iA_j} then
        N is in CR(A_i, A_j)
    else
        if RSS_{A_jN} > RSS_{A_iA_j} then
            N is in CR(A_j, A_i)
        end if
    else
        N is in ER(A_i, A_j)
end if
```

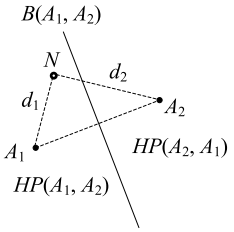

FIGURE 3. The perpendicular bisector of the line segment $\overline{A_1A_2}$,and the distances for unknown node N.

A. DISTANCE RANK VECTORS

Consider an unknown node N with two neighboring anchors A_1 and A_2. The perpendicular bisector of the line segment $\overline{A_1A_2}$, denoted $B(A_1, A_2)$, divides the plane into two half-planes $HP(A_1, A_2)$ and $HP(A_2, A_1)$ containing A_1 and A_2, respectively, as shown in Fig. 3. Let d_1 and d_2 be the distances between N and A_1 and A_2, respectively. The half plane in which N lies in can be determined using the relationship between d_1 and d_2: (1) if $d_1 < d_2, N$ is in $HP(A_1, A_2)$ and (2) if $d_2 < d_1, N$ is in $HP(A_2, A_1)$. Similarly, if node N has m neighboring anchors, $m(m-1)/2$ perpendicular bisectors can be drawn for all combinations of pairs of anchors. Each bisector divides the WSN space into two half-spaces. Let $d_i, 1 \le i \le m$, be the distances between N and A_i. By comparing each pair of distances, a half plane within which N may lie can be determined. The intersection of all $m(m-1)/2$ half planes can be estimated as the residence area of N.

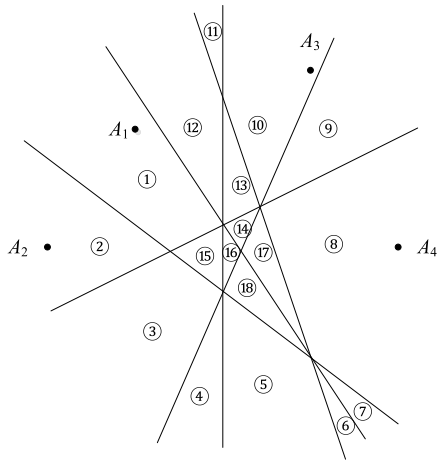

FIGURE 4. The sub-regions obtained from four neighboring anchors.

In fact, the $m(m-1)/2$ perpendicular bisectors divide the WSN space into a number of sub-regions. For example, in Fig. 4, the perpendicular bisectors of the four anchors divide the space into 18 sub-regions. An unknown node can be directly localized into one of these sub-regions by defining a distance rank vector instead of determining all the half planes containing this node individually.

For a specific location, the distance rank vector is defined as a vector whose elements are the ranks of the distances between it and the neighboring anchors. Specifically, let $R'_i = (r'_{i1}, r'_{i2}, \ldots, r'_{im})$ be the distance rank vector of a location C in the ith sub-region divided by the perpendicular bisectors of m neighboring anchors $A_1, A_2, \ldots, A_m$. The jth element r'_{ij} of R'_i denotes the rank of d_{ij} in the sorted distance vector D', where d_{ij} is the distance between C and anchor A_j and D' is constructed by arranging the m distances, $d_{i1}, d_{i2}, \ldots, d_{im}$, in ascending order. Obviously, all the locations in the ith sub-region have the same distance rank vector. Therefore, $R'_i = (r'_{i1}, r'_{i2}, \ldots, r'_{im})$ can also be considered as the distance rank vector of the ith sub-region. Table 1 lists the distance rank vectors corresponding to the sub-regions in Fig. 4.

TABLE 1. Distance rank vectors for the sub-regions in Fig. 4.

Sub-region	②	②	③	④	⑤
Distance rank vector	(1,2,3,4)	(2,1,3,4)	(2,1,4,3)	(2,4,1,3)	(4,2,1,3)
Sub-region	⑥	⑦	⑧	⑨	⑩
Distance rank vector	(4,2,3,1)	(4,3,2,1)	(4,3,1,2)	(3,4,1,2)	(3,1,4,2)
Sub-region	⑪	⑫	⑬	⑭	⑮
Distance rank vector	(3,1,2,4)	(1,3,2,4)	(1,3,4,2)	(1,4,3,2)	(1,2,4,3)
Sub-region	⑯	⑰	⑱		
Distance rank vector	(1,4,2,3)	(4,1,3,2)	(4,1,2,3)		

An m-dimensional distance rank vector, $R'_i = (r'_{i1}, r'_{i2}, \ldots, r'_{im})$, is one of the permutations of the integers $1, 2, \ldots, m$. There are $m!$ possible permutations. However, because of the geometric constraints of the plane, not all permutations can correspond to a sub-region.

Theorem 1: Let λ_m be the maximum number of sub-regions for m neighboring anchors. Then λ_m is given by

$$\lambda_m = \frac{1}{24}(3m^4 - 10m^3 + 21m^2 - 14m + 24). \quad (1)$$

The proof of Theorem 1 is given in the appendix.

For m anchor locations, the number of sub-regions divided by the corresponding perpendicular bisectors is less than or equal to λ_m. Each sub-region has a unique distance rank vector. Therefore, there are at most λ_m feasible distance rank vectors. The other rank vectors are infeasible.

For localization, an unknown node first constructs its distance rank vector and then the sub-region corresponding to this vector is found. RSS measurements are used to obtain the distance information between nodes. In the following section, the RSS rank vectors are first defined, and then unknown node localization using these vectors is presented.

B. RSS RANK VECTORS AND LOCALIZATION USING RSS RANK VECTORS

1) RSS RANK VECTORS

If a node has m neighboring anchors $A_1, A_2, \ldots, A_m$, then the RSS sensing results form a vector, $S = (s_1, s_2, \ldots, s_m)$, where $s_i, 1 \leq i \leq m$, is the RSS value corresponding to anchor A_i. A new vector S' can be obtained by sorting the elements of S in descending order. The RSS rank vector is an m dimensional vector $R = (r_1, r_2, ldots, r_m)$ whose ith element r_i is the rank order of s_i in S'. For example, consider an unknown node with four neighboring anchors A_1, A_2, A_3 and A_4 and RSS vector $S = (16, 35, 9, 22)$. Sorting the elements of S in descending order gives $S' = (35, 22, 16, 9)$. Comparing S with S', the corresponding RSS rank vector is $R = (3, 1, 4, 2)$. If $s_i = s_j, i \neq j$, then the ordering of s_i and s_j is done lexicographically. However, this situation never occurred during the simulations.

2) LOCALIZATION USING RSS RANK VECTORS

In the ideal case, the RSS values decrease monotonically with distance. From the definition of the RSS rank vector, it can be regarded as the distance rank vector of the unknown node and be used for localization directly. In reality, RSS measurements are subject to errors due to channel variations such as multipath fading and shadowing. The RSS rank vector can therefore be regarded as a corrupted distance rank vector. Compared with localization methods that directly use the RSS values, the proposed range-free localization using RSS rank vectors is robust to RSS measurement errors. The reasons for this are given below.

First, consider that the RSS measurement errors are small and the RSS ranks are not changed. Then the unknown node RSS rank vector is still the same as the distance rank vector of the sub-region in which it lies. Hence, the unknown node can be localized in the correct sub-region.

Now consider that the RSS measurement error is large so that the RSS rank vector is changed either to another

feasible distance rank vector or to an infeasible distance rank vector. According to the previous analysis, the number of feasible distance rank vectors for m neighboring anchors is $O(m^4)$ whereas the number of infeasible vectors is $O(m^m)$. Therefore, when m is large, the probability of an RSS rank vector being changed to an infeasible distance rank vector is much higher than to another feasible distance rank vector. Fortunately, when the RSS rank vector is infeasible, it can be detected and corrected. For example, assume an unknown node connected to the four anchors in Fig. 4 is located in sub-region 4. According to Table 1, the distance rank vector of sub-region 4 is $(2, 4, 1, 3)$. Assume that the actual RSS rank vector is $(2, 4, 3, 1)$. Obviously, this is an infeasible vector, so the unknown node should be localized in the sub-region whose distance rank vector is closest to this vector. The Spearman rank order correlation coefficient is introduced in the next section to measure the similarity between vectors. For the example, it can be seen from Table 1 that the distance rank vector of sub-region 4 is closest to the unknown node RSS rank vector. Therefore, the unknown node can still be correctly localized using the erroneous RSS rank vector. On the other hand, the errors cannot be detected if the erroneous RSS rank vector is a feasible distance vector. For example, consider an unknown node located in sub-region 4 in Fig. 4 with RSS rank vector $(2, 1, 4, 3)$, which corresponds to sub-region 3. However, this sub-region is close to the correct sub-region, and this can be expected if the RSS measurement errors are reasonable.

C. ALGORITHM DESCRIPTION

To perform localization, each anchor first broadcasts a beacon message containing its coordinates and ID. If unknown node N receives the beacon signal of anchor A_i, then it is within range of A_i and so A_i is added to its neighboring anchors list denoted by AL_N. Each row in AL_N includes the following information: (1) anchor ID, (2) anchor coordinates, and (3) RSS values corresponding to the received beacon messages from the anchor.

In the CRRV algorithm, an unknown node N is first localized in an initial residence area based on the connectivity constraints. Then this area is refined using the RSS rank vector. Finally, the centroid of the refined residence area is the estimate of the location of the node. Fig. 5 provides an example of this algorithm for an unknown node N with three neighboring anchors A_1, A_2, and A_3. The anchor regions are circles with radius r where r is the communication range of the nodes. The initial residence area obtained using the connectivity constraints, denoted as RA_C, is the overlap of the three circles, and is shown as the shaded area in Fig. 5(a). The residence area of N from the ideal RSS rank vector, denoted as RA_RRV, is the shaded area in Fig. 5(b). Combining the RSS rank vector constraints with the connectivity constraints gives the refined residence area, RA_CRRV, as shown in Fig. 5(c). This is the intersection of RA_C and RA_RRV. The centroid of RA_CRRV, denoted N', is the estimated location of node N.

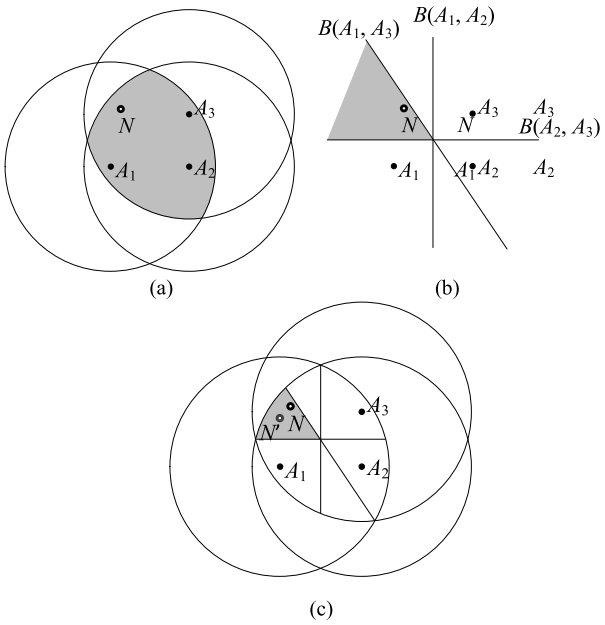

FIGURE 5. The CRRV algorithm (a) residence area using connectivity, RA_C, (b) residence area using the RSS rank vector, RA_RRV, and (c) refined residence area using both connectivity and RSS rank vector, RA_CRRV.

It can be seen from Fig. 5(c) that the estimated location is very close to the actual position of the unknown node. The key question is how to find and describe the residence area obtained from the RSS rank vector and connectivity constraints. For the RSS rank vector, one way to describe the possible sub-regions is to construct a distance rank vector table which includes all feasible distance rank vectors and the vertices of the corresponding sub-regions. To find the sub-region corresponding to an RSS rank vector, an unknown node just needs to search this table. However, the computational complexity and storage requirements for this table are large. Given the locations of the m anchors, it takes $O(m^5 \log_2 m)$ worst-case time and $O(m^5)$ worst-case space to construct the distance rank vector table [30]. Moreover, the connectivity constraint is quadratic. Determining the edges of RA_C is much more complicated. Therefore, a grid-based algorithm is employed here which is suitable for resource constrained nodes.

1) ESTIMATION RECTANGLE

To facilitate the grid-based localization algorithm, the region for an anchor node is first defined as the square that bounds its communication region, as shown in Fig. 6. For simplicity, the sides of these squares are parallel to the axes. Furthermore, for an unknown node, the intersection of the square regions of the neighboring anchors is defined as the estimation rectangle (ER) [25]. Fig. 6 shows that the ER of an unknown node is indeed a rectangle. Compared with RA_C, ER is simple to calculate and divide into grids. Note that ER contains RA_C and is larger than RA_C.

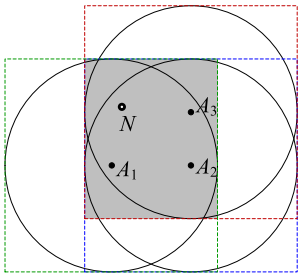

FIGURE 6. The square communication regions of the anchors and ER for unknown node *N*.

2) INITIAL RESIDENCE AREA USING CONNECTIVITY

Assume that an unknown node N has m neighboring anchors, $A_1, A_2, \ldots, A_m$. To obtain the initial residence area using connectivity, first compute the ER and then divide it into square-shaped grids $G = \{G_1, G_2, \ldots, G_k\}$ according to a predefined grid size. As mentioned before, ER is larger than RA_C. Therefore, only part of these grids belongs to RA_C. In order to determine which grids are located in RA_C, the distances between the center of each grid and the neighboring anchors are calculated. Let d_{ij} be the distance between the center of G_i and the neighboring anchor A_j. Let $d_{il}, l \in \{1, 2, \ldots, m\}$ be the largest distance. If $d_{il} > r$, then the center of G_i is not in RA_C so the corresponding grid G_i can be removed from the residence area. These grids are marked with value -2. Otherwise, the center of G_i is in RA_C and G_i is kept. The remaining (valid) grids form an approximation of RA_C.

3) RESIDENCE AREA REFINEMENT USING THE RSS RANK VECTOR

The initial residence area obtained using connectivity is refined using the RSS rank vector. In order to determine which grids are located in RA_RRV, the distance rank vector $R'_i = (r'_{i1}, r'_{i2}, \ldots, r'_{im})$ is computed for each grid G_i. Since the location of G_i is known, its distance rank vector can be computed directly. Assume R is the ideal RSS rank vector of unknown node N. If the center of a grid G_i lies within RA_RRV, its distance rank vector R'_i is the same as the ideal RSS rank vector R. Comparing R'_i with R for the valid grids, the grids belonging to both RA_C and RA_RRV can be determined, and these are marked with value 1. These grids define the refined residence area.

In the ideal case, as long as the grid size is small enough, there always exist grids whose distance rank vector is the same as the RSS rank vector. However, as discussed previously, the actual RSS rank vectors contain errors. In this case, those grids whose distance vectors are close to the corrupted RSS rank vector are considered as the refined residence area. For this purpose, the Spearman rank order correlation coefficient [31] is used to measure the similarity between two rank vectors.

The Spearman rank order correlation coefficient of two m dimensional rank vectors,R and R' is given by

$$\rho = 1 - \frac{6\sum_{i=1}^{m}(r_i - r'_i)^2}{m(m^2-1)}, \tag{2}$$

where r_i and r'_i are the ith elements of R and R', respectively. The range of ρ is $[-1,1]$. The larger ρ is, the greater the correlation between R' and R. ρ is 1 only when R' is identical to R.

For each valid grid, the Spearman rank order correlation coefficient of the distance rank vector and the unknown node RSS rank vector is calculated and added to the grid value. Then, each grid has a value in the range $[-2,1]$. Fig. 7 illustrates the grid scan results. The grids with the largest values define the refined residence area of node N, as shown by the shaded area in Fig. 7.

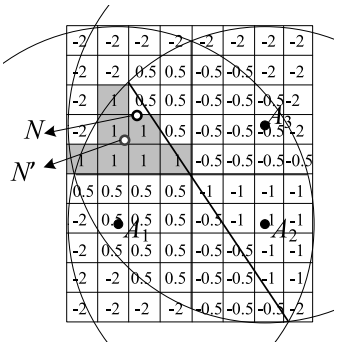

FIGURE 7. The CRRV grid scan results.

Assume that the refined residence area consists of L grids. Then the unknown node coordinates are estimated as

$$N(x, y) = \left(\frac{1}{L}\sum_{i=1}^{L} x_i, \frac{1}{L}\sum_{i=1}^{L} y_i\right),$$

where (x_i, y_i) is the coordinates of the center of the ith grid in the refined residence area.

Fig. 7 shows that the combination of the grids with the largest value (refined residence area) forms an approximation of RA_CRRV. Obviously, reducing the grid size can provide a more precise approximation of RA_CRRV, and hence improve the localization accuracy. However, a small grid size also increases the computational complexity and storage requirements. Therefore, it is crucial to choose an appropriate grid size. An adaptive strategy for determining the grid size is presented in the next section.

IV. PERFORMANCE ANALYSIS AND AN IMPROVED ALGORITHM

To analyze the performance of the CRRV localization algorithm, the average size of the residence area is determined. Obviously, the smaller the size of the residence area, the better the localization accuracy.

A. THE AVERAGE SIZE OF THE RESIDENCE AREA

Assume that an unknown node N has m neighboring anchors at locations $A_1(x_1, y_1), A_2(x_2, y_2), \ldots, A_m(x_m, y_m)$. Let $B_i = [x_i - r, x_i + r] \times [y_i - r, y_i + r]$, $i = 1, 2, \ldots, m$, denote the square regions of A_i, where r is the corresponding communication range. The estimation rectangle (ER) of N is then

given by

$$\mathrm{ER} = Q \cap \bigcap_{i=1}^{m} B_i = [\max(x_+ - r, 0), \min(x_- + r, L)] \times [\max(y_+ - r, 0), \min(y_- + r, L)],$$

where $Q = [0, L] \times [0, L]$ is the operating region of the WSN, $x_+ = \max(x_1, x_2, \ldots, x_m)$, $x_- = \min(x_1, x_2, \ldots, x_m)$, $y_+ = \max(y_1, y_2, \ldots, y_m)$ and $y_- = \min(y_1, y_2, \ldots, y_m)$.

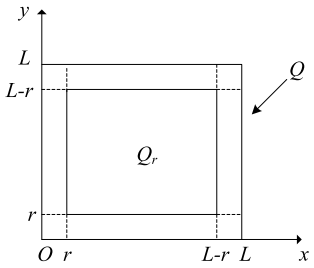

FIGURE 8. The WSN operating region.

Let $Q_r = [r, L - r] \times [r, L - r]$. It can be observed from Fig. 8 that Q_r occupies most of Q since r smaller than L. Let N be a randomly chosen node in Q_r and S_{ER} the corresponding ER area. Then S_{ER} is a random variable given by

$$S_{\mathrm{ER}} = ((X_- + r) - (X_+ - r))((Y_- + r) - (Y_+ - r)), \quad (3)$$

where $X_+ = \max(X_1, X_2, \ldots, X_m)$, $X_- = \min(X_1, X_2, \ldots, X_m)$, $Y_+ = \max(Y_1, Y_2, \ldots, Y_m)$, and $Y_- = \min(Y_1, Y_2, \ldots, Y_m)$ are four random variables.

To determine the average ER size, the expected value of S_{ER} must be obtained. Let the coordinates of N be (x_0, y_0). For simplicity, the neighboring anchors are assumed to be uniformly and independently distributed in the square region $[x_0 - r, x_0 + r] \times [y_0 - r, y_0 + r]$ so that X_i and Y_i are independent. X_i is uniformly distributed in $[x_0 - r, x_0 + r]$ and Y_i is uniformly distributed in $[y_0 - r, y_0 + r]$. Therefore

$$\begin{aligned} E[S_{\mathrm{ER}}] &= E[((X_- + r) - (X_+ - r))((Y_- + r) - (Y_+ - r))] \\ &= E[(X_- + r) - (X_+ - r)]E[(Y_- + r) - (Y_+ - r)] \\ &= (E[X_-] - E[X_+] + 2r)(E[Y_-] - E[Y_+] + 2r) \end{aligned} \quad (4)$$

Next, the expected values of X_+, X_-, Y_+, Y_- are determined. For X_+

$$\begin{aligned} &P\{X_+ \le x\} \\ &= P\{\max(X_1, X_2, \ldots, X_m) \le x\} \\ &= P\{X_1 \le x\}P\{X_2 \le x\} \cdots P\{X_m \le x\} \\ &= \begin{cases} 0, & x < x_0 - r \\ \dfrac{1}{(2r)^m}(x - x_0 + r)^m, & x_0 - r \le x \le x_0 + r \\ 1, & x > x_0 + r \end{cases} \end{aligned} \quad (5)$$

The probability density function (pdf) of X_+ is

$$\begin{aligned} f_{X_+}(x) &= \frac{dP\{X_+ \le x\}}{dx} \\ &= \begin{cases} \dfrac{m}{(2r)^m}(x - x_0 + r)^{m-1}, & x_0 - r \le x \le x_0 + r \\ 0, \text{ otherwise} \end{cases} \end{aligned} \quad (6)$$

So the expected value of X_+ is

$$\begin{aligned} E[X_+] &= \int_{-\infty}^{\infty} x f_{X_+}(x)dx \\ &= \int_{x_0 - r}^{x_0 + r} x \frac{m}{(2r)^m}(x - x_0 + r)^{m-1}dx \\ &= x_0 + \frac{m-1}{m+1}r \end{aligned} \quad (7)$$

For X_-, we have

$$\begin{aligned} &P\{X_- > x\} \\ &= P\{\min(X_1, X_2, \ldots, X_m) > x\} \\ &= P\{X_1 > x\}P\{X_2 > x\} \cdots P\{X_m > x\} \\ &= \begin{cases} 1, & x < x_0 - r \\ \dfrac{1}{(2r)^m}(x_0 + r - x)^m, & x_0 - r \le x \le x_0 + r \\ 0, & x > x_0 + r \end{cases} \end{aligned} \quad (8)$$

and therefore

$$\begin{aligned} &P\{X_- \le x\} \\ &= 1 - P\{X_- > x\} \\ &= \begin{cases} 0, & x < x_0 - r \\ 1 - \dfrac{1}{(2r)^m}(x_0 + r - x)^m, & x_0 - r \le x \le x_0 + r \\ 1, & x > x_0 + r \end{cases} \end{aligned} \quad (9)$$

The pdf of X_- is

$$\begin{aligned} f_{X_-}(x) &= \frac{dP\{X_- \le x\}}{dx} \\ &= \begin{cases} \dfrac{m}{(2r)^m}(x_0 + r - x)^{m-1}, & x_0 - r \le x \le x_0 + r \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (10)$$

and

$$\begin{aligned} E[X_-] &= \int_{-\infty}^{\infty} x f_{X_-}(x)dx \\ &= \int_{x_0 - r}^{x_0 + r} x \frac{m}{(2r)^m}(x_0 + r - x)^{m-1}dx \\ &= x_0 - \frac{m-1}{m+1}r \end{aligned} \quad (11)$$

Similarly, the expected values of Y_+ and Y are

$$E[Y_+] = y_0 + \frac{m-1}{m+1}r, E[Y_-] = y_0 - \frac{m-1}{m+1}r. \quad (12)$$

Substituting (7), (11) and (12) in (4) gives

$$\begin{aligned} E[S_{\mathrm{ER}}] &= (E[X_-] - E[X_+] + 2r)(E[Y_-] - E[Y_+] + 2r) \\ &= \left(\left(x_0 - \frac{m-1}{m+1}r\right) - \left(x_0 + \frac{m-1}{m+1}r\right) + 2r\right) \\ &\quad \cdot \left(\left(y_0 - \frac{m-1}{m+1}r\right) - \left(y_0 + \frac{m-1}{m+1}r\right) + 2r\right) \\ &= \left(\frac{4r}{m+1}\right)^2. \end{aligned} \tag{13}$$

Thus, the expected value of S_{ER} is determined by the communication range r and the number of neighboring anchors m, and is independent of the unknown node coordinates x_0 and y_0. If $m = 1$, this expected value is $4r^2$, which is the area of the square communication region of an anchor. The ER area can be approximated by the average size of the residence area obtained from the connectivity constraints, i.e. $\overline{S_{RA\_C}} \approx E[S_{ER}]$. Table 2 gives the average ER size for different numbers of anchors.

TABLE 2. The average sizes of RA_C and RA_CRRV

	m=1	m=2	m=3	m=4	m=5
$\overline{S_{RA\_C}} \approx E[S_{ER}]$ (r^2)	4	1.78	1	0.64	0.44
$\overline{S_{RA\_CRRV}}_{\min}$ (r^2)	4	0.89	0.17	0.036	0.0097
	m=6	m=7	m=8	m=9	
$\overline{S_{RA\_C}} \approx E[S_{ER}]$ (r^2)	0.33	0.25	0.20	0.16	
$\overline{S_{RA\_CRRV}}_{\min}$ (r^2)	0.0032	0.0013	0.00056	0.00027	

In the proposed localization algorithm, RA_C is divided into sub-regions according to the distance rank vector constraints. From Theorem 1, the maximum number of sub-regions with m-dimensional distance rank vectors is λ_m. Thus, the minimum average size of the residence area with the proposed method can be approximated as

$$\begin{aligned} \overline{S_{RA\_CRRV}}_{\min} &= \frac{\overline{S_{RA\_C}}}{\lambda_m} \approx \frac{E[S_{ER}]}{\lambda_m} \\ &= \frac{384r^2}{3m^6 - 4m^5 + 4m^4 + 18m^3 + 17m^2 + 34m + 24}. \end{aligned} \tag{14}$$

Comparing (13) and (14), it can be concluded that $\overline{S_{RA\_CRRV}}_{\min}$ is much smaller than $E[S_{\mathrm{ER}}]$, which means that the refinement using the RSS rank vector significantly improves the localization accuracy. Table 2 also gives $\overline{S_{RA\_CRRV}}_{\min}$ for different numbers of anchors.

B. ADAPTIVE GRID SIZE

As mentioned in Section III, the grid size is an important parameter for grid-based algorithms. In this section, an adaptive strategy is introduced to determine a grid size which provides a good tradeoff between localization accuracy and computational complexity. Table 2 indicates that the average size of ER and RA_CRRV both decrease as m increases. When m is small, both ER and RA_CRRV are large. In this case, a relatively large grid size can be used to determine RA_CRRV. Conversely, a small grid size results in high computational complexity and storage requirements, and may not improve the localization accuracy. When m is large, a small grid size is needed to determine RA_CRRV. As ER is small in this case, the computational complexity and storage requirements are not substantial with a small grid size. With an adaptive grid size, a large size is employed if the unknown node has a small number of anchors and a small size is employed if there are a large number of anchors. Thus, the number of the grids when partitioning the ER of an unknown node is fixed instead of fixing the grid size. Then the ER of an unknown node is divided into N_0 grids where N_0 is a constant, so the grid size is

$$g_0 = \frac{S_{ER}}{N_0}, \tag{15}$$

with S_{ER} given by (3). Typically, S_{ER} decreases as the number of neighboring anchors increases. Therefore, (15) ensures a large grid size with a small number of neighboring anchors and a small grid size with a large number of neighboring anchors.

It is difficult to choose an appropriate value of N_0 for all the unknown nodes. In order to make localization performance independent of the choice of N_0, each unknown node first scans its ER with grid size g_0. Then the grid size is adjusted according to the Spearman rank order correlation coefficients. In particular, the grid size is reduced by half until the largest Spearman rank order correlation coefficient is larger than a given threshold. The localization method using the above adaptive grid size strategy is referred to as adaptive-CRRV.

V. PERFORMANCE RESULTS

In this section, simulation is used to evaluate the performance of the proposed algorithms. The simulation environment has p anchors and q unknown nodes deployed randomly using a uniform distribution in an area of size $300 \times 300\mathrm{m}^2$. The number of nodes and the communication range are set to 300 and 50 m, respectively. The anchor ratio $p/(p+q)$ is considered to assess the performance. In addition to the localization accuracy, the resource requirements including storage and computational complexity are examined. The results given are for an average of 300 trials. The proposed methods are compared with three existing algorithms, centroid [20], DRLS [25] and HSL [28]. Results are given for the free space and log-normal distance path loss (LDPL) channel models.

A. CRRV PERFORMANCE

The performance of the CRRV algorithm is now evaluated for different anchor ratios. As the grid size is a critical parameter for grid-based algorithms, five different grid sizes are considered and are denoted by the grid length.

1) LOCALIZATION ACCURACY

The localization accuracy is evaluated using the mean error which is defined as

$$meanerror = \frac{\sum_{t}\sum_{i=1}^{M}\sqrt{(\tilde{x}_i - x_i)^2 + (\tilde{y}_i - y_i)^2}}{tMr}, \quad (16)$$

where M is the number of unknown nodes that have been localized, t is the number of trials, $(\tilde{x}, \tilde{y})$ is the estimated position, and (x, y) is the actual position. Note that the mean error is relative to the communication range r.

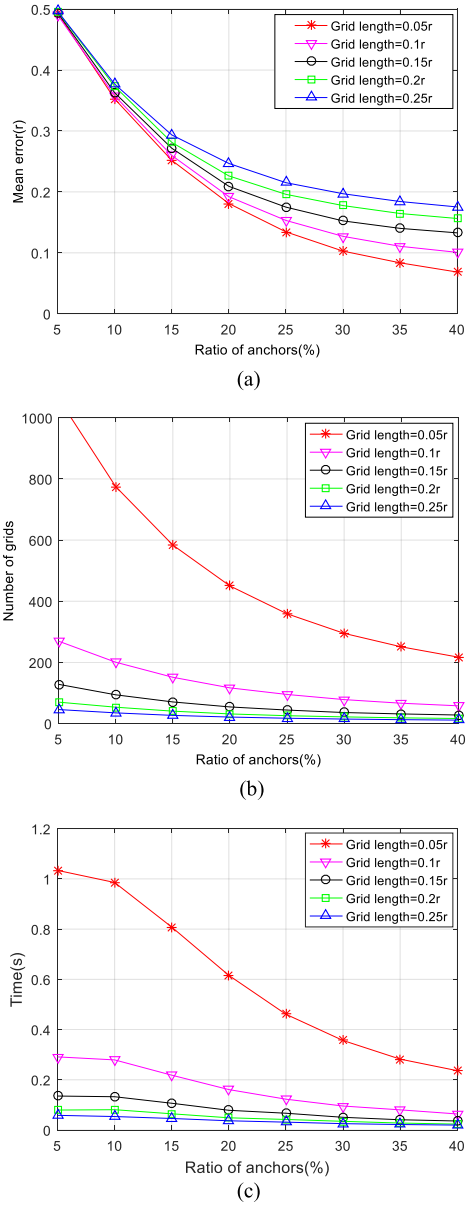

FIGURE 9. Performance of the CRRV algorithm (a) mean localization error, (b) average number of grids, and (c) computation time, CPU: Intel Core i5-6600 3.30 GHz.

Fig. 9(a) presents the mean error versus the anchor ratio for different grid sizes. For a given grid size, the mean error decreases as the anchor ratio increases. This is because the average number of neighboring anchors increases with the anchor ratio. Moreover, as expected, the localization accuracy improves as the grid size is reduced for a given anchor ratio. In addition, the performance improvement obtained from reducing the grid size is less significant when the anchor ratio is low. Conversely, there is a significant difference in localization performance when the anchor ratio is high.

2) STORAGE REQUIREMENTS

With CRRV, an unknown node needs to store the following information.

- The neighboring anchors list, including the neighboring anchor ID, coordinates, and RSS values.
- The ER position, including the coordinates of the four vertices.
- The information of all grids, including grid number and Spearman rank order correlation coefficient.

Compared to the third set of information, the storage cost of the first two can be neglected. Further, the storage requirements are proportional to the number of grids. Therefore, the CRRV storage requirements can be estimated as the number of grids.

Fig. 9(b) presents the average number of grids with the CRRV algorithm versus the anchor ratio. As expected, the number of grids decreases as the grid size increases when the anchor ratio is fixed. For a given grid size, the average number of grids decreases as the anchor ratio increases because the ER area decreases.

3) COMPUTATIONAL COMPLEXITY

With CRRV, each unknown node needs to first calculate its ER and then scans the grids in the ER. Assume that the average number of unknown node neighboring anchors is n. To obtain the ER, the intersection of the n rectangles that bound the communication regions of the n anchors needs to be calculated. Hence, the computational complexity of calculating the ER is $O(qn)$ where q is the number of unknown nodes. In the grid scan stage, an unknown node first computes the distances between each grid and the n neighboring anchors. Then the distance rank vector is constructed by sorting the n distances and the Spearman rank order correlation coefficient is computed. From (2), the complexity of computing this correlation coefficient is $O(n)$. Then the computational complexity of the grid scans is $O\left(qCn(2 + \log_2 n)\right)$, where C is the average number of grids for each unknown node. Therefore, the complexity of the CRRV algorithm is $O\left(qn(1 + 2C + \log_2 n)\right)$. Typically, n is much smaller than C, hence the computational complexity of CRRV can be considered as $O\left(qnC\right)$.

Fig. 9(c) shows the computation time required for localization using the CRRV algorithm versus the anchor ratio. The computer used had an Intel Core i5-6600 3.30 GHz CPU. This relationship is similar to that in Fig. 9(b). This indicates that the computational cost is proportional to the number of grids, which verifies the above analysis, i.e. most of the computational complexity is in the grid scan procedure.

Considering both localization accuracy and resource requirement, it can be concluded from Fig. 9 that for a network with a small number of anchors, a large grid size is appropriate, while a small grid size is preferable for a network with a high anchor density.

Theoretically, CRRV can also be implemented using analytic geometry, except for the grid scan procedure. Analytic geometry-based CRRV is performed as follows. For the RSS rank vector constraints, a distance rank vector table including all feasible distance rank vectors and the vertices of the corresponding sub-regions are constructed. Then RA_RRV is obtained by searching this table. Since using the quadratic connectivity constraints in the simulations is intractable, the square communication regions of the anchors are used instead of the circular regions, i.e. RA_C is replaced with ER.

TABLE 3. Computation time and mean localization error of grid-based CRRV and analytic geometry-based CRRV.

	Anchor ratio	5%	10%	15%	20%	25%	30%	35%	40%
Analytic geometry-based CRRV	Time (s)	0.244	0.306	0.416	0.521	0.676	0.781	0.873	1.084
	Mean error (r)	0.523	0.361	0.272	0.197	0.144	0.103	0.086	0.058
Grid-based CRRV	Time (s)	0.291	0.279	0.218	0.161	0.122	0.096	0.080	0.064
	Mean error (r)	0.515	0.351	0.265	0.183	0.143	0.122	0.112	0.098

Table 3 gives the computation time and mean localization error for grid-based CRRV and analytic geometry-based CRRV. The grid size of grid-based CRRV is set to $0.01r^2$. It can be concluded from these results that the computation time of analytic geometry-based CRRV increases rapidly with the anchor ratio. When the anchor ratio is lower than 10%, the computation time of analytic geometry-based CRRV is comparable to that of grid-based CRRV or even lower. However, the corresponding mean localization error is slightly higher than grid-based CRRV. This is because RA_C is approximated by ER. When the anchor ratio is higher than 15%, the computation time of analytic geometry-based CRRV is much higher than that of grid-based CRRV. For example, it is almost 17 times higher when the anchor ratio is 40%. In contrast, the improvement in localization accuracy using analytic geometry-based CRRV is limited. For example, the largest improvement in localization error is $0.05r$, which is obtained with an anchor ratio of 40%. This improvement can also be achieved using adaptive-CRRV but with a much lower computation time.

B. ADAPTIVE-CRRV PERFORMANCE

In this section, the performance of the proposed adaptive-CRRV is evaluated. Different from CRRV which employs the same grid size for all unknown nodes, adaptive-CRRV employs different grid sizes for these nodes. In particular, each node first employs the grid size obtained from (15) and then determines whether to adjust the grid size according to the largest Spearman rank order correlation coefficient and a predetermined threshold. In the simulations, this threshold is set to be 0.95, and the values of N_0 are 20, 40, 60, 80 and 100.

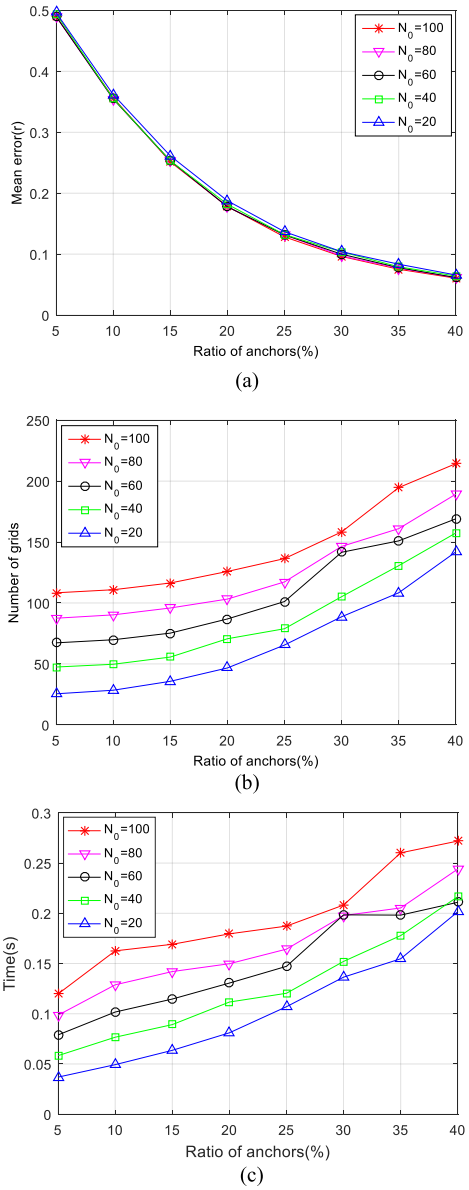

FIGURE 10. Performance of the adaptive-CRRV algorithm (a) mean localization error, (b) average number of grids, and (c) computation time.

Fig. 10(a) presents the mean localization error with the adaptive-CRRV algorithm versus the anchor ratio. Unlike with the CRRV algorithm whose performance is closely linked to the grid size, the performance of adaptive-CRRV has little correlation with N_0. A larger N_0 provides minimal improvement in localization accuracy due to the adaptive grid size strategy.

Fig. 10(b) gives the average number of grids with the adaptive-CRRV algorithm. The storage required for adaptive-CRRV increases with the anchor ratio for a given N_0. When the anchor ratio is low, the average number of grids is close

to N_0, i.e. most unknown nodes have grids with a Spearman rank order correlation coefficient larger than 0.95 using the grid size determined by (15). As the number of anchors increases, the number of sub-regions for an unknown node also increases. Hence, the unknown nodes must decrease the grid size to obtain grids with a Spearman rank order correlation coefficient larger than 0.95. As a consequence, the average number of grids increases with the anchor ratio. For a given anchor ratio, the storage requirements increase with N_0.

The computation time of adaptive-CRRV is plotted in Fig. 10(c). This relationship is similar to that in Fig. 10(b), which indicates that the computation time is dominated by the grid scan procedure.

C. PERFORMANCE COMPARISON WITH THE FREE SPACE CHANNEL MODEL

In this section, the CRRV and adaptive-CRRV algorithms are compared with the HSL, DRLS and centroid methods using the free space model. With this model, the RSS values decrease monotonically with distance. For a fair comparison, the HSL algorithm is implemented using grids. A grid size of $0.01r^2$ is used with the CRRV, HSL and DRLS algorithms, and $N_0 = 80$ for the adaptive-CRRV algorithm.

1) LOCALIZATION ACCURACY COMPARISON

Fig. 11(a) presents the mean error versus the anchor ratio. This shows that the CRRV and adaptive-CRRV algorithms perform better than the centroid and HSL methods regardless of the anchor ratio. This is because the proposed algorithms use both connectivity and RSS rank constraints, while the centroid method uses only connectivity constraints and the HSL technique uses only RSS relationship constraints. Moreover, the adaptive-CRRV algorithm provides the best performance due to the use of an adaptive grid size.

The CRRV and adaptive-CRRV algorithms outperform DRLS when the anchor ratio is high, while DRLS outperforms CRRV and adaptive-CRRV when the anchor ratio is low. DRLS employs both positive and the negative connectivity constraints. The results obtained indicate that the negative connectivity constraints provide less information than the RSS rank constraints when the anchor ratio is high, while the negative connectivity constraints provide more information when the anchor ratio is low.

2) STORAGE REQUIREMENTS COMPARISON

As discussed previously, the storage requirements of grid-based algorithms can be estimated by the number of grids. Fig. 11(b) shows the average number of grids for the CRRV, adaptive-CRRV, HSL and DRLS algorithms. From Fig. 11(b), the number of grids with CRRV and DRLS is similar because they are determined by the estimation rectangle and grid size. The storage requirements with the HSL algorithm is always larger than with CRRV and DRLS. This is because the grid scan region with HSL is based on the Voronoi diagram, and the Voronoi cell sizes are usually larger than the ER.

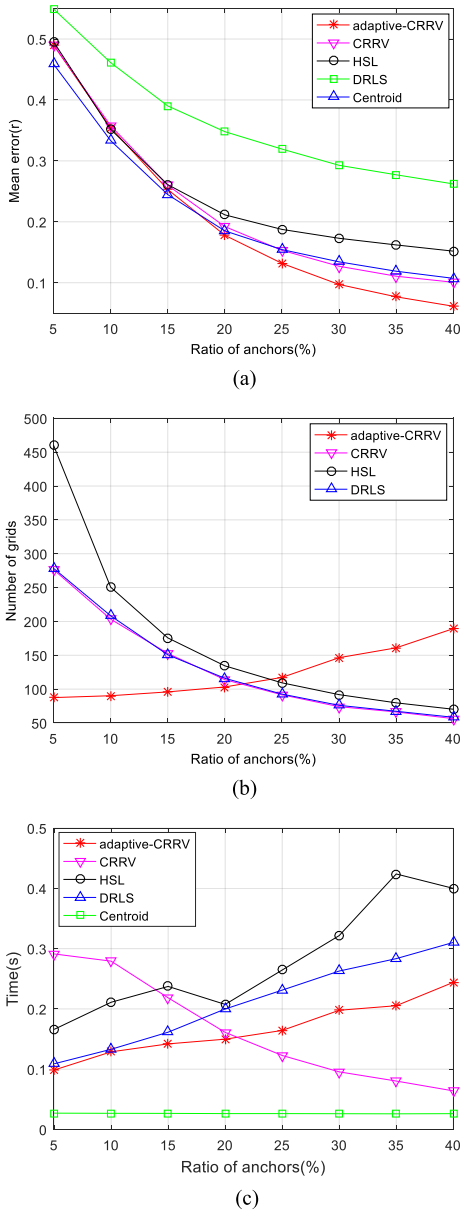

FIGURE 11. Performance with the free space channel model (a) mean localization error, (b) average number of grids, and (c) computation time.

The storage requirements of adaptive-CRRV are the lowest when the anchor ratio is low, and this increases with the anchor ratio. Although the storage requirements of adaptive-CRRV are the highest when the anchor ratio is high, it is still acceptable. Centroid is not a grid-based algorithm, so it has the lowest storage requirements.

3) COMPUTATIONAL COMPLEXITY COMPARISON

The computational complexity of the centroid scheme is $O(q)$, which is the lowest. From [25], the computational complexity of DRLS is $O\left(q(Cn+n+s)\right)$, where n, s, and C are the average numbers of unknown node neighboring anchors, two-hop neighboring anchors, and grids, respectively. For HSL, the Voronoi diagrams of all anchors are first

computed, and then each unknown node conducts a half-symmetric presence test for each pair of neighboring anchors. The computational complexity of determining the Voronoi diagrams is $O(p \log_2 p)$, where p is the number of anchors. There are $n(n-1)/2$ pairs of anchors for each unknown node to calculate the residence area. Therefore, each grid is scanned $n(n-1)/2$ times, and each time HSL computes three distances. Hence, the grid scan computational complexity of HSL is $O(3qCn(n-1)/2)$, so the overall computational complexity of HSL is $O(p \log_2 p + 3qCn(n-1)/2)$.

Figure 11(c) shows the computation time of the five algorithms. As expected, centroid has the lowest time. When the anchor ratio is medium to high, CRRV and adaptive-CRRV have the next two lowest times, while HSL has the highest. This is because the leading term in the computational complexity of CRRV and adaptive-CRRV is $O(qCn)$, while it is $O(qCn^2)$ for HSL. There are two reasons why adaptive-CRRV has a higher computation time than CRRV. The first is that more grids need to be scanned in adaptive-CRRV. The second is that some unknown nodes have to scan their ER more than once due to the adaptive strategy. DRLS has a lower computation time than CRRV when the anchor ratio is low, while it is higher when the anchor ratio is high. This is because the leading term in the computational complexity of DRLS is $O(qCn)$ when the anchor ratio is low, while it is $O(qs)$ while the anchor ratio is high.

D. LDPL CHANNEL MODEL PERFORMANCE

In this section, the log-normal distance path loss (LDPL) channel model is considered. This model characterizes the path loss inside buildings and in densely populated areas [32]. With this model, the RSS can be expressed as

$$P_R(d) = P_R(d_0) - 10\beta \log\left(\frac{d}{d_0}\right) + X_\sigma, \qquad (17)$$

where $P_R(d)$ denotes the received power in dBm at a distance d, $P_R(d_0)$ is the received power in dBm at a reference distance d_0, and β is the path loss exponent. X_σ is a zero-mean Gaussian random variable with variance σ^2 which reflects the attenuation in dB caused by shadowing. A given environment can be modeled by choosing appropriate values of β and σ [32]. In this paper, $\beta = 3$ and $\sigma = 5$ are used to denote an urban area, and $\beta = 4$ and $\sigma = 6.8$ are used to denote an obstructed factory environment.

To improve the localization reliability with the LDPL model, the RSS measurements are averaged over 10 rounds of anchor beacons. The grid size is set to $0.01r^2$ for the CRRV, HSL and DRLS algorithms, and $N_0 = 80$ for the adaptive-CRRV algorithm. Fig. 12 presents the mean error in the urban and obstructed factory environments. The storage requirements and computation time with the LDPL model are similar to those with the free space model, and so are omitted.

The results in Fig. 12 show the negative effects of multipath fading and shadowing, particularly with the DRLS algorithm because the negative connectivity constraints provide

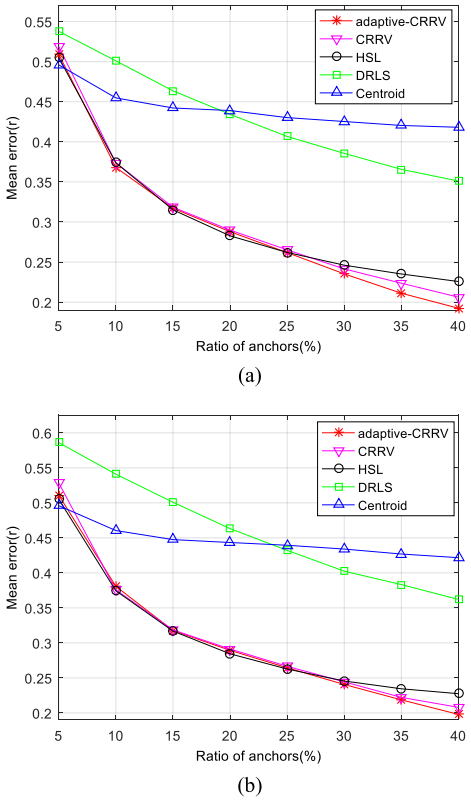

FIGURE 12. Mean localization error using the LDPL channel model (a) urban area, and (b) obstructed factory environment.

poor quality location information. Again, the CRRV and adaptive-CRRV algorithms provide the best localization performance. When the anchor ratio is not high, the performance of HSL is almost the same as with the proposed algorithms. However, as illustrated in Fig. 11, the computation time of HSL is higher than that of the proposed algorithms. Moreover, the calculation of the Voronoi diagrams in HSL must be implemented in a centralized manner. By contrast, the CRRV and adaptive-CRRV algorithms can be implemented in a distributed manner. When the anchor ratio is high, CRRV and adaptive-CRRV both perform better than HSL. This is because CRRV is more robust to RSS measurement errors than HSL. HSL uses the relationships between pairs of RSS values rather than just the RSS. Hence, HSL is robust when the RSS measurement errors are small and the relationships are not changed. However, when these relationships are changed by large measurement errors, HSL cannot detect these errors. In contrast, CRRV and adaptive-CRRV can detect and correct errors in the RSS rank vectors, as discussed in Section III.

VI. CONCLUSION

In this paper, new distributed range-free localization algorithms for wireless sensor networks were proposed. The CRRV algorithm uses connectivity and RSS rank vectors to localize the unknown nodes. The RSS rank vector for an unknown node is formed using the ranks of the RSS values

from neighboring anchors. With this method, an unknown node is first localized in an initial residence area based on the connectivity constraints. Then this area is refined using the RSS rank vector. Finally, the centroid of the refined residence area is used to estimate the node location. To reduce the computational complexity, a grid-scan approach was employed. Further, an improved version of CRRV, denoted adaptive-CRRV, was proposed which uses an adaptive strategy to determine the grid size. This allows for good localization accuracy with low computational complexity. Performance results were presented to demonstrate the superiority of the proposed algorithms.

APPENDIX (PROOF OF THEOREM 1)

In this section, the largest number of sub-regions for m neighboring anchors is derived. In a two-dimensional Euclidean plane, n lines divide the plane into at most $n(n+1)/2+1$ sub-regions. This happens when any two arbitrary lines intersect and any three arbitrary lines do not intersect at a point. For instance, the largest number of sub-regions divided by three lines is seven, as shown in Fig. 13(a). When three lines intersect at one point, the number of sub-regions is reduced by one, as shown in Fig. 13(b).

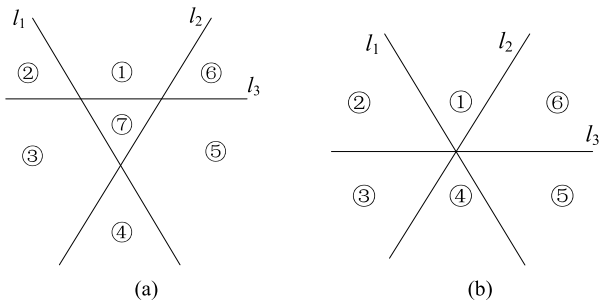

FIGURE 13. The sub-regions determined by three lines (a) seven sub-regions, and (b) six sub-regions.

For an unknown node with m neighboring anchors, any two anchors can determine a perpendicular bisector. If $m \geq 3$, there are C_m^2 perpendicular bisectors. Assume that no three anchors are in a line so that any three arbitrary anchors form a triangle. The perpendicular bisectors of three edges of a triangle intersect at one point which is the circumcenter of the triangle. Then m anchors can form C_m^3 triangles, so there are at least C_m^3 groups of three perpendicular bisectors that intersect at a point. Therefore, C_m^2 perpendicular bisectors divide the plane into at most $C_m^2\left(C_m^2+1\right)/2 - C_m^3 + 1$ sub-regions. In addition, the number of sub-regions is m when m is one or two. Thus, the maximum number of sub-regions for m neighboring anchors is

$$\lambda_m = \begin{cases} 1, m = 1 \\ 2, m = 2 \\ C_m^2\left(C_m^2+1\right)/2 - C_m^3 + 1, m \geq 3 \end{cases}$$

$$= \frac{1}{24}(3m^4 - 10m^3 + 21m^2 - 14m + 24).$$

REFERENCES

[1] G. Mao, B. Fidan, and B. D. O. Anderson, "Wireless sensor network localization techniques," *Comput. Netw.*, vol. 51, no. 10, pp. 2529–2553, 2007.
[2] T. J. S. Chowdhury, C. Elkin, V. Devabhaktuni, D. B. Rawat, and J. Oluoch, "Advances on localization techniques for wireless sensor networks: A survey," *Comput. Netw.*, vol. 110, pp. 284–305, Dec. 2016.
[3] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005.
[4] M. Z. A. Bhuiyan, G. Wang, and A. V. Vasilakos, "Local area prediction-based mobile target tracking in wireless sensor networks," *IEEE Trans. Comput.*, vol. 64, no. 7, pp. 1968–1982, Jul. 2015.
[5] M. Wu, L. Tan, and N. Xiong, "Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications," *Inf. Sci.*, vol. 329, pp. 800–818, Feb. 2016.
[6] H. Mahboubi, J. Habibi, A. G. Aghdam, and K. Sayrafian-Pour, "Distributed deployment strategies for improved coverage in a network of mobile sensors with prioritized sensing field," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 451–461, Feb. 2013.
[7] T. Shu and S. Cui, "Renovating location-based routing for integrated communication privacy and efficiency in IoT," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, May 2017, pp. 1–6.
[8] Z. Zhong and T. He, "RSD: A metric for achieving range-free localization beyond connectivity," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 11, pp. 1943–1951, Nov. 2011.
[9] J.-F. Huang, G.-Y. Chang, and G.-H. Chen, "A historical-beacon-aided localization algorithm for mobile sensor networks," *IEEE Trans. Mobile Comput.*, vol. 14, no. 6, pp. 1109–1122, Jun. 2015.
[10] J. Zhao *et al.*, "Localization of wireless sensor networks in the wild: Pursuit of ranging quality," *IEEE/ACM Trans. Netw.*, vol. 21, no. 1, pp. 311–323, Feb. 2013.
[11] X. Zhu and G. Chen, "Spatial ordering derivation for one-dimensional wireless sensor networks," in *Proc. IEEE Int. Symp. Parallel Distrib. Process. Appl.*, Busan, South Korea, May 2011, pp. 171–176.
[12] C. Bo *et al.*, "Locating sensors in the forest: A case study in green orbs," in *Proc. IEEE Int. Conf. Comput. Commun.*, Orlando, FL, USA, Mar. 2012, pp. 1026–1034.
[13] H. Shen, Z. Ding, S. Dasgupta, and C. Zhao, "Multiple source localization in wireless sensor networks based on time of arrival measurement," *IEEE Trans. Signal Process.*, vol. 62, no. 8, pp. 1938–1949, Apr. 2014.
[14] G. Wang, H. Chen, Y. Li, and N. Ansari, "NLOS error mitigation for TOA-based localization via convex relaxation," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4119–4131, Aug. 2014.
[15] R. Kim, T. Ha, H. Lim, and D. Jung, "TDoA localization for wireless networks with imperfect clock synchronization," in *Proc. Int. Conf. Inform. Netw.*, Phuket, Thailand, Feb. 2014, pp. 417–421.
[16] H.-J. Shao, X.-P. Zhang, and Z. Wang, "Efficient closed-form algorithms for AOA based self-localization of sensor nodes using auxiliary variables," *IEEE Trans. Signal Process.*, vol. 62, no. 10, pp. 2580–2594, May 2014.
[17] Y. Wang and K. C. Ho, "An asymptotically efficient estimator in closed-form for 3D AOA localization using a sensor network," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6524–6535, Jul. 2015.
[18] R. M. Vaghefi, M. R. Gholami, R. M. Buehrer, and E. G. Ström, "Cooperative received signal strength-based sensor localization with unknown transmit powers," *IEEE Trans. Signal Process.*, vol. 61, no. 6, pp. 1389–1403, Mar. 2013.
[19] S. Tomic *et al.*, "RSS-based localization in wireless sensor networks using convex relaxation: Noncooperative and cooperative schemes," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 2037–2050, May 2015.
[20] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less low cost outdoor localization for very small devices," *IEEE Pers. Commun. Mag.*, vol. 7, no. 5, pp. 28–34, Oct. 2000.
[21] D. Niculescu and B. Nath, "DV based positioning in Ad Hoc networks," *Telecommun. Syst.*, vol. 22, nos. 1–4, pp. 267–280, 2003.
[22] Y. Shang and W. Ruml, "Improved MDS-based localization," in *Proc. Joint Conf. IEEE Comput. Commun. Societies*, Hong Kong, Mar. 2004, pp. 2640–2651.
[23] J. Lee, W. Chung, and E. Kim, "A new kernelized approach to wireless sensor network localization," *Inform. Sci.*, vol. 243, no. 10, pp. 20–38, Sep. 2013.

[24] Z. Wang, H. Zhang, T. Lu, Y. Sun, and X. Liu, "A new range-free localisation in wireless sensor networks using support vector machine," *Int. J. Electron.*, vol. 105, no. 2, pp. 244–261, 2018.
[25] J. P. Sheu, P. C. Chen, and C.-S. Hsu, "A distributed localization scheme for wireless sensor networks with improved grid-scan and vector-based refinement," *IEEE Trans. Mobile Comput.*, vol. 7, no. 9, pp. 1110–1123, Sep. 2008.
[26] J.-A. Jiang *et al.*, "A distributed RSS-based localization using a dynamic circle expanding mechanism," *IEEE Sensors J.*, vol. 13, no. 10, pp. 3754–3766, Oct. 2013.
[27] S. Tomic, M. Beko, and R. Dinis, "Distributed RSS-based localization in wireless sensor networks based on second-order cone programming," *Sensors*, vol. 14, no. 10, pp. 18410–18432, Oct. 2014.
[28] N. Lasla, M. F. Younis, A. Ouadjaout, and N. Badache, "An effective area-based localization algorithm for wireless networks," *IEEE Trans. Comput.*, vol. 64, no. 8, pp. 2103–2118, Aug. 2015.
[29] X. Zhang, Z. Duan, and Z. Dong, "RSS-based efficient grid-scan localization algorithm in wireless sensor networks," in *Proc. Int. Conf. Wireless Commun. Signal Process.*, Huangshan, China, Oct. 2012, pp. 1–6.
[30] M. de Berg, O. Cheong, M. V. Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*, 3rd ed. Berlin, Germany: Springer-Verlag, 2008.
[31] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992.
[32] T. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ, USA: Prentice-Hall, 2001.

ZENGFENG WANG received the B.E. degree in electronics engineering from Qingdao University, China, in 2003, and the M.S. degree in signal and information processing from Shandong University, China, in 2006. She is currently pursuing the Ph.D. degree with the Ocean University of China, China. Her research interests include wireless localization, ultra-wideband radio systems, and 60-GHz wireless communications.

HAO ZHANG received the B.E. degree in telecommunications engineering and industrial management from Shanghai Jiao Tong University, China, in 1994, and the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 2004. He is currently a Professor with the Department of Electronic Engineering, Ocean University of China. His research interests include ultra-wideband radio systems, MIMO wireless systems, and cooperative communication networks.

TINGTING LU received the B.S. degree in communication engineering from Hunan University in 2006, and the M.S. degree in communication and information systems and the Ph.D. degree in computer application technology from the Ocean University of China in 2009 and 2013, respectively. She is currently a Lecturer with the Ocean University of China. Her research interests include millimeter wave communication systems, ultra-wideband radio systems, OFDM technologies, and GNSS.

T. AARON GULLIVER received the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 1989. From 1989 to 1991, he was a Defence Scientist with the Defence Research Establishment Ottawa, Ottawa, ON, Canada. He has held various academic positions at Carleton University, Ottawa, and the University of Canterbury, Christchurch, New Zealand. He joined the University of Victoria in 1999, where he is a currently a Professor with the Department of Electrical and Computer Engineering. His research interests include information theory and communication theory, algebraic coding theory, cryptography, multicarrier systems, smart grid, relay and cooperative communication systems, and security. He was a fellow of the Engineering Institute of Canada and the Canadian Academy of Engineering, in 2002 and 2012, respectively.