Analysis of Per-node Traffic Load in Multi-hop Wireless Sensor Networks

Quanjun Chen, Salil S. Kanhere, Mahbub Hassan

School of Computer Science and Engineering,

The University of New South Wales, Sydney, Australia

Email: {quanc, salilk, mahbub}@cse.unsw.edu.au

UNSW-CSE-TR-0804

February 2008

Abstract

The energy expended by sensor nodes in reception and transmission of data packets makes up a significant quantum of their total energy consumption. Consequently, models that can accurately predict the communication traffic load of a sensor node are critical for designing effective and efficient sensor network protocols. In this paper, we present an analytical model for estimating the per-node traffic load in a multi-hop wireless sensor network, where the nodes sense the environment periodically and forwards the data packets to the sink using greedy geographic routing. The analysis incorporates the idealistic circular coverage radio model as well as a realistic model, log-normal shadowing. Our results confirm that, irrespective of the radio models, the traffic load generally increases as a function of the node's proximity to the sink. In the immediate vicinity of the sink, however, the two radio models yield quite contrasting results. The ideal radio model reveals the existence of a *volcano region* near the sink, where the traffic load drops significantly. On the contrary, with the log-normal shadowing model, the opposite effect is observed, wherein the traffic load actually increases at a much higher rate as one approaches the sink resulting in the formation of a *mountain peak*. The results from our analysis are validated by extensive simulations. The simulations also demonstrate that our results are applicable in more realistic environments, which do not conform to the simplifying assumptions made in the analysis for mathematical tractability.

Index Terms

Sensor networks, Traffic modeling, Greedy routing, Energy efficiency, Analysis

I. Introduction

Wireless Sensor Networks (WSN) are being increasingly used in a variety of applications ranging from health care, environmental monitoring to industrial automation [1]. Ensuring energy efficient operation is critical, especially given that a typical WSN is deployed in remote and unaccessible areas and sensor nodes are equipped with a limited battery source. A typical WSN consists of a large number of nodes deployed over a large area. Hence, packets generated at nodes that are outside the communication range of the base station have to be relayed by other nodes. It has been well accepted that the energy expended in transmission and reception of packets forms a significant component of the total energy budget of a sensor node [1] [2]. Consequently, an analytical model that can accurately estimate the traffic load incurred at each sensor node is instrumental in predicting the energy consumption of the nodes and thus the operational lifetime of the entire WSN. In addition, the traffic load characterization can provide important insights for designing and configuring energy efficient network protocols. As an example, the information about the traffic load of nodes in a collision domain can be used to tune MAC parameters such as slot assignments in TDMA-based schemes and also for setting the wake-up and sleep durations in duty-cycling protocols where nodes periodically go to sleep to conserve energy.In addition, knowledge of the energy expenditure of nodes can be useful in planning deployment and maintenance of the WSN. The network designer can deploy redundant nodes in regions where nodes are expected to expend their energy at a higher rate. Maintenance cycles can also be planned for replacing or charging depleted nodes, thus preventing the formation of coverage holes in the network.

The traffic load of a given sensor node depends on several factors. The first and foremost is the relative distance of the node to the sink. In general, the closer the sensor is to the sink, the greater is the traffic load. This is because the nodes closer to the sink have to relay data packets transmitted by other far-off nodes. The traffic load also depends on the routing protocol employed in the network as it determines the selection of the next hop node for relaying the data towards the sink. Lastly, the characteristics of the environment, which affects the radio communication

behavior of the sensor nodes also has an impact on the traffic load.

Analytical models which can accurately characterize the traffic load of nodes in a WSN are rare. The available models, e.g., those presented in [3], [4], [5], are very simplistic and do not incorporate the specifics of the routing protocol used by the sensor nodes. As such, these models can only be used to provide a rough estimate of the mean traffic load over a relative large geographic area. In [6], Esa Hyytia et. al. have analyzed the traffic load of nodes in a very dense wireless multi-hop network with randomly selected communication pairs. However, their measure of traffic load per unit radio coverage can only provide for a coarse-grained characterization of the traffic load. In addition, they have assumed that the shortest path from any node to the sink can be approximated by a straight line segment, which is true only for highly dense deployments. Furthermore, this work and in fact all other aforementioned efforts have adopted the idealistic circular coverage radio model, which is known to be a poor abstraction of the real communication environment.

In this paper, we take a first step towards developing a detailed and precise analytical model for estimating the per-node traffic load in a WSN. Given that the target domain of WSN is quite broad in terms of the type of applications, networking protocols (routing, MAC, etc) employed and characterization of the communication environment, developing a generalised model is extremely challenging. Thus, we have chosen to focus on a important sub-set within this vast space. With regards to the application load, we focus on periodic monitoring applications wherein the sensor nodes sample the environment periodically and forward the collected data (e.g. temperature, humidity, etc) to the base station. A significant portion of WSN deployed today fall into this category (e.g.: Great Duck Island [7], Redwood Forest [8], etc). With regards to the routing strategy, which is important for selecting the next-hop node, we have considered the popular *greedy routing* forwarding scheme [9], [10], [11], [12], [13]. In greedy routing, a sensor node forwards its packets to a neighbor, which is geographically closest to the sink amongst possible neighbors. In doing so, greedy routing can approximately find the shortest path in terms of hops [9] between a sensor node and the sink. Moreover, greedy routing provides a scalable solution for

large sensor networks, because it requires only local (i.e. one hop neighborhood) information for making forwarding decisions. For developing an analytical model of the traffic load it is necessary to use an appropriate model that abstracts the wireless communication characteristics of a realistic environment. In our analysis we have used both the idealistic radio model and the more realistic log-normal shadowing model, thus enabling us to compare the impact of the two on the results. To the best of our knowledge, this work is the first attempt at developing a comprehensive model for characterizing the per-node traffic load in a WSN.

The analytical model is validated by a rich set of simulations. Our results quantitatively confirm that the traffic load of a node increases with the proximity to the sink. More importantly, we observe a peculiar difference in the traffic load in the immediate vicinity of the sink for the two radio models. For the idea radio model, the traffic load declines significantly after a certain knee point as one moves closer to the sink, resulting in the formation of a volcano shape. On the contrary, with the more realistic log-normal shadowing radio model, we observe the opposite effect, i.e., the traffic load of nodes very close to the sink increases quite significantly as a function of the proximity to the sink, which results in a mountain peak pattern. For simplifying our analysis we have had to make several assumptions. We have also investigated the impact of relaxing some of these assumptions and observed that our analytical results are still valid in these circumstances.

The rest of the paper is organized as follows. In Section II, we provide an overview of our model and list some of the simplifying assumptions made in our analysis. In Section III, we present the analysis of the traffic load for both the ideal and more realistic lognormal radio models. Section IV validates the analytical model by comparing the results with those from simulations. Finally, Section V concludes the paper.

II. Overview of the System Model

For mathematical tractability, we make the following simplifying assumptions:

- The sensor nodes are randomly deployed in an infinite plane area. The node distribution follows a homogenous Poisson point process with a density of ρ sensors per unit area, which

can approximate uniform distribution for large area. This assumption has been widely used in analyzing multi-hop wireless ad-hoc networks [14], [15], [16].

- The sensor nodes are deployed in a circular region with the base station (the sink) situated at the centre (also being assumed in previous work [3], [4], [5]). While this assumption simplifies the closed-form expressions in our analysis, the results can be readily applied to other deployment scenarios (e.g., a sink at the edge of the network) with minor modifications.
- All sensors have identical transceivers and the wireless links are assumed to be symmetric.
- Each sensor node periodically generates a data packet containing the relevant sensed data and routes the packet to the sink. This *clock driven data generation model* is typical for many monitoring applications (e.g., monitoring of temperature, soil moisture, etc.) [17]. Our model can be readily extended for an *event-driven data generation model*, wherein the sensor nodes generate packets in response to certain events of interest. However, for the sake of brevity, we only focus on the former case in this paper.
- The network is dense enough such that the greedy routing always succeed in finding a next hop that advances the data packet towards the sink. In other words, we assume that the forwarding strategy does not encounter a *local minima condition* and thus neglect the effect of planar routing, which is employed in these circumstances.
- Sensor nodes forward packets towards the sink without any data aggregation.
- A forwarding node can always successfully delivery packets to the next hop. Therefore, data packets are never retransmitted (no packet lost, no transmission errors or collisions).

In the first part of our analysis we consider an ideal radio model, wherein the signal attenuation between any two nodes is a function of the Euclidean distance separating the nodes. Consequently, in this idealistic environment, the radio coverage of a sensor node is a perfect circular disc with the radius equal to its radio range. However, in reality, the signal attenuation is not solely dependent on the distance. For example, signal reflection or signal noise can also attenuate the signal. To make our analytical results more realistic, we extend our analysis and incorporate the log-normal

shadowing radio model. This model adds a random signal loss component to the purely distance-dependent signal attenuation, and has been widely used to approximate the real environment [14] [15] [16]. As will be elaborated later, we have observed significant differences in the analytical results with the two models.

Excluding packets that are generated by sensor nodes which can directly communicate with the sink, packets generated by all other node are relayed by intermediate nodes as determined by the routing strategy employed. The behavior of a packets' progress towards the destination is therefore important in determining the traffic load at each sensor. We use a discrete Markov chain to model the hop-by-hop progress of a packet from the source to the destination. The state of the Markov chain is defined as the Euclidean distance (measured in some consistent metric unit, e.g. meters) between the current forwarding node that holds the packet and the destination. Ideally, this distance should be modeled as a continuous random variable. However, to simplify our model, we use a discrete state space to approximately represent the continuous distance values. We quantize the distances resulting in a state space of $(0, \varepsilon, 2\varepsilon, ..., n\varepsilon, ...)$, where the parameter ε is the interval of the state space (i.e. the quantization coefficient). When the interval ε is small enough, the discrete state space approximates the original continuous distance metric.

We elaborate on the state transition of the Markov chain using the example illustrated in Fig. 1, Assume that a packet is currently held by node X as it makes its way towards the base station, node D. Since node X is at a distance i from the destination, the current state for this packet is i. Assume that the next hop node chosen by node X using greedy forwarding is node N, which is at a distance of j from the destination. The packet forwarding operation thus results in a state transition from i to j for the packet. In general, the hop-by-hop progress made by a packet towards the destination can be represented by a series of states that the packet transitions through, eventually culminating in state 0 when the packet reaches the sink.

Intuitively, the sensors closer to the sink will have more packets to transmit, because they have to relay the packets originated from other sensors that are distant. Therefore the traffic load of a

sensor is dependent on its distance to the sink. Since we assume a clock driven data generation model, all sensors generate packets with the same periodicity, which is referred to as a *time unit*. We assume that all packets generated in one time unit in the entire WSN are delivered to the base station before the start of the subsequent time unit. This is a reasonable assumption since in a realistic scenario the end-to-end delay incurred in delivering packets from the sensor nodes to the sink is usually much smaller than the packet generation periodicity which is of the order of a few seconds. The goal of our analysis is to determine the traffic load at a sensor, which is defined as the total packets transmitted by the sensor in one time unit. Let $f(d)$ represent the average traffic load incurred by sensor nodes located at a Euclidean distance of d units from the sink. Note that, $f(d)$ includes the packet generated by the sensor and those forwarded on behalf of other sensors. Consider a sensor node located at distance j from the sink. Since greedy forwarding is employed, this node can only receive packets from its one-hop neighbors that are further from the base station than j. In other words, any other node at distance i, where $i > j$, could possibly forward its packets to this particular node. If we denote the state transition probability from state i to j as $P_{i,j}$, then the traffic load of a sensor at distance j is dependent on the traffic load of sensors at distance i and the state transition probability $P_{i,j}$, where $i > j$. Therefore, in our model, we first analyze the state transition probability of the packets. Using the transition probabilities, we can then recursively calculate the traffic load of the individual sensors. The main symbols used in the rest of this paper are listed in Table I.

III. Analytical Model of Forwarding Overhead

We first evaluate the state transition probabilities assuming the ideal circular disc radio model. Next we extend this to include the log-normal shadowing model. Finally, in the last sub-section, we use these transition probabilities to evaluate the traffic load.

A. Evaluating the State Transition Probability for the Ideal Radio Model

For the ideal radio model, a node can only communicate with other nodes that are located within the circular coverage region of this node. We employ an approach that uses geometric computations and probability theory to prove the following,

Theorem 1: In the context of an ideal radio model, the transition probability of a packet from state i to j when employing greedy routing is,

$$P_{i,j} = \begin{cases} 1 & \text{if } i \leq R \text{ and } j = 0, \\ \exp(-\rho A_{i,j}) - \exp(-\rho A_{i,j+\varepsilon}) & \text{if } i > R \text{ and } i - R \leq j < i, \\ 0 & \text{others,} \end{cases} \tag{1}$$

where $A_{i,j}$ is,

$$A_{i,j} = R^2 \arccos \tfrac{i^2+R^2-j^2}{2iR} + j^2 \arccos \tfrac{i^2+j^2-R^2}{2ij} - \tfrac{\sqrt{(R+i+j)(R+i-j)(R-i+j)(i+j-R)}}{2} \tag{2}$$

Proof: Assume that a packet is currently at node X as it makes its way towards the sink. Let node X be at a distance i from the sink as illustrated in Fig. 2. Consequently the packet is currently in state i. The probability that the packet is forwarded to a sensor at distance j and thus resulting in a transition to state j is the probability that node X finds a neighbor at distance j as the next hop.

We start with a simple case, where $i \leq R$, i.e. the destination node is within radio coverage of the current node X. Hence, as the next hop is the destination, the state i must transition to state 0. Consequently, we have,

$$P_{i,j} = \begin{cases} 1 & \text{if } i \leq R \text{ and } j = 0, \\ 0 & \text{if } i \leq R \text{ and } j > 0. \end{cases} \tag{3}$$

Now let us consider the situation where $i > R$. Recall that, we have assumed that greedy routing can always succeed in finding a next hop node which is closer to the sink. Thus the next hop of node X must has a distance that is less than i from the sink. In other words, the probability that the next hop node lies outside distance region $[i - R, i)$, is zero. Therefore, we have,

$$P_{i,j} = 0, \qquad \text{if } i > R \text{ and } (j < i - R \text{ or } j \geq i) \tag{4}$$

Now, we discuss the more complicated and plausible case where, $i - R \leq j < i$. In greedy routing, if the next hop of node X is at distance j, it implies that at least one neighbor of node

X is at distance j and none of its other neighbors are closer to the destination than j. Thus the transition probability is the probability that at least one neighbor of node X lies on the perimeter of the curve of radius j centered at the destination (see Fig. 2) with no neighbors located to the right of this curve. Since we assume a discrete state space with ε as the interval of the state space, we can approximate the curve as a ring of thickness ε, as illustrated in Fig. 2. Let $R_{i,j}$ represent the region of this thin ring that intersects with the radio range of node X (narrow dark region in Fig. 2). We also denote $A_{i,j}$ as the area of the light shaded region in Fig. 2, which is the intersecting region between the radio coverage of node X and a circle of radius j centered at the destination. $R_{i,j}$ and $A_{i,j}$ are also used to represent the area of each region referred.

Now, the transition probability $P_{i,j}$ is the probability that at least one node lies inside region $R_{i,j}$ and the no nodes are within $A_{i,j}$. Let P_1 be the probability that at least one node is within $R_{i,j}$, and P_2 be the probability that no nodes lie within $A_{i,j}$.

Recall that, we have assumed that the node distribution follows a homogenous Poisson point process with density ρ. As a property of this assumption, the number of nodes in any region of area A follow a Poisson distribution with mean of ρA. Thus the number of nodes in region $R_{i,j}$ follow a Poisson distribution with mean $\rho R_{i,j}$, and the number of nodes in region $A_{i,j}$ has a Poisson distribution with mean of $\rho A_{i,j}$. Note that, the area of $R_{i,j}$ can be computed as $A_{i,j+\varepsilon} - A_{i,j}$. Consequently, we have,

$$P_1 = 1 - \text{Prob(no node in } R_{i,j}) = 1 - \exp(-\rho R_{i,j}) = 1 - \exp(\rho A_{i,j} - \rho A_{i,j+\varepsilon}) \tag{5}$$

$$P_2 = \text{Prob(no node in } A_{i,j}) = \exp(-\rho A_{i,j}) \tag{6}$$

In the Poisson point process, the distribution of the number of nodes in any two disjoint region is independent. Thus $P1$ and $P2$ are independent and we have,

$$P_{i,j} = P_1 \cdot P_2 = \exp(-\rho A_{i,j}) - \exp(-\rho A_{i,j+\varepsilon}) \tag{7}$$

Finally, the area $A_{i,j}$ can be computed using simple geometric calculations and expressed as Equation (2). The details are omitted due to space limitations.

Finally, combining Equations (3), (4), (7) and (2), the theorem is proved. ■

B. Evaluating the State Transition Probability for the Log-normal Shadowing Radio Model

Next, we study a more realistic radio model. In the log-normal shadowing radio model, the signal attenuation between two nodes is dependent not only on the distance separating the two nodes, but also a random signal loss. More formally, given a distance s that separates two nodes, the signal attenuation (in dB) from one node to another one is,

$$\beta(s) = \alpha \log_{10}(\frac{s}{1m}) + \beta_1 \tag{8}$$

where α is a path loss rate, and β_1 is a random variable that follows a normal distribution with zero mean and a standard deviation of σ,

$$f(\beta_1) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{\beta_1^2}{2\sigma^2}) \tag{9}$$

Now two nodes are one-hop neighbors, i.e. they have a direct link between them, only if the signal attenuation between them is less than or equal to a predefined attenuation threshold β_{th}. Thus, for two nodes separated by a distance s, the probability that they have a direct link, denoted as $P_\wedge(s)$, is given by,

$$P_\wedge(s) = Prob(\beta(s) < \beta_{th}) \tag{10}$$

The above equation has been solved by Bettstetter in [14] and the result can be represented by,

$$P_\wedge(s) = \frac{1}{2}\left[1 - erf(\frac{10}{\sqrt{2}\xi} log_{10}\frac{s}{R})\right], \quad \xi = \sigma/\alpha \tag{11}$$

where $R = 10^{\frac{\beta_{th}}{\alpha 10}}$, is referred to as the *average radio range*, which is the maximum distance that permits the existence of a link between two nodes in the absence of signal randomness. The function $erf(.)$ is denoted as follows,

$$erf(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-x^2) dx \tag{12}$$

As an illustrative example, Fig. 3 plots the link probability for the log-normal shadowing model for $R = 50meters$ and different values of the random parameter ξ. Note that, the curve has a longer tail for increasing values of ξ, which implies that a node's radio may cover a larger area for larger ξ. Based on the aforementioned characteristics of the log-normal model, we can have the following theorem,

Theorem 2: In the context of the log-normal shadowing radio model, the transition probability of a packet state i to j when employing greedy routing is,

$$P_{i,j} = \begin{cases} P_{\wedge}(i) & \text{if } j = 0 \text{ and } i > 0, \\ 0 & \text{if } j > 0 \text{ and } j \geq i, \\ \big[1 - P_{\wedge}(i)\big] \cdot \exp(-\pi\rho j^2 P_{\wedge}(A_{i,j})) \cdot \big[-\pi\rho\varepsilon(2j+\varepsilon)\exp(1 - \frac{(j+\varepsilon)^2 P_{\wedge}(A_{i,j+\varepsilon}) - j^2 P_{\wedge}(A_{i,j})}{(2j+\varepsilon)\varepsilon})\big] & \text{others,} \end{cases} \tag{13}$$

Where $P_{\wedge}(i)$ is defined in equation (11), and

$$P_{\wedge}(A_{i,j}) = \int_{i-j}^{i+j} P_{\wedge}(s) f_{i,j}(s) ds \tag{14}$$

$$f_{i,j}(s) = \frac{1}{\pi j^2}(j^2\theta^{'} + 2s\phi - i\sin(\phi) + \frac{s^2 + j^2 - i^2}{2}\phi^{'}) \tag{15}$$

$$\theta' = \frac{2s}{\sqrt{4i^2j^2 - (s^2 + j^2 - i^2)^2}} \qquad \phi = \arccos(\frac{s^2 + i^2 - j^2}{2is}) \qquad \phi^{'} = \frac{is}{i\sqrt{4i^2j^2 - (s^2 + i^2 - j^2)^2}}(\frac{i^2 - j^2}{s^2} - 1) \tag{16}$$

Proof:

We start with the simple case when $j = 0$ and $i > 0$. Assume that a packet is currently in state i, while located at a certain node X. The transition probability of the packet from state i to zero is the probability that there is a direct link between node X and the base station. Thus we have $P_{i,j} = P_{\wedge}(i)$ when $j = 0$ and $i > 0$.

Now let us consider the situation where $j > 0$ and $j \geq i$. Since the next hop of node X must has a distance that is less than i from the sink, the probability that the next hop node lies outside distance region $[0, i)$, is zero. Therefore, we have $P_{i,j} = 0$, when $j > 0$ and $j \geq i$.

In other cases where the next state $j > 0$ and $j < i$, the transition probability $P_{i,j}$ is the multiplication of the following three independent probabilities,

- The probability that no direct link exists between node X and the base station (otherwise the packet can be forwarded to the destination directly), which is $1 - P_{\wedge}(i)$.
- the probability that the node X can find at least one neighbor at distance j, denoted as $1 - P_1$, where P_1 is the probability that there is no neighbor at distance j.
- The probability that no neighbor is within the region that is closer to the base station than j, which is denoted as P_2.

Thus, we have,

$$P_{i,j} = \big[1 - P_{\wedge}(i)\big](1 - P_1)P_2, \quad \text{if } j > 0 \text{ and } j < i \tag{17}$$

Similar to the previous sub-section, we use a ring of thickness ε to represent the curve that is located at distance j from the sink, as illustrated in Fig.4. Let $R_{i,j}$ denote the ring area that has distance j to the destination, and Let $A_{i,j}$ represent the shaded disc shaped region that is closer to the base station than j but does not include the location of the base station itself. Note that, the area included under $R_{i,j}$ and $A_{i,j}$ with the log-normal model is much larger as compared with the ideal radio model in section III-B. The reason being that with the realistic log-normal model the one-hop neighbors of X can located anywhere in the network. On the contrary, in the case of the ideal radio model, the one-hop neighbors are restricted in the circular coverage area of node X. Thus, P_1 is the probability that no direct link exists between X and any node in region $R_{i,j}$, and P_2 is the probability that no direct link exists between X and any node in region $A_{i,j}$.

We first calculate P_1. Since the number of nodes within $R_{i,j}$ is a random variable, according to the law of total probability, we have,

$$P_1 = \sum_{k=0}^{\infty} \Big\{ Prob(\text{k nodes in } R_{i,j}) \cdot \ Prob(\text{no direct link from X to any one of k nodes}) \Big\} \tag{18}$$

According to the Poisson point process, the number of nodes within $R_{i,j}$ have a Poisson distribution with mean of $\rho R_{i,j}$ and are independent with regards to the probability of there existing a direct link with node X. Let $P_{\wedge}(A)$ be the probability that there exists a direct link between node X and a node within area A. Equation (18) can be rewritten as,

$$P_1 = \sum_{k=0}^{\infty} \frac{(\rho R_{i,j})^k}{k!} \exp(-\rho R_{i,j})(1 - P_{\wedge}(R_{i,j}))^k = \exp(-\rho R_{i,j} P_{\wedge}(R_{i,j})) = \exp(-\pi\rho\varepsilon(2j + \varepsilon)P_{\wedge}(R_{i,j})) \tag{19}$$

Similar, for P_2, we can have,

$$P_2 = \exp(-\rho A_{i,j} P_{\wedge}(A_{i,j})) = \exp(-\pi\rho j^2 P_{\wedge}(A_{i,j})) \tag{20}$$

Combining Equations (17) (19) and (20), we have,

$$P_{i,j} = \big[1 - P_{\wedge}(i)\big] \cdot \exp\big[-\pi\rho\varepsilon(2j + \varepsilon)P_{\wedge}(R_{i,j})\big] \cdot \exp\big[-\pi\rho j^2 P_{\wedge}(A_{i,j})\big] \tag{21}$$

Next we compute $P_{\wedge}(R_{i,j})$ and $P_{\wedge}(A_{i,j})$. According to the definition of $A_{i,j}$, the combined region of $R_{i,j}$ and $A_{i,j}$ can be represented by $A_{i,j+\varepsilon}$. Therefore $P_{\wedge}(R_{i,j})$ can be represented by

$A_{i,j}$ and $A_{i,j+\varepsilon}$. In the Poisson point process distribution, given that a node is present within $A_{i,j+\varepsilon}$, the node is uniformly distributed in the region and it is either inside region $R_{i,j}$ or $A_{i,j}$. By the law of total probability, we have,

$$\begin{aligned} P_{\wedge}(A_{i,j+\varepsilon}) &= P_{\wedge}(R_{i,j})Prob(\text{the node is within } R_{i,j}) + P_{\wedge}(A_{i,j})Prob(\text{the node is within } A_{i,j}) \\ &= P_{\wedge}(R_{i,j})\tfrac{(2j+\varepsilon)\varepsilon}{(j+\varepsilon)^2} + P_{\wedge}(A_{i,j})\tfrac{j^2}{(j+\varepsilon)^2} \end{aligned} \tag{22}$$

Equivalently, we have,

$$P_{\wedge}(R_{i,j}) = \tfrac{(j+\varepsilon)^2 P_{\wedge}(A_{i,j+\varepsilon}) - j^2 P_{\wedge}(A_{i,j})}{(2j+\varepsilon)\varepsilon} \tag{23}$$

Combining Equations (21) and (23), we have,

$$P_{i,j} = \big[1 - P_{\wedge}(i)\big] \cdot \exp(-\pi\rho j^2 P_{\wedge}(A_{i,j})) \cdot \big[-\pi\rho\varepsilon(2j+\varepsilon)\exp(1 - \tfrac{(j+\varepsilon)^2 P_{\wedge}(A_{i,j+\varepsilon}) - j^2 P_{\wedge}(A_{i,j})}{(2j+\varepsilon)\varepsilon})\big] \tag{24}$$

Finally, we compute the last unknown variable $P_{\wedge}(A_{i,j})$, i.e., given there exists a node within region $A_{i,j}$, the probability that this node has a direct link with node X. Let $f_{i,j}(s)$ represent the probability that this node is located at distance s from the node X. Based on the law of total probability, we have,

$$P_{\wedge}(A_{i,j}) = \int_{i-j}^{i+j} P_{\wedge}(s) f_{i,j}(s) ds \tag{25}$$

Following rigorous geometric calculations, $f_{i,j}(s)$ is computed as indicated in Equation (15). The detailed derivation is omitted here due to the limited space.

Finally, combining Equations (24), (25) and (15), the theorem is proved. ■

We now provide an example to illustrate the state transition probability, $P_{i,j}$. In this example, we assume the following set of parameters, $R = 50meters, \varepsilon = 1meter, \rho = 0.0019$, the current state of a packet is $i = 100$ and the next state varies from 100 to 0. Fig. 5 illustrates the distribution of the transition probability from state i to the next state j for both radio models under consideration. Note that, the log-normal model reduces to the ideal circular coverage model when the random parameter ξ is equal to zero. For the ideal radio model the peak of the distribution is around $j = 57$ and it reduces to zero for all states beyond 50. This is because of the circular coverage assumption (recall that $R = 50m$). With the more realistic log-normal model the distribution is more spread out over the entire range and the peak shifts towards the right, i.e., closer to the sink. This effect is more pronounced as the random parameter ξ increases. This is because higher the randomness in the signal, the greater is the chance that a node closer to the sink is chosen as the next hop.

Note that, while deriving the state transition probability, we have not necessarily made an assumption about the network shape and the traffic pattern. Hence, the above analysis can be applied to any generic network.

C. Analysis of the Per-node Traffic Load

Based on the state transition probability $P_{i,d}$, we now proceed to calculate the traffic load incurred at the individual sensors. The analysis is independent of the radio model under consideration. One simply has to substitute the appropriate state transition probability equations as derived in the previous two sub-sections for the radio model under consideration. Recall that, the traffic load of a sensor refers to the average number of data packets transmitted by the sensor during one time unit (see definition in Section II). By recursive calculation, we have,

Theorem 3: Consider a circular shaped network of radius l with the sink located at the centre. The traffic load of a sensor located at a distance d from the sink, $f(d)$, is given by,

$$f(d) = \frac{S_t(d)}{\pi(2d+\varepsilon)\varepsilon\rho} \tag{26}$$

where $S_t(d)$ is given by,

$$S_t(d) = \begin{cases} \pi(2d+\varepsilon)\varepsilon\rho & \text{if } d = l, \\ \pi(2d+\varepsilon)\varepsilon\rho + \sum_{i\in(d,l]} P_{i,d} S_t(i) & \text{if } d < l, \end{cases} \tag{27}$$

and $P_{i,d}$ denotes the state transition probability for the radio model under consideration. $P_{i,d}$ for the ideal and log normal radio model have been derived in theorem 1 and theorem 2 respectively.

Proof:

Let $n(d)$ be the average number of sensors located at distance d, and $S_t(d)$ be the total average number of packets collectively transmitted by these $n(d)$ nodes. Then we have:

$$f(d) = \frac{S_t(d)}{n(d)} \tag{28}$$

Recall that, we approximate distance as a discrete state space with an interval of ε. As a result, the number of sensors located at distance d are actually the number of sensors within the thin ring of thickness ε located between the two concentric circles of radius d and $d+\varepsilon$. Since the area of the thin ring is $\pi(2d+\varepsilon)\varepsilon$, the average number of sensors at distance d is given by,

$$n(d) = \pi(2d+\varepsilon)\varepsilon\rho \tag{29}$$

The sensors at distance d transmit all the received packets plus the packets originated by these $n(d)$ sensors in one time unit. Let $S_r(d)$ be the average number of packets received by all sensors at distance d in one time unit. We have,

$$S_t(d) = n(d) + S_r(d) \tag{30}$$

We now proceed to calculate $S_r(d)$. When $d = l$, the sensors at distance l are the furthest sensors from the sink and therefore they are not involved in forwarding packets for other sensors. Therefore $S_r(d) = 0$ and

$$S_t(d) = n(d) \quad \text{if } d = l \tag{31}$$

For $d < l$, as discussed in Section II, any sensor at distance farther than d could forward packets to distance d. In other words, the sensors at distance i, where $i \in (d, l]$ could forward packets to sensors at distance d. For a particular distance i, the average number of packets transmitted from the nodes at i to those at d is the traffic load of sensors at distance i multiplied with the corresponding state transition probability $P_{i,d}$, i.e. $S_t(i)P_{i,d}$. By summing up all cases of i, we have,

$$S_r(d) = \sum_{i \in (d,l]} P_{i,d} S_t(i) \quad \text{if } d < l \tag{32}$$

Combining Equations (30) and (32), we have,

$$S_t(d) = n(d) + \sum_{i \in (d,l]} P_{i,d} S_t(i) \quad \text{if } d < l \tag{33}$$

Finally, combining Equations (28, 29), (31) and (33), the theorem is proved. ■

Note that $S_t(d)$ is a recursive function. Since we know $S_t(l)$, the initial value of $S_t(d)$, we can calculate $S_t(l - \varepsilon)$ according to Equation (27). Similarly, $S_t(l - 2\varepsilon)$ and so on can be derived. Finally, for any given d, $S_t(d)$ can be computed, following which, we can calculate $f(d)$ using Equation (26).

IV. Simulation Results

To validate our analytical model, we first developed a custom C++ simulator, which allows us to evaluate the results for the state transition probabilities and the per-node traffic load. Note that, this simulator conforms to the simplifying assumptions made in our analysis as listed in Section II. In fact, for this validation we have not injected traffic in the network, but rather just made passive observations using the network topology.

Recall that, our analysis assumes that nodes are deployed in an infinite plane and that we focus on a circular sub-region of area πl^2 (l being the radius). To realize this we simulated a square network of area $9 \cdot l^2$ and chose a circular sub-network which lies at the center of this square network. Since the square network is double the size of the circular sub-network, from the view of nodes in the sub-network, the whole network can approximate an infinite plane. A similar approach is also used in [14].

In the simulation scenario we assumed a WSN deployed in a circular region of radius $l = 200m$, with a node density of $\rho = 0.0019$ (averagely 240 nodes) and average radio range of $R = 50m$. We simulated three values of the signal randomness parameter ξ, namely, 0,1, and 2, where 0 represents the ideal radio model and other values represent the log-normal shadowing radio model. For each case of ξ, we ran 5000 simulations and the results presented are averaged over all runs.

For an individual run of the simulation, we randomly draw the total number of nodes M from a Poisson distribution with the mean of $\rho(9 \cdot l^2)$. Then these M nodes are uniformly distributed in the whole square network. Once the nodes are placed, we use the appropriate radio model with the particular value of ξ to generate link connectivity over all pairs of nodes. Then for each node in the selected circular sub-network, we employ greedy routing to find the routing path to the centrally located sink. Once the routing paths are established, we can identify the next hop node for each individual sensor. This enables us to determine the next hop state j for each current state i . Grouping the transitions from all nodes located at distance i from the sink gives us the distribution of the transition probabilities from state i.

Recall that each sensor node generates one data packet during one time unit. In other words, each routing path in the sub-network carries one data packet in one time unit. Therefore, for each node, we count the number of routing paths that traverse through the node, which effectively represent the traffic load at this node. Finally, we group all the nodes that are located at the same distance i from the sink and calculate the average per-node traffic load for that particular state, i.

Fig. 5 compares the results from the simulations with the analytical results for the state transition

probability when the initial state, $i = 100$, $R = 50m$ and $l = 200m$. The graph illustrates that our model is quite accurate in predicting the one-hop transition probabilities. The slight inconsistency is caused by the fact that all forwarding paths lead to a common destination, the sink. As a result, the next hop neighbors of two nodes that are relaying two distinct packets towards the sink are likely to overlap (i.e. include common neighbors). Note that, this contradicts our assumption in the analysis that the neighbors of a particular forwarder are completely independent of the neighbours of other forwarders. The resulting correlation causes the slight deviation from the analytically results, as observed in Fig. 5. Note that, this slight inconsistency was not observed in additional simulations (not included) wherein packets were being forwarded to different destinations placed far apart from each other, thus confirming the above hypothesis.

Fig. 6 plots the average per node traffic load as a function of distance to the sink for a network of radius $l = 200m$ and $R = 50m$. The close match between the simulation and analytical results confirms the validity of the analytical model. As intuitively expected, the closer the node is to the sink, the greater is its traffic load. However, the interesting result revealed by the graph is that the traffic load pattern of nodes close to the sink varies significantly with the radio model. A 3-D plot of the traffic load, as depicted in Fig. 7, illustrates this effect more clearly. Results from the analysis have been used to plot these graphs. In these graphs, the z-axis represents the traffic load and the x and y axes represent the two dimensional cartesian coordinate system with the sink located at the origin (0,0). For the ideal radio model, the traffic load declines significantly if nodes are too close to the sink, giving rise to a *volcanic* shape as illustrated in Fig. 7(a). The volcanic shape of the traffic load is a direct consequence of the circular radio range and the shape of the state transition probability distribution (see Section III-B). The sensors located very close to the sink are rarely chosen as next hop forwarders because once a packet reaches a node that contains the sink within its radio range then the packet can be directly delivered to it. We see that the traffic volcano is clearly contained within a circular area (volcano zone) of radius R, the radio range of the sensor nodes, around the sink. The crater signifies a safe area within the volcano zone where

the sensor nodes do not experience the full force of the traffic volcano.

As evident from Fig. 7(b), the situation is quite different with the log-normal shadowing radio. In contrast with the ideal radio model, the traffic load of the nodes increases continuously as we move closer to the sink creating a *mountain peak* effect. This pattern is caused by the radio signal randomness that is prevalent in the log-normal shadowing radio. Recall that, due to this two nodes that are separated by a distance greater than the average radio range R may still be able to communicate with each other. As a result, the nodes closer to the sink tend to receive a large number of packets from other distant nodes. At the same time the signal randomness does not guarantee that two nodes which may be very close to each other will always be able to communicate. Consequently, even the nodes that are close to the sink may still require other nodes to replay their packets. This effect further compounds the traffic load in the vicinity of the sink, thus resulting in the mountain peak effect.

As discussed earlier, the energy spent in transmission and reception of packets contributes significantly to the energy consumption of a sensor node. Consequently, if one were to plot a graph of the energy consumption of the nodes as a function of their position with respect to the sink, a similar shape would emerge as that observed in Fig. 7. When considering the ideal radio model, nodes located at the edge of the volcano, would expend their energy at a higher rate, which would be determined by the height of the volcano. On the contrary, with the log-normal model, nodes located within the mountain peak closest to the sink would have a significantly high energy consumption, again determined by the height of the peak.

In addition to the above topology-driven simulation, we have also carried out simulations using NS-2 [18]. Unlike the previous simulations, the network now actively carries traffic and we have also relaxed several of the assumptions used in our analysis. This enables us to evaluate if the analytical results are applicable in more realistic scenarios. We now assume a bounded circular network of radius, $l = 200m$ with a centrally located sink and incorporate GPSR [9], a popular greedy forwarding routing strategy. The node density is again 0.0019 and the average radio range

is 50 meters. Further, we do away with the assumption of collision-free medium access and simulate a CSMA/CA MAC protocol (similar to the 802.11 MAC). Packet collisions and the consequent retransmissions were neglected in our analysis. We wanted to investigate if these factors significantly affect the observed traffic load. For this we considered two different traffic loads - (i) low traffic load with sensors transmitting a new packet to the sink every 10 seconds and (ii) high traffic load with the packet generation periodicity reduced to 2 seconds. The results presented are averaged over 3000 runs. As seen from Fig. 8, for low load, irrespective of the radio model, the observed traffic load is very close to the analytical results. At higher load, there is a slight deviation from the analytical results for both radio models. This difference can be attributed to the retransmissions caused due to the packets collisions. However, in both cases, the analytical results still provide for an accurate lower bound.

V. Conclusion

Models that accurately predict the energy consumption of each sensor node make valuable contribution towards designing and configuring energy efficient network protocols, planning efficient network deployments, and extending the operational lifetime of the network. We have proposed an accurate mathematical model that analyzes the per-node communication traffic load, the dominating source of energy consumption, in a multi-hop wireless sensor network. Our results confirm that the traffic load of a node increases with the proximity to the sink. In addition, we discover that the radio characterization model has a significant impact on the traffic load pattern of sensors in the immediate vicinity of the sink. The ideal model leads to a volcano effect whereas the log-normal model causes a mountain peak shape. Our analytical model is validated by extensive simulations. Furthermore, the simulations demonstrate that our analytical results are also valid in realistic scenarios where the assumptions made for the analysis have been relaxed (e.g. packet collisions).

In this analysis we have focussed on WSN employing greedy routing. In our future work, we intend to extend our analysis to include other popular sensor routing protocols. In recent years,

experimental measurements of sensor network radio characteristics have led to the development of more realistic wireless communication abstraction models (e.g. [19]). An interesting direction of future work would be to evaluate the effect of these models on the traffic load of a sensor network.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: a survey. *Computer Networks*, 38:393–422, 2002.

[2] W.R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, Hawaii, USA, 2000.

[3] K. Sha and W. Shi. Revisiting the lifetime of wireless sensor networks. In *SenSys 2004: Proceedings of the 2nd international conference on Embedded networked sensor systems*, pages 299–300, Baltimore, MD, USA, 2004.

[4] J. Li and P. Mohapatra. An analytical model for the energy hole problem in many-to-one sensor networks. In *VTC-2005-Fall: Proceedings of IEEE Vehicular Technology Conference*, volume 4, pages 2721–2725, Dallas, Texas, USA, September 2005.

[5] D. Wang, Y. Cheng, Y. Wang, and D. P. Agrawal. Lifetime enhancement of wireless sensor networks by differentiable node density deployment. In *MASS 2006: Preccedings of IEEE International Conference on Mobile Adhoc and Sensor Systems*, pages 546–549, Vancouver, BC, Canada, October 2006.

[6] Esa Hyytiä and Jorma Virtamo. On traffic load distribution and load balancing in dense wireless multihop networks. *EURASIP Journal on Wireless Communications and Networking*, 2007.

[7] Robert Szewczyk, Alan Mainwaring, Joseph Polastre, John Anderson, and David Culler. An analysis of a large scale habitat monitoring application. In *SenSys '04: Proceedings of the 2nd international conference on Embedded networked sensor systems*, pages 214–226, Baltimore, MD, USA, 2004.

[8] G. Tolle, J. Polastre, R. Szewczyk, D. Culler, N. Turner, K. Tu, S. Burgess, T. Dawson, P. Buonadonna, D. Gay, and W. Hong. A macroscope in the redwoods. In *SenSys '05: Proceedings of the 3rd international conference on Embedded networked sensor systems*, pages 51–63, San Diego, California, USA, 2005.

[9] B. Karp and H. T. Kung. GPSR: Greedy perimeter stateless routing forwireless networks. In *MobiCom 2000: Proceedings of 6th Annual International Conference on Mo-bile Computing and Networking*, pages 243–254, Boston, MA, USA, 2000.

[10] Y. Yu, R. Govindan, and D. Estrin. Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks. UCLA Computer Science Department Technical Report, UCLA-CSD TR-01-0023, May 2001.

[11] F. Kuhn, R. Wattenhofer, and A. Zollinger. Worst-case optimal and average-case efficient geometric ad-hoc routing. In *MobiHoc '03: Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pages 267–278, 2003.

[12] S. Ratnasamy, B. Karp, L. Yin, F. Yu, D. Estrin, R. Govindan, and S. Shenker. GHT: a geographic hash table for data-centric storage. In *WSNA'02: Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, pages 78–87, Atlanta, Georgia, USA, 2002.

[13] H. Luo, F. Ye, J. Cheng, S. Lu, and L. Zhang. TTDD: Two-tier data dissemination in large-scale wireless sensor networks. *Wireless Networks*, 11:161–175, March 2005.

[14] C. Bettstetter. Connectivity of wireless multihop networks in a shadow fading environment. *Wireless Network*, 11(5):571–579, 2005.

[15] R. Hekmat and P. V. Mieghem. Connectivity in wireless ad-hoc networks with a log-normal radio model. *Mobile Networks and Applications*, 11(3):351–360, 2006.

[16] D. Miorandi and E. Altman. Coverage and connectivity of ad hoc networks presence of channel randomness. *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, 1:491–502, March 2005.

[17] Y. Chen and Q. Zhao. On the lifetime of wireless sensor networks. *IEEE Communications Letters*, 9(11):976–978, 2005.

[18] The vint project. The UCB/LBNL/VINT Network Simulator-ns 2. http://www.isi.edu/nsnam/ns/.

[19] M. Z. Zamalloa and B. Krishnamachari. An analysis of unreliability and asymmetry in low-power wireless links. *ACM Trans. Sen. Netw.*, 3(2):7, 2007.

TABLE I

LIST OF MAIN SYMBOLS USED IN THE ANALYSIS

Symbol	Definition
l	Radius of circular network
R	Average radio range of sensors
ρ	Node density, i.e. the number of nodes per unit area
ε	Quantization interval
i, j, d	Euclidean distance from a sensor to the centrally located sink
ξ	Signal randomness parameter in log-normal shadowing radio model
$P_{\wedge}(s)$	The probability that two nodes separated by distance s can communicate with each other
$P_{i,j}$	State transition probability. i.e. the probability that a sensor at distance i from the sink can forward its packets to the sensor at distance j
$f(d)$	Traffic load, i.e. the average number of data packets transmitted by a sensor at distance d during one time unit

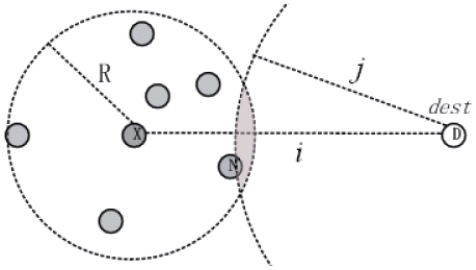


Fig. 1. Example of state transition (from state i to state j)

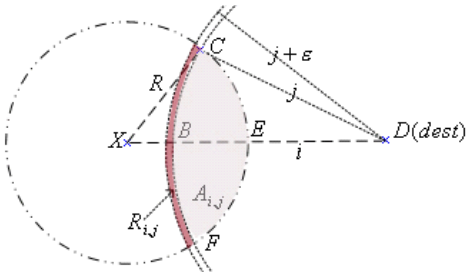


Fig. 2. Illustration used to prove Theorem 1

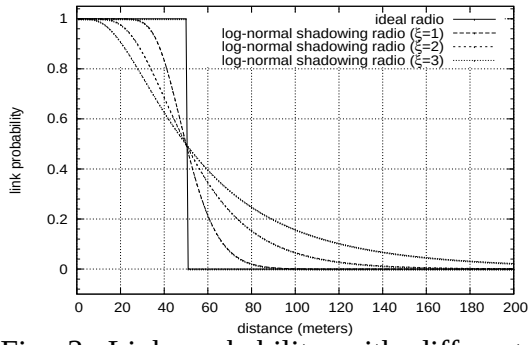


Fig. 3. Link probability with different radio models ($R = 50$)

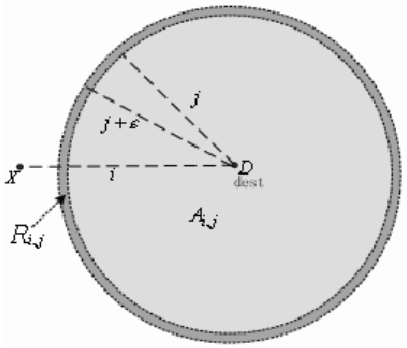


Fig. 4. Illustration used to prove Theorem 2

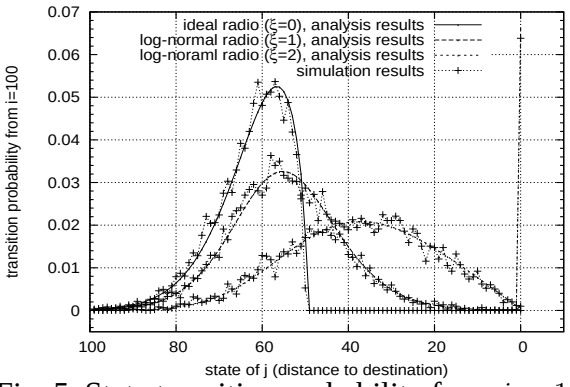


Fig. 5. State transition probability from $i = 100$ ($l = 200,\ R = 50$)

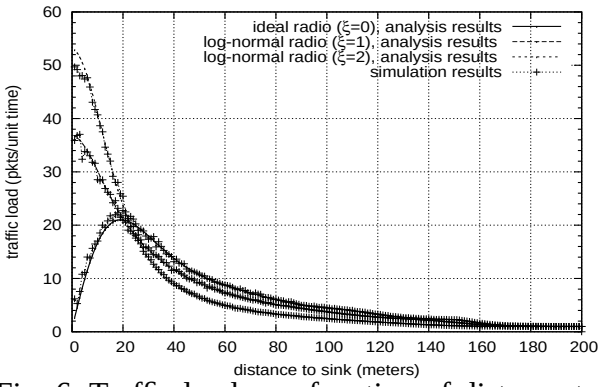
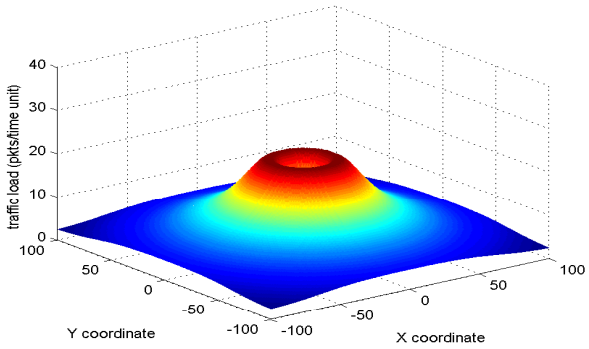
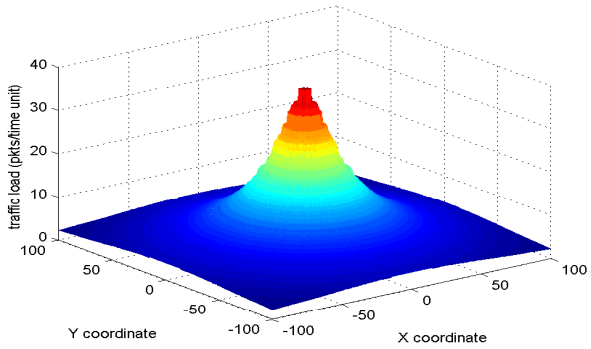


Fig. 6. Traffic load as a function of distance to sink ($l = 200,\ R = 50$)

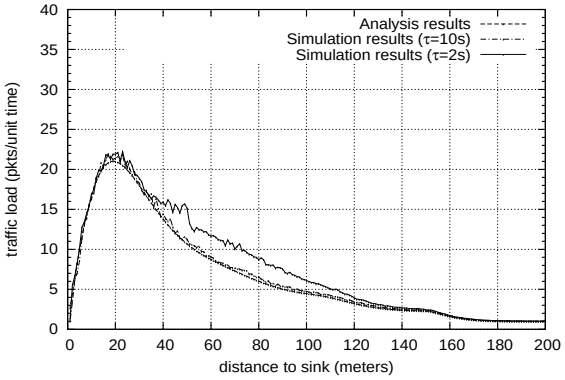


(a) ideal radio

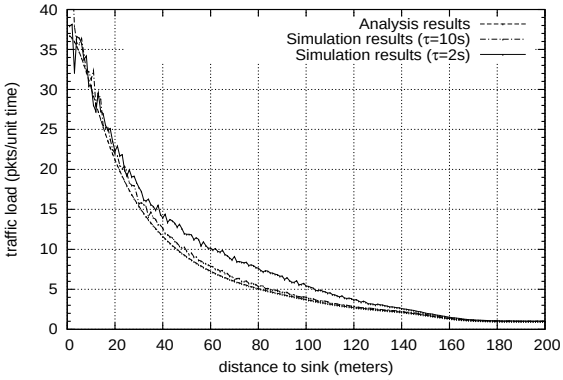


(b) log-normal radio ($\xi = 1$)

Fig. 7. Traffic load as a function of position relative to the sink ($l = 200,\ R = 50$)



(a) ideal radio



(b) log-normal radio ($\xi = 1$)

Fig. 8. Traffic load simulation results using NS-2 ($l = 200,\ R = 50$)