Broadcast with variable data rates in Mobile Ad hoc Networks

JACOB LÖFVENBERG, JIMMI GRÖNKVIST, MATTIAS SKÖLD, ANDERS HANSSON

FOI-R--2582--SE
ISSN 1650-1942

Scientific Report
October 2008

Command and Control Systems

FOI
Swedish Defence Research Agency
Command and Control Systems
P.O. Box 1165
SE-581 11 LINKÖPING

Phone: +46 13 37 80 00
Fax: +46 13 37 81 00

www.foi.se

Jacob Löfvenberg, Jimmi Grönkvist, Mattias Sköld,
Anders Hansson

Broadcast with variable data rates in Mobile Ad hoc Networks

Title	Broadcast with variable data rates in Mobile Ad hoc Networks
Titel	Broadcast med variabla datatakter i mobila ad hoc-nät
Report No. / Rapportnr	FOI-R- -2582- -SE
Report Type	Scientific Report
Rapporttyp	Vetenskaplig rapport
Month / Månad	October / Oktober
Year / Utgivningsår	2008
Pages / Antal sidor	39
ISSN	1650-1942
Customer / Kund	FM
Programme area	7. C^{4}I
Forskningsområde	7. Ledning med MSI
Subcategory	71. Command, Control, Communications, Computers, Intelligence
Delområde	71. Ledning
Project No. / Projektnr	E7108
Approved by / Godkänd av	Anders Törne

FOI, Swedish Defence Research Agency
Command and Control Systems
P.O. Box 1165
SE-581 11 LINKÖPING

FOI, Totalförsvarets Forskningsinstitut
Avdelningeng för Ledningssystem
Box 1165
581 11 LINKÖPING

Abstract

We assess the broadcast transmission cost in ad hoc networks with multipoint relay (MPR) flooding and a variable data rate. Mobile ad hoc networks consists of wireless nodes that build a robust radio network without any pre-existing infrastructure or centralized servers. Tactical mobile communication frequently generates multicast traffic (one-to-many) and broadcast (one-to-all) traffic. Some examples are distribution of status information, position information and voice group calls.

Using variable and adaptive rates on the links in a network will potentially both improve capacity and allow nodes at difficult positions to access the rest of the network with minimal degrading of the network capacity. The routes in the network will also be more robust if its links can degrade gracefully with decreasing rates rater than disappearing at some SNR threshold. In this study, we assume that all nodes use the same single data rate for transmissions and that they can simultaneously vary this rate to improve the performance. A main reason for this approach is to find indications on good strategies for multi-rate broadcast algorithms where different rates are allowed at the same time in the network.

For reference, we also evaluate the transmission cost for unicast traffic (one-to-one) with variable data rates. Here, it seems to be a good strategy to always use the largest possible data rate that does not partition the network. A similar general strategy for the broadcast case is difficult to formulate. However, for networks with low frequency bandwidths and high node densities, there is a transmission cost minimum close to the data rate for which centrally located nodes can reach all other nodes. For these networks it might be a good strategy to chose the data rate so that we have a two-hop broadcast: one hop in to a central node and one hop out to all other nodes in the network. This resembles the transmission behaviour in a cellular base station radio system, which is perhaps not the way we are used to think about ad hoc networks.

Keywords: Ad hoc networks, MPR, multiple datarates, broadcast routing

Sammanfattning

Vi utvärderar transmissionskostnaden för broadcast-trafik i ad hoc-nät med multipoint relay-flooding (MPR-flooding) och variabel datatakt. Mobila trådlösa ad hoc-nät består av ett antal noder, som bildar ett robust radionät utan fast infrastruktur och centraliserade funktioner. Militär mobil kommunikation genererar ofta en stor andel trafik av typen multicast (en-till-många) och broadcast (en-till-alla). Som exempel kan nämnas distribution av statusinformation, positionsinformation och gruppsamtal.

Variabla datatakter på länkarna i nätet har potential att öka kapaciteten samt minska risken att tappa radionoder med svaga länkförbindelser. Rutterna i nätet blir också mer robusta om de ingående länkarna tillåts att gradvis försämras, snarare än att plötsligt försvinna, under ett visst signal-brus-förhållande. I analysen antar vi att alla noder i nätet använder en och samma datatakt och att de simultant kan variera denna datatakt för att förbättra nätets prestanda. Anledningen är att vi vill hitta uppslag till bra strategier för att hantera broadcast i nät med multipla datatakter.

Som referens utvärderar vi också transmissionskostnaden för unicast-trafik (en-till-en) med variabl datatakt. I detta fall verkar det vara en bra strategi att sträva efter så hög datatakt som möjligt utan att nätet delas. I broadcast-fallet är det svårare att formulera en likartad generell strategi. För nät med låg bandbredd och hög nod-täthet ligger emellertid minimum för transmissionskostnaden nära den datatakt som motsvarar att centralt belägna noder når alla andra noder i nätet. I dessa nät kan det alltså vara en god strategi att välja en datatakt som möjliggör “tvåhopps-broadcast”: ett hopp in till en central nod och därifrån ett hopp ut till övriga noder i nätet. Detta påminner om transmissionsbeteendet i cellulära bas-stationsnät, vilket vi normalt inte brukar förknippa med ad hoc-nät.

Nyckelord: ad hoc-nät, MPR, multipla datatakter, broadcast-routing

Contents

1 Introduction

Distribution of status information, position information and voice group calls, are examples of important military tactical communication that generate multicast (one-to-many) or broadcast (one-to-all) traffic in a network. Most methods for broadcast routing are of one of two types. In the first, information describing the network topology is distributed among the nodes so that an appropriate transmission tree can be generated for sending the broadcast traffic. In such a tree, the root is the source node, and all nodes that are not leafs will retransmit the traffic. The other alternative is to use some form of flooding, possibly limited so not all nodes need to resend the traffic. The first approach is more efficient, but the cost for upholding the tree during mobility is high. Flooding techniques are less efficient for the data traffic but need less overhead traffic. Broadcast is not as well studied as unicast in the context of ad hoc networks and the problem is in some senses very different. The trivial solution, using repeated unicast to reach all nodes in the network is inefficient.

1.1 Ad hoc networks

An ad hoc network [1] is a collection of wireless mobile nodes that dynamically form a temporary network without the need for any pre-existing network infrastructure or centralized administration. Due to the limited transmission range of radio interfaces, multiple hops may be needed for one node to exchange data across the network with another node. An ad hoc network is both self-forming and self-healing and can thus be deployed with minimal or no network pre-planning. However, one drawback of this is that the network will not always be connected. A tactical network may be partitioned or fragmented into parts, e.g. due to movements or terrain obstacles. It is therefore necessary that parts of the network can function autonomously, which requires a distributed network control.

1.2 Variable Data Rate

The data rate is an important parameter in any communication situation, with higher rate usually being desired. However, it can sometimes be more efficient to use a lower data rate. The reason for this is that a lower data rate will allow a longer communication range and thus possibly fewer hops to reach the destination. In the case of broadcast traffic, this effect can be very pronounced since with a low data rate (long range) many nodes can be reached with each single transmission. Of course, communication will be inefficient for lower data rates than what is needed to reach all other nodes, and for very high rates the network will be disconnected. This means that there is an optimal data rate for each network. By varying the rate, or using multiple rates, on the links in the network we can potentially keep the network connected and use the available bandwidth more efficiently. The routes in the network will also be more robust if its links can degrade gracefully with decreasing signal-to-noise ratios, rather than disappearing for signal-to-noise ratios below a certain threshold.

We distinguish between variable rate and multirate communication networks. In a network employing variable data rate, the rate is always the same in the entire network, at any given time. The data rate may vary over time, as a result of changing conditions, but since we will not model mobility in this report, dynamic data rate changes will have no impact on our discussion. In a network employing multirate, every node can choose the data rate to use individually for each transmission. This results in a greater degree of freedom and thus possibly a more efficient communication solution. In both the variable rate and the multirate case the rates will be chosen from some pre-defined set of possible data rates. From a theoretical point of view it would be advantageous if this set was very large, so as to allow for a careful optimization. In practice however, this set will probably be rather small.

This report focus on the use of variable data rate for broadcast communication in ad hoc networks. Our ultimate goal is to study the use of multirate communication for broadcast in ad hoc networks. The results we present in this report is a first step in that direction, where we consider how the choice of a single data rate, used by all nodes in the network, impacts the transmission cost of broadcast communication in ad hoc networks.

2 Problem

Using variable, or multiple data rates is an interesting possible way of simultaneously optimizing throughput and connectivity in a wireless ad hoc network. Lower rates allows longer ranges and higher rates will yield a greater throughput. Thus, by choosing the appropriate rate for each transmission it should be possible to achieve a lower total transmission cost.

However, using variable data rates efficiently is a difficult problem, and even more so when used in a broadcast solution. Our ultimate goal is to find effective, distributed algorithms for broadcast communication using variable data rates. In order to achieve this we first have to answer some basic questions. Is there anything to gain by using variable data rates? How much can be gained? How sensitive is the transmission cost to the choice of strategy, i.e., can small changes in the communication strategy yield large changes in cost?

In its general setting these questions are very difficult to answer. Instead we will make some simplifying assumptions and study some more specific problems.

We use multipoint relay flooding as broadcast strategy [2], [3], [4]. As broadcast transmission cost we use the sum of inverted rates for every transmission, that is, the cost is proportional to the total transmission time in the network [5].

We study the impact of the data rate in networks where the data rate is the same in the whole network. This should show if the choice of data rate is important for the broadcast transmission cost. If it is, there will be even more to gain from a more general approach where not all transmissions use the same data rate. The analysis of this special case will also give valuable understanding of the problem that can be used later on, for example in formulating specific strategies with variable rates.

In summary, our questions are: In ad hoc networks with one single data rate,

- how much does the choice of data rate impact the broadcast transmission cost?
- how are data rate and broadcast transmission cost related?
- is it possible to formulate a strategy for choosing the data rate, such that the broadcast transmission cost is minimized (or reasonably good)?

3 Model and Scenario Setup

For theoretical analysis, simulations and experiments to be possible we have to specify what the network model and scenario parameters are and how broadcast communication will work.

3.1 Network

An important part of a scenario is how nodes are placed and how they interact with the terrain in the area in which the network resides.

We will use the following model to describe the network. The nodes in the network will reside in a square area of size $L \times L$ without terrain and terrain effects. We also assume that only one transmission is active at any one time, and that this is solved at some lower layer. This means that the only thing affecting a link between two nodes is the attenuation due to the distance between the two nodes.

Since we do not want to make assumptions on the placement of the nodes, as this will depend greatly on the situation at hand, we will assume a uniform random placement. The networks will further be static, without any movement of the nodes. This can be seen as a snapshot of a network with mobile nodes.

3.2 Communication

Given the knowledge of the nodes' geographical location we also need to model the communication channel and how this is affected by the relation between the nodes. Further, we need to decide on exactly how the broadcast is done, in order to compute a transmission cost measure for the networks.

3.2.1 Radio Interface

We assume that the radio channel is an additive white Gaussian noise (AWGN) channel. The capacity of the channel is then defined by the signal-to-noise ratio (SNR), which can be found using a link budget for the radio channel.

$$\frac{S}{N} = \frac{P_t G_t G_r}{L_b F k T_0 W}, \tag{3.1}$$

where S is the signal level, N is the noise level, P_t is the transmitted power, G_t and G_r are the gains of the transmitter and receiver, L_b is the elementary path loss (signal attenuation between transmitting and receiveing antenna), F is the system noise factor, kT_0 is the noise level at reference temperature $T_0 = 290\text{K}$ and W is the bandwidth.

Before choosing the values of the parameters we will turn to L_b, the elementary path loss. The path loss is often modelled as proportional to distance d raised to the power of α. Free-space propagation would give $\alpha = 2$, while a plane-earth model [6] would give $\alpha = 4$. For most of the discussion in this report, we will chose the plane-earth model for the radio channels between pairs of nodes. This model yields $L_b = d^4/(h_1 h_1)^2$, where d is the distance between transmitter and receiver, and h_1 and h_2 are the elevations of the antennas.

We will use the following values for the parameters: $P_t = 1$, $G_t = G_r = 1$, $kT_0 = 4 \times 10^{-21}$, $F = 25$ and $(h_1 h_2)^2 = 40$ (i.e. $h_1 = h_2 \approx 2.51$). We let W be a parameter and d will be given by the placements of the nodes. Taken together we get the SNR as:

$$\frac{S}{N} = \frac{4 \times 10^{20}}{W d^4}. \tag{3.2}$$

The actual data rate in a given situation will depend on the communication system. For simplicity we will use the channel capacity for AWGN channels [7] as an upper bound on the data rate, given the SNR. The reason for not using the SNR itself in comparisons and graphs is that we believe it to be easier to get an intuitive feel for what the results mean if there are actual data rate values to compare. The data rate values may be overly optimistic in this way (only bounded by the channel capacity), but they will still be in the same range as what is possible to achieve in real systems. Thus, using the expression for the SNR in 3.2 yields an upper bound on the data rate, R, for the radio channel as

$$R \leq W \log_2(1 + \frac{4 \times 10^{20}}{W d^4}), \tag{3.3}$$

which is the expression we will use in the following chapters.

3.2.2 Broadcast Algorithm

To be able to compare solutions we further have to agree on a broadcast algorithm to use in the simulations. We have chosen the multipoint relay (MPR) flooding algorithm [2], [3], [4]. This is a distributed broadcast algorithm which has proven to be both robust and efficient in ad hoc networks.

The MPR broadcast algorithm works in the following way, using graph theoretical notation. Two nodes are considered neighbours if they can communicate with each other. For each node ν we define a set of neighbours N_1 and a set of second-neighbours N_2, such that $\nu \notin N_1$, $\nu \notin N_2$ and $N_1 \cap N_2 = \emptyset$. The MPR set ν_{MPR} of the node is then chosen as a subset of N_1 such that N_2 is a subset of the neighbours of ν_{MPR}. How ν_{MPR} is chosen will affect the transmission cost.

In our experiments the MPR set of a node ν will be chosen in the following, common, way:

- the MPR set ν_{MPR} is heuristically minimized by first adding the neighbours that are alone in reaching some second-neighbours, and then by adding other neighbours in order of how many further second-neighbours are covered,

- in the greedy step above, if there are several neighbours covering the same number of second-neighbours, add the one with more neighbours (regardless of whether they are already covered or not).

If ν wants to broadcast a message it makes one transmission with the message, and other nodes will retransmit the message so that every node receives it. If ν hears a node μ (not necessarily the source of the message) transmitting a broadcast message, ν will retransmit the message if it has not received the message previously and if ν is an MPR-node of μ, $\nu \in \mu_{MPR}$. With this algorithm messages that are broadcast are guaranteed to reach all nodes in the network, as long as the network is connected.

3.2.3 Broadcast Transmission Cost Measure

The need for a communication cost measure to compare different solutions has been mentioned several times. The communication cost measure we will use is the time the communication operation takes in the network for distributing one unit of data, taking into account the fact that we allow only one transmission at a time. This means that for a broadcast operation the cost is defined as the sum of the inverted data rates for all the individual transmissions used. We will not take into account any overhead needed, for example in the routing or network management process.

When comparing solutions using simulations we will randomly create a number of networks. In each network we will compute the broadcast transmission cost for each node and compute an average. These are in turn averaged over the results for the different networks. In this way we will find an approximation of the expected value of the broadcast transmissin cost for this specific combination of simulation parameters and communication strategy, the inverse of which can be interpreted as the expected network broadcast capacity in bits per second.

3.3 Parameter Values

In the simulations we will vary the data rate to see how this relates to the broadcast transmission cost. We will do so for different combinations of n – the number of nodes, L – the size of the network area and W – the bandwidth. For each parameter we will use two "standard values". Simulations will be done with two parameters chosen as standard values and the third parameter varying over a range. All in all this will result in twelve curve families. The parameter values and the simulation runs that will be used are described in Tables 3.1 and 3.2.

Parameter	Standard value 1	Standard value 2	Range
L	10	20	5, 7, . . ., 25
n	50	200	10, 20, 40, 60, . . ., 200
W	100	500	100, 200, . . ., 1000

Table 3.1: Simulation parameter values.

Combination	L (km)	n	W (kHz)
1	10	50	varying
2	10	200	varying
3	10	varying	100
4	10	varying	500
5	20	50	varying
6	20	200	varying
7	20	varying	100
8	20	varying	500
9	varying	50	100
10	varying	50	500
11	varying	200	100
12	varying	200	500

Table 3.2: Combinations of simulation parameters.

4 Theoretical Analysis

In this chapter we will do some theoretical studies of the broadcast problem. We will start in a more general setting, studying how the required number of retransmitters affect when retransmission is preferable. We will then focus on the network scenarios described in the previous chapter to have something comparable for the simulations in the following chapters.

4.1 General Analysis

In this first analysis we will take a slightly more general approach than the scenarios described in previous chapters.

4.1.1 Assumptions

As stated in Chapter 3 we assume that the received power is dependent only on distance, i.e., there are no specific terrain effects. However, unlike when using the base scenarios we will not only look at the plane-earth model when calculating the elementary path loss, but rather use different values of α.

This generalizes the expression 3.2 so that we can express the SNR at the receiver as

$$SNR = 4 \times 10^{20}/Wd^{\alpha}, \tag{4.1}$$

where d is the distance between the nodes and α is the factor that determines how fast the received power decreases with the distance.

An upper limit on the data rate between two nodes will then be

$$R \leq W \log_2(1 + 4 \times 10^{20}/Wd^{\alpha}), \tag{4.2}$$

where W. We let $C = 4 \times 10^{20}/W$, so that we can simplify 4.2 to

$$R \leq W \log_2(1 + C/d^{\alpha}) \tag{4.3}$$

and the SNR to $SNR = C/d^{\alpha}$.

Now, to determine which data rate a node should use, we assume that it only has two data rates to choose from. The first data rate R_1 gives a range of D, the second data rate R_2 gives a range of $D/2$ (assuming we use the maximum data rates possible, i.e., equality, in expression 4.3). Obviously $R_2 > R_1$, however if the node uses the data rate R_2 other nodes need to relay the message to reach as many nodes. Since we can reach nodes at $D/2$ distance we assume that the relaying nodes are at that distance and can also use the R_2 data rate.

This is not entirely correct of course, since there might not always be nodes there; furthermore, in the broadcast case we need to reach all nodes in a circle of radius D, from nodes at distance $D/2$ with range $D/2$, which would require more relays than what would be practical. In reality a slightly lower data rate than what can be achieved at $D/2$ would be used but we will approximate this with R_2 for simplicity. This will make relaying seem better than what it in reality should be, but we are missing its advantages in a more complex terrain where it can be used to avoid obstacles.

We assume that using relaying is as efficient as direct transmission if the costs for sending a message are equal. All nodes are using the same transmission power, so both used energy and transmission resources will be equal in such a case. Furthermore, the interference caused on other nodes will on average be equal due to the use of equal transmission power. Of course, the source will be different, but that should be evened out over a large network.

With these assumptions we have the following inequality for when direct transmission is preferable over relaying.

$$(1+H)/R_2 > 1/R_1, \tag{4.4}$$

where H is the number of needed retransmissions. For unicast H is equal to 1; for broadcast we probably need between 3 and 5 retransmissions if the node is not in the edge of the network surface.

4.1.2 Analysis

The $SNR_{D/2}$ at distance $D/2$ can be rewritten as $(C/D^\alpha)2^\alpha$, i.e., $SNR_{D/2} = 2^\alpha SNR_D$. We use this in 4.4, again using expression 4.3 with equality, and find:

$$(1+H)W\log_2(1+SNR_D) > W\log_2(1+2^\alpha SNR_D), \tag{4.5}$$

which can be simplified to

$$(1+SNR_D)^{1+H} > (1+2^\alpha SNR_D). \tag{4.6}$$

This is not simple to solve for arbitrary choices of H, but for unicast traffic where H is equal to one we get the follow inequality:

$$1+2SNR_D+{SNR_D}^2 > 1+2^\alpha SNR_D. \tag{4.7}$$

Since SNR_D is positive, we can simplify this as

$$SNR_D > 2^\alpha - 2. \tag{4.8}$$

This means that for unicast transmissions it is advantageous to relay if the SNR at direct transmission is smaller than $2^\alpha - 2$. For example, for $\alpha = 4$, this means that we should send a unicast transmission directly if the SNR is greater than , in our scenario corresponing to distances $d = 4111$m for $W = 100$kHz and $d = 2749$m for $W = 500$kHz. For lower SNR it is better to relay it if possible. However, in a free space scenario with $\alpha = 2$ we only need an SNR of 2 at the receiver before direct transmission is preferable.

H/α	2	3	4
1	2 (3)	6 (7.8)	14 (11.5)
2	0 (-∞)	1.19 (0.76)	2.4 (3.8)
3	0 (-∞)	0.48 (-3.2)	1.05 (0.21)
4	0 (-∞)	0.23 (-6.4)	0.6 (-2.2)
5	0 (-∞)	0.11 (-9.6)	0.39 (-4.1)
6	0 (-∞)	0.04 (-14)	0.26 (-5.9)

Table 4.1: SNR values - linear and (dB) values

In a broadcast situation each node will in general have its transmission retransmitted by more than one node. In Table 4.1 we show the highest values of the SNR for which it is worth relaying data for different values of α and different number of relays, H. As can be seen the highest values at which relaying is advantageous rapidly decrease with increasing number of relays.

For unicast traffic, i.e. one relay, we see advantages of relaying even at relatively high values of SNR, especially for high values of α where we can have SNRs of more than 10dB and still gain from relaying.

If several nodes need to relay the message, the highest value of the SNR for relaying becomes very much lower, even at $\alpha = 4$ we are down at 4dB. For more relaying nodes the value is below 0 dB in most cases. For free-space propagation, direct transmission is always preferable.

These results do not take terrain into consideration of course, obstructing objects makes relaying more advantageous than is visible here. We can also see that the fewer relays we use, the higher the SNR values in the table. This suggests that nodes at the edge of a network, which needs fewer relays if it increases the data rate (since a message only is retransmitted in mostly one direction) should probably use a relatively high data rate, while at the centre of a network, in which case a node needs many relays, lower data rates should probably be used. This means that letting a single central node relay (especially if it is highly positioned such as a base station, or have good terrain coverage) a transmission at a low data rate to most of the network may often be a good solution for smaller networks.

For larger networks, relaying is not only advantageous, but absolutely necessary to reach the entire network. However, depending on the shape and connectivity of the network the number of relays for different nodes may vary. For nodes that can reach a large part of the network, in all directions, the results above suggests the use of very low data data rates to avoid relaying. In reality though, practical implementations will often set a lower limit on the data rate for communication of any efficiency. It is also likely that some nodes may be in a better position for doing such a long relaying than others (for example due to terrain) meaning that other nodes will probably want to use as high a rate as possible to reach these nodes.

4.1.3 Comparison with Simulation Parameters

For comparison we will see what this analysis shows if we use the parameter values from Chapter 3. We will study the unicast case, $H = 1$, using expression (3.3) to compute the cost per distance of relaying. In Figure 4.1 is shown the cost per distance for different transmission ranges, d, where we have used equality in expression (3.3).

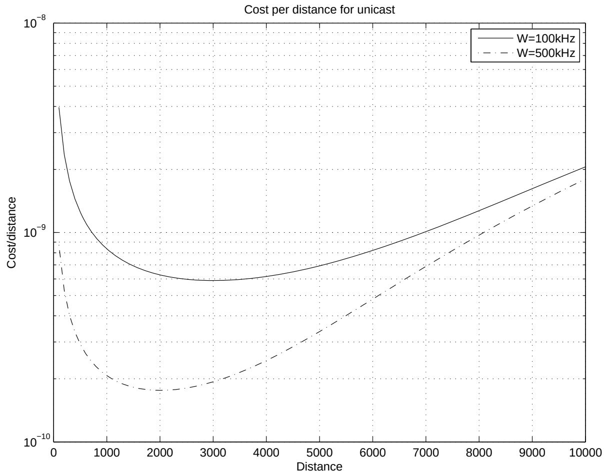

Figure 4.1: Cost per distance for unicasting relaying for two different bandwidths, W.

As we can see there is an optimal distance D_{opt} for which the transmission cost per distance is minimized. This suggests that for communication over large distances we should use a data rate that gives this distance for each relaying transmission.

We can also interpret the results from the previous sections using this figure. For any data rate R_2 with range D_2 in the figure, there is another data rate R_1 with range $D_1 = D_2/2$. If the cost per distance at D_1 is less than the cost per distance at D_2, then relaying as described in this section is advantageous. It is interesting to note that this means that relaying is advantageous for distances D_2 up to, and slightly beyond D_{opt}. At least in the case of the cost/distance-curves in Figure 4.1 the curves are suffiently close to symmetric that D_{opt} is approximately $(D_1 + D_2)/2$.

The same kind of reasoning as we have done here can be done for the broadcast case. However, it is not clear what constitutes a realistic for H, and also, even if the nodes are dense across the surface it is not obvious how much area each transmission covers. This will depend on how the relaying nodes are chosen, and this is a non-trivial problem. For these reason we contend ourselves with showing the results for the unicast case.

4.1.4 Discussion

In this part we have studied the broadcast problem with multiple available data rates in the links. We have shown that in general relaying is more advantageous for unicast communication than for broadcast communication. If packets should be transmitted in all directions from a node, relaying becomes much more expensive and only for very low SNR values, e.g. due to long distances, do we see a gain in using multiple relays rather than using direct transmissions. However, for nodes close to the edge of the network surface, a transmission at a high rate that is relayed towards the centre of the network will often be advantageous since this, locally, is similar to unicast communication.

Terrain and practical consideration will have a large impact on these results, though. Nodes with a good coverage are probably the ones that should use low data rates to reach large portions of the network.

4.2 Scenario Specific Analysis

In this second theoretical analysis we will take a more detailed look at what happens in the specific case described in Chapter 3. We will analyze the broadcast transmission cost in different ways depending on the data rate. We divide the date rate range into three different segments: low, moderate and high rates. These will be discussed in the following subsections.

4.2.1 Low Rates

For low rates the communication range will be large enough for every node to reach all other nodes without the need for retransmission, regardless of their positions on the surface. We call this *low rates*. This will hold as long as the range is greater than $L\sqrt{2}$, i.e., as long as every node can reach the entire surface. In this data rate segment any broadcast will need exactly one transmission, and the broadcast transmission cost will thus be $1/R$. In Table 4.2 below is shown the upper limits for low rates for the simulation parameter standard values.

4.2.2 Moderate Rates

For slightly higher data rates some nodes will not be able to reach the entire surface, and thus, at least in some cases, not every other node. However, as long as the transmission range is at least $L/\sqrt{2}$, a centrally located node will be able to broadcast to every other node, regardless of their position on the surface. These are what we call *moderate rates*. For moderate rates, nodes that do not reach the entire surface will be able to broadcast using a central node as a relay.

An approximate expression for the cost can be found in the following way. We assume that the node density is so high so that there with a high probability will be a dense and close to uniform distribution of nodes on the surface. We let β be the ratio of the surface area from which the entire surface can be reached with a single transmission. Thus, for moderate rates the expected broadcast transmission cost will be approximately $\beta/R + 2(1-\beta)/R$. The value β is in itself dependent of R, so the exact behaviour of this expression cannot be seen without some further analysis. However, we confine ourselves to noting that the cost will go from $1/R_{low}$ to $2/R_{mod}$, where R_{low} and R_{mod} are the upper limits of the low and moderate rates.

In Table 4.2 below is shown the upper limits for moderate rates for the simulation parameter standard values.

L (km)	W (kHz)	Upper limit for low rates (kbit/s)	Upper limit for moderate rates (kbit/s)
10	100	14	138
10	500	14	200
20	100	0.90	14
20	500	0.90	14

Table 4.2: Upper limits for low and moderate rates for combinations of parameter values.

4.2.3 Higher Rates

For *high rates* (everything beyond moderate rates) it is more difficult to find a simple cost expression. We will make an approximation using a surface covering argument. We assume that the entire surface has to be reached by the message and that there are nodes available everywhere for retransmitting the message. Using these assumptions we will analyze the basic behaviour of the cost curve. In order to do this we will start by revisiting the expression we use for the available data rate on the radio channel, the channel capacity given in (3.3). In this case we will use it backwards to see what squared communication range, d^2, that is possible for a given data rate, R, using (3.3) with equality. We find that $d^2 = \sqrt{4 \times 10^{20}/(2^{R/W}-1)W}$.

Each retransmission will cover some area that was already covered by previous transmissions. Most prominently, the area between the retransmitting node and the node that made the previous transmission will to a large extent be surface that has been covered before. We will assume that each transmission will cover a previously uncovered surface area of size d^2. This is a reasonable value and also means that the curves from moderate and high rates will fit together.

The broadcast transmission cost is, with T being the number of transmissions and R the rate,

$$Cost = T/R = (1/R)(L^2/d^2) = L^2\sqrt{(2^{R/W}-1)W}/2 \times 10^{10}R. \tag{4.9}$$

That is, if $2^{R/W}$ is large the cost is approximately proportional to $2^{R/2W}/R$. This approximation is continuous for positive R, decreasing if $R < 2W\log_2(e)$ and increasing if $R > 2W\log_2(e)$.

In the analysis there is nothing that shows that this expression will not be valid for infinitely high data rates. However, in practice the number of nodes is finite, and as the rate increases the range over which communication is possible will decrease. At some point it will become so small that the network will no longer be connected, and broadcast will no longer be possible.

This means that the cost, even though it initially decreases, will grow exponentially for sufficiently high rates. However, due to the networks becoming disconnected at some point, this may not show in practice since the quick cost increase may begin beyond this point.

4.2.4 Combining the Results

We are now at a position where we can combine the results for low, moderate and high rates. In 4.2 below is shown an example of the theoretical transmission cost analysis. Below $R = 13750$ it describes a $1/R$ curve. At $R = 137850$ it is $2/R$ and in between it follows the convex combination of $1/R$ and $2/R$ described in 4.2.2. For high rates the surface covering analysis has been used.

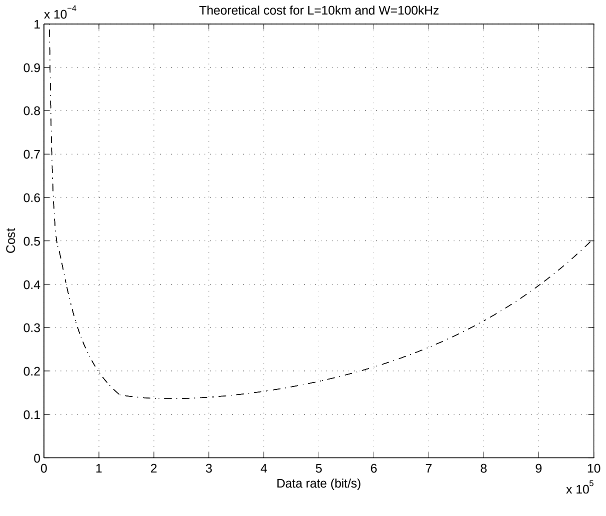

Figure 4.2: Theoretical curve for W=100kHz and L=10km.

5 Simulations

5.1 Simulations

Using the setup described in Chapter 3 we will do a large number of simulations to investigate the behaviour of both unicast and broadcast transmission cost under different circumstances.

In all cases we will randomly generate 500 networks and analyze the transmission cost as a function of the data rate. For each network we will compute the transmission cost for data rates starting with 50kbit/s and increasing until the generated networks are no longer connected. For each rate the cost will be noted, as well as whether or not the network is connected. These results will then be avaraged, yielding the avarage cost and the ratio of non-connected networks.

5.2 Unicast Traffic

For reasons of comparison we have done unicast simulations with the scenario parameters. In this way we will have a wider background with which to compare the broadcast simulations. We have used all eight combinations of standard values from 3 and tried every combination on sender and receiver in the randomly created networks. The cost of these transmissions have then been averaged to show a mean cost of unicast transmission for the parameters in questions. In each case we have randomly generated 500 networks and the results are shown in Figures 5.1 and 5.2 below.

Each graph contains two curves. The red curve shows the average transmission cost in a multirate setting, that is, when the maximum data rate is allowed individually on every link in the network and the path with the smallest overall cost is used between sender and receiver. This represents the lowest possible cost with the given parameters, and since it does not depend on one single data rate, the curve is flat. The multi-colored curves are coded so as to show the ratio of connected networks at every point on each curve. Deep blue means that all networks are connected, green means that 50% of the networks are connected and dark red means that almost no networks are connected. The ratio of connected networks is always strictly increasing.

There are two things that are worth noting in the figures. The first thing is that even though the cost increases in most cases for high data rates, this increase is not very great. The most pronounced increases are always for such high data rates that the ratio of non-connected networks has started to increase. We see that compared to using the optimum data rate, we will not lose very much, in terms of cost, by adopting the strategy to use as high a data rate as possible, as long as the net is connected.

The second thing to notice is that the difference between the variable data rate solution (the same data rate everywhere in the network) and the multirate solution is surpricingly small. If we do choose the optimal data rate, the difference is small in most cases, and depending on the application it may even be insignificant for some parameters.

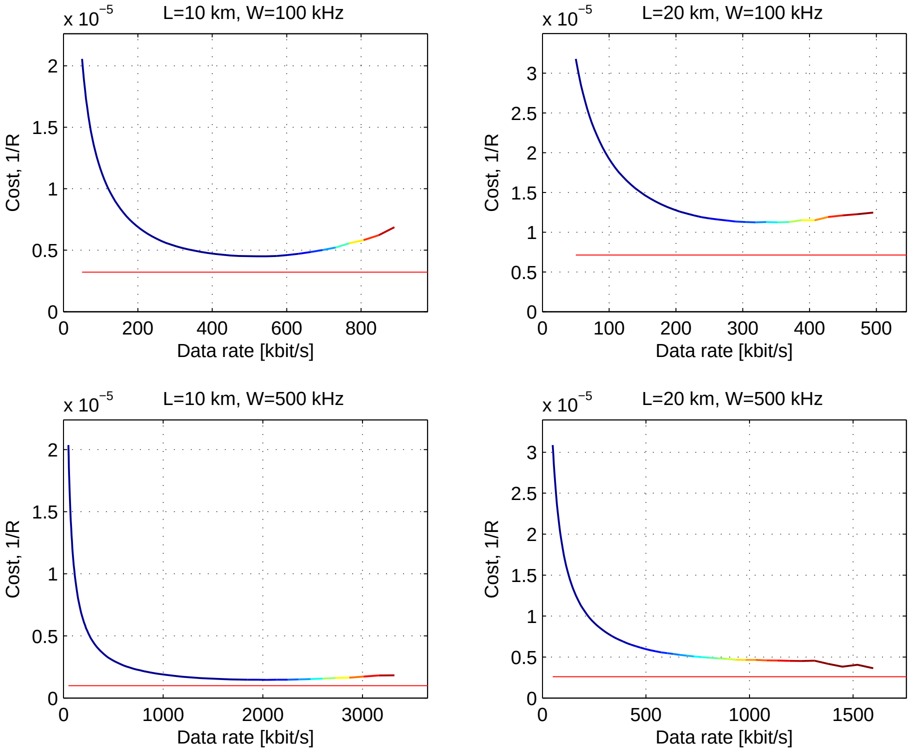

Figure 5.1: Unicast simulation results with number of nodes, $n = 50$. In the multicolored curve deep blue means that all networks are connected, green that 50% are connected and dark red that almost no networks are connected. The ratio of connected networks is always strictly decreasing.

5.3 Broadcast Traffic with Variable Rate

We have investigated the impact of variable data rate on the broadcast transmission cost. Simulation parameters and assumptions have been described in Chapter 3 and a theoretical analysis has been done in Section 4.2. This has hopefully given us some intuitive understanding of the problem at hand and a background against which to interpret the simulation results.

5.3.1 Results

The results are shown in Figures 5.3, 5.4 and 5.5 below. The figures require some explanation. Each subplot shows a family of curves related to a certain combination of two of the three parameters (number of nodes, n, size of the surface, L and bandwidth, W) for different values of the remaining parameter. The curves are color coded so as

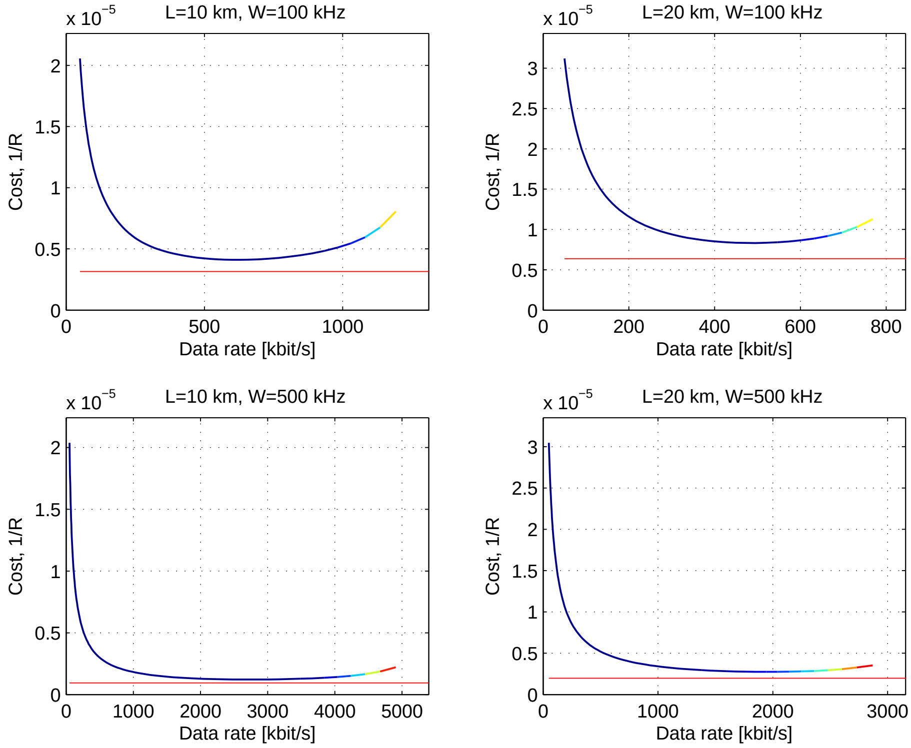

Figure 5.2: Unicast simulation results with number of nodes, $n = 200$. In the multicolored curve deep blue means that all networks are connected, green that 50% are connected and dark red that almost no networks are connected. The ratio of connected networks is always strictly decreasing.

to show the ratio of connected networks at every point on each curve.

Looking at Figures 5.3, 5.4 and 5.5 there are several things to notice. We can see that there is a wide range of different behaviours present. Some curves have a narrow range of distances with low cost close to the minimum point while others are rather flat, with a minimum cost within a wide range of distances with a low cost. There are also curves that decrease with increasing rate, up to the point where the nets become non-connected. We would like to point out that the fact that the curves often exhibit a more “random-like” behaviour at the right end points is an artifact due to the fact that that in the red parts of the curves the avarage cost have been computed over only a few networks (those of the 500 randomly chosen networks that were connected at these rates).

We take a closer look at the two left diagrams in Figure 5.5 and show them in full size, somewhat zoomed, in Figures 5.6 and 5.7. In these figures are shown as crossed circles the rate-cost combinations for when a node placed exactly in the middle can (just) reach the entire surface. As we can see, in all these cases, the local minima are

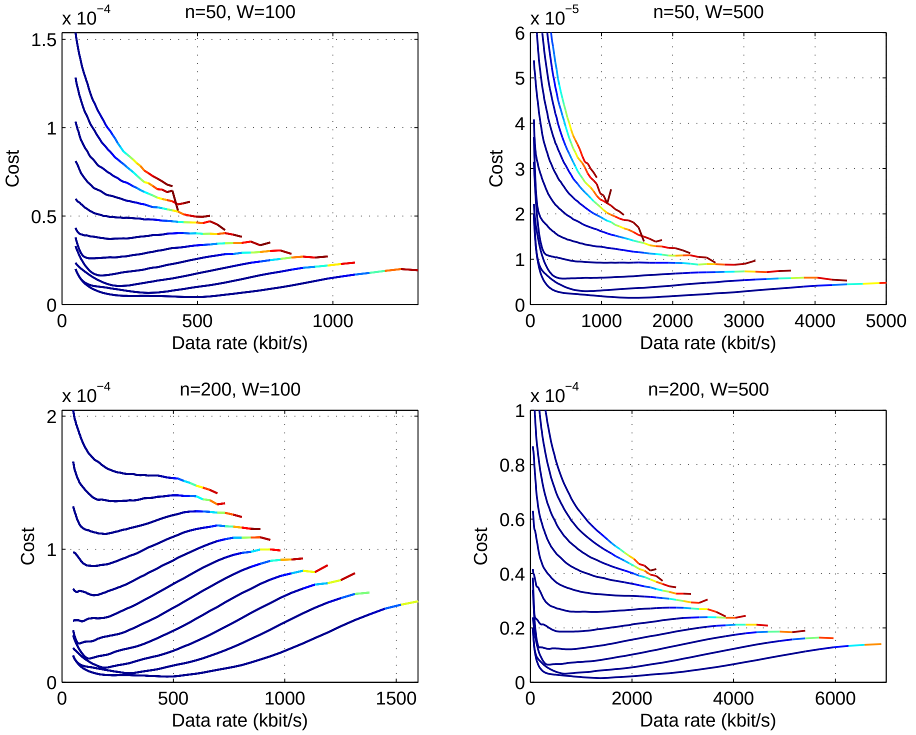

Figure 5.3: Simulation results with $L = 5, 7, \ldots, 25$. Greater L correspond to curves with greater cost. Deep blue means that all networks are connected, green that 50% are connected and dark red that almost no networks are connected. The ratio of connected networks is always strictly decreasing.

only slightly to the right of these data rates. This means that in many cases an efficient solution is to choose the data rate so that one (or several) centrally located nodes can act for other nodes as a relay station with full network reach. This is in a sense similar to a base station solution, where the central node is the base station relaying broadcasts from any source node.

As we have seen, this solution is close to optimal in these situations. There are also other combinations of parameter values for which the base station solution is good, but the ones shown here are the most prominent ones.

5.3.2 Comparison with Theoretical Results

In Figures 5.8, 5.9 and 5.10 below we compare some of the simulation results with what the theoretical analysis in Section 4.2 predicts. We see that the general shape of the simulation curves are very similar to the theoretical ones, although not exactly the

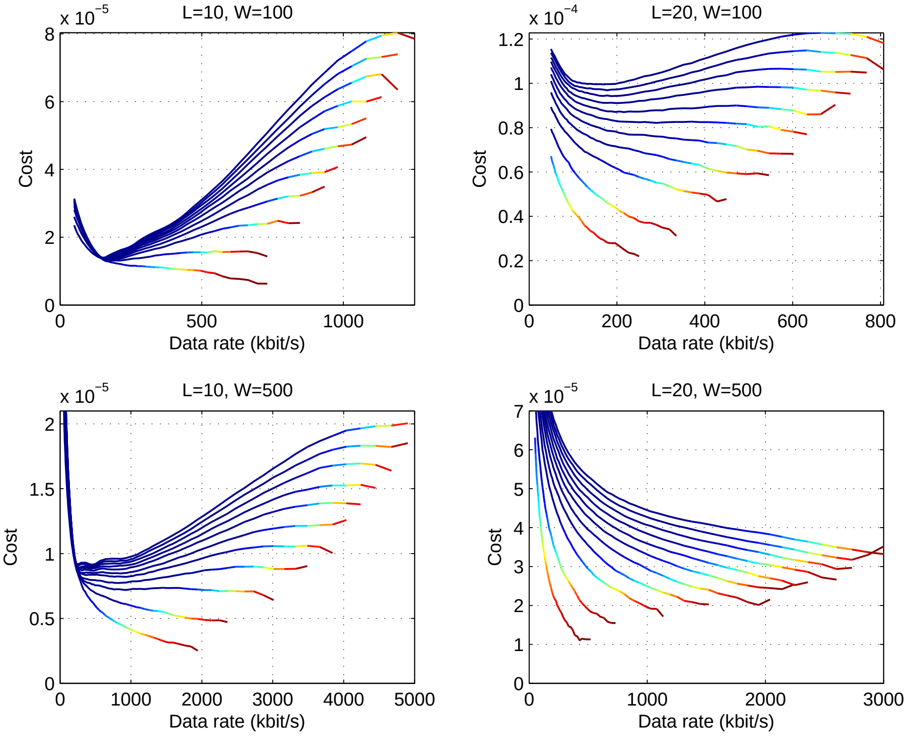

Figure 5.4: Simulation results with $n = 10, 20, 40, 60, \ldots, 200$. Greater n correspond to curves with greater cost. Deep blue means that all networks are connected, green that 50% are connected and dark red that almost no networks are connected. The ratio of connected networks is always strictly decreasing.

same. Our opinion is that they are similar enough for the theoretical analysis to have a value, and for the simulation results to be trustworthy.

5.4 Broadcast Discussion

Studying the figures in the previous section we can see that the cost-data rate curves can show very different behaviours, depending on the parameters used. There is certainly different costs for different data rates, so there is a good reason for carefully choosing the data rate to use in a network. The different behaviours however mean that it may be a difficult problem to find the optimal data rate, or even a good data rate, in a real communication situation, where parameters are unknown, varying or both.

Trying to characterize the curves we find some properties that the curves have in common. All of them have a high cost for small data rates. This cost rapidly decreases

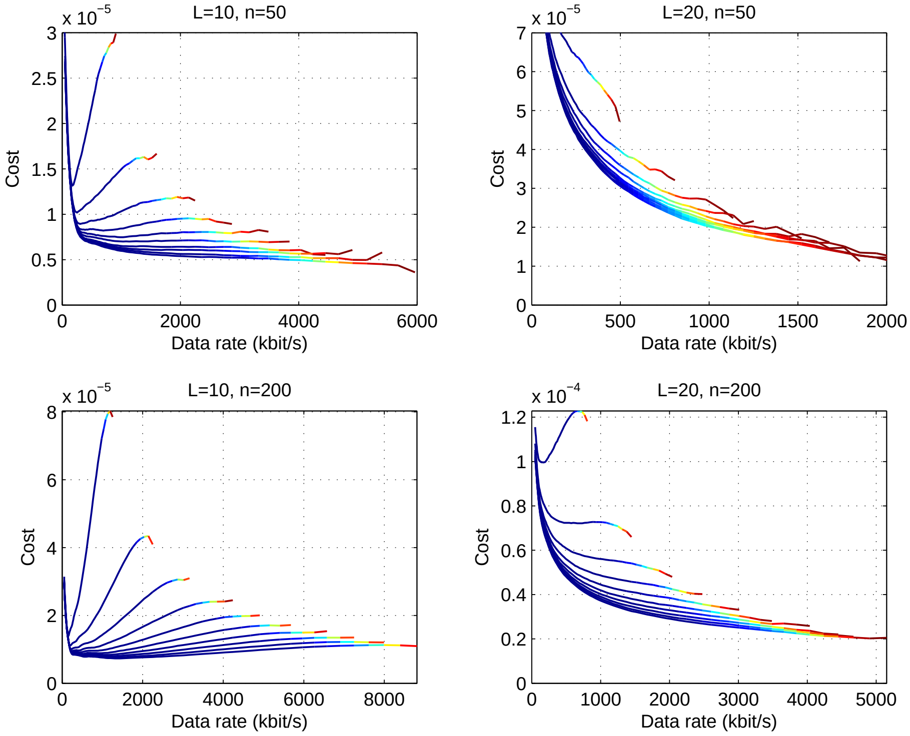

Figure 5.5: Simulation results with $W = 100, 200, \ldots, 1000$. Greater W correspond to curves with lesser cost. Deep blue means that all networks are connected, green that 50% are connected and dark red that almost no networks are connected. The ratio of connected networks is always strictly decreasing.

as the data rate grows. Many curves then show a local minimum before the cost increases again. This corresponds well to what the theoretical analysis predicts and we believe that we have a basic understanding of why the cost curves look the way they do.

The local minima, when they exist, seem to be suitable targets for broadcast transmission in those networks. Even though the data rates differ for the minima, in many cases they correspond approximately to the data rates for which centrally located nodes can reach the entire surface. This is not the way we are used to think about ad hoc networks; we have reduced a multi hop network to a double hop network. However, broadcasting is a special problem, and solutions for this problem do not necessarily have to be similar to unicast solutions.

We note that different simulation results show very different behaviour. In some cases there is a wide range of data rates that yield a low cost. In some other cases the cost decreases with increasing data rate, and in yet other cases there is a narrow range of rates with a low cost close to the optimum data rate, with other data rates yielding

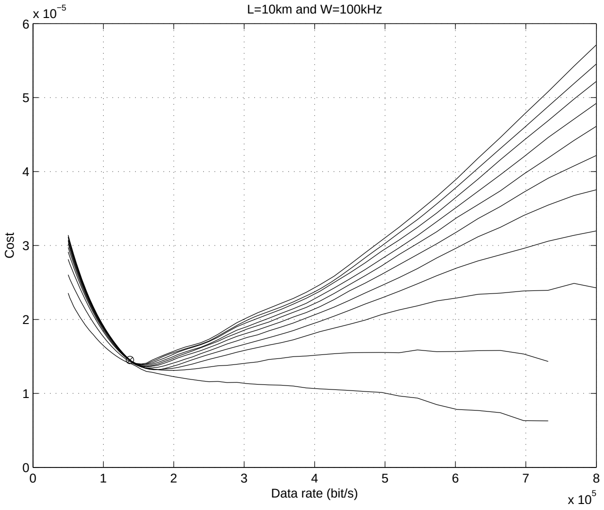

Figure 5.6: Simulation results with $n = 10, 20, 40, 60, \ldots, 200$.

much greater costs. If we try to characterize the parameter values that give the different behaviours we find that low bandwidth, small area and many nodes all tend to result in cost curves with higher costs for high and low data rates and an optimum somewhere in between. We have, however, not found any strict rule by which it is possible to predict the cost curve behaviour without actually doing simulations.

Regarding the theoretical analysis we have seen that the simulated and theoretical cost curves are not very different in absolute values. However, for higher data rates the slope differs and the theoretical curve predicts a higher cost than is actually the case in the simulations. Also, the theoretical analysis does not depend on the number of nodes, something that we have seen in the simulations does make a difference. From this we conclude that there is still room for improvements in the theoretical analysis.

5.4.1 Broadcast Cost Curve Shape

As noted in previously, some broadcast simulation curves have a cost that is strictly decreasing with increasing data rate. Other curves have a local minimum for some optimum data rate, and then grow with increasing data rate. At the same time, both Figure 4.1 and Section 4.2 hint that there *should* be a growing cost for increasing high data rates, or equivalently in the case of Figure 4.1, decreasing short distances. (Figure 4.1 depicts the result from analysing unicast transmissions, but the corresponding result for broadcast transmissions has the same basic shape). An interesting question then, is why the simulation results do not always show the predicted behaviour. One

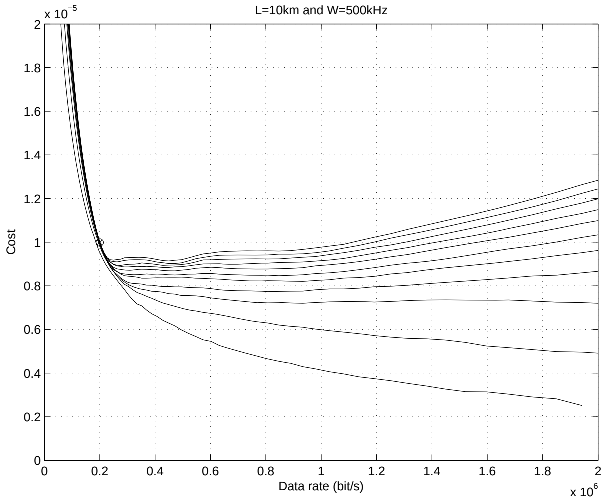

Figure 5.7: Simulation results with $n = 10, 20, 40, 60, \ldots, 200$.

explaination could be that the assumptions underlying the theoretical analyses are too different from the "real" situation in the simulations. One main candidate would be the assumption that the network nodes are dense in the network area, resulting in that the whole area has to be covered and that there in every position is a node available for retransmitting a message. Another candidate is that the theoretical analyses do not take the routing algorithm into account. The discrepancies may be an artifact of the MPR broadcast algorithm. It could also be that the theoretically predicted behaviour is always present in the simulations, but that the networks sometimes become disconnected at a lower data rate than is needed for this to manifest itself. Below we will argue that this last explaination is probably correct.

We will make one more surface covering analysis. The cost of broadcasting a message over an $L \times L$ area can be approximated as: $Cost \propto L^2/d^2R$, where d is the range a node can reach at the rate R, which in turn is the rate used by all nodes in the network. Using the equality in expression (3.3) we can rewrite this as $Cost \propto L^2/d^2W\log_2(1 + K/Wd^4)$, where K is a constant. We study the second term in the log expression and denote it A, such that $A = K/Wd^4$. For large A we can approximate the cost as $Cost \propto L^2/d^2W\log_2(K/Wd^4)$. For small A we can approximate the cost as $Cost \propto L^2/d^2W(K'/Wd^4) = L^2d^2/K'$, where K' is another constant. That is, for large A the cost decreases approximately proportional to $1/d^2$, since the log expression will be dominated by the other factors. For small A the cost grows approximately proportional to d^2. This is essentially the results from the theoretical analyses, but described in terms of range instead of data rate.

With D we denote the cut-off distance at which the network becomes non-connected, i.e., the network will be non-connected if the available range $d < D$. We note that for

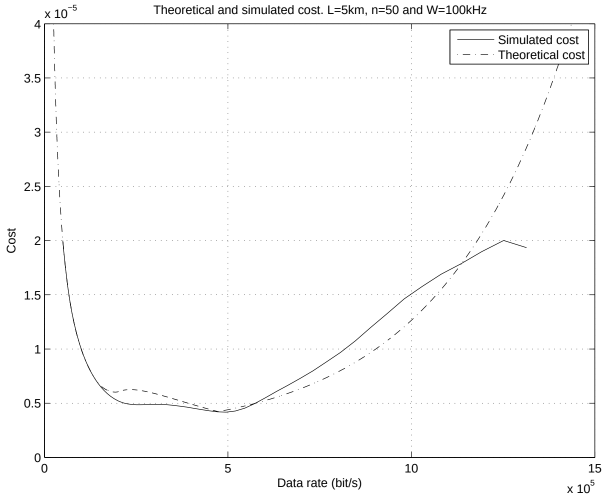

Figure 5.8: Comparison between simulated and theoretical results for L=5, n=50 and W=100.

a given network the value of this cut-off distance is well defined, and does not depend on anything other than the geometry in the network. More specifically, D does not depend on any of the radio or transmission parameters. We now look at Figure 5.11, showing the basic behaviour of the cost as a function of the range, with the cut-off distance visible. We have also marked the range M at which the cost curve attains its minimum value, which will be located somewhere between the ranges for which A is considered large and small respectively.

The cost curve approximation is defined for every positive value of d, but for $d < D$ networks will not be connected, and this part of the curve will not be visible in the simulations. Note that we now have the range on the x-axis, as compared to the figures previously in this chapter where the data rate is used. Data rate and range have an inverse relationship, so that the range d being less than a certain cut-off distance corresponds to the data rate being greater than some cut-off data rate.

There are two interesting things that can be derived from the setup we have done so far. We first note that if the number of nodes in the network is reduced (n decreases), or if the physical size of the network grows (L increases), the node distances in the network will increase, yielding a larger value for D. None of these values change the cost curve approximation, so M does not change. If D grows so that $D > M$ we go from a situation where cost as a function of data rate (not range) has a minimum value followed by increased cost, to a situation where the cost decrease with growing data rate, up until the point where the network become non-connected. This is exactly what was described in Section 5.4 regarding the impact on the curve shape of the number of nodes and the surface area.

We go on and study what happens when W is varied in the approximative cost

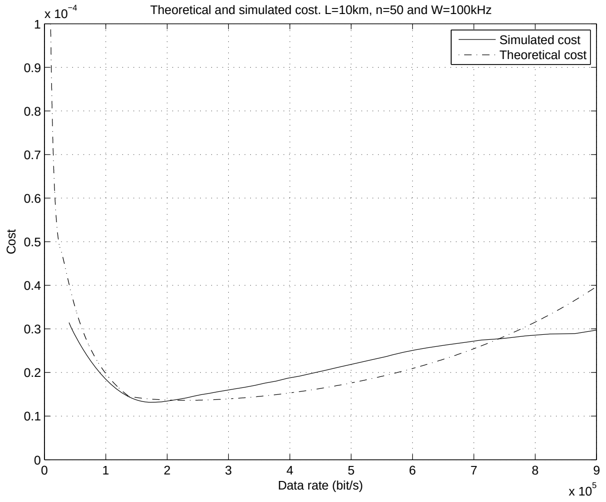

Figure 5.9: Comparison between simulated and theoretical results for L=10, n=50 and W=100.

expressions. Increasing W will not have any impact on D. However, if W is increased d can be smaller and will still yield the same value on the term A. This means that d does not have to be as large as before for A to be considered small. This will compress the cost curve from the right to the left, reducing the value M. If M is reduced so that $M < D$ the situation will change just as described in the previous paragraph, which is exactly what was described in Section 5.4 about the impact of the bandwidth on the curve shape.

We can see that using the surface covering concept we can explain the behaviour described in Section 5.4. We take this as an argument for the value of the concept as such, as well as an argument for the explanation of the behaviour of the cost curves in Section 5.3. In view of this we can describe all of these cost curves as stemming from a single, generic cost curve of which is shown only a section, due to the effect of networks ceasing to be connected.

There are still some parts of the cost curve behaviour that are not explained by the above description. Studying Figure 5.4 we see that the cost grows with increasing number of nodes in the network, something that we have not explained, neither here nor in Chapter 4. However, looking closely at any one subplot in the figure and studying one single data rate, we see that the cost increase is reduced for each higher number of nodes. It looks as if doubling the number of nodes roughly add a constant cost. Our guess is that this increase, slow as it may be, is not a fundamental property of the broadcast problem, but rather an artifact of the imperfections of the MPR broadcast algorithm.

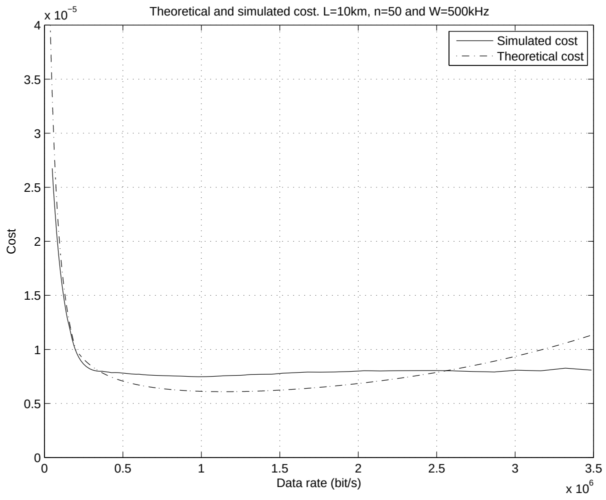

Figure 5.10: Comparison between simulated and theoretical results for L=10, n=50 and W=500.

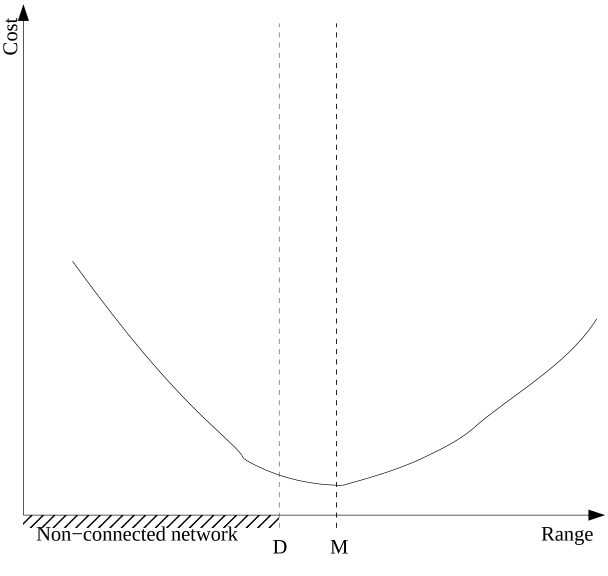

Figure 5.11: Basic shape of cost as a function of range.

6 Discussion and Conclusions

We evaluate the transmission cost in ad hoc networks with a single variable data rate. This is a first step in finding good strategies for multirate broadcast in ad hoc networks, We assess both broadcast traffic with multipoint relay flooding and, for reference, unicast traffic.

A general observation is that the transmission cost is high at very low rates. This is due to the fact that very low rates correspond to a communication range that is greater than necessary, perhaps even covering large areas outside the network range. From the unicast traffic simulations, it seems to be a good strategy to always use the largest possible data rate that does not fragment the network. This gives a transmission cost that is close to its minimal value. Furthermore, the differences betwen the optimal single variable rate and the optimal multirate is not very great in the unicast case.

For the broadcast case, it seems more difficult to formulate a general strategy that reduces the transmission cost. Compared to unicast, we see a much more varied behaviour. For some choices of bandwidth and node density, the transmission cost can be as high as eight times its minimum, depending on the chosen data rate, with a narrow rate range for which the cost is close to its minimum value. In other cases the cost curves looks more like the cost curves in the unicast case. There are also examples of cost curves that are strictly decreasing with increasing data rate, up until the point when the network split. This is the case for networks with a large bandwidth and low node density.

For networks with lower bandwidths and higher node densities, there is a transmission cost minimum close to the data rate for which centrally located nodes can reach all other nodes. For these networks it might be a good strategy to chose the data rate so that we have a two-hop broadcast: one hop in to a central node and one hop to all other nodes in the network. This resembles the transmission behaviour in a cellular base station radio system, which is perhaps not the way we are used to think about ad hoc networks.

6.1 Practical Considerations

All results so far in this report are based on the Shannon capacity for an additive, white, Gaussian channel, which is an upper limit on what can be achieved in practice. In reality, we cannot reach these values of course. With different techniques we may end up rather close to the limit. However, all such techniques require long packet lengths, which may not always be possible in ad hoc networks, which means that we may still be some dBs from the limit. However, this would not really change our results, but the problem from a practical point of view is that it may be more difficult to get close to the limit at low SNRs.

Detecting and synchronizing to a signal gets more difficult the weaker it is. For SNRs below 0 dB, we will mostly need specific methods to detect and synchronize the receiver to the signal. This may require known bit patterns, pilot signals, or other techniques that will lower the effective data rate. For higher data rates, this is much less of a problem and it is easier to reach data rates closer to the Shannon limit.

Since a relayed signal always has a much higher SNR than direct transmission the practical communication problems in the relaying case is much smaller. This effect and consequences of terrain suggests that the results shown in the report underestimate the value of relaying. However, the general result that relaying is more advantageous for unicast communication than for broadcast communication should still be valid.

6.2 Future Directions

We have seen in the previous chapters that the use of a single variable data rate in an ad hoc-network can sometimes have a substantial impact on the broadcast transmission cost. The potential of using multiple rates should be greater. Since different parts of a single ad hoc network can be used in very different surroundings, under very different conditions, we believe there is much to gain from allowing different behaviours in different nodes. However, our experience from the investigation at hand hints that this greater gain may not be visible in the scenario setup used here. When terrain is added into the picture different parts of the network will be more isolated from each other, and will thus have a greater opportunity for gaining from different transmission behaviour.

However, also in the scenario setting of this document multirate solutions show good promise. By adjusting data rate individually at each node during communication it will be possible to dynamically find the locally least costly data rates which the investigations in this documents have shown to exist. If this nodewise optimization succeeds we believe the total gain to be substantially greater than what is possible using a single variable rate. Designing a distributed algorithm for this will be no easy task, but our view is that the potential gain motivates further investigations into this problem.

In our view, the next step is to try a number of specific multirate strategies that have appeared in our discussions during the current project. Evaluating these should make it possible to decide whether the gain from using multirate solutions are big enough in practical situations to make it worth the extra complexity in the communication systems. Some strategies that we would like to try are the following:

- all nodes use the same data rate (as in the current document),
- all nodes start at the same data rate, and then individually choose the highest data rate with the same connectivity as the start rate,
- all nodes individually choose data rates such that they just reaches all other nodes,
- the most centrally located node chooses a data rate such that it just reach all other nodes, while all other nodes individually choose data rates such that they just reach the central node,
- the most centrally located node chooses a data rate such that it just reaches all other nodes, while all other nodes individually choose the cheapest route to the central nodes.

Of course, in large networks, it may be better to use different strategies in different parts of the network.

Bibliography

[1] L. Farman, J. Grönkvist, A. Hansson, E. Johansson, J. Nilsson, K. Persson, M. Sköld, U. Sterner, O. Tronarp, and P. Zeijlon, "Mobila ad hoc-nät - utmaningar och möjligheter," User Report FOI-R--1799--SE, Swedish Defence Research Agency., Div. of Command and Control. Linköping, Sweden, 2003.

[2] ETSI-STC-RES10 Committee, "Radio equipment and systems: High performance radio local area network (hiparlan) type 1," *Functional Specifications ETS*, pp. 300–652, june 1996.

[3] A. Laouti A. Qayyum, L. Viennot, "Multipoint relaying for flooding broadcast messages in mobile wireless networks," in *Proceedings of the 35th Hawaii International Conference on System Sciences*, oct 2002, pp. 3866–3875.

[4] T. Clausen and P. Jacquet, "Optimised link state routing protocol (OLSR)," *RFC 3626*, 2003.

[5] L. Farman, U. Sterner, and O. Tronarp, "Analysis of capacity in ad hoc networks with variable data rates," Technical Report FOI-R--0928--SE, Swedish Defence Research Agency., Div. of Command and Control. Linköping, Sweden, 2003.

[6] L. Ahlin and J. Zander, *Principles of Wireless Communications*, Studentlitteratur, 1997.

[7] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*, Waveland Press, 1965.