This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2007 proceedings.

Capacity Deficit in Mobile Wireless Ad Hoc Networks Due to Geographic Routing Overheads

Nabhendra Bisnik and Alhussein A. Abouzeid
Department of Electrical, Computer, and Systems Engineering
Rensselaer Polytechnic Institute
Troy, NY 12180
bisnin@rpi.edu, abouzeid@ecse.rpi.edu

***Abstract*—Overheads incurred by routing protocols diminish the capacity available for relaying useful data over a mobile wireless ad hoc network. Discovering and understanding the lower bounds on the amount of protocol overhead incurred for routing data packets is important for development of efficient routing protocols, and for understanding the actual (effective) capacity available for network users. In this paper we use an information-theoretic approach for characterizing the minimum routing overheads of geographic routing in a mobile network. We formulate the minimum overhead problem as a rate-distortion problem. The formulation may be applied to networks with arbitrary traffic arrival and location service schemes. We evaluate lower bounds on the minimum overheads incurred for maintaining the location of destination nodes and consistent neighborhood information in terms of node mobility and packet arrival process. We also characterize the deficit caused by the routing overheads in the overall transport capacity of a mobile network.**

I. INTRODUCTION

Mobile ad hoc networks are characterized by dynamically changing network topology which makes routing packets in an ad hoc network a very challenging problem. Routing protocols either fail to cope with the changing topology and yield low packet delivery rate or incur very high overhead. It is important to understand the lower limits of overhead incurred for routing packets with certain level of reliability. Knowledge of such a fundamental overhead limit would not only allow researchers to know how much a protocol deviates from the optimal but also inspire the development of routing protocols that achieve the limit.

The primary goal of a routing protocol is to gather and disseminate *state information* such that a node may take packet forwarding decisions that satisfy certain performance criteria. The state information may comprise link states, node locations, velocity and direction of nodes, queue lengths, etc. The performance criteria could be minimum delay, maximum throughput, maximum lifetime, delivering certain fraction of packets or simply best effort delivery of packets. Information theory provides us with lower bounds on the minimum number of bits required to encode a source. Thus, it is reasonable to expect that information theory would be a suitable tool for developing lower bounds on the amount of overhead incurred by routing protocols for disseminating and gathering state information in a mobile ad hoc network.

In this paper we present an analytical information-theoretic framework for characterizing the minimum overhead incurred by geographical routing protocols. The performance criterion considered is high packet delivery ratio. In geographic routing each node maintains its location information at one or more *location servers* (e.g. see [1], [2]). When a source wants to forward a packet to a destination, it queries an appropriate location server for the location of the destination. The location server replies to the source node with the available location information. Thereafter, the source and intermediate nodes forward the packet according to the location of the destination. It is pointed out in [3] that the fraction of packets delivered by geographical routing varies inversely with the average error in location information stored at the location servers. Thus, maintaining packet delivery ratio above a given threshold corresponds to maintaining location errors below a certain threshold.

We categorize geographic routing overheads into two categories: (i) *Location update overhead*: The overhead incurred in updating the location servers such that the location errors in the reply to location queries is less than ϵ, and (ii) *Beacon overhead*: The overhead incurred in beacon transmission such that the probability that a node has consistent neighborhood information when it needs to forward a packet is greater than $1 - \delta$. We formulate the problems of finding the minimum values of the above-mentioned overheads as rate-distortion problems (please see [4] for an introduction to rate-distortion theory). For location update overheads, the distortion measure used is squared error in the location information stored at the location servers (*squared error distortion measure*). For beacon overheads, the distortion measure is the probability that a perceived neighbor is not an actual neighbor (*Hamming distortion measure*). Using a rate-distortion formulation, we present lower bounds on the minimum geographic routing overhead incurred in terms of node mobility, packet arrival process, and reliability criteria ϵ and δ. First, we consider one-dimensional network case and then extend the results to two-dimensional networks.

We compare the minimum geographic routing overhead with the transport capacity of stationary multihop wireless networks evaluated in [5]. It is observed that when the node mobility is high and the average packet inter-arrival time is

This work was funded in part by the National Science Foundation under grants CNS-322956 and CNS-546402.

0743-166X/07/$25.00 ©2007 IEEE

sufficiently small, the complete transport capacity of an ad hoc network may be consumed by routing overheads. We derive an upper bound on the critical network size above which all the transport capacity of the network would be consumed by the routing overheads and no useful communication would be possible. In this paper we only consider the scenario where the routing protocol initiates the forwarding process as soon as a packet arrives at the source. Thus the potential capacity improvement due to node mobility (achieved at the cost of delay associated with waiting for the destination to move to a nearby location) pointed out in [6], [7] is not applicable to our work.

The main contributions of this paper may be summarized as follows: (i) We present a new information-theoretic formulation for evaluating the minimum routing overhead incurred by geographic routing. The formulation is general so that it may be applied to any node distribution, packet arrival process, and may be extended to any location service scheme and mobility model. (ii) For Brownian mobility model and various packet inter-arrival time distributions, we evaluate lower bounds for the minimum rate at which a node must transmit its location information and beacons such that the packets are routed with desired level of reliability. Combining both overheads, we find a lower bound on the capacity deficit caused by geographic routing overheads in mobile wireless ad hoc networks. (iii) We characterize the effective transport capacity of an ad hoc network after taking into account the minimum routing overheads that must be incurred for reliable geographic routing. (iv) For a given packet arrival process, standard deviation of Brownian motion and reliability parameters (ϵ, δ), we evaluate the upper bound on the number of nodes the ad hoc network can support such that the complete transport capacity of the network is not used up by routing overheads.

The rest of the paper is organized as follows. A brief overview of related work is presented in Section II. The network model is presented in Section III. The rate-distortion formulation and evaluation of a lower bound on the minimum position update and beacon overheads are presented in Sections IV and V respectively. A discussion of the capacity deficit caused by routing overheads is presented in Section VI. We present conclusions and directions for future research in Section VII.

II. RELATED WORK

So far, we believe information theory has not significantly influenced the design and understanding of communication network protocols (an opinion we share with the authors of [8]). One of the earliest (and most significant) attempts in using information theory to enhance the understanding of communication networks was made in [9]. Gallager [9] used an information-theoretic approach in order to characterize a lower bound on the amount of protocol information required to keep track of the sender, receiver and timing of messages for a simple (stationary) network model. It is found that although the introduction of message delay decreases the protocol information, small average message length and high message arrival rate may lead to prohibitively high protocol overhead.

A few relatively recent papers have used information theory to understand the effects of node mobility on wireless networks. An analytical framework, based on entropy of node location, for characterizing delay and overhead associated with paging and routing a call to a mobile station in a cellular environment is provided in [10]. The complexity of tracking a mobile user in a cellular environment is studied using a information-theoretic approach and a position update and paging scheme is proposed in [11]. An entropy based modeling framework for evaluating and supporting route stability in mobile ad hoc networks is proposed in [12]. In [13], the authors propose the entropy of link change as the metric for mobility models against which performance of wireless network protocols could be evaluated.

Our work is along the lines of [14] where the authors use an information-theoretic approach to characterize the minimum routing overhead and memory requirements of topology-based (proactive) hierarchical routing protocols for ad hoc networks. The entropy of ad hoc network topologies as well as the entropy rate are used in [14] to find the above mentioned bounds. However, here, the family of routing protocols considered is geographic routing protocols and the performance constraints are also taken into account. This leads to a new problem formulation as rate distortion which was not considered in earlier work and new results on the effect on transport capacity.

III. NETWORK MODEL

The network consists of n mobile nodes. The nodes perform Brownian motion with variance σ^2. We consider two kinds of network deployments: (i) *One dimensional case:* nodes located along a circle of perimeter L, and (ii) *Two dimensional case:* nodes located over a torus of surface area A. The central and lateral radii of the torus are denoted by R_c and R_l respectively. The closed curve and surface are chosen for the study, instead of a finite line or a rectangle, in order to avoid the complexity of modeling the behavior of Brownian motion at boundary points.

We assume that $L >> \sigma^2$ and $R_c, R_l >> \sigma^2$. The large dimensions ensure that the nodes do not wrap around the curve or surface during small intervals of time. So if we look at the motion of a node during a small interval of time, then with probability almost one the motion is similar to Brownian motion on an infinite line or plane with the initial node position as the origin. Thus, in the rest of the paper, we treat the motion of nodes during the time scale corresponding to packet inter-arrival times as motion on a plane or straight line. Over time the nodes do not drift apart from each other, as they would on a infinite line or plane, but just keep moving around on the circle or the torus.

Conversely, this may be viewed as if we are observing the Brownian motion of the nodes on an infinite line or plane and mapping their positions back on the circle or torus, respectively. For example, consider a node that performs Brownian motion along the x-axis and whose initial position is the origin.

At time t, suppose the node is located at $X(t)$. Then it may be mapped to a point $\mod_L(X(t))$ away from the initial position of the node on the circle, with distance measured in counter-clockwise direction. Similar mapping is possible in the case of torus by considering an infinite plane. Thus, instead of keeping track of the positions of nodes on the circle or torus, we use the coordinates of the nodes on x-axis and infinite plane. This scheme works since we are only interested in the change in positions of nodes during packet inter-arrival periods. The coordinates of nodes are denoted by $X_i(t)$. Hence $X_i(t) = \{X_{i1}(t)\}$ and $X_i(t) = \{X_{i1}(t), X_{i2}(t)\}$ for one and two-dimension case respectively. The location information of node i available at the location server at time t is denoted by $\hat{X}_i(t)$, hence $\hat{X}_i(t) = \{\hat{X}_{i1}(t)\}$ and $\hat{X}_i(t) = \{\hat{X}_{i1}(t), \hat{X}_{i2}(t)\}$ for one and two-dimension case respectively.

The j^{th} packet destined to destination i arrives at a node (source of the j^{th} packet) in the network at time $T_i(j)$, $\forall j \geq 1$. Define $T_i(0) \triangleq 0 \;\forall\; 1 \leq i \leq n$. For all $j \geq 1$, define $S_j \triangleq T_i(j) - T_i(j-1)$ as the packet inter-arrival time which is independently and identically distributed (i.i.d.) according to an arbitrary distribution with probability distribution function (pdf) $f_S(t)$. Similarly let $\tau_i(k)$ denote the time at which the k^{th} packet is forwarded by node i, $\tau_i(0) \triangleq 0$. The forwarded packets include both the packets generated by node i and the packets for which the node acts as an intermediate relaying node. The inter-arrival time of the forwarded packets, $\tau_i(k+1) - \tau_i(k) \;\forall\; k > 0$, are i.i.d. with pdf denoted by $f_\tau(t)$.

The communication radius of each node is r meters. Rendezvous-based location service is used for maintaining locations of destinations at a subset of nodes acting as location servers. When a new packet arrives at a source node, it queries the location server of the packet destination for the location of the destination. The packet is routed to the destination according to greedy geographic forwarding using destination location information returned by the location server. It is assumed the position of a destination does not change significantly while the location server is being queried by the source and the packet is being forwarded through the network. In other words the time scale of forwarding a packet is much smaller than that required for a significant change in position. Also the network is assumed to be always connected such that nodes can communicate with the desired location servers.

IV. LOCATION UPDATE OVERHEAD

In this section we evaluate a lower bound on the minimum rate at which a node must transmit its location information such that the average error in its location stored at the location server is less than ϵ whenever the server is queried. We first introduce the notation and rate-distortion formulation, followed by analysis for one-dimensional and two-dimensional networks.

A. Notation and Rate-Distortion Formulation

Definition 1: $D_i(t)$ is the squared-error in the location information of destination i available at its location server at time t, i.e.,

$$D_i(t) = |X_i(t) - \hat{X}_i(t)|^2 \tag{1}$$

where $|X_i(t) - \hat{X}_i(t)| = \sqrt{\sum_{j=1}^{m} \left(X_{ij}(t) - \hat{X}_{ij}(t)\right)^2}$, $m = 1$ for the 1-dimension case, and $m = 2$ for the 2-dimension case.

Definition 2: $X_i^N = \{X_i(T_1), X_i(T_1), \ldots, X_i(T_N)\}$ is the vector of locations of destination i at time instances T_j, $1 \leq j \leq N$. Similarly $\hat{X}_i^N = \{\hat{X}_i(T_1), \hat{X}_i(T_2), \ldots, \hat{X}_i(T_N)\}$ is the vector of location information at the location server of destination i at time instances T_j, $1 \leq j \leq N$.

Definition 3: $\mathcal{X}_i^N$ and $\hat{\mathcal{X}}_i^N$ are defined as sets of all possible vectors X_i^N and $\hat{X}_i^N$, respectively.

Definition 4: $P_N[x_i^N; \hat{x}_i^N]$ denotes the probability that $X_i^N = x_i^N$ and $\hat{X}_i^N = \hat{x}_i^N$, where $x_i^N \in \hat{\mathcal{X}}_i^N$ and $\hat{x}_i^N \in \hat{\mathcal{X}}_i^N$.

Definition 5: $\overline{D}_{iN}$ is defined as

$$\overline{D}_{iN} \triangleq \frac{1}{N} \sum_{j=1}^{N} E[D_i(T_j)] \tag{2}$$

where $E[D_i(T_j)]$ is given by

$$E[D_i(T_j)] = \sum_{x_i^N \in \mathcal{X}_i^N} \sum_{\hat{x}_i^N \in \hat{\mathcal{X}}_i^N} P_N[x_i^N; \hat{x}_i^N] D_i(T_j) \tag{3}$$

Definition 6: $\mathcal{P}_N(\epsilon^2)$ is defined as the family of probability distribution functions $P_N[x_i^N; \hat{x}_i^N]$ for which $\overline{D}_{iN} \leq \epsilon^2$.

Now we will present a rate-distortion theory based formulation to find the minimum number of bits required to represent the location information such that $\overline{D}_{iN} \leq \epsilon^2$.

Definition 7: $R_N(\epsilon^2)$ is defined as the N^{th}-order rate-distortion function – the minimum rate at which a destination must transmit the location information such that the $\overline{D}_{iN} \leq \epsilon$. According to [4], $R_N(\epsilon^2)$ is given by

$$R_N(\epsilon^2) = \min_{P_N \in \mathcal{P}_N(\epsilon^2)} \frac{1}{N} I_{P_N}(X_i^N; \hat{X}_i^N) \tag{4}$$

where $I_{P_N}(X_i^N; \hat{X_i^N})$ is the mutual information between X_i^N and $\hat{X}_i^N$.

The minimum rate at which a destination must update its location information such that a large fraction of packets are delivered, represented by $R(\epsilon^2)$, is given by

$$R(\epsilon^2) = \lim_{N \to \infty} \min R_N(\epsilon^2) \tag{5}$$

B. One-Dimensional Network

In this section we evaluate a lower bound for the minimum update rate for one-dimensional networks.

Lemma 1: The mutual information between X_i^N and $\hat{X}_i^N$ satisfies the following relationship

$$\inf_{P_N \in \mathcal{P}_N} I_{P_N}(X_i^N; \hat{X}_i^N) \geq N R_1(\epsilon^2) \tag{6}$$

The detailed proof of Lemma 1 is available in [15]. From Lemma 1 and the definition of rate distortion function (4) and (5) it follows that

$$R(\epsilon^2) \geq R_1(\epsilon^2) \tag{7}$$

The following theorem provides a lower bound on the minimum rate at which a destination must update its location information.

Theorem 1: In order to ensure high delivery ratio, the lower bound on the location update rate in bits per packet (bpp) is given by

$$R(\epsilon^2) \geq h(X_{i1}(T_1)) - \frac{1}{2}\log 2\pi e\epsilon^2 \quad (8)$$

where $h(X_{i1}(T_1))$ is the differential entropy of the location of destination i at the time when the first packet destined to it arrives in the network.

Proof: From (7) we know that the minimum update rate is bounded by $R_1(\epsilon^2)$, which in turn is defined as

$$R_1(\epsilon^2) = \inf_{P_1 \in \mathcal{P}_1} I_{P_1}(X_{i1}(T_1); \hat{X}_{i1}(T_1))$$

Now consider $I_{P_1}(X_{i1}(T_1); \hat{X}_{i1}(T_1))$,

$$\begin{aligned} I_{P_1}(X_{i1}(T_1); \hat{X}_{i1}(T_1)) &= h(X_{i1}(T_1)) - h(X_{i1}(T_1)|\hat{X}_{i1}(T_1)) \\ &= h(X_{i1}(T_1)) - h(X_{i1}(T_1) - \hat{X}_{i1}(T_1)|\hat{X}_{i1}(T_1)) \quad (9) \\ &\geq h(X_{i1}(T_1)) - h(X_{i1}(T_1) - \hat{X}_{i1}(T_1)) \quad (10) \\ &\geq h(X_{i1}(T_1)) - h(X_{i1}(\mathcal{N}(0, E[(X_{i1}(T_1) - \hat{X}_{i1}(T_1))^2])) \quad (11) \\ &\geq h(X_{i1}(T_1)) - \frac{1}{2}\log\left(2\pi e\epsilon^2\right) \quad (12) \end{aligned}$$

Here (10) follows from (9) since conditioning does not increase entropy. Equation 11 follows from (10) since for a fixed variance, normal distribution has the highest differential entropy. Equation 12 follows from (11) since for $P_1 \in \mathcal{P}_1$ $E[(X_{i1}(T_1) - \hat{X}_{i1}(T_1))^2] \leq \epsilon^2$. Thus

$$R_1(\epsilon^2) \geq h(X_{i1}(T_1)) - \frac{1}{2}\log\left(2\pi e\epsilon^2\right)$$

and (8) follows directly from it. ∎

Theorem 1 implies that the minimum update rate largely depends on $h(X_{i1}(T_1))$, which in turn depends on two factors: (i) the mobility pattern of the destination node and (ii) the packet inter-arrival process. Let $f_X(x)$ denote the pdf of $X_{i1}(T_1)$ (without loss of generality, $X_{i1}(T_0) = 0$). For Brownian motion with variance σ^2 and packet inter-arrival time distribution $f_S(t)$, $f_X(x)$ is given by

$$f_X(x) = \int_{\tau=0}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2\tau}} e^{-\frac{x^2}{2\sigma^2\tau}} f_S(\tau)d\tau \quad (13)$$

Thus $h(X_{i1}(T_1))$ is given by

$$h(X_{i1}(T_1)) = -\int_{x=-\infty}^{\infty} f_X(x)\log\left(f_X(x)\right)dx \quad (14)$$

The lower bound on the minimum overhead incurred by location update information in bits per second (bps), denoted by $U(\epsilon^2)$, is given by

$$U(\epsilon^2) \geq \frac{1}{E[S]}\left(h(X_{i1}(T_1)) - \frac{1}{2}\log\left(2\pi e\epsilon^2\right)\right) \text{ bps} \quad (15)$$

C. Two-Dimensional Network

In this section we present the update rate analysis for two-dimensional networks, which is based on the analysis for one-dimensional case. We also evaluate the lower bound for various packet arrival processes and discuss the effect of arrival processes on the minimum update rate.

Brownian motion in two-dimensional space may be decomposed into two independent one-dimensional Brownian motions along x and y coordinates each with a variance $\sigma^2/2$. Thus if $X_i(t) = \{X_{i1}(t), X_{i2}(t)\}$ denote the coordinates of destination i at time t, then the distribution of $X_{i1}(t)$ is independent of the distribution of $X_{i2}(t)$. The following Lemma expresses $R\left(\epsilon^2\right)$ in terms of components corresponding to the two coordinates.

Lemma 2: For two-dimensional networks, the rate distortion $R\left(\epsilon^2\right)$ function may be written as

$$R(\epsilon^2) = \min_{0\leq k\leq\epsilon} R^{(1)}(k^2) + R^{(2)}(\epsilon^2 - k^2) \quad (16)$$

where

$$R^{(1)}(\epsilon^2) = \lim_{N\to\infty} \inf_{P_N \in \mathcal{P}_N(k^2)} \frac{1}{N} I_{P_N}(X_{i1}^N; \hat{X}_{i1}^N) \quad (17)$$

$$R^{(2)}(\epsilon^2) = \lim_{N\to\infty} \inf_{P_N \in \mathcal{P}_N(\epsilon^2 - k^2)} \frac{1}{N} I_{P_N}(X_{i2}^N; \hat{X}_{i2}^N) \quad (18)$$

Proof: Recall the rate distortion function is

$$R(\epsilon^2) = \lim_{N\to\infty} \inf_{P_N \in \mathcal{P}_N(\epsilon^2)} \frac{1}{N} I_{P_N}(X_i^N; \hat{X}_i^N)$$

Now consider $I_{P_N}(X_i^N; \hat{X}_i^N)$. For two dimensional networks, this may be written as

$$\begin{aligned} I_{P_N}(X_i^N; \hat{X}_i^N) &= I_{P_N}(X_{i1}^N, X_{i2}^N; \hat{X}_i^N) \\ &= I_{P_N}(X_{i1}^N; \hat{X}_i^N) + I_{P_N}(X_{i2}^N; \hat{X}_i^N|X_{i1}^N) \\ &= I_{P_N}(X_{i1}^N; \hat{X}_i^N) + I_{P_N}(X_{i2}^N; \hat{X}_i^N) \\ &= I_{P_N}(X_{i1}^N; \hat{X}_{i1}^N) + I_{P_N}(X_{i1}^N; \hat{X}_{i2}^N|\hat{X}_{i1}^N) + \\ &\quad I_{P_N}(X_{i2}^N; \hat{X}_{i2}^N) + I_{P_N}(X_{i2}^N; \hat{X}_{i1}^N|\hat{X}_{i2}^N) \\ &= I_{P_N}(X_{i1}^N; \hat{X}_{i1}^N) + I_{P_N}(X_{i2}^N; \hat{X}_{i2}^N) \quad (19) \end{aligned}$$

where $X_{i1}^N = \{X_{i1}(T_1), \ldots, X_{i1}(T_N)\}$, $X_{i2}^N = \{X_{i2}(T_1), \ldots, X_{i2}(T_N)\}$, $\hat{X}_{i1}^N = \{\hat{X}_{i1}(T_1), \ldots, \hat{X}_{i1}(T_N)\}$ and $\hat{X}_{i2}^N = \{\hat{X}_{i2}(T_1), \ldots, \hat{X}_{i2}(T_N)\}$.

We know that $\overline{D}_{iN} \leq \epsilon^2$ implies

$$\frac{1}{N}\sum_{j=1}^{N} E\left[(X_{i1}(T_j) - \hat{X}_{i1}(T_j))^2\right] + \frac{1}{N}\sum_{j=1}^{N} E\left[(X_{i2}(T_j) - \hat{X}_{i2}(T_j))^2\right] \leq \epsilon^2$$

The distortion constraint is satisfied if $\frac{1}{N}\sum_{j=1}^{N} E\left[(X_{i1}(T_j) - \hat{X}_{i1}(T_j))^2\right] \leq k^2$ and $\frac{1}{N}\sum_{j=1}^{N} E\left[(X_{i2}(T_j) - \hat{X}_{i2}(T_j))^2\right] \leq \epsilon^2 - k^2$. Combining this and (19), we get

$$\begin{aligned} R(\epsilon^2) = \min_{0\leq k\leq\epsilon} \lim_{N\to\infty} &\inf_{P_N \in \mathcal{P}_N(k^2)} \frac{1}{N} I_{P_N}(X_{i1}^N; \hat{X}_{i1}^N) + \\ &\inf_{P_N \in \mathcal{P}_N(\epsilon^2 - k^2)} \frac{1}{N} I_{P_N}(X_{i2}^N; \hat{X}_{i2}^N) \quad (20) \end{aligned}$$

which leads to (16). ∎

From Lemma 1 and Theorem 1, it follows that

$$R^{(1)}(k^2) \geq h(X_{i1}(T_1)) - \log\left(2\pi e k^2\right) \quad (21)$$

$$R^{(2)}(\epsilon^2 - k^2) \geq h(X_{i2}(T_1)) - \log\left(2\pi e(\epsilon^2 - k^2)\right) \quad (22)$$

From (21) and (22), it is clear that the right hand side (RHS) of (16) is minimized for $k^2 = \epsilon^2/2$. This leads to the following theorem.

Theorem 2: In order to ensure that the average error in location information used for forwarding packets is less than ϵ, the lower bound on the location update rate (in bits per packet) for a two-dimensional network is given by

$$R(\epsilon^2) \geq h(X_{i1}(T_1)) + h(X_{i2}(T_1)) - \log\left(\pi e \epsilon^2\right) \text{ bpp} \quad (23)$$

and the overhead incurred in bits/sec ($U(\epsilon)$) is given by

$$U(\epsilon^2) \geq \frac{1}{E[S]}\left(h(X_{i1}(T_1)) + h(X_{i2}(T_1)) - \log\left(\pi e \epsilon^2\right)\right) \text{ bps} \quad (24)$$

We now derive the lower bounds for deterministic, uniformly distributed and exponentially distributed inter-arrival times.

1) Deterministic packet arrival: For the deterministic packet arrival process, where packets arrive at $t = kT$, $k = 1, 2, \ldots \infty$, the probability distribution functions of $X_{i1}(T_1)$ and $X_{i2}(T_1)$ are given by

$$f_{X_1}(x) = f_{X_2}(x) = \frac{1}{\sqrt{\pi\sigma^2 T}} e^{-\frac{x^2}{\sigma^2 T}} \quad (25)$$

and $h(X_{i1}(T_1))$ and $h(X_{i2}(T_1))$ is given by

$$h(X_{i1}(T_1)) = h(X_{i2}(T_1)) = \frac{1}{2}\log\left(\pi e \sigma^2 T\right)$$

Thus the lower bounds on location update rate in bits/packet and bits/second are given by

$$R(\epsilon^2) \geq \log\left(\frac{\sigma^2 T}{\epsilon^2}\right) \text{ bpp}, \quad U(\epsilon^2) \geq \frac{1}{T}\log\left(\frac{\sigma^2 T}{\epsilon^2}\right) \text{ bps}$$

2) Uniform distribution of packet inter-arrival time: For the uniform packet arrival process the pdf of packet inter-arrival time, $f_S(t)$ is given by

$$f_S(t) = \begin{cases} 1/T, & 0 \leq t \leq T \\ 0, & \text{otherwise} \end{cases} \quad (26)$$

The probability distribution functions of $X_{1i}(T_1)$ and $X_{2i}(T_1)$ are given by

$$f_{X_1}(x) = f_{X_2}(x) = \frac{1}{T}\int_0^T \frac{1}{\sqrt{\pi\sigma^2 t}} e^{-\frac{x^2}{\sigma^2 t}} \mathrm{d}t$$
$$= \sqrt{\frac{4}{\pi\sigma^2 T}} e^{-\frac{x^2}{\sigma^2 T}} + \frac{2x}{\sigma^2 T}\mathrm{erf}\left(\frac{x}{\sqrt{\sigma^2 T}}\right) - \frac{2|x|}{\sigma^2 T} \quad (27)$$

Let h_U denote the differential entropy of $X_{1i}(T_1)$ and $X_{2i}(T_1)$, i.e., $h_U \triangleq h(X_{i1}(T_1)) = h(X_{i2}(T_1))$, then the lower bound on update rate is given by

$$R(\epsilon^2) \geq 2h_U - \log\left(\pi e \epsilon^2\right), \; U(\epsilon^2) \geq \frac{2}{T}\left(2h_U - \log\left(\pi e \epsilon^2\right)\right)$$

3) Exponential distribution of packet inter-arrival time: For the exponential packet arrival process, the probability distribution functions of $X_{1i}(T_1)$ and $X_{2i}(T_1)$ are given by

$$f_{X_1}(x) = f_{X_2}(x) = \int_0^\infty \frac{\alpha}{\sqrt{2\pi\sigma^2\tau}} e^{-\left(\frac{x^2}{2\sigma^2\tau} + \alpha\tau\right)} d\tau \quad (28)$$

It is not possible to find a closed form expression for the integral in the previous equation. Therefore numerical methods may be applied to evaluate $f_{X_1}(x)$ and $f_{X_2}(x)$. Let h_E denote the differential entropy of $X_{1i}(T_1)$ and $X_{2i}(T_1)$, i.e., $h_E \triangleq h(X_{i1}(T_1)) = h(X_{i2}(T_1))$, then the lower bound on update rate is given by

$$R(\epsilon^2) \geq 2h_E - \log\left(\pi e \epsilon^2\right), \; U(\epsilon^2) \geq \alpha\left(2h_E - \log\left(\pi e \epsilon^2\right)\right)$$

4) Comparison of update rates for various inter-arrival processes: Figure 1 shows the plot of lower bound on $U(\epsilon^2)$ against σ^2 and $E[S]$. It is observed that for high σ^2 and low $E[S]$, the rate at which a source must update its location servers becomes very high. Also it is observed that the rate required for deterministic packet arrival is higher than that required for uniform and exponential arrival processes. In fact the update rate for deterministic packet arrival process is higher than any other packet arrival process with the same mean inter-arrival time. Consider a packet arrival process with pdf $f_S(t)$ and mean $E[S]$, then $\mathrm{Var}\left(X_{i1}(T_1)\right)$ and $\mathrm{Var}\left(X_{i2}(T_1)\right)$ are given by

$$\int_0^\infty \int_{-\infty}^\infty x^2 \frac{1}{\sqrt{\pi\sigma^2 t}} e^{-\frac{x^2}{\sigma^2 t}} \mathrm{d}x f_S(t) \mathrm{d}t = \int_0^\infty \frac{\sigma^2 t}{2} f_S(t) \mathrm{d}t = \frac{\sigma^2 E[S]}{2}$$

This implies that, notwithstanding the packet arrival process, the variance of the change in location between two packet arrival instances depends only on σ and $E[S]$. For the deterministic packet arrival process, $X_{i1}(T_1)$ and $X_{i2}(T_1)$ are Gaussian random variables (25). Since, among random variables with the same variance, a Gaussian random variable has the highest entropy, deterministic packet arrival leads to the highest update rate.

V. BEACON OVERHEAD

In this section we evaluate a lower bound on the minimum rate at which nodes must transmit beacons such that consistent neighborhood information may be maintained at the neighbors. We first introduce the notations and the rate-distortion formulation for the minimum beacon rate problem.

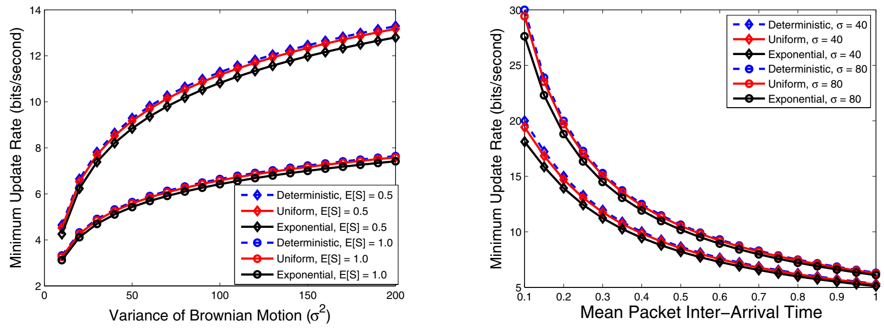
A. Notation and Minimum Beacon Rate Formulation

Definition 8: $N_i(t)$ is the set of nodes that belong to the neighborhood of node i. That is,

$$N_i(t) = \{j : |X_i(t) - X_j(t)| \leq r, 1 \leq j \leq n, j \neq i\} \quad (29)$$

Definition 9: $\hat{N}_i(t)$ is the set of nodes that the node i *perceives* to be its neighbors.

The set $\hat{N}_i(t)$ is constructed by node i based on the beacons it receives. A node that belongs to $\hat{N}_i(t)$ may be excluded from $\hat{N}_i(t+\tau)$ if sufficient beacons are not received from the node during time interval $[t, t+\tau]$. Similarly, a node not belonging to $\hat{N}_i(t)$ may be included in $\hat{N}_i(t+\tau)$ if sufficient beacons are received from the node during time interval $[t, t+\tau]$. The deviation of $\hat{N}_i(t)$ from $N_i(t)$ depends on the rate at which the nodes transmit beacons.



(a) Update rate vs variance of Brownian motion. (b) Update rate vs. avg packet inter-arrival time.

Fig. 1. The minimum update rate increases with increase in σ and decrease in inter-arrival time.

Definition 10: $Z_{ij}(t)$ and $\hat{Z}_{ij}(t)$ ($1 \le i, j \le n$, $i \ne j$) are indicator random variables, defined in the following manner

$$Z_{ij}(t) = \begin{cases} 1, & \text{if } j \in N_i(t) \\ 0, & \text{otherwise} \end{cases} \tag{30}$$

$$\hat{Z}_{ij}(t) = \begin{cases} 1, & \text{if } j \in \hat{N}_i(t) \\ 0, & \text{otherwise} \end{cases} \tag{31}$$

In other words, $Z_{ij}(t)$ equals 1 if node j belongs to the neighborhood of node i at time t. Note that $Z_{ij}(t)$ is a symmetric relation, i.e. $Z_{ij}(t) = Z_{ji}(t)$. On the other hand, $\hat{Z}_{ij}(t)$ is 1 if node i perceives node j to be its neighbor at time t. Unlike $Z_{ij}(t)$, $\hat{Z}_{ij}(t)$ is not symmetric. That is $\hat{Z}_{ij}(t)$ may be 0 although $\hat{Z}_{ji}(t)$ is 1. This may happen because beacon transmission rate of i is high enough to allow j to maintain consistent neighborhood set while beacon transmission rate of node j is not high enough to allow node i to maintain consistent neighborhood set. The deviation of the perceived neighborhood, $\hat{N}_i(t)$, from the actual neighborhood, $N_i(t)$, is reflected by the deviation of $\hat{Z}_{ij}(t)$ from $Z_{ij}(t)$.

Definition 11: The difference of $\hat{Z}_{ij}(t)$ and $Z_{ij}(t)$ is defined as $E_{ij}(t)$, i.e.,

$$E_{ij}(t) = Z_{ij}(t) - \hat{Z}_{ij}(t) \tag{32}$$

$E_{ij}(t) = 0$ implies that node i has accurate information about whether j belongs to its neighborhood or not. It is not necessary that $E_{ij}(t) = 0$ for all t, however it is desirable that $E_{ij}(t) = 0$ with high probability at all time instances when node i has a packet to forward. This is because correct neighborhood information is highly critical for node i to make correct forwarding decisions.

We can now state the minimum beacon rate problem in the following manner.

Minimum beacon rate problem: Find the minimum rate at which node j must transmit beacons such that

$$P[E_{ij}(\tau_i(k)) = 0] \ge 1 - \delta \;\; \forall\; 1 \le i \ne j \le n, k > 0 \tag{33}$$

In order to formulate the above minimum beacon rate problem as a rate distortion problem we present two more definitions.

Definition 12: Let the vectors Z_{ij}^N and $\hat{Z}_{ij}^N$ be defined in the following manner

$$Z_{ij}^N \triangleq \{Z_{ij}(\tau_i(1)), Z_{ij}(\tau_i(2)), \ldots, Z_{ij}(\tau_i(N))\} \tag{34}$$
$$\hat{Z}_{ij}^N \triangleq \{\hat{Z}_{ij}(\tau_i(1)), \hat{Z}_{ij}(\tau_i(2)), \ldots, \hat{Z}_{ij}(\tau_i(N))\} \tag{35}$$

Definition 13: Let $\mathcal{P}_N^{(b)}(\delta)$ denote the family of joint probability distribution functions of Z_{ij}^N and $\hat{Z}_{ij}^N$ such that $P[E_{ij}(\tau_i(k)) = 0] \le 1 - \delta \;\forall\; 1 \le k \le N$.

The superscript in $\mathcal{P}_N^{(b)}(\delta)$ is used in order to distinguish the notation from the one used in the previous section. This superscript will be used for similar purpose in the rest of this section.

Thus the minimum beacon rate, $R^{(b)}(\delta)$, may be expressed in the following manner

$$R^{(b)}(\delta) = \lim_{N\to\infty} \min R_N^{(b)}(\delta) \tag{36}$$

where

$$R_N^{(b)}(\delta) = \min_{P_N \in \mathcal{P}_N^{(b)}(\delta)} \frac{1}{N} I_{P_N}(Z_{ij}^N; \hat{Z}_{ij}^N) \tag{37}$$

and $I_{P_N}(Z_{ij}^N; \hat{Z}_{ij}^N)$ is mutual information between Z_{ij}^N and $\hat{Z}_{ij}^N$.

B. Beacon Rate Analysis for One-Dimensional Networks

In this section we evaluate a lower bound on the minimum beacon rate for one dimensional networks.

Lemma 3: The minimum beacon rate of node j, $R^{(b)}(\delta)$ is greater than equal to $R_1^{(b)}(\delta)$, that is

$$R^{(b)}(\delta) \ge R_1^{(b)}(\delta) \tag{38}$$

The proof of the above Lemma is similar to that of Lemma 1 and is presented in [15]. The next Lemma provides a lower bound on $R_1^{(b)}(\delta)$.

Lemma 4: $R_1^{(b)}(\delta)$ satisfies the following relationship

$$R_1^{(b)}(\delta) \ge H(Z_{ij}(\tau_i(1))) - \mathcal{H}\left(\frac{\delta}{2}, 1-\delta, \frac{\delta}{2}\right) \tag{39}$$

where,

$$\mathcal{H}\left(\frac{\delta}{2}, \delta, \frac{\delta}{2}\right) \triangleq -\delta \log\left(\frac{\delta}{2}\right) - (1-\delta)\log(1-\delta) \tag{40}$$

Proof: Recall that $R_1^{(b)}(\delta)$ is given by

$$R_1^{(b)}(\delta) = \inf_{P_1 \in \mathcal{P}_1(\delta)} I_{P_1}(Z_{ij}(\tau_1); \hat{Z}_{ij}(\tau_1)) \quad (41)$$

Now $I_{P_1}(Z_{ij}(\tau_1); \hat{Z}_{ij}(\tau_1))$ is given by

$$\begin{aligned} I_{P_1}(Z_{ij}(\tau_1); \hat{Z}_{ij}(\tau_1)) &= H(Z_{ij}(\tau_1)) - H(Z_{ij}(\tau_1)|\hat{Z}_{ij}(\tau_1)) &(42)\\ &= H(Z_{ij}(\tau_1)) - H(Z_{ij}(\tau_1) - \hat{Z}_{ij}(\tau_1)|\hat{Z}_{ij}(\tau_1)) &(43)\\ &\geq H(Z_{ij}(\tau_1)) - H(Z_{ij}(\tau_1) - \hat{Z}_{ij}(\tau_1)) &(44)\\ &= H(Z_{ij}(\tau_1)) - H(E_{ij}(\tau_1)) &(45)\end{aligned}$$

We know that the probability distribution of $E_{ij}(\tau_i(1))$ is given by

$$E_{ij}(\tau_i(1)) = \begin{cases} -1, & \text{w.p. } p_1 \\ 0, & \text{w.p. } p_2 \\ 1, & \text{w.p. } p_3 \end{cases} \quad (46)$$

where $p_2 \geq 1 - \delta$, $p_1 + p_3 \leq \delta$ and $p_1 + p_2 + p_3 = 1$ (since $P_1 \in \mathcal{P}_1(\delta)$). Under these constraints $H(E_{ij}(\tau_i(1))))$ is maximized when $p_2 = 1 - \delta$ and $p_1 = p_3 = \delta/2$, when $H(E_{ij}(\tau_1)) = \mathcal{H}\left(\frac{\delta}{2}, 1-\delta, \frac{\delta}{2}\right)$. This leads to (39). ■

Lemmas 3 and 4 imply that lower bound on $R^{(b)}(\delta)$ depends on $H(Z_{ij}(\tau_i(1)))$, which in turn depends on $X_j(0)$ and $f_\tau(t)$. Without loss of generality we assume that $X_i(0) = 0$ and consider two separate cases: (i) $|X_j(0)| \leq r$, i.e., $Z_{ij}(0) = 1$, and (ii) $|X_j(0)| > r$, i.e., $Z_{ij}(0) = 0$.

Suppose $Z_{ij}(0) = 1$. In this case let $X_j(0) = l$ where $-r \leq l \leq r$. From the point of reference of node i, node j performs Brownian motion with variance $2\sigma^2$. The probability that $Z_{ij}(\tau_1) = 1$ is given by

$$p(l) \triangleq P[Z_{ij}(\tau_i(1)) = 1 | X_j(0) = l, |l| \leq r] = \frac{1}{2}\int_{t=0}^{\infty} \text{erf}\left(\frac{r-l}{\sqrt{4\sigma^2 t}}\right) f_\tau(t)dt + \frac{1}{2}\int_{t=0}^{\infty} \text{erf}\left(\frac{r+l}{\sqrt{4\sigma^2 t}}\right) f_\tau(t)dt$$

We know that $H(Z_{ij}(\tau_i(1)))$ is equal to $\mathcal{H}(p(l))$, where $\mathcal{H}(x) = -x\log(x) - (1-x)\log(1-x)$. Depending on σ and $f_\tau(t)$, $\mathcal{H}(p(l))$ is maximized for some l. In order to ensure that $P[E_{ij}(\tau_i(k)) = 0] \leq \delta \; \forall \; i$, the beacon rate must take care of this worst possible case. Also we know that $\mathcal{H}(x)$ is symmetric at $x = 0.5$, symmetric about $x = 0.5$ and is strictly increasing and decreasing in intervals $[0, 0.5)$ and $(0.5, 1]$ respectively. Thus when $Z_{ij}(0) = 1$

$$R^{(b)}(\delta) \geq p(l^\star) - \mathcal{H}\left(\frac{\delta}{2}, 1-\delta, \frac{\delta}{2}\right) \quad (47)$$

where

$$l^\star \triangleq \arg\min_{-r \leq l \leq r} |p(l) - 0.5| \quad (48)$$

That is, $l^\star$ is the value of l that maximizes $\mathcal{H}(p(l))$.

Now consider the case where $Z_{ij}(0) = 0$, i.e., $X_j(0) = l$, $|l| \geq r$. The probability that $Z_{ij}(\tau_1) = 1$ is given by

$$p'(l) \triangleq P[Z_{ij}(\tau_i(1)) = 1 | X_j(0) = l, |l| \geq r] = \frac{1}{2}\int_{0}^{\infty} \text{erf}\left(\frac{|l|+r}{\sqrt{4\sigma^2 t}}\right) f_\tau(t)dt - \frac{1}{2}\int_{0}^{\infty} \text{erf}\left(\frac{|l|-r}{\sqrt{4\sigma^2 t}}\right) f_\tau(t)dt$$

Note that $p'(l) \leq 0.5 \; \forall \; |l| \geq r$ and its value is maximized at $|l| = r$. Since $\mathcal{H}(x)$ is an increasing function of x in the interval $[0, 0.5)$, $|l| = r$ maximizes the value of $\mathcal{H}(p'(l))$. However $p'(r) = p(r)$ and therefore the beacon transmission rate required for the first case is always higher than the second case. This leads to the following Theorem.

Theorem 3: The lower bound on the minimum beacon transmission rate of a node such that the constraint in (33) is satisfied is given

$$R^{(b)}(\delta) \geq \mathcal{H}(p(l^\star)) - \mathcal{H}\left(\frac{\delta}{2}, 1-\delta, \frac{\delta}{2}\right) \text{ beacons/pkt} \quad (49)$$

where $l^\star$ is given by (48).

The minimum overhead in bits per second (bps), denoted by $U^{(b)}(\delta)$, is given by

$$U^{(b)}(\delta) \geq \frac{B}{E[\tau]}\left(\mathcal{H}(p(l^\star)) - \mathcal{H}\left(\frac{\delta}{2}, 1-\delta, \frac{\delta}{2}\right)\right) \text{ bps} \quad (50)$$

where $E[\tau]$ is the expected packet inter-arrival time and B is the size of beacon packet in bits.

C. Beacon Rate Analysis for Two-Dimensional Networks

In this subsection we extend the minimum beacon rate analysis to two dimensional networks. For a arbitrary node pair i and j, we choose an orthogonal coordinate system such that $X_{i1}(0) = X_{i2}(0) = 0$, $X_{j1}(0) = l$, and $X_{j2}(0) = 0$. That is, the origin of the coordinate system corresponds to the position of node i at $t = 0$ and the x-axis of the coordinate system corresponds to the line joining the position of nodes i and j at $t = 0$. It can be easily verified that a Brownian motion with variance σ^2 can be decomposed into two independent Brownian motions with variance $\sigma^2/2$ along each axis. Also note that Lemmas 3 and 4 hold for the two dimensional case as well and may be proved in a similar manner. Thus the minimum beacon rate, $R^{(b)}(\delta)$, satisfies the following relationship

$$R^{(b)}(\delta) \geq H\left(Z_{ij}(\tau_i(1))\right) - \mathcal{H}\left(\frac{\delta}{2}, 1-\delta, \frac{\delta}{2}\right) \quad (51)$$

Similar to the approach in the last section, we proceed by individually considering the cases $Z_{ij}(0) = 1$ and $Z_{ij}(0) = 0$.

For the case when $Z_{ij}(0) = 1$, the probability that j is in the neighborhood of i when i has a packet to send ($p(l)$) is given by

$$p(l) \triangleq P\left[Z_{ij}(\tau) = 1 | X_{j1}(0) = l, X_{j2}(0) = 0, |l| \leq r\right] = \int_{x=-r}^{x=r} P\left[X_{j1}(\tau_i(1)) = x | X_{j1}(0) = l\right] \cdot P\left[-\sqrt{r^2 - x^2} \leq X_{j2}(\tau_i(1)) \leq \sqrt{r^2 - x^2} | X_{j2}(0) = 0\right]$$

Relative to node i, node j performs Brownian motion with variance $2\sigma^2$. Thus

$$P\left[X_{j1}(\tau) = x | X_{j1}(0) = l\right] = \frac{1}{\sqrt{2\pi\sigma^2\tau}} \exp\left(-\frac{(l-x)^2}{2\sigma^2\tau}\right) dx$$

and

$$P\left[-\sqrt{r^2-x^2} \leq X_{j2}(\tau) \leq \sqrt{r^2-x^2}|X_{j2}(0)=0\right] = \operatorname{erf}\left(\frac{\sqrt{r^2-x^2}}{\sqrt{2\sigma^2\tau}}\right)$$

Therefore $p(l)$ is given by

$$p(l) = \int_0^\infty \int_{-r}^{r} \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left(-\frac{(l-x)^2}{2\sigma^2 t}\right) \operatorname{erf}\left(\frac{\sqrt{r^2-x^2}}{\sqrt{2\sigma^2 t}}\right) f_\tau(t) dx\, dt \quad (52)$$

Thus in order to satisfy (33) at all neighbors that are neighbor at time 0, node j must transmit beacon at a rate higher than

$$\mathcal{H}\left(p(l^\star)\right) - \mathcal{H}\left(\frac{\delta}{2}, 1-\delta, \frac{\delta}{2}\right) \quad (53)$$

where $l^\star$ is given by (48).

Now consider the case when j does not belong to the neighborhood of i at $t=0$. It can be easily verified that the probability that $Z_{ij}(\tau_i(1)) = 1$ given that $Z_{ij}(0) = 0$, denoted by$p'(l)$, is given by the same expression as $p(l)$ (equation 52). $p'(l)$ increases with decrease in $|l|$ and is maximized for $|l| = r$. For other values of $l > r$, $p'(l) < 0.5$. Thus similar to the one-dimensional networks, the beacon transmission rate is determined by the rate required to satisfy (33) at the initial neighbors. This leads the following theorem.

Theorem 4: The lower bound of the minimum beacon transmission rate of a node such that the constraint in equation 33 is satisfied is given

$$R^{(b)}(\delta) \geq \mathcal{H}\left(p(l^\star)\right) - \mathcal{H}\left(\frac{\delta}{2}, 1-\delta, \frac{\delta}{2}\right) \text{ beacons/pkt} \quad (54)$$

where $p(l)$ and $l^\star$ are given by (52) and (48) respectively. The beacon transmission overhead in bits per second, $U^{(b)}(\delta)$, is given by

$$U^{(b)}(\delta) \geq \frac{B}{E[\tau]}\left(\mathcal{H}(p(l^\star)) - \mathcal{H}\left(\frac{\delta}{2}, 1-\delta, \frac{\delta}{2}\right)\right) \text{ bps} \quad (55)$$

D. Comparison of Beacon Transmission Rates for Various Arrival Processes

The closed form expression for the integral in (52) cannot be found. So we use numerical computations to evaluate $R^{(b)}(\delta)$ for deterministic, uniform and exponential packet arrival processes. Figure 2 shows plots of minimum beacon rate in bits per second. Figure 2(a) shows the plot of minimum beacon transmission rate against variance of Brownian motion for different mean packet inter-arrival times. It is observed that for low variance the rate is almost constant, while as the variance increases the rate starts decreasing. When the variance of Brownian motion is very small, the variance of the change in position of a node within a packet arrival epoch is also small. For this case, $l^\star = r$ and $p(l^\star) \approx 0.5$ which leads to high beacon rate. As the variance increases, the probability that two neighbors remain neighbors at the end of a packet arrival epoch is very small, no matter what the initial position of nodes might be. That is, when σ^2 is high, $p(l) < 0.5$ $\forall\ l$, which leads to low beacon rate when variance is high. This implies that when nodes are highly mobile they need to transmit beacons less frequently and the membership of nodes in a neighborhood may be more efficiently deciphered by the absence of beacons. Figure 2(b) shows that as the rate of packet arrival increases, the beacon overhead may become prohibitively high. Also, for a given packet arrival rate, it is observed that the rate for a deterministic packet arrival process is smaller than that for exponential and uniform arrivals. This is because the probability that a node leaves the neighborhood of a certain neighbor within a packet inter-arrival duration is the highest ($p(l)$ is close to 1) for deterministic arrival. For the uniform and exponential distributions the probability that packet inter-arrival time is less than the mean inter-arrival time is 0.5 and 0.63 respectively. Thus the probability that a node moves out of neighborhood during an inter-arrival duration is smaller than that for the deterministic arrival process.

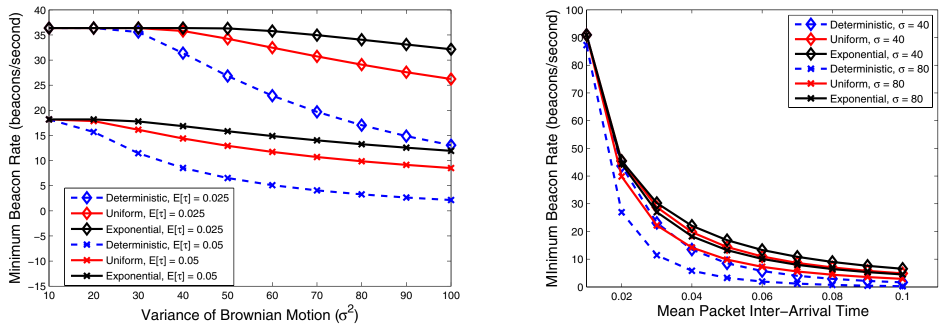
VI. CAPACITY DEFICIT

A wireless ad hoc network is said to transport one bit-meter when a bit is transmitted over a distance of one meter [5]. The transport capacity of a network (in bit-meters per second) is defined as the supremum over the set of feasible rate vectors of the distance weighted sum of rates [16]. The transport capacity is expressed as $\lambda n\overline{L}$, where λ is the average arrival rate at the nodes, n is the number of nodes and $\overline{L}$ is the average distance traveled by the bits. It is shown in [5] that the transport capacity of an arbitrary wireless network is $\Theta\left(W\sqrt{nA}\right)$ where n, W and A are the number of nodes deployed, transmission rate of the nodes and area over which the network is deployed respectively. It is shown in [5] that for a particular interference model known as the *Protocol Model*, the upper bound on the transport capacity of an arbitrary wireless network is given by

$$\lambda n\overline{L} \leq \frac{\sqrt{8}}{\pi}\frac{1}{\Delta}W\sqrt{nA} \text{ bit} - \text{meters/second} \quad (56)$$

Let η denote the expected distance between a node and its location server. Thus, on average, the location update information of a node travels at least η meters before reaching its location server. Thus the average overhead incurred by a node for updating its location information is at least $\eta U(\epsilon^2)$ bit-meters/second and the overhead incurred by location update information on the network equals at least $n\eta U(\epsilon^2)$ bit-meters/second, where $U(\epsilon^2)$ is given by (24). A beacon transmitted by a node travels a distance equal to the communication radius. Thus the overhead incurred by the beacon packets on the network is at least $nrU^{(b)}(\delta)$, where $U^{(b)}(\delta)$ is given by (55). Thus the total transport capacity deficit due to the routing overhead is at least $n\eta U(\epsilon^2) + nrU^{(b)}(\delta)$ bit-meters/second. This leads to the following theorem.

Theorem 5: For the Protocol Model, the upper bound on the *residual transport capacity*, in bit-meters per second, available to an arbitrary network for transmitting data ($\lambda n\overline{L}$) is given



(a) Beacons per second versus variance of Brownian motion.

(b) Beacons per second versus the mean inter-arrival time of packets to be forwarded.

Fig. 2. For moderate mobility rates and high packet arrival rates, the minimum beacon transmission rate may be prohibitively high.

by

$$\lambda n\overline{L} \;\leq\; \frac{\sqrt{8}}{\pi}\frac{1}{\Delta}W\sqrt{nA} \;-\; n\eta U(\epsilon^2) \;-\; nrU^{(b)}(\delta) \quad (57)$$

Theorem 5 has interesting implications. The raw transport capacity of a wireless network scales as $\sqrt{n}$ while the overhead incurred by the routing overheads scales as n. Therefore if the number of nodes deployed in a network increases beyond a certain threshold, denoted by $n^\star$, then no useful information may be transported in the network and the whole capacity is used up by the geographic routing overheads. This leads to the following corollary.

Corollary 1: For geographic routing, the upper bound on the maximum number of nodes that may be deployed in a network while ensuring that it has non-zero residual transport capacity is given by

$$n^\star \leq \left(\frac{\frac{\sqrt{8}}{\pi}\frac{1}{\Delta}W\sqrt{A}}{\eta U(\epsilon^2) + rU^{(b)}(\delta)}\right)^2 \quad (58)$$

Proof: If the residual transport capacity is greater than zero then

$$\frac{\sqrt{8}}{\pi}\frac{1}{\Delta}W\sqrt{nA} - n\eta U(\epsilon^2) - nrU^{(b)}(\delta) \geq 0$$

which implies that

$$\sqrt{n}\left(\eta U(\epsilon^2) + rU^{(b)}(\delta)\right) - \frac{\sqrt{8}}{\pi}\frac{1}{\Delta}W\sqrt{A} \leq 0$$

Rearranging the above equation yields (58). ∎

VII. Conclusion and Future Work

In this paper we presented an information theoretic framework for analyzing the overhead incurred by geographic routing protocols in order to maintain *reliable state information* in a mobile network. We formulated the minimum routing overhead problem as a rate-distortion problem. We evaluated a lower bound on the minimum routing overheads in terms of the node mobility pattern and packet arrival process. We also investigate the effect of routing overheads on the residual capacity available to network users. The extension of this information theoretic framework to analyze other routing paradigms and development of constructive schemes that achieve the lower limit indicated by the analytical results are the focus of future work.

References

[1] B. Karp and H. T. Kung, "GPSR: Greedy perimeter stateless routing for wireless networks." in *MOBICOM '00*, 2000, pp. 243–254.
[2] P. Bose, P. Morin, I. Stojmenovic, and J. Urrutia, "Routing with guaranteed delivery in ad hoc wireless networks," *ACM Wireless Networks*, vol. 7, no. 6, pp. 609–616, 2001.
[3] R. C. Shah, A. Wolisz, and J. M. Rabaey, "On the performance of geographical routing in the presence of localization errors," in *Proc. of ICC'05*, 2005, pp. 2979–2985.
[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & sons, 1991.
[5] P. Gupta and P. R. Kumar, "Capacity of wireless networks," *IEEE Trans. on Information Theory*, pp. 388–404, March 2000.
[6] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 477–486, 2002.
[7] G. Sharma, R. Mazumdar, and N. Shroff, "Delay and capacity trade-offs in mobile ad hoc networks: A global perspective," in *Proc. of IEEE INFOCOM'06*, 2006.
[8] A. Ephremides and B. E. Hajek, "Information theory and communication networks: An unconsummated union." *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2416–2434, 1998.
[9] R. Gallager, "Basic limit on protocol information in data communication networks," *IEEE Transactions on Information Theory*, vol. 22, no. 4, pp. 385–398, July 1976.
[10] C. Rose and R. Yates, "Location uncertainity in mobile networks: A theoretical framework." *IEEE Comm. Magazine*, vol. 35, no. 2, pp. 94–101, 1997.
[11] A. Bhattacharya and S. K. Das, "LeZi-update: An information-theoretic approach to track mobile users in PCS networks," in *MobiCom '99*, 1999, pp. 1–12.
[12] B. An and S. Papavassiliou, "An entropy-based model for supporting and evaluating route stability in mobile ad hoc wireless networks." *IEEE Communications Letters*, vol. 6, no. 8, pp. 328–330, 2002.
[13] Q.-M. Tran, A. Dadej, and S. Parreau, "Characterizing mobility in ad hoc networks: A generalized approach," in *Proc. of ASWN'05*.
[14] N. Zhou and A. Abouzeid, "Routing in ad-hoc networks: A theoretical framework with practical implications," in *Proceedings of INFOCOM'05*, 2005, pp. 1240–1251.
[15] N. Bisnik and A. Abouzeid, "Capacity deficit in mobile wireless ad hoc networks due to routing overheads," Technical Report, Department of ECSE, RPI, Troy, NY. Available at http://www.ecse.rpi.edu/homepages/abouzeid/deficit.pdf.
[16] F. Xue, L.-L. Xie, and P. R. Kumar, "The transport capacity of wireless networks over fading channels." *IEEE Trans. on Information Theory*, vol. 51, no. 3, pp. 834–847, 2005.