A Three-step Methodology for Stochastic Deployment of Wireless Sensor Networks

A. Ejnioui, C. E. Otero, I. Kostanic and L. D. Otero

1Information Technology, University of South Florida Lakeland, Lakeland, Florida, USA
2Electrical and Computer Engineering, Florida Institute of Technology, Melbourne, Florida, USA
3Engineering Systems, Florida Institute of Technology, Melbourne, Florida, USA

Abstract—This paper presents a three-step methodology for stochastic deployment of wireless sensor networks. While the first step uses simulation to elaborate strategies for a given deployment scenario, the second step consists of applying vertical variance trimming techniques to reduce the number of non-redundant alternative strategies. The final step consists of formulating and solving the problem as a multi-attribute decision problem using grey number theory due to the uncertain character of simulation data. Decision-makers can use this methodology to determine the best deployment strategies based on mission specific goals. This methodology can be used to simplify the decision-making process and provide decision-makers the ability to consider all factors involved in the wireless sensor network deployment problem. The methodology can be easily customized to include numerous quality factors to further compare deployment strategies and identify the one that best meet applications requirements.

Keywords: wireless sensor networks, stochastic deployments, statistical analysis, grey numbers, multi-attribute decision.

1 Introduction

Recent advances in micro electro-mechanical systems have led to the development of tiny low-power devices that are capable of sensing the world and communicating with each other. Such devices may be deployed in vast numbers over large geographical areas to form wireless sensor networks (WSN). WSNs provide the means for autonomous monitoring of physical events in areas where human presence is not desirable or impossible. Therefore, they are expected to facilitate many existing applications and bring into existence entirely new ones. Several proposed applications of WSNs include disaster relief, environmental control, military applications and border security [1].

In each application, the sensor nodes are deployed over the area of interest and tasked with sensing the environment and communicating with each other in multi-hop fashion to transmit the information back to a base station, also known as the information sink [2]. From the sink, the information is collected and typically relayed to a central location, across remote sites, where it is processed and analyzed.

For the most part, WSNs are highly application-dependent. This means that details such as node design, form-factor, processing algorithms, network protocols, network topology, and deployment scheme are customized for the proposed application. Among these, the deployment scheme is considered extremely important, since it directly influences parameters such as network complexity, connectivity, coverage, cost, and lifetime.

WSN deployments schemes are classified as deterministic or stochastic. Deterministic deployments typically result in optimal efficiency; however, due to the size and density required to provide appropriate network coverage in large geographical areas, careful positioning of the deployed nodes is impractical. Furthermore, several applications of WSNs are expected to operate in hostile environments [3]. This makes pre-defined deployment in some cases impossible; consequently, stochastic deployments become the only feasible alternative [1]. For these applications, sensor nodes may be dropped from a plane, delivered in an artillery shell, rocket or missile, or catapulted from a shipboard [2]. In these cases, the WSN has the utmost challenge of guaranteeing connectivity and proper area coverage upon deployment [4]. This requires detailed planning to find deployment strategies that meet application requirements in terms of network connectivity, coverage, cost, and lifetime.

This paper presents a decision-making methodology for stochastic deployment of WSNs. The methodology uses simulation, statistical analysis, and a utility-based multi-attribute decision process based on grey number theory to provide an innovative and unique approach that helps decision-makers determine goal-oriented deployment strategies from a set of alternatives. Furthermore, the methodology provides significant contribution to the current body of research by providing an extensible technique that takes into account important parameters, such as connectivity, coverage, cost, lifetime, involved in the deployment of WSNs.

2 Background Work

The deployment problem has been the topic of much research work; however, the majority of the methodologies concentrate on carefully positioning nodes to meet application
requirements [5, 6]. Insufficient work has been done to improve decision-making in stochastic deployment of WSNs. Furthermore, most of the current work provides methodologies that take into account one or two parameters at the expense of other network parameters. In [7], the authors present a methodology to maximize coverage and connectivity in randomly deployed WSN. In [8, 9], the authors present a methodology for decreasing node density (i.e., cost). In [10, 11], the authors point out the lack of research towards the WSN deployment problem and state that “While WSN design, architecture, protocols and performance have been extensively studied, only a few research efforts have studied the device deployment problem”. Furthermore, the authors point out flaws in recent publications by stating, “Most of these works tackle the deployment problem only from a perspective of coverage and/or connectivity. The significance of deployment on lifetime is mostly overlooked”.

These methodologies fail to provide decision-makers with holistic views that consider all parameters involved in the deployment decision-making process and allow them to make customized deployment decisions based on all application requirements.

3 Methodology

WSN are application-specific, therefore it is impractical to expect that the same solution can be used to address deployments in all environments. The proposed methodology is built on this fundamental assumption. For that reason, it requires decision-makers to execute the methodology under settings that provide appropriate characterization of application-specific requirements. Consequently, customized simulations specific to the deployment scenario at hand are required. Simulations of WSN deployments can be carried out using the popular ns2 simulator. Once simulation data are collected, the methodology uses the Vertical Variance Trimming (VVT) technique to eliminate statistically redundant deployment alternatives. Therefore, VVT provides a reduced number of deployment alternatives but equal characterization of the original deployment scenario. Finally, the reduced set of deployment alternatives is analyzed by formulating and solving a multi-attribute decision problem using grey numbers in order to rank deployment alternatives based on deployment goals. Figure 1 shows an overview of this methodology.

4 Vertical Variance Trimming

Vertical Variance Trimming (VVT) is a technique devised for reducing the number of deployment alternatives based on statistical analysis [12]. VVT works by determining the effects of deployment parameters (e.g., number of nodes, radio range, sensor range) on WSN network efficiency. Typical WSN deployment assumptions include: (1) higher network connectivity and area coverage is achieved by deploying higher number of nodes; and (2) higher radio and sensing range result in higher connectivity and area coverage. These are valid assumptions; however, the degree to which they are significant in a stochastic deployment scenario needs to be determined before making deployment decisions. To accomplish this and reduce deployment alternatives, VVT uses three fundamental techniques: Single Factor Analysis of Variance (ANOVA) [13], Least Significance Difference (LSD) [13], and Critical Metric Value (CMV).

![Image of Methodology overview](image)

Figure 1. Methodology overview.

4.1 Single Factor ANOVA

Single factor ANOVA is a statistical procedure that tests the equality of two or more response means. In the deployment case, ANOVA can be used to determine the effects of deploying varying number of nodes at different radio and sensor range, on overall network connectivity and coverage. By using the $F$ statistical test [13], the variance caused by natural error can be compared to the one caused by varying the deployment parameters. When these variances are similar, at least two deployment strategies are considered redundant and minimization of the number of deployment alternatives can be achieved. The test statistic $F$ is computed using equation 1 [13],

$$F_0 = \frac{S^2_T}{S^2_E} = \frac{SS_T/(a-1)}{SS_E/(N-a)}$$

(1)

where $N$ is the total number of observations and $a$ the total number of treatments. A detailed explanation for computing $SS_T$ and $SS_E$ is presented in [13].

There are two ways that single factor ANOVA can be used in the WSN deployment problem. Table 1 displays identified use cases for single factor ANOVA.

<table>
<thead>
<tr>
<th>Case</th>
<th>Number of Nodes</th>
<th>Radio Range</th>
<th>Sensor Range</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Variable</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Connectivity</td>
</tr>
<tr>
<td>2</td>
<td>Variable</td>
<td>Fixed</td>
<td>Fixed</td>
<td>Coverage</td>
</tr>
</tbody>
</table>

In both cases, the effects of varying the number of deployed nodes on network connectivity and area coverage are determined by using a fixed set of values for radio range and sensor range. Once statistical differences have been identified using ANOVA, the LSD test is used to identify which deployment alternatives are different from one another.
4.2 LSD
The LSD test is used on pairs of deployment strategies to determine their statistical equality. Two deployment strategies are different if the difference in their mean response (i.e., connectivity, coverage) is greater than the least significant difference. This is obtained for balanced experimental data as specified in [13]. Otherwise, they are considered statistically similar. When this occurs, decision-makers can eliminate the redundant strategy that results in increased cost. Once all the deployment strategies have been analyzed using LSD, the CMV is used to further reduce the deployment alternative set to match specific application requirements.

4.3 CMV
After redundant deployment strategies have been removed, the Critical Metric Value (CMV) is selected to identify the minimum value for the decision metric that decision-makers are willing to accept for initial deployment. Deployment strategies resulting in values higher than the CMV are considered for initial deployment. Deployment strategies resulting in values below the CMV are considered for re-deployment to fill gaps and correct initial deployment. With this final step, VVT reduces the number of deployment alternatives significantly, which simplifies the application of the utility-based multi-attribute decision making process.

5 Multi-attribute Decision Making Using Grey Systems Theory
Multi-attribute decision making problems occur in situations where a finite set of alternatives need to be evaluated according to a number of criteria or attributes. The evaluation consists of selecting the best alternative or ranking the set of alternatives based on those attributes. The evaluation of various strategies to deploy wireless sensor networks can be approached by finding a set of criteria that provides the optimal benefit by minimizing cost factors. Deployment goals are customized based on specific application requirements. For example, multi-segment WSN require high connectivity and high coverage for the Sensing & Relaying Segment (SRS), but high connectivity and low coverage for the Relaying Segment (RS) [8]. Other examples include WSN that use small autonomous vehicles after deployment to fill existing connectivity gaps in the network. In these applications, high area coverage with extended network lifetime is desired over high connectivity. In general, deployment strategies are composed of a fixed number of deployed nodes, fixed radio range, and fixed sensor range. For a given deployment scenario, there could be a number of deployment strategies, each providing different levels of connectivity, coverage, cost, and network lifetime. However, many decision problems present data that is imprecise or ambiguous leading to conflicting situations in which the evaluation of alternatives becomes difficult. This is the case when deploying WSNs. This information uncertainty has been modeled using fuzzy sets [14] or grey numbers [15]. While the former has been around for some time, only recently has interest been growing in the latter, since uncertainty can be modeled and manipulated in more flexible ways using grey number systems than fuzzy sets [15].

5.1 Grey Numbers and Grey Systems Theory
In practical applications, a grey number represents an indeterminate number that takes its possible value from an interval or a set of numbers. The symbol $\otimes$ denotes a grey number. The most basic types of grey numbers are [15]:
- Grey numbers with only a lower bound: $\otimes \in [a, \infty]$ or $\otimes(a)$, where a is a fixed number representing the lower bound.
- Grey numbers with only an upper bound: $\otimes \in [-\infty, a]$ or $\otimes(a)$ where $\overline{a}$ is a fixed number representing the upper bound.
- Interval grey numbers: $\otimes \in [a, \overline{a}]$ where a and $\overline{a}$ are the lower and upper bounds respectively.
- Continuous and discrete grey numbers: The former numbers can take any values within an interval while the latter can take only a finite number of potential values.
- Black and white numbers: When $\otimes \in [-\infty, \infty]$, that is when $\otimes$ has neither an upper nor lower bound, it is known as a black number. On the other hand, when $\otimes \in [a, \overline{a}]$ and $\overline{a} = a$, it is known as a white number.

After three decades of research, grey systems theory emerged as a new discipline with contributions in [15]:
- Grey algebraic systems, grey equations, grey matrices, etc..
- Sequence operators and generation of grey sequences.
- System analysis based on grey incidence spaces and grey clustering.
- Grey prediction models.
- Decision making using grey target decision models.
- Optimization models using grey programming, grey game theory and grey control.

5.2 Strategy Selection
The first step in grey system decision making approach involves the selection of deployment strategies for a given deployment scenario. These strategies are developed during simulation of specific scenarios. The results are captured in a strategy vector as follows:

$$S = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{bmatrix}$$

where $i = 1, 2, \ldots, n$.

5.3 Scenario Attributes
When simulating a deployment scenario, a number of strategies are developed. Each deployment scenario can be characterized by the following attributes:
- **Nodes:** A deployment strategy consists of deploying a number of nodes scattered over a well-defined geographic area.
• **Radio range**: Each node has an antenna for transmitting and receiving data. This antenna has a well-known radio range.
• **Sensor range**: In addition to an antenna, a node can carry several sensors for monitoring specific parameters in the environment (e.g., temperature, noise, motion, etc.). These sensors have all well-defined sensing ranges.
• **Connectivity**: Each deployment strategy can be defined by the degree to which the deployed nodes are connected. This is known as the connectivity of the network resulting from the deployment strategy. In general, connectivity increases when higher number of nodes is deployed in a given strategy.
• **Coverage**: This is the area covered by the network of nodes in a deployment strategy.
• **Power**: The power consumed by the deployed nodes in a given strategy.

These attributes can be represented in the following vector:

\[ A = [a_1 \ a_2 \cdots a_m] \]

for \( j = 1, 2, \ldots, m \).

### 5.4 The Deployment Scenario Matrix

For each deployment strategy, a value representing a benefit or cost is generally associated with each attribute. Because of the uncertainty of data in deploying WSNs, grey numbers are used to represent benefits or costs in the decision matrix. The overall assessment of a given deployment scenario based on the scenario attributes is captured using the following deployment scenario matrix:

\[
\mathbf{D} = \begin{bmatrix}
\otimes d_{11} & \otimes d_{12} & \cdots & \otimes d_{1m} \\
\otimes d_{21} & \otimes d_{22} & \cdots & \otimes d_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
\otimes d_{n1} & \otimes d_{n2} & \cdots & \otimes d_{nm}
\end{bmatrix}
\]

where the rows represent strategies considered in the deployment scenario while the columns represent the attributes of that scenario. Note that the \( l_{ij} \) and \( u_{ij} \) represent respectively the lower and upper bounds of grey number \( d_{ij} \) for \( i = 1, 2, \ldots, n \) and \( j = 1, 2, \ldots, m \).

### 5.5 Goal Weights

In general, a deployment scenario will also be characterized by very specific goals. For example, the goals of a deployment scenario may consist of maximizing network lifetime, maximizing connectivity, minimizing cost and maximizing coverage in this listed order. The first goal entails minimizing power usage in the deployment scenario while the third goal entails minimizing the number of nodes deployed in the scenario. Optimization goals consist mostly of minimizing or maximizing one or more attributes associated with a deployment scenario. However, these goals may not have the same importance in some cases. As a result, a weight from 0 to 1 indicating the importance of a goal is given to each attribute in each scenario. In the absence of weights, all attributes are assumed to have equal importance. A weight vector is created where \( w_j \) represents the importance of each attribute as follows:

\[
W = [w_1 \ w_2 \cdots w_m] \tag{5}
\]

### 5.6 Normalization of the Deployment Scenario Matrix

The scenario matrix can be normalized by using the core of the grey numbers in each column. If a grey number \( \otimes \in [a, \bar{a}] \) is continuous, then \( \otimes = \frac{1}{2} (a + \bar{a}) \) is the core of \( \otimes \) [14]. Grey numbers in the matrix can be normalized by using the sum of the cores in each matrix column as follows [15]:

\[
\bar{l}_{ij} = \frac{2l_{ij}}{\sum_{i=1}^{n} (l_{ij} + u_{ij})} \tag{6}
\]

\[
\bar{u}_{ij} = \frac{2u_{ij}}{\sum_{i=1}^{n} (l_{ij} + u_{ij})} \tag{7}
\]

for \( i = 1, 2, \ldots, n \) and \( j = 1, 2, \ldots, m \) where \( l_{ij} \) and \( u_{ij} \) are as defined in equation (4). The resulting normalized matrix is \( \bar{D} \).

### 5.7 Weighting of the Normalized Scenario Matrix

The normalized scenario matrix can be weighted by multiplying the bounds of each grey numbers in the matrix by the weight of its attribute. Let \( \otimes \bar{d}_{ij} = \frac{l_{ij}, u_{ij}}{i, j} \) be a grey number in the normalized matrix. Each grey number in the matrix is multiplied by its attribute weight as follows [15]:

\[
\otimes \bar{d}_{ij} = \otimes \bar{d}_{ij} \times w_j = [\bar{l}_{ij}, \bar{u}_{ij} = [\bar{l}_{ij}, \bar{u}_{ij}]] \tag{8}
\]

for \( i = 1, 2, \ldots, n \) and \( j = 1, 2, \ldots, m \). The resulting weighted normalized matrix is \( \bar{D} \).

### 5.8 Benefits and Costs in The Weighed Normalized Matrix

A simple weighted additive approach, similar to the COPRAS-G method, can be used to compute the benefits and costs of the attributes for each strategy in \( \bar{D} \) as follows [16, 17]:

\[
P_i = \frac{1}{2} \sum_{j=1}^{k} (d_{ij} + \bar{a}_{ij}) \tag{9}
\]

\[
R_i = \frac{1}{2} \sum_{j=k+1}^{m} (d_{ij} + \bar{a}_{ij}) \tag{10}
\]

assuming that the first \( k \) attributes are benefits while the remaining \((m-k)\) attributes are costs in \( \bar{D} \).
5.9 Relative Weight of Each Strategy

The importance of each strategy in the weighted normalized matrix can be calculated as follows [16, 17]:

\[ Q_l = P_l + \frac{\sum_{i=1}^{n} R_i}{R_i \sum_{i=1}^{n} 1} \] (11).

5.10 Utility of Each Strategy

The utility degree of each strategy can be calculated based on its relative weight as follows [16, 17]:

\[ U_i = \frac{Q_i}{\max Q_i} \] (12)

for \( i = 1, 2, \ldots, n \). The strategy with the highest utility degree is considered the best deployment strategy of the WSN under consideration given the \( m \) scenario attributes.

6 Case Study

Using simulation, the methodology is executed using a simplified deployment scenario consisting of rectangular deployment area measuring 500 m \( \times \) 500 m; 50 to 100 nodes available for deployment with onboard radio capable of transmitting between 50 to 100 meters and sensors capable of covering between 30 to 60 meters. Cost is assumed to be directly proportional to the number of deployed nodes, and lifetime is related to transmission power. This power is simulated using the Log-Normal RF propagation model, whereby terrain obstructions are modeled as a zero-mean normally distributed random variable with standard deviation proportional to obstructions [18]. VVT was executed to reduce the number of deployment strategies based on the connectivity metric.

The case study identifies deployment strategies composed of specific number of nodes and radio range. Since there are 50, 60, \ldots, n alternatives for the number-of-nodes factor and 40, 45, \ldots, r alternatives for radio range; the initial number of deployment alternatives is \( n \times r = 78 \). Using VVT, the number of deployment strategies was reduced to 51, resulting in 35% reduction of deployment strategies. In addition, a CMV value of 90% was used to further reduce the number of deployment strategies. The deployment alternatives for connectivity are summarized in Tables 2 and 3.

Table 2. Original set of deployment alternatives.

<table>
<thead>
<tr>
<th>Nodes</th>
<th>Radio Ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100</td>
</tr>
<tr>
<td>60</td>
<td>40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100</td>
</tr>
<tr>
<td>70</td>
<td>40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100</td>
</tr>
<tr>
<td>80</td>
<td>40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100</td>
</tr>
<tr>
<td>90</td>
<td>40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100</td>
</tr>
<tr>
<td>100</td>
<td>40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100</td>
</tr>
</tbody>
</table>

Using the trimmed set of deployment alternatives, a deployment scenario in which network lifetime is of highest importance, followed by connectivity, cost, and coverage is created. This scenario is shown in Figure 2. This scenario shows 35 strategies where each strategy is characterized by the following attributes: number of deployed nodes, radio range of these nodes, range of the sensors on these nodes, connectivity of the network deployed by these nodes, coverage of this network, and its power usage. These attributes are represented by grey numbers derived from simulation experiments. The weights at the bottom of Figure 2 illustrate the priority goals of the deployment scenario. As the weight of each attribute shows, power usage has the highest weight, which means that network lifetime is of utmost importance. Next, connectivity, number of nodes, and coverage follow in importance based on their weights.

Table 3. Trimmed set of deployment alternatives.

<table>
<thead>
<tr>
<th>Nodes</th>
<th>Radio Ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100</td>
</tr>
<tr>
<td>60</td>
<td>40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100</td>
</tr>
<tr>
<td>70</td>
<td>40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100</td>
</tr>
<tr>
<td>80</td>
<td>40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100</td>
</tr>
<tr>
<td>90</td>
<td>40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100</td>
</tr>
<tr>
<td>100</td>
<td>40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100</td>
</tr>
</tbody>
</table>

Figure 2. Deployment scenarios.

After normalization and weighting, the weighted normalized matrix appears in Figure 3. Among the attributes of this scenario, radio range, sensor range, connectivity and coverage are maximized whereas number of nodes and power usage are minimized. As such, the benefits of radio range, sensor range, connectivity and coverage are computed using equation (9) while the costs of power usage are computed using equation (10). Next, the relative importance of all strategies and their utility degrees are computed using equation (11) and (12). Figure 4 shows the final results where strategies 176, 153 and 127 are the top three strategies that are highly desirable in the simulated scenario. Although strategy 176 requires a high
number of nodes, it shows a high range of nodes and sensors as well as a high degree of connectivity and coverage.

The approach can be easily customized to include numerous quality factors to further compare deployment strategies.

8 References