Technical University Berlin

Telecommunication Networks Group

Coordination-free Repeater Groups in Wireless Sensor Networks

Andreas Willig

awillig@tkn.tu-berlin.de

Berlin, August 2006

TKN Technical Report TKN-06-004

TKN Technical Reports Series

Editor: Prof. Dr.-Ing. Adam Wolisz

Contents

Chapter 1

Introduction

Sensor networks are not only useful to *observe* the environment, but also to *control* it through actuators [1, 17], for example in building automation applications [21, 13]. In this kind of applications there are both sensor nodes and the actuator nodes. For the sensor nodes many of the considerations usually made for sensor networks [12] apply, including the observation that the individual sensor data packet is not important (and can be suppressed) as long as there are sufficient other sensor nodes which can observe the same data [22]. In contrast, actuators must be individually addressable and often the quality of a control algorithm depends crucially on the network ability to reliably deliver sensor data to the actuator nodes. The focus of this paper is on scenarios where packets should reliably reach selected and individually addressable nodes (henceforth called *destination nodes*) in a sensor network.

In this paper we develop the concept of a *repeater group*. A repeater group is a coordinated and connected group of sensor nodes placed close to the destination node. The group is responsible for receiving *incoming packets* and the members jointly ensure that this packet is received by a destination node (for example the actuator) with high probability. At the same time, the activities of the group members are arranged so that an individual member has enough opportunity to sleep, i.e. can maintain a reasonably low duty cycle. An important characteristic of a repeater group is that here is sufficient geographical separation between members to take advantage of spatial diversity in wireless channels [10]. On the one hand, this arrangement increases the chance that at least a few group members receive and successfully decode an incoming packet.[1] Once this happens, the packet can be communicated to other group members as well. On the other hand, the group members possessing a copy of the incoming packets can decide whether and when they re-transmits the packet. By making sure that these re-transmissions are carried out in an orthogonal (for example in time) manner by many different, geographically separated group members, we can make effective use of the spatial diversity of the wireless channel and give the destination node the possibility to receive the packet over multiple, independently faded wireless channels. Hence, the group can be thought of as *amplifying* incoming packets – this is just what repeaters typically do. The repeater group concept can be regarded as a practical cooperative diversity / cooperative MIMO (single input / multiple output) scheme [14, 15, 9] with additional consideration of node sleeping cycles and transmit/receive operations carried

[1] Due to the diversity gain achievable with multiple receivers, the source of incoming packets can reduce its transmit power while keeping the throughput and target error rate. This argument is especially pronounced in the case when incoming packets are transmitted over a long-haul link [9].

out by the same group of nodes. By relying on a decode-and-forward approach [14, 15] the schemes developed in this paper can be implemented without special support from the physical layer.

Given this core concept of repeater groups, a number of design issues come up. A first design issue concerns the cooperation of the members of the repeater group (henceforth called *repeater nodes* or simply *repeaters*). It has to be ensured that repeaters have sufficient opportunity to spend time in sleep mode, but on the other hand enough repeaters should be awake to pick up and re-transmit the incoming packet. In this paper we consider schemes in which repeaters have some a-priori knowledge about the repeater group (like group size, sleeping policy of nodes and the resulting probability distribution of the number of awake nodes) but do not exchange extra control packets to coordinate their sleep activities or other operational aspects of the group. In general, a repeater has to take all its decisions on the basis this a-priori knowledge and on its observations of the behaviour of its peers. The results of this paper can be regarded as baseline results for more elaborate schemes based on explicit coordination.

A second design issue is to make sure that a sufficient number of awake nodes really do receive the incoming packet and are able to re-transmit it further. Because of channel errors it might well happen that an awake node does not receive the incoming packet – we call such a node a *wave-one* node, whereas the repeaters which have received the incoming packet are *wave-zero nodes*. We present a scheme which allows wave-one nodes to quickly pick up repeated packets from wave-zero nodes and to start their repeating activities later. The scheme performs over a large range of error probabilities almost as good as if there are no channel errors.

A third important issue is how the awake repeaters (wave-zero and wave-one) arrange their transmissions so that the largest possible number of non-overlapping packet re-transmissions coming from different repeaters can potentially be heard at the destination node. To complicate matters, since we avoid explicit coordination in this paper, the number of awake nodes is random, as are the number of wave-zero nodes and the times when wave-one nodes pick up repeated packets and start their activities. In this paper we use a slotted scheme in which each repeater node picks one out of a finite number of slots according to a random distribution. The goal is to maximize the average number of slots in which exactly one repeater transmits. We derive such a scheme and show that it achieves the optimal throughput for slotted ALOHA of $1/e \approx 0.368$ in the case without channel errors, i.e. on average 36% of all slots contain exactly one repeated packet.

A fourth important design issue concerns the handling of immediate MAC-layer acknowledgements for incoming packets. When acknowledgements are required, some coordination is needed between the repeaters to decide who sends the ack. Without MAC layer acknowledgements there is no need for coordination, and since we are interested in coordination-free schemes, we make this assumption throughout the paper.

The paper is structured as follows: In the next Chapter 2 we describe the system model under consideration. In Chapter 3 we develop a baseline scheme for the case without channel errors of incoming packets. Following this, we investigate the impact of channel errors on the baseline scheme in Chapter 4. In Chapter 5 we discuss so-called *quick amplification schemes*, which improve the baseline scheme in case of high error rates. A brief overview on related work is given in Chapter 6 and the paper is concluded in Chapter 7.

Chapter 2

System model

A sketch of the assumed system model is shown in Figure 2.1. The destination node D is shown in the right part of the figure. The overall goal is to transmit at least one valid copy of incoming packets reliably to the target node. The target node as such is not of interest to us, it is assumed to have plenty of energy (this assumption is reasonable for actuators) and other resources, and is awake all the time.[1] The destination node can operate in different modes: it either needs one error-free copy of the incoming packet from any repeater node, or it could be able to combine several erroneous copies [23, 16] coming from different repeater nodes. The schemes in this paper do not take any advantage of packet combining methods, but in general this can be an important design aspect.

We assume a slotted-time model and perfect time synchronization of all the involved nodes. Specifically, we assume that incoming packets arrive periodically at the beginning of so-called *macro slots*. All activities of the repeater group belonging to an incoming packet at the beginning of a macro slot have to end before the beginning of the next macro slot. The incoming packets have all the same size.

The repeater group consists of N nodes. In the figure all the nodes within the grey shape are repeater nodes. Each repeater i decides independently of other nodes at the beginning of a macro slot (before an incoming packet arrives) whether it will sleep dur-

[1] It should be noted that there is nothing in the schemes discussed in this paper which prevents having more than one destination node.

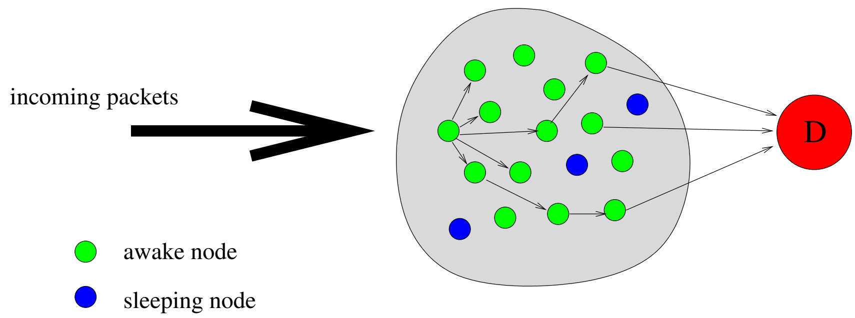

Figure 2.1: System model

ing the macro slot. We assume that this decision is made randomly, and we assume that the numbers $K_1, K_2, K_3, \ldots, K_j, \ldots$ denoting the number of awake nodes in macro slot j forms a sequence of independent and identically distributed random variables. The generic random variable is called K. The common distribution function $F_K(\cdot)$ of $K_1, K_2, \ldots$ is known to all the members of the repeater group, but the realizations of these random variables are not known nor are they tracked by the repeaters. In the specific case where each node decides on the basis of an independent Bernoulli experiment with success (awake-) probability s, the random variables K_i have a binomial distribution with parameters N and s. We call such a repeater group a *binomial repeater group*. It is assumed additionally that the repeaters also know a unique identifier for the repeater group.

The group is assumed to be connected and the members have a reasonably high neighborhood degree. Group members can receive packets of neighbored members with high probability. From a physical perspective, the group members should have a mutual distance of at least half a wavelength. At this distance, the fading observed on different wireless links starts to become independent [19, Chap. 5].

Assume that the incoming packet is received at the beginning of a macro slot (there are no immediate MAC layer acknowledgements). The following time is subdivided into a number M of time slots, numbered from 1 to M. The parameter M is known to all repeaters. The slot size is large enough to accommodate a repeated packet. A repeated packet is generated by a repeater node from adding a small flag and the repeater group identifier to the incoming packet. The flag simply identifies the repeated packet as such. The repeater nodes receiving the incoming packet (which we call *wave-zero* repeaters) pick one of the M slots randomly for transmission or decide to keep quiet. A wave-zero repeater picks time slot $t \in \{1, \ldots, M\}$ with probability p_t^0 or remains quiet with probability p_{M+1}^0. For simplicity we assume that whenever two or more repeaters pick the same slot for transmission a collision arises, rendering the repeated packets useless. Such a slot is called a *collided slot*. When none of the repeaters transmits in a slot, we refer to it as an *empty slot*, whereas when exactly one repeater transmits in a slot we call it a *successful slot*. Now assume that slot t^* is a succesful slot, and repeater R_a is the one transmitting in this slot. It might happen that another repeater R_b which has not received the original incoming packet picks up the repeated packet. Such a repeater R_b is called a *wave-one* repeater. It either picks randomly one of the slots $t \in \{t^* + 1, \ldots, M\}$ (each with probability p_t^1) or remains quiet with probability p_{M+1}^1. In general, p_t^1 and p_{M+1}^1 might also depend on the slot number t^*, but we drop this since we make no further use of this in this paper.

Our goal is to maximize the number of successful slots. However, since the number of awake nodes and their transmission decisions are random variables, in general we want to optimize the **expected** number of successful slots.

Chapter 3

The case without channel errors

We first look at a system where all incoming packets are reliably received by the awake nodes. Stated differently: all of the K nodes are wave-zero nodes and can operate in a very energy-efficient manner: a wave-zero node either decides to remain quiet (and thus can sleep for the remaining macro cycle) or transmits in a single slot without listening to other slots.

The protocol design knobs are the probabilities p_t^0 for $t \in \{1, \ldots, M\}$ and p_{M+1}^0. We abbreviate these probabilities as p_t with $t \in \{1, \ldots, M+1\}$, where p_{M+1} is the probability that the repeater node remains quiet. We abbreviate the vector of probabilities as $\pi = (p_1, p_2, \ldots, p_M, p_{M+1})$ and note that π is a probability distribution. The goal is to choose these probabilities such that the expected number of successful slots is maximized. We now formulate this problem more concisely.

Suppose first that the number of awake nodes K is fixed and known and that all awake nodes receive the incoming packet. Define the random variables $X_{i,j}$ as follows:

$$X_{i,j} = \begin{cases} 1 & : \text{ node } i \in \{1, \ldots, K\} \text{ transmits in slot } j \in \{1, \ldots, M+1\} \\ 0 & : \text{ otherwise} \end{cases}$$

Since we assume that the nodes make their choice independently and all nodes use the same probability distribution π, for each fixed j the $X_{i,j}$ are iid random variables. Furthermore, for fixed i we have

$$\sum_{j=1}^{M+1} X_{i,j} = 1$$

The number Y_j of nodes transmitting in slot j, defined as:

$$Y_j = \sum_{i=1}^{K} X_{i,j}$$

is a sum of iid random variables and hence has a binomial distribution $Y_j \sim \text{Binomial}\,(K, p_j)$. The indicator variable Z_j is defined as:

$$Z_j = \begin{cases} 1 & : & Y_j = 1 \\ 0 & : & Y_j \neq 1 \end{cases}$$

and indicates the (desired) event that in slot j exactly one repeater node transmits. Clearly, we have:

$$\Pr\left[Z_j = 1\right] = \Pr\left[Y_j = 1\right] = b(1; K, p_j) = K \cdot p_j \cdot (1 - p_j)^{K-1}$$

where $b(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$ is the probability mass function of the binomial distribution. The average number of slots in which exactly one repeater node transmits under distribution π is given by:

$$f(\pi) = f(p_1, \ldots, p_{M+1}) = E\left[Z_1 + \ldots + Z_M\right] = \sum_{j=1}^{M} E\left[Z_j\right] = \sum_{j=1}^{M} K \cdot p_j \cdot (1 - p_j)^{K-1}$$

To find the optimal distribution π, we have to solve the following nonlinear constrained optimization problem:

$$\begin{aligned}
\text{maximize} \quad & f(p_1, \ldots, p_{M+1}) \\
\text{subject to} \quad & h(p_1, \ldots, p_{M+1}) = 1 - \sum_{i=1}^{M+1} p_i = 0 \\
& \mathbf{g}(p_1, \ldots, p_{M+1}) = \begin{pmatrix} g_1(p_1, \ldots, p_{M+1}) \\ g_2(p_1, \ldots, p_{M+1}) \\ \cdots \\ g_{M+1}(p_1, \ldots, p_{M+1}) \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ \cdots \\ p_{M+1} \end{pmatrix} \geq \mathbf{0}
\end{aligned}$$

We show in Appendix A that we can simplify this problem: the optimal probability distribution assigns to all slots $1, \ldots, M$ the same probability, i.e. $p_1 = p_2 = \ldots = p_M =: p$, with a true definition on the right hand side. Let further denote $q = p_{M+1}$. Hence, the problem can be reformulated as:

$$\begin{aligned}
\text{maximize} \quad & f(p, q) = M \cdot K \cdot p \cdot (1 - p)^{K-1} \\
\text{subject to} \quad & h(p, q) = 1 - M \cdot p - q = 0 \\
& p \geq 0, q \geq 0
\end{aligned}$$

The parameter p is restricted to the interval $\left[0, \frac{1}{M}\right]$. Obviously, $f(0) = 0$ and

$$f\left(\frac{1}{M}\right) = K \cdot \left(\frac{M-1}{M}\right)^{K-1}$$

It is shown in Appendix A that the value p_{opt} which maximizes $f(\cdot)$ is given by:

$$p_{\text{opt}} = \begin{cases} \frac{1}{K} & : \quad K > M \\ \frac{1}{M} & : \quad K \leq M \end{cases} \tag{3.1}$$

If we fix M, then for $K \to \infty$ the expected fraction of slots in which exactly one node repeats a packet is given by:

$$\lim_{K \to \infty} \frac{f\left(\frac{1}{K}\right)}{M} = \lim_{K \to \infty} \left(1 - \frac{1}{K}\right)^{K-1} = \frac{1}{e} \approx 0.368$$

which confirms that the proposed choice of p_{opt} gives indeed the theoretical maximal throughput of slotted ALOHA for a large population of stations [4, Sec. 4.2].

However, a repeater node does not know K, it knows only M and the distribution of K, but it still has to make a choice of its parameter p. So, instead of maximizing $f(\cdot)$ for known value of K, we choose to maximize:

$$F(p,q) = E\left[M \cdot K \cdot p \cdot (1-p)^{K-1}\right] = M \cdot \frac{p}{1-p} \cdot E\left[K(1-p)^K\right] \tag{3.2}$$

subject to $0 < p < 1$ (the expectation is taken with respect to K). It is shown in Appendix B that $E\left[K(1-p)^K\right]$ can be represented in two different ways:

$$E\left[K(1-p)^K\right] = \sum_{n=1}^{\infty} \frac{E\left[K^n\right]}{(n-1)!} \cdot (\log(1-p))^{n-1} \tag{3.3}$$

provided that all moments $E\left[K^n\right]$ of the distribution of K exist, which, however, is guaranteed for all discrete distributions with finite range. The second representation is:

$$E\left[Ka^K\right] = a\frac{d}{da}\Phi_K(\log(a)) \tag{3.4}$$

evaluated at $a = 1-p$, where $\Phi_K(x) = E\left[e^{xK}\right]$ is the moment-generating function of the random variable K.

A number of strategies can now be used to choose the parameter p optimizing Equation 3.2:

- Motivated by Equation 3.1 one choice could be:

  $$p_* = \frac{1}{M} \cdot \Pr\left[K \leq M\right] + \sum_{k=M+1}^{\infty} \frac{1}{k} \cdot \Pr\left[K = k\right]$$

  which does not depend on K but only on the (known) distribution of K.

- One can determine the optimal p taking Equation 3.4 into account, i.e. from optimizing:

  $$F(p,q) = M \cdot \frac{p}{1-p} \cdot (1-p) \left.\frac{d}{da}\Phi_K(\log(a))\right|_{a=1-p} = M \cdot p \cdot \left.\frac{d}{da}\Phi_K(\log(a))\right|_{a=1-p}$$

  We will call the value that maximizes the previous expression p_{mgf}.

- When the moment-generating function $\Phi_K(\cdot)$ of K is not available or not easily manipulable, the moment-represenation of Equation 3.3 can be exploited. One way is to obtain an approximation by truncating the moment representation after a number n of terms,[1] i.e. to optimize

  $$F_1(p,q) = M \cdot \frac{p}{1-p} \cdot \left(E\left[K\right] + \sum_{i=2}^{n} E\left[K^i\right] \frac{(\log(1-p))^{i-1}}{(i-1)!}\right)$$

  as target function. This requires knowledge of up to n moments of K, and the optimization of p is best done numerically.

[1] At least two terms are required, since for only one term the resulting expression

$$M \cdot \frac{p}{1-p} \cdot E\left[K\right]$$

has no local maximum in $(0,1)$, instead, this expression diverges for $p \to 1$.

We compare these different choices for the special case of a binomial repeater group. Assume that we have N sensor nodes in total, and each of these N nodes makes an independent and time-homogeneous decision whether to sleep (probability $1 - s$) or whether to stay awake (probability s) in the next macro slot. Hence, $K \sim \text{Binomial}(N, s)$. The moment-generating function for the random variable K is:

$$\Phi_K(a) = E\left[e^{aK}\right] = (1 + s(e^a - 1))^N$$

and we have:

$$\begin{aligned}\Psi_K(a) &= \frac{d}{da}\Phi_K(\log(a)) \\ &= \frac{d}{da}\left(1 + s\left(e^{\log(a)} - 1\right)\right)^N \\ &= \frac{d}{da}(1 + s(a-1))^N = s \cdot N \cdot (1 + s \cdot a - s)^{N-1}\end{aligned}$$

and

$$\Psi_K(1-p) = N \cdot s \cdot (1 - sp)^{N-1}$$

Therefore we have:

$$F(p, q) = M \cdot p \cdot N \cdot s \cdot (1 - sp)^{N-1}$$

which for $0 < s < 1$ achieves its maximal value in $(0, 1)$ for

$$p_{\text{mgf}} = \frac{1}{N \cdot s} = \frac{1}{E[K]}$$

However, to satisfy the constraint that $M \cdot p_{\text{mgf}} \leq 1$ we choose:

$$p_{\text{mgf}} = \min\left\{\frac{1}{M}, \frac{1}{E[K]}\right\}$$

To test the quality of approximations based on the moment representation, we have used representations where the series is truncated after the second, third, fourth or fifth moment. For each of these representations the optimal p for a given binomial distribution is obtained numerically.

The results of a numerical study with a group of $N = 100$ repeater nodes and $M = 20$ slots per macro slot are presented in Figure 3.1. We have varied the probability s that a group member stays awake during a macro slot. For each value of s the values for p_*, p_{mgf} and the optimal p-values for the truncated moment representations have been computed and used subsequently to determine the expected number of successful slots under the respective probability parameter. The results show that:

- As expected, p_{mgf} provides indeed the optimal performance, but the differences between p_* and p_{mgf} are quite small. When p_{mgf} is used, the optimal expected number of successful slots converges to ≈ 7.394593 (which is very close to $20/e \approx 7.36$).
- The truncated moment representations do not perform well. Those which are truncated after an even number of moments converge to a constant value for increasing s, but stay below the performance achievable with p_* and p_{mgf}. Truncating after the fourth moment gives better performance than truncating after the second moment. The curves for truncating after the third and after the fifth moment are identical, and they decay for increasing s.

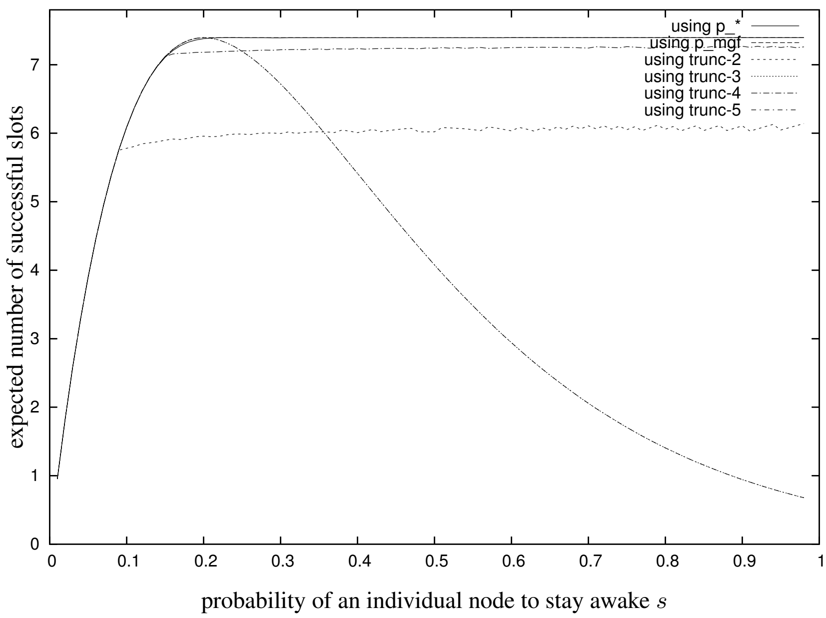

Figure 3.1: Expected number of slots with exactly one repeated packet vs. probability s of an individual node to stay awake in a repeater group of $N = 100$ nodes, $M = 20$.

- Heuristically, all reasonable, i.e close-to-optimal policies have achieved their best throughput consistently for $E[K] \geq M$. This makes sense intuitively, since for $E[K] < M$ on average some slots remain unused.

Chapter 4

The case with channel errors

Next we include channel errors into our considerations. It might happen that a repeater node misses either the incoming packet or even the repeated packets. If a repeater node picks up a repeated packet in the m-th ($m \geq 1$) slot, it might decide to repeat it in the remaining $M - m$ slots, picking each slot with probability p (we assume that p is either p_* or p_{mgf}), or it remains quiet with probability $1 - (M - m)p$. We refer to this scheme as the *baseline scheme*.

We assume that in general the packet error rate for incoming packets is not known to the repeaters.[1] This implies that, although the distribution for the number K of awake repeaters is known to the repeaters, they do not know the distribution of the number W_0 of wave-zero nodes, and of course they do not know the number $W_1 \leq K - W_0$ of wave-one nodes. With respect to energy consumption, the wave-zero nodes receive the incoming packet and transmit their packet in the chosen slot (or remain quiet) and have no disadvantage against the case without channel errors. The wave-one nodes, however, wait until they pick up a repeated packet and then either repeat it in one of the remaining slots or remain quiet. It takes in general a random number of slots before a wave-one node picks up a packet.

We investigate the influence of channel errors on the achievable expected number of successful slots by simulation. Specifically, we assume a binomial repeater group of $N = 100$ nodes with a probability of $s = 0.4$ to be awake during a macro slot, hence there are 40 awake nodes on average. The number of slots is $M = 20$ and each node picks one of each slots available to him with probability $p = p_{\text{mgf}} = 1/40$ as derived for the binomial distribution. It is assumed for simplicity that all repeaters have the same probability P_I to receive an incoming packet and that the different repeaters are independent. The parameter P_I is varied. Furthermore, a repeater node successfully receives a packet from another repeater node with fixed probability $P_R = 0.9$.

For each value of P_I a number of 20.000 macro slots is simulated. The confidence intervals for the average number of successful slots and a confidence level of 99% are quite tight and not shown in the figures. In Figure 4.1 the average number of successful slots is shown versus P_I. It can be seen that the larger P_I becomes, the higher the average number of successful slots, converging to the optimal value ≈ 7.39 as obtained in Chapter 3.

In Figure 4.2 we show for different values of the reception probability P_I and for

[1] Since there can be many sources of incoming packets it is not even meaningful to think about "the" packet error rate, let alone the fact that wireless channel error rates are often time-varying anyway [25].

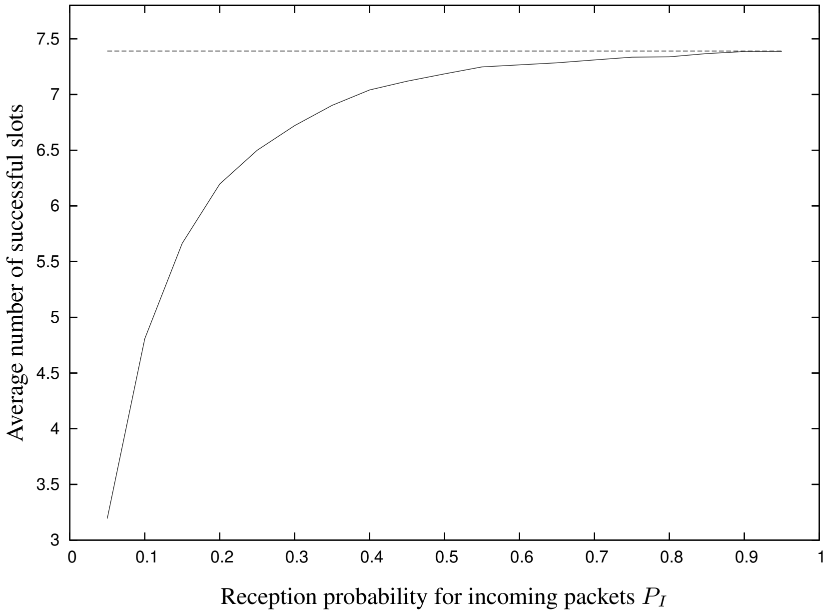

Figure 4.1: Expected number of successful slots vs. probability P_I that an awake repeater node receives the incoming packet ($N = 100$, $s = 0.4$, $M = 20$).

each of the $M = 20$ slots the average number of repeaters sending a packet in the respective slot. The following points are remarkable:

- For small values of P_I the curves display a significantly asymmetric distribution of repeater accesses over the slots, and the optimal value of one repeater on average transmitting in a slot is not reached. The first few of the M slots are rarely occupied. This can be explained as follows: for small P_I the number W_0 of wave-zero nodes is small. Out of these W_0 nodes some decide to keep quiet, others select a random slot out of the M slots. If the first successful slot appears late, all the wave-one nodes can only use the remaining slots, leading to a higher utilization of the late slots.
- When P_I increases towards one, the average number of repeaters in a slot tends towards a uniform distribution over all slots and to an average number of one, just as desired.

Finally, in Figure 4.3 we display the probability that a slot is successful for different reception probabilities P_I. Similar to Figure 4.2, the distribution is asymmetric for small values of P_I and converges towards the uniform distribution as P_I increases.

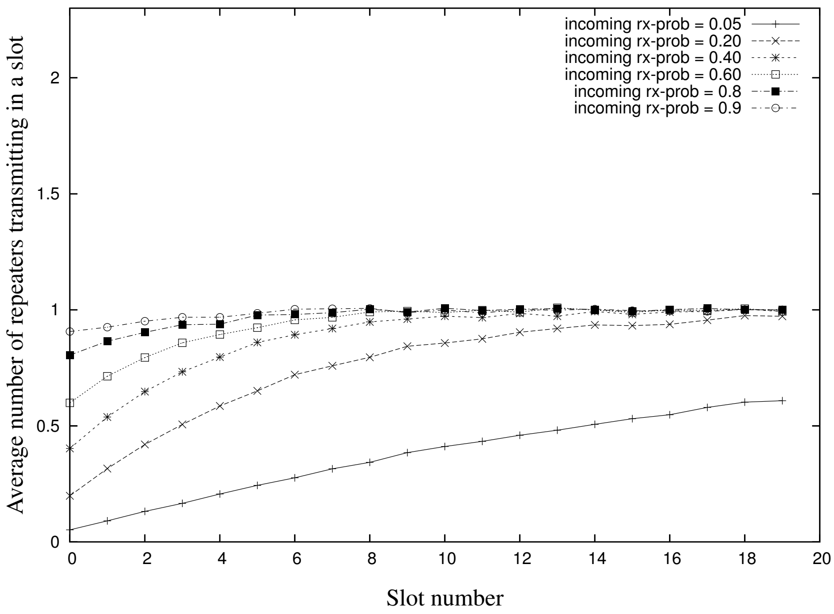

Figure 4.2: Average number of repeaters transmitting in a slot versus slot number for different values of the reception probability P_I to receive an incoming packet ($N = 100$, $s = 0.4$, $M = 20$).

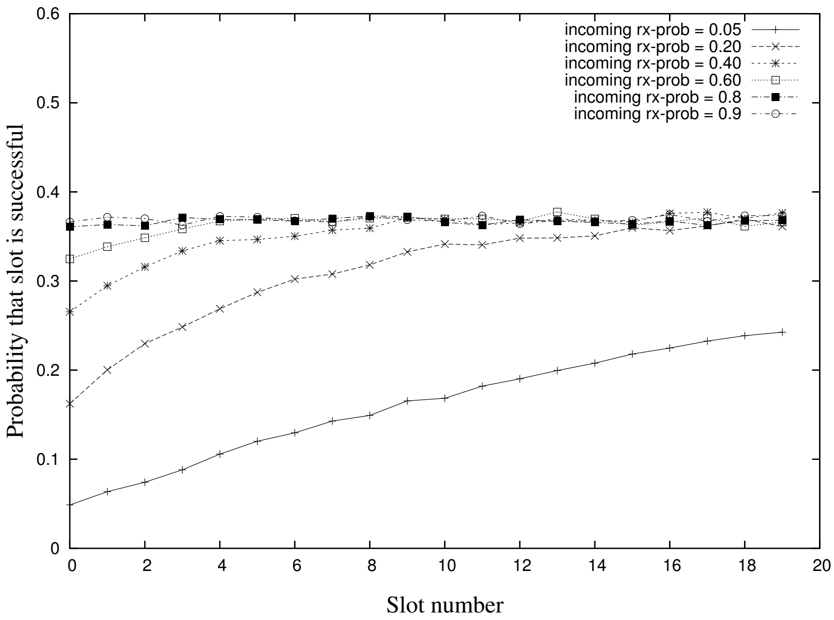

Figure 4.3: Probability that a given slot is successful versus slot number for different values of the reception probability P_I to receive an incoming packet ($N = 100$, $s = 0.4$, $M = 20$).

Chapter 5

Quick Amplification Schemes

In the previous chapter we have applied the baseline scheme to a setup where repeater nodes might fail to receive incoming packets. We have observed that especially for small packet reception probabilities P_I the average number of successful slots is not optimal. This can be attributed to the following reasons:

- Depending on the packet error rate and K, the number W_0 of wave-zero nodes might be quite small. If $E[K]$ is larger than M, then a wave-zero node might decide to remain quiet (with probability $1 - M \cdot p$), again reducing the number of repeated packets. If in addition the first successful slot occurs late, the number of repeated packets reaching the destination will be small.

- When a wave-one node receives the packet in slot m, it then has $M - m$ slots remaining in which it can repeat the packet. If it uses probability p in each of these slots, then the wave-one node remains quiet with probability $1-(M-m)p$, and hence remains quiet with higher probability than the wave-zero nodes.

Therefore, we aim to design what we call *quick-amplification* schemes, satisfying the following goals:

- For small packet reception probabilities P_I the wave-zero nodes should operate in a manner that creates a successful slot as quickly as possible. This way, the wave-one nodes have many remaining slots at their disposal and repeat the packet with almost the optimal probability $M \cdot p$.

- For large packet reception probabilities P_I the operation of the wave-zero nodes should allow to get as close to the theoretical optimum as possible.

- We are interested in schemes that avoid the transfer of separate coordination messages, in order to keep the extra overhead in terms of processing and bandwidth small.

- We are interested in schemes that avoid usage of historical knowledge like estimates of K or W_0 from previous cycles. This is motivated by the fact that wireless channels in general are time-variable [25] and by the consideration that the source of incoming packets might change over time, too.

Any such scheme should work aggressively when W_0 is small in order to activate the wave-one nodes as quickly as possible, but on the other hand, when W_0 is large, its operation should not be so aggressive that too many slots are wasted with collisions.

5.1 Truncated geometric scheme

The first class of schemes, called *truncated geometric scheme*, lets a wave-zero node A observe the channel for a certain number m of slots (m is a design parameter to be determined) and then A makes a decision whether it behaves according to the baseline scheme or in a more aggressive way. Specifically, the scheme is as follows:

- Immediately after receiving the incoming packet a wave-zero node A picks each of the M slots with probability p or remains quiet with probability $1 - M \cdot p$.
- If node A itself chooses one of the first m slots for repeating the packet, then it transmits the packet in this slot and performs no further action, i.e. it behaves according to the baseline scheme.
- If node A has chosen a slot beyond the m-th slot or has chosen to remain quiet, it observes the first m slots. If one of the first m slots is non-empty, node A proceeds according to the baseline scheme. On the other hand, if all m slots are empty, then node A revises its decision to transmit in later slot or to keep quiet and behaves in the following way: node A is guaranteed to transmit and it chooses one of the remaining slots $m+1, m+2, \ldots, M$ according to a probability distribution $\mathbf{r} = r_{m+1}, r_{m+1}, \ldots, r_M$ with $r_{m+1} + \ldots + r_M = 1$.

Unfortunately, the optimal choice of m and $\mathbf{r}$ depends on the distribution of W_0, which in general is not known and hard to estimate in a time-varying environment. Regarding the choice of m, it should be small on the one hand to avoid wasting too much slots for detecting a small value of W_0, but on the other hand it should be large enough so that the probability of false positives (i.e. of large values of W_0 despite the first m slots being empty) is reasonably small. Otherwise, the large number of wave-zero nodes would put too much pressure on the remaining $M - m$ slots (because they transmit with probability one in one of those slots).

Regarding the choice of $\mathbf{r}$ our first intuition is that earlier slots should carry more probability mass to produce the first successful slots quickly. Hence, $r_{m+1} \geq r_{m+2} \geq \ldots \geq r_M$ should hold. Following this intuition, we have specifically looked into truncated geometric distributions, i.e. for r_{m+k} we choose:

$$r_{m+k} = \frac{q^k}{\sum_{i=1}^{M-m} q^i}$$

for some parameter $0 < q < 1$. Smaller values of q shift most probability mass into the first few slots r_{m+1} and r_{m+2}, whereas values close to one let the distribution appear almost uniform.

We have investigated this scheme by simulation for $m = 2$, $m = 3$ and $m = 4$, for different values of q ($q \in \{0.6, 0.7, 0.8, 0.9, 0.99\}$) and for varying probability P_I to receive an incoming packet. The simulation setup was the same as in Chapter 4 ($M = 20$, $N = 100$, $s = 0.4$, simulation for 20000 macro slots, varying P_I). The results for $m = 2$ are shown in Figure 5.1, the results for $m = 3$ are shown in Figure 5.2 and the results for $m = 4$ are shown in Figure 5.3. In each of these figures we have included the results for the baseline scheme (see also Figure 4.1) for easy comparison. The following points are remarkable:

- In all cases, the baseline scheme is the best one for $P_I \geq 0.4$, but the difference between the baseline scheme and the best truncated geometric scheme (attained

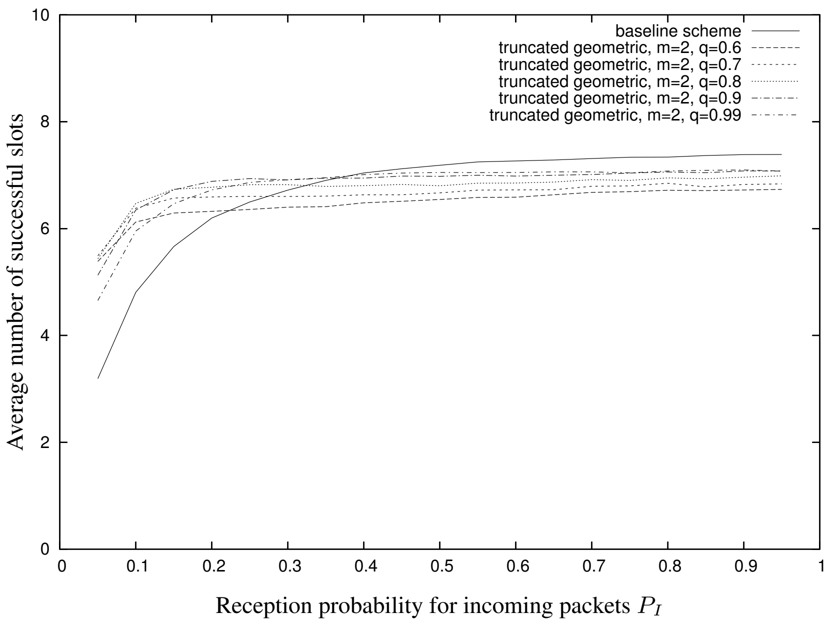

Figure 5.1: Average number of successful slots for the unmodified scheme and the truncated geometric scheme for $m = 2$ and different values of q versus reception probability P_I to receive an incoming packet ($N = 100$, $s = 0.4$, $M = 20$).

for $q = 0.99$ for $m = 2$, $m = 3$ and $m = 4$) becomes smaller as m increases. This can be explained as follows: for $m = 4$ the probability of a false positive is smallest, so that comparatively few wrong decisions are made. In case of a wrong decision the number W_0 of wave-zero nodes is comparably high, and these transmit with probability one in one of the remaining $M - m$ slots, leading to a situation where the (conditional) average number of repeaters transmitting in those slots exceeds the optimal value of one, resulting in an increased number of collisions. The finding that for $P_I \geq 0.4$ always the value $q = 0.99$ is optimal can be explained as follows: in case of a wrong decision about the magnitude of W_0 it is best to distribute the wave-zero nodes uniformly over the remaining slots. For smaller values of q the first few of the $M - m$ slots tend to be wasted in collisions. Consistently, in the range between $P_I = 0.4$ and $P_I = 0.95$ the scheme with $q = 0.9$ is the second-best one.

- In all cases it is true that for $P_I \leq 0.2$ all truncated geometric schemes are better than the baseline scheme. In this regime, for all $m \in \{2, 3, 4\}$, for the smallest values of P_I the schemes with $q = 0.8$, $q = 0.7$ and $q = 0.8$ perform very similar, with varying ranking, in the range between $P_I = 0.15$ and $P_I = 0.2$ the scheme with $q = 0.9$ performs best.

To reduce complexity, we focus the following discussion on the truncated geometric schemes with $q = 0.9$. These are consistently the second-best one, and for intermediate values of P_I they are even the best ones among the truncated geometric schemes. The curves for $q = 0.9$ and $m = 2$, $m = 3$ and $m = 4$ are displayed together with the curve for the baseline scheme in Figure 5.4. This figure highlights another finding:

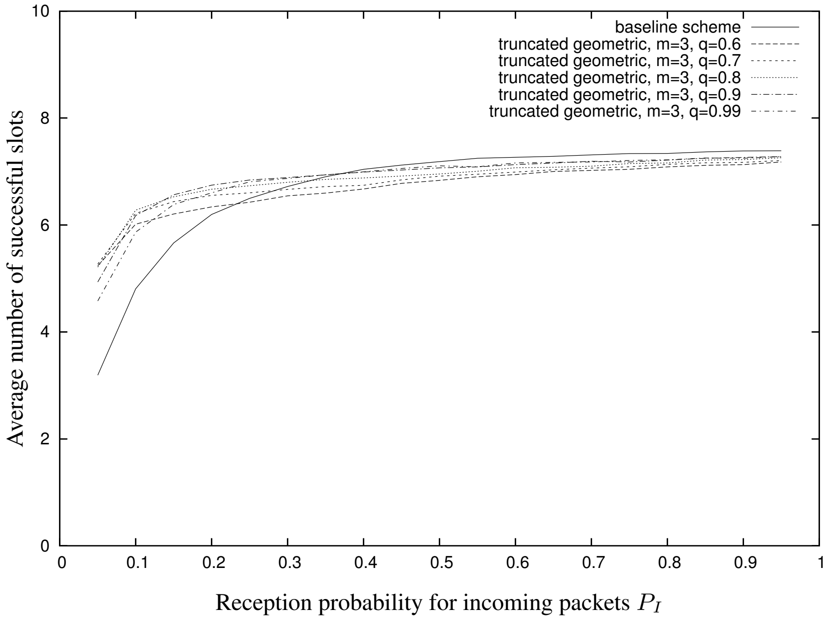

Figure 5.2: Average number of successful slots for the unmodified scheme and the truncated geometric scheme for $m = 3$ and different values of q versus reception probability P_I to receive an incoming packet ($N = 100$, $s = 0.4$, $M = 20$).

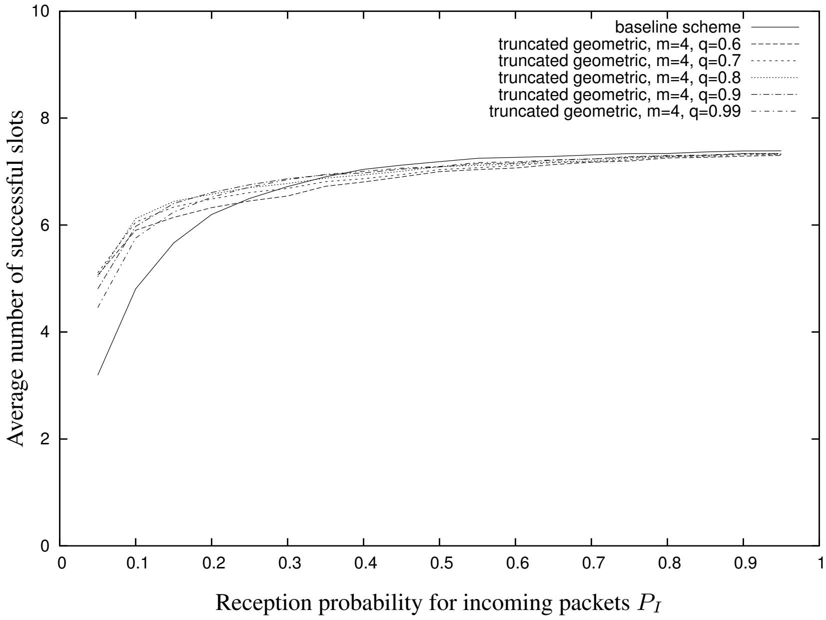

Figure 5.3: Average number of successful slots for the unmodified scheme and the truncated geometric scheme for $m = 2$ and different values of q versus reception probability P_I to receive an incoming packet ($N = 100$, $s = 0.4$, $M = 20$).

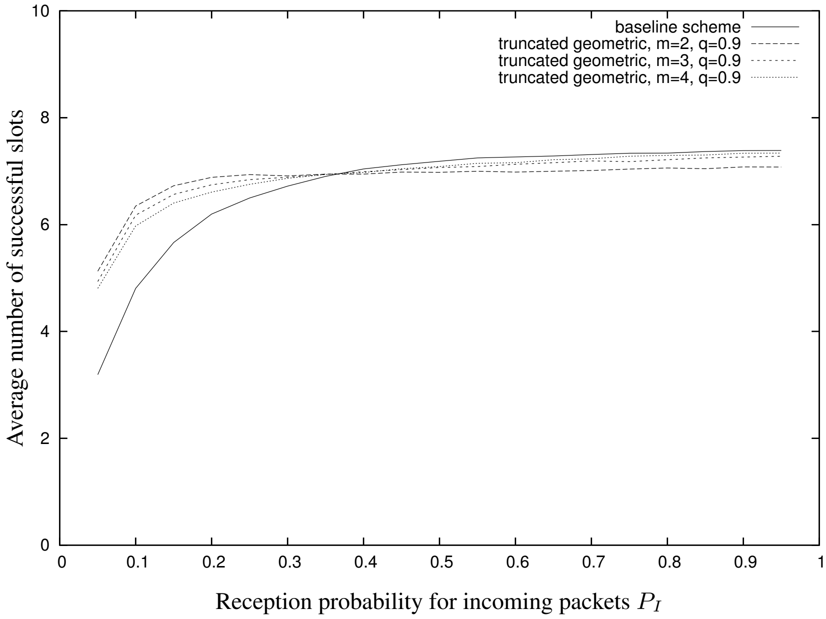

Figure 5.4: Average number of successful slots for the unmodified scheme and the truncated geometric scheme for $q = 0.9$ and different values of $m \in \{2, 3, 4\}$ versus reception probability P_I to receive an incoming packet ($N = 100$, $s = 0.4$, $M = 20$).

for values of $P_I \leq 0.35$, the scheme with $m = 2$ is the best, while it is the worst one for $P_I > 0.35$. Conversely, the scheme with $m = 4$ is the worst one among the truncated geometric schemes for $P_I \leq 0.35$ and the best one for $P_I > 0.35$. The relative advantage of the schemes with smaller m for small values of P_I can be explained as follows: for $m = 2$ and small P_I the absolute rate of false positives is a priori small (since P_I is small and W_0 is hence on average small, too!) and for $m = 2$ simply the number of remaining slots is the largest. In addition, since W_0 is small on average, there is only a minor distortion from the always-transmitting wave-zero nodes to the wave-one nodes.

Does the truncated geometric scheme give the optimal average number of successful slots for small values of P_I? The following small calculation shows that the truncated geometric schemes investigated here are not optimal. Consider as an example the case of $P_I = 0.05$, $m = 2$ and $q = 0.9$. The simulations showed 3110 out of 20000 rounds with no successful slot. Out of those, only 2692 rounds showed no transmission at all, i.e. in 2692 rounds out of 20000 we have $W_0 = 0$. If everything else is optimal, then there should be $\approx 7.39 \cdot \frac{20000-2692}{20000} \approx 6.39$ good slots per round on average. The best that any of the truncated geometric schemes has achieved is ≈ 5.49 average successful slots per round (attained by $m = 2$ and $q = 0.7$). Even if we take the $m = 2$ "wasted" slots into account, the optimally achievable average number of successful slots would be $7.39 \cdot \frac{20000-2692}{20000} \cdot \frac{20-2}{20} \approx 5.75$. One explanation is revealed by further analysis of the data: looking at all rounds where the truncated geometric scheme is triggered (i.e. where the first $m = 2$ slots have been empty), the first successful slot is observed on average later (slot 5.8) than on the total average (slot 5.25). This means that on average almost six slots have gone before the wave-one nodes come into action.

Hence, there is room for improvement.

5.2 Contention scheme

The truncated geometric scheme does not directly estimate the number W_0 of wave-zero nodes, but from passively observing the first m slots a wave-zero node makes an inference about the number $V_0 \leq W_0$ of wave-zero nodes which actually have decided to transmit. However, V_0 can be smaller than W_0 (on average we have $E[V_0] = M \cdot p \cdot E[W_0]$). If W_0 (and therefore V_0) are indeed small, the first m slots of the truncated geometric scheme are likely empty. Therefore, we are interested in schemes, in which the wave-zero nodes do not leave the first m slots empty but try to produce the first successful slot as quickly as possible and in which *all* wave-zero nodes participate in this effort. Hence, the wave-zero nodes should avoid any early decision to remain quiet. However, to avoid excessive collisions when W_0 is large, the majority of the participating nodes (henceforth called *contenders*) should be removed quickly. We have designed a scheme based on these considerations, it is called the *contention scheme*. This scheme aims to eliminate most contenders quickly, somewhat similar in spirit to distributed tree-based contention-resolution schemes [5]. Its operation is as follows:

- Be node A a wave-zero node. Immediately after receiving the incoming packet it starts in the so-called *contention mode*.
- In each slot i out of the first m slots, a contender node A either transmits with probability r_i or decides to listen with probability $1 - r_i$ (all these decisions for subsequent rounds are drawn independently). If A has decided to listen, the following outcomes are possible:
  - If slot i is empty, then node A remains in the contention mode and chooses transmit probability r_{i+1} for the next slot.
  - If node A perceives activity in slot i (it is not necessary that A receives a correct packet), it leaves the contention mode and picks one of the $M - m$ last slots (i.e. it avoids the m slots allocated for the contention phase), each with probability p or decides to remain quiet.
- After the first m slots all remaining contenders leave the contention mode and pick one of the $M - m$ last slots, each with probability p, or it remains quiet.

Of course, the efficiency of this scheme in eliminating contenders and in producing a successful slot quickly depends on the choice of m, $r_1, r_2, \ldots, r_m$. In Appendix C we present a Markov chain model for the contention scheme. It is shown that under a few simplifying assumptions (W_0 is fixed and known, m is large) that the average value of the number T of slots needed until the first successful slot has shown up can be represented as:

$$E[T] = 1 + \sum_{x=2}^{W_0-1} k_x \cdot b(x; W_0, r_1) + k_{W_0} \cdot \left(r_1^{W_0} + (1 - r_1)^{W_0} \right) \tag{5.1}$$

where $b(k; n, p) = \binom{n}{k} p^k (1 - p)^{n-k}$ is the probability mass function of the binomial distribution with parameters n and p, and k_i ($1 \leq i \leq W_0$) is uniquely determined by

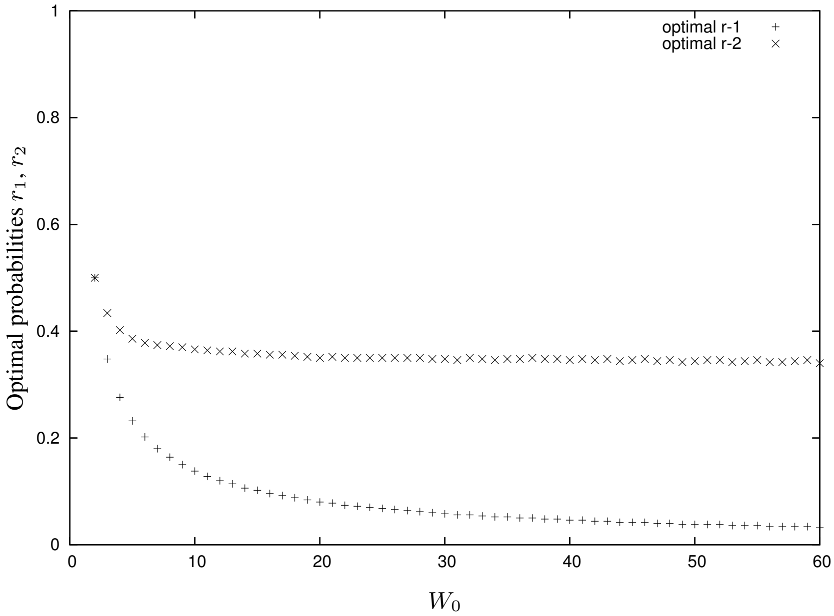

Figure 5.5: Optimal values for transmit probabilities r_1 and r_2 of the contention scheme versus the number W_0 of wave-zero nodes

the following recursive equations:

$$k_1 = 0$$
$$k_i = \frac{1 + \sum_{j=2}^{i-1} b(j; i, r_2) \cdot k_j}{1 - (1 - r_2)^i - r_2^i}$$

This model has been validated by comparing numerical results for $E[T]$ with the result of simulations. The results show an excellent correspondence between theoretical and simulated results.

Please note that Equation 5.1 depends on the three parameters W_0, r_1 and r_2. In fact, $E[T]$ is a rational function of the parameters r_1 and r_2 and can theoretically be minimized for those parameters. The minimum is guaranteed to exist, since $E[T]$ is continuous and r_1, r_2 are taken from compact intervals. However, for larger values of W_0 only numerical optimization is feasible.

In Figure 5.5 we show for varying number W_0 of wave-zero nodes the values r_1 and r_2 minimizing $E[T]$ (with r_1, r_2 sampled as $(r_1, r_2) \in \left\{\frac{k}{\eta} : k = 1, \ldots, \eta - 1\right\}^2$ and the number of samples η chosen as $\eta = 500$). In Figure 5.6 we show $E[T]$ versus W_0 both for the optimal case (individually determined for each W_0) and for the parameter setting used above, i.e. $r_1 = 1/5$ and $r_2 = 1/2$. Some remarks about these results are in order:

- For increasing W_0 the optimal value for r_1 tends to zero. This makes sense: by this choice most of the contenders enter receive mode in the first slot, but with high probability at least one contender transmits. This way, most of the contenders are eliminated already in the first step, reducing the pressure for the subsequent steps.

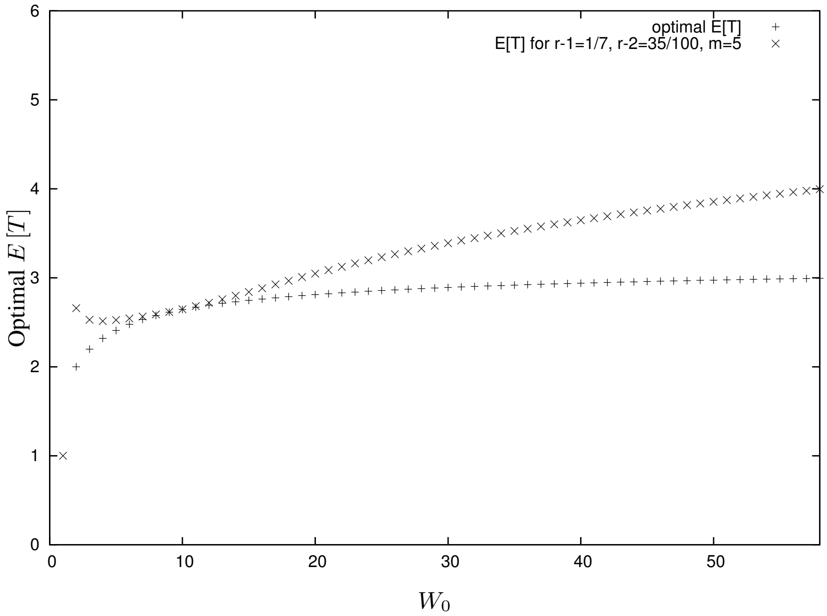

Figure 5.6: Optimal average times $E[T]$ to see the first successful slot in the contention scheme versus the number W_0 of wave-zero nodes

- It is not clear in the moment whether the optimal $E[T]$ remains bounded for $W_0 \to \infty$. We suspect that this is not the case.

Inspired from the numerical results we made the following choices:

$$\begin{aligned} m &= 5 \\ r_1 &= \frac{1}{7} \\ r_2 = r_3 = \ldots = r_m &= \frac{35}{100} \end{aligned}$$

We have investigated this scheme by simulation, using the same setup as for the truncated geometric schemes. The results for the baseline scheme, two truncated geometric schemes ($m = 2$ and $m = 4$, both for $q = 0.9$) and the contention scheme are shown in Figure 5.7. In Figure 5.8 we restrict to a comparison of the baseline scheme and the contention scheme. It can be seen that the contention scheme is a major improvement over the baseline scheme and all the truncated geometric schemes for small values of P_I, whereas for large values of P_I the loss against the baseline scheme is comparably small.

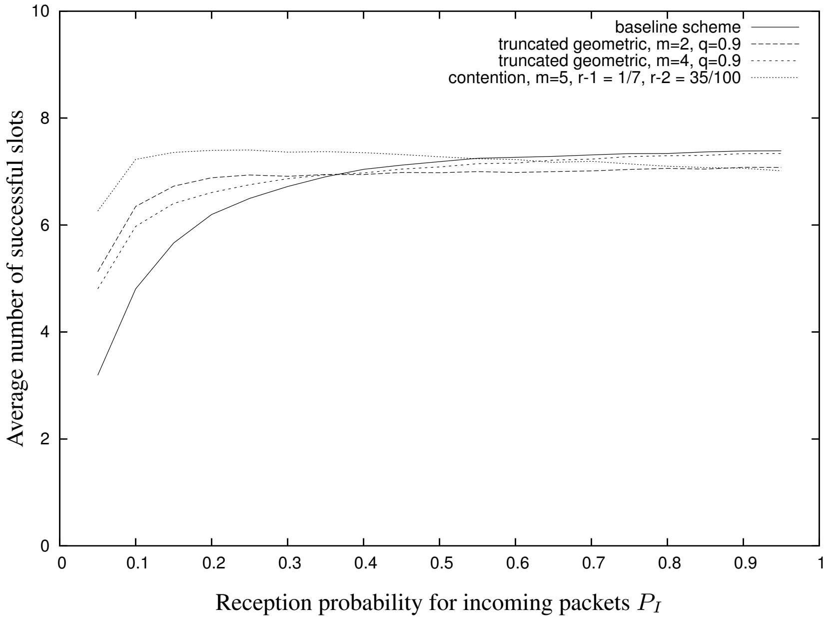

Figure 5.7: Average number of successful slots for the unmodified scheme, the truncated geometric scheme for $q = 0.9$ and different values of $m \in \{2, 4\}$ and for the contention scheme with $r_1 = 1/7$, $r_2 = 35/100$ and $m = 5$ versus reception probability P_I to receive an incoming packet ($N = 100$, $s = 0.4$, $M = 20$).

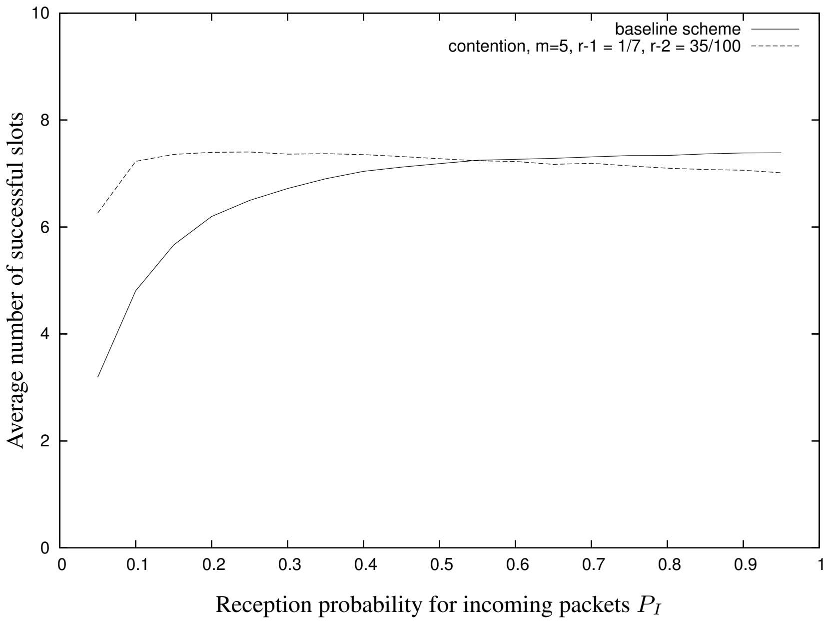

Figure 5.8: Average number of successful slots for the unmodified scheme and the contention scheme with $r_1 = 1/7$ $r_2 = 35/100$ and $m = 5$ versus reception probability P_I to receive an incoming packet ($N = 100$, $s = 0.4$, $M = 20$).

Chapter 6

Related Work

The motivation behind this research comes from the idea of using sensor networks not only to *observe* the environment, but to also *control* it through actuators [1, 17], for example in building automation applications [13]. In this kind of applications we have different types of nodes: we have the sensor nodes and the actuator nodes. For the sensor nodes many of the considerations usually made for sensor networks [12] apply, including the observation that the individual (sensor) node is not important as long as there are sufficient other sensor nodes which can observe the right data [20]. However, this consideration does not apply to scenarios where actuators are present, since these must be individually addressable. Furthermore, the quality of an open-loop or closed-loop control algorithm depends crucially on the network ability to reliably deliver sensor data to the actuator nodes. In our previous terminology, the actuator is a target node, to which the repeater group should deliver the sensor packets successfully. Hence, a repeater group can be placed close to the actuator and by exploiting spatial diversity coming from the transmissions of different nodes in the repeater group, it is possible to adjust the delivery rate of packets at the actuator by proper choice of the number of slots and the size and sleeping discipline of the repeater group.

The repeater group concept presented in this paper can be viewed as a practical incarnation of a decode-and-forward cooperative diversity scheme (compare [14, 15]), which in turn are based on the concept of relaying (see [8] for an information-theoretic treatment, and [24] for practical relaying schemes). In cooperative diversity or cooperative MIMO (multiple input / multiple output) schemes [9] many spatially separated nodes collaborate in transmitting a common signal or in receiving a signal by combining their observations. In general, such multi-antenna techniques can be used to increase capacity or to reduce the error probability for bits / packets [10]. In the realm of sensor networks capacity is typically not much of an issue, but error rates are of importance, especially when actuators are involved. In so-called *amplify-and-forward* cooperative diversity schemes, a relaying node samples incoming waveforms and retransmits them without trying to decode the packet. In *decode-and-forward* schemes a relayer must decode a packet successfully, before it is forwarded. For cooperative diversity / cooperative MIMO schemes information-theoretic bounds for capacity and outage probabilities have been considered[14, 15], but there is yet not so much work on practical schemes and their achievable performance. In [9] the energy consumption of cooperative MIMO systems are compared against single-transmitter/single-receiver systems, balancing the possible reduction of transmit energy needed to satisfy a given target error rate / throughput versus the extra energy needed to run multiple transmit and receive circuits. When it comes to multi-node cooperation, they consider cooperation at the transmitter side (where the packet is communicated to all M_t transmitter nodes by a TDMA scheme, followed by a parallel transmission of all nodes using a modified Alamouti diversity code [2]), and cooperation at the receiver side (where all M_r receivers sample the incoming signal and forward it to the final destination which combines the receivers observations). Between transmitter and receiver groups a long-haul wireless link with Rayleigh fading is used, within the groups the links are of higher quality and AWGN noise are used. It turns out that MISO (many transmitters, one receiver) and SIMO (single transmitter, many receivers) systems are more energy-efficient than SISO systems as soon as the length of the long-haul link exceeds a certain threshold (≈ 15 m for the parameters used in the paper), whereas for the true MIMO case the threshold is slightly larger. Please note that this already takes the additional energy consumption of the local cooperation in the transmit / receive groups and the usage of several instances of transmit / receive circuitry into account.

Finally, we remark that the contention scheme developed in this paper can be modified for usage in settings where a number of N sensors are triggered by the same physical event and make correlated observations. For such a setting the Sift MAC protocol [11] has been designed with the goal of making sure that one of the N sensors can send its observation quickly so that the remaining sensors do not need to send their packets, thus saving energy and reducing interference to others. In Sift a CSMA-based transmission strategy with random backoff times has been adopted, in which the distribution of the backoff time is chosen such that most of the probability mass is concentrated at the end of the admissible time interval. The rationale is that only few nodes will decide for early transmission times and hence there is small risk of collisions at the beginning of the admissible time interval. Our contention scheme can be viewed as complementary to Sift, but designed for the same purpose.

Chapter 7

Conclusions

In this paper we have started the investigation of practical schemes for the construction and operation of repeater groups, which follow the goal of realizing the reliability gains achievable with spatial diversity over wireless channels while at the same time considering the need to let individual nodes sleep and save energy, which is important in sensor networks. Specifically, we have shown that already for schemes without explicit coordination it is possible to ensure that on average a certain number of packet copies indeed reach the destination node successfully. We have demonstrated that this cannot only be done for cases without channel errors, but that it is also possible to construct coordination-free behaviours for repeater nodes which give close-to-optimal performance of the group even when the error probability for incoming packets is high. We are convinced that these results are a good starting point for the search of more efficient schemes.

There is a significant potential for future research. Already for the class of coordination-free schemes a number of issues arises: Which improvements are possible with CSMA-based schemes? Which improvements are possible when the environment is only slowly varying and repeater nodes can obtain estimates of error rates? How can feedback from the destination node be accommodated, for example to stop the repeating activities as quickly as possible? Which gains can be achieved when in addition coding and packet combining at the destination are considered? And what is the performance of these schemes in case of multi-hop repeater groups?

Appendix A

Optimal slot probabilities

For ease of reference we restate the problem. The target function $f(\cdot)$ to optimize depends on the probability distribution π as:

$$f(\pi) = f(p_1, \ldots, p_{M+1}) = \sum_{j=1}^{M} K \cdot p_j \cdot (1 - p_j)^{K-1}$$

This leads to the following nonlinear constrained optimization problem:

$$\begin{aligned}
\text{maximize} \quad & f(p_1, \ldots, p_{M+1}) \\
\text{s.t.} \quad & h(p_1, \ldots, p_{M+1}) = 1 - \sum_{i=1}^{M+1} p_i = 0 \\
& \mathbf{g}(p_1, \ldots, p_{M+1}) = \begin{pmatrix} g_1(p_1, \ldots, p_{M+1}) \\ g_2(p_1, \ldots, p_{M+1}) \\ \ldots \\ g_{M+1}(p_1, \ldots, p_{M+1}) \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ \ldots \\ p_{M+1} \end{pmatrix} \geq \mathbf{0}
\end{aligned}$$

This kind of problems can be solved with the help of the Karush-Kuhn-Tucker (KKT) theorem (see [6, Chap. 20]). To use this theorem, we first must determine which of the constraints $g_i(\cdot)$ are *inactive* (i.e. $g_i(\cdot) > 0$). At least one of the probabilities $p_1, \ldots, p_M$ is nonzero, since otherwise none of the M slots would be used for transmissions. This already implies that *all* probabilities $p_1, p_2, \ldots, p_M$ should be nonzero. To see this, assume without loss of generality that for some $0 < j < M$ we have $p_1 = p_2 = \ldots = p_j = 0$ and that $0 < p_{j+1} \leq p_{j+1} \leq \ldots \leq p_M$ holds. If we now introduce a new probability distribution π' such that

$$p'_1 = p'_2 = \ldots = p'_j = p'_M = \frac{p_M}{j+1}$$

and $p'_i = p_i$ for $i \in j+1, \ldots, M-1$ then indeed

$$f(\pi') > f(\pi)$$

since for $j > 0$ we have

$$1 - p_M < 1 - \frac{p_M}{j+1}$$

which implies

$$(1 - p_M)^{K-1} < \left(1 - \frac{p_M}{j+1}\right)^{K-1}$$

and furthermore

$$p_M(1 - p_M)^{K-1} < (j+1)\frac{p_M}{j+1}\left(1 - \frac{p_M}{j+1}\right)^{K-1}$$

The difference $f(\pi') - f(\pi)$ is just given by:

$$\begin{aligned} & f(\pi') - f(\pi) \\ = & \left(\sum_{i=1}^{j} K\frac{p_M}{j+1}\left(1 - \frac{p_M}{j+1}\right)^{K-1} + \sum_{i=j+1}^{M-1} Kp_i\left(1 - p_i\right)^{K-1} + K\frac{p_M}{j+1}\left(1 - \frac{p_M}{j+1}\right)^{K-1}\right) \\ & - \sum_{i=j+1}^{M} Kp_i\left(1 - p_i\right)^{K-1} \\ & K(j+1)\frac{p_M}{j+1}\left(1 - \frac{p_M}{j+1}\right)^{K-1} - Kp_M\left(1 - p_M\right)^{K-1} \\ > & \ 0 \end{aligned}$$

This implies that all the constraints $g_1, \ldots, g_M$ are inactive. Observing that the total differentials of $f(\cdot)$, $g(\cdot)$ and $h(\cdot)$ are as follows:

$$\begin{aligned} Df(\pi) &= \left(\frac{d}{dp_1}f(\pi), \ldots, \frac{d}{dp_M}f(\pi), \frac{d}{dp_{M+1}}f(\pi)\right) \\ &= \left(K(1-p_1)^{K-2}(1-Kp_1), \ldots, K(1-p_M)^{K-2}(1-Kp_M), 0\right) \\ Dh(\pi) &= \left(\frac{d}{dp_1}h(\pi), \ldots, \frac{d}{dp_M}h(\pi), \frac{d}{dp_{M+1}}h(\pi)\right) \\ &= (-1, \ldots, -1, -1) \\ D\mathbf{g}(\pi) &= \mathbf{I} \end{aligned}$$

The KKT theorem now states that for an optimal vector π^* there exists a Lagrange multiplier λ and a vector $\mu^* = (\mu_1^*, \ldots, \mu_M^*, \mu_{M+1}^*)$ such that:

$$\mathbf{0} \leq \mu^* \tag{A.1}$$

$$\mathbf{0}^T = Df(\pi^*) + \lambda Dh(\pi^*) + \mu^{*T} D\mathbf{g}(\pi^*) \tag{A.2}$$

$$\mathbf{0}^T = \mu^{*T}\mathbf{g}(\pi^*) \tag{A.3}$$

The above shown fact that $p_1 > 0, \ldots, p_M > 0$ together with Equations A.1 and A.3 implies that $\mu_1^* = \ldots = \mu_M^* = 0$. Taking this into consideration, when writing down Equation A.2 component-wise, we obtain:

$$\begin{aligned} 0 &= K(1-p_1)^{K-2}(1-Kp_1) - \lambda \\ & \ldots \\ 0 &= K(1-p_M)^{K-2}(1-Kp_M) - \lambda \\ 0 &= -\lambda + \mu_{M+1}^* \end{aligned}$$

The first M of these equations imply that $p_1 = p_2 = \ldots = p_M$ holds. Hence, an individual node picks each of the M slots with the same probability, say $p := p_1$.

This means that we have reduced our problem to an easier one:

$$\begin{aligned} \text{maximize} \quad & f(p,q) = M \cdot K \cdot p \cdot (1-p)^{K-1} \\ \text{s.t.} \quad & h(p,q) = 1 - M \cdot p - q = 0 \\ & p \geq 0, q \geq 0 \end{aligned}$$

where M and K are fixed and q is the probability that a repeater node does not transmit, whereas p is the probability that one fixed slot i is chosen when the node has decided to transmit. Observe that $f(\cdot)$ is continuous in p. The parameter p is restricted to the interval $\left[0, \frac{1}{M}\right]$. Obviously, $f(0) = 0$ and

$$f\left(\frac{1}{M}\right) = K \cdot \left(\frac{M-1}{M}\right)^{K-1}$$

Hence, we consider the open interval $p \in \left(0, \frac{1}{M}\right)$. The partial derivative of $f(\cdot)$ w.r.t. p is given by (assuming $K \geq 2$, $M \geq 1$):

$$\frac{\partial f(p,q)}{\partial p} = M \cdot K \cdot (1-p)^{K-2} \cdot (1 - K \cdot p)$$

For $p \in \left(0, \frac{1}{M}\right)$ this expression becomes zero when $1 - K \cdot p$ becomes zero, i.e. $p = \frac{1}{K}$. Hence we have:

$$p_{\text{opt}} \in \left\{\frac{1}{K}, \frac{1}{M}\right\} \tag{A.4}$$

For $K \leq M$ we must necessarily have $p_{\text{opt}} = \frac{1}{M}$ since with $p = \frac{1}{K}$ it is not possible to satisfy the constraints.

So, suppose that $K > M$. Observe that:

$$\frac{\partial f\left(\frac{1}{M}, q\right)}{\partial p} = M \cdot K \cdot \left(1 - \frac{1}{M}\right)^{K-2} \cdot \left(1 - \frac{K}{M}\right)$$

becomes negative for $K > M$. Since $f(\cdot)$ is continuous, there must exist at least one point p^* smaller than $\frac{1}{M}$ with $f(p^*) > f\left(\frac{1}{M}\right)$. Because of Equation A.4 and since $\frac{1}{K} < \frac{1}{M}$ this already implies that $p_{\text{opt}} = \frac{1}{K}$.

Appendix B

Moment representation of $E\left[K \cdot a^K\right]$

From the discussion in Section 3 we are interested in finding another representation for the expression

$$f(p) = E\left[K \cdot (1-p)^K\right]$$

which we generalize as

$$f(a) = E\left[K \cdot a^K\right]$$

for some $a \in (0,1)$ and with K being a non-negative discrete random variable for which all moments exist. Then:

$$\begin{aligned} E\left[Ka^K\right] &= E[K] + E\left[K^2\right]\log(a) + \ldots + \frac{E[K^n]}{(n-1)!}(\log(a))^{n-1} + \ldots \\ &= \sum_{n=1}^{\infty} \frac{E[K^n]}{(n-1)!} \cdot (\log a)^{n-1} \end{aligned}$$

This can be seen as follows. We have:

$$\begin{aligned} E\left[Ka^K\right] &= \sum_{k=1}^{\infty} ka^k \Pr[K=k] = a\sum_{k=1}^{\infty} ka^{k-1}\Pr[K=k] \\ &= a\sum_{k=1}^{\infty} \frac{d}{da} a^k \Pr[K=k] \\ &= a\frac{d}{da}\sum_{k=1}^{\infty} a^k \Pr[K=k] = a\frac{d}{da}\sum_{k=0}^{\infty} a^k \Pr[K=k] \\ &= a\frac{d}{da}\sum_{k=0}^{\infty} e^{k\log(a)} \Pr[K=k] \end{aligned}$$

$$= a\frac{d}{da}\Phi_K(\log(a))$$

where $\Phi_K(x) = E\left[e^{xK}\right]$ is the moment-generating function of the random variable

K. [1] One of the well-known properties of moment-generating functions is that [3, Sec. 2.9]:

$$\Phi_K(x) = 1 + xE[K] + \ldots + \frac{x^n}{n!}E[X^n] + \ldots$$

provided all moments exist. From this we have:

$$\begin{aligned}
&a\frac{d}{da}\Phi_K(\log(a)) \\
&= a\frac{d}{da}\left(1 + E[K]\log(a) + E\left[K^2\right]\frac{\log^2(a)}{2!} + \ldots + E[K^n]\frac{\log^n(a)}{n!} + \ldots\right) \\
&= a\left(\frac{E[K]}{a} + \frac{E\left[K^2\right]}{2!}\frac{2\log(a)}{a} + \ldots + \frac{E[K^n]}{n!}\frac{n\log^{n-1}(a)}{a} +\right) \\
&= E[K] + E\left[K^2\right]\log(a) + \ldots + \frac{E[K^n]}{(n-1)!}\log^{n-1}(a) + ...
\end{aligned}$$

[1]To compute the variance of Ka^K the second moment $E\left[K^2 a^{2K}\right]$ is needed, for which can be expressed as:

$$E\left[K^2 a^{2K}\right] = a^2\frac{d^2}{da^2}\Phi_K(2\log(a)) + a\frac{d}{da}\Phi_K(2\log(a))$$

Appendix C

Average time to the first successful slot in the contention scheme

In this appendix we use a Markovian model for the behaviour of the contention scheme described in Section 5.2 to derive the average slot number carrying the first successful slot. In our model we make the following assumptions:

- The parameter m (number of slots after which the contention scheme ends) is disregarded, we assume $m = \infty$. In addition, we assume that all the nodes which have lost contention defer any further transmissions until the contention scheme has been terminated, so as to avoid any interference from outside.
- The number of wave-zero nodes $W_0 > 1$ is known and fixed. The case $W_0 = 1$ is easy to handle, since with the exception of the first slot the average time until the first successful slot is a geometric random variable.
- In the first slot a wave-zero node transmits with probability r_0 (and listens with probability $1-r_0$) and in all the subsequent slots it transmits (listens) with probability r_1 $(1-r_1)$. Please note that we have changed the notation here as compared to Section 5.2 to be more consistent with the following derivation.

We have a slotted, discrete-time system and we model the evolution of the number of contenders $(X_n)_{n\geq 0}$ as a time-homogeneous discrete-time Markov chain. The random variable X_n denotes the number of contenders at the end of the $n-1$-th slot, and X_0 is the initial probability distribution (discussed below). The state space of the Markov chain is given by:

$$\mathcal{I} = \{1, 2, \ldots, W_0\}$$

As discussed in Section 5.2, each of the W_0 wave-zero nodes decides to transmit in the first slot with probability r_0 and to receive with probability $1 - r_0$. If at least one wave-zero node transmits, all the nodes which have chosen are eliminated from contention. For the number X_0 of remaining nodes we then have:

- The event that $X_0 = W_0$ occurs when either *all* wave-zero nodes decide to transmit in the first slot or *all* wave-zero nodes receive in the first slot.

- The event that $X_0 = i$ for some $1 \leq i < W_0$ occurs when exactly i wave-zero nodes decide to transmit and the remaining $W_0 - i$ nodes decide to receive (and are eliminated subsequently).

Taking into account that all wave-zero nodes make their decision independently with the same transmit probability r_0 we have:

$$\Pr[X_0 = i] = \begin{cases} b(0; W_0, r_0) + b(W_0; W_0, r_0) & : \quad i = W_0 \\ b(i; W_0, r_0) & : \quad 1 \leq i < W_0 \end{cases}$$

where $b(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$ is the probability mass function of the binomial distribution with parameters n and p.

For the further evolution of the number of contenders X_n we have to specify the state transition probabilities $p_{i,j}$ for having i contenders at time n and j contenders at time $n + 1$. From the description of the contention scheme we have:

$$p_{i,j} = \begin{cases} 0 & : \quad j > i \\ b(0; i, r_1) + b(i; i, r_1) & : \quad i = j \\ b(j; i, r_1) & : \quad 1 \leq j < i \end{cases}$$

which can be justified as follows. Suppose we are currently in state $i \geq 1$. The number of contenders cannot increase over time, which explains $p_{i,j} = 0$ for $j > i$. To stay in state i, either all contenders have to transmit (with probability r_1) or all contenders have to receive. Because of the independence of the contenders, this event happens with probability $b(0; i, r_1) + b(i; i, r_1)$. Finally, to have $1 \leq j < i$ contenders in step $n + 1$, exactly j of the contenders decide to transmit, which happens with probability $b(j; i, r_1)$. Summarizing, the state transition matrix $\mathbf{P}$ of the Markov chain is given by:

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ b(1; 2, r_1) & b(0; 2, r_1) + b(2; 2, r_1) & 0 & \dots & 0 \\ b(1; 3, r_1) & b(2; 3, r_1) & b(0; 3, r_1) + b(3; 3, r_1) & \dots & 0 \\ \dots & & & & \\ b(1; W_0, r_1) & b(2; W_0, r_1) & b(3; W_0, r_1) & \dots & b(0; W_0, r_1) + b(W_0; W_0, r_1) \end{pmatrix}$$

where the i-th row gives the state transition probabilities for state i, and $\mathbf{P}$ has W_0 rows and columns. It is obvious that state 1 is absorbing, and the other states $2, \dots, W_0$ are transient states. Hence, the Markov chain reaches state 1 with probability one. Let $\mathbf{Q}$ denote the lower-right $W_0 - 1 \times W_0 - 1$ submatrix of $\mathbf{P}$ given by:

$$\mathbf{Q} = \begin{pmatrix} b(0; 2, r_1) + b(2; 2, r_1) & 0 & \dots & 0 \\ b(2; 3, r_1) & b(0; 3, r_1) + b(3; 3, r_1) & \dots & 0 \\ \dots & & & \\ b(2; W_0, r_1) & b(3; W_0, r_1) & \dots & b(0; W_0, r_1) + b(W_0; W_0, r_1) \end{pmatrix}$$

then $\mathbf{Q}$ has the following properties: it is a truly sub-stochastic (non-negative elements with row sum smaller than one) and lower triangular matrix, with non-zero elements $[[Q]]_{i,j}$ for $j \leq i$. Since all diagonal elements are nonzero and distinct, the matrix has $W_0 - 1$ distinct Eigenvalues and is thus diagonalizable. Furthermore, with respect to the row-sum matrix norm we have $\|\mathbf{Q}\| < 1$.

Now, let T be the random variable denoting the first successful slot. For this random variable the following holds:

- When $X_0 = 1$ holds, then the first successful slot is already the first one, and hence we have $\Pr[T = 1 | X_0 = 1] = 1$.

- When $X_0 > 1$ holds, then the first successful slot appears as soon as the chain reaches state 1, i.e.

  $$T = \inf\{n \in \mathbb{N} : X_n = 1\}$$

  that is, T is the first hitting time of the absorbing state 1.

It is well-known [18, Sec. 1.3] that if k_i denotes the average time to reach state 1 if the chain starts in $X_0 = i$, then the vector $\mathbf{k} = (k_i : i \in \{1, \ldots, W_0\})$ is the minimal non-negative solution to the following set of linear equations:

$$k_i = \begin{cases} 0 & : \quad i = 1 \\ 1 + \sum_{j=2}^{i} p_{i,j} \cdot k_j & : \quad 1 < i \leq W_0 \end{cases}$$

In matrix notation, the vector $\mathbf{k} = (k_i : i \in \{2, \ldots, W_0\})$ hence satisfies

$$\mathbf{k} = \mathbf{e} + \mathbf{Q} \cdot \mathbf{k}$$

(where $\mathbf{e}$ is an $W_0 - 1$-dimensional column vector of ones), or differently:

$$(\mathbf{I} - \mathbf{Q}) \cdot \mathbf{k} = \mathbf{e} \tag{C.1}$$

where $\mathbf{I}$ is the $W_0 - 1$-dimensional identity matrix. Since $\|Q\| < 1$, the theorem about the von Neumann series [7, Chap. 1] guarantees that $\mathbf{I} - \mathbf{Q}$ is invertible and Equation C.1 has a unique solution.

However, explicitly computing this solution quickly becomes infeasible. Instead, by utilizing the triangular structure of $\mathbf{P}$, it is easy to derive the following recursive solution:

$$\begin{aligned} k_1 &= 0 \\ k_i &= \frac{1 + \sum_{j=2}^{i-1} b(j; i, r_1) \cdot k_j}{1 - (1 - r_1)^i - r_1^i} \end{aligned}$$

Taking into account the first slot needed to determine X_0 out of W_0 we have:

$$\begin{aligned} E\left[T | X_0 = 1\right] &= 1 \\ E\left[T | X_0 = 2\right] &= 1 + k_2 \\ &\cdots \\ E\left[T | X_0 = W_0\right] &= 1 + k_{W_0} \end{aligned}$$

For the average time needed to see the first successful slot we finally have from using the properties of conditional expectation:

$$\begin{aligned} E\left[T\right] &= E\left[E\left[T | X_0\right]\right] \\ &= \sum_{x=1}^{W_0} E\left[T | X_0 = x\right] \cdot \Pr\left[X_0 = x\right] \\ &= b(1; W_0, r_0) + \sum_{x=2}^{W_0-1} b(x; W_0, r_0) \cdot (1 + k_x) + (b(0; W_0, r_0) + b(W_0; W_0, r_0)) \cdot (1 + k_{W_0}) \\ &= 1 + \sum_{x=2}^{W_0-1} k_x \cdot b(x; W_0, r_0) + k_{W_0} \cdot \left(r_0^{W_0} + (1 - r_0)^{W_0}\right) \end{aligned}$$

Please note that $E\left[T\right]$ depends on three parameters: W_0, r_0 and r_1.

This model has been validated by comparing numerical results for $E\left[T\right]$ with the result of simulations. The results show an excellent correspondence between theoretical and simulated results.

Bibliography

[1] Ian F. Akyildiz and Ismail H. Kasimoglu. Wireless sensor and actor networks: research challenges. *Ad Hoc Networks*, 2:351–367, 2004.

[2] Siavash M. Alamouti. A Simple Transmit Diversity Technique for Wireless Communications. *IEEE Journal on Selected Areas in Communications*, 16(8):1451–1458, October 1998.

[3] Arnold O. Allen. *Probability, Statistics, and Queueing Theory – With Computer Science Applications*. Computer Science and Applied Mathematics. Academic Press, New York, 1978.

[4] D. Bertsekas and R. Gallager. *Data Networks*. Prentice Hall, Englewood Cliffs, New Jersey, 1987.

[5] J. I. Capetanakis. Tree Algorithm for Packet Broadcast Channels. *IEEE Transactions on Information Theory*, 25(5):505–515, September 1979.

[6] Edwin K. P. Chong and Stanislaw H. Zak. *An Introduction to Optimization*. John Wiley & Sons, New York, second edition, 2001.

[7] Richard Courant and David Hilbert. *Methods of Mathematical Physics – Volume I*. Wiley, New York, 1953. Reprint 1989.

[8] Thomas M. Cover and Abbas A. El Gamal. Capacity theorems for the relay channel. *IEEE Transactions on Information Theory*, 25(5):572–584, September 1979.

[9] Shuguang Cui, Andrea J. Goldsmith, and Ahmad Bahai. Energy-efficiency of mimo and cooperative mimo techniques in sensor networks. *IEEE Journal on Selected Areas in Communications*, 22(6):1089–1098, August 2004.

[10] Suhas N. Diggavi, Naofal Al-Dhahir, A. Stamoulis, and A. R. Calderbank. Great Expectations: The Value of Spatial Diversity in Wireless Networks. *Proceedings of the IEEE*, 92(2):219–270, February 2004.

[11] Kyle Jamieson, Hari Balakrishnan, and Y. C. Tay. Sift: A mac protocol for event-driven wireless sensor networks. In *Proc. Wireless Sensor Networks, First European Workshop (EWSN 2006)*, Zürich, Switzerland, January 2006.

[12] Holger Karl and Andreas Willig. *Protocols and Architectures for Wireless Sensor Networks*. John Wiley & Sons, Chichester, 2005.

[13] Wolfgang Kastner, Georg Neugschwandtner, Stefan Soucek, and H. Michael Newman. Communication Systems for Building Automation and Control. *Proceedings of the IEEE*, 93(6):1178–1203, 2005.

[14] J. Nicholas Laneman, David N. C. Tse, and Gregory W. Wornell. Cooperative diversity in wireless networks: Efficient protocols and outage behaviour. *IEEE Transactions on Information Theory*, 50(12):3062–3080, December 2004.

[15] J. Nicholas Laneman and Gregory W. Wornell. Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks. *IEEE Transactions on Information Theory*, 49(10):2415–2425, October 2003.

[16] Hang Liu, Hairuo Ma, Magda El Zarki, and Sanjay Gupta. Error control schemes for networks: An overview. *MONET – Mobile Networks and Applications*, 2(2):167–182, 1997.

[17] Tommaso Melodia, Dario Pombili, Vehbi C. Gungor, and Ian F. Akyildiz. A distributed coordination framework for wireless sensor and actor networks. In *Proc. 6th ACM international symposium on Mobile ad hoc networking and computing (MobiHoc)*, Zürich, Switzerland, January 2005.

[18] J. R. Norris. *Markov Chains*. Cambridge University Press, Cambridge, UK, 1997.

[19] Theodore S. Rappaport. *Wireless Communications – Principles and Practice*. Prentice Hall, Upper Saddle River, NJ, USA, 2002.

[20] Sylvia Ratnasamy, Brad Karp, Li Yin, Fang Yu, Deborah Estrin, Ramesh Govindan, and Scott Shenker. Ght: A geographic hash table for data-centric storage. In *Proc. First ACM Intl. Workshop on Wireless Sensor Networks and Applications (WSNA'02)*, Atlanta, GA, 2002.

[21] Vipul Singhvi, Andreas Krause, Carlos Guestrin, James H. Garret, and H. Scott Matthews. Intelligent light control using sensor networks. In *Proc. 3st Intl. Conf. on Embedded Networked Sensor Systems (SenSys)*, San Diego, CA, November 2005. ACM.

[22] Mehmet C. Vuran and Ian F. Akyildiz. Spatial Correlation-Based Collaborative Medium Access Control in Wireless Sensor Networks. *IEEE/ACM Transactions on Networking*, 14(2):316–329, April 2006.

[23] Stephen B. Wicker. Adaptive rate error control through the use of diversity combining and majority-logic decoding in a hybrid-arq protocol. *IEEE Transactions on Communications*, 39(3):380–385, March 1991.

[24] Andreas Willig. Polling-based MAC Protocols for Improving Realtime Performance in a Wireless PROFIBUS. *IEEE Transactions on Industrial Electronics*, 50(4):806 –817, August 2003.

[25] Andreas Willig, Martin Kubisch, Christian Hoene, and Adam Wolisz. Measurements of a wireless link in an industrial environment using an IEEE 802.11-compliant physical layer. *IEEE Transactions on Industrial Electronics*, 49(6):1265–1282, 2002.