A Statistical Method for Area Coverage Estimation and Loss Probability Analysis on Mobile Sensor Networks

FABIO COCCHI DA SILVA EIRAS; WAGNER LUIZ ZUCCHI
Department of Signal Processing
Escola Politécnica – University of São Paulo
Av Prof Luciano Gualberto – Trav. 03 nº 380– São Paulo - Brazil
BRAZIL
fabio.eiras@usp.br ; wzucchi@lps.usp.br

Abstract: - Sensor networks are formed by fixed or mobile sensor nodes and their functions is to capture the events that occur within a certain area and then relay to a central node. Normally sensor nodes are not able to transmit or receive information over long distances due to the need to use less energy and thus extend their useful life. Therefore, the number of sensor nodes in a given area directly influences the coverage of this area and the ability of information to be relayed by several sensors to the central node. Therefore, if there are many missed messages the application will have its performance compromised. In this paper we use a statistical method-based approach to estimate the probability of message loss and area overage from the variables: node displacement velocity and sensor quantity. The results obtained are compared analytically with simple cases in order to validate the results obtained by the simulations performed.

Key-Words: - Sensor Networks; Computer Communication Network; Mobile Ad-Hoc Network; Wireless Sensor Networks Coverage.

Received: January 21, 2019. Revised: March 26, 2019. Accepted: April 29, 2019.
Published: September 30, 2019

1 Introduction

A sensor networks are characterized by the distribution of sensor nodes able to capture information, such as temperature, positioning, depth, wind speed or even capturing an image and transmitting the information to a system able to interpret and process the collected data. The sensors can be distributed in previously established positions or randomly in a given area, known as sensing area. They can also be fixed or mobile, depending on the application.

Network sensors are currently put to the most diverse applications. In the field of civilian applications, we can highlight agricultural sensing, flock monitoring, power distribution line monitoring, climate monitoring and highway and railway monitoring. In the field of military applications, the most important are remote border monitoring and space monitoring, whether aerial, from the coast line, from enemy territory or regarding troop locations.

These applications define the characteristics of the places where the sensor network should be installed. In certain cases, these sensors are best installed in remote terrains, such as forests, for example, or agricultural fields, or even in battle fields. For civilian applications, these sensors can be installed in more urbanized places, such as universities or industries, for example.

Generally speaking, a sensor network can be systemically divided into three components, systems, communication protocols and services, as shown in Figure 1 [1]. The first component is the system. Each sensor node is a system in itself and contains specific characteristics, such as memory capacity, battery life, type of sensing and technology utilized. The second component are the communication protocols that allow communication between the nodes and the application, between a node and the base station and/or between nodes. This component contains the protocols responsible for routing the network, for instance. The last component are the services developed with the purpose of handing a network with the expected performance and efficiency over to the application.

The characteristics or requirements of a sensor network are directly correlated with the requirements of the application that will be using this network and will therefore vary according to the purpose it will be used for. Among these characteristics, one can single out Ad-hoc networks, infrastructure networks, networks containing extra sources of energy, such as solar panels, if the sensors are mobile or fixed or if the sensors are distributed in a predetermined fashion or totally randomly

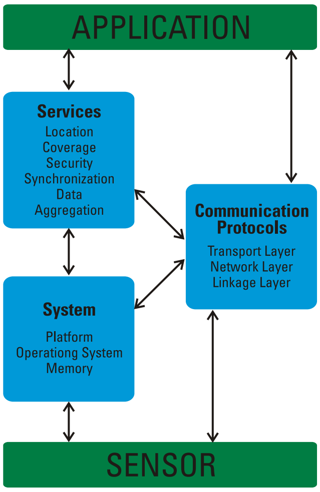

Fig. 1 –Sensor Network Structure [1]

Considering the aspect of mobility, sensor networks will present specific characteristics influenced by the speed at which the sensor moves about, and can suffer interference from natural obstacles, such as landforms, or non-natural obstacles, such as buildings. They can also suffer electromagnetic interferences from other kinds of devices, such as radio and television antennas, or even radars used both in civilian applications for air traffic control and in military applications, such as target or enemy aircraft detection. Another source of interference is the climate.

Mobile sensor networks are called MANETS (Mobile Ad-hoc Networks) and present the following characteristics: networks without any kind of infrastructure and no central point controlling the network or node input/output. Its architecture is not specific to any kind of application. This kind of network was developed to control and monitor a large spectrum of events and applications. Within the group of MANETS, we have the VANETS (Vehicle Ad-hoc Networks). These are vehicle networks and can be considered a specific type of MANET. There are basically two kinds of VANETS: Vehicle-to-Vehicle (V2V) and Vehicle-to-Road (V2R). V2V networks are networks between vehicles, which exchange information directly between themselves. In this kind of network, a given vehicle may, for example, inform the other vehicles that it is slowing down and when these vehicles receive this information, they also slow down, thus reducing the odds of collisions between them. In this kind of network, communication is limited to the range of the radio installed in the vehicle. V2R networks are formed both by the vehicles as well as by the road they are driving on. This means that a vehicle that detects and accident can transmit this information to all the other vehicles by communicating with the highway they are driving on, regardless of whether they are within range or not.

Characteristics such as changes in the topology, mobility standard and what speed the nodes are moving at are important points to be considered for VANETS [2].

Following the same premise of segmentation, sensor networks formed by remotely-piloted aircraft are called FANETS (Flying Ad-hoc Networks). These networks present specific characteristics such as mobility standard, what speed the nodes are moving at, network density, topology changes, propagation model of the communication signals and node processing power [3].

Regardless of the characteristics of the application, the sensor network must be able to meet the minimum service requirements, such as: capacity of delivery, transmission delay, power consumption, fault tolerance and sensing capacity. Other two requirements that must be considered part of quality of service in a sensor network, and that have a direct impact on the consumption of power are: connectivity and the coverage provided by the distribution of the sensor nodes [4]. According to [5], the attention of researchers has been increasingly drawn to this last parameter, since many times the sensor nodes are distributed randomly in the sensing area, and therefore the correct calculation of the ratio between coverage and number of nodes is essential in order to reach the objectives of the application.

Two specific examples for coverage analysis are: 1) meteorological monitoring by government agencies. According to Portal Brasil [6], there are 29 meteorological radars installed in Brazil, whereas the *National Oceanic and Atmospheric Administration* [7], an agency directly connected to the US Department of Commerce, has 160 meteorological radars installed, approximately 5.5 times more; 2) a study in India [8] in which soil sensors were distributed in an area to preview the occurrence of landslides and thus minimize the impacts caused by this kind of catastrophe on society.

However. in both these examples, the fundamental issue for remote sensing to perform as expected is for the event occurring within an area under surveillance to be detected and for this information to be transmitted to a processing system or to the sink node, otherwise, there is absolutely no point in deploying a sensor network. In further consideration of coverage, in both cases, there are fixed sensors, which simplifies the analysis of sensor distribution and, consequently, of coverage capacity. For a broad range of applications, it is impossible to divide the network into disconnected sets of nodes, making the problem of connectivity a crucial issue. For the purpose of this paper, the coverage and the connectivity of the sensor nodes are considered conjointly.

The scope of this paper is to present a mathematical model based on the calculation of the probability of a given sensor in movement to transmit or receive captured information or control messages, such as synchronization messages, using the sensing area, the sensor transmission range and the number of active nodes distributed in the target area as input parameters. To this end, the mathematical model that was developed is presented, considering sensor nodes moving at low speeds (e.g. a flock in a field), and at high speeds (e.g. drones over a forested or natural disaster area) and the deployment of the software model.

The analysis of the exchange of synchronization messages is particularly interesting, since sensor network applications need temporal synchronization and the possibility of accomplishing this synchronization through messages means minimizing the cost and power consumption of the sensors, since no mechanisms, such as GPS, are needed.

This article is organized as follows: Section 2 provides an overview of papers related with the sensor network coverage issue, section 3 describes the problem and the mathematical model that was developed, while section 4 provides details regarding the simulation, the software developed to perform this simulation, the validation of the software and a discussion regarding the achieved results. And finally, section 5 presents the conclusions of the paper.

2 Related Works

Coverage, connectivity and power consumption in a sensor network are very closely correlated. While leaving many sensor nodes in a state of dormancy increases the power efficiency of the network as a whole, it also reduces both the sensing area coverage as well as its connectivity [9].

An area is considered as having coverage when each point in the area to be monitored is under the surveillance of a sensor node, while a wireless sensor network is considered connected if each pair of sensor nodes is able to communicate directly or indirectly with other sensor nodes, with the purpose of discovering a minimum subset of active sensor nodes for the captured data to be sent to the processing system or the sink node [10].

According to [11], the coverage of a sensor network can typically be classified in three ways: A) coverage of an event - where the objective is for sensor to capture any events that take place; b) area coverage – where the objective is to cover the greatest possible percentage of the area being monitored (covered) by a sensor; c) coverage by barrier – the sensors are placed in such a way as to form a barrier, meaning any intrusion penetrating the sensing area is monitored. Still according to [11], the interest in researching coverage and connectivity in sensor networks has increased recently, however these studies have focused on a single aspect, such as mobility [12] and [13], coverage quality or the installation model of the sensors [14].

The *Particle Swarm Optimization (*PSO) is an algorithm widely used in sensor networks due to the balance it provides between complexity and high-quality optimization [15]. The work developed in [16] uses PSO to find the distribution of sensors that ensure the best coverage, while the Voronoi diagram [17] is used to evaluate how adequate the solution is. The algorithm is evaluated by simulating different scenarios, and the simulation results suggest that the proposed algorithm ensures good coverage with the best temporal efficiency. However, due to the iterative characteristics of the algorithm and its use in real time applications that require frequent optimizations, it is not recommended.

In [18] the *Harmony Search Algorithm* (HS) is used to improve the distribution of sensors within a specific sensing area in order to identify the optimal number of sensors, as well as their best location to maximize sensor network coverage. The results achieved by the authors demonstrate that the use of the HS algorithm ensures a 25 percent improvement in coverage compared to randomly distributed sensor nodes. However, the simulations performed did not take network connectivity, power consumption and the existence of obstacles that might interfere in the distribution of sensor nodes into consideration.

The study developed by Romoozi et al. [19] sets forth an optimization of power consumption, while preserving the area covered by the sensing node. To accomplish this, the *Genetic Algorithm* [20] was used for a better distribution of the sensor nodes, and the *Fuzzy C-means Algorithm* [21] was used to group the nodes into clusters. The obtained results demonstrate a balanced consumption of power, with no detriment to the desired coverage.

The study developed by [22] presents a model with mobile instead of static sensors to monitor a segment of the sensing area, forming a kind of dynamic barrier. For this model to achieve satisfactory results, that it, to actually detect events occurring along the barrier, the sensors must have previous knowledge of the probabilities of intrusions taking place, thus making it possible to determine the mobility strategy of each sensor more precisely. However, this kind of prior knowledge is impossible for many applications, resulting in a very low detection rate.

Unlike the above-mentioned studies, this paper presents a model for analyzing sensor network coverage based on two combined factors: mobility and connectivity. The developed model considers node movement at high and low speeds in an area coverage and preserved connectivity model.

In this case, for an event to be considered detected, it is not enough for it to have taken place within the range of a given sensor. The sensor responsible for its detection must also have connectivity with the sink node or processing station. Connectivity means the ability of a given sensor node to transmit and/or receive messages from a sink node directly or by means of other sensor nodes.

3 Description the Problem and Methodology

In a mobile sensor network, the means of communication are variable, because neighboring relationships are adjusted dynamically according to the spatial distribution of the sensor nodes.

Depending on the degree of mobility of the sensor nodes, this distribution may be more intense. For example, in a FANET, the dislocation speed of the nodes can be as high as 900 Km/h (around 560 mph). This means that a remotely-piloted aircraft, at a given instant in time, might be in a position where there is no viable route to the sink node, with a direct impact on the ability of the system as a whole to receive the information captured by this sensor. Therefore, a part of the area will not be covered during a given instant of time, t. Another factor that can impact the ability of transmitting messages is the state a given sensor might find itself in at a given instant in time, since in sensor networks, power consumption is a critical factor and one of the ways to save power is turning off the sensor node's radio for certain intervals of time.

3.1 Mathematical Description of the Model

The mobility of a sensor node is treated based on the calculation of the probability of a given sensor node being able to transmit or receive a message to or from the sink node, with the sensor nodes in movement.

The capacity of a message being delivered on the network depends on the existence of a path defined by the routing protocol. This dependence is related to the distance between the sensor nodes and the speed that these nodes are moving at. The implementation of this model is based on the assumption that the mobility of the sensor nodes always takes place within a defined sensing region and a known area.

Considering the t_0 and t_1 steps of a discrete event simulator, there is a probability that the sensor node will be connected at t_1, assuming that it was connected at t_0. Figure 2 presents two distributions of sensor nodes in a sensing area. At t_0, the sensor nodes were distributed randomly. Therefore, there is a probability that a sensor node will be connected at t_0, defined as Pt_0. At the next t_1 instant, due to the mobility of the sensor nodes, the distribution of the sensor nodes in the sensing area will not be the same. Therefore, there is a new probability that a sensor node will be connected, defined as Pt_1.

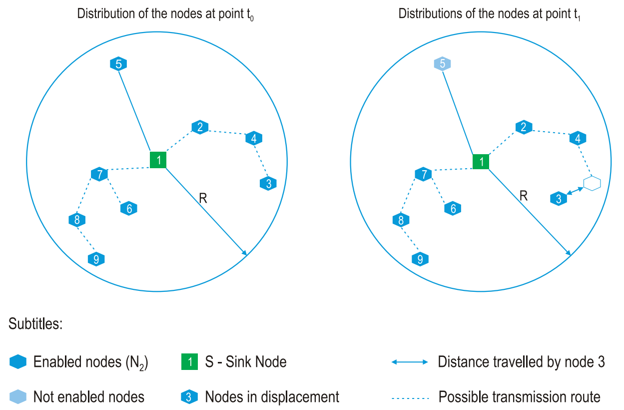

Fig. 2 – Displacement of Sensor Node 3 at Times t_0 and t_1

This probability is analyzed in two cases: high sensor node dislocation speed and low sensor node dislocation speed. In the case of high speeds, this probability is a random variable that depends on the coverage area, while for low dislocation speeds, the probability of the sensor node being connected depends on the coverage area and the state of connectivity in the previous step.

The characterization of high and low dislocation speeds and the development of a model to calculate these probabilities is presented in more detail in the following sections.

3.1.1 High Sensor Node Dislocation Speed

To determine if the dislocation speed of a sensor node, called V_d, is high, the distance traveled, called D_p, by this sensor node during a simulation step is compared with the range of the sensor node's radio transmission, called r, in a time interval, Δt (simulation interval). Figure 3 illustrates a possible sensor node distribution in a sensing area and the radio transmission range of these same sensor nodes.

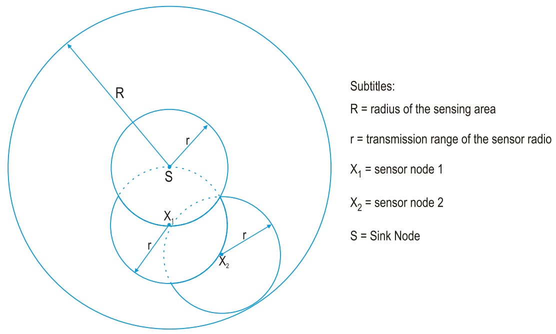

Fig. 3 – Sensor Node Transmission Range

The distance traveled by the node (D_p) is calculated by the following equation (1):

$$Dp = Vp / \Delta t, \text{ where } \Delta t = (t1 - t0) \quad (1)$$

If D_p is greater than r ($D_p > r$), than the dislocation speed is considered high and the probability of a given x_i sensor node receiving a synchronization message at t_0 is the same as the probability of this same sensor node receiving a message at t_1. This is called the *Prandom* Probability, as shown in equation (2) below:

$$Pt0\ (xi) = Pt1\ (x1) = Prandom = Ptk\ (x1) \quad (2)$$

This means that the probability of there being a valid path for forwarding the message to the x_i sensor node at t_0 is the same at t_1, and therefore, it is as if at each instant of time t_i the sensor nodes were randomly distributed again in the defined area.

3.1.2 Low Sensor Node Dislocation Speed

For the sensor node dislocation speed to be considered low, the distance traveled by the sensor node at Δt ($t_1 - t_0$) must be less than the distance of the radio transmitter's range ($D_p < r$) of the same sensor node, considering that at t_0, there is a valid communication route between the sensor node and the sink node.

The probability of sensor node x_1 receiving a message (e.g. synchronism message) at t_k, assuming that this message was received at t_0, is calculated based on Figure 4.

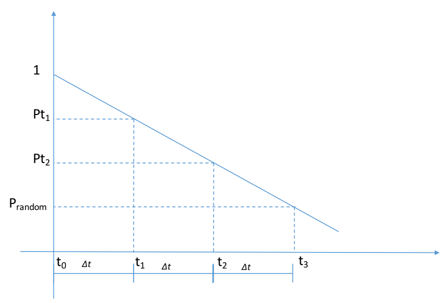

Fig. 4 – Probability of Receiving Messages

Considering that the number of steps needed for a given sensor node to lose the communication due to the traveled distance is given in equation (3):

$$n = rouddowns\ \frac{r}{Dp} \quad (3)$$

and that the declivity coefficient (m) of the line is given by the equation (4):

$$m = \frac{1 - P_{random}}{3\ \Delta t} \quad (4)$$

and Δt is the integration time interval, that is, $\Delta t = t_1 - t_0 = t_3 - t_2 = t_i - t_{i-1}$. The probability of sensor node x_1 receiving a message (e.g synchronism message) at t_k, for n=1 and n=2 is given by equations (5) and (6) respectively, and its generalization is given by equation (7).

$$Pt_1(x_1) = \frac{2 + P_{random}}{3} \quad (5)$$

$$Pt_3(x_1) = 1 - \frac{1 - P_{random}}{3\ \Delta t}\ .\ 3.\Delta t = P_{random} \quad (6)$$

$$Pt_k(x_i) = 1 - \frac{1 - P_{random}}{n\ \Delta t}\ .\ k.\Delta t \qquad 0 < K \leq n + 1 \quad (7)$$

3.1.3 Communication Probability - Prandom

The probabilities of an x_i sensor node establishing communication with the sink node when moving at high or low speeds is given by P_{random}. To calculate P_{random}, the following hypotheses are assumed:

- *Hyphotesis 1)* Given a circular area with an R radius and sensor nodes with r transmission range, there will be N_1 sensor nodes distributed in this area and N_2 ENABLED sensor nodes, where $N_2 \leq N_1$.
- *Hypothesis 2)* If the distance between two sensor nodes is less than r, the communication is possible. Otherwise, the sensor node can neither receive or transmit the message.
- *Hypothesis 3)* The sensor nodes communicate with a central sensor node (sink node). This central sensor node is located at the center of the circular area, with radius R and is fixed. This central node exists in all simulation and modeling condition and is not included in the subset of active nodes (N2).

An expression for P_{random} is deduced based on some simple cases, where the Theory of Probabilities is used to validate the results.

Case 1 – Given the fact that there is a one central sensor node (sink) and one x_1 sensor node is distributed randomly in an area with R radius, what is the probability that this x_1 sensor node will establish direct communication with the central sensor node, if the positioning of this sensor node is determined randomly ? Figure 5 represents the described situation.

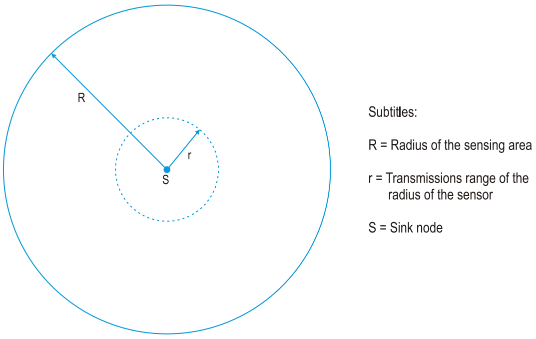

Fig. 5 – Sink Node Transmission Range

In this case, the probability of the x_1 sensor node establishing a communication with the sink sensor node is the probability of the x_1 sensor node being within the central node's transmission range, radius r, is given by equation (8).

$$P(x_1) = \left(\frac{r}{R}\right)^2 \qquad (8)$$

Case 2 – Considering 2 ENABLED sensor nodes (N_2 = 2), what is the probability of node sensor x_2 communicating with the sensor node (S) through sensor node x_1?
Figure 6 represents the described situation. The illustration (a) represents the transmission range from the sink sensor node (S) and the x_1 sensor node.
Considering that sensor node x_1 is at a distance of ρ from the sink sensor node (S), meaning that $0 \leq \rho \leq$ r, a second node sensor, x_2, can establish communication with the sink sensor node by of sensor node x_1 if, and only if, sensor node x_2 is positioned highlighted area in Figure 6 (b).

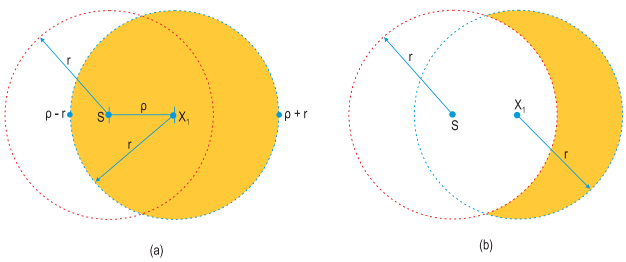

Fig. 6 - Total Transmission Range (Sink Node + Node X_1)

If Ds(Xk) is the distance from a given *Xk* sensor node to the sink sensor node (S), then for the situation presented in Figure 6 $Ds(x_1)$ = □. Since □ is a continuous random variable, the probability of $Ds(x_1)$ being equal to zero can be determined as null. Therefore, D_S can be considered a small distance of width $\partial\rho$.
Equation (9) present the probability of the x_2 sensor node communicating with S via sensor node x_1

$$P\{\rho \leq D_s(x_1) \leq \rho + \partial\rho \mid x_1 \text{ communicates with S}\} \,.\, P\{x_2 \text{ communicates with } x_1\} = \frac{2\,\pi\,\rho\partial\rho}{\pi r^2} \cdot \frac{r^2}{R^2} = \frac{2\,\rho\partial\rho}{R^2} \qquad (9)$$

If E_p is the event for which the above probability was calculated and presented in equation (9), $X_2 \rightarrow X_1$ the event of sensor node X_2 communicating with sensor node x_1, and $X_2 \rightarrow X_1 \rightarrow X_S$ the event of sensor node X_2 communicating with central sensor node S exclusively by means of sensor node X_1. Then $P\{X_2 \rightarrow X_1 \rightarrow X_S \mid E_p\}$ is equal to the probability of sensor node X_2 being in the highlighted area in Figure 6 illustration (b).
The calculation of the highlighted area is achieved by the area of the circle formed by the range of sensor node X_1, which is equal to $A_{X1} = \pi.r^2$ minus the area established by the intersection of the circles formed by the range of sensor nodes S e X_1, called $A_{intersec}$.

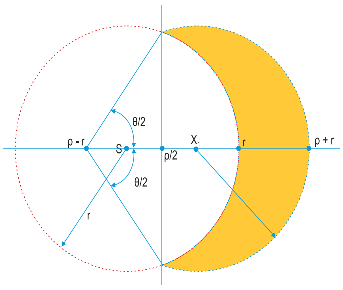

Fig. 7 – Area added by node X_1 to the range of Sink Node

The area highlighted in Figure 7 is the result of equation (10) below.

$$A_{highlighted} = \pi\, r^2 - r^2\,(\theta \text{ - sen } \theta\,) \qquad (10)$$

where:
$\theta = 2$ arccos $(\rho / 2r)$ ou $\cos(\theta/2) = (\rho / 2r)$

The intersection area ($A_{intersec}$) is calculated by the equation (11)

$$A_{inteserc} = r^2\,(\theta \text{ - sen } \theta) \qquad (11)$$

For borderline cases, $\rho = 0$ and $\rho = r$, the intersecting area will have the following values:
For $\rho = 0$, then $\theta = \pi$ and, therefore, the intersecting area will be calculated by the equation (12):

$$A_{intersec} = \pi r^2 \qquad (12)$$

For $\rho = r$, then $\theta = (2\pi / 3)$ and, therefore, the intersecting area will be calculated by the equation (13):

$$A_{intersec} = 1{,}22 \cdot r^2 \qquad (13)$$

One can thus remove the condition of the probability presented in equation (9) and eliminate parameter □□, and obtain from this probability, based on parameters R (radius of the sensing area) and r (transmission range of the sensor node), resulting in equation (14), which shall be used to validate the model.

$$P\{X_2 \rightarrow X_1 \rightarrow X_S\} = \left(\frac{\pi r^4}{R^2} + \frac{4}{3}\frac{r^4}{R^2}(\sqrt{3} - \pi) + \frac{r^4}{R^2}\left[\frac{16}{3} - 2\sqrt{3}\right]\right)\frac{1}{\pi R^2} \qquad (14)$$

3.1.4 Generalization of P_{random}

Case 2, presented and calculated in the previous section, is based on the assumption of a single x_1 sensor node, positioned within the transmission range of the *sink* transmission node. However, this is just one of the possibilities, since several sensor nodes can be positioned and enabled at a given instant t, within the transmission range of the *sink* sensor node, given by $\pi \cdot r^2$
In this case, the x_2 sensor node, positioned outside the range of the sink sensor node, would be able to receive the message transmitted by S, through an X_i sensor node positioned anywhere within the S range.
For the generalization of the P_{random} calculation, four cases of sensor nodes positioned within the range of central sensor node S are presented.
Assuming there are several sensors nodes that can communicate directly with the sink node S, let's call these sensors nodes from $X_{1'}$, $X_{1''}$, $X_{1'''}$ and so on. Figure 8 presents two of the four cases, which are: case (a) which is the same situation as the one discussed in the previous section; and case (b), where a new sensor node positioned within the range of central sensor node S, called $X_{1'}$ is presented, resulting in a new possibility

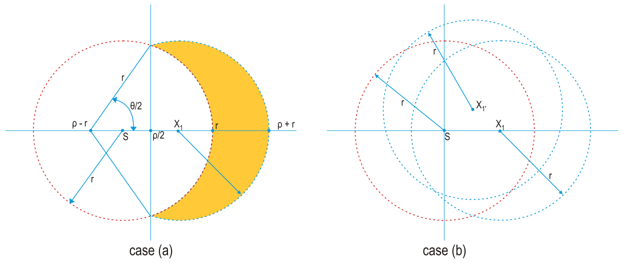

Fig. 8 – Possible Positioning of the node X_1 - Cases (a) and (b)

For both cases (a) and (b), the area added to the area determined by the range of sink sensor node S needs to be determined.
For case (a), the calculation of the total range of transmission is given by the area of central node S, called of fundamental area, plus the area incorporated by the range of sensor node X_1 (highlighted area). Therefore, the added area is equal to the area determined by the range of sensor node X_1, which is $A_{x1} = \pi \cdot r^2$, minus the intersecting area, or common area between sensor node X_1 and the sink sensor node S, called $A_{intersec(X1/S)}$.
The $A_{intersec}(X_1/S)$ because of the distance between sink sensor node S and sensor node $X1$ (parameter □), is given by equation (15).

$$A_{intersec(X_1/S)} = r^2\left(2\ arcocos\frac{\rho}{2r} - sen\left(2\ arcocos\frac{\rho}{2r}\right)\right) \qquad (15)$$

Therefore, the total range will be calculated by equation (16):

$$A_{total\ range} = A_{fundamental} + A_{X1} - A_{intersec}(X_1/S) \qquad (16)$$

For case (b), the range added by sensor node $X_{1'}$ to the total range, considering sensor nodes S, X_1 and $X_{1'}$, is calculated by equation (17):

$$A_{added\ by\ X1'} = A_{X1'} - A_{intersection(X1'/S)} - A_{intersection(X1'/X1)} + A_{intersection(X1/X1'/S)} \qquad (17)$$

The calculation of the intersecting area of the three circles is an issue that has already been solved. The first stage must be the calculation of the triangle inside the intersecting area. This triangle is presented in Figure 9 and its area is calculated by the Heron formula, presented in equations (18) and (19).

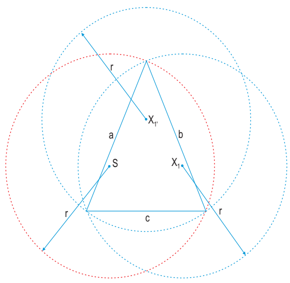

Fig. 9 – Triangle formed by the Intersection of three circles

$$A_{abc} = \sqrt{S\,(S-a)(S-b)(S-c)} \qquad (18)$$

$$S = \frac{1}{2}\,(a+b+c) \qquad (19)$$

Edges a, b and c of the triangle can be calculated based on the intersecting points, equation (20) presents the calculation for edge a, and edges b and c are calculated in the same manner.

$$a = \sqrt{(X_{i1} - X_{i2})^2 + (Y_{i1} - Y_{i2})^2} \qquad (20)$$

The total intersecting area of the three circles is given by equation (21)

$$A_{intersection(X1/X1'/S)} = \sum_{n=1}^{3} r_n^2 sen^{-1}\left(\frac{a_n}{2r_n}\right) - \sum_{n=1}^{3} \frac{a_n}{4}\sqrt{4r_n^2 - a_n^2} + \sqrt{S\,(S-a)(S-b)(S-c)} \qquad (21)$$

The third case, called case (c) is presented in Figure 10. For this case, a third sensor node called $X_{1''}$ is added to the range area of central sensor node *S*.

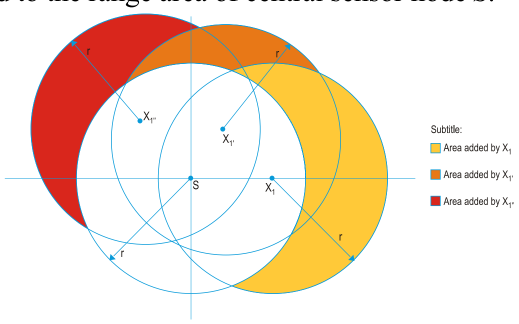

Fig. 10 - Possible positioning of node X_1 – case (c)

One can observe that the area added by sensor node $X_{1''}$ does not depend on the area added by the first sensor node X_1, and is determined by equation (22).

$$A_{added\ by\ X1''} = A_{X1''} - A_{intersection(X1''/S)} - A_{intersection(X1''/X1')} + A_{intersection(X1''/X1'/S)} \qquad (22)$$

The fourth and last case is presented in Figure 11, where a fourth sensor node, called $X_{1'''}$ is added within the range of sink sensor node *S*.

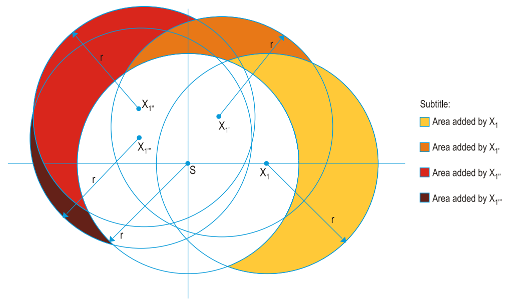

Fig. 11 – Possible positioning of node X_1 – Case (d)

As in the case of the previously mentioned cases, the added area for this new sensor node is given by equation (23).

$$A_{added\ by\ X1'''} = A_{X1'''} - A_{intersection(X1'''/S)} - A_{intersection(X1'''/X1'')} + A_{intersection(X1'''/X1''/S)} \qquad (23)$$

Based on the analysis of the four cases presented, one can determine that the area added by a given X_{1i} sensor node, positioned within the range of the S central sensor node, can be generalized by equation (24), displayed below, as long as all X_{1i} sensor nodes are ordered ascending from angle α, as shown in Figure 12.

$$A_{added\ by\ X1i} = A_{X1i} - A_{intersection(X1i/S)} - A_{intersection(X1i/X1i-1)} + A_{intersection(X1i/X1i-1/S)} \qquad (24)$$

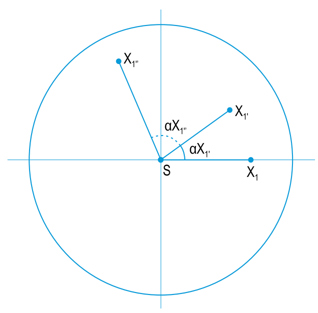

Fig. 12 – Angle αorder

Thus, the generalization of P_{random} to establish communication between sensor node X2, positioned outside the range of central sensor node S, and the sink sensor node, by means of sensor node X1, positioned within the range of the sink sensor node, is calculated by equation (25)

$$P_{random} = \frac{\sum \text{Added Areas} + \text{Fundamental Area}}{A_{SR}}$$

(25)

where A_{SR} is the sensing area and is calculated by: $A_{SR} = \pi R^2$

4 Simulation

To analyze the probability of a given node to establish communication with the sink sensor node, a software was developed in MATLAB, based on the Monte Carlo Method. This statistical simulation method uses a random sequence of numbers to develop simulations. In other words, it is considered a universal numerical method to solve problems by means of random sampling (approximating the solution).

There is no need to write down the differential equations describing the behavior of complex systems for this method. The only requirement is that the physical or mathematical system be described (modeled) in terms of functions of density of probability distribution (FDP). Once these distributions have been established, the Monte Carlo Simulation can begin random sampling, based on them. This process is repeated innumerable times, and the desired result is obtained by means to statistical techniques on a given number of executions (samples) that can vary from dozens to millions of times [23].

4.1 How the simulator works

To execute a simulation, five input parameters are required. They are: 1) sensing area radius, defined in the software as the *radius_area* variable; 2) transmission range radius of the sensor nodes, the *radius_range* variable; 3) number of nodes to be randomly distributed in the sensing area, the *qtd_sensors* variable; 4) percentage parameter of the active sensor nodes, variable *A* (for example, if A is equal to 1, then all N_1 nodes are active); and 5) number of simulations executed for utilizing the Monte Carlo method, the *qtd_executions* variable.

The software is divided into three major execution blocks. The first block defines the positioning of the sensor nodes, chosen randomly. A pair of coordinates (x, y) is chosen by randomly for this. For each pair of randomly chosen coordinates, a check is run to determine if the sensor node is inside or outside the sensing area (Asr) with a radius of *R*. If the node is outside the sensing area, it is discarded and a new draw is performed. If the node is within the sensing area, then the coordinates are stored in a matrix, called the Active Node Matrix.

The purpose of the second block is to ascertain which if the randomly distributed nodes communicate directly with the sink node and which nodes do not (active node x sink node test - Figure 14).

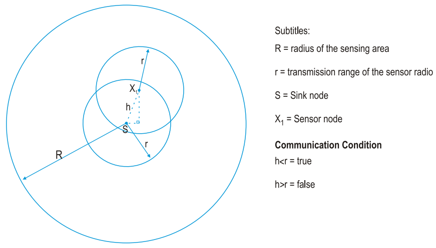

Fig. 13 – Validating Communication between two sensors nodes

After this first test, the nodes are separated into two groups, called C0 (nodes with direct communication with sink node) and NC1 (nodes without direct communicate with sink node), respectively.

The third and last execution block checks if the NC1 nodes communicate with the C0 nodes.

For communication to be considered existent, a given NC1 node must be within the transmission range of a C0 node, as shown in Figure 13.

This communication test is run node by node. In other words, let us say that nodes 1 to 5 communicate directly with the sink sensor node (C0), and nodes 6 to 10 do not communicate directly with the sink node (NC1). The software will then use the node 6 coordinates to validate if there is communication with any C0 node. The first test will be run on node 6 versus node 1, the second, node 6 against node 2, and so forth.

If this test condition is true, in other words, if there is communication, the execution loop is interrupted, the node coordinates are stored in C1 and the next NC1 node will begin to be validated (e.g. node 7 against node 1, node 7 against node 2, and so forth). If a given NC1 node is unable to hold real communication with at least one C0 node, this means the node is positioned outside the range of the C0 nodes, and its coordinates are stored in NC2, which will be tested against the C1 nodes, as represented in Figure 14.

This communication loop test is run until the communication condition of all the nodes distributed in the Sensing Area have been tested, with the final result showing how many nodes communicate with the sink node, regardless of how many hops are needed, and how many nodes do not communicate with the sink node. For nodes that do communicate with the sink node, the software calculates how many hops were needed to establish this communication.

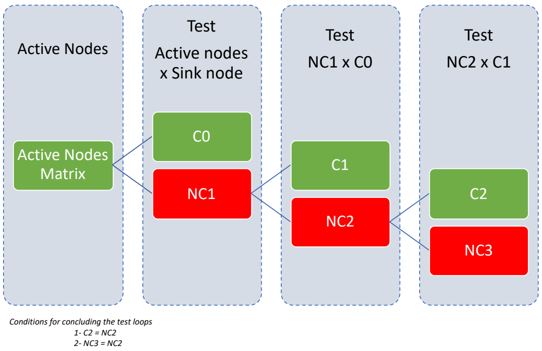

Fig. 14 – Logic of the Communication Test

4.2 Validation

Considering two sensor nodes, X_1 and X_2 with a transmission range of r, randomly distributed in a circular sensing area A_{SR}, with a radius of R and an S sink node, positioned in the center of the A_{SR} area, there is a Probability $P(E_i)$ of 4 possible events occurring, as described in table 1

Table 1 – Possible Events

Event	Sensor X_1	Sensor X_2	Description
E_1	$X_1 \times S$	$X_2 \times S$	X_1 and X_2 do not communicate directly with S node
E_2	$X_1 \rightarrow S$	$X_2 \times S$ $X_2 \times X_1$	X_1 communicates directly with S node X_2 does not communicate with X_1
E_3	$X_1 \rightarrow S$	$X_2 \rightarrow X_1$	X_1 communicates directly with S node; X_2 communicate directly with X_1
E_4	$X_1 \rightarrow S$	$X_2 \rightarrow S$	X_1 and X_2 communicates directly with S node

For each event in table 1, simulations were carried out using the developed software described in section 4, item 4.1. The results were compared with the results returned from the mathematical solution developed and presented in section 3.

Table 2 – Validation Parameters

Parameters	Values
radius_area	R = 4r
radius_range	r
qtd_sensors	2 (X1 and X2 sensors plus sink node)
A	1 (all nodes are actives)

Event E3 is mathematically modeled in section 3, and its extrapolation was implemented in software, as described in section 4, item 4.1 – *How the Simulator Works*, so that fort a given X_i sensor node to establish communication with the sink node through n hops.

Thus, to validate the software, 22 simulations were executed, with a fixed number of repetitions defined for each simulation. The first simulation was repeated 8,000 times and the last 60,000 times. The probability obtained for each simulation is presented in Figure 15 and compared with the calculated probability. The calculated and simulated results for the other events, as well as the mean of the simulated results achieved for event 3 are presented in Table 3

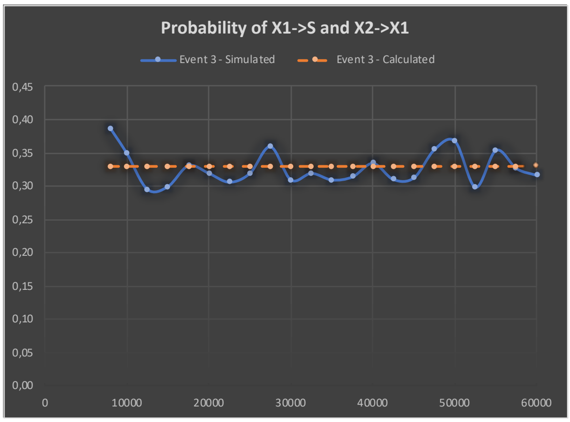

Fig. 15 – Probability of occurrence of the events

Table 3 – Probability per event

Event	Calculated Results	Simulated Results
E_1	87,89 %	87,74 %
E_2	11,39 %	11,56 %
E_3	0,33 %	0,33 %
E_4	0,39 %	0,37 %

The event analyzed in Figure 15 is E3, since this event represents the probability of a sensor node communicating with the sink node through another sensor node ($X_2 \rightarrow X_1 \rightarrow S$). As observed in the displayed graph and in Table 3, the average probability obtained with the simulation of the developed software was 0.33%, while the mathematically calculated probability is 0.33% (using equation 14). This shows that it is possible to use software to achieve simulation results for extrapolated scenarios, in other words, with quantities higher than 2 sensor nodes, without the need of developing complex mathematical models.

4.3 Experiments and Measurements

To carry out the simulation, an application of sensors in an agricultural area was considered, e.g. cattle monitoring. In order to simulate a scenario as similar as possible to a real-life situation, we used a circular sensing area equivalent to a 4000-hectare farm [24], sensors with a 100-meter transmission range and the sink node placed in the center of the Area. Table 4 shows all the input parameters used in the simulation.

In the simulation scenario, the delay in communication between the sensor and sink sensor node is not taken into consideration.

Table 4 – Simulation Parameters

Parameter	Value
radius_area	3570 m
radius_range	100 m
qtd_sensors	Variation between 4000 and 10000
A	1 (All nodes are actives)
qtd_executions	200

The achieved results, shown in figure 16, demonstrate that if 4000 sensors are distributed over an area to be monitored, the probability of a new sensor randomly distributed in this same area being able to establish contact with the sink node is 0.96%. However, if in this same area, there are 7000 sensors, the probability of communication soars to 92%.
This means that for an increase of 75% in the number of sensors, the probability of there being communication is 96.8% higher. On the other hand, to achieve a 99.77% probability of communication, the number of sensors must be increased 2.5 times.
Based on the data achieved with 8000 sensors, the probability of communication is 97.59%. However, to achieve a probability of communication 1.0223 times higher, the number of sensors would have to be increased by 25%. This shows that, for a given number of distributed sensor nodes, there is a saturation of the probability of communication, which leads to the conclusion that any gains from increased sensor node distribution, without evaluating the coverage, is minimal.

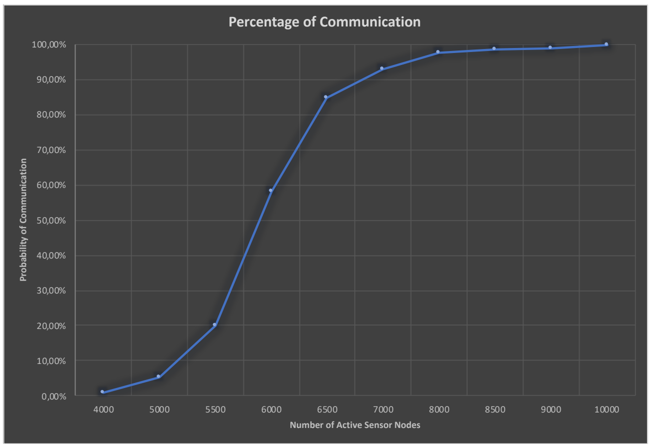

Fig. 16 – Number of Nodes x Communication Percentage

In order to evaluate the behavior of the probability of communication in relation to the variation of the coverage radius of the sensor (*r/R*), another simulation was carried out. For this study, the number of sensor nodes was fixed, and the variation of the sensor node transmission range (r) was executed in steps of 50 meters, starting at 100m (base scenario).
To establish the number of sensor nodes to be used in the simulation, a new simulation was made, with the transmission range of the sensor node fixed at 0.1 of the Radius of the Sensing area, that is, $r = 0,1R$, and the number of active sensor nodes was varied from 100 sensors up, with additions of 100 units. The achieved results are displayed in Figure 17.

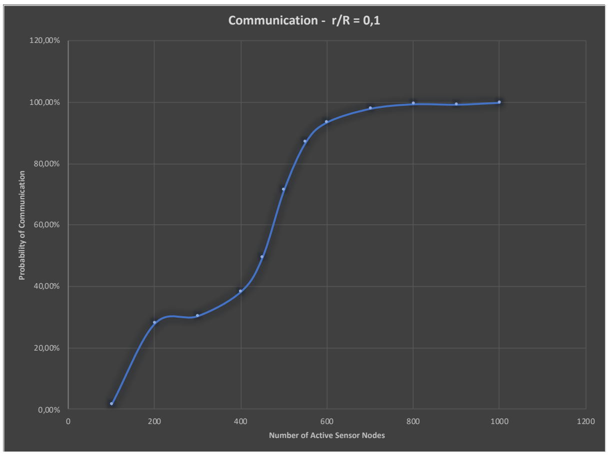

Fig. 17 – Number of Active Sensor Nodes x Probability of Communication

Based on the results achieved with this simulation, it was established that with 450 active sensors, the probability of communication is 49.48% (approximately 50%), and, therefore, this was the amount of sensor nodes chosen to verify the behavior of the sensor nodes based on the relation *r/R*. The behavior of the probability of communication based on the relation (*r/R*) is presented in Figure 18 the summary of the data *R; r; r/R; non communication probability (PNC) and Communication Probability (Pcom)* is presented in Table 5.

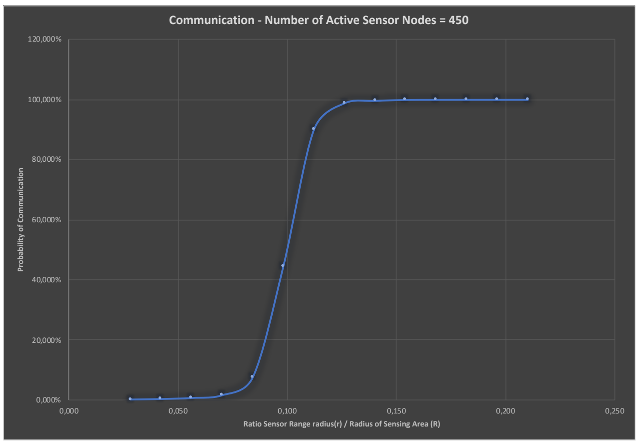

Fig. 18 – Relation r/R x Probability of Communication

An analysis of the results shows that the relation *r/R* is a determining factor only in one segment of the relation *r/R* . Based on the approximate value of *r/R = 0.12*, the probability of communication remains practically the same. This is relevant from the viewpoint of sensor network designers who need specific equipment that have both optimal coverage, which means higher transmission ranges, and sensors with long lasting batteries.

Table 5 – Summary of the r/R data simulation

R(m)	r(m)	r/R	PNC (%)	Pcom
3570	100	0.028	99.894	0.106
3570	150	0.042	99.726	0.274
3570	200	0.056	99.386	0.614
3570	250	0.070	98.460	1.540
3570	300	0.084	92.500	7.500
3570	350	0.098	55.540	44.460
3570	400	0.112	9.998	90.002
3570	450	0.126	1.170	98.830
3570	500	0.140	0.360	99.640
3570	550	0.154	0.068	99.932
3570	600	0.168	0.020	99.980
3570	650	0.182	0.002	99.998
3570	700	0.196	0.001	99.999
3570	750	0.210	0.000	100.000

5. Conclusions

This paper establishes a way to evaluate the probability of communication between a given sensor node and the sink node. The achieved results demonstrate an increasing relationship between number of sensor nodes and the probability of communication, by which it is possible to evaluate the gains in communication by increasing the number of nodes in a sensing area.

This interpretation can be extrapolated, bringing one to the conclusion that the probability of communication demonstrates the coverage percentage of the monitored area.

With the software developed to execute the simulation, it is possible to reduce the complexity of mathematical models to analyze the coverage and communication in sensor networks that require large amounts of sensor nodes.

Another relevant contribution lies in the analysis of the relation between the range of the sensor, the sensing area and the number of nodes, since these factors bear a significant impact on the life of the sensor batteries. This them has been studies over the yea, and the analysis of the *r/R* relation from this perspective contributes to increase the efficiency and minimize the deployment costs of sensor networks.

The proposed model can also be used for different applications, such as FANETs, military applications, the monitoring of vehicles on highways, border surveillance, to name a few.

It is the author's intention, based on the proposed model, to evaluate the impact of the level of synchronization on a sensor node, considering the delay in communication variable and the mobility variable when nodes are engaged in low-speed dislocations.

6. Conflict of Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

7. References

[1] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Comput. Networks*, vol. 52, no. 12, pp. 2292–2330, Aug. 2008.

[2] M. Conti and S. Giordano, "Mobile ad hoc networking: Milestones, challenges, and new research directions," *IEEE Commun. Mag.*, vol. 52, no. 1, pp. 85–96, 2014.

[3] İ. Bekmezci, O. K. Sahingoz, and Ş. Temel, "Flying Ad-Hoc Networks (FANETs): A survey," *Ad Hoc Networks*, vol. 11, no. 3, pp. 1254–1270, 2013.

[4] M. Ilyas and I. Mahgoub, *Handbook of Sensor Networks : Compact Wireless and Wired Sensing Systems*. Boca Raton, London, New York, Washington D.C.: CRC Press, 2005.

[5] C. Zhu, C. Zheng, L. Shu, and G. Han, "A survey on coverage and connectivity issues in wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 35, no. 2, pp. 619–632, 2012.

[6] Portal Brasil, "Portal Brasil - Ciência e Tecnologia," 2014. [Online]. Available: http://www.brasil.gov.br/ciencia-e-tecnologia/2012/11/rede-de-radares-meteorologicos-do-brasil-sera-ampliada-a-partir-de-2013. [Accessed: 07-Apr-2017].

[7] NOAA, "NOAA," 2016. [Online]. Available: https://data.noaa.gov/dataset/noaa-next-generation-radar-nexrad-level-iii-products. [Accessed: 07-Apr-2017].

[8] M. V. Ramesh, "Design, development, and

deployment of a wireless sensor network for detection of landslides," *Ad Hoc Networks*, vol. 13, pp. 2–18, Feb. 2014.

[9] P. K. (Jayoti V. W. U. Hirani and M. (Jayoti V. W. U. Singh, "A Survey on Coverage Problem in Wireless Sensor Network," *Int. J. Comput. Appl.*, vol. 116, 2015.

[10] N. Gupta, N. Kumar, and S. Jain, "Coverage Problem in Wireless Sensor Networks : a Survey," in *International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)-2016*, 2016, pp. 1742–1749.

[11] Y. Wang, S. Wu, Z. Chen, X. Gao, and G. Chen, "Coverage problem with uncertain properties in wireless sensor networks: A survey," *Comput. Networks*, vol. 123, pp. 200–232, 2017.

[12] B. Wang, H. B. Lim, and D. Ma, "A survey of movement strategies for improving network coverage in wireless sensor networks," *Comput. Commun.*, vol. 32, no. 13–14, pp. 1427–1436, 2009.

[13] M. R. Senouci, A. Mellouk, K. Asnoune, and F. Yazid Bouhidel, "Movement-Assisted Sensor Deployment Algorithms : A Survey and Taxonomy," *IEEE Commun. Surv. Tutorials*, vol. 17, no. April 2016, pp. 2493–2510, 2015.

[14] D. S. Deif and Y. Gadallah, "Classification of Wireless Sensor Networks Deployment Techniques," *IEEE Commun. Surv. Tutorials*, vol. 16, no. 2, pp. 834–855, 2014.

[15] R. V Kulkarni, S. Member, and G. Kumar, "Particle Swarm Optimization in Wireless-Sensor Networks: A Brief Survey," *IEEE Trans. Syst. Man, Cybern. - PART C Appl. Rev.*, vol. 41, no. 2, pp. 262–267, 2011.

[16] N. A. B. A. Aziz, A. W. Mohemmed, and M. Y. Alias, "A Wireless Sensor Network Coverage Optimization Algorithm Based on Particle Swarm Optimization and Voronoi Diagram," in *2009 International Conference on Networking, Sensing and Control*, 2009, pp. 602–607.

[17] F. Aurenhammer, "Voronoi diagrams - A Survey of a Fundamental Geometric Data Structure," *ACM Comput. Surv.*, vol. 23, no. 3, pp. 345–405, 1991.

[18] O. M. Alia and A. Al-Ajouri, "Maximizing Wireless Sensor Network Coverage With Minimum Cost Using Harmony Search Algorithm," *IEEE Sens. J.*, vol. 17, no. 3, pp. 882–896, 2017.

[19] M. Romoozi, M. Vahidipour, M. Romoozi, and S. Maghsoodi, "Genetic Algorithm for Energy Efficient & Coverage-Preserved Positioning in Wireless Sensor Networks," *Proc. - 2010 Int. Conf. Intell. Comput. Cogn. Informatics, ICICCI 2010*, pp. 22–25, 2010.

[20] J. H. Holland, "Genetic Algorithms and Adaptation," *Scientific American*, pp. 66–72, Jul-1992.

[21] R. J. Hathaway and J. C. Bezdek, "NERF c-means: Non-Euclidean Relational Fuzzy Clustering," *Pattern Recognit.*, vol. 27, no. 3, pp. 429–437, 1994.

[22] S. He, J. Chen, X. Li, X. Shen, and Y. Sun, "Cost-effective barrier coverage by mobile sensor networks," *Proc. - IEEE INFOCOM*, pp. 819–827, 2012.

[23] R. B. Nasser, "McCloud Service Framework : Arcabouço para desenvolvimento de serviços baseados na Simulação de Monte Carlo na Cloud," Pontifícia Universidade Católica do Rio de Janeiro - PUC-Rio, 2012.

[24] EMBRAPA, "EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária," 2018. [Online]. Available: https://www.embrapa.br/. [Accessed: 10-Mar-2018].

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001