Design and Analysis of Delay-Tolerant Sensor Networks for Monitoring and Tracking Free-Roaming Animals

Samina Ehsan, Kyle Bradford, Max Brugger, Bechir Hamdaoui, Yevgeniy Kovchegov, Douglas Johnson, Mounir Louhaichi[†]

Oregon State University
[†] International Center for Agricultural Research in the Dry Area (ICARDA)
ehsans,bradfork,bruggerm,hamdaoub,kovchegy,douglas.e.johnson@onid.orst.edu;M.Louhaichi@cgiar.org

Abstract—This paper is concerned with the design and analysis of delay-tolerant networks (DTNs) deployed for free-roaming animal monitoring, wherein information is either transmitted or carried to static access-points by the animals whose movement is assumed to be random. Specifically, in such mobility-aided applications where routing is performed in a store-carry-and-drop manner, limited buffer capacity of a carrier node plays a critical role, and data loss due to buffer overflow heavily depends on access-point density. Driven by this fact, our focus in this paper is on providing sufficient conditions on access-point density that limit the likelihood of buffer overflow. We first derive sufficient access-point density conditions that ensure that the data loss rates are statistically guaranteed to be below a given threshold. Then, we evaluate and validate the derived theoretical results through comparison with both synthetic and real-world data.

I. INTRODUCTION

The deployment of extremely versatile sensor networks in a variety of real world applications is progressing from concept to reality. Wildlife monitoring is an important example that has involved intensive research activity in the past few years [1]. In particular, biologists have long recognized the need for insight into animal habitat, the need for monitoring endangered species, and the need for studying animals' behaviors and movements, as these are all necessary to understand their physiology, behavior, and ecology [2]. However, many wild species are, by nature, free-roaming and wide-ranging which makes it too difficult to track and monitor their behavior and mobility directly through human intervention, thus calling for automated monitoring systems which demand less human presence in the field [1, 3]. In addition, human presence may disrupt animals' normal activities. After initial attempts which, at times, provided inconsistent and invalid outcomes, efforts have been made towards the deployment of sensor networks with enormous potential for this kind of application [4, 5].

In these sensor networks, light-weighted, battery powered, small tracking devices, called collars, are attached to animals' necks to collect and save spatio-temporal data (e.g., location information, biometric, timing information, etc.) without disrupting animals' normal behaviors and movement. These collars are typically equipped with small memory chips with limited capacity, and are designed to operate inconspicuously. At regular intervals, the collar transmits its data to a device where data

storage is not an issue. However, due to the continuous and random movement of the animals and due to their sparsity nature, traditional MANET techniques cannot be applied. Instead, a number of static nodes (with unlimited or enough power and data storage capacity) is deployed in the free-ranging area, and whenever the animal comes close to one of these static nodes, the collar attached to it downloads its generated data to the static node. A special class of sensor networks, known as delay-tolerant networks (DTNs), is considered to be well suited for these wildlife monitoring applications that are typically underserved by traditional networks [6].

Such DTNs can then be thought of as sparse networks composed of mobile nodes (i.e., free-roaming animals), equipped with buffers of limited capacity, and static access-points, equipped with buffers of virtually unlimited (or enough) capacity. Since every mobile node in the network stores an amount of data that increases with time, and there is no guarantee of when mobile nodes will reach the coverage area (the area surrounding the access-point where data transfer can be performed), the buffer may overflow, leading to data loss which may hamper the reliability of the system. It is therefore important to understand how the reliability of such networks depends on the density of the static nodes, as this density relates to both the frequency with which a mobile node visits coverage areas and the probability of buffer overflow. Although considerable research efforts have focused on protocol design [7–9], connectivity analysis [10, 11], delay modeling and characterization [12–14], and mobility analysis [15, 16], the effect of access-point density on data loss is still not well investigated.

In this paper, we derive sufficient conditions on access-point node density of partially covered, intermittently connected DTNs deployed for wildlife monitoring/tracking so that the data loss rate does not exceed a given threshold. To the best of our knowledge, there is no previous work addressing the issue of critical density from this perspective. Due to the limited coverage (the network is disconnected in the traditional sense), data delivery is only possible through animals whose movements are assumed random (henceforth referred to as mobile nodes), which store and carry data until they come close to a fixed node with no power constraints (henceforth referred to as an access-point), where gathered data is then fully downloaded. The focus of this work is then on partially-covered DTNs whose node density and

coverage ratio are both low.

We explore the linear (1-D) and two two-dimensional (2-D) access point deployment structures, and use a mathematical model based on Brownian motion to analyze the movement of mobile nodes. The following summarizes our main contributions:

- Derivation of sufficient access-point density conditions of partially covered, intermittently connected DTNs, consisting of both mobile and static nodes, deployed for wildlife monitoring, to ensure that data loss rates are bounded by a given threshold;
- Asymptotic behavioral analysis of the access-point density when varying the buffer size and/or data loss rate threshold;
- Verification of the derived theoretical results through intensive simulations; and
- Validation of the theoretical analysis and the considered model through real/field data, collected from free-roaming horses via GPS collars.

The remainder of the paper is organized as follows. The network model is introduced in Section II. The sufficient bounds on access-point density is derived in Section III. Verification and validation results are presented in Section IV. Finally, the article is concluded in Section V.

II. DTN MODEL

We consider a delay-tolerant sensor network used for tracking and monitoring free-roaming animals in their natural environment. For this, we study DTNs that can experience long data transmission delays and frequent disconnection, and that consist of access-points laid out on a grid structure and a set of mobile nodes (animals) that can freely move in an open area. We assume that mobile nodes in these sensor networks independently and continuously generate data (for example, the animal's position and speed provided via GPS) at rate c, and buffer it in their collars' memory chips. Whenever a mobile node comes within the coverage area of an access point, it can then download its generated/buffered data to the access point. Each mobile node is assumed to be equipped with a memory chip that has a buffer with limited size of B bits, and when the buffer is full the newly generated data is dropped. Let $\tau = B/c$, which basically represents the minimum amount of time required to overflow the buffer of the mobile node. Also, let $\bar{\epsilon}$ denote the data loss rate threshold that mobile nodes (i.e., the monitoring system) can tolerate.

In these delay-tolerant sensor networks, the access points can only cover the network partially, due to the large size of the sensed area. Therefore, throughout this paper, we assume that the *coverage ratio*[1] of the studied DTNs is relatively low. Here, mobile nodes rely on their mobility to maintain connectivity with the access points. As mobile nodes move, they will eventually traverse an area covered by an access point, allowing them to download their buffered data.

Finding an optimal deployment structure of access points requires an understanding of animal behavior and movement. Wildlife scientists study animals for years and would optimally place the access points based on data from statistical observational studies. However, there is still no good understanding of

[1]The coverage ratio is defined as the fraction of the area covered by access-points' communication ranges relative to that of the network area.
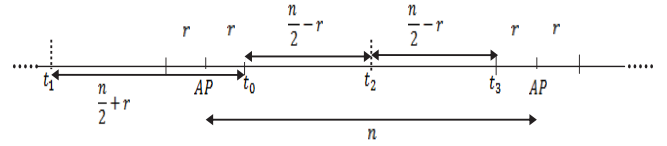

Fig. 1. Each access-point has communication range of length r surrounding it, and is of distance n away from its closest two neighbors in the line structure.

movement behavior of free-roaming animals [17]. This being the case, in this work, we place our access points with a uniform spacing. More specifically, we investigate a one-dimensional (1-D) access point deployment structure with equal spacing, and two two-dimensional (2-D) access point deployment structures in an open space: the square grid and the hexagonal grid. Before describing the studied access-node deployment models, it is worth mentioning that although the main focus of our analysis is on the 2-D deployment models, we are studying the 1-D model here, so as to help us analyze and understand the 2-D case, which is more accurate and realistic than the 1-D case.

We describe the movement pattern of the free-roaming animal as Brownian motion. Random walk and Brownian motion models have already been used in many works that study mobile networks (e.g., [12, 14, 18]). In this work, we validate our mobility model choice via real data traces collected by fitting six free-roaming horses with collars that are equipped with GPS units capable of recording positional fixes at 1-second intervals (Fig. 7). The horses were part of larger herds of about 20 mares, geldings and stud, which were pastured together (data field collection details are given in Section IV-B).

A. One-dimensional Node Deployment Model

We now define the one-dimensional (1-D) node deployment model, where the movement pattern of a mobile node can be described as 1-D Brownian motion on a straight line divided into segments, and each segment has length of n. In line deployment structure, as shown in Fig. 1, each access-point is located at the center of the line segment and has two access-points of distance n away on both sides. From a practical viewpoint, this 1-D model can be used for designing DTNs for tracking and monitoring animals that roam along for e.g. long rivers.

We consider that each access-point has coverage area of length r on each of the two sides. Note that if a mobile node is on any line segment, its closest access-point is at most of distance $\frac{n}{2}$ away from it (as an access-point is placed at the center of the line segment). The node density, ν, and coverage ratio, η, can respectively be expressed as $\frac{1}{n}$ and $\frac{2r}{n}$.

B. Two-dimensional Node Deployment Model

We also consider node deployment structures where access-points are placed via a grid structure, and mobile nodes are free to move within the plane—their paths are modeled by a 2-D Brownian motion. In this paper, we study two node deployment structures: the square grid and the hexagon grid. In the square grid structure, shown in Fig. 2 (left), each access-point is surrounded
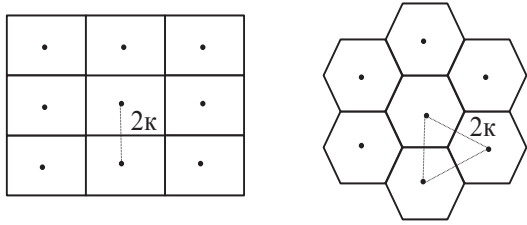
Fig. 2. Each access-point has a communication disk of radius ρ surrounding it, and is distance 2κ away from its four closest neighbors in the square grid structure and its six closest neighbors in the hexagon grid structure.

by four access-points of distance 2κ away. We assume that each access-point is surrounded by a communication disk of radius ρ. We draw a square around each access-point, of side length 2κ, and note that if an animal is anywhere in the square, the access-point at the center is its closest access-point. Thus, the node density, ν, and coverage ratio, η, can respectively be expressed as $1/(4\kappa^2)$ and $\pi\rho^2/(4\kappa^2)$.

In the hexagon grid structure, shown in Fig. 2 (right), access-points are placed in the plane to form a hexagon grid—each access point is surrounded by six access-points in each direction, each 2κ away. We identify an arbitrary point's closest access-point by drawing a hexagon around each access-point. We assume that each access-point is surrounded by a small communication disk of radius ρ, and that each hexagon has *apothem* κ, where the *apothem* is defined to be the length of the shortest line from the center to an edge (the radius is the length of the longest such line). In the hexagon grid structure, the node density ν and the coverage ratio η can be expressed as $1/(2\sqrt{3}\kappa^2)$ and $\pi\rho^2/(2\sqrt{3}\kappa^2)$, respectively.

III. ACCESS-POINT DENSITY ANALYSIS

It is clear that the density of access-points in a given network affects the data loss rate of a mobile node. In this section, we derive and provide sufficient conditions on the access-point density for both 1-D and 2-D access-point deployment models to ensure that the data loss rate does not exceed a given threshold.

Before delving into our analysis, it is important to reiterate that although the 1-D model is not as practical/realistic as the 2-D model, we are studying it here because its mathematics can easily be understood, and hence, it will give us useful insights as well as can facilitate the study of the 2-D model that we investigate later.

A. One-Dimensional Analysis

We begin our analysis by investigating the one-dimensional model. Recall that for this case we delineate the movement of a mobile node by a Brownian motion on a straight line of equal length segments of length n and the access-point is located at the center of each segment and has coverage area of length r surrounding it.

In the 1-D case we define the hitting time as the time it takes a mobile node that just left the edge of the coverage area to reach one of the edges of the line segment and then return to the coverage area again. As the problem is symmetric, it suffices to consider one area between two neighbouring access-points.

Proposition 3.1: The expected hitting time $\bar{\sigma}$ of a mobile node in a line deployment is $n^2/2 - nr$.

Proof: The expected hitting time to the endpoints of interval $[a, b]$, starting at a point x such that $a < x < b$ is given by $\mathbb{E}T(a, x, b) = (b - x)(x - a)$ [19]. Thus, the expected time it takes a mobile node to hit the endpoint (e.g., t_1 or t_2 as shown in Fig. 1) of the line segment when it has just left the coverage area (e.g., t_0 as shown in Fig. 1) is

$$\mathbb{E}T(t_2, t_0, t_1) = \left(\frac{n}{2}\right)^2 - r^2,$$

and the expected time it takes a mobile node to return back to the coverage area (e.g., t_0 or t_3 as shown in Fig. 1) from the end-point (e.g., t_1 as shown in Fig. 1) of the segment is

$$\mathbb{E}T(t_0, t_1, t_3) = \left(\frac{n}{2} - r\right)^2.$$

Hence, the length of the total block T_0, or the expected time it takes a mobile node to leave a coverage area and reach the edge of the line segment plus the time it leaves the edge of a segment and enter a coverage area is $\mathbb{E}T_0 = \frac{n^2}{2} - nr$. ∎

Recall that $\nu = \frac{1}{n}$ denotes the access-point density. We now provide a sufficient condition on the critical density for the 1-D access-point deployment model.

Theorem 3.2: For sufficiently small r, when $\tau \geq r^2$, the data loss rate is guaranteed to remain below the threshold, $\bar{\epsilon}$, if the following condition on access-point density holds:

$$\nu \geq \frac{-r - r\ln\frac{1}{\bar{\epsilon}} + \sqrt{\left(r - r\ln\frac{1}{\bar{\epsilon}}\right)^2 + 2\tau\ln\frac{1}{\bar{\epsilon}}}}{2\tau - 4r^2}$$

Proof: Recall that because of symmetry, it suffices to consider the motion of a mobile node in between two neighbouring access nodes on a line. Let C be a random variable representing the amount of time the mobile node spends inside the communication range r (i.e., inside the coverage area). For small coverage ratio η, we can write $\eta \approx \mathbb{E}C/\bar{\sigma}$ or equivalently $\mathbb{E}C \approx \eta\bar{\sigma}$, where $\bar{\sigma}$ is the expected hitting time and $\mathbb{E}C$ is the expectation of C; i.e., the mean time a mobile node spends in the coverage area.

Now, note that buffer overflow occurs when T minus $\mathbb{E}C$ exceeds τ. Hence, the probability of overflow, P_O, can be expressed as $P(T - \mathbb{E}C > \tau)$ or $P(T > \tau + \mathbb{E}C)$. Since T can be approximated with an exponential distribution with parameter $1/\bar{\sigma}$ [19], P_O can be written as $P_O = \exp\left\{-\frac{\tau}{\bar{\sigma}} - \frac{2r}{n}\right\}$.

To ensure that the probability of overflow P_O does not exceed the threshold $\bar{\epsilon}$, it suffices to prove that $\exp\left\{-\frac{\tau}{\bar{\sigma}} - \frac{2r}{n}\right\} \leq \bar{\epsilon}$. Now, replacing n by $\frac{1}{\nu}$ and for $\tau \geq r^2$ yields (after some algebraic simplification),

$$\left(2\tau - 4r^2\right)\nu^2 + \left(2r^2 + 2r\ln\frac{1}{\bar{\epsilon}}\right)\nu - \ln\frac{1}{\bar{\epsilon}} \geq 0.$$

Solving the above quadratic equation provides the sufficient condition on ν stated in the above theorem. ∎

B. Two-Dimensional Analysis

This section deals with the derivation and analysis of the two two-dimensional models: square and hexagonal grids. We first observe the symmetric structures of the square and hexagon grid deployments: as a mobile node reaches the edge of a square
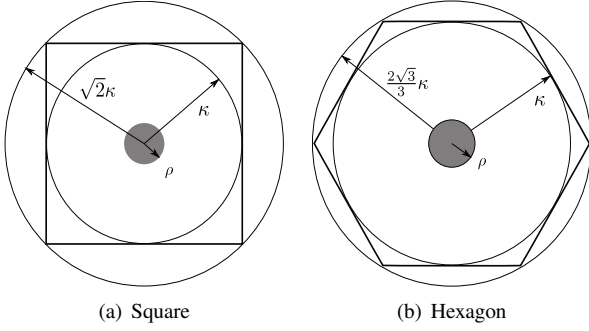
3

Fig. 3. We approximate the square/hexagon by two circles, one inscribed inside, and the other circumscribing the outside, in order to calculate bounds on the hitting time.

or an hexagon, there is no difference between approaching a new communication disk (i.e., coverage area), or returning to the same communication disk, in terms of the time spent outside of coverage areas. Therefore, in our analysis, we restrict our focus to one square in the square grid deployment and to one hexagon in the hexagon deployment.

In this work, we use 2-D Brownian motion to model the movement of the mobile nodes. In the 2-D case we define the hitting time as the time it takes a mobile node that just left the edge of the coverage area to hit the edge of the square or hexagon and then to hit (return to) the coverage area again.

We now provide lower and upper bounds on the hitting time for both 2-D deployment models. Recall that the coverage area has a circular shape and the boundary region has either a square or an hexagonal shape. We then derive a lower bound by considering a mobile node traveling in the circle inscribed in the square (Fig. 3(a)) or in the hexagon (Fig. 3(b)) with radius κ and centered at our access point. Similarly, we find an upper bound on the hitting time by considering a mobile node travelling in the outer circle that circumscribes the square or hexagon. In the case of the square grid, this circle has radius $\sqrt{2}\kappa$ and in the case of the hexagon grid, this circle has radius $2\sqrt{3}\kappa/3$, as shown respectively in Figs 3(a) and 3(b).

Proposition 3.3: For sufficiently small η, the expected hitting time $\bar{\sigma}$ in the square grid deployment is bounded below and above by

$$\kappa^2 \ln\left(\frac{\kappa}{\rho}\right) \leq \bar{\sigma} \leq 2\kappa^2 \ln\left(\frac{\sqrt{2}\kappa}{\rho}\right)$$

Proof: Using the theorem given in [20], it follows that when the coverage area has circular shape of radius ρ and the boundary region also has circular shape of radius $R > \rho$, the expected hitting time can be expressed as $h(\rho, R) = R^2 \ln\left|\frac{R}{\rho}\right|$. Hence, applying this result to the inner boundary region of radius κ and to the outer boundary region of radius $\sqrt{2}\kappa$, as shown in Fig. 3(a), the expected hitting time $\bar{\sigma}$ in the square grid deployment can then be lower bounded by $h(\rho, \kappa)$ and upper bounded by $h(\rho, \sqrt{2}\kappa)$. ∎

Proposition 3.4: For sufficiently small η, the expected hitting time $\bar{\sigma}$ in the hexagon grid deployment is lower and upper bounded as

$$\kappa^2 \ln\left(\frac{\kappa}{\rho}\right) \leq \bar{\sigma} \leq \frac{4}{3}\kappa^2 \ln\left(\frac{2\sqrt{3}\kappa}{3\rho}\right)$$

Proof: Similar to the previous proof, provided that the expected hitting time, when the coverage area has circular shape of radius ρ and the boundary region has circular shape of radius $R > \rho$, can be expressed as $h(\rho, R) = R^2 \ln\left|\frac{R}{\rho}\right|$, the expected hitting time $\bar{\sigma}$ in the hexagon grid deployment can be lower bounded by $h(\rho, \kappa)$ and upper bounded by $h(\rho, 2\sqrt{3}\kappa/3)$. ∎

1) Sufficient Access-Point Density of Square Deployment: Recall that $\nu = \frac{1}{4\kappa^2}$ denotes the access-point density of the square grid deployment. We now provide a sufficient condition on the critical density when deploying the access points in the square grid structure.

Theorem 3.5: For sufficiently small ρ, when $\tau \geq \frac{\pi\rho^2}{8}$, the data loss rate is guaranteed to remain below the threshold, $\bar{\epsilon}$, if the following condition on access-point density holds:

$$\nu \geq \frac{-\pi - 2\ln\frac{1}{\bar{\epsilon}} + \sqrt{\left(\pi + 2\ln\frac{1}{\bar{\epsilon}}\right)^2 + 8\left(\frac{4\tau}{\rho^2} - \pi\right)\ln\frac{1}{\bar{\epsilon}}}}{16\tau - 4\rho^2\pi}$$

Proof: Recall that because of symmetry, it suffices to consider the motion of a mobile node on a single square of the grid. In this case, let C be a random variable representing the amount of time the mobile node spends inside the communication disk of radius ρ (i.e., inside the coverage area), and T be a random variable representing the total amount of time spent inside the square of length 2κ. Following the similar definition provided in proof of the Theorem. 3.2, for this case, the probability of overflow P_O can then be expressed as

$$P_O = \exp\left\{-\frac{\tau}{\bar{\sigma}} - \pi\rho^2\nu\right\} = \exp\left\{-\frac{\tau}{\bar{\sigma}} - \frac{\pi\rho^2}{4\kappa^2}\right\}.$$

From Proposition 3.3, it then follows that P_O is lower and upper bounded as

$$\exp\left\{-\frac{\tau}{h(\rho, \kappa)} - \frac{\pi\rho^2}{4\kappa^2}\right\} \leq P_O \leq \exp\left\{-\frac{\tau}{h(\rho, \sqrt{2}\kappa)} - \frac{\pi\rho^2}{4\kappa^2}\right\}$$

where again $h(x, y) = y^2 \ln\left|\frac{y}{x}\right|$.

We know that for all $z > 0$, $\ln z \leq z - 1$ holds. Hence, it follows that $h(\rho, \sqrt{2}\kappa) \leq f(\rho, \sqrt{2}\kappa)$, where $f(\rho, \sqrt{2}\kappa) = \kappa^2\left(\frac{2\kappa^2}{\rho^2} - 1\right)$, which implies that

$$\exp\left\{-\frac{\tau}{h(\rho, \sqrt{2}\kappa)} - \frac{\pi\rho^2}{4\kappa^2}\right\} \leq \exp\left\{-\frac{\tau}{f(\rho, \sqrt{2}\kappa)} - \frac{\pi\rho^2}{4\kappa^2}\right\} \quad (1)$$

Eq. (1) implies that P_O is also bounded above by

$$\exp\left\{-\frac{\tau}{f(\rho, \sqrt{2}\kappa)} - \frac{\pi\rho^2}{4\kappa^2}\right\}.$$

To ensure that P_O does not exceed the threshold $\bar{\epsilon}$, it suffices that

$$\exp\left\{-\frac{\tau}{f(\rho, \sqrt{2}\kappa)} - \frac{\pi\rho^2}{4\kappa^2}\right\} \leq \bar{\epsilon}.$$

For $\tau \geq \frac{\pi\rho^2}{8}$, replacing κ^2 by $\frac{1}{4\nu}$ yields (after some algebraic simplification),

$$\left(8\tau\rho^2 - 2\rho^4\pi\right)\nu^2 + \left(\pi\rho^2 + 2\rho^2\ln\frac{1}{\bar{\epsilon}}\right)\nu - \ln\frac{1}{\bar{\epsilon}} \geq 0.$$

Solving the quadratic equation provides the sufficient condition on ν proposed in the theorem for square grid deployment. ∎

2) Sufficient Access-Point Density of Hexagon Deployment:
Now, we apply the same approach to derive a sufficient condition on the access-point density for the case of hexagon grid deployment. Recall that in the hexagonal deployment, the access-point density ν can be expressed as $\frac{1}{2\sqrt{3}\kappa^2}$ (as mentioned in Section II).

Theorem 3.6: For a sufficiently small ρ, when $\tau \geq (\pi\rho^2)/(6\sqrt{3})$, data loss rates are guaranteed to remain below a given threshold, $\bar{\epsilon}$, if the following condition on access-point density holds:

$$\nu \geq \frac{-2\pi - 3\sqrt{3}\ln\frac{1}{\bar{\epsilon}} + \sqrt{(2\pi + 3\sqrt{3}\ln\frac{1}{\bar{\epsilon}})^2 + 24\sqrt{3}(\frac{2\sqrt{3}\tau}{\rho^2} - \pi)\ln\frac{1}{\bar{\epsilon}}}}{54\tau - 6\sqrt{3}\rho^2\pi}$$

Proof: Following similar derivation to that given in the proof of Theorem 3.5, P_O can be written as $P_O = \exp\{-\frac{\tau}{\sigma} - \frac{\pi\rho^2}{2\sqrt{3}\kappa^2}\}$.

Now from Proposition 3.4, P_O can be bounded below by

$$\exp\left\{-\frac{\tau}{h(\rho,\kappa)} - \frac{\pi\rho^2}{2\sqrt{3}\kappa^2}\right\},$$

and bounded above by

$$\exp\left\{-\frac{\tau}{h(\rho, \frac{2\sqrt{3}}{3}\kappa)} - \frac{\pi\rho^2}{2\sqrt{3}\kappa^2}\right\}.$$

Again, as $\ln z \leq z - 1$ holds for all $z > 0$, it follows that $h(\rho,\kappa) \leq f(\rho,\kappa)$ where $f(\rho, \frac{2\sqrt{3}}{3}\kappa) = \frac{2\kappa^2}{3}(\frac{4\kappa^2}{3\rho^2} - 1)$. This implies

$$\exp\left\{-\frac{\tau}{h(\rho, \frac{2\sqrt{3}}{3}\kappa)} - \frac{\pi\rho^2}{2\sqrt{3}\kappa^2}\right\} \leq \exp\left\{-\frac{\tau}{f(\rho, \frac{2\sqrt{3}}{3}\kappa)} - \frac{\pi\rho^2}{2\sqrt{3}\kappa^2}\right\}$$

which in turn implies that P_O can be bounded above by

$$\exp\left\{-\frac{\tau}{f(\rho, \frac{2\sqrt{3}}{3}\kappa)} - \frac{\pi\rho^2}{2\sqrt{3}\kappa^2}\right\}.$$

To keep the probability of overflow from exceeding the data loss rate threshold $\bar{\epsilon}$, it also suffices that the above bound to be less than or equal to $\bar{\epsilon}$. Then, replacing κ^2 by its expression $\frac{1}{2\sqrt{3}\nu}$ and after some algebra, it yields that for $\tau \geq \frac{\pi\rho^2}{6\sqrt{3}}$,

$$\left(27\tau\rho^2 - 3\sqrt{3}\rho^4\pi\right)\nu^2 + \left(2\pi\rho^2 + 3\sqrt{3}\rho^2\ln\frac{1}{\bar{\epsilon}}\right)\nu - 2\ln\frac{1}{\bar{\epsilon}} \geq 0.$$

Solving the above quadratic equation provides the stated sufficient condition on ν for hexagonal access-point deployment. ∎

C. Asymptotic Analysis

We are also interested in studying the asymptotic behaviors of the access-point density for the studied DTNs. Note that to ensure that the probability of overflow does not exceed the required threshold, it suffices that the density remains below a certain value; which we derived and proposed in Theorems 3.2, 3.5 and 3.6—we call this value the sufficient access-point density and denote it as ν^s. Note that ν^s depends on the communication range r or radius ρ, the time to overflow the buffer τ, and the given threshold $\bar{\epsilon}$.

Corollary 3.7: For fixed $\bar{\epsilon}$, the sufficient density ν^s in line, square and hexagon grid deployments is $\Theta(1/\sqrt{\tau})$ as $\tau \to \infty$.

Thus for all deployments, as the buffer size increases to infinity, the sufficient access-point density decreases asymptotically as fast as the inverse of the square root of the buffer size.

Corollary 3.8: For fixed τ, the sufficient density ν^s in the line, square and hexagon grid deployments is $\Theta(\ln\frac{1}{\bar{\epsilon}})$ as $\bar{\epsilon} \to 1$.

In other words, as the data loss rate threshold $\bar{\epsilon}$ goes to 1, the sufficient density ν^s decreases asymptotically as fast as $\ln\frac{1}{\bar{\epsilon}}$.

IV. VERIFICATION AND VALIDATION

The derived results are first verified via simulations using MATLAB, and then validated via real-field traces collected by equipping six free-roaming horses with GPS units capable of gathering and recording position and speed information.

A. Verification Through Simulated Data

We use MATLAB to verify the derived sufficient conditions presented in Theorems 3.2, 3.5 and 3.6 by mimicking 1-D Brownian motion on a line and 2-D Brownian motion in the bounded square and hexagonal regions.

Because of the symmetry and equidistant placement of access points on a line, for the 1-D case, we consider a unit line segment of length n, with one access point located at its middle. To simulate a 1-D Brownian motion, we generate a normal random variable with parameters ($\mu = 0, \sigma^2 = 1$) for the displacement of random variable on the line segment. An access point with communication range of r on its both sides is placed in the middle of the line segment of length n. We use it to simulate hitting time which is as previously defined: the time it takes a mobile node that just left the edge of the coverage area to reach one of the edges of the line segment and then return to the coverage area again. The hitting time is in turn used to measure the average data loss rate (ϵ_m), which is a function of the communication range r, access-point density ν, and time required to overflow the buffer τ. In our simulation, we consider three different values of $r = 10$, 20 and 30. We also consider different values of ν (which depends only on r) for measuring the average data loss rate (for each r). We set the data loss rate threshold $\bar{\epsilon} = 0.9$ (but other values can also be tried). For a given r, we calculate the value of τ as $\tau = r^2$. Also, for a fixed r, from Theorem 3.2, we calculate the theoretical sufficient density, ν^s, which is a function of r, τ, and $\bar{\epsilon}$. For each ν^s, we simulate and measure the average data loss rate for various values of ν.

Similarly, we perform the 2-D Brownian motion simulation by taking two normal random variables with distribution ($\mu = 0, \sigma^2 = 1$) at each time step for the distance the mobile node travels in a unit time interval in the x- and y-directions. In the 2-D cases, we then use it to simulate the hitting time as previously defined: the time it takes the Brownian motion to leave a communication disk (i.e., coverage area), having radius ρ centered at the origin, then hit the boundary of the square or hexagon, and then return to the communication disk. Here also, the hitting time is in turn used to measure the average data loss rate (ϵ_m), which is a function of the radius of the communication disk ρ, access-point density ν, and time required to overflow the buffer τ. Similar to the 1-D case, in the 2-D simulation we consider three different values of ρ, $\rho = 10$, 20, and 30 and ν (which depends only on κ) for measuring the average data loss rate (for each ρ). We set the data loss rate threshold $\bar{\epsilon} = 0.9$ for each of the two grid deployments. For a given ρ, we calculate the value of τ as $\tau = \pi\rho^2/8$ for square deployment and $\tau = \pi\rho^2/(6\sqrt{3})$ for the hexagon deployment (for both cases, it
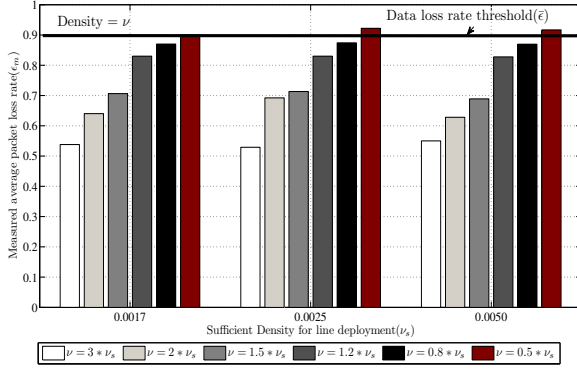
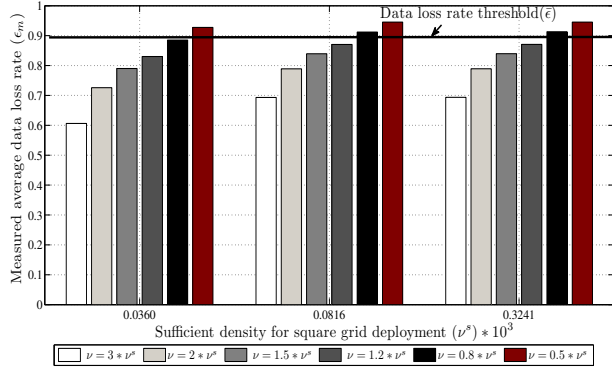Fig. 4.   Measured average data loss rate for the line deployment

Fig. 5.   Measured average data loss rate for the square grid deployment

is notable from Theorem 3.5 and Theorem 3.6 that for a given ρ, any two fixed values greater than the above values can be used). Also, for a fixed ρ, from Theorem 3.5 and Theorem 3.6, similar to the 1-D model verification, we measure the theoretical sufficient density, ν^s, as a function of ρ, τ, and $\bar{\epsilon}$ and various values of ν.

Figs. 4, 5 and 6 illustrate the verification of Theorems 3.2, 3.5 and 3.6. There are three observations that we make from these three figures. First, observe that when ν is higher than the theoretical sufficient density, ν^s, the measured average data loss rate is below the given threshold, regardless of the deployment structure. Second, also observe that when the measured data loss rate is above the given threshold, then the corresponding density violates the sufficient density condition. For example,

Fig. 6.   Measured average data loss rate for the hexagon grid deployment

Fig. 7.   The GPS collar/unit attached to a gelding

the first bar from the right (in both figures) corresponds to a measured data loss rate that exceeds the threshold, but note that the corresponding density ν does not meet the sufficient density condition either; i.e., $\nu = 0.5\nu^s < \nu^s$. Third, note that when ν is lower than the sufficient density, ν^s, the average data loss rate may or may not exceed the given threshold, since our derived conditions are sufficient (for both 1-D and 2-D models). For example, the measured data loss rate exceeds the required threshold in the case of square grid deployment, as shown in Fig. 5 for $\nu = 0.8\nu^s$ (second bar from the right) and $\nu^s = 0.0816 \times 10^{-3}$. Whereas, the measured data loss rate does not exceed the threshold in the case of line deployment and hexagon grid deployment even when $\nu = 0.8\nu^s$, as shown in Fig. 4 for $\nu^s = 0.0025$ and Fig. 6 for $\nu^s = 0.1084 \times 10^{-3}$. This means, as mentioned earlier, that the conditions provided in the theorems are sufficient, but not necessary.

B. Validation Through Real Field Data

In the previous section, we verified our derived sufficient conditions by generating and simulating Brownian motions using MATLAB. We now use real position information (traces) of free-roaming horses (obtained from an interdisciplinary project undertaken in collaboration with biologists at the Rangeland Ecology and Management Department at Oregon State University) to not only verify our derived results, but also to validate our Brownian motion assumption.

Field data collection. GPS data (position, speed, and time) is gathered by fitting three mares and three geldings with collars that contained GPS units capable of collecting and recording positional fixes at 1-second intervals (Fig. 7). The horses were part of larger herds of about 20 mares, geldings and stud which were pastured together. GPS information was collected for up to 149 hours between 26 March 2009 and 1 April 2009. Animals with GPS collars were free-roaming throughout this period. The study pasture used for this investigation are located on the Catlow Rim of Harney County, Oregon, and consisted of a series of longitudinal valleys drained by Skull Creek and Echart Creek. The area of the pasture is approximately 8,100 ha (20,000 acres). We take the GPS data of each horse individually and plot them in x- and y-directions with the aim of validating our theoretical results with real data. For illustration, we consider and show in Fig. 8 the movement data of one horse among the six.
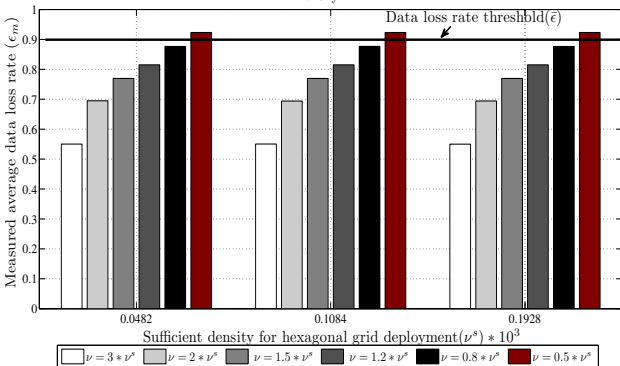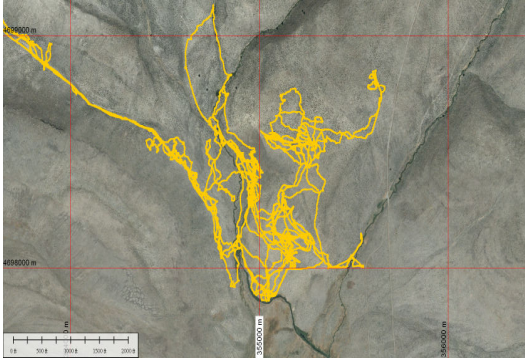
Fig. 8. Tracklogs of a gelding carrying a GPS Unit between 26 March 2009 and 1 April 2009
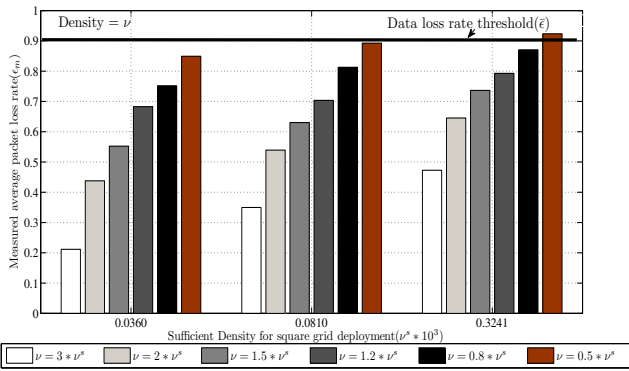

Fig. 9. Measured average data loss rate based on real-world data for the square grid deployment

Analysis. We use the same validation set-up used in Section IV-A for the square grid deployment. We set ρ = 10, 20, and 30. We also consider different values of ν (derived from the theoretical sufficient density ν^s, which is a function of ρ, τ, and $\bar{\epsilon}$) for measuring the average data loss rate (for each ρ). We set the data loss rate threshold $\bar{\epsilon} = 0.9$ and for a given ρ, we calculate the value of τ as $\tau = \pi\rho^2/8$, the same as performed for the square grid deployment in the previous section.

Fig. 9 clearly shows that in all cases, the proposed model of the sufficient node density conditions show fairly accurate performance when applied to the real horse data. Note that when ν is higher than the theoretical sufficient density, ν^s, the measured average data loss rate remains below the given threshold. Again, when the measured loss rate surpasses the given threshold, the corresponding density violates the sufficient density condition. Moreover, when the sufficient condition is being violated, it either remains below (for example, $\nu = 0.8\nu^s$ (second bar from the right) and $\nu^s = 0.0360 \times 10^{-3}$) or goes above (for example, $\nu = 0.5\nu^s$ (first bar from the right) and $\nu^s = 0.03241 \times 10^{-3}$) the data loss rate threshold. Note that the validation with hexagonal grid deployment yields a similar conclusion.

To summarize, in this section, we first verified our derived results via MATLAB simulations; i.e., we validated the derived theoretical conditions, provided the Brownian motion assumption. Second, since our derived theoretical results and analysis are based on the Brownian motion model, as the real data is shown to follow the sufficient conditions (stated in the derived theorems), we conclude that our assumption that the free-roaming animal movement can be approximated by Brownian motion is realistic for free-roaming horses.

V. CONCLUSION

We derived sufficient conditions on the density of access-points of DTNs deployed specifically for wildlife tracking. We also analyzed the asymptotic density behavior under various design parameters. Finally, we validated our results via synthetic and real field data. By verifying and validating our model with real data, we found that our analysis can be useful for modelling and designing delay-tolerant sensor networks for tracking and monitoring free-roaming and wide-ranging animals.

REFERENCES

[1] M. Louhaichi, P. E. Clark, M. D. Johnson, D. G. Ganskopp, R. C. Cook, M. Vavra, and D. E. Johnson, "Using high frequency GPS to determine spatial-temporal activity of ungualtes," in *Proc. of the American Society for Photogrammetry and Remote Sensing*, May 2008.

[2] M. Louhaichi, K.D. Wilson, L.L. Larson, M.D. Johnson, D.E. Johnson, and P.E. Clark, "Comparison of the diurnal pattern and magnitude of velocities of goats (capra hircus), sheep (ovis aries), horses (equus caballus) and cattle (bos taurus)," in *62nd Annual Meeting Soc. for Range Manage*, 2010.

[3] M. Rutishauser, V.V. Petkov, T. Williams, C. Wilmers, J. Boice, K. Obraczka, and P. Mantey, "CARNIVORE: A Disruption-Tolerant System for Studying Wildlife," in *Proc. of 19th Int'l Conf. on Computer Communications and Networks*, 2010, pp. 1–8.

[4] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proc. of the ACM Workshop on Wireless Sensor Networks and Applications*, 2002, pp. 88–97.

[5] P. Juang, H. Oki, Y. Wang, M. Martonosi, L.S. Peh, and D. Rubenstein, "Energy-efficient computing for wildlife tracking: Design tradeoffs and early experiences with zebranet," in *ACM SIGOPS Operating Systems Review*, 2002, vol. 36, pp. 96–107.

[6] J-H. Huang, Y-Y. Chen, Y-T. Huang, P-Y. Lin, Y-C. Chen, Y-F. Lin, S-C. Yen, P. Huang, and L-J. Chen, "Rapid Prototyping for Wildlife and Ecological Monitoring," *IEEE Systems J.*, vol. 4, no. 2, pp. 198–209, 2010.

[7] Q. Li, S. Zhu, and G. Cao, "Routing in socially selfish delay tolerant networks," in *Proc. of IEEE INFOCOM*, 2010, pp. 1–9.

[8] P. Jacquet, B. Mans, P. Muhlethaler, and G. Rodolakis, "Opportunistic routing in wireless ad hoc networks: Upper bounds for the packet propagation speed," *IEEE JSAC*, vol. 27, no. 7, pp. 1192–1202, 2009.

[9] T. Spyropoulos, T. Turletti, and K. Obraczka, "Routing in delay-tolerant networks comprising heterogeneous node populations," *IEEE Tran. on Mobile Computing*, pp. 1132–1147, 2009.

[10] F. De Pellegrini, D. Miorandi, I. Carreras, and I. Chlamtac, "A graph-based model for disconnected ad hoc networks," in *Proc. of INFOCOM*, 2007.

[11] Z. Kong and E.M. Yeh, "Connectivity and latency in large-scale wireless networks with unreliable links," in *Proc. of INFOCOM*, 2008.

[12] R. La, "Distributional Convergence of Inter-Meeting Times Under Generalized Hybrid Random Walk Mobility Model," *IEEE Tran. on Mobile Computing*, 2010.

[13] P. Jacquet, B. Mans, and G. Rodolakis, "Information propagation speed in mobile and delay tolerant networks," *IEEE Tran. on Information Theory*, vol. 56, no. 10, pp. 5001–5015, 2010.

[14] T. Karagiannis, J.Y. Le Boudec, and M. Vojnovic, "Power law and exponential decay of inter contact times between mobile devices," *IEEE Tran. on Mobile Computing*, 2010.

[15] M. Garetto, P. Giaccone, and E. Leonardi, "On the capacity of ad hoc wireless networks under general node mobility," in *Proc. of INFOCOM*, 2007, pp. 357–365.

[16] M. Abdelmoumen, E. Dhib, M. Frikha, and T. Chahed, "Impact of mobility patterns on the performance of routing protocols in delay tolerant networks," in *Internal Report, SupCom*, June 2010.

[17] T. A. Patterson, M. Basson, M. V. Bravington, and J. S. Gunn, "Classifying movement behaviour in relation to environmental conditions using hidden markov models," *Journal of Animal Ecology*, vol. 78, 2009.

[18] P. Jacquet, B. Mans, and G. Rodolakis, "On space-time capacity limits in mobile and delay tolerant networks," in *Proc. of INFOCOM*, 2010.

[19] D.J. Aldous, *Probability approximations via the Poisson clumping heuristic*, Springer-Verlag New York, 1988.

[20] K. Bradford, M.Brugger, S. Ehsan, B. Hamdaoui, and Y. Kovchegov, "Data loss modeling and analysis in partially-covered delay tolerant networks," in *Proc. of IEEE ICCCN*, August 2011.