Final version of this paper appeared in IEEE Transactions on Mobile Computing, 11(6), 2012.

On the Cost of Knowledge of Mobility in Dynamic Networks: An Information-Theoretic Approach

Di Wang and Alhussein A. Abouzeid

Abstract—In this paper, we extend an information-theoretic approach for characterizing the minimum cost of tracking the motion state information, such as locations and velocities, of nodes in dynamic networks. A rate-distortion formulation is proposed to solve this minimum-cost motion-tracking problem, where the minimum cost is the minimum information rate required to identify the network state at a sequence of tracking time instants within a certain distortion bound. The formulation is applicable to various mobility models, distortion criteria, and stochastic sequences of tracking time instants and hence is general. Under Brownian motion and Gauss-Markov mobility models, we evaluate lower bounds on the information rate of tracking the motion state information of nodes, where the motion state of a node is 1) the node's locations only, or 2) both its locations and velocities. We apply the obtained results to analyze the geographic routing overhead in mobile ad hoc networks. We present the minimum overhead incurred by maintaining the geographic information of nodes in terms of node mobility, packet arrival process, and distortion bounds. This leads to precise characterizations of the observation that given certain state-distortion allowance, protocols aimed at tracking motion state information may not scale beyond a certain level of node mobility.

1 Introduction

In this paper, we study communication networks whose state keeps changing over time. Here, the state information of a network may be composed of link states, node locations, velocity of nodes, etc. In many cases, keeping track of the state information of a dynamic network so as to maintain a timely view of the network is a crucial task. For example,

- In a mobile ad hoc network, a geographic routing protocol requires to maintain node locations (and velocities in some cases, e.g., [2]) in order to make proper routing/forwarding decisions for data packets.
- In a cellular network, a central station may need to collect the locations and velocities of mobile users in order to implement efficient paging and handoff mechanisms.
- In a vehicular network, a traffic control center may want to track the movements of various vehicles so that traffic congestions or accidents may be detected and then communicated to vehicles/drivers in time.
- In a cognitive radio network, a secondary user may have to know the locations of primary users as well as the location of its intended receiver such that it can decide whether to transmit over the primary spectrum or not.
- In a military wireless network (or a rescue and recovery mission), a commander may need to know the locations of all the (mobile) soldiers in a dynamic chaotic environment.

As networks continue to grow in size and become more dynamic, the network state information may change rapidly over time and hence significant control overhead may be incurred by tracking the network state. Such an overhead, in our opinion, has yet to be fully understood.

The premise of this work is that, instead of trying to find solutions that produce the least overhead under certain conditions for each of the aforementioned examples and other network scenarios, we aim to extend a general framework that may bound the overhead while satisfying a constraint that the network state information is maintained within some given accuracy. In this paper, we consider dynamic networks with mobile nodes where the state information being tracked is the *motion state* of nodes. Here, the motion state of a node can be (any combinations of) the location, velocity and/or acceleration of the node. The key departure point of this framework is to treat a node's motion state as a random (vector) process that exhibits random changes. The minimum overhead is the minimum amount of state information rate needed such that the current state of nodes of the network can be identified within a certain distortion bound. This motivates the use of information theory and rate-distortion theory as a tool for developing lower bounds on such an overhead.

We consider the scenario where the motion state of nodes of a dynamic network is tracked by a network master at a sequence of time instants. Here, the network master

- *D. Wang is with Juniper Networks, Inc., 1194 N. Mathilda Ave., Sunnyvale, CA 94089. E-mail: wangdi.thu00@gmail.com.*
- *A.A. Abouzeid is with Electrical,* Computer, and Systems Engineering *Department, Rensselaer Polytechnic Institute, Room 6038, Troy, NY 12180-3590.*
E-mail: abouzeid@ecse.rpi.edu.

© 2010 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

could be any node of the network and is also called "the master node." In our analysis, the *actual* motion state of a node and the motion state *perceived* by the master node are both treated as random processes. The discrepancy between the actual and perceived state is evaluated under some predefined distortion measure, for example, the mean-squared distortion measure. Then, under the given distortion measure, the information rate of tracking the motion state of nodes in a dynamic network is the bit rate at which the master node must receive such that the distortion evaluated at the time instants of interest is bounded. We formulate the problems of finding the minimum value of these bit rates as rate-distortion problems, wherein the tracking time instants are regarded as "side information" (please see Section 3). Using the rate-distortion formulation with side information and under Brownian motion and Gauss-Markov mobility models [3], we obtain lower bounds on the minimum bit rate of tracking the motion state of dynamic networks, where the motion state of a node is, either 1) the node's location information only, or 2) the node's location and velocity information.

Further, we show how the derived results can be applied to one of the aforementioned network scenarios: the protocol overhead of geographic routing protocols in packet-switched wireless networks. We characterize the protocol overhead incurred in terms of node mobility, packet arrival process, and distortion bound ϵ. We then connect the results obtained here with the results on the throughput capacity of stationary multihop wireless networks evaluated in [4], in order to characterize the *effective* throughput capacity available for users. It is shown that, under the Gauss-Markov mobility model, in order to prevent the total traffic from being overwhelmed by protocol overhead, the average speed of nodes in the network must scale down as the total number of nodes grows. This implies that, within a certain state-distortion bound, protocols aimed at tracking motion state information may not be scalable beyond a certain level of node mobility. Please note that the routing protocol initiates the forwarding process as soon as a packet is generated 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 [5] is not applicable to our work.

The contributions of this paper are summarized as follows:

1. We present an information-theoretic formulation for evaluating the overhead incurred by tracking the motion state information of dynamic networks. The formulation is general in that it can be applied to any stochastic sequences of tracking time instants, any distortion measures, and any mobility models, as shown in the analysis.
2. For Brownian motion and Gauss-Markov mobility models and *arbitrary* tracking time distributions, we derive lower bounds for minimum information rate at which a master node must receive motion state information of nodes in the network such that the errors in nodes' motion state are bounded (Theorems 1 and 2 and Corollary 3).
3. The obtained results are applicable to the analysis of protocol overhead of geographic routing protocols in mobile ad hoc networks. We derive the minimum overhead incurred by maintaining the geographic information of nodes in terms of node mobility, packet arrival process, and distortion bounds. It can be observed that given certain state-distortion allowance, protocols aimed at tracking motion state information (such as geographic routing protocols) are not scalable beyond a certain level of node mobility (Theorem 4 and Corollary 5).

The rest of the paper is organized as follows. A brief overview of related work and the relations with this paper are presented in Section 2. The problem definition and its mathematical formulation are presented in Section 3. The evaluation of lower bounds on the information rates of tracking motion state information of dynamic networks is presented in Sections 4 and 5, where in Section 4 the motion state information being tracked is the location information only, while in Section 5 the location and velocity information is tracked jointly. Applications to analyzing the protocol overhead of geographic routing protocols in mobile ad hoc networks is introduced in Section 6. In Section 7, we briefly discuss issues and extensions that are not yet addressed in the previous sections. We present conclusions and future research directions in Section 8.

2 Related Work

We share the viewpoint with the authors in [6] that information theory and networks are two usually disparate fields, and information-theoretic techniques have not yet been widely used in the domain of networking with random user activity. One of the earliest efforts of using information theory to understand the overhead of network protocols dates back to a seminal work by Gallager [7]. Under a simple (stationary) network model and using information-theoretic methods, Gallager [7] quantified the amount of protocol information per packet that is needed for reconstruction of the packets at the destination within a specified mean delay. It is shown in [7] that if the packet lengths are small compared to the packet interarrival times, then the protocol information can far exceed the amount of data bits carried by the packets.

A number of recent studies have used information theory to understand network protocols or applications, and we review the most related results below.

Authors in [8] provide a general approach for analyzing the cost of the registration and paging schemes in cellular networks with the use of information theory. A more concrete information-theoretic approach is used in [9] to study the complexity of tracking a mobile user in a cellular environment and a position update and paging scheme is proposed, wherein the entire network area under study is divided into a finite number of predefined zones, and the node location information being tracked is roughly determined as one of those zones. In [10], the authors propose the entropy of link change as the metric for mobility models against which performance of wireless network protocols could be evaluated. Authors of [11] use rate-distortion theory to investigate the optimal timing for updating a certain link state metric (e.g., bandwidth) such that the update rate is minimized within a given cost constraint. Distributed information discovery in P2P systems is studied from an information-theoretic perspective in [12], through which the overhead induced by the task of information discovery is bounded. Authors in [13] studied the scheduling problem in a TDMA-based network and used rate-distortion theory to characterize the basic limits on the amount of network information that should be transmitted in order to achieve a given level of network performance. An information-theoretic approach is proposed in [14] in order to characterize the minimum routing overhead and memory requirements of a specific two-level hierarchical routing protocols for ad hoc networks. The overhead of generic link state routing protocols is studied using a rate-distortion approach in [15]. The authors in [16] study the family of geographic routing protocols for mobile wireless networks and the rate-distortion lower bounds for the geographic routing overheads that are evaluated under the Brownian-motion mobility model and an i.i.d. assumption about the interarrival times of data packets.

While our work is in some sense along the lines with [14], [15], [16], there are some major differences that underline the contributions of this paper and are summarized as follows:

1. In this paper, we study the network state tracking problem where the network state is the *motion state* of nodes and hence has a broader meaning than the location information studied in [16].
2. The problem formulation used in this paper allows the tracking node to utilize the tracking time information to infer the movement of the node being tracked. The benefits of such a modification include: a) achieving lower information rates than those derived in [16] (see Section 4); b) allowing our analysis to be applied to tracking time sequences with *arbitrary* probability distributions; hence, the i.i.d. assumption about tracking time intervals used in [16] no longer exists in our analysis.
3. While the previous work [16] relies solely on the use of Brownian motion mobility model to deliver mathematically tractable results, in this paper we extend the analysis to the Gauss-Markov mobility model, which is considered to be more realistic, and more widely used in the literature. This allows us, when studying the protocol overhead of geographic routing protocols, to derive some interesting results, e.g., the maximum scaling of the average speed (in terms of total number of nodes) that can be supported by ad hoc networks (Corollary 5).

3 Problem Definition and Mathematical Formulation

We study a general motion-tracking problem described as follows. A number of n mobile nodes are deployed over a 2D plane. For each node i, let $S_i(t)$ denote its motion state at time t, where $S_i(t)$ can be node i's location $X_i(t)$, or its velocity $V_i(t)$, or its acceleration $A_i(t)$, or any possible combinations of these three,[1] all at time t. Assume node i encodes its motion state information and updates to a master node so that the master node can track node i's motion at time instants $\{T_j : j = 1, 2, \ldots\}$ with tracking rate

$$\lambda := \lim_{N\to\infty} \frac{N}{E[T_N]}, \tag{1}$$

such that the actual motion states $\{S_i(T_j) : j = 1, 2, \ldots\}$ of node i are perceived as $\{\hat{S}_i(T_j) : j = 1, 2, \ldots\}$ by the master node. Now, we want to find the minimum information rate at which node i must update its motion state information so that the information $\{\hat{S}_i(T_j) : j = 1, 2, \ldots\}$ decoded by the master node is within a certain accuracy as compared to the actual motion state information $\{S_i(T_j) : j = 1, 2, \ldots\}$. A strict mathematical formulation of this problem is presented through the following definitions.

1. Although the acceleration $A_i(t)$ is mentioned here as one of the possible components of the motion state information, it is not included in the analysis of this paper. We leave this as part of our future work.

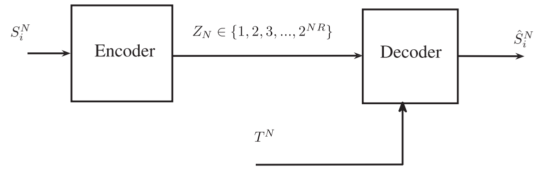


Fig. 1. Rate-distortion formulation with side information.

Definition 1. *$d(S_i(t), \hat{S}_i(t))$ is the distortion measure between node i's actual motion state and the state available at the master node at time t.*

Definition 2. *$S_i^N = [S_i(T_1), S_i(T_2), \ldots, S_i(T_N)]$ is the vector of motion states of node i at time instants T_j, $1 \le j \le N$. Similarly, $\hat{S}_i^N = [\hat{S}_i(T_1), \hat{S}_i(T_2), \ldots, \hat{S}_i(T_N)]$ is the vector of motion states of node i perceived by the master node at time instants T_k, $1 \le k \le N$.*

Definition 3. *$T^N = [T_1, T_2, \ldots, T_N]$ is the vector of the first N time instants that node i's motion state information is tracked by the master node.*

Definition 4. *$\mathcal{S}_i^N$, $\hat{\mathcal{S}}_i^N$, and $\mathcal{T}^N$ are defined as the sets of all possible vectors S_i^N, $\hat{S}_i^N$, and T^N, respectively.*

Definition 5. *P_N denote the probability measure on the sample space $\mathcal{S}_i^N \times \hat{\mathcal{S}}_i^N \times \mathcal{T}^N$. Then, the distortion between S_i^N and $\hat{S}_i^N$, denoted by $D_N(S_i^N, \hat{S}_i^N)$, is defined by*

$$D_N\big(S_i^N, \hat{S}_i^N\big) = \frac{1}{N}\sum_{k=1}^{N} E[d(S_i(T_k), \hat{S}_i(T_k))], \tag{2}$$

where the expectation is taken with respect to the probability measure P_N.

Definition 6. *Let $Z_N \in \mathcal{Z}_N$ denote the encoded version of the actual motion states S_i^N, where $\mathcal{Z}_N$ is the alphabet of Z_N with cardinality $|\mathcal{Z}_N| = 2^{NR}$.*

Without loss of generality, we assume that $\mathcal{Z}_N = \{1, 2, 3, \ldots, 2^{NR}\}$. Note that T^N is statistically correlated with S_i^N; therefore, S_i^N is a useful piece of information for the master node to decode/reconstruct S^N. By the terminology of information theory, T^N is known as *side information*.

Now, we can present this problem formulation based on rate-distortion theory to find the minimum information rate at which node i must update its motion state information such that $D_N(S_i^N, \hat{S}_i^N) \le \epsilon$. Fig. 1 illustrates the formulation schematically, while Definition 7 states it formally.

Definition 7. *The Nth-order rate-distortion function with side information T^N, $R_N(\epsilon)$, is defined as the minimum rate required to achieve distortion ϵ if the side information T^N is available to the decoder. Precisely, $R_N(\epsilon)$ is the infimum of all rates R such that there exists maps $f_N : \mathcal{S}_i^N \rightarrow \{1, 2, \ldots, 2^{NR}\}$, $g_N : \mathcal{T}^N \times \{1, 2, \ldots, 2^{NR}\} \rightarrow \hat{\mathcal{S}}_i^N$ such that*

$$D_N\big(S_i^N, g_N\big(T^N, f_N\big(S_i^N\big)\big)\big) \leq \epsilon. \quad (3)$$

Over a long period of time, the minimum information rate $U(\epsilon)$ (in bits/sec) at which the master node must receive such that the distortion between the actual and perceived motion state information is bounded by ϵ, is given by

$$U(\epsilon) = \lambda \lim_{N\rightarrow\infty} R_N(\epsilon). \quad (4)$$

Note that the problem formulation presented above is general in that it is not restricted to any particular mobility models or any particular distortion measures. In this paper, the mobility models under study will be 1) Brownian motion, and 2) Gauss-Markov mobility model. These two models are introduced as follows:

1. *Brownian motion:* Let $X_i(t) = \{X_{i1}(t), X_{i2}(t)\}$ denote the Cartesian coordinate of a node i's location at time t. Node i is said to be performing Brownian motion with drift velocity $\mu = \{\mu_1, \mu_2\}$ and location variance σ^2, if and only if $X_{i1}(t)$ and $X_{i2}(t)$ are two independent Wiener processes with variance $\frac{1}{2}\sigma^2$, and with drift velocities μ_1 and μ_2, respectively.
2. *Gauss-Markov mobility model:* Let $V_i(t) = \{V_{i1}(t), V_{i2}(t)\}$ denote the velocity (vector) of a node i at time t. Hence, the node i's location at time t is given by

$$X_{ij}(t) = X_{ij}(0) + \int_0^t V_{ij}(s)ds, \ j = 1, 2. \quad (5)$$

Let $W_1(t)$ and $W_2(t)$ denote two independent Wiener processes (without drift) with unit variance. Then, Node i is said to be moving according to the Gauss-Markov mobility model with drift velocity $\mu = \{\mu_1, \mu_2\}$, velocity variance η^2, and relaxation time τ, if and only if $V_{i1}(t)$ and $V_{i2}(t)$ are two independent Ornstein-Uhlenbeck processes[2] with drift velocities μ_1 and μ_2, respectively, and satisfying the following stochastic differential equations

$$dV_{ij}(t) = -\frac{1}{\tau}V_{ij}(t)dt + \frac{\eta}{\sqrt{\tau}}dW_j(t), \quad j = 1, 2. \quad (6)$$

Here, the parameter τ in the Gauss-Markov mobility model indicates how strong the current velocity of a node is correlated to the past, with a large τ indicating a strong correlation. Note that when $\tau \rightarrow 0$ while fixing the product term $2\eta^2\tau$ to be σ^2, i.e., asymptotically there is no correlation between the current and past velocities, the Gauss-Markov model with parameters $\{\mu, \eta^2, \tau\}$ will gradually become the Brownian motion with parameters $\{\mu, \sigma^2\}$.[3]

2. Please refer to [17, Chapter 11-1] for detailed definitions of Wiener processes and Ornstein-Uhlenbeck processes.
3. Please refer to [18] for a detailed explanation on this.

Please note that, while the mobility models used in this paper are all *continuous time*, the Gauss-Markov mobility model first introduced in [3] and used by many others are *discrete time*, which is sampled from the continuous-time model with some time interval Δt. The discretized Gauss-Markov model used in [3] has parameters $\{\mu, \eta^2, \alpha\}$,[4] in which the correlation parameter has the relation $\alpha = e^{-\Delta t/\tau}$, while other parameters still hold the same meanings as those in our model.

Figs. 5, 6, and 7 provide a visualization of the Gauss-Markov mobility model by plotting a number of sample traces of a node's movement under different mobility parameter settings. These figures will help to illustrate some of the analytical results derived in later sections.

4 Information Rates of Location Tracking

In this section, we study the motion-tracking problem where the motion state of interest is the location information only. The mobility model under study is the Brownian motion model introduced in Section 3. The distortion measure used here is the mean-squared error,[5] i.e.,

$$d(X_i(t), \hat{X}_i(t)) := |X_i(t) - \hat{X}_i(t)|^2 = \sum_{j=1}^{2}(X_{ij}(t) - \hat{X}_{ij}(t))^2. \quad (7)$$

We will present our analytical results in Section 4.1 and numerical evaluations in Section 4.2.

4.1 Analytical Results

A lower bound on the minimum information rate of tracking a node's location information under mean-squared distortion measure is presented in the following theorem.

Theorem 1. *Assume a node i is performing Brownian motion with drift velocity μ and location variance σ^2. Then, the lower bound on the minimum information rate $U(\epsilon)$ of tracking a node's location information such that the mean-squared distortion between the actual and perceived location information at time instants $\{T_j\}_{j=1}^{\infty}$ is bounded by ϵ, is given by*

$$U(\epsilon) \geq \lambda\left(E[\log_2 \widetilde{Y}] + \log_2\left(\frac{\sigma^2}{\epsilon}\right)\right), \quad (8)$$

where $\widetilde{Y}$ is a random variable following the mixture pdf

$$f_{\widetilde{Y}}(t) = \lim_{N\rightarrow\infty} \frac{1}{N}\sum_{j=1}^{N} f_{Y_j}(t), \quad t \geq 0, \quad (9)$$

where $f_{Y_j}(t)$ is the pdf of tracking time interval $Y_j := T_j - T_{j-1}$, $j = 1, 2, \ldots$

Proof. Please see Appendix A, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TMC.2011.117. □

4. The notation for the velocity variance in [3] was σ^2. In this paper, we use η^2 instead, because here σ^2 is used to represent the location variance of Brownian motion.
5. Although here we only study the Brownian motion with MSE distortion, the derivation method used here is general and can be applied to other mobility models and other distortion measures. This will be elaborated in Section 7.

TABLE 1
Minimum Location Information Rate (w/o Using Side Information) with Different Tracking-Time Distributions

	$U(\epsilon)$	$U'(\epsilon)$	notes
Case 1	$\lambda \log_2\left(\frac{\sigma^2}{\lambda\epsilon}\right)$	$\lambda \log_2\left(\frac{\sigma^2}{\lambda\epsilon}\right)$	$=$
Case 2	$\lambda \log_2 \frac{2\sigma^2}{\lambda e \epsilon}$	$-\int_{-\infty}^{\infty}\left(\int_0^{2/\lambda} \frac{e^{-\frac{x^2}{\sigma^2 t}}}{\sqrt{\pi\sigma^2 t}}dt\right)\log_2\left(\int_0^{2/\lambda} \frac{e^{-\frac{x^2}{\sigma^2 t}}}{\sqrt{\pi\sigma^2 t}}dt\right)dx - \lambda\log_2(\pi e\epsilon)$	$<$
Case 3	$\lambda \log_2 \frac{\sigma^2}{\lambda e^{\gamma}\epsilon}$	$-2\lambda^2\int_{-\infty}^{\infty}\left(\int_0^{\infty} \frac{e^{-\lambda t-\frac{x^2}{\sigma^2 t}}}{\sqrt{\pi\sigma^2 t}}dt\right)\log_2\left(\int_0^{\infty} \frac{e^{-\lambda t-\frac{x^2}{\sigma^2 t}}}{\sqrt{\pi\sigma^2 t}}dt\right)dx - \lambda\log_2(\pi e\epsilon)$	$<$

Case 1: Periodic; Case 2: Uniform; Case 3: Exponential.

From (8), we notice that the minimum information rate $U(\epsilon)$ does not depend on the drift velocity μ. This is obvious because, while changing μ may affect the expected locations of a node, it has no effects on the uncertainty of node locations. Apart from this observation, we can also make several remarks on the results of Theorem 1.

Remark 1. *On the tightness of the lower bound.* The lower bounds presented in Theorem 1 (8) is known as the *Shannon lower bounds*, which are asymptotically tight as the distortion decreases to zero (see [19] and the references therein). This justifies the evaluation of the information rates using the lower bounds for small ϵ.

Remark 2. On the impact of the statistical properties of tracking time sequence T^N. Notice the fact that the tracking rate $\lambda = E[\widetilde{Y}]^{-1}$, therefore we can rewrite (8) as

$$U(\epsilon) \geq E[\widetilde{Y}]^{-1}\left(E[\log_2 \widetilde{Y}] + \log_2\left(\frac{\sigma^2}{\epsilon}\right)\right). \quad (10)$$

This equation suggests that, for fixed mobility parameter σ and fixed distortion parameter ϵ, the minimum information rate $U(\epsilon)$ only depends on the mixture density of all tracking time intervals $\{Y_j\}_{j=1}^{\infty}$. This result applies to all possible distributions of $\{Y_j\}_{j=1}^{\infty}$ and is not just limited to the i.i.d. case as opposed to [16]. This result also indicates that, the information rate $U(\epsilon)$ cannot be affected by the degree of correlations in tracking times, as long as the mixture density (or the marginal distributions) of $\{Y_j\}_{j=1}^{\infty}$ is known. This observation is important from an engineering point of view, since the mixture density $f_{\widetilde{Y}}(\cdot)$ can be estimated by the empirical distribution of tracking time intervals; hence, $U(\epsilon)$ can be calculated by simply plugging the empirical pdf into (10) without any hassle of dealing with tracking-time correlations.

Remark 3. *On the use of side information T^N.* Using T^N at the decoder helps to reduce the information rates. To see this, note that, according to the proof of Theorem 1 we know that the minimum information rate $U(\epsilon)$ of tracking a node's location information using the side information T^N is given by[6]

$$U(\epsilon) \geq 2\lambda(H(X_{i1}(T_1)|T_1) - \log_2(\pi e\epsilon)). \quad (11)$$

Meanwhile, according to the previous work [16], the minimum information rate of tracking a node's location information *without* utilizing the side information T^N, denoted by $U'(\epsilon)$, is given by

$$U'(\epsilon) \geq 2\lambda(H(X_{i1}(T_1)) - \log_2(\pi e\epsilon)). \quad (12)$$

Since conditioning does not increase entropy, it is implied by (11) and (12) that $U(\epsilon) \leq U'(\epsilon)$, which means that the side information T^N can only help to reduce the information rate. The benefit of using side information will be evaluated numerically in the following section.

4.2 Numerical Evaluations and Discussions

Theorem 1 provides analytical expressions for the minimum information rate $U(\epsilon)$ with the time-interval sequence $\{Y_j\}_{j=1}^{\infty}$ having a general mixture pdf $f_{\widetilde{Y}}(t)$. In this section, we study the information rate $U(\epsilon)$ numerically by considering several specific mixture probability distributions for $\{Y_j\}_{j=1}^{\infty}$, and make comparisons with the results in [16] where the side information T^N is not utilized by the tracker. Please note that we are evaluating the results of Theorem 1 without the need of assuming $\{Y_j\}_{j=1}^{\infty}$ are i.i.d.; however, we have to limit our comparisons with [16] to the i.i.d. case.

Fixing the tracking rate λ, we evaluate $U(\epsilon)$ with (8) for the following cases:

1. Case 1: the tracking time instants $\{T_j\}_{j=1}^{\infty}$ is (deterministically) periodic, i.e., the distribution of $\{Y_j\}_{j=1}^{\infty}$ follows a trivial pdf $f_{\widetilde{Y}}(t) = \delta(t - \lambda^{-1})$.
2. Case 2: The mixture density of $\{Y_m\}_{m=1}^{\infty}$ is the uniform distribution with pdf

$$f_{\widetilde{Y}}(t) = \begin{cases} \frac{\lambda}{2}, & 0 \leq t \leq 2\lambda^{-1}; \\ 0, & \text{otherwise.} \end{cases} \quad (13)$$

3. Case 3: The mixture density of $\{Y_m\}_{m=1}^{\infty}$ is the exponential distribution with pdf

$$f_{\widetilde{Y}}(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0; \\ 0, & \text{otherwise.} \end{cases}$$

For conciseness, the results are summarized in Table 1. The derivations are placed in Appendix B, available in the online supplemental material.

Figs. 2 and 3 show the plot of the minimum information rates $U(\epsilon)$ and $U'(\epsilon)$ against the location variance of Brownian motion σ^2 and the tracking rate λ (by a master node) for the three different cases. Several conclusions can be drawn according to these figures:

6. In order to compare with the results in [16], we have to apply Theorem 1 to the special case that the tracking time intervals $\{Y_j\}_{j=1}^{\infty}$ are i.i.d.

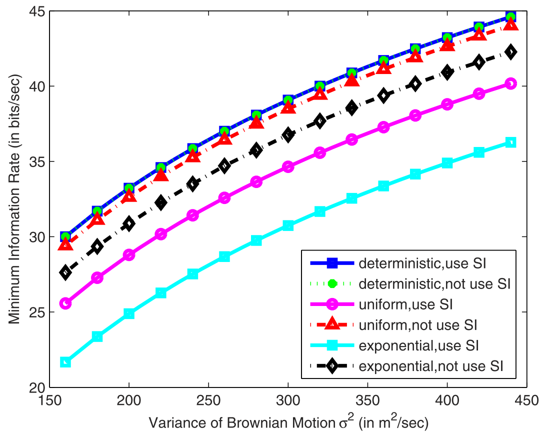


Fig. 2. Minimum location information rate (w/o using side information) versus location variance of Brownian motion. $\lambda = 10, \epsilon = 2$.

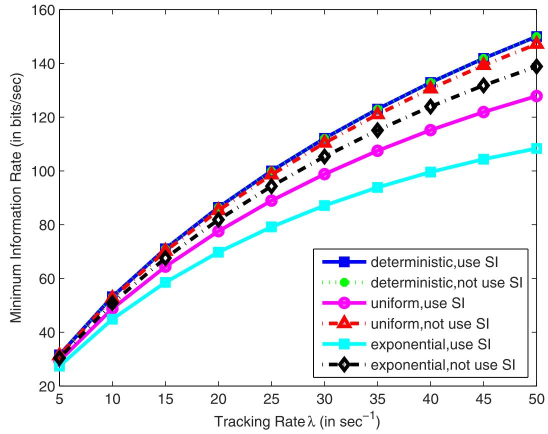


Fig. 3. Minimum location information rate (w/o using side information) versus tracking rate by a master node. $\lambda = 10, \epsilon = 2, \sigma^2 = 400$.

- It is observed that for high location variance σ^2 and high tracking rate λ, the minimum information rate required to identify the locations of a node in dynamic networks becomes very high.
- For Case 1, where the side information $\{T_m\}_{m=1}^{\infty}$ bears no randomness, it is shown that the presence of side information will not reduce the information rate. However, for Cases 2 and 3, it is observed that side information does reduce the information rate thanks to its random nature.
- It is shown that the rate required for periodical observations (Case 1) is higher than that required for uniform and exponential cases (Cases 2 and 3). In fact, the rate for periodical location tracking is the highest of any probability distributions of $\widetilde{Y}$ with fixed tracking rate λ. This can be shown as follows. By Jensen's inequality, the minimum information rate $U(\epsilon)$ for an arbitrary probability distribution of $\widetilde{Y}$

$$\begin{aligned}\lambda\Bigg(\int_0^{\infty} \log_2(t) f_{\widetilde{Y}}(t)dt + \log_2\left(\frac{\sigma^2}{\epsilon}\right)\Bigg) &= \lambda E[\log_2 \widetilde{Y}] + \lambda \log_2\left(\frac{\sigma^2}{\epsilon}\right) \\ &\leq \lambda \log_2 E[\widetilde{Y}] + \lambda \log_2\left(\frac{\sigma^2}{\epsilon}\right) \\ &= \lambda \log_2\left(\frac{\sigma^2}{\lambda\epsilon}\right)\end{aligned}$$

is shown to be no greater than the minimum information rate of periodical tracking provided by Case 1 from Table 1.

5 Information Rates of Location-Velocity Joint Tracking

In this section, we study the motion-tracking problem where a node's location and velocity are being tracked jointly, i.e., the motion state of interest is composed of a node's location information and its velocity information; hence, the motion state of a node i at time t is described by $S_i(t) = \{X_i(t), V_i(t)\}$. The mobility model used here is the Gauss-Markov model introduced in Section 3. We still use the mean-squared error as the distortion measure, i.e.,

$$d(S_i(t), \hat{S}_i(t)) = \{d_1(S_i(t), \hat{S}_i(t)), d_2(S_i(t), \hat{S}_i(t))\}, \tag{14}$$

where

$$\begin{cases} d_1(S_i(t), \hat{S}_i(t)) := |X_i(t) - \hat{X}_i(t)|^2 \\ \quad = \sum_{j=1}^{2}(X_{ij}(t) - \hat{X}_{ij}(t))^2 \\ d_2(S_i(t), \hat{S}_i(t)) := |V_i(t) - \hat{V}_i(t)|^2 \\ \quad = \sum_{j=1}^{2}\big(V_{ij}(t) - \hat{V}_{ij}(t)\big)^2. \end{cases} \tag{15}$$

Compared to Section 4, a new component/dimension of the motion state information is introduced in the analysis of this section. This allows us to demonstrate the generality of the formulation proposed in this paper. In Section 5.1, we will present our analytical results and the connections to the results presented in Section 4. Then, we provide numerical evaluations in Section 5.2 and give insights on how the mobility parameters may affect the information rates.

5.1 Analytical Results

A lower bound on the minimum information rate of tracking a node's location and velocity information jointly under mean-squared distortion measure is presented in the following theorem.

Theorem 2. *Assume a node i is moving according to the Gauss-Markov mobility model with drift velocity μ, velocity variance η^2, and relaxation time τ. Then, the lower bound on the minimum information rate $U(\epsilon)$ of* joint tracking a node's location and velocity information, *such that the mean-squared distortion between the actual and perceived motion information at time instants $\{T_j\}_{j=1}^{\infty}$ is bounded by $\epsilon = \{\epsilon_1, \epsilon_2\}$, is given by*

$$U(\epsilon) \geq \max\{L_1, L_2\}, \tag{16}$$

where

$$\begin{cases} L_1(\epsilon) = \lambda\Big(E\Big[\log_2\Big(\frac{2\widetilde{Y}}{\tau}\big(1 + e^{-\widetilde{Y}/\tau}\big) - 4\big(1 - e^{-\widetilde{Y}/\tau}\big)\Big)\Big] \\ \quad + E[\log_2\big(1 - e^{-\widetilde{Y}/\tau}\big)] + \log_2 \frac{\eta^4\tau^2}{\epsilon_1\epsilon_2}\Big) \\ L_2(\epsilon_1) = \lambda\Big(E\Big[\log_2\Big(\frac{2\widetilde{Y}}{\tau}\big(1 + e^{-\widetilde{Y}/\tau}\big) - 4\big(1 - e^{-\widetilde{Y}/\tau}\big)\Big)\Big] \\ \quad - E\big[\log_2\big(1 + e^{-\widetilde{Y}/\tau}\big)\big] + \log_2 \frac{\eta^2\tau^2}{\epsilon_1}\Big), \end{cases}$$

where $\widetilde{Y}$ is a random variable following the mixture density $f_{Y_j}(t)$ of tracking time intervals $Y_j := T_j - T_{j-1}$, $j = 1, 2, \ldots$

Proof. Please see Appendix C, available in the online supplemental material. □

From Theorem 2, we notice that, the lower bound L_2 on $U(\epsilon)$ is dependent on the location error ϵ_1 and is independent on the velocity error ϵ_2. In fact, it is actually the lower bound for the minimum rate of tracking a node's location information only when the node's movement is governed by the Gauss-Markov mobility model. Therefore, we have the following corollary.

Corollary 3. *Assume a node i is moving according to the Gauss-Markov mobility model with drift velocity μ, velocity variance η^2, and relaxation time τ. Then, the lower bound on the minimum information rate $U(\epsilon)$ of tracking a node's* location information only, *such that the mean-squared distortion between the actual and perceived location information at time instants $\{T_j\}_{j=1}^{\infty}$ is bounded by ϵ, is given by*

$$U(\epsilon) \geq L_2(\epsilon), \tag{17}$$

where L_2 is from (16).

Here, the parameter τ in the Gauss-Markov mobility model indicates how strong the current velocity of a node is correlated to the past, and a large τ indicates a strong correlation. Note that the Gauss-Markov model with parameters $\{\mu, \eta^2, \tau\}$ will gradually become the Brownian motion with parameters $\{\mu, \sigma^2\}$ if we allow $\tau \to 0$ while fixing the product term $2\eta^2\tau$ to be σ^2. This allows us to verify the consistency between Corollary 3 and Theorem 1 as follows.

Notice that for $\tau \to 0$, $(1 + e^{-\widetilde{Y}/\tau})$ and $(1 - e^{-\widetilde{Y}/\tau})$ tend to 1, and $E[\log_2(1 + e^{-\widetilde{Y}/\tau})]$ approaches 0, then by enforcing $\sigma^2 = 2\eta^2\tau$, (17) becomes

$$\begin{aligned} U(\epsilon) &\geq \lambda \left(E\left[\log_2 \frac{2\widetilde{Y}}{\tau}\right] + \log_2 \frac{\eta^2\tau^2}{\epsilon} \right) \\ &= \lambda \left(E[\log_2 \widetilde{Y}] + \log_2 \left(\frac{\sigma^2}{\epsilon} \right) \right), \end{aligned} \tag{18}$$

which conforms to (8) in Theorem 1.

The above analysis that reaches (18) does provide an alternate route for reaching the conclusions of Theorem 1. The above derivation relies on the special connection between Brownian motion and the Gauss-Markov Mobility Model and hence does not preserve the generality of the approach used in deriving Theorem 1 in Section 4. Moreover, the special trick used here cannot yield a meaningful understanding about the effect of the side information newly introduced in this work. Therefore, we cannot regard this as a replacement of Theorem 1.

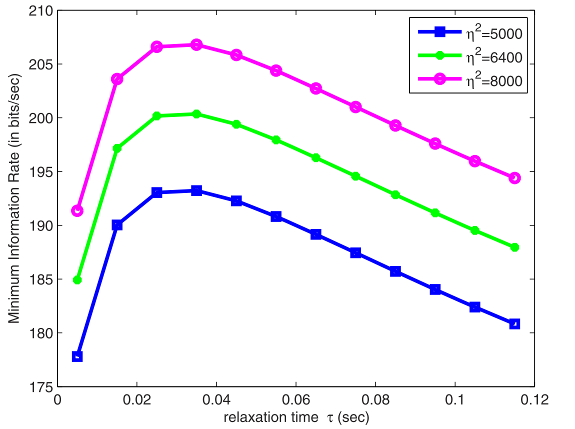


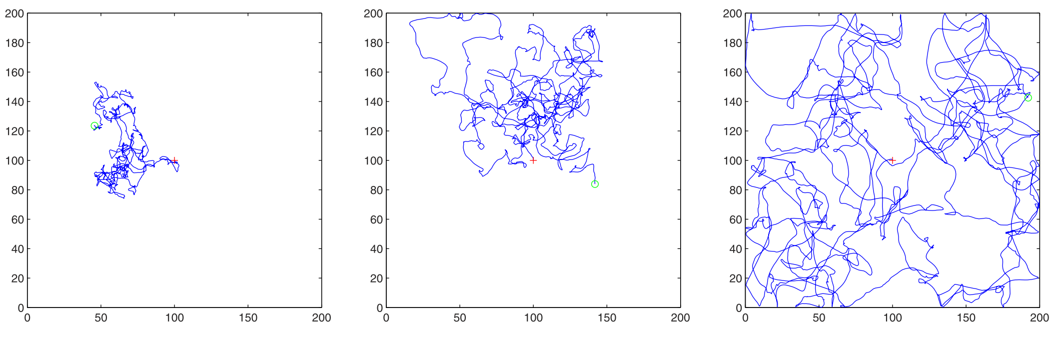
Fig. 4. Minimum location information rate versus relaxation time. $\lambda = 10, \epsilon_1 = 0.01, \epsilon_2 = 10$.

5.2 Numerical Evaluations and Discussions

In this section, we evaluate the results of Theorem 2 numerically. Due to mathematical complexities, we only consider the case where the tracking time instants $\{T_j\}_{j=1}^{\infty}$ is (deterministically) periodic, i.e., the distribution of $\{Y_j\}_{j=1}^{\infty}$ follows a trivial pdf $f_Y(t) = \delta(t - \lambda^{-1})$. In this case, (16) becomes

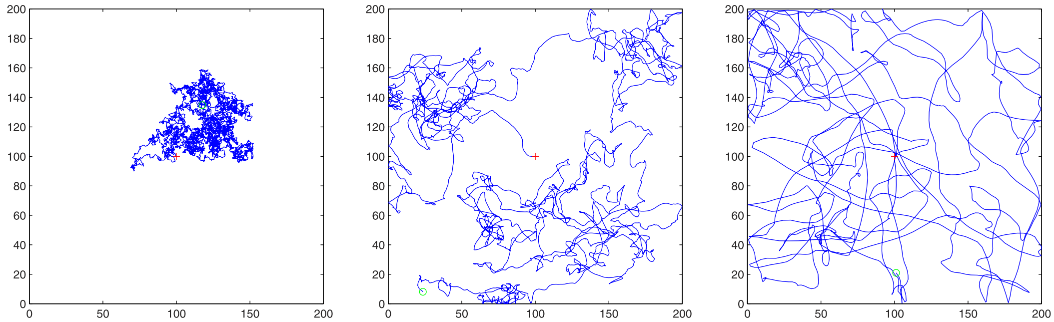
$$U(\epsilon) \geq \begin{cases} \lambda\Big(\log_2\Big(\frac{2}{\lambda\tau}\big(1+e^{-\frac{1}{\lambda\tau}}\big) - 4\big(1-e^{-\frac{1}{\lambda\tau}}\big)\Big) \\ \quad + \log_2\Big(1-e^{-\frac{1}{\lambda\tau}}\Big) + \log_2\frac{\eta^4\tau^2}{\epsilon_1\epsilon_2}\Big), & \text{if } \frac{\epsilon_2}{\eta^2} + e^{-\frac{1}{\lambda\tau}} \leq 1; \\ \lambda\Big(\log_2\Big(\frac{2}{\lambda\tau}\big(1+e^{-\frac{1}{\lambda\tau}}\big) - 4\big(1-e^{-\frac{1}{\lambda\tau}}\big)\Big) \\ \quad - \log_2\big(1+e^{-\frac{1}{\lambda\tau}}\big) + \log_2\frac{\eta^2\tau^2}{2\epsilon_1}\Big), & \text{otherwise.} \end{cases}$$

Fig. 4 shows the plot of the minimum information rate $U(\epsilon)$ against relaxation time τ and velocity variance η^2 under finite distortion $\{\epsilon_1, \epsilon_2\}$. It is observed that for large η^2 the information rate of tracking a node's motion becomes very high. In fact, if we take a look at Fig. 5 wherein sample



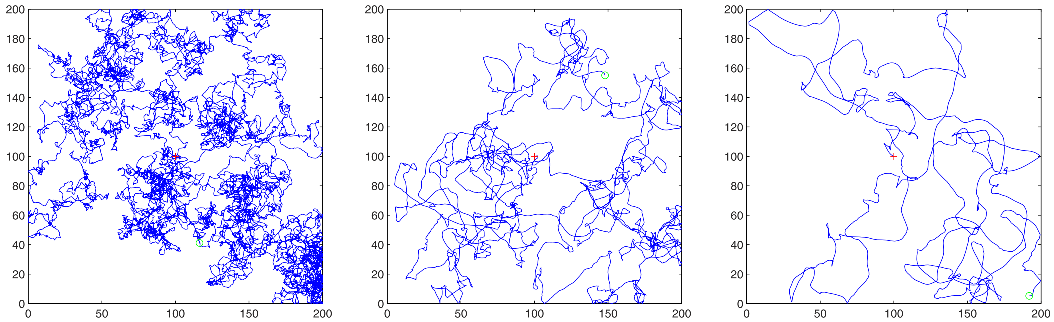
(a) $\eta^2 = 200$ (b) $\eta^2 = 1000$ (c) $\eta^2 = 6000$

Fig. 5. Sample traces of Gauss-Markov mobility model for different values of η^2. ($\mu = 0$; $\tau = 0.2$; simulation time: 100.)



(a) $\tau = 0.01$ (b) $\tau = 0.10$ (c) $\tau = 0.50$

Fig. 6. Sample traces of Gauss-Markov mobility model for different values of τ. ($\mu = 0$; $\eta^2 = 5{,}000$; simulation time: 100.)



(a) $\tau = 0.01$ (b) $\tau = 0.10$ (c) $\tau = 0.50$

Fig. 7. Sample traces of Gauss-Markov mobility model for different values of τ while fixing $\eta^2\tau$ as constant. ($\mu = 0$; $\eta^2\tau = 500$; simulation time: 100.)

traces of a node's movement are plotted for different values of η^2, we can notice that for small η^2 the node's movement can be very restricted (Fig. 5a); with the increase of η^2, the node's movement can become more and more unrestricted (Figs. 5 and 5c), indicating an increasing uncertainty of the node's motion state.

From Fig. 4, we also observe that, for large relaxation time τ, with the decrease of τ, the information rate $U(\epsilon)$ will be increasing. This fact can be understood straightforwardly, since a decrease in τ suggests that a node's current and past velocities are becoming less and less correlated, and hence it will increase the uncertainty of the node's motion state. However, for small τ, it is shown in Fig. 4 that, with the decrease of τ, the information rate $U(\epsilon)$ is actually decreasing. The reason behind this seemingly counter-intuitive phenomenon is that, when τ is small, the location variance of node i is

$$\mathrm{var}[X_{ij}(t)] \approx \eta^2 t\tau, \quad j = 1, 2. \tag{19}$$

Therefore, the location uncertainty of a node will diminish with the decrease of τ and hence yield low information rate $U(\epsilon)$. In fact, if we take a look at Fig. 6 wherein sample traces of a node's movement are plotted for different values of τ, we can notice that the node's movement can become very restricted if we simply decrease τ while keeping η^2 unchanged.

In order to remove the effect of τ on the location uncertainty, we have to fix $\eta^2\tau$ to be some constant. Now, we take a look at Fig. 7 wherein sample traces of a node's movement are plotted for different values of τ while keeping $\eta^2\tau$ as constant. In this case, we can observe that the node's movement is no longer restricted even if τ takes on a very small value (e.g., $\tau = 0.01$, Fig. 7a).

By letting $2\eta^2\tau = \sigma^2$, we plot the curves of $U(\epsilon)$ against τ and σ^2 under finite distortion $\{\epsilon_1, \epsilon_2\}$ in Fig. 8. In this case, the information rate $U(\epsilon)$ is always increasing with the decrease of τ. In fact, it is easy to verify that for this case the rate $U(\epsilon)$ will become unbounded with $\tau \to 0$. This is because for $\tau \to 0$, a node will asymptotically perform Brownian motion and hence the node's instantaneous velocity become unbounded,[7] and hence it is impossible to keep track of the velocity of Brownian motion under a finite distortion error.

6 Application: Analyzing the Overhead of Geographic Routing Protocols

In this section, we present how the theories developed in the above sections can be applied to analyze the overhead incurred by geographic routing protocols over mobile ad hoc networks. First, we present a brief introduction of geographic routing protocols in Section 6.1. Then, in Section 6.2, we characterize the overhead of geographic routing protocols for mobile wireless networks and investigate how the protocol overhead interacts with the data traffic.

6.1 Geographic Routing Protocols

In packet-switched communication networks, routing is the process of forwarding data packets from one host to another such that packets may eventually reach their destinations. Geographic routing is a routing technique

7. Please see [17] for detailed explanation.

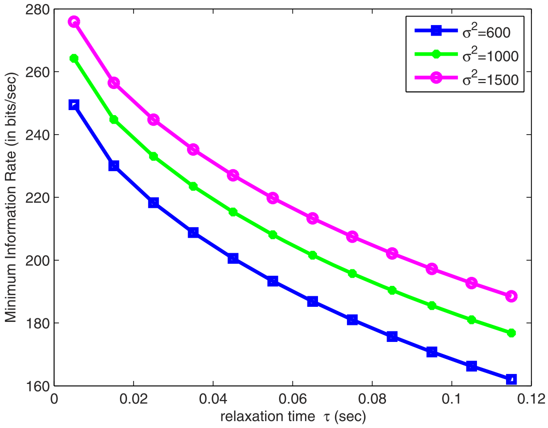


Fig. 8. Minimum location information rate versus relaxation time by fixing $2\eta^2\tau$ to be σ^2. $\lambda = 10, \epsilon_1 = 0.01, \epsilon_2 = 10$.

that uses the position of destination nodes in order to make routing decisions [20]. Geographic routing requires nodes to know their locations. This may be accomplished using GPS [21] or other mechanisms [22]. Also, geographic routing requires a distributed location service which allows source nodes to collect the location information of the destination nodes [23]. All geographic routing protocols function in the following manner. When a new packet arrives at a source node, it queries the location service in order to discover the current position of the destination. The position of the destination is added to packet headers and the source and intermediate nodes forward the packet to a neighbor that is closer to the destination.

It is clear that, in geographic routing, the process of collecting the destinations' location information by the source nodes produces an overhead. Such a process can be abstracted as follows. For any source-destination pair (k, i), destination node i will encode its location information and send out as control packets, which may be received by the source node k such that it can make routing decisions upon the arrivals of data packets at node k. Therefore, *the control message overhead associated with S-D pair* (k, i) is formed by the bit streams of all such control packets received by node k. In the following section, we will apply the results derived in previous sections to bound such an overhead.

6.2 Overhead-Throughput Analysis of Mobile Ad Hoc Networks Using Geographic Routing

Consider a mobile ad hoc network of n nodes that are randomly distributed over a torus with surface area of $A(n) = \Theta(n)$.[8] The movement of each node is governed by the Gauss-Markov mobility model with drift velocity $\mu(n)$, velocity variance $\eta^2(n)$ and relaxation time $\tau(n)$. The transmission range of each node is assumed to be fixed and is denoted by $r(n)$. We pick uniformly at random a matching of source-destination pairs, so that each node is the destination of exactly one source. The source node k is constantly generating data packets (with packet size c bits) destined to i at rate $\lambda(n)$ packets/sec such that the jth packet destined to destination i is generated at time $T_j = j/\lambda(n), \forall j \geq 0$.

8. This network area setting keeps the node density constant, and is termed as "extended network" in some works in the literature.

We assume that both the drift velocity $\mu(n)$ and the velocity variance $\eta^2(n)$ of a node can only scale as $O(1)$. Since the network radius scales as $\Theta(n)$, this ensures that the nodes do not wrap around the surface during the time scale corresponding to packet interarrival times. So, if we look at the movement of a node during a small time interval, then with high probability the motion is similar to the Gauss-Markov model on an infinite 2D plane. Therefore, we study the movement of nodes over the torus as if they are moving over an infinite plane and hence results from previous sections can be applied here.

We also assume that the relaxation time $\tau(n)$ scales as $\Theta(1)$. This is because, either $\tau(n) = o(1)$ or $\tau(n) = \omega(1)$ will make the Gauss-Markov mobility model become trivial: 1) when $\tau(n) \to 0$, nodes are performing Brownian motion with zero location variance; 2) when $\tau(n) \to \infty$, nodes are just moving with constant velocity $\mu(n)$.

According Corollary 3, in order to have node i's location error bounded by $\epsilon(n)$ at node k, the minimum bit rate at which node k must receive node i's location information is given by

$$U(\epsilon) \geq \lambda(n)\Bigg(\log_2\bigg(\frac{2}{\lambda(n)\tau}\big(1+e^{-\frac{1}{\lambda(n)\tau}}\big) - 4\big(1-e^{-\frac{1}{\lambda(n)\tau}}\big)\bigg) - \log_2\big(1+e^{-\frac{1}{\lambda(n)\tau}}\big) + \log_2\frac{\eta^2(n)\tau^2}{2\epsilon(n)}\Bigg). \tag{20}$$

Note that the maximum per-node throughput of a random multihop wireless network scales as $\Theta(1/\sqrt{n\log n})$ [4], and it is indicated by Corollary 3 that in mobile wireless networks, the control message overhead produced by geographic routing protocols reduces the per-node throughput at least by $U(\epsilon(n))$. Therefore, the bit rate of transmitting data packets from a source to its destination, $c\lambda(n)$, which is termed as *effective per-node throughput*, can only scale as $O(1/\sqrt{n\log n})$, and it remains to be seen whether the effective per-node throughput $c\lambda(n)$ can actually attain the order $\Theta(1/\sqrt{n\log n})$. This question is answered in the following theorem.

Theorem 4. *Consider a mobile ad hoc network where node movement follows the Gauss-Markov mobility model. Then, in order to have the effective per-node throughput* $c\lambda(n)$ *achieve the optimal scaling* $\Theta(1/\sqrt{n\log n})$, *the velocity variance of a node* $\eta^2(n)$ *must scale as* $O(\epsilon(n)/\sqrt{n\log n})$.

Proof. Please see Appendix D, available in the online supplemental material. □

The results of Theorem 4 determines the level of node mobility (in terms of total number of nodes in the network) that geographic routing protocols can ultimately support. The exact order of node mobility, which is for this case the velocity variance $\eta^2(n)$, however, is still not explicit and is dependent on the location error $\epsilon(n)$. It is clear that the location error $\epsilon(n)$ cannot be too large since otherwise the routing protocols will not function properly. Since the determination of location error $\epsilon(n)$ in order to achieve satisfactory network performance (e.g., high delivery ratio) is out of the scope of this paper, we need to borrow some results of such a study in the literature, e.g., [29] and [30]. It is observed in [29] that for geographic routing protocols using greedy forwarding, 20 percent error in node's location information (relative to its transmission range r) may cause substantial loss in performance. An improvement, comprising of maintaining two hop neighborhood information, is proposed which makes geographic routing tolerant to 40 percent error in location. Similar observations regarding loss in performance due to location errors induced by increased mobility and beacon interval are made in [30]. Based on these observations, for the mean-squared location error $\epsilon(n)$ with the dimension of unit squared distance, we use the order $\epsilon(n) = O(r^2(n))$ as a rule of thumb to determine the location error that a geographic routing protocol can tolerate. In this case, we have the following corollary.

Corollary 5. *Assume a geographic routing protocol can only tolerate a mean-squared location error $\epsilon(n) = O(r^2(n))$, and let the drift velocity $\mu(n)$ in the Gauss-Markov model for each node to be 0. Then, in order to have the effective per-node throughput $c\lambda(n)$ achieve the optimal scaling $\Theta(1/\sqrt{n \log n})$, the average speed of a node $\overline{|V(n)|} := E[|V(n)|]$ must scale as*

$$O\left(\sqrt{\frac{\log n}{n}}\right). \tag{21}$$

For the case $\overline{|V(n)|} = \Theta(1)$, the effective per-node throughput $c\lambda(n)$ can only scale as

$$O\left(\frac{1}{\sqrt{n}\log^{1.5}(n)}\right). \tag{22}$$

Proof. Please see Appendix E, available in the online supplemental material. □

The results of Corollary 5 indicate that, under Gauss-Markov mobility model without drift, in order to have the data rate of each S-D pair achieving the maximum per-node throughput, the average speed of each node must scale down with the increase of the number of node n. If we let the average speed of each node stay constant, then, as compared to the maximum per-node throughput scaling $\Theta(1/\sqrt{n \log n})$ that can be achieved in static multihop wireless networks, the data rate of each S-D pair must be reduced by a factor of $\log n$ in order to let the network accommodate the protocol overhead incurred by geographic routing protocols.

It is worth noting that, under the same network setting, results in [24] indicate that for topology-based proactive routing protocols (e.g., DBF [25] and OLSR [26]), the maximum mobility degree (measured in node's average speed) can be supported by the network is $\overline{|V(n)|} = O(1/(n \log n)^{3/2})$. On the other hand, for reactive routing protocols that maintain the route state or vector state of a network (e.g., DSR [28] and AODV [27]), the maximum mobility degree must scale as $\overline{|V(n)|} = O(\frac{1}{n^2})$. By comparison we understand that, asymptotically, geographic routing protocols may tolerate a relatively higher degree of mobility and hence have better scalability than the routing protocols examined in [24] that maintain other types of state information of the network.

7 Discussions

In this section, we discuss issues yet to be addressed in this paper. First, we discuss how the formulation and analysis may be extended to the absolute difference distortion measures. We then address why the cost of tracking local neighboring information is not considered in the analysis of geographic routing.

7.1 Absolute Difference Distortion

In previous sections, all the rate-distortion bounds are derived under the MSE distortion. However, we do realize other distortion criteria may also be of great interest. For example, in some cases, we do want to bound node locations nonprobabilistically, and consequently the absolute difference distortion measure makes more sense. Precisely, the per symbol distortion is given by[9]

$$d(X_i(t), \hat{X}_i(t)) := |X_i(t) - \hat{X}_i(t)| = \sqrt{\sum_{j=1}^{2} (X_{ij}(t) - \hat{X}_{ij}(t))^2}. \tag{23}$$

And the Nth-order total distortion is defined by

$$D_N\left(X_i^N, \hat{X}_i^N\right) := \max_{1 \le k \le N} \{d(X_i(T_k), \hat{X}_i(T_k))\}. \tag{24}$$

With this new distortion measure, we let

$$D_N\left(X_i^N, \hat{X}_i^N\right) \le \sqrt{\epsilon}, \tag{25}$$

which is clearly a more stringent requirements for maintaining location accuracy. Under Brownian motion, the result of Theorem 1 is changed to the following:

$$U(\epsilon) \ge \lambda\left(E[\log_2 \widetilde{Y}] + \log_2\left(\frac{e\sigma^2}{\epsilon}\right)\right). \tag{26}$$

The proof of this derivation can be found in Appendix F, available in the online supplemental material.

Compared to the bound under MSE distortion, the new bound under absolute difference distortion measure is larger by $\lambda \log_2 e$.

7.2 Why the Cost of Tracking Neighboring Information Not Considered

In Section 6, we analyzed the protocol overhead of geographic routing protocols. In particular, the overhead under study is the cost of source node in knowing destination node's position estimates. Admittedly, some geographic routing algorithms require nodes to exchange beacons to obtain neighboring information and hence incur an overhead as known as "beacon overhead." The rate-distortion bound on beacon overhead is well studied in the previous work [16]. However, the reason we do not study or revisit such an overhead in this paper is that, beacon exchanges are not absolutely required by all geographic routing protocols. For example, Beaconless Geographic Routing [31], [32] does not require such a mechanism and hence does not incur such an overhead. Since our protocol overhead analysis mainly focus on investigating the scaling

9. For notation simplicity, we only discuss tracking the location information only.

properties rather than giving precise characterization of the actual overhead for all scenarios, therefore we prefer to only analyze the part of the overhead necessary for all geographic routing protocols.

8 Conclusions and Future Work

In this paper, we extended an information-theoretic framework for bounding the motion-tracking cost in dynamic networks. In this problem, a tracking node wants to maintain the motion state information, such as locations and velocities, of other nodes in the network. The objective is to find the minimum amount of information rate (i.e., overhead) required by the tracking node such that the distortion between the actual and the perceived motion state information of other nodes is bounded. We formulate this minimum overhead problem as a rate-distortion problem with side information. As compared to previous work [16], such a formulation not only helps to yield better lower bounds on the information rate (protocol overhead), but also extend the analysis to more realistic mobility models, and incorporate more general stochastic models for the sequence of tracking time instants. It is shown that our theoretical analysis can be applied to analyze the protocol overhead of geographic routing protocols in mobile ad hoc networks. The generality of our proposed framework leads to some interesting results for large-scale mobile ad hoc networks, including finding the maximum scaling of average node speed that can be supported by networks that use geographic routing to route packets. Directions of future research may include further extending our proposed information-theoretic framework to incorporate additional or other types of network state.

Acknowledgments

This work was supported by US National Science Foundation Grant Nos. 322956 and 546402. Initial work [1] was accepted to the INFOCOM 2010 conference.

References

[1] D. Wang and A.A. Abouzeid, "On the Cost of Knowledge of Mobility in Dynamic Networks," *Proc. IEEE INFOCOM,* 2010.
[2] K.-T. Feng, C.-H. Hsu, and T.-E. Lu, "Velocity-Assisted Predictive Mobility and Location-Aware Routing Protocols for Mobile Ad Hoc Networks," *IEEE Trans. Vehicular Technology,* vol. 57, no. 1, pp. 448-464, Jan. 2008.
[3] B. Liang and Z.J. Haas, "Predictive Distance-Based Mobility Management for Multidimensional PCS Networks," *IEEE Trans. Networking,* vol. 11, no. 5, pp. 718-732, Oct. 2003.
[4] P. Gupta and P.R. Kumar, "The Capacity of Wireless Networks," *IEEE Trans. Information Theory,* vol. 46, no. 2, pp. 388-404, Mar. 2000.
[5] M. Grossglauser and D. Tse, "Mobility Increases the Capacity of Ad Hoc Wireless Networks," *IEEE/ACM Trans. Networking,* vol. 10, no. 4, pp. 477-486, Aug. 2002.
[6] A. Ephremides and B. Hajek, "Information Theory and Communication Networks: An Unconsummated Union," *IEEE Trans. Information Theory,* vol. 44, no. 6, pp. 2416-2434, Oct. 1998.
[7] R.G. Gallager, "Basic Limits on Protocol Information in Data Communication Networks," *IEEE Trans. Information Theory,* vol. 22, no. 4, pp. 385-398, July 1976.
[8] C. Rose and R. Yates, "Location Uncertainty in Mobile Networks: A Theoretical Framework," *IEEE Comm. Magazine,* vol. 35, no. 2, pp. 94-101, Feb. 1997.
[9] A. Bhattacharya and S.K. Das, "LeZi-Update: An Information-Theoretic Approach to Track Mobile Users in PCS Networks," *Proc. ACM MobiCom,* 1999.
[10] Q.M. Tran, A. Dadej, and S. Parreau, "Characterizing Mobility in Ad Hoc Networks: A Generalized Approach," *Proc. Workshop Applications and Services in Wireless Networks (ASWN '05),* 2005.
[11] G. Cheng and N. Ansari, "Rate-Distortion Based Link State Update," *Computer Networks,* vol. 50, pp. 3300-3314, 2006.
[12] H. Hong, "Distributed Information Discovery in Networks from the Perspective of Information Theory," http://ccr.sigcomm.org/online/files/distr_info.pdf, 2007.
[13] J. Hong and V.O.K. Li, "Impact of Information on Network PerformanceCAn Information-Theoretic Perspective," *Proc. GlobeCom,* 2009.
[14] N. Zhou and A.A. Abouzeid, "Routing in Ad Hoc Networks: A Theoretical Framework with Practical Implications," *Proc. IEEE INFOCOM,* 2005.
[15] D. Wang and A.A. Abouzeid, "Link State Routing Overhead in Mobile Ad Hoc Networks: A Rate-Distortion Formulation," *Proc. IEEE INFOCOM,* 2008.
[16] N. Bisnik and A.A. Abouzeid, "Capacity Deficit in Mobile Wireless Ad Hoc Networks Due to Geographic Routing Overheads," *Proc. IEEE INFOCOM,* 2007.
[17] A. Papoulis and U.S. Pillai, *Probability, Random Variables and Stochastic Processes.* McGraw-Hill Science/Engineering/Math, 2001.
[18] D.T. Gillespie, "Exact Numerical Simulation of the Ornstein-Uhlenbeck Process and Its Integral," *Physical Rev. E,* vol. 54, pp. 2084-2091, 1996.
[19] T. Linder and R. Zamir, "On the Asymptotic Tightness of the Shannon Lower Bound," *IEEE Trans. Information Theory,* vol. 40, no. 6, pp. 2026-2031, Nov. 1994.
[20] M. Mauve, J. Widmer, and H. Hartenstein, "A Survey on Position-Based Routing in Mobile Ad-Hoc Networks," *IEEE Network Magazine,* vol. 15, no. 6, pp. 30-39, Nov./Dec. 2001.
[21] E. Kaplan, *Understanding GPS: Principles and Applications.* Artech House, 1996.
[22] S. Capkun, M. Hamdi, and J. Hubaux, "GPS-Free Positioning in Mobile Ad-Hoc Networks," *Proc. 34th Ann. Hawaii Int'l Conf. System Sciences (HICSS '01),* 2001.
[23] S. Das, H. Pucha, and Y. Hu, "Performance Comparison of Scalable Location Services for Geographic Ad Hoc Routing," *Proc. IEEE INFOCOM,* 2005.
[24] Z. Ye and A.A. Abouzeid, "A Unified Model for Joint Throughput-Overhead Analysis of Mobile Ad Hoc Networks," *Proc. Int'l Symp. Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM '08),* 2008.
[25] D. Bertsekas and R. Gallager, *Data Networks.* Prentice Hall, 1992.
[26] P. Jacquet et al., "Optimized Link State Routing Protocol," IETF MANET Working Group Internet draft, draft-ietf-manetolsr-05.txt, 2000.
[27] C. Perkins, E. Royer, and S.R. Das, "Ad Hoc On-Demand Distance Vector (AODV) Routing," IETF Mobile Ad Hoc Networking Working Group Internet draft, draft-ietf-manet-aodv-13.txt, 2003.
[28] D.B. Johnson, D.A. Maltz, and J. Broch, "DSR: The Dynamic Source Routing Protocol for Multi-Hop Wireless Ad Hoc Networks," *Ad Hoc Networking,* Pearson Education, 2001.
[29] R.C. Shah, A. Wolisz, and J.M. Rabaey, "On the Performance of Geographical Routing in the Presence of Localization Errors," *Proc. IEEE Int'l Conf. Comm. (ICC '05),* 2005.
[30] D. Son, A. Helmy, and B. Krishnamachari, "The Effect of Mobility-Induced Location Errors on Geographic Routing in Mobile Ad Hoc and Sensor Networks: Analysis and Improvement Using Mobility Prediction," *IEEE Trans. Mobile Computing,* vol. 3, no. 3, pp. 233-245, July/Aug. 2004.
[31] J. Sanchez, P. Ruiz, and R. Marin-Perez, "Beacon-Less Geographic Routing Made Practical: Challenges, Design Guidelines, and Protocols," *IEEE Comm. Magazine,* vol. 47, no. 8, pp. 85-91, Aug. 2009.
[32] H. Zhang and H. Shen, "Energy-Efficient Beaconless Geographic Routing in Wireless Sensor Networks," *IEEE Trans. Parallel and Distributed Systems,* vol. 21, no. 6, pp. 881-896, June 2010.