

Probability Distributions for Channel Utilisation

Christian Schindelhauer^{1,*} and Kerstin Voß^{2,**}

¹ Heinz Nixdorf Institute, Paderborn University
schindel@uni-paderborn.de

² Paderborn Center for Parallel Computing - PC², Paderborn University
kerstinv@uni-paderborn.de

Abstract. Sensor nets have many undisputed fields of application. A paradigm of communication is the use of **one** control channel in the MAC layer. We challenge this paradigm for nodes with very restricted hardware resources. In our model nodes support the use of different channels and use clock synchronisation.

We present a simple probabilistic synchronised channel utilisation scheme for wireless communication. The main features are its simplicity, robustness against radio interference, the high throughput caused by less interfering signals, and predictable energy consumption. For this, the channel selection is based on a carefully chosen probability distribution maximising the expected number of successfully delivered packets up to a constant factor.

Combined with a standard synchronisation scheme it provides a novel energy-efficient, robust, and fast message delivery service for sensor networks where data gathering is not available due to memory restrictions.

Topics: Access control, Sensor Networks.

Keywords: Sensor network, frequency selection, medium access control.

1 Introduction

The growing research interest in sensor nets is motivated by the broad applicability of such systems. The primary task of the sensor nodes is to send observed sensory attributes of their environment to a central station [1]. Sensor nodes consist of a processor (CPU), storage (RAM and FLASH), radio transceiver, sensor, and an autarkic energy source. They need to be inexpensive and small and due to this their hardware resources are sparse [2,3], i.e. low battery energy, tiny RAM, and small computational power. So, algorithms for sensor nets must be time and memory efficient. Due to the fact that nodes run on batteries, the range for sending and receiving is very limited if they communicate wirelessly. Consequently, in order to send data to the central station, nodes have to cooperate with each other by providing multi-hop-connections. In wireless communication radio interference occasionally occurs and connections can fail. Furthermore,

* Partially supported by the DFG-Sonderforschungsbereich 376 and by the EU within 6th Framework Programme under contract 001907 "Dynamically Evolving, Large Scale Information Systems" (DELIS).

** Contact author: Kerstin Voß, PC², Fuerstenalle 11, 33102 Paderborn, Germany, Phone. +49 5251 606321, Fax: +495251 60-6297.

nodes can move or drop out due to loss of energy. According to these characteristics, routing protocols have to manage communication with dynamic partners.

Some state of the art radio transceivers offer only a small number of frequencies. For example, the *Chipcon CC1000* supports at most 13 frequencies of 64 kHz, if the safety section is as large as the interval for a channel [4]. However, if the message transmission does not need much time, a higher number of frequencies can be simulated using internal rounds. Typically, most energy is consumed for sending and receiving [5]. According to this, during inactive rounds only little energy will be consumed. The energy consumption in such waiting periods should be neglected.

In case of high communication traffic, a greater number of frequencies will increase the throughput. However, nodes have to select carefully the channels in order to avoid as much interference as possible. Finally, if two nodes try to send on the same channel, they normally interfere with another. This problem leads to the following question: Which probability distribution in the frequency selection would attain best results? For many cases the uniform distribution is the best choice. However, for very few or a large number of senders and receivers there are better probability distributions. According to the Data Link Layer in the OSI Model (see IEEE P802.3ae) this is a typical problem of the Medium Access Layer because the question relates to the use of the communication medium.

This paper discusses which probability distribution should be used in the MAC layer to maximise the number of successful delivered messages within one parallel sending attempt. The structure of the paper is as following: At first, we take a look at related work. Afterwards the explicit problem and model are described. Section 4 presents the results of our research which includes analysis and empirical examples. A perspective of the future work follows and the conclusion forms the end.

2 Related Work

Only special routing algorithms can be used in sensor nets because of the nodes' sparse resources and the network dynamic. In the past several algorithms have been developed. Few protocols work with data aggregation but these cannot be used with every node specification (if the nodes have only little RAM, e.g. 256 Byte). We refer to [6] and mention the LEACH protocol [7] as examples for protocols gathering data. Additionally, the Data Collection Protocol (DCP) is notable [8].

Rudiments of routing in mobile ad-hoc networks (reactive and proactive) are rarely transferred for sensor networks. Indeed resources are limited for mobile ad-hoc networks but these resource-restrictions are certainly stronger for sensor nets. Anyhow, the Pulse Protocol [9,10] based on a reactive approach has been used for sensor nets.

Usually, different communication directions are considered in sensor networks. We assume that the clock synchronised sensor nodes deliver packets to and from a central station. This communication pattern is the standard communication mode for sensor nets. Sometimes the communication between several nodes in the net is also of interest as for example in the direct diffusion mechanism [7,11].

In sensor nets also different methods are proposed for the underlying MAC layer. Are collisions avoided or not? To avoid interference, nodes arrange the data transfer

with several control messages (See PAMA [12] and MACA [13,14]). If the transmitted messages are very small, it can be advantageous that no control messages are used [15]. This method was already integrated in ALOHA [16]: In ALOHA nodes send their messages at any time. If collisions appear, the packages are repeated after a randomised waiting-time. A variant is the slotted ALOHA [17,18,19]: Sending is exclusively allowed in special defined slots.

Several MAC protocols with coordinated wake up have been developed. Very energy-efficient is the *time-slotted* PMAC [20] because the sleep time depends on the traffic in the net. If there nodes recognise no communication activities, they sleep longer to save energy. The sleep-wakeup pattern can be used in nearly every protocol and also with our new channel selection.

In the considered sensor network all nodes use the same medium. Following, multiplexing is needed in the MAC layer [14] which can be achieved in four dimensions: space, time, frequency, and code. Our work is based on frequency division multiplexing (FDM) because the used frequency band is divided in several channels. Additionally, time division multiplexing (TDM) can be integrated if the radio transceiver does not support as many channels as required in the protocol. TDM simulates channels with the aid of internal rounds. However, messages can collide without access control.

All these discussed sensor net protocols only work with one or a few number of channels. No protocol takes advantage of the number of supported frequencies. A clever utilisation of this multitude increases the throughput which further leads to energy savings and a longer lifetime of the nodes. This is our motivation to discuss the integration of this feature into the MAC layer.

As a basis principle we choose a probabilistic channel selection. Our mechanism needs no control channel on which the communication participants arrange a rendezvous on another channel and initialise the pseudo-random channel hopping. This approach is realised in Bluetooth for every Piconet: The clients follow the pattern of the master [21,22]. Bluetooth's method of frequency hopping was also transferred especially for smart devices with Bluetooth like special sensor nodes in [23].

3 The Problem

To increase throughput towards the central node, we use zoning which classifies nodes according to their hop distance. Then, multi-hop communication towards the central station can be reduced to series of single hops with an unknown number of possible senders n and receivers m . Senders and receivers can choose among C channels. A node is a sender if it carries a packet to be delivered. If a node is allowed and capable to accept such a packet, it is a receiver. Because of the tininess of our packets, we do not use a sophisticated medium access protocol. Instead we implement an easy three message hand-shake for a selected channel at some synchronised point of time (possibly simulated by a TDM).

- Every sender sends one packet, which includes the message and some control information, on a random channel.
- A receiver chooses a random channel for listening a time duration sufficient to receive one packet. If one sender sends on the same channel synchronously, the

receiver answers on the same channel with an acknowledgement packet containing the sender's ID.

- If a sender receives an acknowledgement, it erases the packet. Otherwise it tries to re-transmit the message in an subsequent round.

The numbers n and m of sending and receiving nodes are unknown to the protocol and to the nodes. Nodes are informed only about parts of their neighbourhood because of their memory restrictions. The communication consists of a single packet on a channel at a specific point of time. To reach a good throughput the nodes are synchronised, so that sending and listening take place at the same time. For the beginning we concentrate on one sending and receiving attempt, i.e. one point of time.

3.1 The Model

The significant measure for the throughput is the expected value of the successful transmitted messages in one sending phase.

Definition 1. *Given n senders, m receivers and C channels. Each sender chooses channel $i \in C$ independently with probability p_i and each receiver chooses this channel with probability p'_i . Define the random variable $M_{p_i, p'_i}^{n, m}$ as the number of forwarded messages in the considered round. This value corresponds to the number of channels chosen by exactly one sender and at least one receiver.*

The optimisation objective is to maximise the expected number of delivered messages during a single round for all numbers of senders from a given range $[n_0, N_0]$ and receivers from $[m_0, M_0]$:

$$\min_{n \in [n_0, N_0]} \min_{m \in [m_0, M_0]} \mathbf{E} \left[M_{p_i, p'_i}^{n, m} \right].$$

Considering the expectation this is equivalent to minimise the following term:

Lemma 1.

$$\mathbf{E} \left[M_{p_i, p'_i}^{n, m} \right] = \sum_{i=1}^C n \cdot p_i \cdot (1 - p_i)^{(n-1)} \cdot (1 - (1 - p'_i)^m)$$

Proof. follows from the independence of the channel selection.

3.2 Candidates for Probability Distributions

In this paper we consider the following probability distributions:

Definition 2.

1. Uniform distribution: $p_i^{uni} := \frac{1}{C}$
2. Geometric distribution: $p_i^{geo} := \begin{cases} \frac{1}{2^i}, & \text{if } i < C \\ \frac{1}{2^{C-1}}, & \text{if } i = C \end{cases}$
3. Factorised geometric distribution with parameter $s \in \mathbb{N}$ uses the geometric distribution stretched with factor C/s , assuming C is a multiple of s .

$$p_i^{s-geo} := \frac{1}{s} p_{\lceil \frac{i}{s} \rceil}^{geo} = \frac{1}{s \cdot 2^{\lceil \frac{i}{s} \rceil}}$$

4. Pareto distribution with parameter $\alpha \geq 1$: $p_i^\alpha := \frac{k_{\alpha,C}}{i^\alpha}$,
 with $k_{\alpha,C} = \left(\sum_{i=1}^C \frac{1}{i^\alpha}\right)^{-1}$. Note that for $\alpha > 1$: $k_{\alpha,C} = O\left(\frac{1}{\alpha-1}\right)$ and for $\alpha = 1$:
 $k_{\alpha,C} = \Theta(\ln C)$.

Both, senders and receivers, can use independently from another one of these probability distributions.

The *uniform distribution* chooses each channel with the same probability resulting in an expected number of $\frac{m}{C}$ senders and $\frac{m}{C}$ receivers on each channel. For many cases the uniform distribution seems to be a good choice. We will see that this intuition is correct if there are approximately as many receivers as senders and if these numbers fits approximately to the number of channels. If there are significantly less senders than available channels, other distributions outperform the uniform distribution.

The *geometric distribution* performs best for one sender. The expected number of messages decreases rapidly with more senders. Also it never reaches a higher value than one. So, it is very surprisingly that the closely related *factorised distribution* performs extraordinarily well in simulations. For our envisaged scenario of a range of [10, 100] senders and receivers, a scaling factor of $s \in [12, 14]$ outperforms all other probability distributions. The paper we will explain this phenomenon. We conjecture that the factorised geometric distribution is optimal if the number of channels is large enough.

The *Pareto distribution* with factor one achieves better results than distributions based on factors $\alpha = 2, 3, \dots$. It turns out that these distributions using any factor outperform all other probability distributions if the number of channels is smaller than the number of senders.

4 Results

In this section we give a mathematical analysis of the above defined distributions, present a lower bound, case differentiations for upper bounds, and graph plots for the practical relevant scenarios.

The communication characteristics leads to three analysis cases: First, a sender chooses an empty channel with at least one receiver. This is the only case in which messages are assumed to be transmitted successfully. Second, a sender selects a channel without any receiver. Third, two senders collide on a channel, then we assume no message is delivered to a potential listening receiver.

For the message delivery the presence of multiple receivers is not negative. However, this may result in duplicating a message which may cause problems in possible subsequent rounds. Delimiting such effects is part of ongoing work.

4.1 Approximating the Expected Number of Messages

We start our analysis with a *master-lemma* giving asymptotic tight bounds for the expected number of messages.

Lemma 2. For constants $c = (1 - \frac{1}{e}) \min_i \{1 - p_i\}$ and $c' = e^{\max_i \{p_i\}}$ we have

$$c \sum_{i=1}^C H\left(\frac{(n-1) \cdot p_i}{1 - p_i}\right) \cdot \min\{1, mp'_i\} \leq \mathbf{E} \left[M_{p_i, p'_i}^{n, m} \right] \leq c' \sum_{i=1}^C H(n \cdot p_i) \cdot \min\{1, mp'_i\},$$

where $H(x) := xe^{-x}$.

Proof. First note that for all $p \in [0, 1]$ and $m > 1$:

$$\left(1 - \frac{1}{e}\right) \min\{1, pm\} \leq (1 - (1 - p)^m) \leq \min\{1, pm\}.$$

Further, observe that $(1 - p)^{n-1} \leq e^{-pn+p} \leq e^{-pn}e^p$ which implies the upper bound.

To proof the lower, note that $\forall n > 1: (1 - \frac{1}{n})^{n-1} \geq \frac{1}{e}$ which implies $(1 - p) \geq e^{-\frac{p}{1-p}}$. Therefore: $np(1 - p)^{n-1} \geq npe^{-\frac{np}{1-p}} \geq (n - 1)pe^{-\frac{np}{1-p}}$. From this the lower bound follows straight-forward. \square

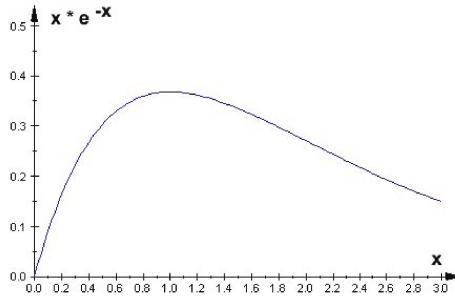


Fig. 1. The function $x \mapsto x \cdot e^{-x}$

We investigate probability distributions p with $0 \leq p_i \leq \frac{1}{2}$. So, $c \geq \frac{1}{2} (1 - \frac{1}{e}) \approx 0.316$ and $c' \leq e^{\frac{1}{2}} \approx 1.649$. For the analysis it is crucial to understand the meaning of the function $H(x) = xe^x$ visualised in Fig.1. Additionally for the estimation of the performance of the probability distributions, we will use the following Lemma.

Lemma 3. It holds

$$\forall x \in [0, \frac{1}{2}] : \frac{1}{\sqrt{e}}x \leq xe^{-x} \leq x, \quad \forall x \in [\frac{1}{2}, 2] : \frac{2}{e^2} \leq xe^{-x} \leq \frac{1}{e},$$

$$\forall x \geq 2 : 2 \cdot e^{-x} \leq xe^{-x} \leq 2 \cdot 2^{-x}, \quad \int_{x=0}^{\infty} xe^{-x} dx = 1$$

Proof. follows by applying straight-forward mathematical analysis. \square

4.2 The Performance of the Distributions

In this section we assume senders and receivers to use the same probability distribution, i.e. $p'_i = p_i$.

Theorem 1. *Uniform distribution:*

$$\mathbf{E} \left[M_{p^{uni}, p^{uni}}^{n,m} \right] = \begin{cases} \Theta\left(\frac{nm}{C}\right), & \text{if } n \leq C \text{ and } m \leq C \\ O(m \cdot 2^{-\frac{n}{C}}), & \text{if } n \geq C \text{ and } m \leq C \\ \Theta(n), & \text{if } n \leq C \text{ and } m \geq C \\ O(C \cdot 2^{-\frac{n}{C}}), & \text{if } n \geq C \text{ and } m \geq C \end{cases}$$

Proof. follows by applying Lemma 2 and Lemma 3 to an extensive case study. \square

This result underlines the aforementioned intuition: If the number of senders approximates the number of channels and sufficient receivers are available, then the uniform distribution is asymptotically optimal. It performs worse if less or more senders want to communicate.

Theorem 2. *Geometric distribution:*

$$\mathbf{E} \left[M_{p^{geo}, p^{geo}}^{n,m} \right] = O(1).$$

Proof. Consider the terms $p_i n = \frac{1}{2}n, \frac{1}{4}n, \dots$. Using Lemma 3 there are three cases for the term $H(p_i n)$:

1. $p_i n \leq \frac{1}{2}$: Then $H(p_i n) \leq p_i n$ and this implies $\sum_{i: p_i n \leq \frac{1}{2}} H(p_i n) \leq 1$.
2. $\frac{1}{2} < p_i n \leq 2$: Then $H(p_i n) \leq \frac{1}{e}$. For those i 's the following estimation has to be valid: $i - 1 < \log n \leq i + 1$. Here, we have $\sum_{i: \frac{1}{2} < p_i n \leq 2} H(p_i n) \leq \frac{2}{e}$.
3. $p_i n \geq 2$: Then $1 + n2^{-i} = 1 + p_i n \leq p_{i-1} n = n2^{-i+1}$. Therefore

$$\sum_{i: p_i n \geq 2} H(p_i n) \leq \sum_{j=2}^{\infty} H(j) \leq 1.$$

Combining all cases gives an upper bound of $2 + \frac{2}{e}$. \square

Theorem 3. *Factorised geometric distribution:* Let $s \leq n$ and $C \geq s \log n$:

$$\mathbf{E} \left[M_{p^{s-geo}, p^{s-geo}}^{n,m} \right] = \begin{cases} \Theta\left(s \frac{m}{n}\right), & \text{if } m \leq n \\ \Theta(s), & \text{if } m \geq n \end{cases}$$

Proof. From $s \leq \frac{C}{n}$ and $s \leq n$ follows that there exist $2s$ indices i with $\frac{1}{2} \leq np_i \leq 2$. These indices satisfy $\frac{1}{2} \leq \frac{n}{s2^{\lceil \frac{i}{s} \rceil}} \leq 2$ which is equivalent to

$$s(\log n - \log s) - s \leq i \leq s(\log n - \log s) + s$$

For these s indices the term $\min\{1, mp_i\}$ is $\Theta(1)$ if $m \geq n$, otherwise $\Theta\left(\frac{m}{n}\right)$.

From Lemma 3 follows that for each of these indices $\frac{2}{e^2} \leq H(np_i) \leq \frac{1}{e}$ is valid.

This already implies a lower bound of $(1 - \frac{1}{e})\frac{2}{e^2}s \min\{1, mp_i\}$. Within this range of indices the sum is upperbounded by $\frac{1}{\sqrt{e}}s \min\{1, mp_i\}$.

Now consider $i \leq s(\log n - \log s) - s$ then $H(np_i) \leq np_i$ combined with $\sum_{i \leq s(\log n - \log s) - s} H(np_i) \leq s$. For the case $i \geq s(\log n - \log s) + s$ note that $np_{i+s} \leq np_i - 1$. Since $np_i \geq 2$, $H(np_i) \leq 2 \cdot 2^{-np_i}$ is valid (from Lemma 3) which also implies an upper bound of $\sum_{i \geq s(\log n - \log s) + s} H(np_i) \leq s$.

For $m \geq n$, this implies the claim.

In the case $m \leq n$ for $i \leq s(\log n - \log s) + s$ we have $\min\{1, mp_i\} = O(\frac{m}{n})$. For $i \geq s(\log n - \log s) + s$ we have to consider the sum

$$\begin{aligned} \sum_{i \geq s(\log n - \log s) + s} H(np_i) \min\{1, mp_i\} &\leq \sum_{i \geq s(\log n - \log s) + s} H(np_i) mp_i \\ &\leq \frac{m}{n} \sum_{i \geq s(\log n - \log s) + s} 2 \cdot 2^{-np_i} np_i \leq 2s \frac{m}{n} \sum_{i=1}^{\infty} 2^i 2^{-2^i} = O\left(\frac{sm}{n}\right) \end{aligned}$$

Then the claim follows by Lemma 2. □

Theorem 4. Pareto distribution: For $\alpha \geq 1$, $C \geq (nk_{\alpha,C})^{\frac{1}{\alpha}}$, and $k_{\alpha,C} = \frac{1}{\sum_{i=1}^C i^{-\alpha}}$ we observe

$$\mathbf{E} [M_{p^{\alpha}, p^{\alpha}}^{n,m}] = \begin{cases} \Theta\left((nk_{\alpha,C})^{\frac{1}{\alpha}} \frac{m}{n}\right), & \text{if } m \leq n \\ \Theta\left((nk_{\alpha,C})^{\frac{1}{\alpha}}\right), & \text{if } m > n \end{cases}$$

Proof. Let $k := k_{\alpha,C}$. First we take a look at the case $m > n$. For the senders we consider three cases:

1. $np_i = \frac{nk}{i^{\alpha}} \leq \frac{1}{2} \Leftrightarrow i \geq (2nk)^{\frac{1}{\alpha}}$. From Lemma 3 it follows $H(np_i) \leq np_i$. This implies

$$\sum_{i \geq (2nk)^{\frac{1}{\alpha}}} \frac{nk}{i^{\alpha}} \leq \frac{1}{2} \sum_{i \geq 1} \frac{1}{i^{\alpha}} \leq \frac{1}{2} k$$

2. $\frac{1}{2} \leq np_i \leq 2 \Leftrightarrow (\frac{nk}{2})^{\frac{1}{\alpha}} \leq i \leq (2nk)^{\frac{1}{\alpha}}$. Hence, we have from Lemma 3

$$\sum_{(\frac{nk}{2})^{\frac{1}{\alpha}} \leq i \leq (2nk)^{\frac{1}{\alpha}}} H(np_i) \leq \left(\frac{3}{2}nk\right)^{\frac{1}{\alpha}} \frac{1}{e}.$$

3. $np_i \geq 2 \Leftrightarrow i \leq (2nk)^{\frac{1}{\alpha}}$. Then $\frac{nk}{i^{\alpha}} - \frac{nk}{(i+1)^{\alpha}} \geq \frac{\alpha nk}{i^{\alpha+1}} \geq \frac{2\alpha}{i} \geq \frac{2\alpha}{(2nk)^{\frac{1}{\alpha}}}$. From Lemma 3 it follows $H(np_i) \leq 2 \cdot 2^{-np_i}$. This implies

$$\sum_{i=1}^{(2nk)^{\frac{1}{\alpha}}} H(np_i) \leq \sum_{i=1}^{(2nk)^{\frac{1}{\alpha}}} 2 \cdot 2^{-np_i} \leq \sum_{i=1}^{\infty} 2 \cdot 2^{-\frac{2\alpha i}{(2nk)^{\frac{1}{\alpha}}}} \leq \frac{2}{1 - 2^{-\frac{2\alpha}{(2nk)^{\frac{1}{\alpha}}}}} \leq \frac{(2nk)^{\frac{1}{\alpha}}}{\alpha}$$

If $m \geq n$, then in the second and third case the sums change only by a constant factor which implies the claim.

We now consider the case $m \leq n$ differentiate as following:

$$1. \quad np_i = \frac{nk}{i^\alpha} \leq \frac{1}{2} \Leftrightarrow i \geq (2nk)^{\frac{1}{\alpha}}$$

Analogous to the above, we get an upper bound for the expectation of

$$\frac{1}{2} mp_i k \leq \frac{1}{2} \frac{m}{n} np_i k \leq \frac{1}{2} \frac{m}{n} k.$$

$$2. \quad \frac{1}{2} \leq np_i \leq 2 \Leftrightarrow \left(\frac{nk}{2}\right)^{\frac{1}{\alpha}} \leq i \leq (2nk)^{\frac{1}{\alpha}}$$

For the expectation this implies an upper bound of

$$\left(\frac{3}{2}nk\right)^{\frac{1}{\alpha}} \frac{1}{e} mp_i \leq \frac{m}{n} \left(\frac{3}{2}nk\right)^{\frac{1}{\alpha}} \frac{2}{e}$$

$$3. \quad np_i \geq 2 \Leftrightarrow i \leq (2nk)^{\frac{1}{\alpha}}$$

This is the only non-trivial case. Then $\frac{nk}{i^\alpha} - \frac{nk}{(i+1)^\alpha} \geq \frac{\alpha nk}{i^{\alpha+1}} \geq \frac{2\alpha}{i} \geq \frac{2\alpha}{(2nk)^{\frac{1}{\alpha}}}$

$$\sum_{i=1}^{(2nk)^{\frac{1}{\alpha}}} H(np_i) \min\{1, mp_i\} \leq \underbrace{\sum_{i=1}^{(2mk)^{\frac{1}{\alpha}}} H(np_i) mp_i}_A + \underbrace{\sum_{i=(2mk)^{\frac{1}{\alpha}}}^{(2nk)^{\frac{1}{\alpha}}} H(np_i)}_B$$

From Lemma 3 follows $H(np_i) \leq 2 \cdot 2^{-np_i}$. This implies for the sum A:

$$\sum_{i=1}^{\infty} H(np_i) mp_i \leq \frac{m}{n} \sum_{i=1}^{\infty} H(np_i) np_i \leq \frac{m}{n} \frac{(2nk)^{\frac{1}{\alpha}}}{\alpha}$$

For the sum term B we get the following upper bound

$$\begin{aligned} \sum_{i=(2mk)^{\frac{1}{\alpha}}}^{(2nk)^{\frac{1}{\alpha}}} H(np_i) &\leq \sum_{i=(2mk)^{\frac{1}{\alpha}}}^{(2nk)^{\frac{1}{\alpha}}} 2 \cdot 2^{-np_i} \leq 2 \cdot 2^{-\frac{n}{2m}} \cdot \sum_{i=0}^{\infty} 2 \cdot 2^{-np_i} \\ &\leq 4 \frac{m}{n} \sum_{i=0}^{\infty} 2 \cdot 2^{-np_i} \leq 4 \frac{m}{n} \frac{(2nk)^{\frac{1}{\alpha}}}{\alpha} \end{aligned}$$

For this estimation we used $x \leq e^x$ and $2^{-\frac{n}{2m}} = \frac{1}{e^{\frac{\ln 2}{2} \frac{n}{m}}} \leq \frac{m}{n} \cdot \frac{2}{\ln 2}$.

□

Summary. All distributions perform best when the number of receivers is at least the number of senders. The uniform distribution is the best choice if $n \approx C$, yet such a ratio cannot be taken for granted. The geometric distribution cannot be recommended, while the factorised geometric distribution achieves the best results with few senders and many channels. The success of the Pareto distributions depends on the factors C, α

Table 1. Performance for Receivers > Senders

Distribution	Performance	Use in Case
Uniform	$\Theta(n)$	sender \approx channels
Pareto	$\Theta\left(\left(n \cdot k_{\alpha,n}\right)^{\frac{1}{\alpha}}\right)$	senders > channels
Factorised	$\Theta(s)$	senders \ll channels
Geometric	$\Theta(1)$	one sender

and the resulting k . So if s is larger than $O\left(\left(n \cdot k_{\alpha,C}\right)^{\frac{1}{\alpha}}\right)$, then the factorised geometric distribution should be used. For an overview see Table 1.

From the analytical insights of the proofs there is high evidence that in fact the factorised geometric distribution is even optimal for an unlimited number of channels. Very recently, we have verified the following conjecture which is proven in a subsequent paper currently under submission.

Conjecture 1. *The factorised geometric probability distribution is asymptotically optimal if the number of channels is unlimited, i.e. there exists $c > 0$ for all $n_0 \leq N_0 \leq m_0 \leq M_0$ such that for $s = \lfloor \frac{n_0}{2} \rfloor$*

$$\min_{n \in [n_0, N_0]} \min_{m \in [m_0, M_0]} \mathbf{E} \left[M_{p^s\text{-geo}, p^s\text{-geo}}^{n,m} \right] \geq c \cdot \sup_{p, p'} \min_{n \in [n_0, N_0]} \min_{m \in [m_0, M_0]} \mathbf{E} \left[M_{p_i, p'_i}^{n,m} \right].$$

We have not explained yet why senders and receivers should use the same probability distribution. If the senders choose the uniform distribution, each channel has the same probability. So, there exists no frequency on which receivers meet a sender with a higher probability than on other frequencies. Such channels are important if there are only few senders. However, situations with many receivers may result in duplicates. Additionally, if the receivers use the factorised geometric and the senders the uniform distribution, with a increasing number of senders more messages will be unheard. That is why the receivers should also use the uniform distribution.

If senders use the factorised geometric distribution, the first s frequencies have the highest probability. So, in situations with few senders and few receivers the communication participants meet with a higher probability than if both use the uniform distribution. However, in situations with many senders interference occurs on these channels with the highest probability. If additionally only few receivers are available, the factorised geometric distribution should be factorised with a higher factor, e.g. factor $3s$ for receivers and s for senders, and of course, the same channel numbering. If there are many receivers a uniform distribution achieves best results because receivers have the same probability for each channel. So receivers also choose channels on which interference is not very probably.

As described above it depends on the ratio between senders and receivers whether or not nodes should use the same distribution for sending and receiving. Due to the fact that this ratio is unknown and can vary widely, all nodes should use the same probability distribution.

4.3 Empirical Results

In this section we compare the results of the uniform, the factorised geometric and Pareto one distribution in different situations. With the aid of a defined ratio (success ratio) of senders and expected values as well as of receivers and channels the performance is compared in different situations. The average success ratio was calculated for 96, 120 or 196 channels¹ and from 10 to 60 senders and 5-35 receivers. We have only studied situations in which the number of senders never rises above the number of channels.

The factorised geometric distribution outperforms the uniform distribution. As mentioned before, the factor s in the factorised geometric should lay close-by 12 and 14.

Example 1. In figure 2 the success ratio of the factorised geometric distribution for 96 channels is shown on the vertical axis. On the horizontal axis the stretch factor proceeds. The Pareto one distribution achieves an average success of 0.58097 where $k = 0.1943$. An average success ratio of 0.6468 has the uniform distribution. For dilation factor 12 maximises the factorised geometric distribution with an success ratio of 0.7608.

These results argue for the use of the factorised geometric distribution. However, if there are many communication participants, the uniform distribution seems to be the best choice.

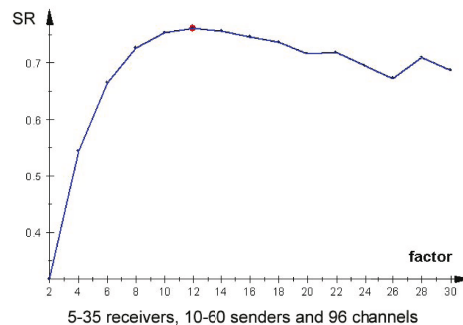


Fig. 2. Success Ratio of the Factorised Geometric Distribution

Graph plots a) and b) in Figure 3 show the expectations on the vertical axis and the number of senders on the horizontal axis. They compare the uniform distribution, the factorised geometric distribution, and the Pareto distribution with $\alpha = 1$.

The figure shows that the factorised geometric distribution achieves worse results when the number of senders increase. For this reason a combination of the uniform and the factorised distribution should be used in a routing protocol. The Pareto one distribution outperforms the factorised if there exist many senders. However, in those situations the uniform distribution performs at its best.

¹ These values are estimations for a real-world supermarket scenario.

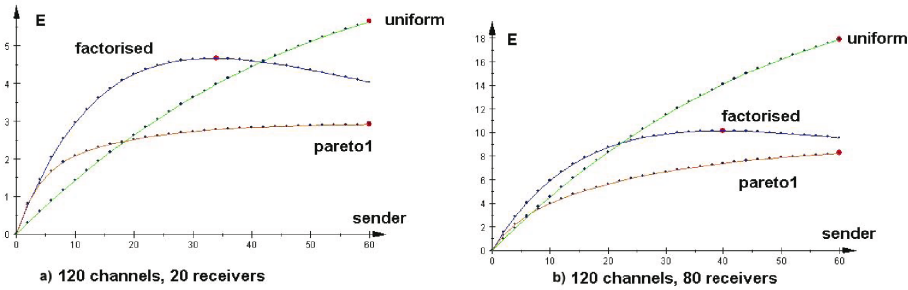


Fig. 3. Expectations with few (a) and many (b) Receivers

4.4 Distributions and Synchronisation

For synchronised nodes the usage of the factorised geometric distribution offers many advantages: In a synchronised network all senders send at the same time. So all nodes in sending range are potential senders. The throughput of messages can be increased if a dynamic measure for the hop distance to the server is used. Furthermore, it should realise that nearly the same number of nodes are in each distance.

To reduce the probability of an interference caused by senders in different hop distances, the factorised geometric distribution is a very good instrument: The permutation of the channel numbering can depend on the distance. For example, the sending range includes only nodes of two other distances: We take a look at a node of distance 5, it can communicate only with nodes of distance 4 and 6 (see figure 4). To reduce the probability that an interference occurs due to two senders of different distances, the channel numbering permutes. The first s channels have the highest probability to be chosen for sending if the node has an even distance. Otherwise the last s channels have the highest probability. If nodes try to receive a message, the reverse order of channel numbering has to be used to reach a good throughput. With this modification interference caused by senders of different distances are significant less probably than those caused by senders of same distance.

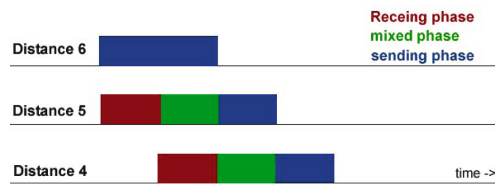


Fig. 4. Modified global synchronisation

A modification of the (global) synchronisation can realise that only nodes of two different distances are sending at the same time (see figure 4). For this feature, the wake up time depends on the distance and if a message is in the buffer. The first rounds

are only for receiving data. If a node has a not delivered messages in sleep mode, it can continue sleeping in these receiving phases. After some receiving rounds all nodes which have a message try to send while other nodes still listen for a message. Therewith a **local** synchronisation of the wake up time is realised.

5 Conclusion

We have seen that the throughput of messages can be increased by a combination of the uniform and the factorised geometric distribution. In doing so, a high number of channels is desirable. Currently, only few frequencies are supported by radio transceivers. This restriction, however, is merely caused by the current paradigms in wireless networking and national laws on radio frequency utilisation, which should not hinder research from considering wider frequency spectra.

The developed channel selection method can be used in every MAC Layer. We have only assumed that nodes are synchronised with a standard synchronisation scheme. If collisions are not explicitly avoided and messages are small, sending and receiving actions take a short time. This is advantageous regarding latency and energy consumption. Because of the short transmission, the energy costs of internal rounds are negligible. The shorter the messages the more channels can be simulated with the same energy consumption.

The uniform distribution for the channel selection does not always achieve the best expectations. The factorised geometric distribution has advantages in situations with only few communication participants and supports multiple round protocols by the permutation of channel ordering: Nodes with different distances to the central station use different probabilities for the same channels. In that way the probability of interference caused by senders of different distances is reduced. The throughput rises with a more probabilistic successful transmission. Messages need less sending attempts for successful transmitting and so energy is saved.

Here, we have achieved one of the important aims by developing algorithms for sensor networks: Messages need less sending attempts for successful transmitting and so energy is saved. A recent, yet unpublished, paper of the authors proves the conjecture stated in section 4.2 to be true: The factorised geometric distribution is asymptotically optimal for one sending attempt if the number of channels are unlimited. In the same paper, the choice of probability distribution is improved for a multiple round model.

References

1. Shih, E., Cho, S., Ickes, N., Min, R., Sinha, A., Wang, A., Chandrakasan, A.: Physical Layer Driven Protocol and Algorithm Design for EnergyEfficient Wireless Sensor Networks. International Conference on Mobile Computing and Networking, Proceedings of the 7th annual international conference on Mobile computing and networking, 272-287 (2001)
2. Warneke, B., Last, M., Liebowitz, B., Pister, K.S.J.: Smart dust: Communicating with a cubic-millimeter computer. *Computer* **34** (2001) 44–51
3. Hill, J., Szewczyk, R., Woo, A., Hollar, S., Culler, D., Pister, K.: System architecture directions for networked sensors. In: ASPLOS-IX: Proceedings of the ninth international conference on Architectural support for programming languages and operating systems, New York, NY, USA, ACM Press (2000) 93–104

4. Jeong, J., Ee, C.T.: Forward Error Correction in Sensor Networks. UCB Technical Report (2003)
5. Langendoen, K., Halkes, G. In: *Embedded System Handbook*, R. Zurawski (editor). CRC Press (2005) Chapter: Energy-efficient Medium Access Control.
6. Han, Q., Lazaridis, I., Mehrotra, S., Venkatasubramanian, N.: Sensor data collection with expected reliability guarantees. In: *PERCOMW '05: Proceedings of the Third IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOMW'05)*, Washington, DC, USA, IEEE Computer Society (2005) 374–378
7. Heinzelman, W.R., Chandrakasan, A., Balakrishnan, H.: Energy-efficient communication protocol for wireless microsensor networks. In: *HICSS '00: Proceedings of the 33rd Hawaii International Conference on System Sciences-Volume 8*, Washington, DC, USA, IEEE Computer Society (2000) 8020
8. Handy, M., Grassert, F., Timmermann, D.: Dcp: A new data collection protocol for bluetooth-based sensor networks. In: *DSD '04: Proceedings of the Digital System Design, EUROMICRO Systems on (DSD'04)*, Washington, DC, USA, IEEE Computer Society (2004) 566–573
9. Awerbuch, B., Holmer, D., Rubens, H.: The Pulse Protocol: Energy Efficient Infrastructure Access. (IEEE Infocom)
10. Awerbuch, B., Holmer, D., Rubens, H., Wang, I.J., Chang, K.: The Pulse Protocol: Sensor Network Routing and Power Saving. *MILCOM* (2004)
11. Intanagonwiwat, C., Govindan, R., Estrin, D.: Directed diffusion: a scalable and robust communication paradigm for sensor networks. In: *MobiCom '00: Proceedings of the 6th annual international conference on Mobile computing and networking*, New York, NY, USA, ACM Press (2000) 56–67
12. Singh, S., Raghavendra, C.: PAMAS - Power Aware Multi-Access protocol with Signalling for Ad Hoc Networks. *SIGCOMM Comput. Commun. Rev.* **28** (1998) 5–26
13. Karn, P.: MACA- a New Channel Access Method for Packet Radio, S.134-140. *ARRL/CRRL Amateur Radio 9th Computer Networking Conference* (1990)
14. Schiller, J.: *Mobile Communications*, Pages 37-42 and 61-81. (2000)
15. Schindelhauer, C., Liu, M.J., Ruehrup, S., Volbert, K., Dierkes, M., Bellgardt, A., Ibers, R., Hilleringmann, U.: Sensor Networks with more Features using less hardware. *GOR International Conference Operations Research*, Page 63 (2004)
16. Roberts, L.G.: Aloha packet system with and without slots and capture. *SIGCOMM Comput. Commun. Rev.* **5** (1975) 28–42
17. Vanderplas, C., Linnartz, J.P.M.: Stability of mobile slotted ALOHA network with Rayleigh fading, shadowing and near-far effect. *IEEE Trans. Vehic. Technol.* **39**, 359-366 (1990)
18. Namislo, C.: Analysis of mobile radio slotted ALOHA networks. *IEEE J. Selected Areas Commun.* **2**, 199-204 (1984)
19. Borgonovo, F., Zorzi, M.: Slotted aloha and cdpa: a comparison of channel access performance in cellular systems. *Wirel. Netw.* **3** (1997) 43–51
20. Zheng, T., Radhakrishnan, S., Sarangan, V.: Pmac: An adaptive energy-efficient mac protocol for wireless sensor networks. In: *IPDPS '05: Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS'05) - Workshop 12*, Washington, DC, USA, IEEE Computer Society (2005) 237.1
21. Group, B.S.I.: Specifications of the Bluetooth System Vol. 1, v. 1.OB 'Core' and Vol.2, v. 1.OB 'Profiles'. (1999)
22. Golmie, N.: Bluetooth dynamic scheduling and interference mitigation. *Mob. Netw. Appl.* **9** (2004) 21–31
23. Siegemund, F., Rohs, M.: Rendezvous layer protocols for bluetooth-enabled smart devices. *Personal Ubiquitous Comput.* **7** (2003) 91–101