Effect of network parameters on delay in wireless ad-hoc networks

Srihari Narasimhan and Srisankar Kunniyur
Department of Electrical and Systems Engineering
University of Pennsylvania
{srihari,kunniyur}@seas.upenn.edu

August 29, 2004

Abstract

In this paper we characterize the channel access delay and the end-to-end delay experienced by a message in a wireless ad-hoc network. We show that the delay experienced is a function of four network parameters: 1) channel access probability, 2) transmission power or radius, 3) network load and 4) density of nodes. We characterize the effect of each of these parameters on the delay and show that there exists an optimal transmission radius and channel access probability for given load and node density that delivers the best delay performance for the network.

1 Introduction

Ad hoc networks are autonomous systems of devices (nodes) that communicate with each other using wireless links without a fixed infrastructure. In an ad-hoc network, the nodes can be mobile or static. In this paper we consider static ad hoc networks in which nodes are immobile. Ad hoc networks have a very diverse region of application from biological spheres to ubiquitous computing. These systems can support some specific applications including Personal communications like cell phones, laptops, PDA; Group communications such as communication set-up in exhibitions, conferences, presentations, meetings, lectures; Military / emergency / discovery / civil communication. An ad hoc network has several advantages over traditional wireless networks, including ease and speed of deployment and a reduced dependence on a fixed infrastructure. It is attractive because it provides an instant network formation without the presence of fixed base stations and system administrators. At the same time, ad-hoc networks currently suffer from low bandwidths (compared to wired networks) and inefficient channel access due to the distributed nature of the wireless medium. As a result, quality of service (QoS) metrics like delay needs to be explicitly provisioned in ad-hoc networks. Sensor networks are prime examples and form an important class of ad hoc networks. These types of networks are already being deployed in many projects such as environmental sensing, enemy intrusion detection in war zones and traffic monitoring. Sensor networks supporting heterogeneous applications need to support the diverse delay requirements of diverse data. For example, consider a sensor network deployed to sense temperature in a forest. An abnormally high temperature in a particular location may be an indication of a fire. As a result, such messages have more stringent delay requirements than messages reporting temperatures in the normal range. A sensor network monitoring environmental conditions like pressure, temperature and seismic activity is another example of an ad-hoc network supporting heterogeneous data flows between the sources and sinks. Irregularities in the seismic measurements have a stronger delivery constraints that require the network to provide delay differentiation among sensed data. Similarly one can think of many situations where a delay differentiation mechanism must be in place. However, the first step in providing delay differentiation is to understand the effect of various network parameters on delay.

Delays incurred in a wireless network can be attributed to three main sources:

1. *Multi-hop nature of the network:* The multi-hop nature of ad-hoc networks forces a message to traverse several hops to reach the destination. At each hop, the message incurs channel access delay, processing and queuing delays, and aggregation delays. As a result, the delay incurred by a message increases as the number of hops to the destination increases. The design of ad-hoc networks advocates the use of low power for transmissions to maximize throughput and lifetime of the nodes. As a result, the number of hops between the source and destination can be quite large leading to huge delays. The hop count between the source and destination is usually a function of the transmission power at each node. One can typically use a higher transmission power at each node to reach nodes farther away thereby reducing the hop count from the source to the destination.

2. *Channel access delay:* The wireless channel is contention based. As a result, nodes have to contend to gain access to the channel. A popular mode of channel access is the CSMA/CA mode set by the IEEE 802.11 standard. Even though optimized channel access can be designed for specific ad-hoc networks, it is believed that the more common 802.11 access mechanism will be used to create a plug and play environment. As a result each node in the network uses a CSMA/CA protocol to access the channel. The CSMA/CA protocol by design introduces delays in channel access. For example, the 802.11 standard specifies that a node should transmit only when the channel has been idle for a specific amount of time. In the event of a collision, the wait time of the node is increased exponentially. As a result, even if the hop count of the path is low, the total delay experienced by the message might be large due to the channel access delays. Channel access delays are usually functions of load on each node, the node density or the number of nodes in the network and the transmission power.

3. *Aggregation and queuing delays at intermediate nodes:* It is conceivable that nodes in ad-hoc networks aggregate information before transmitting the data to the next hop. This is especially true in sensor networks where data aggregation or compression

(where nodes remove the redundancy in aggregated data) is put forward as a possible mechanism to conserve energy. In addition, data aggregation also helps in improving the overall throughput of the network by reducing the node's channel access. As a result, aggregation and compression at intermediate nodes can lead to large delays in the transmission of a message. Additionally, a message might incur queuing delays at intermediate nodes. Aggregation and queuing delays are a function of the load on the network and the routing protocols used.

The three sources of delay described above are tightly coupled and any one source of delay cannot be optimized individually without impacting the other two. In this paper, we look at the impact of network parameters on the end-to-end and the channel access delay characteristics of an ad-hoc network. We assume that the network provides no aggregation and we ignore queuing delays in this paper. The multi-hop nature of the network forces a message (or a packet) to travel several hops from the source to the destinations. At each hop, the packet incurs MAC access delay. As a result, the end-to-end delay incurred by the packet increases as the number of hops to the destination increases. The channel access delay for a packet arises from the channel contention mechanism of the node. One can employ higher transmission powers to reduce the number of hops to the sources to the destination. However such an increase in transmission power leads to a commensurate increase in the number of neighbors and hence increased channel contention. This therefore increases the channel access delay at each hop. The decrease in the number of hops due to a higher transmission power may still lead to better overall delay performance if the load on the network is not too large or if the number of of new neighbors added is not too big. As a result, there exists a trade-off between the load on each node, the density of nodes and the transmission power in determining the best delay performance of the network. In this paper, we study this trade-off using simple models to capture the delay experienced by a message in terms of the network load, node density and transmission power.

A different type of ad-hoc network where delay plays a big role is in distributed coordination and control of agents over ad-hoc networks. A prime example of such a system is the coordination of groups of mobile autonomous agents using nearest neighbor communications. In such a network, a group of autonomous agents (for example unmanned fighter aircraft) move at the same speed but with different headings. Each agent updates its heading based on its heading and the average of its neighbors' headings. It can be shown that the headings converge to common value under this scheme. However, the speed of convergence depends upon the number of local neighbors and the delay taken to exchange the headings. A small transmission power reduces the channel access delay for the message exchange process while a higher transmission power propagates the information to a larger subset of nodes. As a result, there exists a tension between power control and MAC layer control to provide small delays in the transmission of messages. Such a trade-off is also seen in gossip and belief propagation based schemes for ad-hoc and sensor networks. As a result, understanding the relationship between delay and the network parameters is an important first step in providing delay differentiation.

2 Related work

In the seminal paper of Gupta and Kumar [4] on the capacity of wireless networks, it has been shown that throughput of a network of n nodes is asymptotically equal to $\Theta(\frac{\lambda}{\sqrt{n}})$ where λ is the rate of transmission of the nodes in bits/sec. A conclusion that can be drawn from this result is that the optimal network throughput is obtained at the lowest transmission power that allows connectivity. The intuition behind this is that with smaller transmission ranges, the interference caused is very little. Thus more nodes are able to communicate data more effectively. This however increases the delay as the number of hops required to reach the destination increases. While the above result constrains the total throughput to scale as $\frac{1}{\sqrt{n}}$, Grossglauser and Tse [3] show that a constant $\Theta(1)$ throughput per source-destination pair can be achieved when nodes are mobile. However no delay guarantees are provided. The trade-off between the delay experienced and the throughput possible in the network was shown by Sharma and Mazumdar in [11] and by Gammal et al. in [2]. However, the channel access mechanism and the effect of collisions on delay were not considered in both the analysis.

Many MAC layer protocols (cf. [10,12,13] to name a few) and power control protocols (cf. [1,5–7,9] to name a few) has been proposed in the ad-hoc network literature to provide better throughput and energy efficiency. However, the most common MAC layer protocol is the CSMA/CA protocol proposed in the 802.11 standard of the IETF.

The 802.11 standard defines two different MAC access methods, the Distributed Coordination Function (DCF) and the Point Coordination Function (PCF). The basic access mechanism, called the Distributed Coordination Function, is basically a Carrier Sense Multiple Access with Collision Avoidance mechanism (CSMA/CA). A station wanting to transmit initializes a random countdown timer and senses the medium. If the medium is busy it defers. If the medium is free for a specified time (called Distributed Inter Frame Space (DIFS) in the standard), then the station decrements its timer and repeats this process and transmits the packet when the value of the timer hits zero. A successful transmission is indicated by an acknowledgment from the receiver. If the sender does not receive an acknowledgment for the transmitted packet, it assumes a collision and backs off. The packet is retransmitted till the sender receives an acknowledgment. The packet is discarded after a given number of retransmissions.

3 System Model

In this paper, we will consider two models. In the first model, we will assume that nodes are distributed in a plane using a Poisson distribution with intensity Λ. In the second model, we assume that n nodes are distributed uniformly in a torus[1] of unit area. The first model will relate our results to the node density, while the second model relates our result to the number of nodes.

Transmission model: Each node is able to transmit information to and receive information from nodes within a distance r of it. We will refer to r as the transmission radius

[1]We consider a torus to remove the edge effects that make the analysis more complicated.

or the communicating radius of the network. Known results in percolation theory and spatial random graphs assert that there exists a threshold function for the transmission radius at which network connectivity appears abruptly (cf. Gupta and Kumar [4] for the result in the current framework; for modest tightening of the results and extensions see Kunniyur and Venkatesh [8]). We are mainly concerned with the situation when the network is connected.

Channel access model: The channel access mechanism in the network is assumed to be a CSMA/CA access mechanism. We assume a simplified CSMA/CA channel access mechanism as described below. Time is divided into slots and each node with a packet for transmission contends for the channel at the beginning of a slot with probability p. If a node successfully captures the channel, it transmits for the slot. In other words, we assume that the maximum transmission opportunity (TXOP) is set to one slot[2]. If there is a collision, the node retries the access in the next slot with the same probability p. That is, a node always contends for the channel with probability p. The back-off feature of the 802.11 protocol is not modeled to preserve simplicity[3]. We also consider a model in which a node holds the channel for an exponential amount of time after gaining access to it. However, contention for the channel occurs only at slot boundaries. The parameter p will be referred to as the channel access probability throughout this paper.

Network Throughput: We assume that a throughput of λ is feasible if every node in the network is able to transmit at a rate of λ bits per second [2–4]. We define λ to be the network throughput.

Assume a feasible network throughput of λ with the source-destination pairs separated by an average distance of D units. A typical session must then involve on an average $\frac{D}{r}$ hops from the source to the destination. Therefore, the total load on each node can be approximated by $\lambda\frac{D}{r}$. As a result, the average throughput of λ per node can be supported by the network only if $\lambda\frac{D}{r} \leq P\{\text{Successful transmission}\}$, i.e., the rate at which data is generated at each node is smaller than the rate at which the node can transmit a packet. In this paper we will constrain ourselves to the stable region. Note that a small transmission radius increases the relaying load on each node due to the increased hop count. Such a formulation captures the effect of the transmission radius on the load that can be supported by the network.

Interference model: We assume a model similar to the Protocol model described in [4] for a successful transmission. A transmission from node i to node j is successful if and only if:

- Node i and j are within a distance r of each other, i.e., $d(i, j) \leq r$
- For every other simultaneous transmission from Node k, $d(i, k) \geq (1 + \Delta)r$ for some $\Delta > 0$.

The first condition requires node j to be within the transmission range of node i. The second condition requires a $(1 + \Delta)r$ neighborhood of node i to be silent for a successful

[2]We implicitly assume that the acknowledgment for the transmitted packet is relayed in the same slot.

[3]The channel model considered here closely resembles the SEEDEX MAC protocol proposed by Rozovsky and Kumar in [10].

transmission. In the 802.11 access scheme, a one hop neighborhood of both the transmitter and receiver is made silent. If $\Delta = 0$, this forces only a one hop neighbor of the transmitter to be silent. This serves as a lower bound to the channel access delay experienced. Similarly we can derive an upper bound to the channel access delay experienced by letting $\Delta = 1$ in which we require all nodes in a two-hop neighborhood to be silent. Deriving the exact expression for the access delay is a subject of future research.

4 Delay analysis: Poisson distribution of nodes

Assume that nodes are distributed using a Poisson distribution with intensity Λ. Let r be the transmission radius of each node in the network. Before we derive an expression for the expected channel access delay, we define the stable operation of the network.

Definition 1 *A network is stable if the average rate at which data is generated at each node is smaller than the node's probability of successful transmission, i.e.,*

$$\frac{\lambda D}{r} \leq P\{\textit{Successful Transmission}\}.$$

Lemma 1 *A network is stable for a given* λ, p *and* r *if:*

$$\frac{\lambda D}{r} \leq p\left[e^{-\Lambda\pi f^2 \frac{p\lambda D}{r}} - e^{-\Lambda\pi f^2}\right],$$

where $f = (1+\Delta)r$.

Proof: See Appendix ■

The results presented in this paper assume a stable network operation. The channel access delay (or the MAC delay)[4] is a random variable due to the probabilistic nature of the contention algorithm. However, we can define an expected channel access delay and derive its expression. The following theorem describes the expected channel access delay.

Theorem 1 *Assume that nodes are distributed using a Poisson distribution of intensity* Λ. *Given a channel access probability* p, *network throughput* λ, *and transmission radius* r *the expected channel access delay* $\hat{d}_c$ *in slots is given by:*

$$\mathbb{E}[d_c] = \frac{e^{-\Lambda\pi f^2}}{p}\left[e^{\frac{\Lambda\pi f^2}{1-\frac{p\lambda D}{r}}} - 1\right], \quad (1)$$

where $f = (1+\Delta)r$.

Proof: Consider an arbitrary node i with a packet to transmit. Let $\overline{T_X}$ denote the event that node i does not transmit in a slot. Let

$$P_k(\overline{T_X}) := P(\overline{T_X}|\text{ node } i \text{ has } k \text{ neighbors}).$$

[4]In addition to channel access delay there might be a fixed delay overhead which is ignored in this analysis

¿From our channel access model node i does not transmit in a slot if it does not capture the channel or if a collision occurs when it captures the channel. That is,

$$P_k(\overline{T_X}) = (1-p) + p.P_k(\text{collision}). \tag{2}$$

A collision occurs in a slot when two or more nodes transmit in the same slot. Using the protocol model described in Section 3, a collision occurs when any node within a $(1+\Delta)r$ neighborhood of node i transmits a packet. Note that a node transmits in a slot with probability p only when it has data to transmit. Let S be the number of nodes within a $(1+\Delta)r$ neighborhood of node i that have data to transmit at the beginning of the slot. Conditioning on $S = s$,

$$P_k(\text{collision}|S = s) = (1-(1-p)^s).$$

Taking expectation with respect to S to get rid of the conditioning we have

$$P_k(\text{Collision}) = \sum_{s=0}^{k}(1-(1-p)^s)P(S = s).$$

We assume that data is generated at each node independently [5] with probability $\frac{\lambda D}{r}$. As a result, S has a Binomial distribution with parameters $\frac{\lambda D}{r}$ and k. Therefore,

$$P_k(\text{Collision}) = \sum_{j=0}^{k}(1-(1-p)^j)\binom{k}{j}\left(\frac{\lambda D}{r}\right)^j\left(1-\frac{\lambda.D}{r}\right)^{k-j}.$$

Simplifying the above equation, we get

$$P_k(\text{collision}) = 1-\left(1-p\frac{\lambda.D}{r}\right)^k. \tag{3}$$

Substituting this in (2) we get the probability of unsuccessful transmission in a slot given k neighbors to be

$$P_k(\overline{T_X}) = (1-p) + p\left[1-\left(1-p\frac{\lambda.D}{r}\right)^k\right].$$

Simplifying the above equation yields

$$P_k(\overline{T_X}) = 1-p\left(1-p\frac{\lambda.D}{r}\right)^k. \tag{4}$$

Now let $\hat{q}_k = P_k(\overline{T_X})$. The time it takes node i to access the channel given that node i has k neighbors is a geometric random variable with success probability $(1-\hat{q}_k)$. Therefore,

$$\mathbb{E}_k[d_c] = \frac{1}{1-\hat{q}_k}.$$

[5] In general, data generated at each node is not independent since a collision in a previous slot indicates that at least one node in the interference region has a packet to transmit. Similarly, when a packet is successfully relayed from node i to node k, node k has a packet to transmit in the next slot. However for analytical tractability we make the assumption of independence in this paper.

We then take expectations with respect to the number of neighbors to remove the conditioning. By our protocol model, the number of neighbors is equal to the number of nodes that fall within a interference radii of $f = (1+\Delta)r$. Taking expectations, we get the channel access delay to be

$$\begin{aligned}
\mathbb{E}[d_c] &= \sum_{k=1}^{\infty} \frac{1}{p(1-\frac{p\lambda D}{r})^k}.\frac{e^{-\Lambda\pi f^2}(\Lambda\pi f^2)^k}{k!} \\
&= \frac{e^{-\Lambda\pi f^2}}{p}\sum_{k=1}^{\infty}\frac{(\frac{\Lambda\pi f^2}{(1-\frac{p\lambda D}{r})})^k}{k!} \\
&= \frac{e^{-\Lambda\pi f^2}}{p}\left[e^{\frac{\Lambda\pi f^2}{(1-\frac{p\lambda D}{r})}} - 1\right].
\end{aligned}$$

Thus,

$$\mathbb{E}[d_c] = \frac{e^{-\Lambda\pi f^2}}{p}\left[e^{\frac{\Lambda\pi f^2}{1-\frac{p\lambda D}{r}}} - 1\right].$$

Hence proved. ■

We note that as the channel access probability goes to one, i.e., $p\frac{\lambda D}{r} \to 1$ in (1), the expected channel access delay $\mathbb{E}[d_c] \to \infty$. Such a relationship is expected as increasing the load increases the number of collisions which to a concomitant increase in the channel access delay. Similarly, when the channel access probability given that the node has a packet to transmit goes to zero ($p \to 0$), the expected channel access delay, $\mathbb{E}[d_c] \to \infty$. The effect of load λ on the channel access delay can also be inferred from (1). As the load on each node increases (conditioned on the fact that $\frac{\lambda D}{r} \leq p$,), the expected channel access delay also increases. Similarly, when the load goes to zero, the channel access delay $\mathbb{E}[d_c] \approx \frac{1}{p}$ i.e., the channel access delay is determined by the access probability p.

Remark: As mentioned in Section 3, in a 802.11 setting, a one hop neighborhood of the transmitter and receiver is made silent by the RTS/CTS mechanism. Choosing $\Delta = 0$ and $\Delta = 1$ provides a lower and upper bound respectively to the delay experienced by the message. However, one can approximate the delay experienced by the message where a one hop neighborhood of both the transmitter and the receiver is made silent by letting $\Delta = \sqrt{2} - 1$.

We now describe the total end-to-end delay incurred by a message from its source to its destination L units away.

Lemma 2 *Given a channel access probability* p, *network throughput* λ, *transmission radius* r *and number of nodes* n, *the expected end-to-end delay* (d_t) *in slots that a message experiences from the source to the destination* L *units apart is given by:*

$$\mathbb{E}[d_t] \geq \frac{Le^{-\Lambda\pi f^2}}{pr}\left[e^{\frac{\Lambda\pi f^2}{1-\frac{p\lambda D}{r}}} - 1\right], \tag{5}$$

where $f = (1+\Delta)r$.

Proof: Let η be the number of hops taken by the message. The total end-to-end delay d_t incurred by the message from the source to the destination is given by

$$d_t = \sum_{j=1}^{\eta} d_c(j),$$

where $d_c(j)$ is the channel access delay experienced by the message at the j^{th} hop. Since the transmission radius of each node is r, we have $\eta \geq \frac{L}{r}$. Therefore,

$$d_t \geq \sum_{j=1}^{\frac{L}{r}} d_c(j).$$

We assume that the channel access delay experienced at each hop is independent. Therefore,

$$\mathbb{E}[d_t] \geq \frac{L}{r}\mathbb{E}[d_c].$$

Using Theorem 1, we get

$$\mathbb{E}[d_t] \geq \frac{Le^{-\Lambda\pi f^2}}{pr}\left[e^{\frac{\Lambda\pi f^2}{1-\frac{p\lambda D}{r}}} - 1\right].$$

Hence proved. ■

We now try to untangle the complex relationship between the channel access probability p, the network throughput λ, the transmission radius r, and the intensity of the node distribution in the next few sections. To study the relationship between these parameters we fix all the parameters except one and study its effect on the delay characteristic of the network.

4.1 Effect of p and r on total delay

In this section, we will look at the effect of the channel access probability p and the transmission radius r on the delay incurred by the message as well as on the stable operation of the network. If the data generation rate at each node is greater than the channel access probability, i.e., if $\frac{\lambda D}{r} \geq P\{\text{Successful Transmission}\}$, then the system is unstable. Therefore, in steady state each node always has a message to transmit and the queuing delays at the node goes to infinity. As a result, we will assume that the delay incurred by a message in the unstable region of operation is infinity (since this regime is not applicable).

We will divide the delay experienced by a message into two parts: the channel access delay and the total delay incurred by the message to traverse a distance $L = 1$ unit. We assume the intensity of the Poisson distribution Λ to be 100 and the network throughput λ to be equal to 0.025. We set $\Delta = 0$ in all our results[6]. We fix the channel access probability p and vary the transmission radius of the network. The effect of varying the transmission

[6]In essence, we look at the lower bound on the expected delay experienced by a message

radius on the channel access delay for different values of the channel access probability is shown in Figure 1. In all the figures, the high delay values on the left and right hand side of the graph indicate unstable regions where the total load at each node $\frac{\lambda D}{r}$ is greater than the probability of successful transmission.

¿From Figure 1 we see that for small values of the channel access probability p, the stable region of operation is limited to only high values of the transmission radius r. The intuition behind this is that for small channel access probabilities, the message requires higher number of attempts to successfully gain access to the channel. In order to keep the network stable, the number of hops must be minimized. This is done by having a larger transmission radius r. At the same time, an extremely large transmission radii reduces the probability of successful transmission (due to a bigger neighbor set) and hence the system becomes unstable. These two effects can clearly be seen in Figure 1.

Another conclusion that we can obtain from Figure 1 is that in the stable operating regime the channel access delay increases monotonically as the transmission radius increases. This is due to the fact that an increase in transmission radius adds more neighbors and therefor more collisions. The decrease in the load ($\frac{\lambda D}{r}$) on each node due to the increased transmission radius is not enough to offset the interference brought by the new neighbors.

The total expected end-to-end delay is shown in Figure 2. As seen in Figure 1 we can see the instability region for the different values of p. However, a major difference is that in the stable operating regime the total end-to-end delay is not a increasing function of the transmission radius. To illustrate this more clearly, the left hand portion of the figure is expanded in detail in Figure 3 for each value of p. From Figure 3, we can see that there exists an optimal transmission radius that yields the optimal end-to-end delay. Note that the optimal value of the transmission radius decreases as the channel access probability increases. This is due to the fact that a higher channel access probability leads to more collisions, so the optimal choice of the transmission radius tries to reduce the number of neighbors to reduce the collisions. Too small a transmission radius increases the load on each node which diminishes the effect of the reduced neighbor set. Therefore, the delay exhibits a convex shaped curve as shown in Figure 3.

Similarly, for fixed transmission radii r, the relationship between the end-to-end delay and the channel access probability is shown in Figure 4. Once again we see that there exists an optimal channel access probability that leads to the optimal delay characteristic. Therefore, given the network throughput λ, the intensity of node distribution Λ, the optimal channel access probability p can be computed for a given transmission radius r and similarly an optimum transmission radius can be calculated for a given channel access probability p.

A three dimensional graph showing the total end-to-end delay experienced as a function of both the channel access probability p and transmission radius r is shown in Figure 5. Note that the unstable region delays are set to a constant value of 50 slots. This is done so as to clearly show the valleys where the operation of the network is desired. We can see a narrow valley between two unstable regions where network is stable. The optimal region of operation (smallest delay) also lies in the valley of the delay mesh as shown in Figure 5. We can see that a high transmission radius usually results in a small channel access probability and a high channel access probability results in small transmission

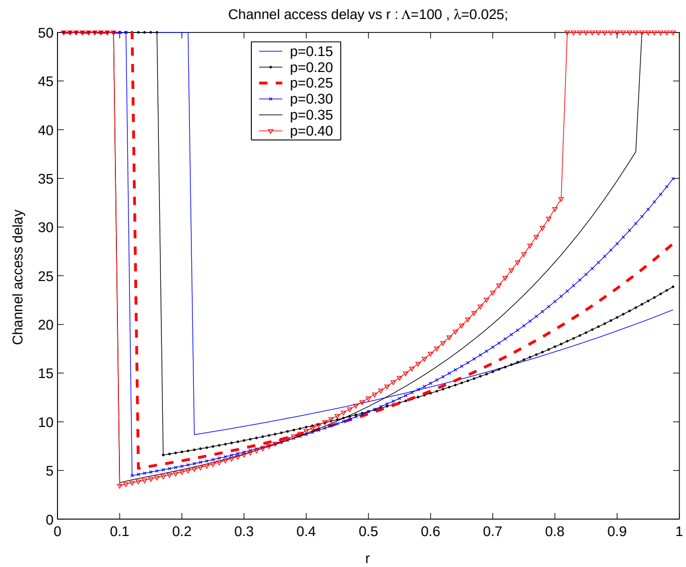

Figure 1: *Expected channel access delay*

radius.

4.2 Effect of node density Λ on the total delay

In the previous section we looked at the effect that the transmission radius and channel access probability had on the channel access delay and the total delay. In this section we look how the intensity of distribution affects the delay characteristics of the network. Figure 6 shows the total end-to-end delay experienced by a message for two different values of the channel access probability when nodes are distributed with intensity $\Lambda = 50$. Here again, the unstable regions are indicated by a flat constant high value. Figure 7 shows a closer look at Figure 6. We can once again see that there exists an optimal transmission radius for a given channel access probability that provides the best delay performance. However, the optimal transmission radius does not lead to a significant delay improvement. This is due to the fact that the intensity and the load on each node is small enough that the delay is not sensitive to the choice of the transmission radius or the channel access probability. Note that the optimal transmission radii is bigger than the optimal transmission radii when $\Lambda = 100$. This is due to the fact that when the transmission radii is increased, the increase in the number of neighbors is smaller in this case since the intensity is smaller. As a result, using a higher power to reduce the number of hops make more sense in this case. Similarly, Figure 8 shows the relationship between the end-to-end delay and the channel access probability for a fixed transmission radius. Comparing this with Figure 4 where $\Lambda = 100$ we can see that the optimal channel access probability occurs at a higher value. Once again we can explain this by pointing to the fact that a smaller node density leads to smaller number of neighbors thereby accommodating a higher channel access probability. A three dimensional graph showing the total end-to-end delay experienced as a function of both the channel access probability p and transmission radius r is

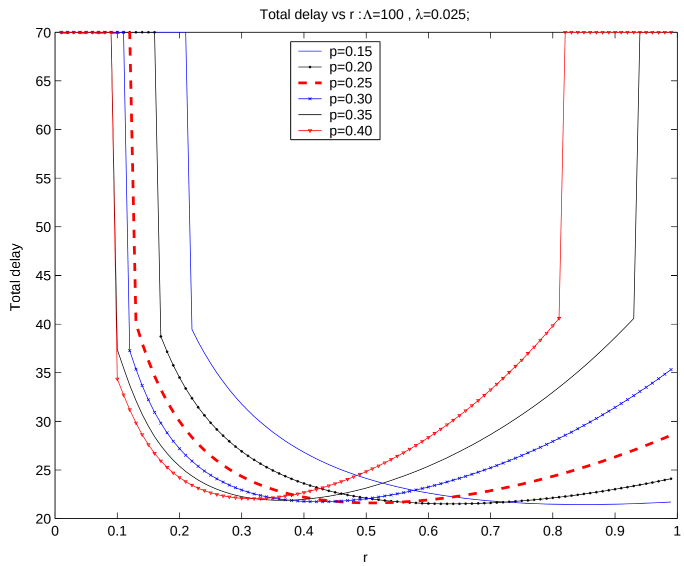

Figure 2: *Expected end-to-end delay*

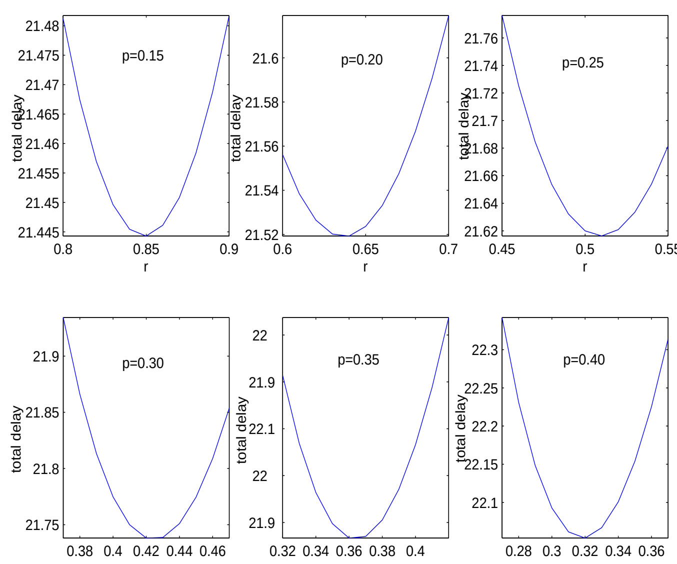

Figure 3: *Trade-off between transmission radius and total end-to-end delay*

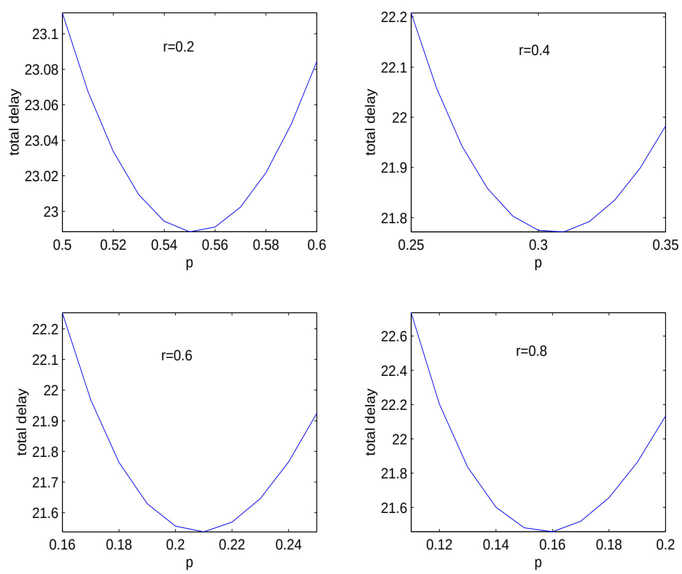

Figure 4: *Trade-off between channel access probability and total end-to-end delay*

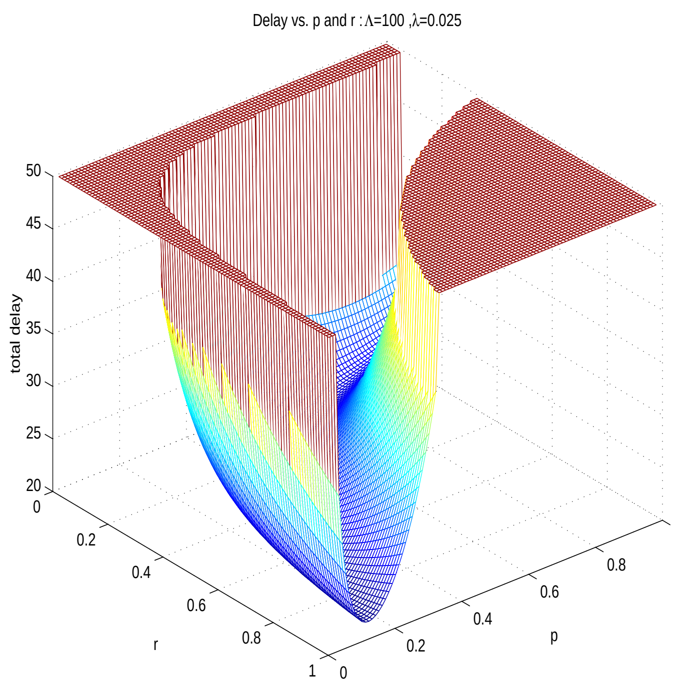

Figure 5: *End-to-end delay as a function of the transmission radius and the channel access probability*

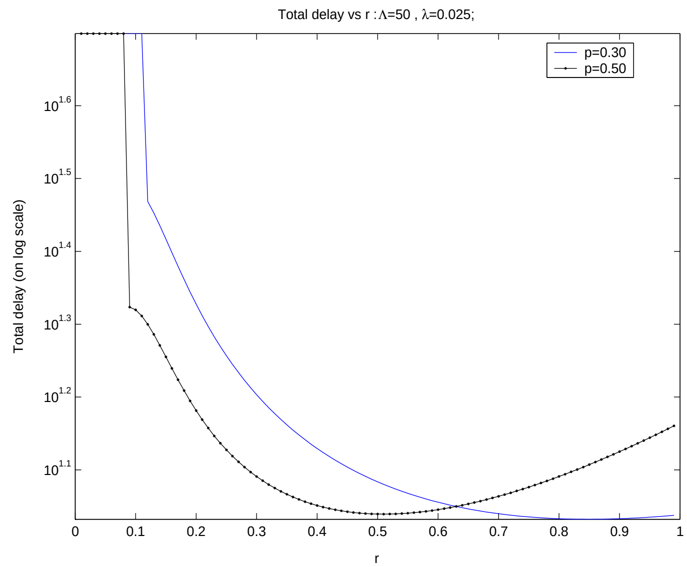

Figure 6: *Expected end-to-end delay when* $\Lambda = 50$

shown in Figure 9. We can see that the stable region of operation (valley between the two plateaus) is larger in this case. The intuition behind this is that as node density decreases, the channel contention reduces thereby leading to an increased region of operation. The relationship between the total end-to-end delay, the channel access probability and the transmission radius when $\Lambda = 150$ and $\Lambda = 200$ is shown in Figures 10 and 11. We can see that the stable operating region becomes more and more constrained as the intensity of node distribution increases till there exist no stable region[7] when the the intensity $\Lambda = 200$.

4.3 Effect of λ on the total delay

In the previous sections we looked at the effect of transmission radius r, channel access probability p and the intensity of node distribution on the total delay. In this section we will look at the effect that the network throughput λ has on the total delay. We fix $\Lambda = 100$ in all the simulations. Graphs showing the total end-to-end delay experienced as a function of both the channel access probability p and transmission radius r are shown in Figure 12 when $\lambda = 0.01$ and in Figure 13 when $\lambda = 0.03$. Comparing Figures 12, 5 and 13, we can see that the the stable region of operation for low delay (valley between the two plateaus) decreases as the network throughput increases. The intuition behind this is that with increasing network throughput, more nodes contend for the channel leading to increased contention.

[7]This is shown by a flat surface with a high delay value of 70 units in Figure 11.

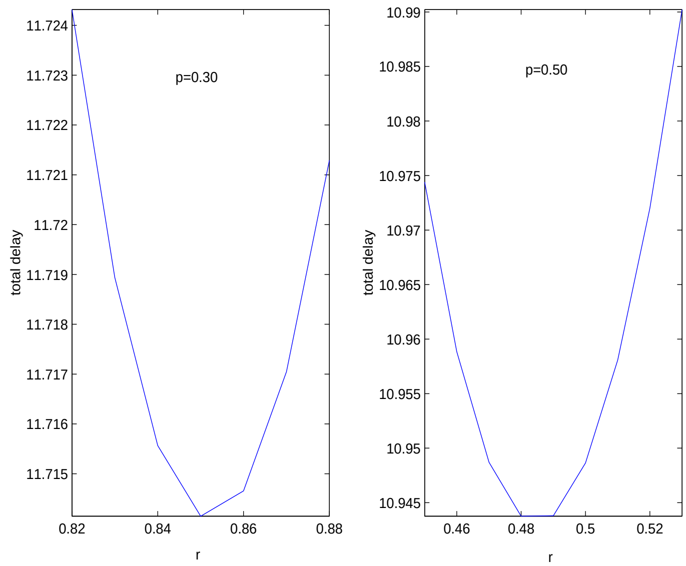

Figure 7: *Trade-off between transmission radius and total end-to-end delay when* $\Lambda = 50$

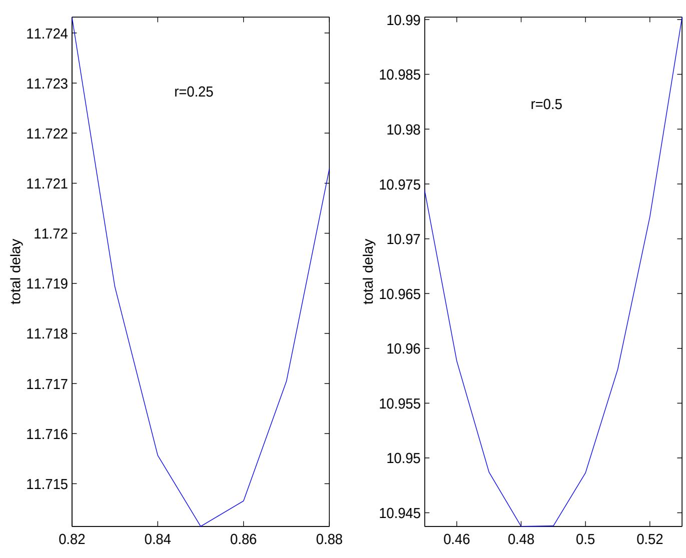

Figure 8: *Trade-off between channel access probability and total end-to-end delay when* $\Lambda = 50$

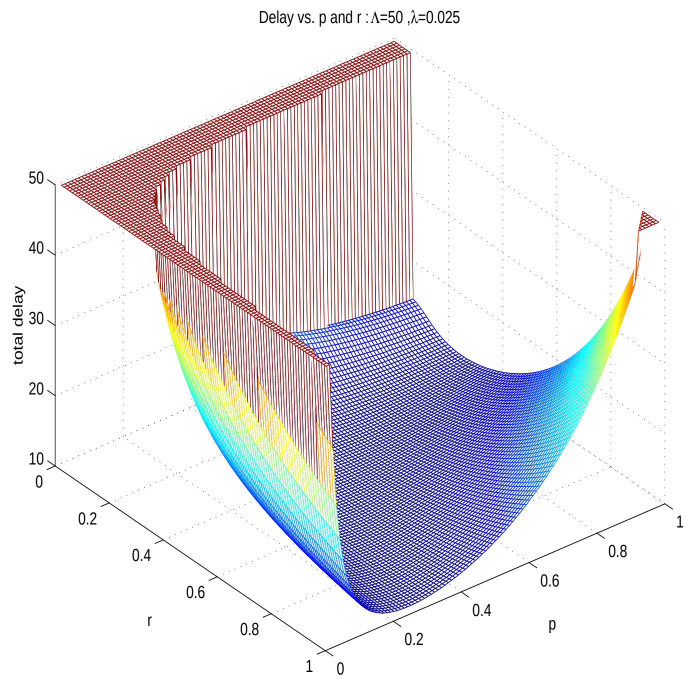

Figure 9: *End-to-end delay as a function of the transmission radius and the channel access probability when* $\Lambda = 50$

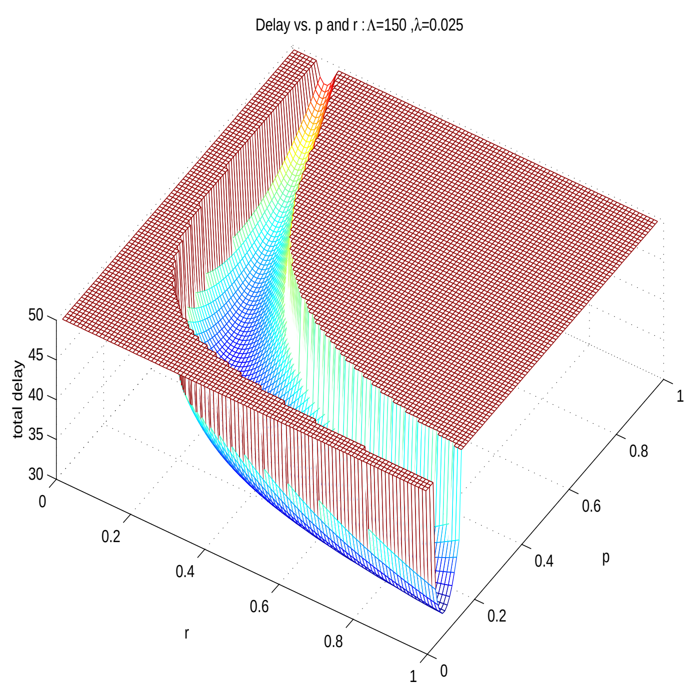

Figure 10: *End-to-end delay as a function of the transmission radius and the channel access probability when* $\Lambda = 150$

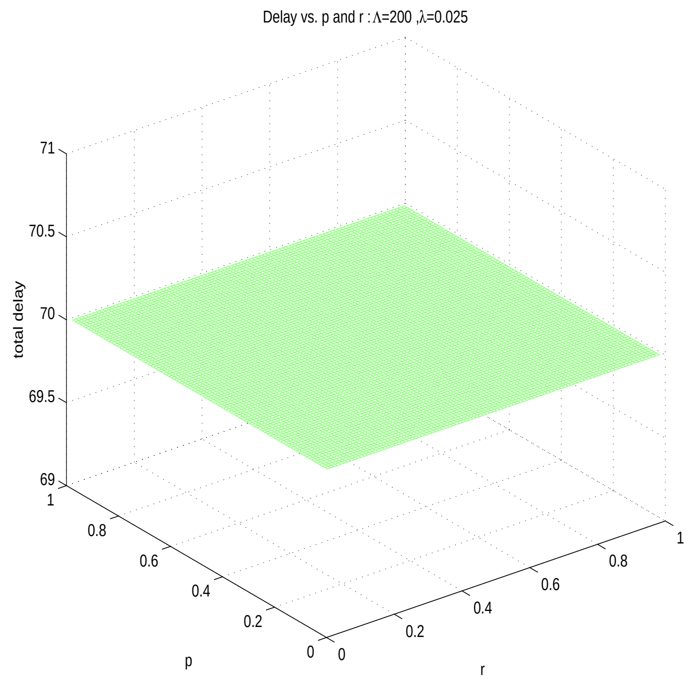

Figure 11: *End-to-end delay as a function of the transmission radius and the channel access probability when* $\Lambda = 200$

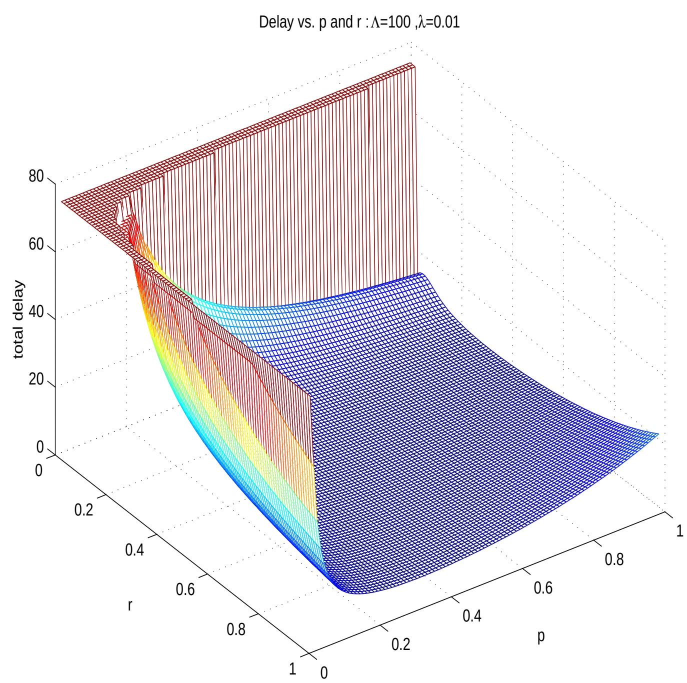

Figure 12: *End-to-end delay as a function of the transmission radius and the channel access probability when* $\lambda = 0.010$

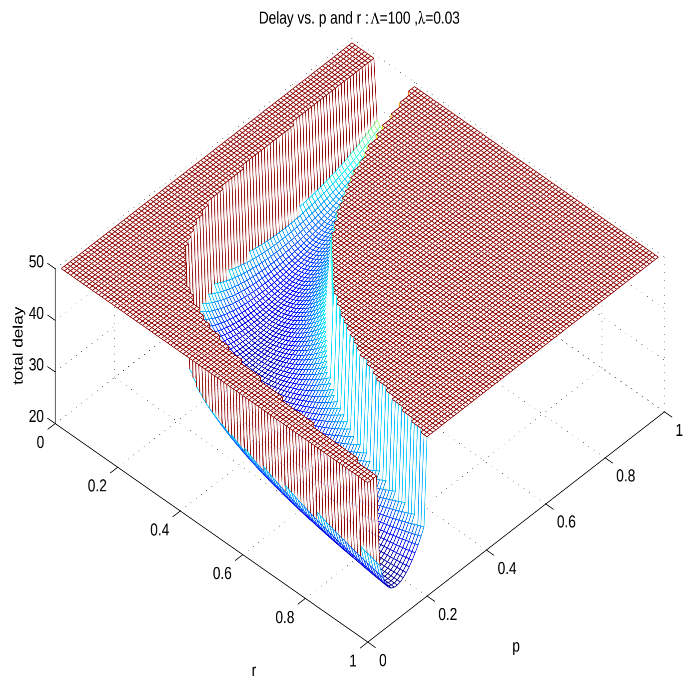

Figure 13: *End-to-end delay as a function of the transmission radius and the channel access probability when* $\lambda = 0.030$

4.4 Exponentially distributed packet sizes

In the last section we had assumed that the transmission time equals one slot. In this section we generalize the previous result when transmission frames are exponentially distributed. We assume that once a node gets access to the channel, it transmits for a duration that is exponentially distributed with mean $\frac{1}{\mu}$. However, nodes contend for the channel only at the beginning of the next slot.

Lemma 3 *Given a channel access probability* p*, network throughput* λ*, and an exponentially distributed transmission times with mean* $\frac{1}{\mu}$*, the probability that node* i *finds a slot busy given that node* i *has* k *neighbors is given by:*

$$\hat{q}_k = \frac{kz(1-z)^{k-1}e^{-\mu}}{1 - e^{-\mu} + kz(1-z)^{k-1}e^{-\mu}}, \tag{6}$$

where $z = p\frac{\lambda D}{r}$.

Proof: Condition on k neighbors. Let $q_k(m)$ denote the probability that the channel is busy at the start of slot m. Let T be an exponentially distributed $r.v.$ with mean $\frac{1}{\mu}$. We know that slot m is busy if slot $m-1$ has an ongoing transmission that continues in slot m or if any of node i's neighbor captures the channel in slot $m-1$ and has a transmission that continues in slot m. Simply put,

$$q_k(m) = \left\{ \begin{array}{l} q_k(m-1)P(T>1) + (1 - q_k(m-1)) \\ P\{\text{one successful} T_X\} P(T>1) \end{array} \right. . \tag{7}$$

Note that we exploit the memoryless property of the exponential distribution to arrive at the above expression. We assume that the data is generated at each node independently with probability $\frac{\lambda D}{r}$. Therefore, the probability that a node will contend for the channel in any given slot is given by $z = p\frac{\lambda D}{r}$, where p is the channel access probability. Therefore conditioned on the fact that node i has k neighbors, we have

$$P\{\text{one successful transmission}\} = kz(1-z)^{k-1}. \tag{8}$$

Since the traffic generated at each slot and at each node is independent, the probability that a slot is busy is independent and identically distributed across slots, i.e.,

$$q_k(m) = \hat{q}_k \qquad \text{for all } m. \tag{9}$$

Substituting (8) and (9) in (7), we have

$$\hat{q}_k = \hat{q}_k P(T > 1) + (1 - \hat{q}_k)kz(1-z)^{k-1}P(T > 1).$$

Noting that $P(T > 1) = e^{-\mu}$, we have

$$\hat{q}_k = \frac{kz(1-z)^{k-1}e^{-\mu}}{1 - e^{-\mu} + kz(1-z)^{k-1}e^{-\mu}}.$$

Hence proved. ■

Due to complex dependency on the number of neighbors, a closed form expression for stability is not possible. The stability condition is now redefined to take into account the exponential length of the packet as follows.

Lemma 4 *Given a network throughput* λ, *transmission radius* r, *channel access probability* p, *exponentially distributed packet length with mean* μ, *the condition for stability is given by*

$$\frac{\lambda D}{r} < p\sum_{k=1}^{\infty}(e^{-\Lambda\pi r^2}\frac{\Lambda\pi r^2}{k!})\left(\hat{q}_k + (1-\hat{q}_k)(1-\frac{p\lambda D}{r})^k\right) \tag{10}$$

where $\hat{q}_k$ *is given by (6)*

Proof: Let $\overline{T_X}$ denote the event of an unsuccessful transmission from a node. From (2) we have that

$$b_k = P_k(\overline{T_X}) = (1-p) + pP_k(\text{collision}).$$

In this case a collision can occur only when the channel is sensed to be idle and more than one neighbor attempts to transmit simultaneously. So conditioned on the channel being idle the probability of collision is given by (3). Thus the probability of a collision for a node having k neighbors,

$$P_k(\text{collision}) = (1-\hat{q}_k)(1-(1-p\frac{\lambda.D}{r})^k).$$

Substituting for $P_k(\text{collision})$, we get,

$$b_k = (1-p) + p(1-\hat{q}_k)(1-(1-p\frac{\lambda.D}{r})^k). \tag{11}$$

The probability of successful transmission is $(1-b_k)$. Thus from our definition of stability, we require that

$$\frac{\lambda D}{r} < p \sum_{k=1}^{\infty} (e^{-\Lambda \pi r^2} \frac{\Lambda \pi r^2}{k!}) \left(\hat{q}_k + (1-\hat{q}_k)(1-\frac{p\lambda D}{r})^k \right)$$

Hence proved. ■

Once we determine the probability that a slot is busy, we can write down the expected delay experienced by a message. However, due to the complex dependency on the number of neighbors, a closed form is not possible.

Theorem 2 *Given a network throughput* λ*, transmission radius* r*, channel access probability* p*,and an exponentially distributed transmission times with mean* $\frac{1}{\mu}$*, the expected channel access delay is given by*

$$\mathbb{E}[d_c] = \sum_{k=1}^{\infty} \frac{1}{b_k} . \frac{e^{-\Lambda \pi f^2} (\Lambda \pi f^2)^k}{k!}, \tag{12}$$

where b_k *is the conditional probability of unsuccessful transmission given by (11),* $f = (1+\Delta)r$ *and* $z = p\frac{\lambda D}{r}$. ■

5 Delay analysis: n nodes uniformly distributed

In this section we will assume that the nodes are uniformly distributed over an unit area. The earlier condition on the stability of the system is modified to take into account the uniform distribution of nodes.

Lemma 5 *A network with* n *nodes distributed uniformly is stable for a given* λ, p *and* r *if:*

$$\frac{\lambda D}{r} \leq p \left(1 - \frac{f^2 p \lambda D}{r}\right)^n$$

where $f = (1+\Delta)r$.

Proof: See Appendix. ■

We now state the relationship between the expected channel access delay, transmission power, network throughput, number of nodes and channel access probability.

Theorem 3 *Given a channel access probability* p*, network throughput* λ*, transmission radius* r *and number of nodes* n*, the expected channel access delay (*d_c*) in slots is given by:*

$$\mathbb{E}[d_c] \approx \frac{1}{p} \left[\left(\frac{f^2}{1-p\frac{\lambda D}{r}} + 1 - f^2 \right)^{n-1} - (1-f^2)^{n-1} \right], \tag{13}$$

where $f = (1+\Delta)r$.

Proof: From Theorem 3, we know that the expected channel access delay conditioned on the number of neighbors is given by

$$\mathbb{E}[\hat{d}_c^k] = \frac{1}{p(1-\frac{p\lambda D}{r})^k}$$

Taking expectations to remove the conditioning on the number of neighbors k we get,

$$\mathbb{E}[d_c] = \sum_{k=1}^{n-1} \frac{1}{1-\hat{q}_k} \cdot \binom{n-1}{k} g^k (1-g)^{n-1-k}, \tag{14}$$

where g is the probability of finding a node within the interference region. Since nodes are distributed uniformly, g is proportional to the area of a nodes interference range, i.e., $g = ((1+\Delta)r)^2 = f^2$. Substituting the value of $\hat{q}_k$ in (14) we have

$$\mathbb{E}[d_c] = \sum_{k=1}^{n-1} \frac{1}{\left(1-\frac{p\lambda D}{r}\right)^k} \cdot \binom{n-1}{k} g^k (1-g)^{n-1-k}.$$

Simplifying the above expression yields

$$\mathbb{E}[d_c] \approx \frac{1}{p}\left[\left(\frac{f^2}{(1-\frac{p\lambda D}{r})} + 1 - f^2\right)^{n-1} - (1-f^2)^{n-1} \right].$$

Hence proved. ■

Lemma 6 *Given a channel access probability* p*, network throughput* λ*, transmission radius* r *and the Poisson parameter of the node distribution* Λ*, the expected delay (*d_t*) in slots that a message experiences from the source to the destination* L *units apart is given by:*

$$\mathbb{E}[d_t] \geq \frac{L}{pr}\left[\left(\frac{f^2}{(1-\frac{p\lambda D}{r})} + 1 - f^2\right)^{n-1} - (1-f^2)^{n-1} \right]. \tag{15}$$

Using a similar argument as in the Poisson distribution case, the total delay when nodes are uniformly distributed is lower bound by the number of hops η times the channel access delay at each hop. Thus we get,

$$\mathbb{E}[d_t] \geq \frac{L}{r}\mathbb{E}[\hat{d}_c]$$

$$\mathbb{E}[d_t] \geq \frac{L}{pr}\left[\left(\frac{f^2}{(1-\frac{p\lambda D}{r})} + 1 - f^2\right)^{n-1} - (1-f^2)^{n-1} \right].$$

Hence proved. ■

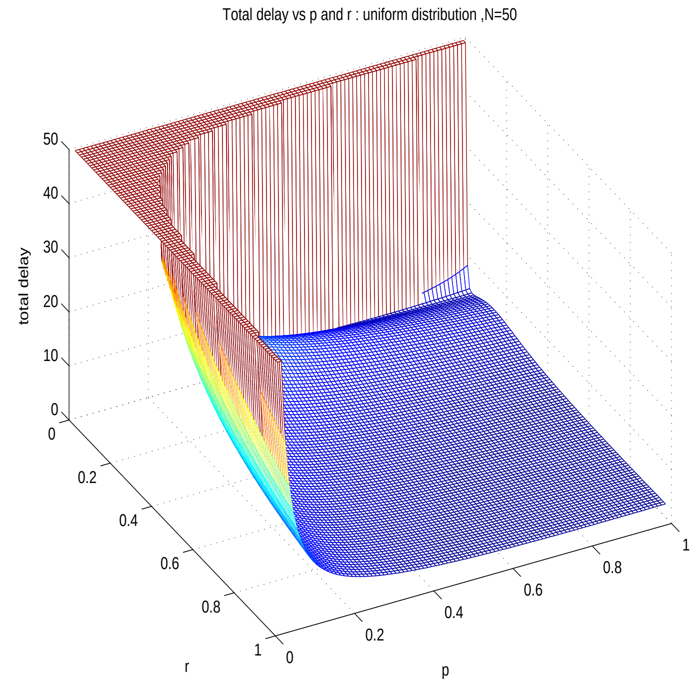

Figure 14: *End-to-end delay as a function of the transmission radius and the channel access probability for 50 nodes uniformly distributed*

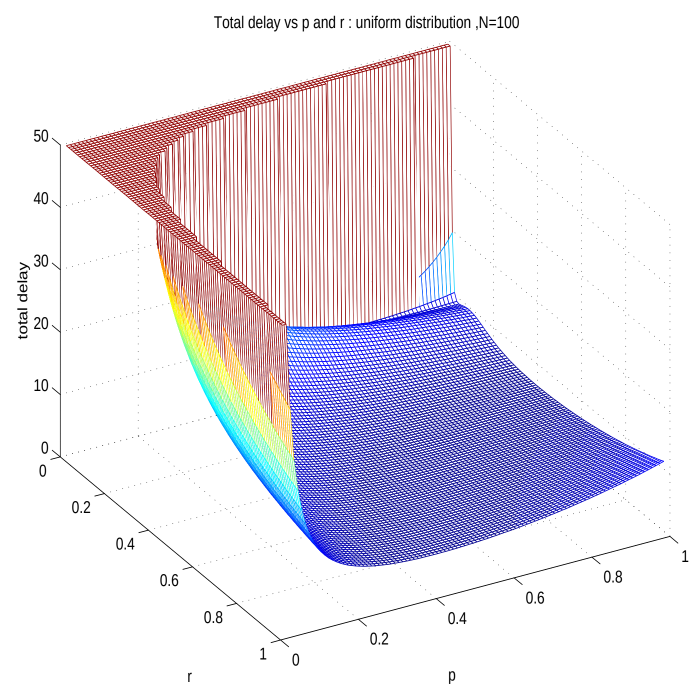

Figure 15: *End-to-end delay as a function of the transmission radius and the channel access probability for 100 nodes uniformly distributed*

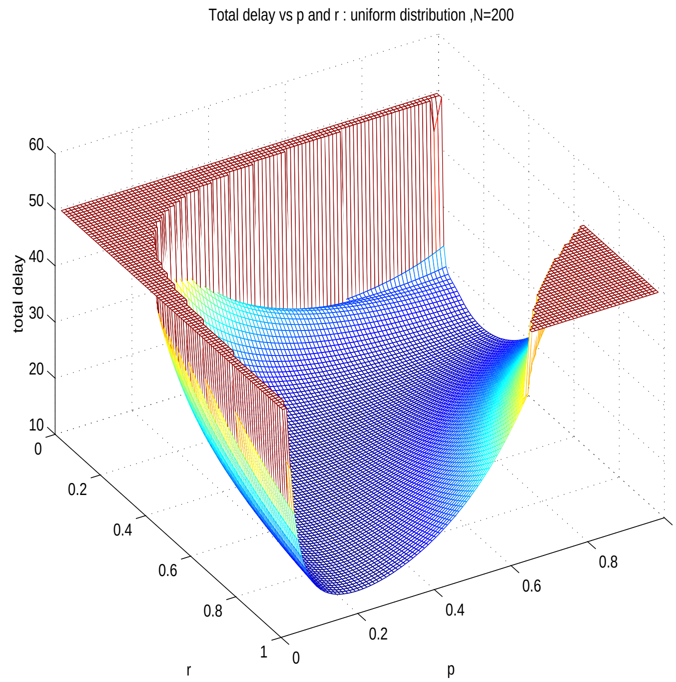

Figure 16: *End-to-end delay as a function of the transmission radius and the channel access probability for 200 nodes uniformly distributed*

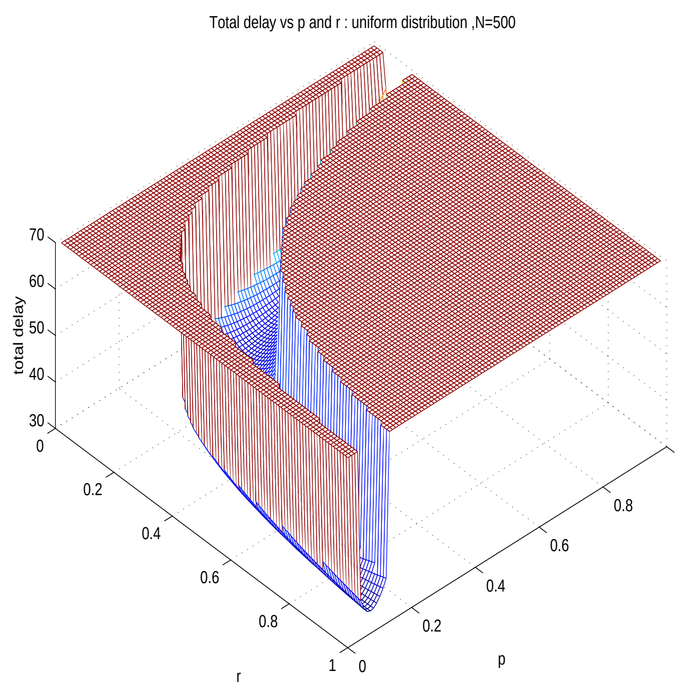

Figure 17: *End-to-end delay as a function of the transmission radius and the channel access probability for 500 nodes uniformly distributed*

6 Conclusions and Discussions

In this paper we look at the delay experienced by messages in an ad-hoc network. In particular we derive expressions for the channel access delay and the total end-to-end delay experienced by a message. The derived expression is a function of four network parameters: 1) channel access probability, 2) transmission radius, 3) network throughput and 4) density of nodes. We analyze the effect of each of these parameters on the channel access and the end-to-end delay. We show that there exists an optimal channel access probability and transmission radius that delivers the best delay performance for the network while guaranteeing stable operation. Finally, one can conclude that:

1. The stable regime of the network is a function of the transmission radius (r), channel access probability (p), node density (Λ) and load on each node (λ). In particular, the network may not be stable for a particular choice of r, p, Λ and λ.

2. There exists a transmission radius and channel access probability in the stable regime that provides the best end-to-end delay in the network. Given a node density, load and channel access probability, one can find the transmission radius that delivers the best end-to-end delay. Similarly, given a node density, load and transmission radius, one can find the channel access probability that delivers the best end-to-end delay.

In this paper we have assumed the channel access probability p to be a non-adaptable quantity per se. However, one can choose a node's channel access probability as a function of the number of its neighbors and the network throughput. A choice of $p = \frac{\alpha}{m+1}$, where m is the number of neighbors and $\alpha \geq 0$ was advocated by Rozovsky and Kumar in [10]. Deriving an exact analytical expression for the case when the channel access probability depends upon the number of neighbors is analytically cumbersome due to the dependencies that arise when deriving the collision probability. Deriving a good bound for the delay when the channel access probability is a function of the number of neighbors is a subject of future research.

7 Appendix

7.1 Proof of Lemma 1

Proof: Consider an arbitrary node i. From Section 3 we know that the average rate at which data is generated for transmission at node i is given by $\frac{\lambda D}{r}$. Let T_X denote a successful transmission from node i. Let

$$P_k(T_X) = P(\text{Successful transmission}|\text{kneighbors}).$$

¿From (4) in Theorem 1 we have:

$$P_k(T_X) = p\left(1 - p\frac{\lambda.D}{r}\right)^k. \tag{16}$$

Taking expectation with respect to the number of neighbors, we have:

$$\begin{aligned} P(T_X) &= \sum_{k=1}^{\infty} p(1-\frac{p\lambda D}{r})^k \frac{e^{-\Lambda\pi f^2}(\Lambda\pi f^2)^k}{k!} \\ &= pe^{-\Lambda\pi f^2} \sum_{k=1}^{\infty} \frac{(\Lambda\pi f^2(1-\frac{p\lambda D}{r}))^k}{k!} \\ &= p\left[e^{-\Lambda\pi f^2 \frac{p\lambda D}{r}} - e^{-\Lambda\pi f^2}\right]. \end{aligned}$$

Hence proved. ■

7.2 Proof of Lemma 5

Proof: With similar notation as in Lemma 1, from (4) , we have

$$P_k(T_X) = p\left(1 - p\frac{\lambda.D}{r}\right)^k. \tag{17}$$

Taking expectation with respect to the number of neighbors, we have :

$$\begin{aligned} P(T_X) &= \sum_{k=1}^{n} p(1-\frac{p\lambda D}{r})^k \binom{n}{k} f^{2k}(1-f^2)^{n-k} \\ &= p\sum_{k=1}^{n} \binom{n}{k} \left(f^2(1-\frac{p\lambda D}{r})\right)^k (1-f^2)^{n-k} \\ &= p\left(1 - \frac{f^2 p\lambda D}{r}\right)^n. \end{aligned}$$

Hence proved. ■

References

[1] Gomez J.and Campbell A.T., Naghshineh M., and C.Bisdikian. Paro : Suporting dynamic power controlled routing in wireless ad-hoc networks. *ACM/Kluwer Journal on Wireless Networks (WINET)*, to appear 2003.

[2] Abbas El Gamal, James Mammen, Balaji Prabhakar, and Devavrat Shah. Throughput-delay trade-off in wireless networks. *Proc. INFOCOM*, 2004.

[3] M Grossglauser and D Tse. Mobility increases the capacity of ad hoc networks. *Proc. INFOCOM*, 2001.

[4] Piyush Gupta and P.R.Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, March 2000.

[5] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In *Proc. HICSS*, January 2000.

[6] Vikas Kawadia and P.R.Kumar. Power control and clustering in ad-hoc networks. *Proc. INFOCOM*, April 2003.

[7] K.Kar, M.Kodialam, T.V.Lakshmanan, and L.Tassiulas. Routing for network capacity maximization in energy-constrained ad-hoc networks. *Proc. INFOCOM*, April 2003.

[8] Srisankar Kunniyur and Santosh Venkatesh. Network devolution and the growth of sensory lacunae in sensor networks. In *Proceedings of Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOPT 04)*, March 2004.

[9] Srihari Narasimhan and Srisankar Kunniyur. Delay differentiation in sensor networks using power control. In *Proc. CISS*, March 2004.

[10] R. Rozovsky and P. R. Kumar. SEEDEX: A MAC protocol for ad hoc networks. In *Proceedings of The ACM Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc 2001*, October 2001.

[11] G Sharma and R R Mazumdar. On achievable delay/capacity trade-offs in mobile ad hoc networks. *Workshop on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, March 2004.

[12] Suresh Singh and C.S.Raghavendra. Power efficient mac protocol for multihop radio networks. *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 1998.

[13] Wei Ye, John Heidemann, and Deborah Estrin. An energy-efficient mac protocol for wireless sensor networks. *Proc. of the 21st International Annual Joint Conference of the IEEE Computer and Communications Societies*, June 2002.