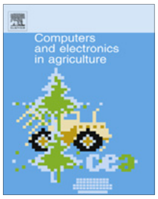
Computers and Electronics in Agriculture 87 (2012) 74–84

Contents lists available at SciVerse ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag



# Performance assessment of a kinetically-powered network for herd localization

Nelson I. Dopico *, Álvaro Gutiérrez, Santiago Zazo

*ETSI Telecomunicación, Universidad Politécnica de Madrid, Avda. Complutense 30, 28040 Madrid, Spain*

A R T I C L E  I N F O

*Article history:*
Received 20 June 2011
Received in revised form 23 April 2012
Accepted 24 April 2012

*Keywords:*
Localization
GPS
Reindeer
Modeling
Herding
Energy efficiency

A B S T R A C T

Developing a herd localization system capable to operate unattended in communication-challenged areas arises from the necessity of improving current systems in terms of cost, autonomy or any other facilities that a certain target group (or overall users) may demand. A network architecture of herd localization is proposed with its corresponding hardware and a methodology to assess performance in different operating conditions. The system is designed taking into account an eventual environmental impact hence most nodes are simple, cheap and kinetically powered from animal movements – neither batteries nor sophisticated processor chips are needed. Other network elements integrating GPS and batteries operate with selectable duty cycles, thus reducing maintenance duties. Equipment has been tested on Scandinavian reindeer in Lapland and its element modeling is integrated into a simulator to analyze such localization network applicability for different use cases. Performance indicators (detection frequency, localization accuracy and delay) are fitted to assess the overall performance; system relative costs are enclosed also for a range of deployments.

## 1. Introduction

Animal localization has been an active research area for years. Animal behavioral studies (Bergman et al., 2000; Zhang et al., 2004), cattle monitoring (Schlecht et al., 2004), improvement of livestock techniques (Bailey, 2000) and other applications perform animal localization (Nadimi et al., 2008; Huircán et al., 2010). Different devices and strategies have been used throughout the years and, in particular, the use of GPS collars has been extended to moose (Rempel and Rodgers, 1997), camels (Grigg et al., 1995), goats (Buerkert and Schlecht, 2009) and other species. Frequently, accuracy is a must since such collars are required for fine localization or tracking. However, user requirements are not always the same and the aforementioned devices may exceed the budget of herders who could be satisfied with equipment offering less features. Thus, researchers and engineers may face a challenge consisting of developing alternative systems lowering costs and/or providing users with additional facilities. Developing such new systems comprises a number of stages besides the mere electronic design of collar-like devices. These stages can be, for example, on-field testing, modeling, simulation, performance characterization, etc., and developers can go through them several times before they come up with an acceptable system.

This paper studies and models an architecture for herd localization which can be adjusted to users' needs and extends node autonomy by replacing batteries with a kinetic generator. The authors' effort was primarily driven towards sustaining a traditional lifestyle in a natural environment: Saami herders and their semidomesticated Scandinavian reindeer, *Rangifer tarandus tarandus* L. Highlights and outcomes of their work are enclosed hereafter covering the system developed and a methodology and its tools for performance assessment. In such a context, the scheme followed can be particularly interesting for those working on herd localization systems, who can use it to evaluate new developments in preliminary phases of their work or refine the system presented.

The paper provides information regarding system elements along with their modeling. Likewise, a generic application scenario is modeled from measurements taken by the authors and from other studies in the same region and conditions (Mårell et al., 2002). Results outputted by an agent simulator tailored for such framework are studied for a wide variety of experiments, which allows to evaluate a wide range of operational modes. Simulation outcomes are then used to characterize statistically the system further by obtaining closed forms for a number of statistical estimations. Such knowledge is then used to determine system adjustments for two scenarios with different requirements as examples of its applicability. The system is, therefore, defined conceptually, analytically, and is enough detailed so that other scenarios–species, season, etc. – could be integrated and assessed.

## 2. Materials and methods

### 2.1. System architecture and operation

Three different kind of elements make up the network: primary nodes, secondary nodes and hotspots – see Fig. 1 or (Gutiérrez

---

* Corresponding author. Tel.: +34 91 5495700x4006.
*E-mail addresses:* nelson@gaps.ssr.upm.es (N.I. Dopico), aguti@etsit.upm.es (Á. Gutiérrez), santiago@gaps.ssr.upm.es (S. Zazo).

0168-1699/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.compag.2012.04.015

et al., 2009) for a thorough system description. The equipment described hereafter is designed and manufactured specifically for our application. However, components are commercial off the shelf which can be purchased. Primary and secondary nodes are mounted on animals and are therefore mobile nodes, whereas hotspots may be static (then named base stations) or mobile – carried by an individual or vehicle.

Secondary nodes are the simplest elements in the network and are powered up by kinetic energy from animal movements. Their kinetic generator is made up of a hollow tube with a magnet inside and two coils on its ends – see Fig. 2. The generated voltage follows the Lenz law. The kinetic generator depends on the swing of the magnet inside the tube; therefore, its efficiency depends on its placement and the movement transferred to the generator. Moreover, secondary nodes do not integrate a sophisticated CPU, but just a PIC, a radio transmitter and circuitry to broadcast a unique ID or beacon.

Primary nodes, in turn, integrate a Global Positioning System (GPS) device – which can be switched on and off depending on the final application needs, a CPU, a receiver of secondary-node beacons and a transceiver to communicate with hotspots. They are battery powered.

A hotspot is a battery-powered node which has access to the Internet or another network that makes data available to the end user by means of a monitoring system. It has the same CPU, receiver and transceiver as primary nodes.

### 2.1.1. Operation

The system operates in two stages, namely *Detection* and *Notification*. The first one corresponds to the operation up to the time on which a primary node detects the presence of a secondary node, whereas the latter refers to the notification of the previous detection to a monitoring system by communication between primary node and hotspot.

The aforesaid network components communicate over radio links which operate in two different bands. Secondary-primary links (secondary links hereafter) modulate their information in the 433 MHz band with a data rate of 4.16 Kbps. Primary-hotspot links (hotspot links hereafter) operate in the 166 MHz band and are able to reach up to 200 Kbps.

Communication over secondary links is enabled stochastically by animal movements. Such behavior is a consequence of the secondary-node simplicity which allows to have devices which are not battery powered. As the kinetic generator on secondary nodes harvests enough energy from animal movements, it powers up both circuitry and transmitter in order to achieve an ID beacon transmission. If a primary node is then within the transmission range, it receives and stores the transmitted ID. Communication over secondary links is unidirectional without medium access control. However, transmissions from secondary nodes are not expected to be very frequent (as it is explained later), hence beacon collisions do not impact dramatically on system performance.

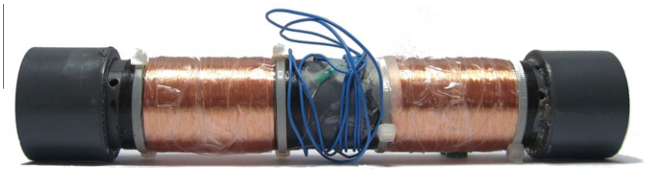


**Fig. 2.** Kinetic generator in secondary nodes.

A primary node which receives a transmission from a secondary node approximates such secondary-node position through its own location – indeed its last GPS reading. This stage is called *Detection*. While such primary node is moving in an environment, it fills out a table with the different secondary-node IDs received, assigned approximated position and the time when the transmission took place. *Notification* happens later: as the aforementioned primary node enters a hotspot communication range, it transmits to such hotspot all the information acquired from secondary nodes along with its own trajectory. A light protocol stack operates on each terminal of a hotspot link which allows for collisions, losses and manages retransmissions.

Depending on the characteristics of the final deployment, hotspots can be on fixed locations, can be carried by an individual or both. Once a hotspot receives data dumped by a primary node, it sends them to the monitoring system. Such monitoring system is expected to receive information from different hotspots, hence it will merge all the information and provide it to the end user.

Thanks to the aforementioned operation, the final system has information about position estimates of primary and secondary nodes. Therefore, it will be able to reconstruct roughly the trail of different animals carrying either a primary node or a secondary node. Note that such secondary-node trail reconstructions are approximations of the real ones as a consequence of the stochastic transmission of secondary nodes, the probabilistic reception of the primary nodes and, to a lesser extent, the discretization of the GPS readings. As already depicted, system operation and architecture pave the way for animal monitoring in outdoor environments, however derivative use to localize other goods, items or individuals

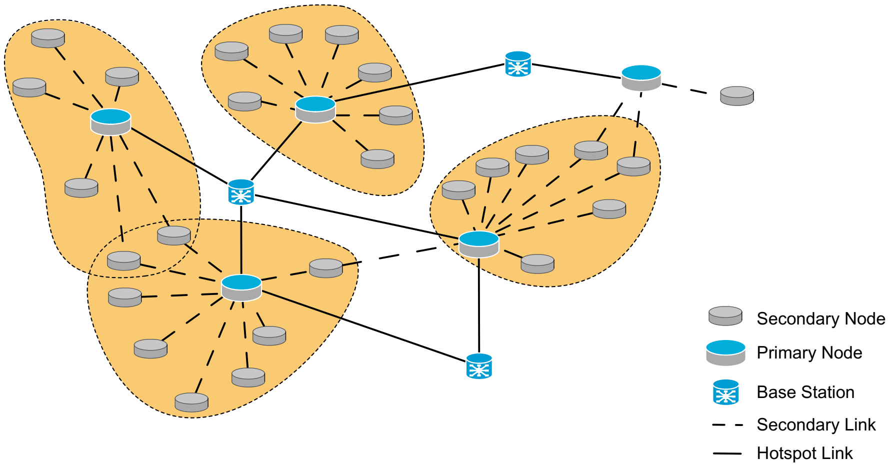


**Fig. 1.** Network architecture.

could be possibly achieved. Nonetheless, in order to work correctly in a specific and real environment, some parameters such as the GPS duty-cycle – which affects primary-node lifetime –, the ratio between primary and secondary nodes and the hotspot deployment, among others, must be defined for each specific application. Since secondary nodes are non-battery powered, simpler and cheaper than primary nodes, designers' goal should be generally to reduce the number of primary nodes and hotspots in favor of secondary nodes.

#### 2.1.2. Equipment trials

One of the first questions a herder or end user may ask regarding such a localization system – in which transmissions are stochastic and depend on someone else *listening* – is how much information will (s)he collect from his(her) herd. Different experiments were developed both in Lapland and Spain with tailored hardware to validate system localization and communication capabilities and serving as the basis for system characterization. Early tests started with its first stages in late 2008 and lasted until 2011. Trials were held in university laboratories, in outdoors in Spain, in Jokkmokk and its surroundings (Sweden) in winter 2009, 2010 and 2011. Hereafter, some test results are highlighted in order to provide readers with some proof of the system feasibility, readers seeking further information should check (Gutiérrez et al., 2009; Dopico et al., 2011) for other details and complementary information.

A first experiment involving the GPS module was carried out in order to estimate the time required for primary nodes to obtain both current time and position. Device capabilities were tested upon a cold start, which means such device is switched on after being off for a long period. On a cold start, a GPS device regularly needs to synchronize with at least four satellites without any previous information regarding their signals and timing. A period of 2 minutes on average is needed for self-localization after a cold start. Position updates can be then performed every second if the GPS module is not completely switched off. Consequently, in the event that the application design fixes a GPS activation period shorter than 2 minutes, there is little difference from an energy point of view between power it on periodically and a permanent operation.

Preliminary trials of energy generation in secondary nodes were conducted in 2009 to test if animal movements were capable to power up nodes – an early version of secondary nodes and mostly the trigger and storage circuit were evaluated. Although the result was limited in scope, it was quite satisfactory first with a dog in southern Madrid (Spain) and later in Jokkmokk with Scandinavian reindeer. Every test on reindeer took place over day time, one reindeer per day over a whole week. Different positions and setups were tested on reindeer body and the best place for the secondary-node location turned out to be its neck because it maximized the energy production from the kinetic generator while, at the same time, eased secondary-node attachment. As animals move their neck (e.g. on search for food on the ground) the generator is swung. Secondary nodes could operate both with two generators in parallel or one, but obviously two leveraged the number of successful transmissions. Results showed that reindeer were able to generate one frame every three minutes in average with two generators.

After that, a proof-of-concept study of the network was performed to validate a first version of the system on live animals. Nodes were operationally tested at temperatures as low as −12 °C in Jokkmokk and −25 °C in a cold experimental chamber. It was then checked if animal movements were able to swing the kinetic generator as much as to yield enough energy for the secondary ID transmission and if node dimensions were suitable for, at least, some mammals like dogs or reindeer (see Fig. 3).

Tests in outdoors comprised also experiments in forest areas of Madrid with individuals (humans) moving while swinging periodically their kinetic generator and carrying a secondary node and a primary node – results were successful over the test (3 hours). Trials in Sweden were carried out with reindeer in a bounded arena of $100 \times 100\ \mathrm{m}^2$ and some other times moving freely in the neighborhood of Jokkmokk (see Fig. 4). Hotspots, repeaters, primary nodes and secondary nodes were tested again during two different weeks. The whole system operated both at night and daytime with a variable number of reindeer across several tests of variable duration – up to five reindeer carried both nodes at a time and the longer test lasted for 24 h. Temperatures as low as −35 °C were reached in testing. Results showed that alkaline batteries were not suitable for our application since a typical 4-AA battery pack, used to power the primary node, was not able to last longer than 12 h in operation. Therefore, lithium batteries are preferred because of the low temperatures. Unlike battery-powered nodes, kinetic generators exhibited no problem at such temperatures.

Previous tests assisted in refining, validating and modeling the hardware supporting our application. From now on, attention is focused on system modeling and the assessment of network operational parameters: ratio of secondary over primary reindeer, GPS acquisition period and the number of hotspots.

### 2.2. Simulation software

In order to study the overall network behavior, a simulator was developed with a reindeer mobility pattern (detailed later) and a time granularity as low as 0.1 s which allows for *collisions* by making the agents (animals) change their trajectory and speed if two are about to collide. The term collision relates here to the need of changing the trajectory of an animal if it perceives another one on its way, i.e., a flight disruption as it is stated in Section 2.3. Changing trajectory may turn out to be a directional change, a speed reduction or simply stopping. For comparison between the simulator used and other network simulators see (Dopico et al., 2011).

### 2.3. Simulation model

Animal paths in our model meet the basic assumptions of a correlated random walk (CRW), in particular independence between movement length and turning angle (no cross-correlation) and symmetric distribution of turning angles around 0 – an equal probability of turning left or right (Kareiva and Shigesada, 1983; Marsh and Jones, 1988; Mårell et al., 2002).

**Fig. 3.** Collar with primary and secondary nodes mounted on a reindeer.

**Fig. 4.** Reindeer with equipment during 2010 trials in Lapland.

A ‘patchy’ environment was not considered, but rather one with equal probability of hosting *a priori* a reindeer on any spot at any time. Such an assumption seems to be feasible according to the conclusions by Mårell et al. as few indications were found that reindeer in their study were performing area-restricted searching behavior. It is consistent as well with observations by Ball et al. (2000).

The scenario upon consideration involves semi nomadic Saami and unlike Bergman’s study (Bergman et al., 2000), our modeled mammals are herded similarly to the description of semidomesticated reindeer by Mårell et al. (2002). Such reindeer husbandry is done in such a fashion that the possibility of escaping is not null as there may be the case in which no fences or barriers do exist. Hence our scenario resembles animal herding not restricted to a corral or, similarly, free animals living in a certain area. Moreover, our mobility pattern fits into the category specified by Marsh and Jones (1988) as ‘Unoriented Movement with No Length-Direction Correlation’ since reindeer are expected to move within certain boundaries that might occasionally traverse, though such a traversing behavior does not take place frequently.

Mårell et al. (2002) studied three time periods over summer 1999 which Edwards (2011) fitted statistically. The third one, which corresponds to late August (namely E), was chosen as a reference for the mobility model in this paper. On the one hand, reindeer seem to move farther distances in E than in any other period, which guarantees localization error not to be underestimated. On the other hand, the frame transmission frequency that arises from the number of recorded up-and-down head movements in E falls between the two other data sets – see Section 2.3.1. In addition, Edwards’s fitting for data set E obtained the best score in the goodness-of-fit test compared to any other fit for any of the three periods.

Overall, the synthetic model detailed below can be sorted as defined in Marsh and Jones (1988) (that is ‘Unoriented Movement with No Length-Direction Correlation’) while its analytical characteristics are borrowed and supported from five sources: turning angles come from Mårell et al. (2002), the definition of a non-patchy environment is supported by Mårell et al. (2002), Ball et al. (2000), the move length distribution deducted from Mårell et al. (2002), Edwards (2011) and the overall CRW considerations from Mårell et al. (2002), Bergman et al. (2000), and Kareiva and Shigesada (1983).

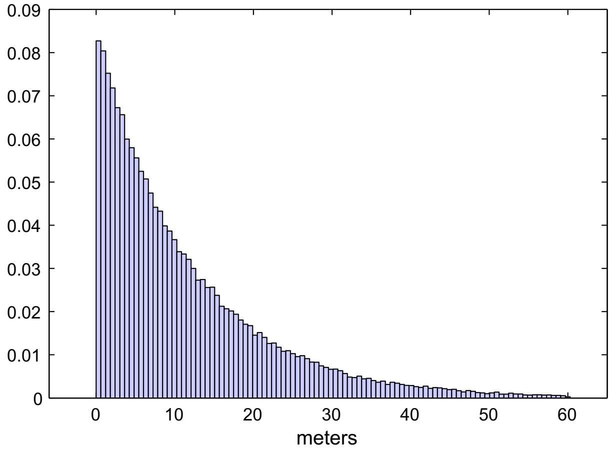
Regarding our study, two parameters make up the animal movement pattern: length of flight distribution (speed distribution) and turning angle distribution.

*Length of flight*. Animal movements have been modeled as Lévy flights many times. Actually, our primary source of statistical information based on data sets (Mårell et al., 2002) assumed them as a plausible option. However, Edwards (2011) performed a study focused on the analysis of Lévy flights as a model for animal movement patterns and concluded that the data set E should be better fitted as an exponential law rather than a power law as it is the common case for pure Lévy flights. Therefore, the Probability Density Function (PDF) of the length of flight is defined as (Edwards, 2011):

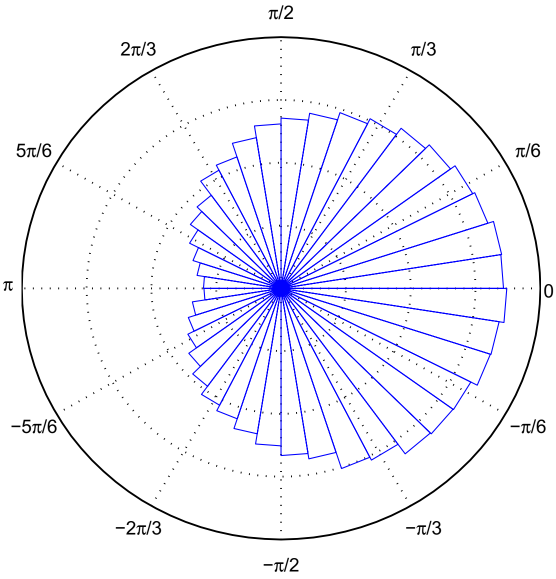
$$f_{Lf}(d) = 0.085e^{-0.085d}, d > 0 \tag{1}$$

Fig. 5a shows the PDF according to Eq. (1). Reindeer were sampled every 30 s, consequently the aforementioned lengths are traversed over 30 s unless one is about to collide with another individual which will make it change its trajectory and interrupt the *flight*.

*Turning angle*. It is modeled as a random variable which keeps to a bounded Gaussian distribution: $\Theta \sim \mathcal{N}(0, \frac{2}{3}\pi),\ -\pi \leqslant \Theta \leqslant \pi$. See Fig. 5b for a polar histogram. Although other fittings could be feasible according to the information provided in the literature, the chosen one was considered plausible due to the reasons referred hereafter. Mårell et al. (2002) analyzed the turning angle ($\Theta$) and checked that it was symmetrically distributed and its sampled



(a) Length of Flight



(b) Turning Angles

**Fig. 5.** Histograms of reindeer mobility pattern.

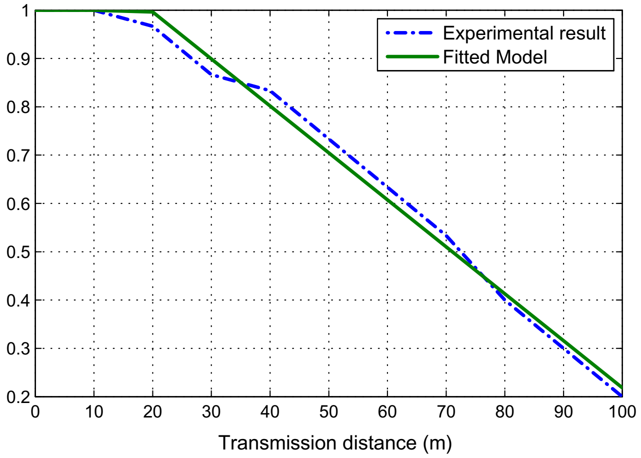


**Fig. 6.** Secondary-link probability of successful reception.

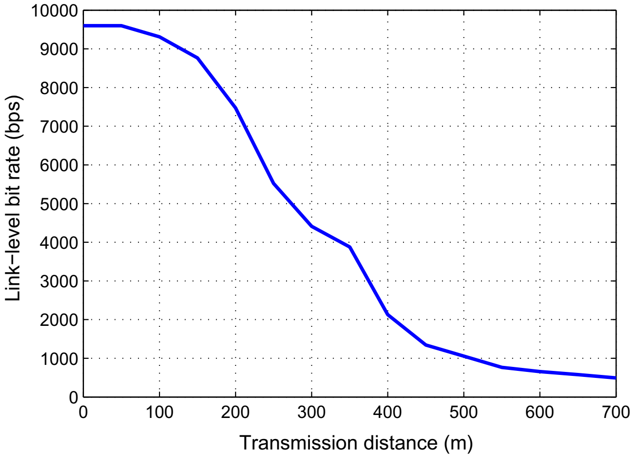


**Fig. 7.** Hotspot link throughput.

mean ($\overline{\Theta}$) was 0. Turning angles were non-uniformly distributed and they were centered around a mean angle of 0. No preference for turning left or right was shown and their distribution was symmetric around 0. Reindeer spent most of their time walking or running and little feeding during the data set collection of Mårell et al., therefore, straight movements ($\Theta_i = 0$) were more frequent rather than continuous turns in the opposite direction ($\Theta_j = \pi$) from the previous direction of movement (Mårell et al., 2002; Bergman et al., 2000).

### 2.3.1. Secondary transmission pattern

From the information provided by Mårell et al. (2002) reindeer move their heads up and down once every 5 min as average over daytime. Such figure has been chosen as it seems to be a fair trade-off from the available information. Likewise, our model makes any simulated discrete time instant equiprobable for a given reindeer to move its head given an average period of 5 min: frames are then transmitted accordingly.

### 2.3.2. Secondary link

Secondary-link model stems from the probability of successful reception by a primary node of beacons sent out by a secondary node placed in an obstacle-free environment on 10 m steps from 0 m to 120 m – see (Gutiérrez et al., 2009) for further information. Frames were received up to 100 m, but no frames were received beyond (on 110 and 120 m). Fig. 6 shows the test results and the model fitted. Transmissions within 20 m are considered to be always successful while from 20 to 100 m the model is fitted by means of the minimum-least-square method. The probability of successful reception is analytically defined as:

$$p_{Rx}(d) = \begin{cases} 1 & \text{if } d < 20 \\ -0.0097d + 1.1907 & \text{if } 20 \leqslant d \leqslant 100 \\ 0 & \text{if } d > 100 \end{cases} \quad (2)$$

### 2.3.3. Hotspot link

Unlike secondary links, hotspot links support a light protocol stack which manages collisions, lost packets and subsequent retransmissions. The farthest communication distance achieved in tests was 720 m (Fig. 7) – see (Gutiérrez et al., 2009) for more information regarding this link. Hotspot links are modeled as disks of radius 600 m wherein communication is always achieved, whereas is unsuccessful from any other spot beyond such boundaries. This way, the existence of a protocol stack is taken into account and, at the same time, extremely low link-layer throughputs (<500 bps) are neglected.

## 3. Performance evaluation

System performance is studied in this section based on three random variables: detections per day, localization error and localization delay. For each of them two statistics from 10th percentile, mean and 90th percentile are estimated and their behavior is compared with system parameters for different configurations: number of primary nodes, GPS activation period, system geometry, etc. Such estimations of statistics from simulations are then regarded as quality indicators which are fitted by means of polynomials or rational functions with respect to the previous parameters. Previous fittings yield analytical expressions which are provided along with their corresponding plot. Since different approximation degrees are possible for every analytical expression, the criterion followed states to use the polynomial of lower degree which warranties a relative error on each estimation below 5% being the relative mean error equal or smaller than 1%.

### 3.1. Detections per day

A secondary node is detected as long as it harvests enough energy to broadcast its ID and a primary node is within its transmission range. Two or more primary nodes may receive the same beacon; such overlap is, nevertheless, accounted as a single detection. This way, the number of detections per secondary node which take place over a day is an intuitive measure of the amount of information that a herder may retrieve from the system.

Dopico et al. (2011) coined the term *active area* as the area within primaries' reception range – any place where a secondary may broadcast from so that a primary may receive its beacon. They concluded that even though the active area mattered, the area size in which animals were expected to move on was more relevant; therefore, it has been found more convenient and intuitive to use the density of primary reindeer per km$^2$ as the conditioning variable to base on the study of the number of daily detections.

More than 50 simulations were performed in order to study just the detection dependence over other parameters like the density of primary reindeer. Detections depend on the mobility pattern and primary-node density. The whole density (primaries plus secondaries) might impact also since it may rise the collision occurrence probability – see Section 2.2. However, simulations show that it can be disregarded in most real cases as no influence was observed in statistical distributions for densities lower than 85 reindeer/km$^2$. Due to computational reasons, tests could not go systematically farther than that ratio to have solid conclusions. We found, nevertheless, 85 reindeer/km$^2$ meaningful for the intended

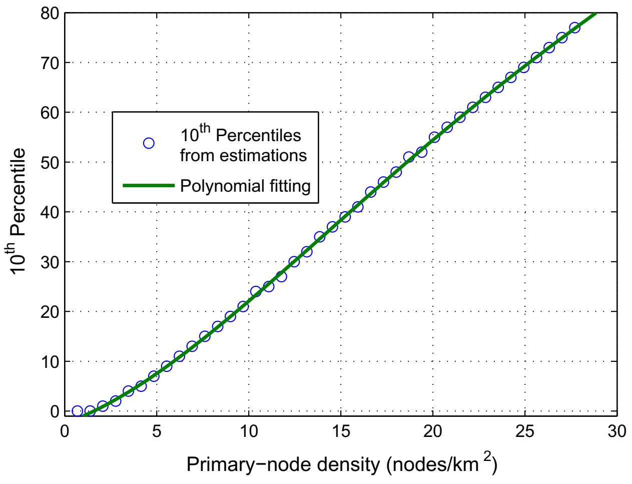


**Fig. 8.** Detections per day – 10th percentiles and fitting over primary density.

proof-of-concept application as the system was primarily developed for semidomesticated reindeer herded extensively. Detections are therefore studied based on their dependence on primary-node density.

### 3.1.1. Quality indicators

As stated earlier, the number of detections a herder can be notified depends on the number of animals with a primary node mounted and the ratio to the area where animals can move. Simulations show a monotonic dependence of statistics such as mean (see (3)) and 10th percentile (see (4) and Fig. 8) on primary-node density. A density as low as 10 prim/km$^2$ can, thus, rise detection mean as many as 44 detections (or 20 in 10th percentile).

$$\overline{Dt}(x) = -3.36 \cdot 10^{-5}x^3 - 3.09 \cdot 10^{-2}x^2 + 4.75x + 6.18 \cdot 10^{-1},\ x > 0 \tag{3}$$

$$Dt_{10}(x) = \begin{cases} 0 & \text{if } 0 < x \leqslant 3 \\ -6.24 \cdot 10^{-5}x^4 - 5.7 \cdot 10^{-3}x^3 & \text{if } x > 3 \\ \quad - 1.65 \cdot 10^{-1}x^2 + 1.31x + 2.48 & \end{cases} \tag{4}$$

## 3.2. Localization error

*Localization error* is defined as the absolute difference between secondary-node actual position and the position stored on a primary node's records upon their encounter – i.e., a primary node that is within such secondary-node transmission range when sending out a beacon.

Understanding how the error comes up may help understand the factors that it depends on and its relation with other system characteristics. By the time a detection takes place, the primary node receiving the ID beacon has read its GPS some time earlier and, consequently, has moved a certain distance ($D_p$). In addition, the beacon has traversed over the air another distance which depends on the propagation conditions ($D_{\text{Tx-sec}}$). Therefore, the system makes two assumptions (or approximations) which lead to the localization error. $D_p$ depends on the mobility pattern and the time elapsed from the last GPS reading. Since power consumption is a fundamental issue in our system, GPS readings are conditioned by the GPS activation period. Consequently, the time elapsed from the last GPS reading (*a priori* a uniform random variable) depends on the GPS activation period. This way, localization error depends on the mobility pattern, GPS activation period and propagation conditions (Dopico et al., 2011). Both the first one and the latter are modeled in previous sections. It may happen that the mobility pattern changes over the year (it actually does as reported by Mårell et al., 2002) and propagation depends on a number of factors such as, for instance, collar position on reindeer – actually the secondary antenna. Since we are assuming that our modeling is valid as a proof of concept that helps us understand system performance, GPS duty cycle is therefore the parameter to focus on the error dependence. As it happened with detections and their timing, *collisions* might influence on the error behavior. However, we have performed additional studies on their contribution and concluded that there was no impact according to our simulations – even in the worst case of a hypothetical closed square field of 1-km side with 160 reindeer inside.

Localization error is closely related to the energy consumption on primary nodes and will influence on their lifetime, hence it is not trivial since it turns out to be a trade-off between autonomy and accuracy.

### 3.2.1. Quality indicators

As it was already explained, localization error is fitted according to the GPS activation period which, in turn, can make secondary-link propagation distance and straight net displacement heavier or lighter with respect to the overall error. Quality indicators display a polynomial fitting each: mean is given in (5) and 90th percentile in (6) – see Fig. 9 for its plot –, but both can be used to assist designers in tasks prior deployment like it is done in Section 5 with case studies.

$$\bar{E}(t) = -1.6 \cdot 10^{-3}t^2 + 9.61 \cdot 10^{-1}t + 67.85,\ t > 10 \tag{5}$$

$$E_{90}(t) = -3.1 \cdot 10^{-3}t^2 + 1.89t + 117.1,\ t > 10 \tag{6}$$

## 3.3. Localization delay

Herders are aware of secondary nodes detected by primary nodes as the latter ones enter a hotspot communication range. If detections take place within hotspot communication range, they consequently have zero notification delay, whereas those off hotspot range are dependent on the primary-node trajectory until connection can be settled with a hotspot or base station. It is therefore possible to sketch localization delays as a mixture of two distributions: one arising from instantaneous notifications and another one due to the primary-node trip times. The probability of instantaneous notifications emerges from the ratio of hotspot transmission range and the overall simulation area, named hereafter coverage ratio – ratio of hotspot coverage area to total area. However, in real cases, it does not just matter such ratio, but environment will play a relevant role too. Soil composition, springs, rivers, hills, cliffs, etc., will determine how animals move across the environment and where they tend to be more often. The same way, if herd and herders are migrating, the scenario changes completely since it does not seem feasible to have base stations all along the way as might happen if a herd is staying on a meadow. Having a herd near herder's residence may change things as well since a hotspot could be installed by the residence and could be carried by the herder over day time. All these examples show the complexity that the study of localization delay exhibits.

Two setups are analyzed in the current subsection, each with a different number of static hotspots (1 and 4) which ease the comprehension as well as the study due to their geometry. Having more may turn out to be an endless effort to solve an np hard problem; that is the reason for considering these two as examples of deployments with reduced number of base stations installed. Extrapolation to other environments is reasonable for the basic conclusions.

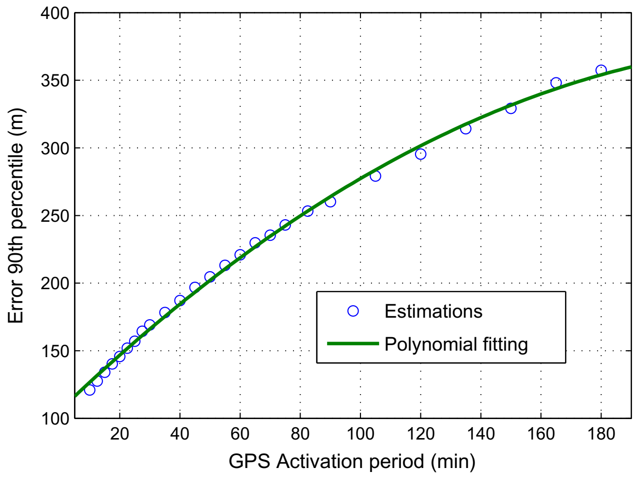
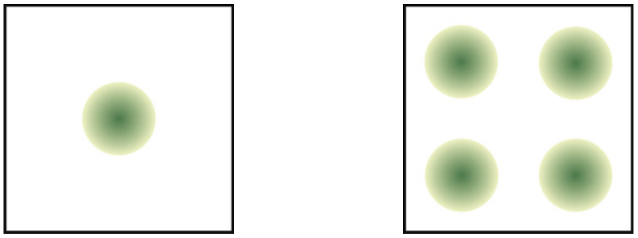


**Fig. 9.** Localization error – 90th percentile and fitting.



(a) H1 (b) H4

**Fig. 10.** Coverage area.

While previous variables perform regardless the simulation-field geometry as long as the primary-node density or GPS duty cycle are taken into account, now such geometry is a key issue as it affects the path that primary reindeer walk from a detection instant to a place with hotspot coverage. Consequently, it is appropriate to point out the simulation area: a square 3.8 × 3.8 km$^2$. Hotspots were placed to keep symmetry, as Fig. 10 shows, and to prevent biasing the outcomes. Due to the size of the simulation area, an additional setup initially considered (16 hotspots) was no further studied since it turned out to be the case in which 99% of the localizations had no delay – consequence of covering most of the simulation area with hotspots. It is important, however, to take it into account as it supports the hypothesis stated later regarding the relationship between null delays and coverage ratio.

Although hotspot transmission range is fixed in our model (see Section 2.3.3), two series of simulations were performed to analyze the impact on delay of a range of coverage ratios leading us to confirm our hypothesis as Fig. 11 shows. In both cases, the expected coverage ratio and the ratio of notifications with null delay ($\alpha$) was similar, though not the same. That is mainly attributed to the simulation border effect making the center (where 1 hotspot is located) slightly more likely to host a reindeer compared to the borders – from which the 4 hotspots are nearer.

Localization delay may impact on any deployment depending on environmental regulations, user's constraints and his(her) needs. If 100% of coverage ratio (or nearly) can be reached, then users will have information in real time. However, not every user may really need or demand such real time operation, for instance, trials reported by Gutiérrez et al. (2009) were conducted in the framework of the project N4C which deals with delay tolerant applications. On the other hand, the less coverage ratio reached, the heavier weight will have random walk lengths up to a coverage area – such random walks will depend not only on the coverage ratio, but on its distribution. Hence, knowledge on places frequently visited by animals or roaming paths may leverage deployment efficiency while reducing its cost and environmental impact.

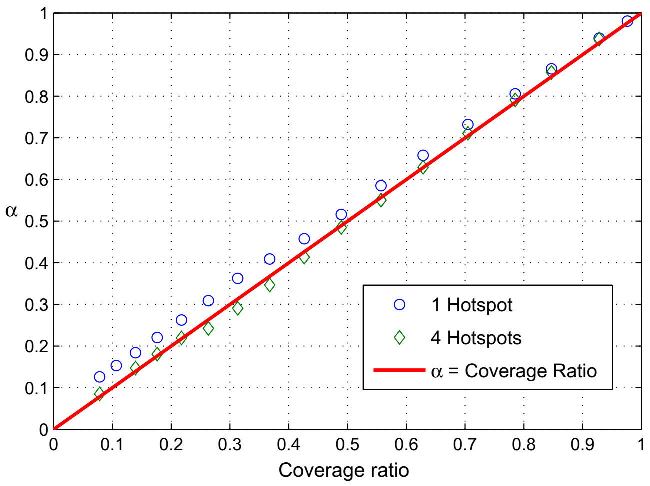


**Fig. 11.** Localization delay - dependence of zero-delay detections on coverage ratio.

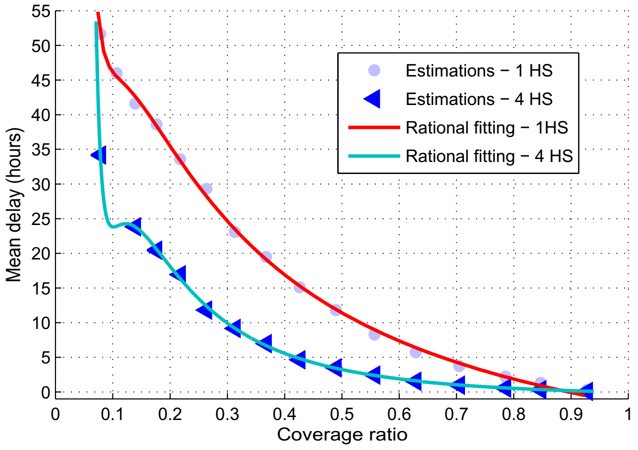


**Fig. 12.** Localization delay – means and fittings for 1 and 4 hotspots.

### *3.3.1. Quality indicators*

Mean delay (7) is computed as the indicator for system latency. Results are given in terms of coverage ratio (defined between 0 and 1) for two hotspot configurations: one hotspot in the center of the area and four hotspots placed in the center of every equally-sized subarea. In both cases, mean functions are monotonically decreasing as the coverage area rises – see Fig. 12.

$$\overline{Dl}(x) = \begin{cases} -2.9\cdot10^{-3}x^{-4} + 1.49\cdot10^{-1}x^{-3} - 2.67x^{-2} + \\ \quad +21.53x^{-1} - 22.3, \;\; 0 < x \leqslant 1, \quad \text{if 1 hotspot} \\ 1.5\cdot10^{-2}x^{-4} - 3.7\cdot10^{-1}x^{-3} + 2.69x^{-2} - \\ \quad -2.44x^{-1} + 0.04, \;\; 0 < x \leqslant 1, \quad \text{if 4 hotspot} \end{cases} \tag{7}$$

## 4. Economic analysis

It is not possible to compare the cost of deploying a prototype, for scale production reasons, with the expenses of mounting GPS

**Table 1**
Node costs.

| Device | Prototype cost (€) | Relative cost |
|---|---|---|
| Primary N. | 200 | 1 |
| Secondary N. | 20 | 0.1 |
| Hotspot | 400 | 2 |
| GPS collar | N/A | 1 |

collars as available on the market. Yet, an analogy is posed hereafter between primary nodes and regular GPS collars to develop a generic economic analysis.

Table 1 displays the production costs of prototypes of the secondary node, primary node and hotspot boards. The third column shows costs relative to a single primary node – GPS collars are considered equivalent to a primary node and used as a currency unit. From a functional perspective, any regular GPS collar and a primary node have similar components: GPS module, CPU, memory, transceiver, battery, antenna and electronics required to interconnect them. Primary nodes however integrate an additional functionality, a receiver for the secondary link and its antenna, which can be considered of little impact (€7) on the overall cost of a primary-node prototype (€200). Although the board of a hotspot and a primary node are essentially the same (the former does not integrate a GPS), hotspot cost doubles primary-node cost because both its (larger) antenna and batteries are more expensive.

For the sake of comparison, seven cases are plotted in Fig. 13. Three with GPS collars mounted on a population of 100, 500 and 1000 heads without support from hotspots – these are the zero-slope thin lines. Three equal to the previous, but with hotspots deployed, which are thickened. The last case is plotted with asterisks and consists of a number of primary nodes which guarantees 14 detections per secondary node as an average – it also includes hotspots. All of them are plotted over a range of square areas from 4 to 500 km$^2$ and hotspot deployments are considered to form a grid (similar to the case of 4 hotspots in Section 3.3) with a density of 1 hotspot per 3 km$^2$ which leads to a mean localization delay of 5 h. Geometrically, the asterisk line is the bound between two planes: the upper left one which is the set of configurations (area and number of primary collars) that make the system analyzed more economical (and regular GPS collars more expensive) and the lower right, on which GPS collars pay off compared to the analyzed solution.

Since the asterisk slope depends monotonically on the required number of detections (given a specific delay), a smaller number of detections will imply a larger upper meaningful surface (more configurations for which our system is worthier) and *vice versa* for larger detection numbers.

From a formal perspective, the system described so far pays off as the number of heads rises, but stating just that would be economical with the truth since the spatial scale and number of required primary collars may still lead the system to be more expensive in a wide range of real cases. As any other solution, deploying an architecture as it has been detailed will be users' decision upon their needs.

On the other hand, deploying hotspots rises costs, but may assist users in retrieving information timely about herd positions regardless GPS collars or primary-secondary nodes are used. Nevertheless, the former (GPS collars) do not really need base stations to operate while the latter are intended to rely on them. In Fig. 13 one can see that for 100 heads the threshold for GPS collars to pay off without hotspots is 30 km$^2$, for 500 it is 140 km$^2$ and for 1000 heads is 275 km$^2$. In the event that a herder wishes to have hotspot infrastructure, given the performance imposed previously, the threshold is shifted upwards – i.e. 35, 165 and 330 km$^2$ respectively.

## 5. Case studies

Two case studies are referred hereafter in order to show the applicability of the previous adjustments.

The first example considers a user with a herd of 200 head which can move within a rectangular area of 3.85 km$^2$ (2.85 × 1.35 km). It is required to detect every reindeer at least 26 times per day (with 90% probability) with a maximum mean error of 100 m. Environmental characteristics allow to reach 100% of coverage ratio, consequently information is obtained in real time.

The previous problem statement leads to use two quality indicators: 10th percentile of daily detections and mean of localization error. Both were defined in Sections 3.1.1 and 3.2.1 respectively. If (4) is solved for $Dt_{10} = 26$, it gives $x = 11.22$ prim/km$^2$ which would

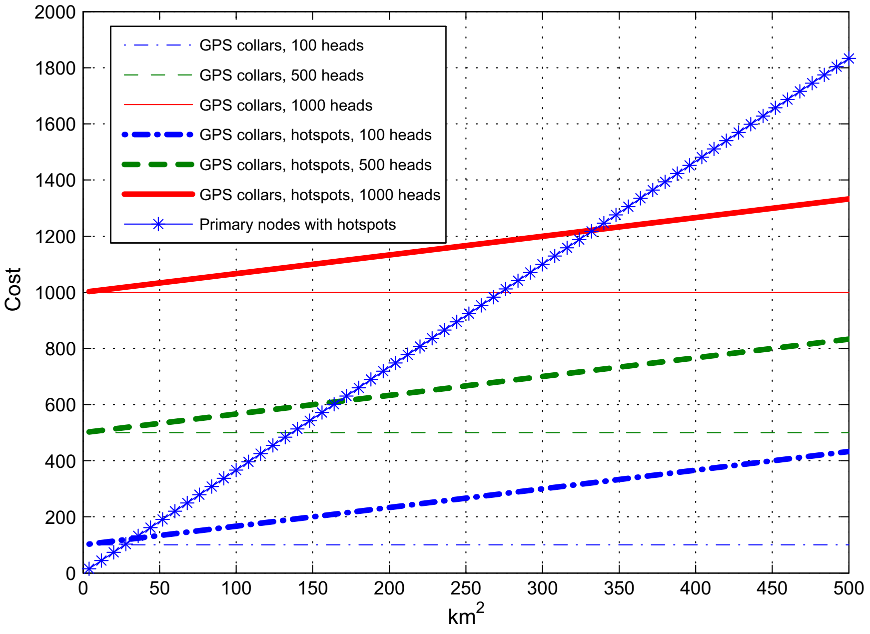


**Fig. 13.** Cost comparison – deployments with regular GPS collars vs. primary-secondary deployment.

**Table 2**
Case study I – performance summary.

| 45 Prim—GPS-36' | Required | Simulation |
|---|---|---|
| Detections (90th percentile) | 26 | 24 |
| Error (mean) | 100 | 100.06 |

**Table 3**
Case study II – performance summary.

| 82 Prim—GPS-120' 4 base stations | Required | Simulation |
|---|---|---|
| Detections (90th percentile) | 12 | 9 |
| Error (90th percentile) | 300 | 268.13 |
| Delay (mean) | 8 h | 6 h 50 min |

make 43.12 primary nodes for the given scenario – 45 were eventually chosen to round it and set a guard band. On the other hand, GPS activation period is obtained by solving (5) for $\overline{E} = 100$ m which gives $t = 35.49$ min. Therefore, GPS activation period must be equal or lesser than 36 min to fulfill the requirements imposed.

Such scenario was simulated over 15 simulation days (initial 4 days accounted as transient period) with the given system setup: 45 primary nodes, 155 secondary nodes and a 36-min activation period. Results were at least 24 detections per day per secondary node with 90% confidence and a localization mean error of 100.06 m – see Table 2.

The second example depicts a herder with not very high requirements, but more head of herd in a larger area – $3.5 \times 3.5$ km. The herder demands a system which allows him to see any secondary reindeer at least 12 times per day with 90% probability and to have a maximum error of 300 m with the same probability.

Quality indicators to be used now are: 10th percentile of detections per day, 90th percentile of localization error and mean of localization delay. Detections determine the density of primary nodes according to (4) – or by Fig. 8 – which turns out to be 6.65 prim/km$^2$ and, for the example considered, it means 82 primary nodes. GPS activation period is the solution of (6) for $E_{90} = 300$ m – graphically can be done finding the $t$-value in Fig. 9 – which gives $t = 118.63$ min – 120 min were then selected. Regarding localization delay, (7) can be used to check if it is possible to fulfill requirements either with one or four hotspots. For one hotspot, mean delay requires 58% of coverage ratio, however it is not possible to reach it due to the area size under study – 12.25 km$^2$ would demand 7.1 km$^2$ of coverage while one hotspot just reaches 1.13 km$^2$. The case for four hotspots can be calculated either from the aforementioned mathematical expression or, graphically, from Fig. 12. In case of four hotspots, it is required a coverage ratio at least of 33% which means 4.04 km$^2$ while these four actually reach 37%; it is therefore possible. According to (7) our setup (four hotspots in 12.25 km$^2$) determines a mean delay of 6.58 h and it is expected to have 37% of localizations notified with null delay.

Detections are a bit different to what was expected, however the usefulness of the prediction becomes obvious. The same way, localization-error simulation result is similar to the constraint imposed. While localization delay is tough to fit as it was outlined in Section 3.3, the formula in (7) has assisted us in finding a proper hotspot setup with four hotspots which covers 37% of the surface. See Table 3 for a summary of this case study.

## 6. Results

Previous sections detailed the process of fitting and modeling system performance. They are an example of how modeling and simulations can support statistical analysis to foresee system behavior for configurations which may have not been even simulated. This way, it can be the basis for future work either on improving the same system or analyzing other systems with a similar methodology. The system on study is thus characterized by three variables, namely detections per day, localization error and localization delay.

The number of detections of a given secondary node is determined by the density of primary nodes moving across the field on study. Secondary transmission pattern and mobility pattern play a relevant role as well. However, based on outcomes of previous works, they are fixed so that the current study may become more comprehensible and tractable. The number of daily detections depends monotonically on the density of primary nodes, thus, 10th percentile is 4 detections per day for 3.46 prim/km$^2$ while 14 detections are reached just doubling density to 6.92 prim/km$^2$.

The system on study is focused on performing rough localization which is affected by the trade-off with system cost. As a consequence, localization error is higher than if just a set of primary-like collars – all integrating a GPS device – were used, however error can be tuned according to user's chances (or will) to replace batteries. For some animals (primary nodes) error depends exclusively on the GPS activation period while others (carrying secondary nodes) are impacted as well by the propagation distance. The study has focused on the latter case (secondary nodes) in which propagation-distance influence becomes diluted as the activation period becomes larger. This way, periods shorter than 1 min suffer clearly from propagation effects while from 10 min on, they are not noticeable. Error evolution over different GPS periods (see Fig. 9) shows a system capable to be tuned according to user's needs while maintaining acceptable figures – 90th percentile from around 120 m (for 10 min) to 350 m for (3 h). Therefore, localization error is tight to GPS duty cycle.

Localization delay depends on the area covered by hotspots, the probability for such area to host animals and the ratio with respect to the whole area where they can move. Environmental constraints may limit hotspot deployments, consequently their location may turn out to be fundamental in some cases. If it is possible to identify frequent places where animals pass by, then deployments can become cost effective. Similarly, if individuals or vehicles may pass by common animal places – or simply traverse the area where they can be – they can become 'mules' who collect the information and relay it to a communication network. In cases where herders travel with their herds, hotspots can be perfectly carried by them providing information with null delay. A study on delay distribution for two deployments of static hotspots is performed concluding that the amount of localizations with null delay is proportional to the coverage ratio while non-zero delays are scaled according to the primary-node trip length which depends on system geometry. In case of one base station and a geometry as the one studied in Section 3.3 the expected value is 51 h (7.8% coverage ratio), but if four base stations are deployed it is lowered to 9 h (31.3% coverage ratio).

## 7. Discussion

According to the reindeer movement considered (i.e. independent from each other) some realistic large scales such as 2000 km$^2$ may turn out the system to be economically unfeasible as it is budgeted currently. In addition, it is possible to argue that previous analyses should be based on simulations which comprised gregarious movements of reindeer rather independent movements from each other. The study is, however, susceptible to be considered as an approximation to such a case as long as one thinks on the simulation arena as a camera objective focused on the herd – i.e., tracking a hypothetical gravity center of the herd regardless whether it moves or not. Anyway, such swarm movement would require precise knowledge on reindeer behavior for different roles within the herd (such as their movement closed form) which is not documented to such extent. Still, gregarious movements do not harm system detection capabilities, but at the least, leverage them; thus lowering the amount of primary collars initially determined for a certain performance – such movement pattern implies dependence between animal movements hence secondary transmissions may be more likely to happen in the neighborhood of other animals carrying a primary node. Herders can then download information relative to where and when each individual was detected for the last time and, thus, have information potentially useful to find sub-herds or individuals which split off from the main group. Such necessity may arise from historical problems derived from the national borders which divide Lapland or arrangements to prevent reindeer from going beyond certain boundaries (Elbo, 1952) – nowadays it still continues to be timely as herders pointed out within N4C project.

Radio links in lower bands may suffer less from propagation attenuation hence secondary transmissions might reach farther distances if they transmitted over lower frequencies. Secondary links operate in a license free band available worldwide; radio band licensing varies however in every country, therefore chances to improve the system may arise in the form of national exceptions as well as future changes in frequency regulations. Changes on secondary links would affect mainly the system capability to detect animals and its accuracy, whereas primary links deal with transmission rates and localization delay.

The system is designed allowing for an infrastructure (hotspot deployment) which assists users in retrieving information. Yet, users can carry personal hotspots and download information directly from primary nodes by themselves – approaching 600 m may suffice to enter their communication range. In case of migration, herders may transport their hotspots and deploy them again in the new environment, thus reducing costs and decreasing environmental impact. The system architecture presented yields a localization delay when there is not 100% of coverage area. Simulations assumed the same probability for any spot to host a reindeer within the simulation arena. In a real case, animals will prefer certain areas more than others for a wide variety of reasons. Therefore, as it was pointed out previously, maximizing the probability for a primary reindeer to reach a place covered by a hotspot after a detection will enable system optimization regarding delay. Prior knowledge on the environment may assist in design as information regarding common animal paths, areas where they use to pass by – lake sides for instance – or any other regarding usual locations can lead to efficient hotspot deployments in which delay be minimized by covering areas of predefined interest. Besides localization delay, hotspots can support an additional purpose in the event that herds should not traverse certain boundaries. Since they can communicate not only with primary nodes, but can also receive frames from secondary nodes, it is possible to record a log comprising any animal passing by their neighborhood regardless it carried either a primary node or a secondary node – transmission distance of every link applies though.

Relative production costs are estimated in Section 4 based on prototype manufacturing; nevertheless, alternative kinetic generators with cost-effective production techniques might lower even more secondary-node cost with respect to primary nodes.

Last but not least, in the event that a user is just concerned about his herd as a whole rather than missing reindeer, it is then possible to restrict the system to primary collars carried by a number of animals which are expected to gather or lead the group. In such a case, GPS duty cycles can be still set in order to extend battery lifetime hence the study on localization error is applicable yet.

## 8. Conclusions

Previous sections have presented modeling tools which may assist in gaining an insight of the applicability of communication technologies on a herding problem. Careful element modeling, simulations and statistics are combined to reach mathematical expressions which may assess system performance regarding the information collected (animal detections), its accuracy and the delay in making such information available to users.

Herders in need of information regarding particular herd individuals may benefit from the system presented, but its overall economical feasibility strives on the area in which the system should operate, the information that it should provide to users with and the facilities demanded.

The operational system area is key because at certain large scales such as 2000 km$^2$ it may turn out to be too expensive.

The information required by users is also important because those just interested in the herd as a whole may dispense with secondary nodes and mount a number of primary collars (similar to regular GPS collars), whereas those interested in single individuals as well should use secondary nodes too. Such latter case may arise because reindeer can be liable to cross national borders or break certain regulations for reindeer husbandry as it is referred in Section 7. In both cases herders may find useful the analyses on localization error or the hotspot infrastructure.

Besides the information that users want to retrieve, the way they do it makes all the difference between deploying hotspots and disregarding any infrastructure off the animals themselves. The system presented comprises hotspots which ease herd information retrieval, but it may be the case that either users cannot deploy them or simply can do without them. In addition, other environments may enable alternatives to static hotspots – they are considered static for simplicity and clarity, but their electronic board is hand held and, therefore, can become portable. This way, users migrating with their herds could transport with themselves a number of hotspots both static (to be left unattended) and hand held, thus making a small deployment cost effective.

Still, as most systems, the system presented is susceptible to be enhanced by reducing the budgeted costs, spanning battery lifetime while maintaining localization accuracy or, contrariwise, improving its accuracy keeping at the same time primary-node autonomy. On the other hand, more efficient generators are another issue which could be addressed in the future to raise detection frequency – impacting on costs derived from primary-node density. The study enclosed in this paper is self-contained so that users may assess the feasibility of the system for their own particular case, whereas researchers and engineers may use the suite of analytical tools and models described as a worked basis for future alternative designs.

## Acknowledgments

N. Dopico thanks his colleagues at C-303 and M. García for their proofreading and comments. This work has been partly funded by the following Grants: TEC2009-14219-C03-01; CSD2008-00010 COMONSENS; FP7-ICT-2009-4-248894-WHERE-2; and TEC2010-21217-C02-02-CR4 HFDVL.

## References

Bailey, D., 2000. Evaluating new approaches to improve livestock grazing distribution using GPS and GIS technology. In: Proceedings of the First National Conference on Grazing Lands, pp. 91–99.

Ball, J.P., Danell, K., Sunesson, P., 2000. Response of a herbivore community to increased food quality and quantity: an experiment with nitrogen fertilizer in a boreal forest. Journal of Applied Ecology 37, 247–255.

Bergman, C.M., Schaefer, J.A., Luttich, S.N., 2000. Caribou movement as a correlated random walk. Oecologia 123, 364–374.

Buerkert, A., Schlecht, E., 2009. Performance of three GPS collars to monitor goats' grazing itineraries on mountain pastures. Computers and Electronics in Agriculture 65, 85–92.

Dopico, N.I., Gutiérrez, A., Zazo, S., 2011. Performance analysis of a delay tolerant application for herd localization. Computer Networks 55, 1770–1783.

Edwards, A., 2011. Overturning conclusions of Lévy flight movement patterns by fishing boats and foraging animals. Ecology 92, 1247–1257.

Elbo, J., 1952. Lapp reindeer movements across the frontiers of northern Scandinavia. Polar Record 6, 348–358.

Grigg, G., Pople, A., Beard, L., 1995. Movements of feral camels in central Australia determined by satellite telemetry. Journal of Arid Environments 31, 459–469.

Gutiérrez, A., González, C., Jiménez-Leube, J., Zazo, S., Dopico, N., Raos, I., 2009. A heterogeneous wireless identification network for the localization of animals based on stochastic movements. Sensors 9, 3942–3957.

Huircán, J.I., noz, C.M., Young, H., Dossow, L.V., Bustos, J., Vivallo, G., Toneatti, M., 2010. Zigbee-based wireless sensor network localization for cattle monitoring in grazing fields. Computers and Electronics in Agriculture 74, 258–264.

Kareiva, P.M., Shigesada, N., 1983. Analyzing insect movement as a correlated random walk. Oecologia 56, 234–238.

Marsh, L., Jones, R., 1988. The form and consequences of random walk movement models. Journal of Theoretical Biology 133, 113–131.

Mårell, A., Ball, J.P., Hofgaard, A., 2002. Foraging and movement paths of female reindeer: insights from fractal analysis, correlated random walks, and Lévy flights. Canadian Journal of Zoology 80, 854–865.

Nadimi, E., Sgaard, H., Bak, T., Oudshoorn, F., 2008. Zigbee-based wireless sensor networks for monitoring animal presence and pasture time in a strip of new grass. Computers and Electronics in Agriculture 61, 79–87.

Rempel, R.S., Rodgers, A.R., 1997. Effects of differential correction on accuracy of a GPS animal location system. The Journal of Wildlife Management 61, 525–530.

Schlecht, E., Hlsebusch, C., Mahler, F., Becker, K., 2004. The use of differentially corrected global positioning system to monitor activities of cattle at pasture. Applied Animal Behaviour Science 85, 185–202.

Zhang, P., Sadler, C.M., Lyon, S.A., Martonosi, M., 2004. Hardware Design Experiences in ZebraNet. Proceedings of the Second International Conference on Embedded Networked Sensor Systems. ACM Press, New York, NY, pp. 227–238.