Mobile Wireless Sensor Networks: Modeling and Analysis of Three-Dimensional Scenarios and Neighbor Discovery in Mobile Data Collection

TOMÁS ROBLES¹, BORJA BORDEL¹,*, RAMÓN ALCARRIA² AND DIEGO MARTÍN¹

¹Departamento de Ingeniería de Sistemas Telemáticos. Universidad Politécnica de Madrid. Avenida Complutense nº 30, “Ciudad Universitaria”. 28040 - Madrid (España). Tel. 91 549 57 00 ext. 3035
²Departamento de Ingeniería Topográfica y Cartografía. Universidad Politécnica de Madrid. UPM Campus Sur, Km 7.5 de la Autovía de Valencia. 28031 Madrid, Spain. Tel. 91 549 57 00 ext. 3035
Emails: tomas.robles@upm.es, bbordel@dit.upm.es, ramon.alcarria@upm.es, diego.martin.de.andres@upm.es

Received: June 13, 2015. Accepted: February 3, 2016

WSNs, as important part of the Internet of Things, became an essential field of study in recent years. However, most works use simplified scenarios for its analysis, disregarding variables that affect the data collection process. In this paper, we present a new model for data collection, considering the connection time among the nodes as a work’s main part. In addition, a real three-dimensional model is described. The analysis of this new model will allow us to discover new problems for the WSN operation. We start describing both two-dimensional and three-dimensional scenarios, and continue comparing the operation of different types of mobile nodes (such as walking people or drones). Later, we analyze the connection time, the rate and the range of the most used Wireless communication technologies on WSN (like Bluetooth or 802.15.4). Joining both results, we created a model for data collection and analyzed how to build more efficient WSNs.

Keywords: Neighbor discovery; mobile wireless sensor networks; mobile data collection; three-dimensional scenarios; wireless technologies; analytical model

*Corresponding author. bbordel@dit.upm.es
1 INTRODUCTION

Data collection process from a set of fixed and distributed nodes is, nowadays, one of the most critical topics in communications research [1], [2], [3]. Maximizing the transmission capability with the minimum consumption requires a deep study that has not been fully addressed in the literature.

With the incredible improvement of communication technologies and integration capability, new applications and network configurations have been investigated. In particular, the progress of sensor manufacturing techniques and the increase of the wireless communication efficiency enable the birth of the so called Wireless Sensor Networks (WSN) [4].

Initially, WSN were composed by a collection of fixed nodes, connected in a persistent way by means of any wireless communication technology (e.g. WiFi or Bluetooth). The problem of this configuration lies in the fact that these networks usually are isolated, that is, there is often a lack of Internet connectivity (which is required if, for example, we want to deploy services based on the cloud). To solve this situation most proposals are focused on adding, at least, one mobile node to WSN [5]. In that way, both elements together are named as Mobile Wireless Sensor Networks (MWSN).

Mobile nodes, as they are moving around, cannot establish persistent connections with the fixed nodes. This fact provides a dynamic topology to the network. To face that problem, MWSN implement Delay Tolerant Networks (DTN) principles [5].

As an important component of the revolution of the Internet of Things, MWSN allow the deployment of a wide range of applications [6]: from crop control to Smart Grids environment or public transport surveillance. To make these services work properly, the fixed nodes must be capable to realize two actions while one mobile node is inside its coverage area: to establish a temporal connection with it, and transmit as much stored data as possible.

It is clear that if the connection time (which depends on the wireless technology employed) is too high, there is no time for data collection. That is, in fact, a great problem because, at the end, it can produce the loss of most of the data. For this reason, in the last years, many schemes and algorithms for efficient data transmission (between one fixed node and one mobile node) have been presented [7].

However, all those works are focused either on increasing the transmission rate, or on optimizing the routing algorithms, assuming that, with these improvements, the connection time does not affect the data collection process (it is insignificant compared to total available time) [8]. Furthermore, due to its complexity, three-dimensional scenarios are usually reduced to two-dimensional ones, which is practical, but provides less precise results.

Thus, the novelty of this this work is the proposal and use of a new and more realistic model, for analyzing the data collection process from MWSN. Basically, we improve the model’s precision by using a real three-dimensional
mobile WiReless sensoR neTWoRks 3

geometrical study, and by considering the impact of the connection time between sensor network and mobile layer. Finally, a numerical study will allow us detecting new problems and limitations in the deployment and operation of MWSN.

Characterizing the time necessary to establish the ephemeral connections among the nodes is a task depending on too many variables: the protocol stack, the communication environment, the code optimizations, etc.; for what to verify these assumptions is a really difficult topic. However, if we consider the data link and physical levels, where the number of variables is greatly reduced, it is easier to realize a comparison where clearly appears that, in real scenarios, even only with the neighbor discovery process in those levels, most of the proposed mobile nodes (specially aerial ones) remain in the coverage area too little time to be able to complete the data collection process.

With this philosophy, in this article we propose a comparison in two levels: on the one hand we will study the main mobile nodes considered nowadays (walking people, terrestrial vehicles, drones and Zeppelins). On the other hand, we will obtain the time necessary for neighbor discovery in different wireless technologies, such as Bluetooth, Bluetooth Low-Energy, 802.15.4 and WiFi (only considering the link and physical levels).

From the combination of this information, we will obtain a new model, in which connection time among the nodes does not have to be insignificant, and it is applicable directly in three-dimensional scenarios. With that, our objective is to bring out new challenges in MWSN, such as the use of low-energy technologies in three-dimensional scenarios.

As we will see, only slowest mobile nodes (e.g. walking people) or highest range radio technologies (e.g. WiFi) guarantee the transmission of great quantities of data. In other scenarios, a neighbor discovery process has to be studied because it can be decisive (especially in three-dimensional networks).

The rest of the paper is organized as follows: in Section 2 we present, classify and review the main works over MWSN and neighbor discovery in wireless technologies. Section 3 presents general model construction. In Section 4, the two-dimensional and the three-dimensional cinematic models are deducted, the involved variables are presented and the different mobile nodes, and their application to MWSN, are showed and studied. In that way, Section 5 presents a study of wireless communication technologies commonly used on WSN. Section 6 evaluates all the previous deductions and joins the obtained results, whereas Section 7 provides some conclusions and future work.

2 RELATED WORKS

In the last years, data collection from MWSN has received a lot of attention [7]. However, most works are focused on network level. For example, there is a wide range of options in relation to the routing protocols: from the ones
based on TCP/IP principles [9], to the ones which employ geographical information from the scenario in which the network is placed to select the proper path [10] [11].

Regardless, some works about data collection modeling can be found [12] [13]. With respect to this previous works, we include in this article two main new contributions: on the hand, we consider physical variables such as the mobile node’s speed and make our calculations over a real three-dimensional scenario; on the other hand, we take into account the time need for neighbor discovery. Moreover, related articles used to study one specific scenario, generally based on high node density environments [14] [15]. Nevertheless, in these cases, the probability of interference and/or losses is really high, therefore the time needed to stablish a connection also increases, and it may not be possible to disregard it (even if it has done for simplicity). Definitely, we need a model which considers the neighbor discovery process.

Other articles, though, are focused in empirical studies over real devices [16], introducing sometimes new variables as the power consumption [17] or the processing time [18]. In these cases, it is assumed an interference free scenario, where the node density is low enough for only trying to stablish one connection at the same time [19] (if sensors are dispersed, it is reasonable to think that only one node is in the mobile node’s coverage area at the same time). Although we are going to study a similar scenario, these works do not often analyze the communication systems or processes, so they do not consider main physical variables as the mobile node speed or the quantity of data that it is possible to transmit. The objective of our work is to contribute to solve this pending challenge.

Furthermore, in the lasts months, the study of neighbor discovery process in MWSN is becoming more important. Some works, as [20], are focused on studying the connection time in MWSN from an electronic point of view (power consumption, shared resources, etc.). Others center new proposals and standard reviews (theoretical and practical), in the field of wireless communications. The objective of these reviews is always reducing the connection time, although this can be made either by adjusting the variable proposed in the standard [21] or modifying the way in which the protocol works [22]. Despite all this, any work has analyzed how these new studies or standards affect the amount of data which can be collected from a WSN. This article tries to fill that void.

Finally, with respect to three-dimensional scenarios, around seven years ago they received a lot of attention, and many articles about their connectivity and coverage problems appeared ([23], [24]). However, nowadays, most articles use projected systems which reduce one dimension of the scenario (becoming it in a two-dimensional one). In this work, we want to return to the initial approach, and we are using a real three-dimensional scenario in which, for first time, neighbor discovery process is going to be considered. Moreover, the concept of three-dimensional MWSN has been really little
worked (probably due to the necessity of using physical models to study them), and most works related to three-dimensional scenarios only cover three-dimensional WSN [25] [26] (where mobile nodes are not considered).

In conclusion, it is greatly difficult to find, in present publications, the knowledge for being able to analyze neighbor discovery in the data collection process from MWSN deployed in three-dimensional-scenarios. We assume, however, that this study is basic to discover new problems in MWSNs, and to deploy networks based on next-generation wireless technologies (as, for example, Bluetooth Low Energy). Therefore, we present a three-dimensional model for mobile data collection, which takes into account the time employed in neighbor discovery.

3 MODEL CONSTRUCTION

In this section, we discuss about data collection process in MWSN. We describe the environment where we considered the deployment of the WSN and propose a logical model in three layers for this scenario. Besides, we present the basic assumptions about the fixed and mobile nodes. Finally, in order to identify the variables our model has to take into account, we study the time distribution among the task carried out in the data collection process.

Our scenario is designed for natural environment where sensors are deployed for gathering information such as temperature, humidity, presence, etc. The node density is low, because the natural environment does not justify a higher number of elements. In this initial study we do not consider the transmission of images, video, sounds or other information requiring large amount of data.

We are also assuming we need a constant monitorization of the environment, but we do not need real-time processing (as, for instance, in some agricultural management systems [27]). Therefore, deploying an ad hoc sensor network and getting data collected by a mobile node (one per day, per week... depending on the sensors’ memory size, the sensorized area, etc.) is the most efficient and least expensive technical solution. Other solutions, such as including the sensors in the mobile node, are more costly, even unrealizable in practice, so they will not be considered.

Three layers are identified in this scenario: sensor network, mobile layer and data processing (Figure 1).

Sensor network: The sensor network is distributed over a wide area, providing sensors, data gathering and support for wireless communications, but without communication infrastructure. In general, sensors may be distributed in a random way (if all geographical points are equally important) or in a planned way (if some points are more important as, for example, if we want to control the trees in a plantation). Therefore, we cannot guarantee that we
know every position of the fixed nodes. In this work we are assuming we do not know neither the sensors’ position nor network topology or the number of nodes.

Finally, in relation with the sensors’ characteristics, our model allows employing any type of device. That means we do not have to be worried about values such as the sensing range, the sensing period or the monitored variable; because our proposal is independent from sensors’ features, and only considers communication parameters.

**Mobile layer:** considers a mobile device (or mobile node) that is moving around a sensorized area and, using DTNs based communications, collects available sensor data. The mobile node can be single persons, cars (SUV) or air equipment such as drones or airships. This mobile equipment collects information and translates them to the data processing layer.

Due to its position, mobile layer must implement two different communication interfaces. On the one hand, it must be capable of using the sensor’s communication technology (in which we focused in this work); and, on the other hand, it has to establish a successful communication with the processing layer. In relation with this second interface, many technologies can be used, but it seems clearly enough mobile technologies are the most effective solution [28] (especially for its great communication range).

**Data processing layer:** this layer receives data from mobile layer, stores and processes it. This layer can be based on any standard technology like central-
ized services or cloud services, because in our scenarios it is not required real
time communication neither feedback among sensors and data processors. A
deep analysis of all possible solutions should be based on the definition and
characterization of different scenarios, but in our case, we are going to work
over some common characteristics for all cases, which are:

1. Lack of infrastructure for supporting communications and power supply.
2. WSN deployed for communication with sensors and actuators.
3. Usage of DTN technologies for communication between sensor layer
   and mobile layer.
4. Usage of specific protocols for establishing WSN’s topology in an auto-
   nomous way, and for routing data from the origin to the nodes which can
   transmit them to the mobile nodes.
5. Modeling of a three dimensional space: we are going to consider mobile
   nodes at ground level, and others that move on the air such as drones.
6. The movement of the mobile nodes can be defined as: (a) free, when
   there is not external control; (b) controlled, when it is possible to plan the
   and route for them and; (c) limited to one or several specific routes inde-
   pendently if they are free or controlled.

The first and second points limit the kind of technologies we can use,
and more specifically the lack of power supply limits the service time of the
network, which is a key element for the usability of such kind of solutions
in real [29].

The third and fourth points explain how the mobile nodes can collect data
from sensors. On the one hand, as mobile nodes are not available all time, we
must include Delay-Tolerant Network (DTN) technologies to allow reliable
and efficient communication. On the other hand, as sensors are distributed in
a random way, we cannot guarantee the mobile nodes pass near all fixed
nodes. That is why we need a protocol able to identify the sink nodes (those
can communicate with the mobile nodes), and able to route data to them
when one mobile node appears (and all this with minimal time consumption).
Concepts as compression of information may be the key element [30].

Finally, the fifth and sixth points are directly generated by the require-
mements imposed by scenarios for agriculture and natural areas. There, move-
ment of mobile nodes is not limited to the ground, and we can take advantage
of drones and other aerial equipment. In addition, there is a great diversity
of mobile nodes from tourists or workers performing their specific activities
to specific equipment deployed and operated for performing as mobile
nodes [10].

In general, any equipment with wireless communication capabilities,
when none transmission is running, can behave in two different ways: or it
maintains the interface powered and actively hearing for a new communica-
tion, or it turns off the communication interface (waiting for a signal which wakes it up). In the first case three tasks must be executed: establish a connection, transmit data and close the connection (see Figure 2a). In the second case, to the three previous tasks, one more has to be added: the awakening process (see Figure 2b).

Although turning off the communication module is the most efficient behavior in terms of power consumption, not all wireless access technologies support this possibility by default. Then, in order to obtain comparable results, we are going to consider the wireless interfaces of the fixed and mobile nodes are always available.

The total time in which a fixed sensor and a mobile node can communicate ($T_{in-cover}$) is a function, basically, of three variables: maximum range for the Radio Access Technology (RAT) selected, called as $R$; the speed of the physical node ($v(t)$), which can vary with time, and the geometry of the scenario. In Section 4 we will discuss the analytical expression of this function.

The time needed to establish a connection, $L_{open\ RAT}$, depends on the Radio Access Technology selected. Usually, the connection can be considered as established when a physical signal (beacon) is received. However, other times, it is necessary to send a MAC message.

At last, the time for a node to close a connection ($L_{close\ RAT}$) depends also on Radio Access Technology selected. Nevertheless, all protocol stacks implement procedures to close a connection even when the connection is down. Therefore, in our model, we are going to suppose fixed nodes close the connection when the mobile node is outside their coverage area (in order to exploit the available time to transmit as much data as possible).

FIGURE 2
Distribution of the $T_{in-cover}$ time among different tasks. (a) In the case of keeping the wireless interface on all the time. (b) In the case of turning off the communication module when none transmission is running.
Then it is possible to obtain (1):

\[ T_{\text{in-cover}} = T_{\text{data collected}} + L_{\text{open RAT}} \]  

(1)

Now, considering the real rate offered, at each moment, at physical layer by the RAT \( C(t) \), we have the total amount of data collected is (2):

\[ Data_{\text{collected}} = \int_0^{T_{\text{data collected}}} C(t) \, dt \]  

(2)

Thus, finally, the expression which models the amount of data it is possible to transmit between a fixed sensor and a mobile node is (3):

\[ Data_{\text{collected}} = \int_0^{T_{\text{in-cover}}(R, v, \text{geometry}) - L_{\text{open RAT}}} C(t) \, dt \]  

(3)

In the next section we are describing the value of \( T_{\text{in-cover}}(R, v(t), \text{geometry}) \), and in Section 5 we are discussing the value of \( C(t) \) and \( L_{\text{open RAT}} \).

4 CINEMATIC MODEL AND MOBILE NODES STUDY

Once we have described the natural environment and the network logical structure, it is necessary to design a model to calculate the geometrical variables for the mobile nodes movement, and to estimate the total available time. This model includes the structure and number of nodes of the WSN, the number and location of the sensor and actuators, and the number, frequency and speed of the mobile nodes.

We base the physical model on a structure [31], where: the underlying topology on which sensors, mobile nodes and access points are placed is assumed to be a discrete and finite two-dimensional grid; only a fraction of the grid points are occupied by sensors and access points. The access points are modelled to be uniformly spaced on the grid while the sensors are randomly distributed. The mobile node motion is modelled as a simple random walk on the grid; the mobile nodes communicate with the sensors or access-points only when they are co-located at the grid points. Mobile nodes do not have problem of memory and we are going to assume in this first analysis that nodes have enough memory space for storing the data.

Specifications over the general initial model:

1. We assume that the mobile nodes can connect directly with the service layer by the use of suitable technology. Therefore, we assume every packet successfully transmitted to the mobile node reaches the service layer.
2. As power resources, components’ size and weight and, in general, on board equipment’s capabilities are higher than the sensor nodes’ resources; we are assuming the uplink is the limiting.

3. We are going to use a three dimensional model as described below

For ground mobile nodes, we use a two-dimensional model, where the sensor node is stationary and the mobile node moves around it. In the most general case (see Figure 3a), the mobile node will describe a random trajectory. However, our study has local character, and only analyzes the mobile behavior inside the sensors’ coverage area. Typically, the absence of infrastructure for supporting power supply causes the coverage area of the sensor nodes is small compared with the total area occupied by the MWSN. Then, we may assume that in the vicinity of the fixed nodes, the trajectory of the mobile nodes is almost linear (see Figure 3b).

Defining a planar Cartesian coordinate system centered on the position of the sensor (see Figure 4a), the trajectory of the mobile node inside the coverage area can be expressed as a pair of functions varying in time, as in (4).

\[
\mathbf{r}(t) = (x(t), y(t))
\]  

On the one hand, the function \( y(t) \) expresses the ideal linear movement at a speed \( v_{\text{linear}}(t) \). The mathematical expression of \( y(t) \) is (5).

\[
y(t) = \int v_{\text{linear}}(t) \, dt + K
\]  

FIGURE 3
Mobile nodes’ trajectory in a two-dimensional scenario. (a) Seeing the whole random trajectory (b) Considering only the almost linear trajectory inside each coverage area
In the typical case, $v_{\text{linear}}(t)$ will be almost constant, and it would only be affected by small variations due to lack of speed stability of the mobile node. Then, $v_{\text{linear}}(t)$ can be understood as the addition of a constant speed $v$ and a small vibrational movement with random parameters varying in time, as can be seen in (6):

$$v_{\text{linear}}(t) = v + B(t)\cos(\beta(t)t)$$  \hspace{1cm} (6)

Where $B(t)$ and $\beta(t)$ are stochastic processes.

$B(t)$ and $\beta(t)$ values depends on the speed stability of the mobile node, but in general they will vary smoothly and slowly (compared to the constant speed $v$). Therefore, in first approximation, we can consider that, within a sensor node’s coverage area, they maintain a fixed value.

$$B(t) = \omega_1 \cdot \epsilon_1$$ \hspace{1cm} (7)

$$\beta(t) = \omega_1$$ \hspace{1cm} (8)

The specific values of $\epsilon_1$ and $\omega_1$ will be experimentally obtained later for each mobile node.

Then, the final expression of $y(t)$ can be seen in (10):

$$v_{\text{linear}}(t) = v + \omega_1 \cdot \epsilon_1 \cos(\omega_1 t)$$ \hspace{1cm} (9)

$$y(t) = \int v_{\text{linear}}(t)dt + K = vt + \epsilon_1 \sin(\omega_1 t) + K$$ \hspace{1cm} (10)

Where $K$ is an integration constant.
On the other hand, the function \( x(t) \) contains the small spurious movements along the second axis (the horizontal axis in our deduction). Thus, the function \( x(t) \) can be mathematically expresses as fixed distance modified by a small vibrational movement with random parameters varying in time (11).

\[
x(t) = d + A(t)\sin(\alpha(t)t)
\]

(11)

Where \( A(t) \) and \( \alpha(t) \) are stochastic processes.

\( A(t) \) and \( \alpha(t) \) values depend on the stability of the mobile node and on how rugged the ground is where the MWSN is deployed. However, as a general consideration, both parameters vary smoothly and, it is foreseeable, slowly (compared to the linear velocity of the mobile node, \( v_{\text{linear}}(t) \)). Therefore, in first approximation, we can consider that, within a sensor node’s coverage area, they maintain a fixed value.

\[
A(t) = \varepsilon_2
\]

(12)

\[
\alpha(t) = \omega_2
\]

(13)

The specific values of \( \varepsilon_2 \) and \( \omega_2 \) will be experimentally obtained later for each mobile node.

Now, to calculate the \( T_{\text{in-cover}} \) time it is necessary to consider the mathematical expression of the coverage area. Considering it as a perfect circle we have that all the points belonging to the coverage area must verify the equation (14):

\[
x^2 + y^2 \leq R^2
\]

(14)

Seeing Figure 4a, we can see there exists two points which belong both to the node’s trajectory and to the circumference that limits the coverage area. Considering the time starts running when the mobile node gets in the coverage area, the mobile node will reach these temporal points at (15) and (16):

\[
T_1 = 0
\]

(15)

\[
T_2 = T_{\text{in-cover}}^{2-\dim}
\]

(16)

Substituting in the expressions for \( T_1 = 0 \), we can calculate the value of \( K \) as in (17).

\[
x^2(0) + y^2(0) = d^2 + K^2 = R^2
\]

(17)

\[
K = -\sqrt{R^2 - d^2}
\]

(18)
Therefore, the trajectory of the mobile nodes, and its total speed, will be defined as in (19) and (20).

\[
\overrightarrow{r}(t) = \left( d + \varepsilon_2 \sin(\omega_2 t), \nu t + \varepsilon_1 \sin(\omega_1 t) - \sqrt{R^2 - d^2} \right) \quad (19)
\]

\[
v(t) = \left( \varepsilon_2 \omega_2 \cos(\omega_2 t), \nu + \omega_1 \varepsilon_1 \cos(\omega_1 t) \right) \quad (20)
\]

Finally, substituting in the expressions for \( T_2 = T_{in-cover}^{2-dim} \), we can calculate the value of \( T_{in-cover}^{2-dim} \) following (21) and (22).

\[
x^2 \left( T_{in-cover}^{2-dim} \right) + y^2 \left( T_{in-cover}^{2-dim} \right) = R^2 \quad (21)
\]

\[
\left( d + \varepsilon_2 \sin(\omega_2 T_{in-cover}^{2-dim}) \right)^2 + \left( \nu T_{in-cover}^{2-dim} + \varepsilon_1 \sin(\omega_1 T_{in-cover}^{2-dim}) - \sqrt{R^2 - d^2} \right)^2 - R^2 = 0 \quad (22)
\]

Solving this transcendent equation in the general case is practically impossible, and numerical methods must be used. However, a simplified expression can be obtained if we consider the ideal case (Figure 4b). In that situation, all the spurious effects are cancelled, and the \( T_{in-cover}^{2-dim} \) can be directly obtained following eq. (23).

\[
\nu^2 T_{in-cover, ideal}^{2-dim} - 2\nu \sqrt{R^2 - d^2} = 0 \quad (23)
\]

\[
T_{in-cover, ideal}^{2-dim} = \frac{2\sqrt{R^2 - d^2}}{\nu} \quad (24)
\]

For a three-dimensional model, the cinematic deduction is pretty similar to the previous one, for two-dimensional scenarios. The base geometrical scheme can be seen in Figure 5.

Imaging a three-dimensional Cartesian coordinate system centered on the position of the sensor, the trajectory of the mobile node inside the coverage area can be expressed as a shortlist of functions varying in time, as in (25).

\[
\overrightarrow{r}(t) = \left( x(t), y(t), z(t) \right) \quad (25)
\]

The expression of \( x(t) \) and \( y(t) \) is the same than the presented for the two-dimensional case. Then:

\[
y(t) = \int v_{linear}(t) dt + cte = \nu t + \varepsilon_1 \sin(\omega_1 t) + K \quad (26)
\]
\[ x(t) = d + \varepsilon_2 \sin(\omega_2 t) \]  

(27)

Where \( \varepsilon_1 \) and \( \omega_1 \) represent the speed stability of the mobile node; and \( \varepsilon_2 \) and \( \omega_2 \) reflect the horizontal stability of the aerial vehicle.

Now, to calculate the expression of \( z(t) \), we are following a similar process to which we followed to obtain the expression of \( x(t) \).

The function \( z(t) \) contains the small spurious movements along the third axis (the vertical axis in our deduction). Thus, the function \( z(t) \) can be mathematically expresses as fixed altitude modified by a small vibrational movement with random parameters varying in time (28).

\[ z(t) = h + C(t) \sin(\gamma(t) t) \]  

(28)

Where \( C(t) \) and \( \gamma(t) \) are stochastic processes.

\( C(t) \) and \( \gamma(t) \) values depends on the vertical stability of the aerial node. However, as in the other cases, both parameters vary smoothly and, it is foreseeable, slowly. Therefore, in first approximation, we can consider that, within a sensor node’s coverage area, they maintain a fixed value.

\[ C(t) = \varepsilon_3 \]  

(29)

\[ \gamma(t) = \omega_3 \]  

(30)

The specific values of \( \varepsilon_3 \) and \( \omega_3 \) will be experimentally obtained later for each mobile node.

Now, to calculate the \( T_{in-cover} \) time it is necessary to consider the mathematical expression of the coverage area. Considering it as a perfect sphere we have that all point belonging to the coverage area must verify the equation (31).
\[ x^2 + y^2 + z^2 \leq R^2 \]  

(31)

One more time, there are two temporal points (eq. (32) y (33)) in which the trajectory crosses the limit of the coverage area.

\[ T_1 = 0 \]  

(32)

\[ T_2 = T_{in-cover}^{3-dim} \]  

(33)

Substituting in the expressions for \( T_1 = 0 \), we can calculate the value of \( K \). See eq. (34)

\[ x^2(0) + y^2(0) + z^2(0) = d^2 + K^2 + h^2 = R^2 \]  

(34)

\[ K = -\sqrt{R^2 - d^2 - h^2} \]  

(35)

Therefore, the trajectory of the mobile nodes, and its total speed, will be defined as in equations (36) and (37): 

\[
\vec{r}(t) = \left( d + \varepsilon_2 \sin(\omega_2 t), vt + \varepsilon_1 \sin(\omega_1 t) - \sqrt{R^2 - d^2 - h^2}, h + \varepsilon_3 \sin(\omega_3 t) \right)
\]  

(36)

\[
\vec{v}(t) = \left( \varepsilon_2 \omega_2 \cos(\omega_2 t), v + \omega_1 \varepsilon_1 \cos(\omega_1 t), \varepsilon_3 \omega_3 \cos(\omega_3 t) \right)
\]  

(37)

Finally, substituting in the expressions for \( T_2 = T_{in-cover}^{3-dim} \), we can calculate the value of \( T_{in-cover} \). See (38).

\[
x^2 \left( T_{in-cover}^{3-dim} \right)^2 + y^2 \left( T_{in-cover}^{3-dim} \right)^2 + z^2 \left( T_{in-cover}^{3-dim} \right)^2 = R^2 
\]  

(38)

\[
\left( d + \varepsilon_2 \sin(\omega_2 T_{in-cover}^{3-dim}) \right)^2 + \left( v T_{in-cover}^{3-dim} + \varepsilon_1 \sin(\omega_1 T_{in-cover}^{3-dim}) - \sqrt{R^2 - d^2 - h^2} \right)^2 \\
+ \left( h + \varepsilon_3 \sin(\omega_3 T_{in-cover}^{3-dim}) \right)^2 - R^2 = 0
\]  

(39)

As in the two-dimensional case, if we suppose the trajectory is ideal (a perfect straight line), the transcendent equation can be strongly simplified.

\[
v^2 T_{in-cover, ideal}^{3-dim} - 2v \sqrt{R^2 - d^2 - h^2} = 0
\]  

(40)
And then, the solution can be expressed as in (41).

\[
T_{\text{in-cover, ideal}}^{3-\text{dim}} = \frac{2\sqrt{R^2 - d^2 - h^2}}{v}
\]  

Finally, in three-dimensional scenarios, it is also typical to eliminate the parameter representing the horizontal distance between the flying node and the fixed node \(d\), due to difficulty of measuring this value. To do that, one value for this parameter (expressed as function of the altitude and the maximum wireless technology rage) is applied. For example, a good value to be employed is the medium value.

The maximum horizontal distance for an aerial vehicle which flies at \(h\) meters, and must keep inside a sphere of radius \(R\) meters can be calculated following (42).

\[
x^2 + y^2 + z^2 = d_{\text{max}}^2 + 0 + h^2 = R^2
\]  

\[
d_{\text{max}} = \sqrt{R^2 - h^2}
\]  

And the minimum is elemental to calculate.

\[
d_{\text{min}} = 0
\]  

Therefore, the medium value is shown in Eq. (45):

\[
d_{\text{medium}} = \frac{\sqrt{R^2 - h^2}}{2}
\]  

And, finally, (46) and (47) show the practical expression obtained.

\[
\left(\frac{\sqrt{R^2 - h^2}}{2} + \varepsilon_2 \sin(\omega_2 T_{\text{in-cover}}^{3-\text{dim}})\right)^2 \\
+ \left(\sqrt{\frac{3}{4} \left(R^2 - h^2\right)} \right)^2 \sin(\omega_1 T_{\text{in-cover}}^{3-\text{dim}}) + h + \varepsilon_3 \sin(\omega_3 T_{\text{in-cover}}^{3-\text{dim}}) - R^2 = 0
\]  

In the general case, or the ideal one:

\[
T_{\text{in-cover, ideal}}^{3-\text{dim}} = \frac{\sqrt{3 \left(R^2 - h^2\right)}}{v}
\]
Now, a final element must be considered in our cinematic model: the precision with which \( T_{\text{in-cover}}^{2\text{-dim}} \) and \( T_{\text{in-cover}}^{3\text{-dim}} \) are obtained. As \( T_{\text{in-cover}}^{2\text{-dim}} \) and \( T_{\text{in-cover}}^{3\text{-dim}} \) are calculated from secondary variables, it is important to express the precision in time as a function of errors in the secondary variables. This objective is practically impossible in the general case (using eqs. (22) and (46)), and numerical methods should be used to limit the error value. However, in the ideal cases, we would be able to calculate the interval in which the real result is located, from errors made by measuring the others variables (such as speed or altitude). Using error propagation theory, and considering the committed errors are small comparing with the variables’ value, we can obtain eqs. (48) and (49).

\[
\Delta T_{2\text{-dim}} = \frac{2}{v} \left( \frac{R}{\sqrt{R^2 - d^2}} \Delta R + \frac{d}{\sqrt{R^2 - d^2}} \Delta d + \frac{\sqrt{R^2 - d^2}}{v} \Delta v \right) \quad (48)
\]

\[
\Delta T_{3\text{-dim}} = \frac{\sqrt{3}}{v} \left( \frac{R}{\sqrt{R^2 - h^2}} \Delta R + \frac{h}{\sqrt{R^2 - h^2}} \Delta h + \frac{\sqrt{R^2 - h^2}}{v} \Delta v \right) \quad (49)
\]

In the final part of this section, we are going to discuss about the mobile node’s speed and the geometrical parameters of our models, in order to evaluate \( T_{\text{in-cover}}^{2\text{-dim}} \) and \( T_{\text{in-cover}}^{3\text{-dim}} \). In the three-layer model for using MWSN, we can identify different types of mobile nodes, which one can be used in various contexts and address different situations:

1. **Walking people**: any person carrying one communication device using the suitable technology can perform the task of mobile node. Such person can act as passive walker such as tourist in a natural area, or he/she can follow specific routes or perform controlled tasks. Typically, a walking person has a medium speed of 4 Km/h [32]. We will use this type of mobile node in home applications (as garden monitoring). Considering a typical garden, it is reasonable supposing a walking person is never further than 2 meters of a sensor.

2. **Cars, off-road vehicles and in general any type of vehicle**: those vehicles can move through the sensors area loaded with a node for data collecting, and as in the case of walking people, they can follow non determinstic paths or can be controlled for performing specific ways. Controlling quite big semi-urban areas (such as golf courses) is likely the most common application for these mobile nodes. 20 km/h is a typical speed for golf carts, and, taking into account the dimensions of the touristic courses (6400 meters), we can consider 10 meters as the minimum distance between a terrestrial vehicle and a sensor.
3. Aerial vehicle and more specifically unmanaged aerial vehicles (drones) can be used for data collection. In this group, we have to distinguish two types of drones: plane-type drones (which cannot be stopped on air) and helicopter-type drones (which can remain on air with zero speed). Plane-type drones are commonly used to control great natural areas (such as forest, lakes, etc.), while helicopter-type drone are employed in urban environment (as they can remain in air controlling people, cars, etc.). The speed of these mobile nodes depends on the model selected, but the most usual value for cruise speed is 60 Km/h for helicopter-type drones, and 80 Km/h for plane-type drones. Finally, drones’ altitude is function of the skyline altitude [2]. Thus, in networks deployed in low-height and large-scale crops, drones can be operated at 5 meters; for monitoring dense forest or metropolitan areas drones may be fly at 10 meters (more or less); and for collecting data from the financial district (where buildings have more than 10 floors), drones must have an altitude around 100 meters.

4. Zeppelins, such type of mobile nodes can be used stay in the air for long time periods or moving over extensive areas with low energy consumption. The most common application of these mobile node is verifying the state of the mountain parks (like Rocky Mountain Park), where the mobile node must operate at 3000 meters or more. Their speed, however, due to the environmental risks is much lower than the aerial vehicles’ one. Typically, a zeppelin does not move faster than 5 Km/h.

Finally, it is necessary to evaluate the parameters related to the stability ($\varepsilon_1$, $\omega_1$, $\varepsilon_2$, $\omega_2$, $\varepsilon_3$ and $\omega_3$) for each mobile node. To do that, an experimental evaluation has been conducted.

First, an electronic system to determine in a very precise way the position of each mobile node is designed. This module consisted of the following elements (Figure 6 shows an electronic diagram of the module):

- Acentrimetricprecision GPS module. Recently, small form factor GPS receivers have been developed. In particular, Piksilow-cost GPS receiver, from Swift Navigation [33], has a 53x53mm form factor and supports GPS, GLONASS and Galileo signals. By means of this sensor, latitude and longitude of the location of the mobile node can be known with centimetricprecision. This device connects to the outside via serial port (UART).
- Abarometricsensor. The considered GPS module calculates very precisely latitude and longitude of the location of the mobile node. However, if using aerial nodes, we wish to obtain the same precision in
height. For this, a high precision barometric sensor (BMP180 in this case [34]) is included. The device’s output interface is serial (I2C).

- An Arduino Nano microcontroller. It receives data from GPS module by UART and from barometric sensor by I2C. It encapsulates both data in an application message and transmits the message by a second UART at 19200 bauds.

- A Bluetooth module. A HC-06 Bluetooth slave module receives data by UART at 19200 bauds and transmits them through a Bluetooth 3.0 interface.

Second, a Bluetooth receiver was built and connected to a general purpose computer. This receiver is made of a Bluetooth master module (HC-05 module), a microcontroller and a RS-232 adapter to connect the microcontroller with the computer. When the microcontroller receives a message from the mobile node, it sends it by RS-232 to the computer, where a MATLAB application is listening.

Finally, in the MATLAB implementation, the data are processed and stability values calculated. For that the geographic coordinates (latitude and longitude) are converted to Cartesian XY coordinates. Moreover, barometer data are processed to obtain the height of the node as the third Cartesian coordinate Z. Later all sequences are low pass filtered to remove high frequency numerical noise introduced by processing algorithms and sensor accuracy. Then, X and Z sample sequences are transformed by means of the FFT algorithm for expression in the frequency domain. The values of horizontal and vertical stability are given by the amplitude and pulsation of the main harmonic. The third sequence (Y coordinate according to Figures 4 and 5) is numerically derived for obtaining the linear velocity on said component.
Finally, the resulting sequence is transformed by means of the FFT algorithm for expression in the frequency domain. The speed stability is given by the amplitude and pulsation of the main harmonic.

For each mobile node type the experiment was performed 16 times. This number of times has been selected taking into account the experiment’s restrictions. The final stability values are the median of all the particular obtained results. Finally, due to the complexity associated with zeppelins, this type of vehicle is not considered in the experiment.

The next table (Table 1) summarizes speed, distance to fixed nodes and the stability results for different mobile nodes.

5 WIRELESS TECHNOLOGIES STUDY

Once we have described the characteristics of the main mobile node types, and the work scenario, in this section we are studying the wireless technologies commonly used on WSN, as well as some variants of them. Then, we will obtain the value of $L_{RAT}$ and model the data rate $C(t)$.
For the scenarios described, we need the wireless technology that consumes as less power as possible. Due to the fact that the nodes lack of a permanent power connection, so we need technologies with a reduce coverage range to be capable of getting this purpose. The study will be focused in personal area networks technologies, to which the WiFi technology will be added with the objective of establishing a comparison with the more extended wireless standard. The complete list of the considered technologies is: Bluetooth 3.0, Bluetooth 4.0 (particularly the Low-Energy core or BLE), 802.15.4, a variant of the 802.15.4 standard called Fast 802.15.4 proposed in [22], WiFi, considering two different protocol implementations.

In our calculations we present the particular process of neighbor discovery employed in each technology (process which determines the connection time needed).

5.1. Neighbor discovery in Bluetooth 3.0 technology

The neighbor discovery process of Bluetooth 3.0 is the most complex and least efficient. In this version, the neighbor discovery procedure described in the standard is called “inquirer process”, and supposes that they exist two active devices for all time: on the one hand we find the master (which is looking for new neighbors), and on the other we find the slave (which tries to be discovered to be integrated in one network).

Periodically, the master broadcasts some special packets (named as ID packets), whose mission is to discover new devices. The packets are transmitted using 32 different frequencies, following a procedure known as Frequency Hopping (FH). The FH procedure divides the 32 available channels in two groups of 16, called trains and labeled as A and B.

In each cycle, four sequences are broadcasted, each one made up for 256 frequency trains. In the first and third frequencies A train is transmitted, and in the second and fourth B train is transmitted. The time necessary to broadcast the ID packets in each channel is one slot of 625 ms [35], so 16 slots are necessary to transmit one train of frequencies (see Figure 7).

We suppose now a scenario where there are no interferences, thus, it’s reasonable to think the discovery of one new device will be completed in one cycle. Then, if the discovery goes on in the n hop, the total time spent will be (50):

\[ L_{open\ BT} = n \cdot t_{slot} \] (50)

FIGURE 7
Graphic explanation of Bluetooth 3.0 neighbor discovery process
5.2. Neighbor discovery in Bluetooth Low-Energy technology

As we said in Section 5.1, Bluetooth 3.0 neighbor discovery is really expensive in terms of power. Thus, when Bluetooth Special Interest Group decided to create a Low-Energy core for Bluetooth 4.0 (called BLE), they had to propose a totally renovated procedure.

In BLE, again, it exists two types of devices: scanner or initiator (which tries to discover new devices), and the advertiser (which tries to be discovered). The initiator periodically scans three special channels allocated for the advertisement broadcast (labeled as 37, 38 and 39) in a sequential way, looking for advertising packets of others. One the other hand, the advertiser transmits advertising packets in the same channels (using the FH procedure), and wait for an answer sent by the scanner. One the advertiser receives the scanner respond, it is considered the advertiser has been discovered; even the advertiser also has to respond to confirm the connection [36].

BLE standard defines two different procedures for neighbor discovery, depending on the knowledge stored in the scanner: if the scanner knows the advertiser it uses Direct advertising, if not, it uses an Indirect advertising. In our case, as the mobile node is known, we are evaluating the Indirect advertising type.

The time a device that plays the role of advertiser needs to complete an advertising event (the name used in the standard for a search cycle), is computed by (51):

$$T_{event} = T_{adver} + delay$$  \hspace{1cm} (51)

The $T_{adver}$ is an integer multiple of 0.625 ms, ranging from 20 ms to 10.24 seconds, and $delay$ is a pseudo-random value ranging from 0 ms to 10ms (generated by the Link Layer for each advertising event).

Inside each advertising event, for each frequency an advertising packet is sent; later, the advertiser will be listening on the same channel for a while to

FIGURE 8
Graphic explanation of BLE neighbor discovery process
check if there is response coming from any initiator. According to the specification, the total time spent by channel shall be less than or equal to 10 ms (see figure 8).

In conclusion, and if we suppose that the time spent by channel is the same for the three frequencies and that there aren’t any interferences, the discovery only needs one advertising event. The time necessary to complete a discovery of a device that transmits in the $n$ advertising channel is (52):

$$L_{open\ BLE} = n \cdot (T_{tx} + T_{delay\ intra-channel} + T_{RX}) + (n-1) \cdot T_{delay\ intra-channel}$$  (52)

### 5.3. Neighbor discovery in 802.15.4 technology

Although BLE strongly reduces the power consumption (compared to Bluetooth 3.0), all Bluetooth technologies are designed for transmitting a great data flow, which not agrees at all with the assumptions exposed in Section 2. Therefore, the third technology we are considering is 802.15.4 standard, destined to support low-energy and low-rate communications.

In 802.15.4 standard [37], networks have a hierarchical structure (unlike the Bluetooth technology). Thus, in all networks must be a coordinator responsible for managing the new associations.

The new device has to send an association solicitation to the corresponding coordinator for being incorporated into a network; for which, it is necessary that the new device knows both coordinator’s MAC and the frequency where the network is working (among the 16 available channels).

If we suppose the new device has no information about the network, it must scan iteratively the 16 available channels looking for the packets the coordinator broadcasted (where it is included the whole information necessary for building the association message). Once this message is sent, the device is considered to be associated to the network.

In conclusion and supposing the network operates over the $n$ channel, the time spent in adding a new device to the network is (53):

$$L_{open\ 802.15.4} = n \cdot t_{scan} + t_{asso\ mess}$$  (53)

$$= n \cdot aBaseSuperFrameDuration \cdot (2^{BO} + 1) \cdot t_{sym} + t_{asso\ mess}$$

Where $aBaseSuperFrameDuration$ and $BO$ are two integer parameters.

### 5.4. Neighbor discovery in Fast 802.15.4 technology

As we can see in (53), connection time in 802.15.4 technology has a linear relation with the channel number where the network coordinator is transmitting. This can be very problematic if we consider a big set of channels, so in the last standard version (802.15.4e) the committee has tried to optimize the scan time, although the truth is the impact of these measures is limited.
Therefore, it is so important a standard’s modification proposed recently [22], called Fast 802.15.4, which reaches to take a great advantage of the switching frequency capability of 802.15.4 networks.

They propose, then, to allocate one channel (named as Beacon Channel) to transmit over it the associating messages, and all the information needed to build this solicitation. The others channel will be used for data transmission. In that way, the time necessary to add a new device is fixed, and it is equal to (54):

\[
L_{open\ fast 802.15.4} = t_{scan} + t_{asso\ mess} \\
= a Base SuperFrame Duration \cdot (2^{BO} + 1) \cdot t_{sym} + t_{asso\ mess}
\]  

(54)

5.5. Neighbor discovery in WiFi technology

Finally, we are going to analyze neighbor discovery in WiFi technology. With these data, our model will also include LAN technologies, what is basic in order to obtain remarkable results in Section 6.

In 802.11 standard, neighbor discovery (known as Access Point discovery) is performed through the scanning process, in which a mobile node (called mobile station -MS- in this standard) sends a Probe Request management frame and waits for a Probe Response on each channel.

This waiting time is managed by two timers in the scanning process, namely MinChannelTime (MinCT) and MaxChannel-Time (MaxCT).

If no Probe Response is received before MinCT expires, the MS switches to the next channel and sends a new Probe Request. This may happen if no AP is operating on the channel, if the Probe Response is sent after the timer expiration, or if there is a transmission error. Otherwise, if at least one Probe Response was received, the MS waits for a longer timer (namely MaxCT). In this additional time, the MS is listening for more responses from other Access Points operating in the same channel (see Figure 9).
It is very important to say that although the 802.11 standard proposes 16 available channels, in Europe, only 13 channels are available [38]; so, in the following section we only consider those 13 frequencies.

Taking into account this process, the time spent in locating a new device which uses the $n$ channel is (55):

$$L_{\text{open} \text{802.11}} = \text{MaxCT} + (n - 1) \cdot \text{MinCT} \quad (55)$$

### 5.6. Probabilistic model of $L_{\text{open RAT}}$ delay

As we saw in Equations (50), (52), (53), (54) and (55) the $L_{\text{open RAT}}$ delay depends on $n$ (the number of the channel where the neighbor device is found). Knowing a priori this value is impossible, since it depends on the specific scenario to be considered: topography of the land, interfering systems, etc. Then, it is a random variable. Therefore, to characterize completely $L_{\text{open RAT}}$ it is not only necessary to show the mathematical expression as a function of $n$, but also the probability of each specific value (i.e. the associated probability density function).

Thus, two different probability density functions must be considered. First, one of the functions represents the probability of finding one device in a specific given channel (called $f_n[n]$ onwards). Second, the function that relates the value of $L_{\text{open RAT}}$ calculated for each value of $n$ with the probability of needing this time for finding a neighbor (called $f_L[L]$ onwards). Since the channel number is a discrete variable (ranging from $n = 1$ to $n = \text{number of channels}$), both probability density functions are also discrete.

The functions $f_n[n]$ and $f_L[L]$ are not independent, but only $f_n[n]$ is independent, being possible to calculate $f_L[L]$ from it. In protocols in which both devices which want to communicate are moving along the frequency spectrum (such as Bluetooth 3.0 and Bluetooth Low-Energy), each channel tracking is independent of the previous, so the probability of stopping the neighbor discovery process in a specific channel is equal to the probability of the device to be on that channel. Then, the relation between $f_n[n_i]$ and $f_L[L_i]$ where $L_i$ is the $L_{\text{open RAT}}$ delay suffered if one neighbor is found in the $n_i$ channel is (56).

$$f_L[L_i] = f_n[n_i] \quad (56)$$

However, in protocols in which one of the elements remains permanently transmitting in the same channel (802.15.4, 802.11, etc.), the $n + 1$ channel is evaluated only if the device is not found in the $n$ channel (search is ordered). Thus, the relation between $f_n[n_i]$ and $f_L[L_i]$ where $L_i$ is the $L_{\text{open RAT}}$ delay suffered if one neighbor is found in the $n_i$ channel is (57).

$$f_L[L_i] = f_n[n_i] \quad \text{if} \quad i = 1 \quad (57)$$
Finally, $f_n[n]$ must be calculated considering the characteristics of the deployment scenario, the sensor programming, etc. For example, if other radio services cover the area of the MWSN, probably it would be necessary to evaluate the interferences following the ITU-R recommendations (P.452 in this case [39]).

5.7. Upstream radio rate $C(t)$ evaluation

Finally, the last parameter that we must evaluate on wireless access technologies is the upstream radio rate $C(t)$. The randomness of this parameter (along with its time dependence), makes $C(t)$ in a stochastic process, therefore lacks deterministic mathematical expression. Thus, integrating $C(t)$ in time becomes a complicated task, but essential to our model.

In order to partially solve this problem, our model only considers the expected value of the $Data_{collected}$ parameter. However, for calculating this value some considerations have to be taken into account. First, let’s consider $C(t)$ as a stationary stochastic process, i.e. the probability distribution is the same for every instant of time. Additionally, as it is common in communications systems, we will assume that $C(t)$ has ergodic character (thus the process statistic parameters are invariant over time).

Now, we calculate the expected value of the Equation (3)

$$\mathbb{E}[Data_{collected}] = \mathbb{E}\left[\int_{0}^{T_{in-cover}(R,v,geometry)-L_{open RAT}} C(t)dt \right]$$

Considering the definition of the Riemann integral we can obtain expression (59)

$$\mathbb{E}\left[\int_{0}^{T_{in-cover}(R,v,geometry)-L_{open RAT}} C(t)dt \right] = \lim_{n \rightarrow \infty} \frac{T_{in-cover}(R,v,geometry)-L_{open RAT}}{n} \sum_{k=1}^{n} C \left( \frac{k(T_{in-cover}(R,v,geometry)-L_{open RAT})}{n} \right)$$

Being the expected operator continuous and linear, it can be interchangeable with the limit operator. Moreover, as $L_{open RAT}$ and $C(t)$ are physical independent processes, they are also independent statistics variables. Then it is possible to obtain expression (60).

$$\mathbb{E}\left[\lim_{n \rightarrow \infty} \frac{T_{in-cover}(R,v,geometry)-L_{open RAT}}{n} \sum_{k=1}^{n} C \left( \frac{k(T_{in-cover}(R,v,geometry)-L_{open RAT})}{n} \right) \right]$$
Furthermore, as \( C(t) \) is a stationary stochastic process, the expected value is constant in time. Thus, the expression (61) becomes true.

\[
\mathbb{E} \left[ \sum_{k=1}^{n} \frac{k \left( T_{in\text{-}cover} (R, v, geometry) - L_{open\text{ RAT}} \right)}{n} \right] = \mathbb{E} \left[ \frac{k \left( T_{in\text{-}cover} (R, v, geometry) - L_{open\text{ RAT}} \right)}{n} \right]
\]

Joining expressions (58), (59) and (61) we can obtain expression (62):

\[
\mathbb{E} \left[ \int_{0}^{T_{in\text{-}cover} (R, v, geometry) - L_{open\text{ RAT}}} C(t) \, dt \right] =
\]

\[
= \lim_{n \to \infty} \frac{T_{in\text{-}cover} (R, v, geometry) - \mathbb{E} \left[ L_{open\text{ RAT}} \right]}{n} \sum_{k=1}^{n} \mathbb{E} \left[ \frac{k \left( T_{in\text{-}cover} (R, v, geometry) - \mathbb{E} \left[ L_{open\text{ RAT}} \right] \right)}{n} \right]
\]

And considering one more time the definition of the Riemann integral (see (63))

\[
\mathbb{E} \left[ \int_{0}^{T_{in\text{-}cover} (R, v, geometry) - L_{open\text{ RAT}}} C(t) \, dt \right] =
\]

\[
= \int_{0}^{T_{in\text{-}cover} (R, v, geometry) - \mathbb{E} \left[ L_{open\text{ RAT}} \right]} \mathbb{E} \left[ C(t) \right] \, dt
\]

Integrating expression (63) and considering Equation (58), it is possible to calculate the expression (64).

\[
\mathbb{E} \left[ Data_{collected} \right] = \mathbb{E} \left[ C(t) \right] \left( T_{in\text{-}cover} (R, v, geometry) - \mathbb{E} \left[ L_{open\text{ RAT}} \right] \right)
\]

The value of \( \mathbb{E} \left[ L_{open\text{ RAT}} \right] \) can be calculated from the discussions presented in the previous subsections, but \( \mathbb{E} [C(t)] \) remains unclarified.
In general, to estimate $\mathbb{E}[C(t)]$ several realizations of $C(t)$ would be necessary, but as the process is supposed to be ergodic, the expected value can be replaced (since it coincides in value) by a time average (so that even a single embodiment can characterize the process).

To obtain the expected value of the upstream radio rate in all radio technologies described above, another experiment was conducted. Twelve fixed nodes, implementing two to two the same radio technology, were deployed in an agricultural environment (Figure 10). All nodes included only a program being able to calculate the instantaneous rate by performing a continuous data sending. So, it can be assumed that the calculated value perfectly matches the rate at radio level. Each couple of nodes was exchanged data for two hours, measuring the upstream radio rate each 500ms (in total they were taken around 14500 measurements).
At the end of the experiment, data was gathered by means a debugger board (Figure 11) and processed using numerical calculation software. Table 2 shows the expected value, the most probable value and the standard deviation obtained for each radio access technology.

### 6. MODEL UTILIZATION. ANALYTICAL EVALUATION

In this section, we are going to join the results of the previous sections to create the final model. We are also describing the process which must be followed in order to obtain an estimation of the total amount of data collected. Later, we are proposing typical values for the parameters in the model to obtain some numerical results. With the study of these results new problems and limitations for the MWSNs will arise.

#### 6.1. Final model. Model utilization

Considering (64), the final model can be expressed as (65).

\[
\mathbb{E}[Data_{collected}] = \mathbb{E}[C(t)](\tau_{in-cover}^{i-dim} - \mathbb{E}[L_{open\ RAT}])
\]

\[i \in \{1,2\}\]

\[RAT \in \{BT, BLE, 802.15.4, fast802.15.4, 802.11\}\]

To obtain an estimation of the data collected by means of the expression (65), the following steps must be followed.

1. Having chosen the deployment scenario, the radio access technology for the communications and the type of mobile node which will be

---

<table>
<thead>
<tr>
<th>Radio technology</th>
<th>Expected value (Mbps)</th>
<th>Standard Deviation (Mbps)</th>
<th>Most probable value (Mbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bluetooth 3.0</td>
<td>0.844</td>
<td>0.03</td>
<td>0.852</td>
</tr>
<tr>
<td>Bluetooth Low-Energy</td>
<td>0.874</td>
<td>0.04</td>
<td>0.861</td>
</tr>
<tr>
<td>802.15.4</td>
<td>0.205</td>
<td>0.01</td>
<td>0.207</td>
</tr>
<tr>
<td>Fast 802.15.4</td>
<td>0.204</td>
<td>0.01</td>
<td>0.204</td>
</tr>
<tr>
<td>WiFi (fixed temporizers)</td>
<td>8.774</td>
<td>0.97</td>
<td>9.057</td>
</tr>
<tr>
<td>WiFi (Smart temporizers)</td>
<td>8.894</td>
<td>1.11</td>
<td>8.855</td>
</tr>
</tbody>
</table>

**TABLE 2**

Statistics description of the upstream radio rate for each RAT
used as data mule, said mobile node must be characterized in relation to its horizontal stability, speed stability (and vertical stability, if an aerial node). Values such as the minimum distance at which the mule will be from the fixed sensors must be also obtained. In Section 4 we described this process (and presented some results) for some typical mule types.

2. With the parameters calculated in the previous step, and expressions (22) and (46) (depending on if we consider a two-dimensional scenario or a three-dimensional one), the value of $T_{in-cover}^{i-dim}$ may be numerically calculated.

3. Known the geometrical parameters of the deployment, a radio evaluation of the scenario must be performed. With it, and knowing the RAT employed, the probability function to find a node in a channel ($f_n[n]$) may be estimated.

4. Considering the calculated $f_n[n]$ the probability density function $f_L[L]$ may be obtained. In Section 5 (expression (57)) the process is described.

5. A deep analysis of the hardware and programming employed in the fixed nodes must be conducted in order to estimate the parameters which appear in the expression of the neighbor discovery delay for the selected RAT. Taking them and $f_L[L]$ into account, the value of $E[L_{open \, RAT}]$ can be calculated.

6. In order to know the value of $E[C(t)]$, a temporal evaluation of the evolution of upstream radio rate should be performed. The time average of this evaluation will give us the sought value. In Section 5 we described this process (and presented some results) for the considered RAT.

7. Applying expression (65) the amount of collected data may be calculated.

6.2. Analytical evaluation

In this section, we are obtaining the amount of collected data using the proposed model, for all combination of the selected RAT and considered mobile nodes.

In Section 4, the stability of the considered mobile nodes was evaluated. In Section 5, the value of $E[C(t)]$ for all selected RAT was also calculated. Then, to be able to evaluate the amount of data collected it is necessary, firstly, to fix the variables’ values in the neighbor discovery equations (expression (50), (52), (53), (54) and (55)).

Bluetooth 3.0 technology fixes all the values of neighbor discovery process’ parameters, so it is enough with the information given in Section 5.

On the contrary, in BLE technology the standard describes the neighbor discovery as a function of several variables (see expression (52)). For them, it gives only a maximum value. So, in order to fix these variables’ values, we are going to employ as reference the BLE implementation used in the CC2540 device from Texas Instruments (the most common system-on-chip for this technology). The specific values are (see table 3):
Considering, 802.15.4 and Fast 802.15.4 solutions, Table 4 contains all necessary data to calculate the results [13]. These values have been set comparing various implementations of 802.15.4 protocol, such as IEEE 802.15.4 from Texas Instruments, open-ZB or ZBOSS.

Determining the temporizer values of WiFi technology, the standard does not suggest any solution so the selected policy for setting these values strongly impacts in the results. For this reason, we are studying two different solutions: the first one based on fixed values for the temporizers; and the second one based on a smart algorithm for selecting the values for each channel.

For the first implementation we have selected a value of 15 ms for the temporizers (most common commercial value), and for the second proposal we are going to use the algorithm described in [12].

Later, it is necessary to estimate the probability distribution function \( f_n[n] \). In this case, due to the fact that there are no interferences in the environment, we can suppose all channel have the same probability. Therefore, \( f_n[n] \) is a uniform probability distribution between \( n = 1 \) and \( n = N \) (both included) where \( N \) is the number of the available channels. Mathematically:

\[
 f_n[n] = \frac{1}{N} (u[n] - u[n-N-1]) 
\]

(66)

Being \( u[n] \) is the Heaviside-step function of discrete type.
Now, we may obtain the value of \( \mathbb{E}[L_{\text{open RAT}}] \). In Table 5, we calculate the main values that can be obtained from the random variable \( L_{\text{open RAT}} \). There, we also show the expected value and nominal value of the upstream radio rate (\( \mathbb{E}[C(t)] \)) and the maximum range (\( R \)) for every RAT. As we saw, to obtain these values we supposed a low-density natural environment and a standard transmission power (considering the nodes implement a system-on-chip, as CC25XX series from Texas Instruments).

As we said in Section 2, it is important to highlight that, in some technologies, fixed nodes, usually spend most of the time in a sleeping mode (in order to save as much power as possible). However, in this work we are going to suppose all nodes are active all the time (so values shown in Table 5 are final).

Considering numerical data from Tables 1 and 5, the final model expressed in (65), and solving numerically the expressions (22) and (46), it is possible to calculate the amount of data collected from sensor for each technology and mobile node type. See Table 6 and Figure 12.

Before extracting any conclusion, it is important to notice that the data quantity expressed in Figure 12 is calculated at radio level. The real data quantity from application level is lower. For example, in 802.15.4 standard only 2 kbps of the 250 kbps available at radio level can be used at application level.

Any case, a great importance result (which we should obtain) is the impact of considering real effects (such as the mobile node’s stability) on the amount of data calculated. Figure 13 shows, as percentage, the difference between the amounts of data shown in Figure 12, and those obtained without considering stability parameters and taking, instead of the expected value of the upstream radio rate, its nominal value. As can be seen, in general, the percentage is positive, which means that the amount calculated considering “real effects” is lower than which could be obtained without considering these effects.

In Figure 13, two overlapping effects can be shown: one due to consider a more realistic mobility model and another due to consider a “real” value for the upstream radio rate. A study of both effects separately may show that the variation due to consider the stability parameters is the order of 5%, being the remainder of the variation due to consider an upstream radio rate different to the nominal.

The variation in the amount of data due to consider a value different from the nominal upstream radio rate is always positive (i.e., the amount is lower because of considering a non-nominal value). However, the effect due to consider a more complex mobility model can take both signs. In particular, in the case of using a plane-type drone, the time in coverage is much higher if stability parameters are considered. Thus, in P5 and P10 scenarios, the variation is below the average; even it can be seen that the total data transmitted considering the “real” effect is greater than that if we do not consider it (in the case of using RAT 802.15.4).
<table>
<thead>
<tr>
<th>Wireless Technology</th>
<th>Minimum connection time</th>
<th>Maximum connection time</th>
<th>Medium connection time</th>
<th>Most probable connection time</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Expected upstream Radio rate</th>
<th>Nominal upstream Radio rate</th>
<th>Maximum range ((R))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bluetooth 3.0</td>
<td>625 ms</td>
<td>10.24 s</td>
<td>5.12 s</td>
<td>---</td>
<td>5.12 s</td>
<td>3.6 s</td>
<td>844 kbps</td>
<td>1 Mbps</td>
<td>10 m</td>
</tr>
<tr>
<td>BLE</td>
<td>600 ms</td>
<td>2.1 ms</td>
<td>1.35 ms</td>
<td>---</td>
<td>1.35 ms</td>
<td>612 ms</td>
<td>874 kbps</td>
<td>1 Mbps</td>
<td>50 m</td>
</tr>
<tr>
<td>802.15.4</td>
<td>0.5 s</td>
<td>2.7 s</td>
<td>952.21 ms</td>
<td>0.5 s</td>
<td>1.67 s</td>
<td>1.027 s</td>
<td>205 kbps</td>
<td>250 kbps</td>
<td>70 m</td>
</tr>
<tr>
<td>Fast 802.15.4</td>
<td>628.24 ms</td>
<td>628.4 ms</td>
<td>628.24 ms</td>
<td>628.24 ms</td>
<td>628.24 ms</td>
<td>0</td>
<td>204 kbps</td>
<td>250 kbps</td>
<td>70 m</td>
</tr>
<tr>
<td>WiFi (fixed temporizers)</td>
<td>15 ms</td>
<td>210 ms</td>
<td>62.94 ms</td>
<td>15 ms</td>
<td>195 ms</td>
<td>70.14 ms</td>
<td>8.9 Mbps</td>
<td>11 Mbps</td>
<td>100 m</td>
</tr>
<tr>
<td>WiFi (smart temporizers)</td>
<td>15 ms</td>
<td>121 ms</td>
<td>6.9 ms</td>
<td>15 ms</td>
<td>93 ms</td>
<td>5.49 ms</td>
<td>8.8 Mbps</td>
<td>11 Mbps</td>
<td>100 m</td>
</tr>
</tbody>
</table>
7 CONCLUSIONS AND FUTURE WORK

Our objective in this research work is to present a new model for data collection in MWSN. A numerical analysis of this model allowed us to identify new problems in the operation of WSN, by studying the impact of the connection time and scenario’s geometry in the process of data collection. The realized analysis includes both two-dimensional and three-dimensional scenarios, so terrestrial and flying mobile nodes can be considered.

Furthermore, the realistic parameter’s values we have used allow us to extract some interesting conclusions:

<table>
<thead>
<tr>
<th>Wireless Technology</th>
<th>WP</th>
<th>TV</th>
<th>H5</th>
<th>H10</th>
<th>H100</th>
<th>P5</th>
<th>P10</th>
<th>P100</th>
<th>DIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bluetooth 3.0</td>
<td>12.58</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BLE</td>
<td>90.90</td>
<td>17.47</td>
<td>5.15</td>
<td>5.07</td>
<td>0</td>
<td>3.8</td>
<td>3.82</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>802.15.4</td>
<td>126.45</td>
<td>23.75</td>
<td>6.28</td>
<td>6.25</td>
<td>0</td>
<td>4.5</td>
<td>4.45</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fast 802.15.4</td>
<td>126.5</td>
<td>24.08</td>
<td>6.61</td>
<td>6.55</td>
<td>0</td>
<td>4.83</td>
<td>4.78</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>WiFi (fixed temporizers)</td>
<td>181.62</td>
<td>35.43</td>
<td>10.28</td>
<td>10.24</td>
<td>0</td>
<td>7.73</td>
<td>7.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>WiFi (smart temporizers)</td>
<td>181.7</td>
<td>31.49</td>
<td>10.34</td>
<td>10.30</td>
<td>0</td>
<td>7.79</td>
<td>7.76</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**TABLE 6**
Medium time available for data collection ($T_{in-cover}^{dim} - \mathbb{E}[L_{open\ RAT}]$)

**FIGURE 12**
Transmittable data quantity during the available time (Y-limit 100 Mb). For WP and WiFi technologies, we found a maximum of 1.6Gb and for TV and WiFi a maximum of 315 Mb.
In general, the connection time strongly impacts the total available time along which the fixed and mobile nodes are in their coverage area. However, some technologies (like WiFi) are immune to the degradation due to the connection time, although these standards do not belong to the PAN set, so they are not useful in scenarios with scarce power supply.

Moreover, the values we showed in Figure 12 are calculated in a controlled scenario. In a real situation, higher layers in the protocol stack will also increase the time needed for establishing an ephemeral connection, the interferences in the radio channel can make the devices need more than one search cycle for discovering a new neighbor, the weather changes may affect the communication quality, etc. Even in this ideal scenario, the radio rate destined to data might be not enough. Considering the simplest applications, it is possible to achieve the requirements easily, but applications (and overall services) with more value added need a rate at top level that is not clear we can get.

As we can see in Figure 12, the amount of data collected is remarkable only for WiFi technology and for the slowest mobile nodes (namely, walking people and terrestrial vehicles). Nonetheless, Bluetooth 3.0 is so inefficient that only walking people can be considered as a valid mobile node. The use of drones only can be taken into account in scenarios where 802.11 technology is available, although if the application does not require a great data rate, BLE technology is also possible in low and medium altitude scenarios. Finally, it is clear that in high altitude scenarios, any PAN or LAN technology can be used. In those situations mobile technologies (such as LTE or the next 5G solutions) could be the only choice.

Therefore, in general, the main conclusion is that there are two possible scenarios: one where slow mobile nodes are employed; and other where LAN technologies (e.g. WiFi, WiMAX…) are deployed. Then, the role of

![FIGURE 13](image)

FIGURE 13
Difference between the amounts of data shown in Figure 12, and those obtained without considering stability parameters and considering the nominal upstream radio rate.
distance and mobile node speed as key element is clear, so they should be the center of our future research. In that sense, it is really important to optimize the drones’ trajectories as essential measure for improving the results of these systems.

Once we have described the problems associated with the necessity of establishing an ephemeral connection, a lot of new analysis and proposals can be made. On the one hand, new standards’ versions and special neighbor discovery procedures can be studied, focused on reducing the connection time. On the other hand, many extensions of this works may be made: from new analysis considering the time needed for waking up the nodes to studies based on more complicated physical scenarios, for example, one where the mobile nodes consider random paths. Furthermore, it will be very interesting a similar work that considers others wireless communication technologies or standards such as 802.11v, LTE or WiMAX.

ACKNOWLEDGMENTS

Borja Bordel has received funding from the Ministry of Education through the FPU program (grant number FPU15/03977). Additionally, the research leading to these results has received funding from the Ministry of Economy and Competitiveness through SEMOLA project (TEC2015-68284-R) and from the Autonomous Region of Madrid through MOSI-AGIL-CM project (grant P2013/ICE-3019, co-funded by EU Structural Funds FSE and FEDER).

REFERENCES


[38] European Table of Frequency Allocations and Applications http://www.erodocdb.dk/Docs/doc98/official/pdf/ercrep025.pdf