Distributed Data Delivery in Partially Connected Networks

Mikkel Gade Jensen , Janne Dahl Rasmussen , Anders Grauballe, Achuthan Paramanathan
Supervisor: Tatiana K. Madsen
7th Semester, Group 720, Institute of Electronic Systems 2007

AALBORG UNIVERSITY

Introduction

The basic scenario in this project is a cluster of sensors which are partially wirelessly connected with a gateway. A sensor is a small device which can measure environmental conditions. It is also able to communicate and relay messages to a control unit or gateway, see **Figure 1**. Each sensor and the gateway is partially connected because they are not active at the same time, they are e.g. out of range or powered off. The gateway will in this scenario not always be able to receive each measurement from the sensors.

Therefore the objective of this project is to develop a distribution model, see **Figure 2** such that data from each sensors can be received by a gateway even if not all sensors are online / in range.
Both online time and data storage capacity are considered.

Methods

A solution to this problem could be to distribute a full copy of the measurement to other sensors in the cluster, but because it is necessary to consider storage capacities on the devices it is not optimal to store the entire message from a device on several others. Therefore a suitable method for adding redundancy and making the system fault tolerant towards random failures is required. Therefore the Reed-Solomon coding scheme is proposed, this coding scheme adds *m* redundancy units allowing *m* arbitrary failures in the system [1].

Proposed algorithm

A binomial distributed probabilistic system model has been derived to establish the relationship between the online probability p_{on} of the devices and the system reliability p_R [2].

Data distribution probability:

$$p_d = 1 - \sum_{l=1}^{n} \binom{n}{l} (1-p_{on})^{N \cdot l} (1-(1-p_{on})^{N})^{n-1}$$

Reconstruction probability:

$$p_r = \sum_{k=n-m}^{n} \binom{n}{k} p_{on}^{k} (1-p_{on})^{n-k}$$

System reliability:

$$p_R = p_d \cdot p_r$$

where *n* = total number of devices, *m* = redundancy devices, *N* = number of device cycles between gateway presence.
This model has been tested using a simulation implemented in Java, see **Figure 3**.
To observe how the system performs in a real life scenario it has been implemented in sensors provided by the university, see **Figure 4**.

Scenario

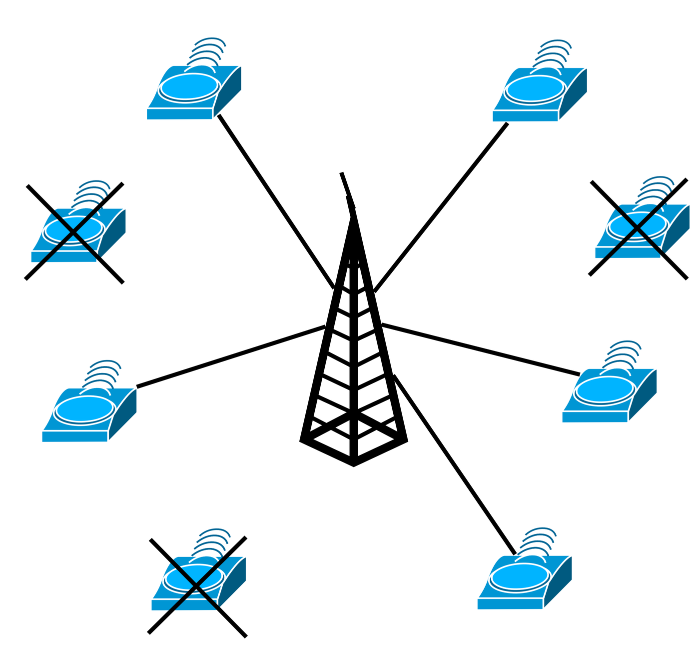

Figure 1: The basic scenario consisting of a gateway and sensors all in range of each other. Sensors are randomly unavailable to the gateway.

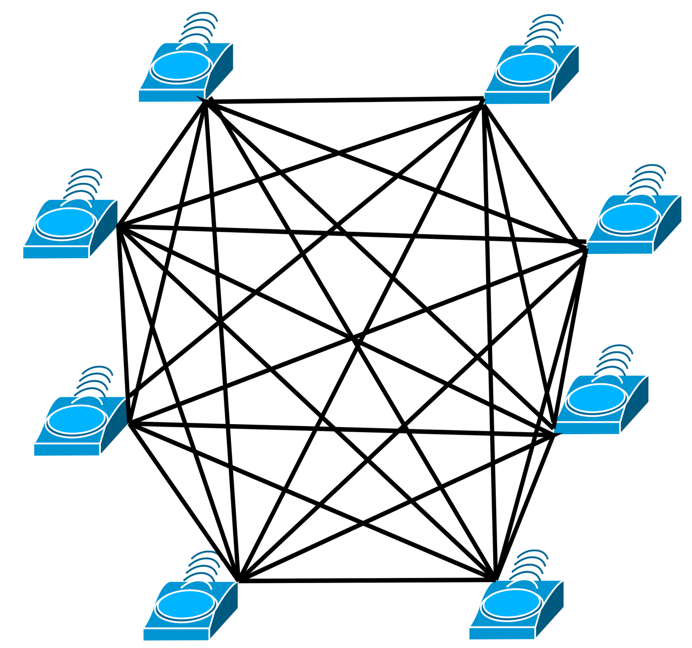

Figure 2: All sensors are in range of each other and are able to distribute data in order to increase the reliability.

Methods

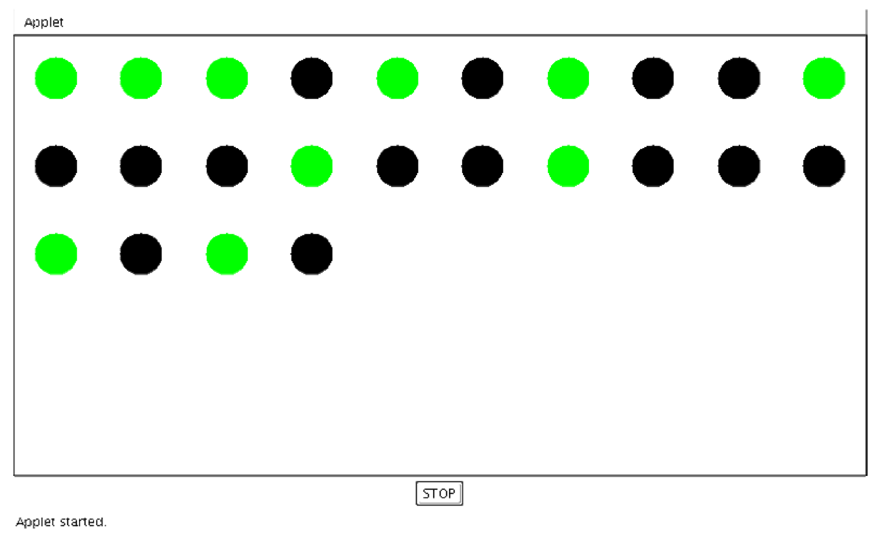

Figure 3: Screen shot showing the visual part of the simulation. Each dot represent a sensor, where green indicates online and black offline. In this case the number of sensors are 24.

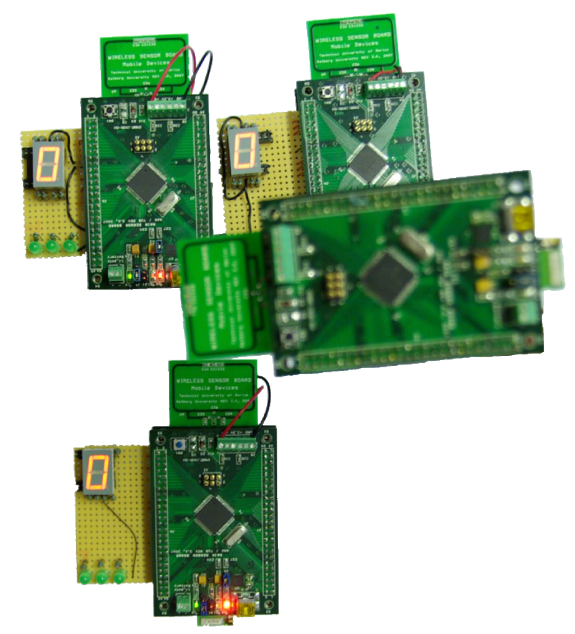

Figure 4: System prototype consisting of 3 sensors and one gateway.

Results

Online probability	Probabilistic model (reliability)	Simulation (reliability)	Prototype (reliability)
50 %	50 %	49.4%	42.5 %
80 %	90 %	90.3%	60.5 %
90 %	97.5 %	96.9%	79.5 %

Table 1: Result from model, simulation and prototype of the system for three online probabilities.

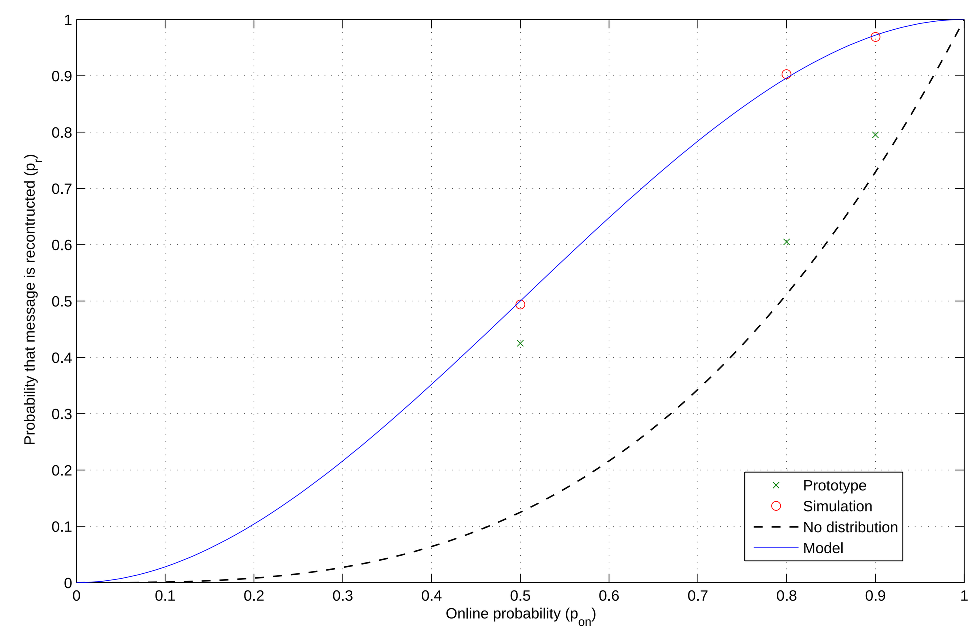

Figure 5: The system reliability as a function of online probability for Reed-Solomon coding with two data unit and one redundancy unit in comparison to no distribution.

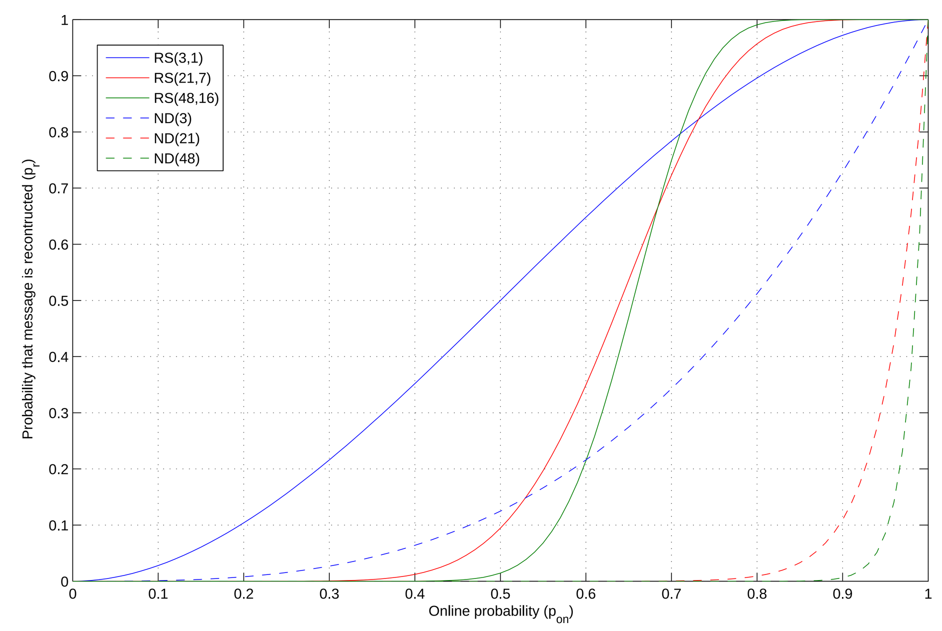

Figure 6: Each solid line represents Reed-Solomon coding with a ratio of 3:1. RS(3,1) as an example shows Reed-Solomon coding with three units, where one acts as redundancy unit. The corresponding dashed lines shows no distribution (ND) models.

Results

The following results have been obtained during this project:

Probabilistic system model describing the system reliability as a function of the online probability.
Simulation which verifies the probabilistic model and visualizes the changes in the system. The simulator is able to simulate up to 50 sensors, see **Figure 3**.
Prototype consisting of three sensors and one gateway has been implemented, see **Figure 4**. It is possible for the sensors to distribute data in the cluster and for the gateway to recover it. Furthermore it is possible to see the state changes for each sensor.

The results for the system reliability at three different online probabilities are stated in **Table 1** and outlined in **Figure 5**.

Discussion

As seen from the results the methods have proven to be suitable for the chosen scenario.

The results from the simulation verifies the system model.

The results from the prototype test shows a deviation from the system model and simulation, see **Table 1** and **Figure 5**. This is mainly due to a MAC protocol not suited for the scenario and a shared wireless medium.

From **Figure 6** it can be seen that the system reliability p_R is higher than the online probability p_{on} of each device, for large p_{on}. Furthermore it can be seen that the Reed-Solomon coding scheme can maintain the same level of reliability, when p_{on} is decreased and the number of devices in the system is increased.

It can be concluded from **Figure 5 and 6** that the distribution model with use of Reed-Solomon coding provides higher system reliability than the no distribution model.

References

[1] James S. Plank, “A Tutorial on Reed-Solomon Coding for Fault-Tolerance in RAID-like Systems”,
[2] Sheldon M. Ross, “Introduction to probability and statistics for engineers and scientists”