

# Optimizing Control Overhead for Power-aware Routing in Wireless Networks

Anand Seetharam<sup>1</sup>, Bo Jiang<sup>1</sup>, Dennis Goeckel<sup>2</sup>, Jim Kurose<sup>1</sup>, Robert Hancock<sup>3</sup>

<sup>1</sup>Department of Computer Science, University of Massachusetts, Amherst, USA

<sup>2</sup>Department of Electrical and Computer Engineering University of Massachusetts, Amherst, USA

<sup>3</sup> Roke Manor Research, UK

{anand, bjiang, kurose}@cs.umass.edu, goeckel@ecs.umass.edu, robert.hancock@roke.co.uk

**Abstract**—We analyze the tradeoff between the amount of signaling overhead incurred in path selection in a MANET with time-varying wireless channels and the application-level throughput and end-to-end power expended on the selected path. Here, increased overhead increases the accuracy of the link-state estimates used in path selection but decreases the amount of bandwidth available for application use. We develop an information-theoretic, bounding approach to quantify the signaling overhead. Specifically, we investigate (i) the time granularity at which link state is sampled and communicated, and (ii) the minimum number of bits needed to encode this link state information, such that the expected power consumption within a sampling interval is minimized subject to a fixed source-destination goodput constraint. We formulate an optimization problem that provides a numerically computable solution to these questions, and quantitatively demonstrate that short sampling intervals incur significant overhead while long intervals fail to take advantage of the temporal correlation in link state. Additionally, we find that using a small number bits per sample do not provide sufficient information about the network while using too many bits provide little additional information at the expense of increased overhead. Our work can be used by network operators as a tool to determine parameters like the optimal state update frequency and the number of bits per sample.

## I. INTRODUCTION

The overhead of gathering state/control information (e.g., link states, node locations, queue lengths) can be significant in a mobile ad-hoc wireless network (MANET) when bandwidth is limited and network structure and state may change frequently. In such dynamic scenarios, it is still advantageous to collect state information, provided that this information leads to better decisions that more than compensate for the additional overhead incurred. For example, the decrease in available path bandwidth as a result of state gathering overhead may be more than compensated for by the choice of better paths for routing data packets. Efficient bandwidth use is not the only metric of concern in ad hoc networks; since nodes are typically battery powered, minimizing power consumption is also important.

Understanding the tradeoff between the cost incurred in state information collection in a network and the resulting performance is a fundamental, yet largely unexplored problem. In this paper, we analyze this tradeoff between the amount of state information collected (at what precision?, how often?) and overhead incurred, and the resulting performance in wireless networks while providing goodput guarantees. We develop an information-theoretic, bounding approach to analyze the

tradeoff between the amount of signaling overhead incurred in path selection in a MANET with time-varying wireless channels and the application-level throughput and end-to-end power expended on the selected path. Most prior work characterizing the impact of control overhead on performance in wireless networks has relied on simulation [1] or analysis of specific topologies [2]. We take a more abstract, information-theoretic approach to characterize this tradeoff. Our work is closest to Wang *et.al* [3], which adopts an information-theoretic approach to characterizing the overhead of link state routing. We differ from [3] in that we consider the path selection problem and analyze the tradeoff between the signaling overhead (state update frequency and the number of bits per sample) and power consumption in time-varying channels while satisfying goodput constraints.

We consider a network of  $n$  nodes with multiple source-destination pairs. We assume each source has  $m$  disjoint paths to the destination with  $k$  links on each path and that time is divided into intervals. At the beginning of every interval, each source collects ‘noisy’ estimates about the links in the network. By ‘noise’ we refer to the quantization error arising from finite precision representation of link states. The link state estimates in our model characterize the (time-varying) effect of shadowing on the received power. Shadowing is the variation in signal strength at the seconds timescale caused by large objects (e.g. buildings, trees) between the transmitter and receiver and is modeled in the standard fashion as being independent of the distance between the transmitter and receiver.

We use the information-theoretic rate-distortion approach to quantify the noise in the link measurements - as we use more bits to encode time-varying link state, the fidelity of the estimates increase, but the control overhead also increases. Moreover, we assume each source also desires to achieve a fixed amount of goodput, which is defined as the total throughput (including control and data) minus the control overhead. The goal of the source is to select a path  $i$  among the  $m$  paths such that the expected power consumed in the interval is minimized. The problem can be then stated in the following manner.

*At what time granularity should links be sampled and at what rate (bits) should link values be encoded such that the expected power in any interval is minimized subject to a fixed source-to-destination goodput constraint?* We formulate an optimization problem which provides a numerically computable

solution to these questions. The optimization problem takes as input the desired goodput, and leverages the distribution and autocorrelation of the shadowing process to determine the optimum value of the sampling interval and the number of bits per sample such that minimum power is consumed. Our optimization problem is solved off-line and provides network operators a tool for determining optimal operating points (state update frequency, bits per sample).

As expected, our evaluation quantitatively demonstrates that short sampling intervals incur significant overhead while long intervals fail to take advantage of the temporal correlation in link state. We also observe that using a small number bits per sample do not provide sufficient information about the network while using too many bits provides little additional information at the expense of increased overhead. Additionally, we simulate a network with varying link states and compare the performance of the numerical and simulation results.

The rest of paper is organized as follows. In Section II, we discuss related work and then provide a brief overview of rate distortion in Section III. We describe our network model and the optimization problem in Sections IV and V respectively. We then provide a solution for the optimization problem in Sections VI and VII. We present the numerical and simulation results in Sections VIII and IX respectively and finally conclude the paper in Section X.

## II. RELATED WORK

Most prior work has adopted simulation-based techniques to study the overhead of routing protocols in mobile wireless networks. In [1], the authors study via simulation the impact of mobility (i.e. maximum speed of nodes) on the overhead and reliability of AODV and OLSR. Similarly, simulation has been used to study the performance of these protocols in VANETs [4]. The authors compare overhead, packet delivery ratio and delay of OLSR and AODV with respect to data traffic rate, velocity and density of nodes and conclude that globally OLSR outperforms AODV in VANETs. An earlier study [5] done along similar lines for MANETs concludes that none of these protocols outperforms the other; the two protocols complement each other and provide benefits in different domains. Viennot *et. al* [6] provide an analysis of control traffic for reactive and proactive protocols in MANETs. They derive a simple model by considering the average degree per node, the average number of routes created/sec and the number of simultaneous active routes to model the control overhead. To model mobility they consider link breakage and creation while the average length of each route is considered to model the shape of the network. They then compare this analytical approach with simulation results for AODV, DSR and OLSR.

Theoretical studies characterizing the overhead of routing protocols in MANETs has been done by Abouzeid *et. al* [7], [8], [9], [3]. Zhou and Abouzeid [7] mathematically analyze the overhead of the reactive routing protocols and estimate the overhead associated with route discovery and route failure. Their analysis is developed in the context of an unreliable network modeled by: 1) an unreliable Manhattan (i.e., degree 4) grid and 2) a random Poisson point distribution of nodes each having equal coverage radius. They validate

their numerical results via simulations of regular and random topologies. Information-theoretic techniques have been used to obtain lower bounds on memory requirements and routing overhead for hierarchical proactive routing in mobile ad hoc networks in [8]. The authors study the overhead of cluster-based routing protocols in MANETs in [9] and demonstrate the importance of traffic patterns in determining the protocol scalability. In [10] the authors capture the tradeoff in the scaling between the transport capacity and the size of the routing table while the tradeoff between network properties such as connectivity, unpredictability and resource contention and state (control or data or both) information collection has been studied by Manfredi *et. al* [11].

Our work is closest to [3] where the authors use rate-distortion techniques (an information-theoretic approach) for analyzing the protocol overhead of link state MANET routing. They derive lower bounds on the minimum bit-rate at which a node must receive link state information in order to route data packets with a guaranteed delivery ratio. We differ from the above mentioned works because we consider the path selection problem and analyze the tradeoff between the signaling overhead (state update frequency and the number of bits per sample) and power consumption in time-varying channels while providing goodput guarantees.

Power consumption in wireless networks is also a well explored field [12], [13], [14]. In [12] the authors consider the problem of joint routing, scheduling and power control in wireless networks and provide an approximate algorithm with performance guarantees to address it. Liu *et.al* [13] study the optimal power allocation scheme which maximizes the throughput with delay and average power consumption constraints. Network lifetime maximization for an arbitrary data-gathering tree of wireless nodes has been explored in [14] and the authors propose an optimal binary search algorithm for power allocation to achieve the objective. The primary difference between existing literature on power optimization and our work is that we model state gathering overhead/costs and are interested in determining the optimal sampling frequency and number of bits for encoding samples so as to minimize the power dissipation while maintaining a fixed goodput.

## III. BACKGROUND

We begin with a brief overview of information-theoretic rate-distortion theory. A thorough description of this approach is available in [15]. Our goal here is to introduce the reader to this technique and understand its application to our problem.

Rate distortion theory describes the minimum rate (bits) required to achieve a particular distortion, where distortion is defined as the expected distance between a random variable and its reconstruction from its representation in bits (i.e, quantization). The theory also tells us that given a sequence of  $n$  i.i.d. random variables it is possible to achieve a lower rate at a given distortion if we represent the sequence of the  $n$  variables jointly instead of considering them individually.

Let  $X$  be the (source/encoded) alphabet and  $\hat{X}$  be the (receiver/decoded) alphabet. Similarly  $X^n$  and  $\hat{X}^n$  denote the encoded and decoded sequences and  $f$  and  $g$  are the encoding and decoding functions respectively. The distortion

$d(x, \hat{x})$  is a measure of the cost of representing the symbol  $x$  by the symbol  $\hat{x}$ . The distortion between two sequences  $x^n$  and  $\hat{x}^n$  is denoted by  $d(x^n, \hat{x}^n)$  is defined as  $d(x^n, \hat{x}^n) = \frac{1}{n} \sum_{i=1}^n d(x_i, \hat{x}_i)$ . For a given encoding and decoding scheme,  $D = \mathbb{E}[d(X^n, g(f(\hat{X}^n)))]$  where the expectation is calculated over  $X$ .

The rate distortion function  $R(D)$  for an i.i.d. source  $X$  with distribution  $p(x)$  and bounded distortion function  $d(x, \hat{x})$  is equal to the associated information rate distortion function  $R^{(I)}(D)$  and is defined by equation (1).

$$R(D) = R^{(I)}(D) = \min_{p(x|\hat{x}): \mathbb{E}[d(X, \hat{X})] \leq D} I(X; \hat{X}) \quad (1)$$

where  $I(X; \hat{X})$  is the mutual information between  $X$  and  $\hat{X}$  and the minimization is taken over all possible distributions  $p(x|\hat{x})$ .  $R(D)$  thus represents the minimum number of bits required to encode each symbol  $X$ , given that the entire sequence  $X^n$  is encoded. The rate distortion function thus tells us that there exists some  $f$  and  $g$  such that the expected distortion is bounded by  $D$  if rate  $R(D)$  is employed.

#### IV. NETWORK MODEL

In this section we describe our network model and assumptions. We consider a network of  $n$  nodes with multiple source-destination pairs where each source has  $m$  disjoint paths to the destination with  $k$  links on each path. We assume time is divided into intervals of duration  $T_s$  and at the beginning of every interval, each source collects ‘noisy’ estimates about the links in the network.

In our model these link state estimates characterize the (time-varying) effect of shadowing on the received power. Shadowing is assumed to be a lognormally distributed random process (in dB it is normally distributed) [16]. Consider any sampling interval and let  $t$  be a time of interest in that interval,  $0 \leq t < T_s$ . Let us consider the  $i^{\text{th}}$  path and the  $j^{\text{th}}$  link along this path at some time  $t$ .

Let  $L_{ij}(t)$  be the lognormal shadowing process and  $X'_{ij}(t) = 10 \log_{10} L_{ij}(t)$  be its value in dB.  $X'_{ij}(t)$  is assumed to be a stationary Gaussian random process with mean  $\mu = 0$  and autocorrelation function  $R_{X'}(\tau) = \sigma'^2 e^{-\lambda\tau}$  [17]. The autocorrelation coefficient function for any stationary random process  $X'(t)$  may be defined as  $\rho'(\tau) = \frac{R_{X'}(\tau) - \mu^2}{R_{X'}(0) - \mu^2}$ . Thus for the shadowing process, the autocorrelation coefficient function is given by (2).

$$\rho'(\tau) = e^{-\lambda\tau} \quad (2)$$

For ease of analysis we express  $\ln L_{ij}(t) = \frac{\ln 10}{10} X'_{ij}(t) = X_{ij}(t)$  replacing the logarithm to base 10 with the natural logarithm. Hence,  $X_{ij}(t)$  is also Gaussian random process with mean 0 and autocorrelation function  $R_X(\tau) = \sigma^2 e^{-\lambda\tau}$  where  $\sigma^2 = (\frac{\ln 10}{10})^2 \sigma'^2$ . The autocorrelation coefficient function  $\rho(\tau) = \rho'(\tau)$ . The correlation of  $X_{ij}(t)$  indicates how the link state varies during the sampling interval, given its value at the beginning of the sampling interval. Knowledge of

the correlation is essential for computing the expected power spent in an interval.

At the beginning of the sampling interval the source receives  $\hat{X}_{ij}(0)$ , which are ‘noisy’ estimates of  $X_{ij}(0)$ . As  $X_{ij}(0)$  are drawn from a continuous distribution, encoding them exactly will require an infinite number of bits. The ‘noise’ therefore corresponds to the quantization error and thus  $\hat{X}_{ij}(0)$  are finite precision representation of  $X_{ij}(0)$ . The number of bits used to encode the values of  $X_{ij}(0)$  determines the closeness of  $\hat{X}_{ij}(0)$  to  $X_{ij}(0)$ ; thus, the inaccuracy in  $\hat{X}_{ij}(0)$  decrease as more bits are used for encoding. If  $\epsilon$  is the noise or quantization error then,

$$\hat{X}_{ij}(0) = X_{ij}(0) + \epsilon \quad (3)$$

We model  $\epsilon$  as Gaussian noise with mean 0 and variance  $\sigma_e^2$  [18], [19]. We considering that all the link state values are encoded together and sent to the source. We thus use rate-distortion techniques to upper bound  $\sigma_e^2$ . In particular, define the distortion as the squared-error distortion,

$$d(x, \hat{x}) = (x - \hat{x})^2 \quad (4)$$

Then  $\sigma_e^2 = E[(\hat{X}_{ij}(0) - X_{ij}(0))^2] \leq D$ . The rate distortion function  $R(D)$  for any  $N(0, \sigma^2)$  source with squared-error distortion is given in [15]:

$$R(D) = \begin{cases} \frac{1}{2} \log_2 \frac{\sigma^2}{D} & 0 \leq D \leq \sigma^2 \\ 0 & D > \sigma^2 \end{cases} \quad (5)$$

Equation (5) thus represents the minimum number of bits required to encode each shadowing sample. It is also clear that  $\hat{X}_{ij}(0)$  is a Gaussian random variable with mean 0 and variance  $\sigma_D^2$  given by

$$\sigma_D^2 = \sigma_e^2 + \sigma^2 \quad (6)$$

We assume that the path loss and thus the distance between any two pairs of nodes in the network is the same. Later in section VII we discuss how to relax this assumption.

#### V. MINIMUM POWER PROBLEM

In this section we describe the *Minimum Power Problem*. Each source desires a goodput  $G$ . Let  $C_b$  and  $C_t$  be the control overhead and the overall throughput (combined control and data) respectively. Therefore we have  $C_t = G + C_b$ . At the beginning of each sampling interval, the source collects noisy link state estimates. The objective of the source is to minimize the expected power spent in any interval to achieve goodput  $G$ . Based on the noisy link state estimates collected, the source calculates the expected power consumed along each of the  $M$  paths to the destination in that sampling interval. It then selects the path  $i$  for which the expected power consumed is least.

The goal of the *Minimum Power Problem* is to determine  $T_s$  and  $D$  such that over all possible instantiations of link estimates the expected power consumed (for transmitting both control and data) in any sampling interval to achieve a goodput requirement  $G$  is minimized.

Let  $Q_i$  be the expected power dissipated along the  $i^{\text{th}}$  path in a sampling interval, given the sampling interval  $T_s$ ,

the distortion  $D$  and the link state estimates  $\hat{X}_{ij}(0)$  at the beginning of the interval. As the source selects the path which dissipates the minimum expected power in the sampling interval we can formally state the *Minimum Power Problem* in the following manner.

$$\begin{aligned} \text{Objective: } & \min_{T_s, D} \mathbb{E}[\min_i Q_i] \\ \text{subject to the constraint:} & \\ & C_t - C_b = G \end{aligned}$$

## VI. POWER CONSUMPTION AND CONTROL OVERHEAD

In this section, we begin by modeling the transmit power expended along each path needed to achieve a fixed throughput during the sampling interval. We then model the control overhead as a function of the total number of links in the network and the rate distortion function. These models for power, control overhead and shadowing are then used to obtain an approximate solution to the *Minimum Power Problem* in Section VII.

### A. Power Consumption

The transmitted power  $P_i(t)$  along the  $i^{\text{th}}$  path at time  $t$  to achieve a total throughput  $C_t$  (data and control) is obtained by summing the per-link power of each hop. Let  $P_{ij}^W(t)$  be the transmit power on the  $j^{\text{th}}$  link along the  $i^{\text{th}}$  path at time  $t$  when  $W$  and  $B$  are the transmission rate at any node and the available channel bandwidth in  $Hz$  respectively. Let  $d$  denote the distance between any two nodes in the network. Further let us consider a reference distance  $d_0$  and let  $P_t(d_0)$  and  $P_r(d_0)$  be the transmit and received power between two nodes separated by  $d_0$ . Shannon's formula [15] in (7) relates the transmission rate, the shadowing, the AWGN and the power.

$$W = B \log_2 \left( 1 + \frac{P_{ij}^W(t) L_{ij}(t)}{F N_0} \right) \quad (7)$$

where  $N_0$  is the noise,  $F = \frac{P_t(d_0)}{P_r(d_0)} \left( \frac{d}{d_0} \right)^\alpha$  and  $\alpha$  is the path loss exponent. Hence we can transform the above equation in the following manner:

$$P_{ij}^W(t) = \frac{2^{W/B} - 1}{L_{ij}(t)} F N_0 \quad (8)$$

There is a subtle point to be noted here. Although the transmission rate is  $W$ , the source can only achieve a lower throughput  $C_t$ , as the wireless medium is a shared resource - if multiple nodes transmit together, interference and packet loss can occur. We assume that there is a scheduling algorithm that determines the time periods during the sampling interval when each source gets the opportunity to transmit. Each source transmits for only a fixed fraction of time during a sampling interval, e.g., it is allocated a fixed number of transmission slots in an interval. Let  $T_1$  be the amount of time a source transmits in an interval of duration  $T_s$ .

We abstract away the scheduling details and define the scheduling factor as  $S = \frac{T_1}{T_s}$ .  $S$  depends on the scheduling algorithm and the number of nodes and is a parameter in our model. Further, we consider a MANET with fast moving nodes such that  $C_t$  is much smaller than  $W$ . We also note that any

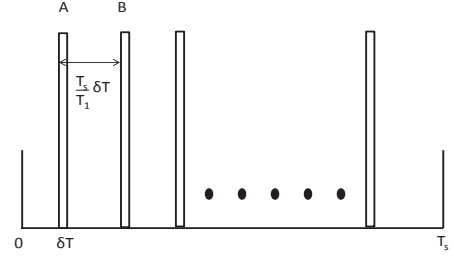


Fig. 1. Power Transmitted in a sampling interval

arbitrary value of  $C_t$  is not achievable, e.g., the achievable  $C_t$  is bounded by results such as the Gupta-Kumar result [20].

Each source transmits for a duration  $T_1$  in a sampling interval. Abstracting away the scheduling details, we assume that the source transmits at rate  $W$  uniformly for small durations ( $\delta T$ ) throughout the sampling interval and the time between two consecutive transmissions is  $\frac{T_s}{T_1} \delta T$ . This is shown in Figure 1. The bars in the figure depict time periods when transmissions take place. This abstract modeling approach is necessary as one cannot assume that the source transmits continuously at rate  $W$  for a duration  $T_1$  in the sampling interval. This would lead to an incorrect estimate of the expected power expended during the sampling interval, because the effects of the correlation of the shadowing process would be incorrectly accounted for if an interval  $T_1$  is considered instead of  $T_s$ .

Our objective is to derive an expression for  $P_{ij}(t)$ , the transmit power on the  $j^{\text{th}}$  link required to achieve a constant throughput ( $C_t$ ) for the entire sampling interval  $T_s$  similar to (8). We model  $P_{ij}(t)$  by:

$$P_{ij}(t) = \frac{a}{L_{ij}(t)} F N_0 \quad (9)$$

The value of  $a$  should be such that the total energy consumed and the total number of bits transmitted in the sampling interval are the same when transmitting at  $W$  for time  $T_1$  and at  $C_t$  for time  $T_s$ . Ensuring that the total number of bits transmitted in both cases are the same, leads to (10).

$$W = C_t \frac{T_s}{T_1} = \frac{C_t}{S} \quad (10)$$

We must also ensure that the total energy consumed is equal. Consider any two consecutive transmission time periods, i.e., points A and B in Figure 1. We assume that as  $\delta T$  is very small, the shadowing value remains constant during the time interval  $\frac{T_s}{T_1} \delta T$ . Therefore we have,

$$\begin{aligned} a &= (2^{W/B} - 1) \frac{T_1}{T_s} \\ &= (2^{C_t/SB} - 1) S \end{aligned} \quad (11)$$

Substituting (11) in (9) we obtain the expression for  $P_{ij}(t)$ . The total power  $P_i(t)$  expended along the  $i^{\text{th}}$  path is the

additive sum of the per-link power of each hop and thus:

$$\begin{aligned} P_i(t) &= \sum_{j=1}^k P_{ij}(t) \\ &= \sum_{j=1}^k \frac{2^{C_t/SB} - 1}{L_{ij}(t)} SFN_0 \end{aligned} \quad (12)$$

### B. Control Overhead

Following [3], we model the minimum overhead for gathering link state information as,

$$C_b = \frac{n(n-1)}{2} \frac{R(D)}{T_s} \quad (13)$$

The rationale behind this abstract model is that the total number of links must be less than  $\frac{n(n-1)}{2}$  (the total number of links is  $O(n^2)$ ), and that a source must know the state of all network links to compute its best path to the destination. Hence, following [3]  $\frac{n(n-1)}{2} \frac{R(D)}{T_s}$  represents the minimum control overhead.

## VII. SOLVING THE OPTIMIZATION PROBLEM

In this section we approximately solve the *Minimum Power Problem*. All Lemmas used in this section are stated in the Appendix. We begin by expressing  $P_i(t)$  (12) as:

$$P_i(t) = \sum_{j=1}^k CY_{ij}(t) \quad (14)$$

where  $C = (2^{C_t/SB} - 1)SFN_0$  and  $Y_{ij}(t) = \frac{1}{L_{ij}(t)}$ . Therefore,  $Y_{ij}(t)$  is also a lognormal random process and we have  $\ln Y_{ij}(t) = -X_{ij}(t)$ .

Recall that  $Q_i$  is the expected power consumed in a sampling interval, given the sampling duration  $T_s$ , the distortion  $D$  and the link state estimates  $\hat{X}_{ij}(0)$ .  $Q_i$  can be formally expressed as,

$$Q_i = \frac{1}{T_s} \int_0^{T_s} \mathbb{E}[P_i(t) | \hat{X}_{i1}(0), \hat{X}_{i2}(0), \dots, \hat{X}_{ik}(0); T_s, D] dt \quad (15)$$

Note that  $T_s$  and  $D$  are model parameters and are not random variables: we thus omit them while expressing conditional expectations. The expression for  $Q_i$  can be rewritten as,

$$\begin{aligned} Q_i &= \frac{1}{T_s} \int_0^{T_s} \mathbb{E}[\mathbb{E}[P_i(t) | \hat{X}_{i1}(0), \dots, \hat{X}_{ik}(0), \\ &\quad X_{i1}(0), \dots, X_{ik}(0)] | \hat{X}_{i1}(0), \dots, \hat{X}_{ik}(0)] dt \\ &= \frac{1}{T_s} \int_0^{T_s} \mathbb{E}[\mathbb{E}[P_i(t) | X_{i1}(0), \dots, X_{ik}(0)] \\ &\quad | \hat{X}_{i1}(0), \dots, \hat{X}_{ik}(0)] dt \end{aligned} \quad (16)$$

The above simplification can be done because given  $X_{ij}(0)$ ,  $P_i(t)$  is independent of  $\hat{X}_{ij}(0)$ , i.e, the underlying process itself does not depend on the observation  $\hat{X}_{ij}(0)$ . We first determine  $H_i = \mathbb{E}[P_i(t) | X_{i1}(0), \dots, X_{ik}(0)]$  which can be done

in the following way (17). At any given time  $t$ ,  $X_{ij}(t) | X_{ij}(0)$  is a Gaussian random variable with mean  $\mu_x(t) = \rho(t)X_{ij}(0)$  and variance  $\sigma_x^2(t) = \sigma^2(1 - \rho^2(t))$  (Lemma 1). Hence at any given time  $t$ ,  $Y_{ij}(t) | X_{ij}(0)$  is a lognormal random variable with mean  $e^{-\mu_x(t) + \frac{\sigma_x^2(t)}{2}}$  (Lemma 2).

$$\begin{aligned} H_i &= C \sum_{j=1}^k \mathbb{E}[Y_{ij}(t) | X_{i1}(0), \dots, X_{ik}(0)] dt \\ &= C \sum_{j=1}^k \mathbb{E}[Y_{ij}(t) | X_{ij}(0)] dt \\ &= C \sum_{j=1}^k e^{-\mu_x(t) + \frac{\sigma_x^2(t)}{2}} dt \\ &= C \sum_{j=1}^k A(t) e^{-\rho(t)X_{ij}(0)} dt \end{aligned} \quad (17)$$

where  $A(t) = e^{\frac{\sigma^2}{2}(1 - \rho^2(t))}$ . Substituting (17) in (16) we have,

$$\begin{aligned} Q_i &= \frac{C}{T_s} \sum_{j=1}^k \int_0^{T_s} \mathbb{E}[A(t) e^{-\rho(t)X_{ij}(0)} | \hat{X}_{ij}(0)] dt \\ &= \frac{C}{T_s} \sum_{j=1}^k \int_0^{T_s} A(t) e^{-\rho(t)\hat{X}_{ij}(0)} \mathbb{E}[e^{\rho(t)\epsilon}] dt \\ &= \frac{C}{T_s} \sum_{j=1}^k \int_0^{T_s} A(t) e^{-\rho(t)\hat{X}_{ij}(0)} e^{\frac{\rho^2(t)\sigma_\epsilon^2}{2}} dt \\ &= \frac{C}{T_s} \int_0^{T_s} A(t) e^{\frac{\rho^2(t)\sigma_\epsilon^2}{2}} \sum_{j=1}^k e^{-\rho(t)\hat{X}_{ij}(0)} dt \end{aligned} \quad (18)$$

Equation (18) uses (3) and the fact that  $\epsilon$  is independent of  $\hat{X}_{ij}(0)$ . Moreover, at any given time  $t$ ,  $\rho(t)\epsilon$  is a Gaussian random variable with mean 0 and variance  $\rho^2(t)\sigma_\epsilon^2$ . Therefore, any given time  $t$ ,  $e^{\rho(t)\epsilon}$  is a lognormal random variable with mean  $e^{\frac{\rho^2(t)\sigma_\epsilon^2}{2}}$  (Lemma 2).

We would like to further simplify the expression for  $Q_i$  and do so using Lemma 3 which approximates the sum of lognormal random variables by a lognormal random variable. In (18), at any given time  $t$ ,  $Y'_{ij}(t) = e^{-\rho(t)\hat{X}_{ij}(0)}$  is a lognormal random variable with mean  $\mu_{y'}(t) = e^{\frac{\rho^2(t)\sigma_D^2}{2}}$  and variance  $\sigma_{y'}^2(t) = (e^{\rho^2(t)\sigma_D^2} - 1)e^{\rho^2(t)\sigma_D^2}$ . Therefore,  $Y'_i(t) = \sum_{j=1}^k Y'_{ij}(t)$  is approximated by a lognormal random variable with mean  $\mu_1(t) = k\mu_{y'}$  and variance  $\sigma_1^2(t) = k\sigma_{y'}^2$ . Let  $Z_i(t)$  be the Gaussian variable corresponding to  $Y'_i(t)$ . From Lemma 3 we can express its variance  $\sigma_z^2(t) = \ln \left[ \frac{e^{\rho^2(t)\sigma_D^2} - 1}{k} + 1 \right]$  and mean  $\mu_z(t) = \ln k + \frac{\rho^2(t)\sigma_D^2}{2} - \frac{\sigma_z^2(t)}{2}$ . Further, let  $A_1(t) = A(t) e^{\frac{\rho^2(t)\sigma_\epsilon^2}{2}} = e^{\frac{\rho^2(t)\sigma_\epsilon^2 + \sigma^2(1 - \rho^2(t))}{2}}$ . We then express (18) as,

$$Q_i = \frac{C}{T_s} \int_0^{T_s} A_1(t) \sum_{j=1}^k Y'_{ij}(t) dt$$

$$\begin{aligned}
&\approx \frac{C}{T_s} \int_0^{T_s} A_1(t) Y_i'(t) dt \\
&= \frac{C}{T_s} \int_0^{T_s} A_1(t) e^{Z_i(t)} dt \quad (19)
\end{aligned}$$

We define  $H' = \min_i Q_i$ .  $H'$  can be expressed as,

$$\begin{aligned}
H' &= \min_i \frac{C}{T_s} \int_0^{T_s} A_1(t) e^{Z_i(t)} dt \\
&> \frac{C}{T_s} \int_0^{T_s} A_1(t) e^{-\max\{-Z_i(t)\}} dt \quad (20)
\end{aligned}$$

The inequality is due to the fact that minimum of a summation is greater than the summation of the minimum. For solving the *Minimum Power Problem* we then need to determine  $\mathbb{E}[H']$  which can be written as,

$$\begin{aligned}
\mathbb{E}[H'] &> \mathbb{E}\left[\frac{C}{T_s} \int_0^{T_s} A_1(t) e^{-\max\{-Z_i(t)\}} dt\right] \\
&> \frac{C}{T_s} \int_0^{T_s} A_1(t) \mathbb{E}[e^{-\max\{-Z_i(t)\}}] dt \quad (21)
\end{aligned}$$

The next step is to determine the distribution of  $U = \max\{-Z_i(t)\}$ . It is clear that  $\{-Z_i(t)\}$  are i.i.d Gaussian random variables with mean  $-\mu_z(t)$  and variance  $\sigma_z^2(t)$ . The maximum of i.i.d Gaussian random variables follows a Gumbel distribution asymptotically, as  $M$  the number of paths goes to  $\infty$  with scaling factor  $a_M = \frac{\sigma_z(t)}{\sqrt{2 \ln M}}$  and location factor  $b_M = \sigma_z(t) \left( \sqrt{2 \ln M} - \frac{\ln \ln M + \ln(4\pi)}{2\sqrt{2 \ln M}} \right) - \mu_z(t)$  respectively (Lemmas 4 and 5). Let us consider the random variable  $V$  such that  $\ln V = U$ .  $V$  follows a log-Gumbel distribution with the same parameters as  $U$  (Lemma 6). From Lemmas 5 and 6 we can say that as  $Z_i(t)$  are Gaussian, the mean of the log-Gumbel distribution exists and it follows a gamma function multiplied by an exponential.

But, we are interested in  $U' = -U$  which follows a negative Gumbel distribution. Define  $\ln V' = U'$ . We adopt a similar approach as Lemma 6 to determine the expectation of  $V'$ . It can be easily shown that  $\mathbb{E}[V'] = e^{-b_M} \Gamma(1 + a_M)$ .

$$\begin{aligned}
\mathbb{E}[H'] &\approx \frac{C}{T_s} \int_0^{T_s} A_1(t) \mathbb{E}[e^{U'}] dt \\
&= \frac{C}{T_s} \int_0^{T_s} A_1(t) \mathbb{E}[V'] dt \\
&= \frac{C}{T_s} \int_0^{T_s} A_1(t) e^{-b_M} \Gamma(1 + a_M) dt \quad (22)
\end{aligned}$$

$\mathbb{E}[H']$  computed from (22) will be approximation to  $\mathbb{E}[\min_i Q_i]$ . The optimization problem thus reduces to  $\min_{T_s, D} \frac{C}{T_s} \int_0^{T_s} A_1(t) e^{-b_M} \Gamma(1 + a_M) dt$ , which can easily be computed numerically.

Equation (22) holds for the equal path loss scenario. But if this assumption is relaxed, the above analysis holds with minor modification until (21) - we only need to model the Gaussian variable  $Z_i(t)$  to take into account the different values of  $C$  for the different links resulting from the unequal path loss assumption. If the *Minimum Power Problem* is to be solved in an unequal path loss scenario, one can obtain the distribution of  $\max\{-Z_i(t)\}$  numerically (which is easy as  $Z_i(t)$  are Gaussian) and then determine  $\mathbb{E}[H']$ . However, note that such a procedure will be computationally expensive.

## VIII. EVALUATION

In this section we present numerical results obtained by solving the optimization problem using (22). We first study the tradeoff between the sampling interval and the number of bits per sample for a specific set of parameters and then proceed to investigate the impact of the various parameters on this tradeoff. We consider a network of 100 nodes with  $G = 75 Kbps$ ,  $B = 10 MHz$ ,  $S = 0.05$  and  $\lambda = \frac{1}{5}$ . The variance of shadowing is 25 dB. Further, we assume  $m = 5$  and  $k = 5$ , i.e., the source has 5 disjoint paths with 5 links each. The results are obtained by increasing the number of bits per sample at a granularity of 0.5. We also use this same configuration when we study the effect of the different parameters on the sampling interval and bits per sample (except for the parameter under investigation).

Figure 2 shows the variation of the transmit power with the number of bits per sample for different values of sampling interval. We observe that with a small number of bits per sample (very little information about the network), the expected power consumed is high irrespective of the length of the sampling interval. In particular, when the number of bits per sample is 0 (equivalent to choosing a path at random), the power consumed is very high. Conversely, when the number of bits per sample is high, the additional information is of marginal use in determining the minimum power path, but the overhead expended in transmitting these control bits is high.

We are interested in obtaining the global minima of the power consumed considering the entire range of the sampling interval and bits per sample. We observe that for the parameter values considered, the optimal value of the sampling interval is 1 second and the number of bits per sample is 1.5. Although the results in Figure 2 is obtained for  $S = 0.05$ , similar figures were obtained for other values of  $S$ . In the throughput range of interest (when  $C_t$  is small), the factor  $(2^{C_t/SB} - 1)$  in (12) linearizes, making the power almost independent of  $S$  and vary linearly with  $C_t$ .

We next study the impact of the various parameters (number of nodes, shadowing correlation ( $\frac{1}{\lambda}$ ), goodput, number of links in a path, number of paths) on the tradeoff between the number of bits per sample and the sampling interval. As these results are obtained by increasing the sampling interval and the

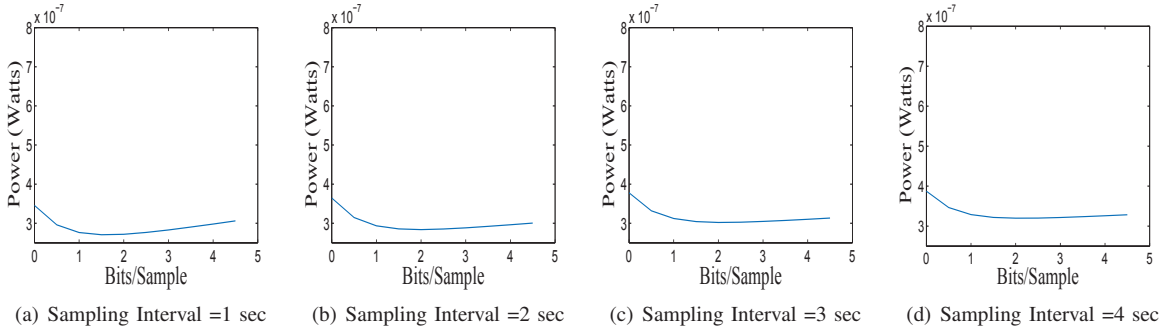


Fig. 2. Numerical: Bits per sample vs. Sampling interval tradeoff

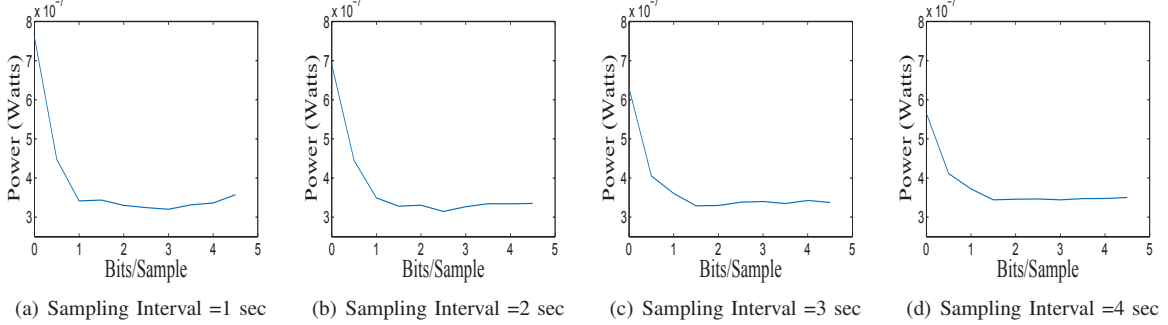


Fig. 3. Simulation: Bits per sample vs. Sampling interval tradeoff

number of bits per sample at a granularity of 0.5, the graphs are discontinuous.

Figures 4(a) and 4(b) show the variation of the number of bits per sample and the sampling interval with the number of nodes. We observe that as the number of nodes increases, the number of bits per sample decrease and the sampling interval increases. This is intuitive since as the number of nodes in the network increases, the control overhead also increases (roughly as  $O(n^2)$ ). Therefore, when the number of nodes is low the optimum decision is to have a small sampling interval (i.e., to sample the network frequently) and encode the samples using a greater number of bits. On the other hand when the number of nodes is large, increased overhead results in the optimum sampling interval being high and number of bits per sample being low. Note that when the number of nodes is very high, the optimal strategy is to select a path at random - this corresponds to the case when the number of bits per sample is equal to zero in Figure 4(a).

We study the variation of the number of bits per sample and the sampling interval with the correlation of the shadowing process ( $\frac{1}{\lambda}$ ) in Figures 4(c) and 4(d) respectively. Figures 4(c) and 4(d) show that both the number of bits per sample and the sampling interval increase with the shadowing correlation. This is because as shadowing correlation increases, the optimal configuration takes advantage of this by sampling at a lower frequency (longer sampling interval). Simultaneously, the number of bits per sample also increases as the decrease in overhead due to a longer sampling interval provides the network an opportunity to gather high fidelity samples.

Figures 5(a) shows that with increasing goodput, the number of bits per sample increases. This is because as the goodput is much larger than the overhead, additional bits can be used to

encode link state values. At the same time, Figure 5(b) shows that as the goodput increases the sampling interval decreases, which can also be attributed to the fact that the overhead is smaller in comparison to the goodput.

In Figures 5(c) and 5(d), we observe that the number of bits per sample and the sampling interval increases with the number of links in a path. As the number of links in a path increases, the probability that at least one of these links is in a bad state (i.e., requiring high power to meet the goodput requirement) increases. Because of the exponential dependence of power on the quality of links, the power consumed along the entire path will be dominated by the bad links. Further as the error in estimating the expected power over a path in an interval increases with the number of links in it, it is advantageous to use more bits for encoding the samples, so that the correct decision is taken i.e., that path with the minimum power is chosen. The increased overhead resulting from the larger number of bits used can then be compensated for by choosing a larger sampling interval.

We also studied the variation of transmit power with the number of paths and found that the number of bits per sample decreases and then becomes zero. As the number of paths increases, the chance of selecting a good path goes up and thus the number of bits per sample decreases. When there are many available paths, selecting a path at random suffices and there is no need to collect state information.

## IX. SIMULATION

In this section we report on our use of simulations using (18) to drive the simulation, to validate our numerical results. Specifically, we study the impact of the inequality in (20) and the two main assumptions of the model - (i) approximating

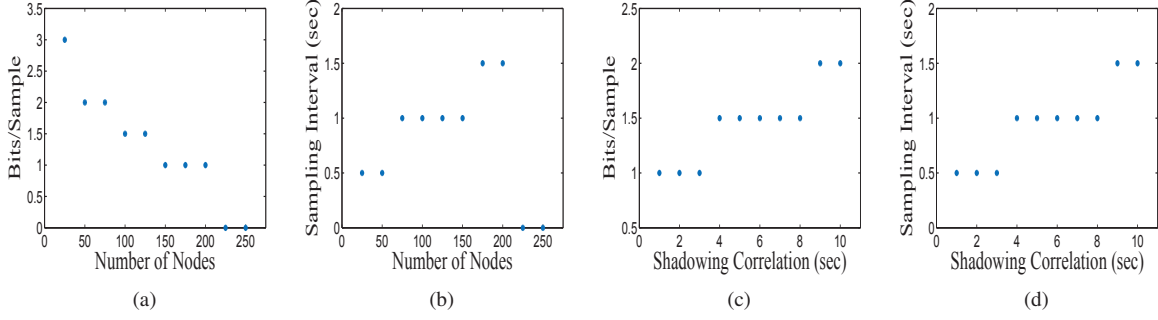


Fig. 4. Variation of Bits/Sample and Sampling Interval with Number of Nodes and Shadowing Correlation ( $\frac{1}{\lambda}$ )

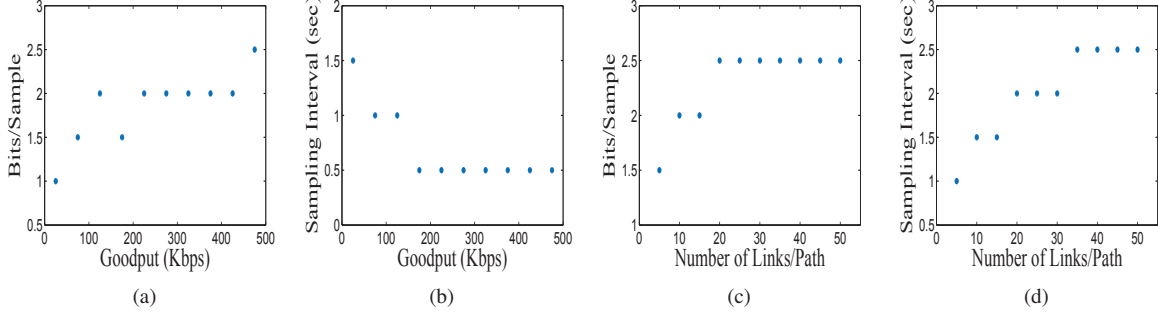


Fig. 5. Variation of Bits/Sample and Sampling Interval with Goodput and Links/Path

the sum of lognormal random variables by a lognormal and (ii) approximating the maximum of a i.i.d Gaussian random variables by a Gumbel distribution - on the accuracy of our numerical results.

We consider the same set of parameters used in the numerical evaluation. For a particular value of sampling interval and number of bits per sample, we generate shadowing measurements for each of the  $k$  links on the  $M$  paths to emulate the link state values collected at the beginning of the sampling interval. We then determine the expected power consumed along each of the  $M$  paths. We select the path for which the expected power consumed for the entire interval is minimum. For each pair of values of sampling interval and number of bits per sample, we repeat this process 500 times to obtain the mean power consumed.

Simulation results depicting the tradeoff between the number of bits per sample and sampling interval with the transmit power are shown in Figure 3 and should be comparable to the numerical results in Figure 2. As in the case of our numerical evaluation, the simulation results also show that the expected power decays rapidly with an increasing number of bits per sample and then begins increasing again.

We note that the power consumption is higher in case of simulation, particularly so for a small number of bits per sample (approaching 0). This is because our numerical analysis is an approximation that becomes better as the number of bits per sample increases. A careful examination of Figures 2 and 3 reveals that when the number of bits per sample is 0, the expected power consumed increases for the numerical evaluation and decreases for simulation with an increasing sampling interval. The intuitive explanation as to why the expected power decreases with an increase in the sampling

interval in case of a real system (i.e, in our simulation) can be explained as follows.

Let us consider for the sake of simplicity that paths are of two types - good and bad; paths are classified as good when the power consumed at the beginning of the sampling interval is low and bad when it is high. The expected power consumed in any sampling interval is thus the additive sum of the conditional expected power consumed given a path of a specific quality (good or bad), multiplied by the probability that the selected path is of the specified quality. The above fact holds true irrespective of the duration of the sampling interval.

Let us next consider the probability of selecting a good or bad path. As shadowing is Gaussian distributed, the probability of any path being good or bad will be same and is independent of the sampling interval. As the number of bits per sample is zero (equivalent to selecting a path at random), the chance of selecting good and bad paths is the same. The next point to note is, because of the exponential dependence of power on the path quality, the expected power expended during a sampling interval is much higher when the selected path is bad in comparison to when it is good.

So far we have only considered the effect of the quality of the path on the expected power consumption. We will now reason about the impact of the sampling interval on the expected power consumption. When the selected path is bad, the expected power expended during the sampling interval will be higher for a shorter sampling interval than for a longer sampling interval since shadowing correlation decays exponentially. Similarly, when a good path is selected, the expected power expended during the sampling interval will be lower for a shorter sampling interval.

But, the positive difference in the expected power expended



between small and large sampling interval when the selected path is bad, is not compensated by the negative difference in expected power expended between them when the selected path is good. Thus, when the number of bits per sample is zero, the expected power consumed when the sampling interval is small is higher than when the sampling interval is long.

Note that, although there is a mismatch between the numerical and simulation results when the number of bits per sample is small, our goal is not to study any specific scenario, but rather to determine the optimal sampling interval and the number of bits per sample. From our simulation, we find that the minimum expected power is consumed for bits per sample=2.5 and sampling interval=2 seconds, which is comparable to the numerical results (bits per sample=1.5; sampling interval=1 second). Hence we conclude that the approximations in Section VII help in modeling the system accurately.

We have also studied the tradeoff between the number of bits per sample and sampling interval for a network with unequal path loss via simulation and observed that a tradeoff similar to the equal path loss case.

## X. CONCLUSION

In this paper, we formulated an optimization problem to determine the frequency at which a source should gather link state estimates and the number of bits used to encode these estimates such that the expected power consumed over a sampling interval is minimized subject to goodput constraints. We observe that long sampling intervals fail to take advantage of the temporal correlation of link state estimates while short sampling intervals incur significant overhead. Similarly, small number of bits per sample provide very little information about the network state while large number of bits provide marginal additional information. Our work can be used by network designers as a tool for determining optimal operating points (state update frequency, bits per sample).

## XI. APPENDIX

*Lemma 1:* Let  $X_1$  and  $X_2$  be two Gaussian random variables with means  $\mu_1$  and  $\mu_2$  and variances  $\sigma_1^2$  and  $\sigma_2^2$  respectively. Let  $\rho$  be the correlation between them. Then,  $X_2|X_1$  is also a Gaussian random variable with mean  $\mu_2 + \rho \frac{\sigma_2}{\sigma_1}(x_1 - \mu_1)$  and variance  $\sigma_2^2(1 - \rho^2)$  [21].

*Lemma 2:* Let  $X$  be a Gaussian random variable with mean  $\mu$  and variance  $\sigma^2$ . Let  $\ln Y = X$ . Then  $Y$  is a lognormal random variable with mean  $\mu_y = e^{\mu + \frac{\sigma^2}{2}}$  and variance  $\sigma_y^2 = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$  [22].

*Lemma 3:* The sum of  $N$  lognormal random variables can be approximated by a lognormal random variable whose first and second moments are equal to the sum of the first and second moments of the  $N$  lognormal random variables [23].

*Lemma 4:* Let  $X_1, X_2, \dots, X_N$  be  $N$  i.i.d random variables. Then  $M_N = \max\{X_1, X_2, \dots, X_N\}$  converges to a Gumbel, Frechet or Weibull distribution depending on the right tail of  $X_i$  [24].

*Lemma 5:* Let  $X_1, X_2, \dots, X_N$  be  $N$  i.i.d Gaussian random variables with mean  $\mu$  and variance  $\sigma^2$ . Then  $M_N =$

$\max\{X_1, X_2, \dots, X_N\}$  converges to a Gumbel distribution with cumulative probability density  $F(x; b_N, a_N) = e^{-e^{-\frac{(x-b_N)}{a_N}}}$ , where  $a_N = \frac{\sigma}{\sqrt{2 \ln N}}$  and  $b_N = \sigma(\sqrt{2 \ln N} - \frac{\ln \ln N + \ln(4\pi)}{2\sqrt{2 \ln N}}) + \mu$  [24]. We observe that  $b_N = \sigma\sqrt{2 \ln N}$  as  $\ln N \rightarrow \infty$ .

*Lemma 6:* Let  $Y$  have a Gumbel distribution with scale parameter  $a$  and location parameter  $b$ . Let  $\ln X = Y$ .  $X$  follows a logGumbel with cumulative distribution function  $F(x; b, a) = e^{-e^{-\frac{(\ln x - b)}{a}}}$ . The  $r^{th}$  moment ( $m_r$ ) of  $X$  is given by  $m_r = (e^b)^r \Gamma(1 - r, a)$  [25].

## REFERENCES

- [1] A. Klein, "Performance comparison and evaluation of aodv, olsr, and sbr in mobile ad-hoc networks," in *IEEE ISWPC*, 2008.
- [2] H. Tripp, R. Hancock, and J. Kurose, "Self-tuning routing algorithms," in *ACITA*, 2010.
- [3] D. Wang and A. Abouzeid, "Link state routing overhead in mobile ad hoc networks: A rate-distortion," in *IEEE Infocom*, 2008.
- [4] J. Haeri, F. Felali, and C. Bonnet, "Performance comparison of aodv and olsr in vanets urban environments under realistic mobility patterns," in *Med-Hoc-Net*, 2006.
- [5] T. Clausen, P. Jacquet, and L. Viennot, "Comparative study of routing protocols for mobile ad-hoc networks," in *MedHoc*, 2002.
- [6] L. Viennot, P. Jacquet, and T. Clausen, "Analyzing control traffic overhead versus mobility and data traffic activity in mobile ad-hoc network protocols," *Wireless Networks*, vol. 10, no. 4, 2004.
- [7] N. Zhou, H. Wu, and A. Abouzeid, "The impact of traffic patterns on the overhead of reactive routing protocols," *IEEE Trans on Selected Areas of Comm*, vol. 23, no. 3, 2005.
- [8] N. Zhou and A. Abouzeid, "Routing in ad hoc networks: a theoretical framework with practical implications," in *IEEE Infocom*, 2005.
- [9] H. Wu and A. Abouzeid, "Cluster-based routing overhead in networks with unreliable nodes," in *IEEE WCNC*, 2004.
- [10] R. Ganti and M. Haenggi, "Throughput versus routing overhead in large ad hoc networks," in *WICON*, 2008.
- [11] V. Manfredi, M. Crovella, and J. Kurose, "Understanding stateful vs stateless communication strategies for ad hoc networks," in *ACM Mobicom*, 2011.
- [12] R. Bhatia and M. Kodialam, "On power efficient communication over multi-hop wireless networks: Joint routing, scheduling and power control," in *IEEE Infocom*, 2004.
- [13] X. Liu and A. Goldsmith, "Optimal power allocation over fading channels with stringent delay constraints," in *ICC*, 2002.
- [14] S. Vasudevan, C. Zhang, D. Geockel, and D. Towsley, "Optimal power allocation in wireless networks with transmitter-receiver power trade-offs," in *IEEE Infocom*, 2006.
- [15] T. Cover and J. Thomas, *Elements of Information Theory*. A John Wiley and Sons, INC., Publication, 2006.
- [16] T. Rappaport, *Wireless Communications: Principles and Practice*. Prentice Hall, 2002.
- [17] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electronic Letters*, vol. 27, no. 23, 1991.
- [18] D. Jimenez, L. Yang, and Y. Wang, "Pcm quantization errors and the white noise hypothesis," *Oversampling and Coarse Quantization for Signals*, University of Maryland, vol. 27, no. 23, 2005.
- [19] J. Mazo, "Quantization noise and data transmission," *The Bell Labs Technical Journal*, 1968.
- [20] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. on Information Theory*, vol. 26, no. 02, 2000.
- [21] H. Derin, "Probability and random processes for engineers," in *Lecture Notes, Dept. Of ECE, Univ. of Massachusetts, Amherst*.
- [22] "Relationships between mean and variance of normal and lognormal distributions," in *Lecture Notes, Dept. Of Civil Engg, MIT*.
- [23] F. Fenton, "The sum of log-normal probability distributions in scatter transmission systems," *IEEE Trans on Communication Systems*, vol. 8, no. 1, 1960.
- [24] A. Bovier, "Extreme values of random processes," in *Lecture Notes, Institute for Applied Mathematics, Bonn Germany*.
- [25] J. Heo and J. Salas, "Estimation of quantiles and confidence intervals for the log-gumbel distribution," *Stochastic Hydrology and Hydraulics*, vol. 10, no. 3, 1996.