

MOBILE AD-HOC NETWORK PERFORMANCE ASSESSMENT BASED ON SIMULATION WITH SCENARIO SPECIFIC PROPAGATION AND MULTI-LINK MODELLING

Designing a communication system for disaster relief in the Philippines

By: Rasmus Liborius Bruun



AALBORG UNIVERSITY
STUDENT REPORT



AALBORG UNIVERSITY
STUDENT REPORT

Institute of Electronic Systems

Fredrik Bajers Vej 7

DK-9220 Aalborg Ø

Title:

Mobile ad-hoc network performance assessment based on simulation with scenario specific propagation and multi-link modelling

Theme:

Networks and Distributed Systems

Project Period:

10th semester, spring 2018

Project Group:

NDS10-1020

Participant:

Rasmus Liborius Bruun

Supervisor:

Troels Bundgaard Sørensen

Copies: 1

Pages: 53 (70)

Date of Completion:

June 7th, 2018

Abstract:

Natural disasters can destroy cellular and wired network infrastructure, subsequently leaving impacted areas without means of communication. The Reachi system combine mobile ad-hoc network devices with a gateway to aid the work of relief coordinators, by enabling communications from the catastrophe epicenter.

A multi-link model which describe the wireless links in the deployment area of the Reachi system is specified. The model considers spatial correlation of links, which is a commonly overlook factor in network evaluations. The model is utilized for computer simulations. Simulation results using the model show consistently lower network performance in comparison to the equivalent model without spatial correlation.

Simulations with the spatial correlation model is an important tool for design of link layer and network layer protocols for the Reachi system. Link layer protocols should synchronize the network and rely on randomization. Further evaluation is necessary to decide between a proactive or reactive routing scheme.

The content of this report is freely available, but publication (with reference) may only be pursued in agreement with the author

Preface

This report is the master thesis under the study program *Networks and Distributed Systems* at Aalborg University. The work was carried out in the spring semester 2018.

The thesis is inspired by challenges and tasks from a collaboration project titled Reachi. The Reachi project is a collaboration of the companies LinkAiders ApS, NeoCortec A/S and Attensys.io GmbH and two departments on Aalborg University; Wireless Communication Networks (WCN) and Center for Embedded Software Systems. LinkAiders has kindly provided measurements, which are of key importance to the thesis. The thesis address modelling of wireless links, which is beneficial for evaluation of the Reachi system through simulations.

Troels Bundgaard Sørensen, Associate Professor at the WCN section, has supervised the creation of the thesis, through fruitful discussion, vital observations and well considered recommendations. Tatiana Kozlova Madsen, Associate Professor at the WCN section, has contributed with knowledge and brainstorming related to the mobile ad-hoc network topics in the thesis.

The reader of this thesis is expected to have fundamental knowledge in algebra, calculus, probability theory, computer science and communication systems engineering. Familiarity with common network related terminology is assumed. Appendix A gives a brief brush up of overlapping network terminology.

References are made using the IEEE referencing, where a number in a square bracket in the text, e.g. [1], refers to an entry in the references on page 54. Each of the categories *equations*, *figures*, *tables* and *algorithms* are enumerated sequentially after appearance. Unless otherwise stated, the figures are made by the author.

neo.cortec

attensys.io



AALBORG UNIVERSITET



Table of Contents

1	Introduction.....	1
2	Preliminary analysis	3
2.1	Disaster relief in the Philippines.....	3
2.1.1	Networking conditions.....	4
2.2	Designing the right solution	5
2.2.1	Satellite and balloon networks.....	6
2.2.2	Recovering/extending existing infrastructure	6
2.2.3	Mobile ad-hoc network.....	7
2.3	Problem formulation	8
2.3.1	Methodology	8
3	Link modelling.....	10
3.1	Pathloss	10
3.1.1	Distance dependent path loss.....	12
3.1.2	Slow fading	13
3.1.3	Fast fading.....	13
3.1.4	Combined fading.....	14
3.2	Correlation models	14
3.2.1	Temporal correlation.....	15
3.2.2	Spatial correlation.....	16
4	Empirical modelling.....	20
4.1	Measurement campaign.....	20
4.1.1	Equipment	21
4.1.2	Procedure.....	23
4.1.3	Data.....	23
4.2	Data analysis and model regression.....	24
4.2.1	Pathloss exponent and standard deviation	25
4.2.2	Correlation parameters.....	27
5	Simulation	29

5.1	Simulation description.....	29
5.1.1	Computational aspects/other observations.....	30
5.2	Model verification and validation.....	31
5.2.1	Verification.....	31
5.2.2	Validation.....	32
6	Performance evaluation.....	36
6.1	Evaluation scenarios.....	36
6.1.1	Static scenarios.....	37
6.1.2	Mobile scenarios.....	40
6.2	Results and discussion.....	43
6.2.1	Static scenarios.....	43
6.2.2	Mobile scenarios.....	45
6.3	MANET protocol design considerations.....	47
6.3.1	The layered protocol stack.....	47
6.3.2	Data link layer protocol design considerations.....	48
6.3.3	Network layer protocol design considerations.....	50
7	Conclusion.....	52
7.1	Future work.....	53
8	References.....	54
	Appendix.....	58
A	Terminology.....	58
B	Autocorrelation sequence of the first order autoregressive function.....	59
C	Data format.....	61
8.1.2	Peculiarities.....	62

1 | Introduction

Disaster relief is an important humanitarian act. Natural disasters like droughts, floods, tsunamis, typhoons, earthquakes and forest fires destroys communities, and require quick action in the local area to aid the injured, access the damage and start restoring houses or otherwise restore the life of those affected. Globally, tens of thousands of people are killed, and hundreds of millions are affected by natural disasters every year [1]. Humanitarian organizations have resources to step in and help during disasters. However, coordination of the effort is crucial to minimize the casualties, damages and waste of available resources. Overview of the situation and communication between relief workers and coordinators is vital as illustrated by the many initiatives initiated to incorporate information and communication technology (ICT) into the works of humanitarian organizations [2], [3]. Two technologies are the main enablers of communication across large areas: cellular and satellite communication. However, during many natural disasters traditional communication infrastructure is not available. Cellular infrastructure might not be deployed, have suffered damage or power loss and satellite equipment might be too expensive for sufficient adoption. An inexpensive ad-hoc solution is desired.

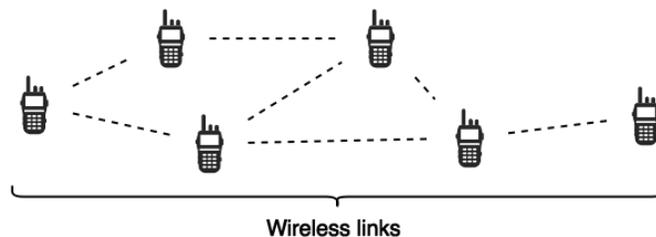


Figure 1 Mobile ad-hoc networks consists of mobile nodes connected by wireless links.

Mobile ad-hoc networks (MANET) is a class of networks with the potential to be the desired inexpensive solution. MANETs are characterized by being infrastructure-less and self-configuring. Figure 1 illustrates how MANETs consist of wireless devices. Participating nodes are wireless and can be highly mobile. In contrast, largely adopted wireless networks like cellular and WiFi experience only mobility of the user equipment and typically only the last link is wireless, see Figure 2. Lots of theory and empirical knowledge have been accumulated to continually improve coverage, capacity, delay, etc. of especially cellular networks, incentivized by telecom providers desire to offer the best product, which draws customers to their services. MANETs has a desired flexibility and in contrast to cellular, WiFi and satellite networks, MANETS require no infrastructure. This is a major advantage in the response to natural disasters, as networks form where needed without need for specialist intervention. Additionally, equipment cost is contained to the price of individual terminals. Unfortunately, the flex-

ibility of MANETs is also their biggest weakness. The goal of any network is to enable communication between the participating nodes. Guaranteeing network connectivity in a network, where every node is mobile, constitutes a challenge.

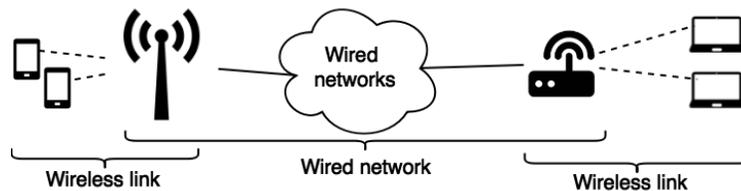


Figure 2 Traditional wireless systems are only wireless in the last link from the user equipment to the access point. Mobility is limited to the user equipment.

The prospect of utilizing MANETs in disaster relief aid is interesting, because it has potential to provide communications capabilities to every person involved in the relief effort with minimal delay and cost. The company LinkAiders has conceptualized the idea of using MANET technology in disaster relief in a project called *Reachi*. Reachi has won a social entrepreneurship award [4] and received funding towards further development, optimization and production of the Reachi system [5]. LinkAiders have partnered up with Philippine Red Cross (PRC), who will be the first client on the system. In collaboration with PRC volunteers, LinkAiders have performed field tests with the wireless Reachi prototype devices.

The goal of this project is to aid the evaluation and optimization of the Reachi system. Particularly, with a basis in the Reachi field tests and supported by current literature on MANETs, this project investigates the performance and how to optimize the MANET technologies, which constitute the Reachi system. The contributions of the present work are the following:

- Scenario specific propagation model based on empirical data
- Multi-link model including correlation between links, parametrized based on empirical data
- Assessment of link connectivity in specific MANET scenarios based on simulation
- Recommendations for design of data link layer and network layer protocols for a MANET in a specific scenario

2 | Preliminary analysis

Mobile ad-hoc networks are often specialized networks with specific purposes. Applications include real-time voice communication between military units, reporting of measurements from smart meters and control of smart home IoT-devices. Obviously, the requirements for the network depend on the application. Latency is a concern in real-time voice communication but is less critical in sensor measurements. Throughput requirements are larger for video surveillance than activating a door lock. While requirements are commonly easy to specify, evaluating whether a particular network fulfils the requirements is a task made complex by the many variables affecting wireless mobile ad-hoc networks.

This chapter will identify the variables relevant for communications in a disaster scenario, with the purpose of devising a model which can be useful for analysis of mesh network protocols. The focus will be on a specific use case of relief work in the Philippines. To generalize the findings in this project and justify specific decision, a broader outlook is given where suitable.

2.1 Disaster relief in the Philippines

Philippines is an island nation situated in the Western Pacific Sea. The geography of the Philippines makes them prone to mortal and destructive earthquakes and typhoons. To combat these natural disasters, Philippines Red Cross (PRC) has identified four stages on which they handle disasters. The four stages are [6]:

1. Disaster risk reduction. Assessment of vulnerable areas lead to focused effort to either reinforce vulnerable points or establish appropriate early warning systems and plans.
2. Disaster preparedness. Training staff and local volunteers in emergency response. Stocking equipment and supplies in anticipation of disasters.
3. Disaster response. Assessing the situation, planning the effort of search, rescue, healthcare and supply distribution.
4. Disaster recovery. Providing shelter for those who lost their homes and reestablishing the community.

PRC has a strong presence among the local community, and a big base of volunteers counting 0.5 million in 2011 with expectation to reach 1.8 million [7]. In the event of a disaster, the volunteers will work together with Red Cross staff and coordinators to focus the relief effort. Despite the impressive workforce, the impact of relief work is held back by the volunteer's inability to communicate the situation to the relief coordinators. Without updated information it is a challenge for relief coordinators to dispatch resources timely and in the right amounts to the correct locations. There is a need for an ad-hoc communication solution to improve the

coordination of the disaster response stage. The focus of the present work is stage 3. Disaster response.

2.1.1 Networking conditions

To provide the optimal network for disaster communications it is vital to know the operating conditions of the network. Network conditions include the following:

- **Environmental conditions**

For any wireless communication, the environment in which the network is deployed plays a vital role. Here, the term environment is used only to refer to physical objects, terrain and phenomena, which affect the propagation of electromagnetic waves in the radio frequency spectrum. Different things are to be considered in e.g. communication in a barren desert versus communication in a hilly city like San Francisco. Although obstruction of the signal path is arguably the main environmental concern, weather conditions will also affect the network performance. Though most prevalent at radio frequencies above 11 GHz, atmospheric precipitation can cause signal attenuation. For disaster relief in the Philippines the focus is populated urban areas. This encompass a variety of environments with varying building density and height, terrain undulation and vegetation.

- **Equipment performance**

Naturally, a network is bounded by the capabilities of the networking equipment. Equipment performance describes the range, battery life, robustness, etc. of the equipment itself.

- **Equipment mobility**

Equally important is the placement and mobility of the equipment. The placement of equipment determines which wireless links can be established, which in turn determines the topology of the network, and whether there are available paths between nodes in the network. Mobility introduces further complexity as it allows isolated nodes to reconnect to the network and establishment of new paths. Conversely, nodes are at risk of disconnection from the network and routes might be broken.

Realistic disaster mobility is difficult to specify, but compared to the simplest mobility models e.g. random walk, a few things are common for the mobility in most disaster scenarios [8]:

- Entities are heterogeneous, i.e. some entities move much more than others and at different speeds.
- Repetitiveness is common among certain entities, e.g. transport of people or supplies between points of interest.
- Group mobility is common as both civilians and rescuers travel in groups to increase their safety.
- Other relationships between entities, e.g. a squad of search and rescue workers scan through different neighborhoods in search of survivors.

Multiple disaster mobility models have been specified, to capture these aspects in certain scenarios. The post-disaster mobility (PDM) model [7], the large-scale disaster mobility model (LSDMM) [9] and the disaster area mobility model [10] rely on dedicated sectors with a particular purpose. PDM and the disaster area mobility model utilize well defined zones, e.g. the main coordination center, relief camp, evacuation center, hospital and police station of the PDM model. Nodes are assigned certain roles, and their mobility is determined from these roles. LSDMM instead rely on cellular division of a large area, where each cell is either an activity cell, an obstacle cell (dense or sparse) or an empty cell. Nodes are either static or move towards a randomly selected activity cell.

The mobility in a disaster scenario in the Philippines is difficult to specify. The situation in the Philippines is further complicated by the fact that the Philippines Red Cross relief workers are mainly local volunteers. As the volunteers are likely to be affected by the disaster, it is reasonable to assume that some will fail to follow the training that has been provided to them. The strength of recruiting numerous volunteers is their presence in the areas where disaster might strike. If means of communication can be provided to the local volunteers, disaster relief could be aided by e.g. letting relief coordinators know the needs of specific neighborhoods or communicate the locations of medical care and supplies to victims.

2.2 Designing the right solution

The following properties are important for any post disaster communication system. No two disasters are the same. Therefore, there is no perfect all-round solution and every application must find the right tradeoff between the properties.

- Coverage. The network should be available in the areas where people need help. Coverage can be of widespread geographical areas, but also related to specific outdoor or indoor environments.
- Deployment speed. Communication enables informed decision making. It is vital that means of communication is established as soon as possible following a disaster.
- Lifetime. It is vital that communication can be maintained in the first 72 hours of a disaster, but in certain cases communications might be necessary for months. The network must be delivering suitable means of communication for the entire period.
- Flexibility. In the lifespan of the network, traffic patterns may change. Both in terms of the amount of traffic and geographic location of source and destination.
- Robustness. Post disaster conditions might be hostile. Resilience to water, ash, shock and high winds are just some of the considerations.
- Network performance (latency, throughput, connectivity, etc.). Depending on the offered services, the network must live up to certain performance requirements. Services might range from simple message applications to video surveillance or chat.

Different solutions for disaster communication have been proposed. In the following, 4 potential solutions are presented

2.2.1 Satellite and balloon networks

Line of sight is a key enabler of long distance wireless communication. Airborne infrastructure has a clear advantage when it comes to coverage.

Multiple satellite network operators offer satellite phone plans, which enable communication from anywhere on earth. Obviously, infrastructure and equipment costs are high which leads to expensive phone plans. High fixed and acquisition cost is a main prohibitor of mass adoption for communication during relief work. Additionally, the satellite networks are not dimensioned to handle thousands of calls from the same relatively small area on earth. Therefore, satellite communication might be a suitable solution for few high value relief workers, e.g. medical professionals.

In comparison to launching satellites into orbit, inflating a balloon is inexpensive. Arimura et al. [11] developed and evaluated an ad-hoc network infrastructure of balloons based on 802.11. With an allowable distance up to 600 meters between balloons they were able to transmit video streams over 5 hops and maintain VoIP communication. Challenges for the system includes power consumption and robustness. Power is supplied to the balloons from the ground and balloons are naturally vulnerable to strong winds.

2.2.2 Recovering/extending existing infrastructure

In areas where communication infrastructure is damaged by a disaster, recovering or rebuilding the infrastructure is a possible solution. This strategy has previously been deployed, e.g. following the strike of the super typhoon *Haiyan* in the Philippines in 2013 [12]. 108 Huawei employees and subcontractors were dispatched in helicopter to recover the network in 3 major cities. The drawback of the recovery process is the high expense and long rebuild time. Full recovery took more than 4 months. [12]

An inexpensive variation using existing infrastructure is proposed by Minh et al. [13]. They conceptualize a system, where user equipment connects to surviving WiFi access points (AP) and extends the range by becoming virtual access points. In this way, ad-hoc networks are extending the connectivity of surviving infrastructure. The authors show that the solution is fast to setup and provides a quality of service suitable for web browsing and voice calls over several hops of virtual APs. However, the demonstrated network was setup on laptops inside a campus building, and the laptops participated by manually downloading the required software. Several issues must be addressed before the system is suitable for real life disaster communications, including power consumption, scalability, security, multiple surviving APs and handling of mobility/link failure.

2.2.3 Mobile ad-hoc network

Disaster communications is an obvious application area of MANETs. The mobility and flexibility of ad-hoc networks are highly compatible with the networking needs in disaster communication. Multiple MANET based systems have been proposed for disaster recovery. Lu et al. [14] designed the smartphone application *TeamPhone*, which seamlessly combines cellular, ad-hoc and delay tolerant networking to provide a messaging service between relief workers. Implementation on off-the-shelf hardware demonstrated the feasibility of the application under certain test conditions. The fact that the system runs on smartphones is its greatest advantage. However, the utility in actual disaster recovery is challenged by the power consumption of consumer smartphones and achievable range of 2.4GHz WiFi communication.

LinkAiders are working on a dedicated disaster communication system [15]. The vision is to equip volunteers with small robust devices, which will form an ad-hoc network. For every 1000 handheld/wearable (on a strap around the neck) devices a satellite, GSM and WiFi enabled, battery powered gateway will be deployed. Through a backhaul link (be it satellite, cellular or WiFi based), the gateway relays the information to a coordinator, who can dispatch available resources to the distressed areas, as illustrated in Figure 3.

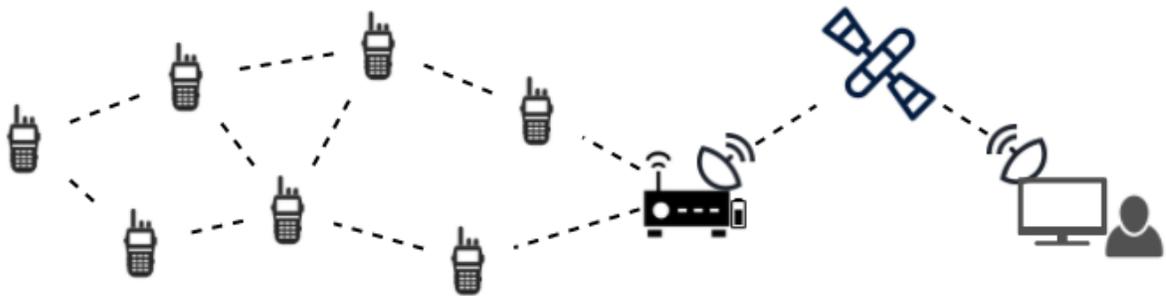


Figure 3 Disaster communication concept. Handheld/wearable devices carried by volunteers form a MANET and are used for reporting the current situation from the disaster area to the relief coordinator.

The Reachi system is designed in collaboration with the Philippines Red Cross. Thus, exact communication requirements and specification of the network conditions of the Reachi system is an ongoing effort.

The Reachi system is a potentially viable business case, which provides benefits for relief workers. MANET technology eliminates the need for dedicated infrastructure, which reduce both the required knowledge and skill for deployment of the network, but more importantly it reduces the cost of the system. The system is financed through the sales of handheld Reachi devices directly to PRC volunteers. Thus, the cost of gateways can be divided onto the cheaper handheld devices, while keeping the devices affordable - allegedly.

NeoCortec are supplying the radio-modules which will facilitate the communication. Their technology has proven itself in IoT-networks, where nodes are expected to be more static than in the disaster scenario. The big challenge will be adapting the protocols to the inherently mobile disaster scenarios. Ensuring connectivity and optimizing the network to the conditions in the Philippines is vital to ensure satisfying performance of the system. Therefore, thorough evaluation of the system is necessary.

2.3 Problem formulation

The performance of a MANET is limited by the participation of nodes. Their placements determine the coverage, connectivity and link conditions. Additionally, node mobility result in continuously changing network conditions. The penalty of being a flexible technology is the increased complexity introduced by relying on participants to form the network infrastructure.

This project work will contribute to performance enhancement of the Reachi system. The basis of the project is the following problem formulation:

How can the network performance of the Reachi system be optimized for deployment in the Philippines during a disaster scenario?

To answer the problem formulation, it is necessary to understand the performance of the Reachi system in the network environment of the Philippines, which can be divided into two subproblems:

- How does the Reachi system work?
- What implications does the network environment in the Philippines have on the network performance?

Specific details on the Reachi system are unavailable. Therefore, this thesis will focus on specifying the networking environment and giving a general insight to the performance of MANET systems in the specific environment.

2.3.1 Methodology

Conceptually, evaluation of the network performance could be carried out by either full-blown implementation, analytical evaluation or simulation. The disadvantage of a full-blown implementation is the effort required to setup and run tests, as well as the cost of scaling the network. The disadvantage of a pure analytical evaluation is difficulty of expressing the temporal protocol behavior in an analytical expression. Therefore, the method for network evaluation is simulation. However, computer simulation of the network requires appropriate models of the physical aspects of the network simulation. In this case, the necessary models are (1) a node mobility model and (2) a link model describing the radio propagation on every wireless link of the network.

The realism of the models is vital for the validity of the evaluation. Therefore, the choice/design of the models will rely heavily on measurements made on location in the Philippines. The specification of the link model will additionally draw on the accumulated knowledge in the field of propagation modelling derived from both theory and empirical observations. Specifying node mobility in a disaster scenario is beyond the scope of this project. General static and mobile node scenarios is used in place of a specific mobility model. Performance metrics are defined and the performance of two multi-link models is evaluated. The performance in the simple scenarios is the basis for a presentation of data link layer and network layer protocol considerations.

3 | Link modelling

The performance of a MANET comes down to the network conditions and the capabilities of the technology, i.e. protocols managing the system. To enable performance evaluation, firstly a means of evaluation must be established. Evaluation of any wireless networks comes down to the properties of the individual wireless links in the networks. Therefore, a specification of the wireless links in the conditions of the Philippines is needed. Once the wireless links have been specified, analysis of network protocols can commence.

Mathematical modelling is an indispensable tool for system analysis, development and evaluation. Mobile ad-hoc networks are no exception. No single model can encompass networks between military drones and networks of indoor IoT-devices. Customization for the specific application is necessary.

In this chapter, with a basis in the current literature, common link models are revised. In connection with the review, the proposed extension to include correlation in the link modelling is presented.

3.1 Pathloss

Wireless communication relies on the transmission and reception of electromagnetic waves. The wireless link can be described by the signal loss inflicted by the propagation from transmitter to receiver. The term path loss is used to describe this loss.

Accurate description of the path loss of electromagnetic waves is possible through deterministic and theoretical models e.g. the free space model or the 2-ray model. The free space model specifies the path loss in terms of frequency and distance between transmitter and receiver. In addition, the 2-ray model includes transmitter and receiver height, as it assumes a ground reflection of the signal.

While these simple models are applicable for certain studies, often it is the case that scattering, reflection, refraction and diffraction make analytical prediction of the propagation loss infeasible. The phenomena might be manageable individually and in simple scenarios but in practice uneven surfaces of everything from the grass field to the paved streets and obstacles like hills, trees and buildings call for other ways to quantify path loss. This has led to the development of empirical models, e.g. the Okumura model or the COST Hata model. The empirical models are often valid for some specific environments and frequency ranges. The Okumura model is valid for cellular like systems, where one transceiver is placed at least 30 meters above the ground level and link distances are between 1 and 100km.

The aforementioned models are deterministic in the sense, that for a given input they always return the same path loss. In the case of the empirical models, the path loss is specified as a mean or any percentile value. This is valuable in cellular network planning, where the goal is to guarantee a level of service in a percentile of the coverage area. However, for simulation purposes it is necessary to generate realizations of the path loss – not just the mean. Therefore, the models need to include stochastic elements, which can be generated by a computer.

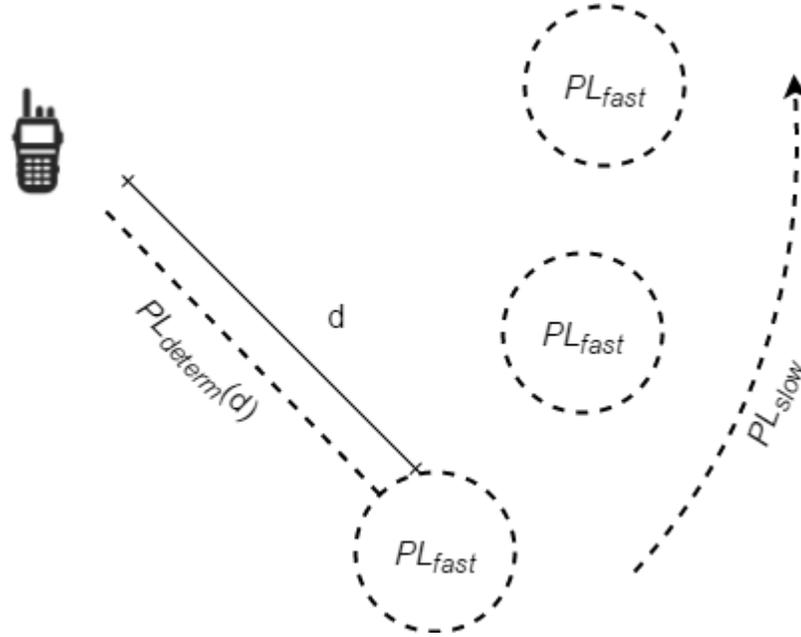


Figure 4 The three typical pathloss components. One component models the attenuation caused by distance, the slow fading describes the variation in local mean and the fast fading component models the fast fluctuations caused by multipath signal components.

A single link path loss model is typically broken into three components, as illustrated in Figure 4:

1. $PL_{deterministic}(d)$: A deterministic distance dependent part, which describes the mean signal attenuation at any given link distance (distance between transmitter and receiver).
2. PL_{slow} : A stochastic slow (/shadow) fading part, which specifies the local mean of the signal. The slow fading variable is approximately constant in a local area.
3. PL_{fast} : A stochastic fast (/multipath) fading part, which describes the rapid fluctuations in signal strength in a small area – within a few wavelengths.

Combined, the path loss components compose the total path loss, PL :

$$PL_{dB} = PL_{deterministic,dB}(d) + PL_{slow,dB} + PL_{fast,dB} \quad (1)$$

The path loss equation (1) can be used in simulations of the empirical path loss models. One simply uses the mean path loss of the empirical model in place of the deterministic path loss

and specify the stochastic (zero mean) elements for the given model. In the following subsections, the elements of (1) are explain in depth.

3.1.1 Distance dependent path loss

The free space path loss model [16] describes the relationship between transmitted, P_t , and received power, P_r , of a wave with wavelength λ , propagating a distance, d .

$$\frac{P_r}{P_t} = D_t D_r \left(\frac{\lambda}{4\pi d} \right)^2 \quad (2)$$

where D_t and D_r is the transmitter and receiver antenna directivity, respectively. Often, it is more convenient to give the pathloss in decibel (dB),

$$PL_{dB} = 10 \log_{10} \left(\frac{P_t}{P_r} \right) \quad (3)$$

where the transmitted power is in the numerator to give a positive path loss value. The practical usability of the free space path loss model is for predicting the path loss between antennas raised a significant distance above the ground and other reflectors. Theoretical path loss models as the free space model or the 2-ray model can be approximated in simple scenarios, i.e. on a football field. For deployable systems, e.g. cellular or MANETs the possible link scenarios are more complex and infeasible to describe using ray-tracing. This is where the empirical models have their strengths.

The Okumura path loss model [17] was built on measurements from Tokyo, Japan. It specifies the path loss in terms of the free space path loss, PL_{fs} , an attenuation factor A , heights of the mobile and base station, h_{MS} , h_{BS} , and some correction factors, $K_{correction}$, which signify the environments. Correction factors include terrain undulations, terrain slope and relation between land and sea.

$$PL_{Okumura,dB} = PL_{fs} + A - h_{MS} - h_{BS} + \sum K_{correction} \quad (4)$$

The complexity and expressiveness of the distant dependent path loss models vary. However, the common feature is that the distance dependent path loss follows a power law:

$$PL \propto (d)^\gamma \Leftrightarrow PL_{dB} \propto \gamma \log_{10} d \quad (5)$$

The path loss exponent γ is determined by the environment. In the Okumura model, the path loss exponent is contained in the attenuation factor and the correction factor. The free space model specifies a path loss exponent of 2. Table 1 gives path loss exponents obtained in measurements of device to device systems.

Environment	Pathloss exponent, γ
Urban line of sight	2.4 - 5

Urban non-line of sight	3.6 – 7.6
Dense forest	4

Table 1 Typical values for the pathloss exponent in different environments for device to device communication. [18], [19], [20]

3.1.2 Slow fading

When a wireless device is moving around in a city, the signal attenuation will vary with the place. On a hill top the signal will be good but large buildings might shadow the path between transmitter and receiver. These large-scale variations in the environment cause variation in the path loss. The variation caused by large obstacles is called slow or shadow fading. It is common to model the stochastic fading caused by large scale obstacles with a log-normal distribution [21]. The log normal distribution of the pathloss means, that when expressed in dB , the slow fading, $PL_{slow,dB}$, follows a normal distribution with zero mean and variance $\sigma_{S,dB}^2$.

$$PL_{slow,dB} \sim \mathcal{N}(0, \sigma_{S,dB}^2) \quad (6)$$

The standard deviation of the slow fading is determined by the environment. Common values for cellular systems lie between 5 and 12 [22]. Measurements have also been made for device to device systems. Standard deviations found in measurements of device to device systems are presented in Table 2.

Environment	Shadow fading standard deviation, $\sigma_{S,dB}$ [dB]
Urban line of sight	2.9 – 7.6
Urban non-line of sight	4.2 – 10.3

Table 2 The standard deviation of the shadow fading component measured in device to device systems. [18], [19]

The standard deviation is an expression for the spread of the measured or simulated values. Large standard deviation characterizes systems which experience a large variability in performance over the specified area.

3.1.3 Fast fading

In urban environments and other places where a transmitted signal can take multiple different paths, the communications is prone to fast fading. When the receiver moves a distance shorter than a few wavelengths, it will experience vast variations in received signal power due to the interference between delayed signal components. When the phase of the signal components is out of sync, they will have a dampening effect on the resulting signal.

The signal envelope caused by fast fading is commonly modeled as a Rayleigh distributed variable with probability distribution function [23]

$$p_{Rayleigh}(x) = \begin{cases} \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{\sigma^2}\right) & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (7)$$

It is assumed that the phase of the individual signal components is uniformly distributed and has the same amplitude. This is a reasonable assumption in non-line of sight scenarios. However, in line of sight scenarios, the direct signal component is bound to be stronger than those reflected of e.g. a building or the ground. In the presence of a strong signal component, the Rician distribution more accurately models the signal envelope.

3.1.4 Combined fading

The common fading models log normal shadowing and Rayleigh fast fading are combined in the Suzuki distribution. However, when the standard deviation of the log normal fading is high ($\sigma_{S,dB} > 6dB$), the lognormal distribution approximates the Suzuki distribution [23]. It might seem that merging the two types of fading is a degradation of the model. However, in reality the fast fading is highly unpredictable because minute movements of either receiver or sender can drastically change the fast fading. Therefore, knowing the fast fading value is not useful in practical systems – in contrast to knowing the shadow fading and distance dependent path loss, which is used in e.g. power control schemes. Consequently, in design of wireless systems it is common to account for the fast fading by increasing the link margin – the difference between minimum expected received power and the receiver sensitivity.

On this ground, the fading terms are combined, such that the overall expression for path loss can be simplified to

$$PL_{dB} = PL_{determin,dB}(d) + PL_{fading,dB} \quad (8)$$

3.2 Correlation models

The path loss model described so far relates to a single communication link in a single instance. MANETs consists of multiple links dispersed in a common environment and the links are typically considered over a period of time. In a realistic scenario, two links sharing similar propagation paths will experience similar fading. Also, the path loss on the same link at two consecutive time steps, are bound to relate to each other. This relationship, be it over time or space, is called correlation. The correlation is a measure between 1 and -1, which describe the similarity of two stochastic variables/processes. Here, the correlation describes the similarity in path loss on two links (spatial correlation) or on the same link at different times (temporal correlation). A correlation coefficient of 1 would mean that the path loss is the same on the

links, and conversely a correlation coefficient of 0 means that the path loss values are independent of each other. Thus, there is a need to specify the temporal and spatial link correlation. Omitting correlation would imply independence of the path loss values, which would lead to unrealistic simulations and therefore questionable results.

3.2.1 Temporal correlation

The fading on a link will change over time, as the environment is dynamic, and the transceivers are assumed to move around. A dynamic environment is likely to have no influence on the shadow fading, and therefore little impact on the total fading. The crucial influence on the temporal correlation of a link is the movement of the transceivers. As the transceivers move, the environment of the link changes, thus the path loss develops.

Consider the stochastic process, $\{X_{PL}\}$, describing the path loss on a certain link in discrete time steps, i.e. $X_{PL}(0)$ is the path loss at the link at time 0. Because the process describes the path loss over time, the autocorrelation function of the process describes the temporal correlation of the process. The autocorrelation function, $R(k)$, is defined as (assuming wide sense stationary)

$$R(k) = \frac{E((X(t) - \mu)(X(t+k) - \mu))}{\sigma^2} \quad (9)$$

Where μ and σ^2 is the mean and variance of $\{X_{PL}\}$. The path loss of a given link can be made stationary by subtracting the distance dependent part. Leaving only the fading term, which has constant mean and variance.

Gudmundson modeled the temporal correlation of the shadow fading on a link between a static base station and a mobile terminal moving at a constant speed in [24]. He proposed a simple decreasing autocorrelation function:

$$R(k) = \sigma^2 \phi^{|k|} \quad (10)$$

$$\phi = \varepsilon_D^{\frac{vT}{D}} \quad (11)$$

Where σ^2 is the variance of the shadow fading, ε is the correlation at distance D , v is the speed of the receiver (the transmitter is assumed stationary) and T is the sampling interval. $\varepsilon \leq 1$ as the correlation cannot increase indefinitely. The exponent is unitless and increases with the speed and sampling interval, which makes sense as these would increase the physical distance between measurements and therefore lower the correlation. In most literature the utilized distance is the decorrelation distance, defined as the distance where the correlation drops to 0.5.

To impose correlation on a stochastic shadow fading process (white Gaussian noise), $w(t)$, one can simply apply a first order autoregressive filter to the process [25]. The autoregressive filter is defined as

$$y(t) = \phi y(t-1) + w(t) \quad (12)$$

Which has autocorrelation function (see derivation in Appendix B)

$$R_y(k) = \frac{\sigma_w^2}{1 - \phi^2} \phi^{|k|} \quad (13)$$

I.e. to recreate the autocorrelation function of the Gudmundson model, one must choose the variance of the white Gaussian process as $\sigma_w^2 = (1 - \phi^2)\sigma^2$.

While the Gudmundson model is widely popular, it is not suited for MANETs, where both transceivers are mobile. An extension of the Gudmundson model was made by Wang et al. [26] to deal with this issue. They specify the autocorrelation function as a function of the distance traveled by transmitter and receiver independently:

$$R(d_t, d_r) = e^{-\frac{|d_t| + |d_r|}{d_{corr}} \ln(2)} \quad (14)$$

The constant d_{corr} is the decorrelation distance. Wang et al. investigated the decorrelation distance for handheld devices, by analyzing a dataset collected in Bristol, England. They describe the city as having 3 story buildings with an average height of 12 meters and an average road width of 20 meters. A decorrelation distance of **20m** was found. Interestingly, the decorrelation distance was constant for a variety of scenarios including increasing the height of one device, and at different transmission frequencies.

3.2.2 Spatial correlation

In multi-link networks the spatial correlation is important, as it describes the correlation in path loss on different links. The inter link correlation in MANETs is intuitive. When two links share a propagation path down a street or around a building, the shadow fading effect on the links are bound to be similar. Often in research on MANETs, the spatial correlation is overlooked, see e.g. [27]. The lack of spatial correlation leads to optimistic conclusions [28]. However, few researchers have put effort into creating a realistic spatial correlation model. The spatial correlation can be considered as the correlation in path loss on two links. Let $\{X\}$ be the stochastic variable describing the path loss on one link and $\{Y\}$ be the stochastic variable describing the path loss on another link. The spatial correlation of the links is defined as

$$\rho_{X,Y} = \frac{E((X - \mu_X)(Y - \mu_Y))}{\sigma_X \sigma_Y} \quad (15)$$

This leaves the question of what constitutes a link. Naturally, a link is specified by its endpoints, and the endpoints are associated with a given position. Unfortunately, this leaves an infinite or at least infeasibly large amount of different links if endpoints can be placed arbitrarily in a given region. Therefore, there is a need for specifying the correlation coefficient of two links, based on something other than absolute positioning of endpoints.

Agrawal et al. [28] formulates a spatial correlation model and compares it to the Gudmundson model. They translate the Gudmundson model for use as spatial correlation by using the following observation. Consider a link between node 1 and node 2. Node 1 remains stationary while node 2 moves at constant velocity for s seconds. Using the autocorrelation model specified by Gudmundson, the correlation coefficient of the link at time 0 and the same link at time s , can be found. Realizing that the same link at different time lags correspond to two links at different spatial positions, Agrawal et al. has now obtained the spatial correlation of the two links. This approach is only valid for determining the correlation of links which share a common node.

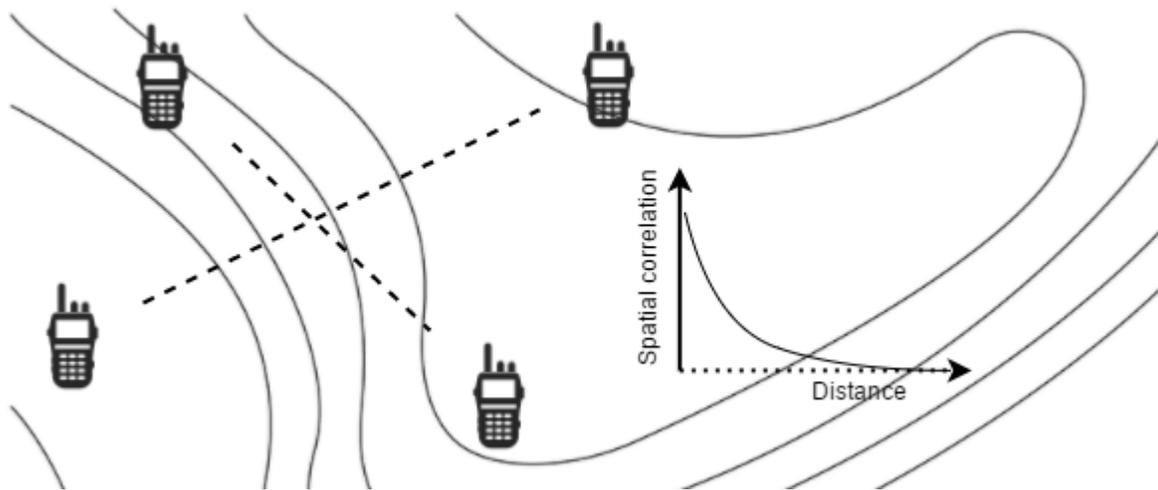


Figure 5 Contour plot of the spatial field. Correlation between two links can be found by integrating the correlation of the spatial field along the two links.

Agrawal et al. propose a model based on a spatial field. The spatial field specifies the underlying loss conditions as a wide sense stationary white Gaussian field with exponentially decaying spatial correlation. A contour plot of a spatial field is shown in Figure 5, where 2 links are illustrated. By integrating along the lines of the two links, the correlation is found (see [28] for details). The model is compared to the Gudmundson model and validated against real life measurements. Measurements were taken by placing 16 nodes in a 4×4 grid with 1 meter spacing. To construct different fading scenarios, they placed cardboard boxes within the grid, based on a Poisson spatial stochastic process. Based on the measurements, they specified the

link correlation in 28 different link geometries and compared the measured correlation to their own model and the Gudmundson model.

While the Agrawal model is able to capture spatial correlation more accurately than the Gudmundson model, the drawback of the Agrawal model is the specification of the spatial field as wide sense stationary Gaussian. Intuitively, in an urban environment the shadowing obstacles will not be Poisson distributed and therefore the assumed exponential spatial correlation might not hold. As illustrated in Figure 6, the correlation between the links drawn with stiped and solid lines is vastly dependent on whether the links are propagating in an urban canyon or blocked by buildings.

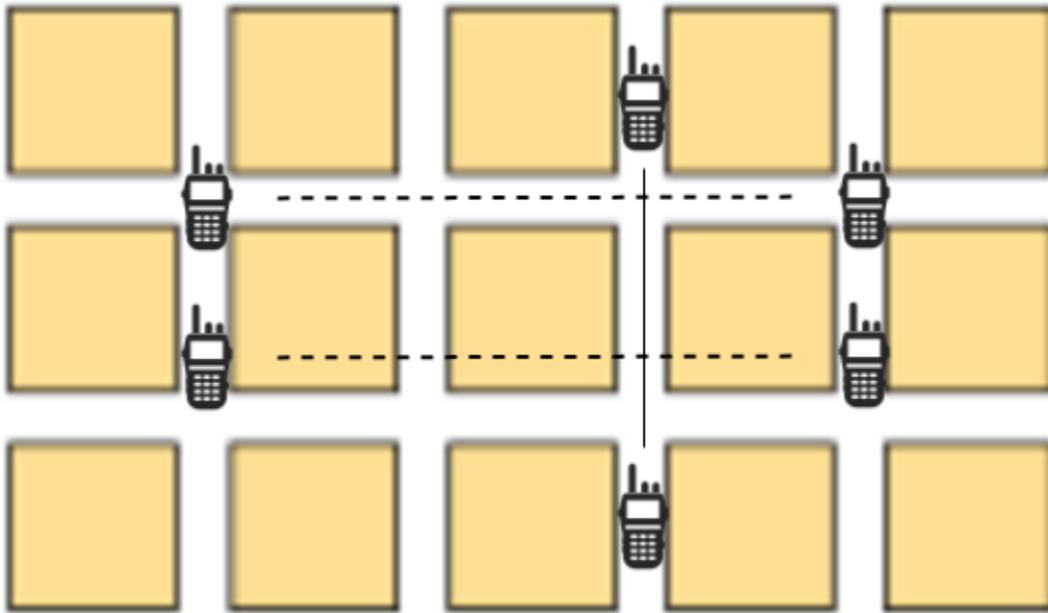


Figure 6 Links between nodes in a Manhattan grid can experience very different fading despite having overlapping propagation path.

Instead, consider the straight line between transmitter and receiver. A simple measure of the similarity between two links is the angular separation between two links sharing a common node. Without knowing the specifics of the environment, it is safe to assume that links with low angle separation will share paths and therefore experience correlated fading. The same observation has been used for determining the spatial correlation in cellular networks, where the common node is the mobile station and the two endpoints are base stations [29].

As with the Gudmundson model, the angle dependent spatial correlation model works only for links with a common node. However, the placement of the common node is important for the angle separation. Additionally, the angle separation relates to the distance between endpoints, but unlike the Gudmundson model, the correlation on parallel links (zero angle) will be equal despite the length of the links. Considering an urban environment this feature seems

possible, as the correlation on links between 3 nodes moving down the same straight road, is likely to stay constant. Figure 7 illustrates the concept of the angle dependent spatial correlation function. In this model, the correlation is determined based on how close the links are to each other. Here, closeness is described by the difference in angle between the two links. The structure of the correlation function is left open for the moment being. Some model it as a piecewise function where the correlation is linear at shallow angles and constant when the angle surpasses a threshold [30]. However, the correlation function seems to depend on the environment. Determination of the structure is left for section 4.2, where data analysis is performed on measurements from the Philippines.

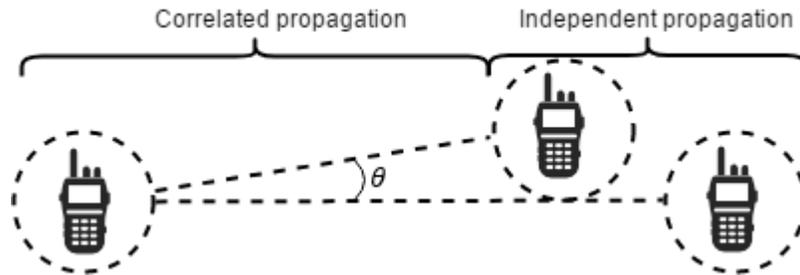


Figure 7 Inter link correlation is caused by the shadow fading, which is correlated for links that are close to each other. Correlation occur along the propagation path, but also in the clutter at the transceiver, which is common to both links.

4 | Empirical modelling

In any type of modelling, the purpose is to mimic some real-life phenomenon. In link modelling the acknowledged approach is to base the overall structure of the model on theory like the exponential path loss over distance. However, the specifics are best determined by studying the actual (or a representative) network environment. This approach has led to standard models like the COST-231 [31], which has a broad range of utilization in cellular networks. Propagation channel models for device to device communication is still an active research field. Standards organization are in the progress of performing comprehensive measurements to characterize the D2D propagation conditions in a wide variety of environments. One example is the US National Institute of Standards and Technology (NIST), whose public safety communications research division are in the progress of performing extensive measurements [32].

In the absence of propagation models suitable for the scenario of relief work in the Philippines, it is natural to base the model on actual measurements from the environment. Luckily, as part of the development of the relief work communication devices, measurements were taken on site in the Philippines. While this gives a valuable insight to the conditions in the Philippines, the measurements should be used with care, as the measurements campaign was relatively short, and was not repeated.

This chapter presents how the measurements were carried out. Secondly, the measurements are processed and the regression of the model to fit the specific scenario is presented.

4.1 Measurement campaign

The purpose of the measurements was to establish data, which can characterize the radio environment in the Philippines. The equipment used for the measurements were prototype Reachi devices containing the radios, which will be used in the final product. Additionally, the devices were equipped to PRC volunteers. Tailoring the measurement setup to the specific scenario is an easy way of ensuring that the model is applicable to the system. That is, it eliminates the need to account for the orientation of the device and the position of the operator relative to the device, when using the model for evaluation of the system, as these effects are naturally incorporated in the measurements and therefore the model.

The downside to including the effects of device orientation and possible operator shadowing is, that the variability of the measurements is bound to increase. Another limitation of using the production radios relates to how measurements are registered. The medium access control (MAC) protocol of the radios are responsible for reporting the signal strength measurements. However, the protocol is only capable of reporting the measurements of signals received from

neighboring nodes. This is an issue, because there is a limit to the possible number of neighbors. Furthermore, the exact conditions for being registered as a neighbor are proprietary and kept secret by the producer of the radios, NeoCortec.

The following subsections will explain how and what data was gathered.

4.1.1 Equipment

33 prototype devices equipped with production radios were used in the campaign (see Figure 8). Each device contains 3 functions relevant for the data collection.

1. The radio module is measuring the received signal strength from other devices.
2. The GPS module provides time synchronization and positioning for the device.
3. The data storage saves collected data for post processing.



Figure 8 The Reachi device is made of high quality plastic. The antenna is placed in the trapezium extending from the device (on the right in the figure). Picture borrowed from [33].

The radio module is the Neocortec nc1000lr module. The characteristics of the module are outlined in Table 3.

NEOCORTEC nc1000lr radio characteristics [34]		
Parameter	Value	Comment
Receiver sensitivity	-99 dBm	At PER 1%
Receiver saturation	-15 dBm	
Transmission power	26 dBm	
Frequency band	US 915 MHz ISM	
Channel spacing	333.252 KHz	
Antenna	Pulse Electronics W5017	Only antenna for which the device is FCC approved
Antenna gain	2 dBi	
RSSI variation	$\pm 2.5\text{ dB}$	Based on production test

Table 3 Characteristics of the radio module used by the Reachi devices. [34]

There is no available data on the GPS-module, nor is it specified how many satellites the GPS-module connected to. The performance of a GPS is specified in terms of its accuracy and precision. The accuracy specifies the expected error between the actual position of the GPS and the position estimate provided by the GPS module. The precision describes the spread of position estimates measured at the same position. See Figure 9 for a graphical interpretation of accuracy and precision.

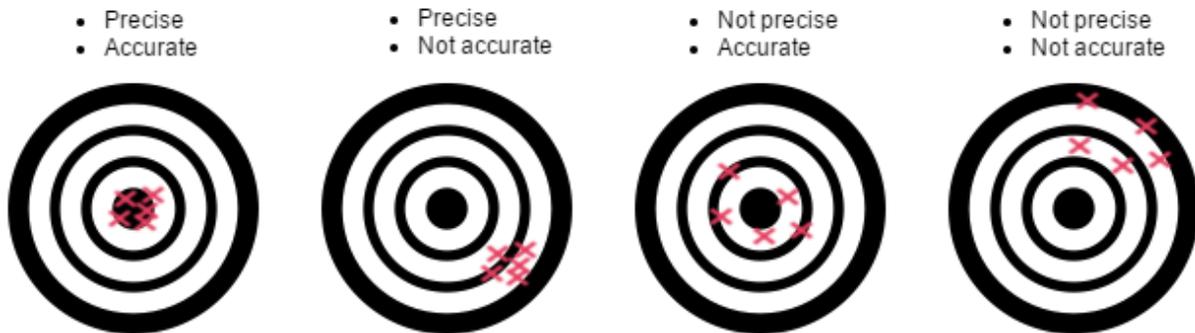


Figure 9 The concept of accuracy and precision illustrated. The true position of the GPS module is on the bullseye. The red crosses represent position estimates made by the GPS module.

It is not possible to determine the accuracy of the GPS module used in the measurements, because the true positions of the devices are unknown. It is a common claim that GPS is accurate to 5 meters. In a study with 1000 participants, the mean error was measured to 4.9 m in open areas and 16.8 m in cities [35]. No comments were made on whether the direction of the errors was consistent. However, it was noted that the height of surrounding building was highly correlated with the GPS accuracy, as explained in their ‘*law of urban multipath*’ [35]:

$$Accuracy [m] = BuildingHeight_{[floors]} + 5 \tag{16}$$

The precision of the GPS module can be estimated from the measured positions. In the measurements, one node is approximately stationary for 9 minutes outside a public library. Assuming that the person is stationary, due to a conversation with an acquaintance or simply taking a rest, the data series can be used to estimate the precision of the GPS. A measure describing the GPS precision is the circular error probability (CEP) [36]. The CEP is the radius of a circle within which the position estimate falls with 50% probability. For the aforementioned 9 minutes period of measurements, the CEP is 2.85m.

Time synchronization of devices are carried out via GPS. The simplest GPS time synchronization technique is *One way GPS time transfer*, which provides an uncertainty less than 20ns [37].

4.1.2 Procedure

33 PRC volunteers were equipped with a prototype Reachi device. The data collection was initiated. The volunteers walked different paths in the local area. Every path started and ended at the same spot marked with green in Figure 10. The routes of the volunteers were arbitrary. The full range of recorded GPS-locations are illustrated in Figure 10 and covers a rectangular area of 1 by 1.5 km. The duration of the data collection was approximately 1 hour.

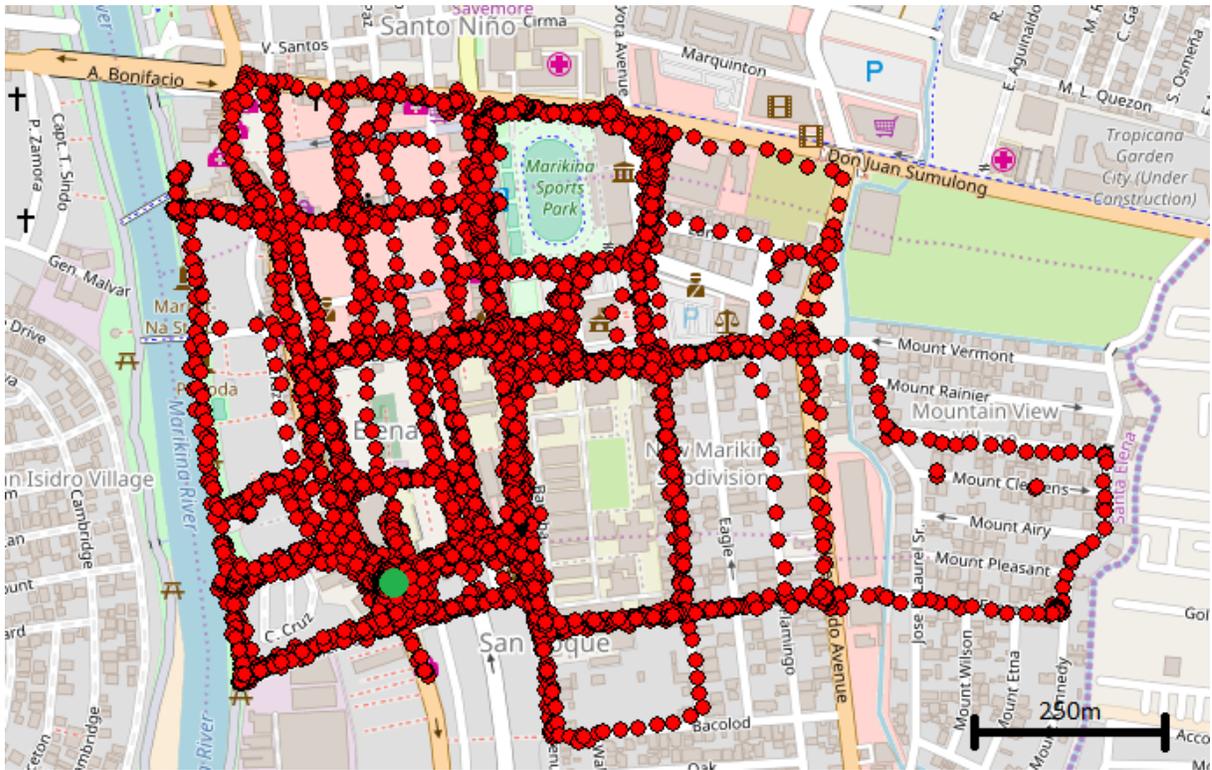


Figure 10 All GPS-positions recorded during measurements in the Philippines. The green spot marks the starting and ending position.

4.1.3 Data

Measurements were logged in 20 second intervals on the individual devices. Upon completion of the campaign, data was extracted from devices and put through a pre-processing. The function of the pre-processing is formatting and collecting the data in a manageable file. The format of the resulting comma separated (csv) file is given in Table 4. Appendix C explains in depth the formatting of each field.

Field 1	Field 2	Field 3	Field 4	Field 5		
Timestamp	ID	UID	Latitude	Longitude		
Field 6	Field 7	Field 8	...	Field 39	Field 40	Field 41
Neighbor 1 ID	Neighbor 1 RSSI	Neighbor 1 distance	...	Neighbor 12 ID	Neighbor 12 RSSI	Neighbor 12 distance

Table 4 Format of the comma separated data file after pre-processing.

The interesting parameter regarding link characterization is the received signal strength indicator (RSSI). The value is given in *dBm* by a 1-byte integer value, i.e. it can take on integer values in the range $[-128; 127]$. However, as the receiver has a saturation point of -15dBm and a sensitivity of -99dBm , this is the actual range of the reported RSSI values [34]. These two limitations pose a distortion to the measurement data, as data points outside the limits are bound to exist but will not be registered. Further, the characteristics of the measurement device means that the recorded RSSI value is associated with some inaccuracy, which becomes extra pronounced under -85 dBm . It should be expected that the inaccuracy is similarly pronounced close to the saturation point.

Table 4 show that a maximum of 12 neighbors are reported in the measurements. With 33 nodes taking part in the measurement, the devices need to decide which nodes are included as neighbors in the reported measurements. Consequently, the measurements are influenced by the protocol, which decides which nodes to include in the neighbor list.

4.2 Data analysis and model regression

For link modelling, the important data is the received signal strength indicator (RSSI) in combination with the coordinates of the endpoints of each link. Having conducted the measurements with identical radios and GPS' allow comparison of all links. Because of the pronounced inaccuracy of the RSSI close to the saturation and sensitivity levels, the measured data is filtered, so that data point outside the RSSI interval, $[-15, -90]$ are discarded. Additionally, the GPS data contains uncertainty. Even though the individual GPS accuracy is several meters, the link length depends on the relative distance between devices. As the factors which contributes to the GPS inaccuracy are largely environmental, GPS-modules in close vicinity will likely experience the same bias. Therefore, datapoints where links are less than 1 meter in length are discarded. The data points are plotted in Figure 12, where the contour overlay describing the density reveals that 60% of the data points are situated in the distance interval $[100\text{m}, 300\text{m}]$.

4.2.1 Pathloss exponent and standard deviation

In Section 3.1.1 it was argued how the distance dependent path loss is exponentially decreasing with the link distance. Finding the pathloss exponent, γ , can be done by linear regression on the data described by the equation

$$RSSI_{dBm} = -10\gamma \log_{10}(d) + c \quad (17)$$

In the regression of the measured data to (17), d is the independent variable and $RSSI_{dBm}$ the dependent variable. The constant offset, c , is the pathloss at 1m. As both variables are subject to error, the proper regression is the total least squares, also known as orthogonal or errors-in-variables regression. In contrast, least squares regression assumes error in the dependent variable only. Figure 11 illustrates the conceptual difference between total least squares and least squares regression.

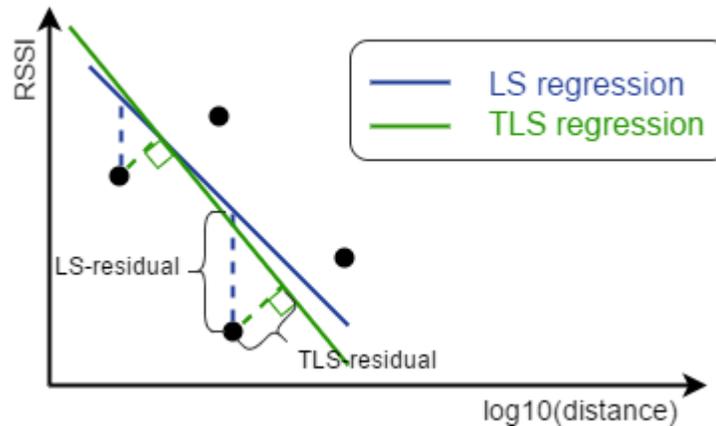


Figure 11 Comparison of the total least squares and the least squares regression methods. The dotted lines illustrated the errors, whose summed squares are subject to minimization.

It may seem straight forward to go with the TLS, because of the error sources in both variables. However, the choice is difficult for two reasons. Firstly, the accuracy of the RSSI and distance measurements are both numerically different, quantified differently and measured in different units. Secondly, the accuracy of the distance is measured in meters, but the independent variable in the regression problem is measured in $\log_{10}(\text{meters})$. Optimally, the variables should be converted into unitless variables by some normalization process. This could be done by normalizing the variables to their respective accuracies. Nevertheless, we accept the slight drawbacks of TLS regression, because the exact accuracies are unknown.

The two regression methods yield significantly different results. Least squares regression results in a pathloss exponent of $\gamma = 3.2$ and offset of $c = 0.4$, while total least squares regression yields $\gamma = 5.5$, $c = 44.8$. The results of the two regression methods are hugely different. One could be expected in an urban line of sight scenario (LOS), while the other would be likely in

a non-line of sight (NLOS) scenario. The actual scenario is likely a mixture of LOS and NLOS. The offsets, c , are equivalently different, and notably both unrealistic. The offset value specifies the signal strength of a 1m link. Obviously, the signal strength of the transmitted signal must be lower than the transmitted signal strength of **26dBm**. Using free space path loss from Equation (2), a path loss of **32.44dB** is expected at **1m**, resulting in an expected offset at **1m** of $c = 26\text{dBm} - 32.44\text{dB} = -6\text{dBm}$.

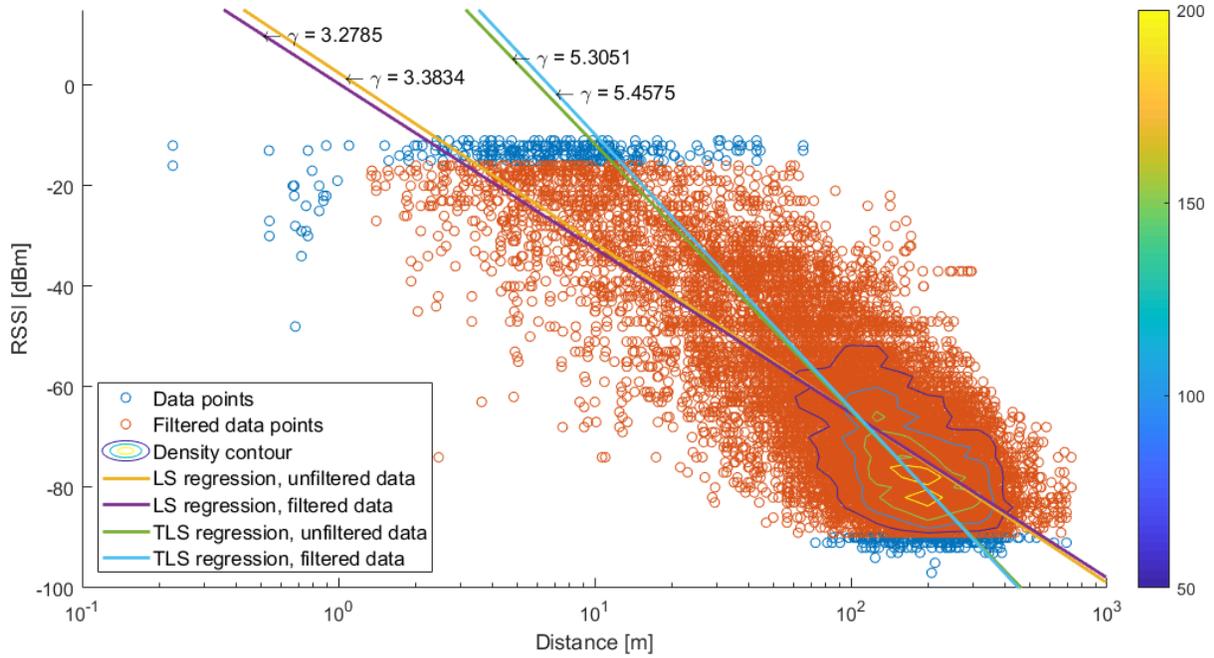


Figure 12 RSSI versus link length measured in the Philippines. A contour plot is overlaid the datapoints to highlight the variation in density of the datapoints. Trendlines illustrate the path loss exponent found with LS and TLS regression

Once the pathloss exponent is determined, the distance dependent path loss can be subtracted from the measurements. Graphically this is equivalent to computing the LS-residual as illustrated in Figure 11. The overall shadow fading standard deviation can be found, by computing the sample variance of the residuals. The resulting standard deviations are in the case of LS and TLS respectively: $\sigma_{LS} = 10.9\text{dB}$, $\sigma_{TLS} = 14.2\text{dB}$. Unsurprisingly, the TLS standard deviation is larger, which is intuitive as LS minimizes the residuals subject to the variance measure. However, both values exceed what has previously been observed by other, as stated in Table 2. The large values have to be attested to the measurement procedure, where neither the device orientation, nor the placement of device relative to operator were constant factors.

4.2.2 Correlation parameters

Temporal and spatial correlation needs to be specified. The measured data is unsuited for analysis of the temporal correlation. The sampling frequency at $1/20\text{Hz}$ prevents calculation of local (within tens of wavelengths) means – which would filter the effects of fast fading. Additionally, the proper measurements series to determine decorrelation distance would be keeping one device stationary, while the other moves away. This is not available from the measurement data. Therefore, the decorrelation distance of 20m measured by Wang et al. [26] is adopted in the present model.

The spatial correlation between links is determined based on the angle between links. To determine the angular correlation, link pairs were separated into bins based on the angle between them. This excludes link pairs without a common node, but because up to 12 links may attach to one node and every link attach 2 nodes, the possible combination of link pairs in the dataset sharing a common node extends above $40,000$. In each bin the links that constitute a link pair are regarded as realizations of two random variables, X and Y . The assignment of a link to one of the variables is randomized, i.e. the same link may be represented in both X and Y , though as part of different link pairs. This randomization is important as it will affect the mean and variance of the random variables which in turn affect the correlation. The correlation coefficient of X and Y is determined as

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X\sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X\sigma_Y} \quad (18)$$

Where σ and μ signifies the sample standard deviation and sample mean respectively.

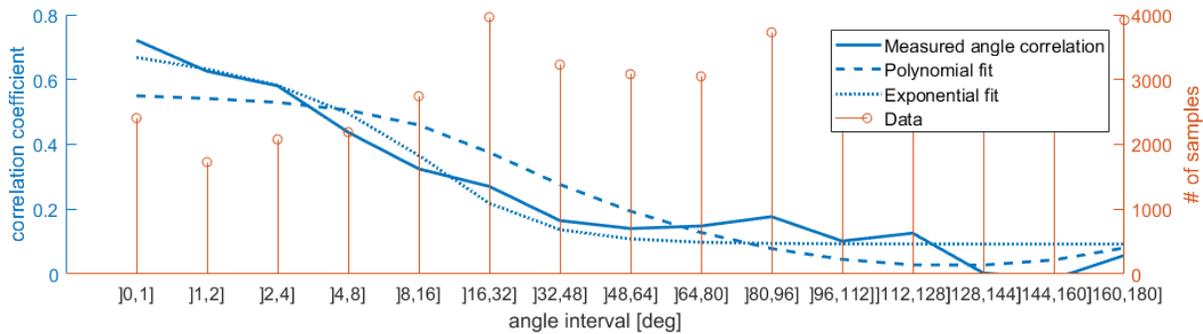


Figure 13 Correlation and number of samples in each bin. Bin edges are given on the first axis.

The nature of the dataset places on overweight of samples on the angles lower than 10° . In order to equalize the number of samples in each bin, the bin edges have been adjusted, as illustrated on the first axis in Figure 13. Polynomial and exponential least squares regression to the calculated correlation coefficient was performed. The use of least squares is justified by the binning of angular values, which minimizes the inaccuracy of the independent variable.

The root mean square (rms) of the polynomial and exponential regressions were $rms_{poly} = 0.09$ and $rms_{exp} = 0.05$ respectively. Besides a lower rms, the exponential fit better captures the high correlation at lower angles, as illustrated in Figure 13. Therefore, the spatial correlation between link i and link j are described in terms of the angular separation as

$$r_{i,j}(\theta) = 0.595e^{-0.064\theta} + 0.092 \quad (19)$$

5 | Simulation

Implementation of the model is the next step to enable simulation of the link conditions in the Philippines. This chapter will present the methods used to create simulations with the multi-link model. Additionally, the computational complexity involved is discussed.

5.1 Simulation description

Let a network consist of N nodes. The link matrix, \mathbf{L} , with entries,

$$l'_{i,j} = l'_{j,i} \mid i, j = \{0, 1, \dots, N\} \quad (20)$$

describes the path loss in dB on the link between node i and node j . Clearly, links are undirected and exhibits reciprocity – equal loss in both directions. The diagonal, $l'_{i,i} = 0$, is zero as there is not link from a node to itself. The link matrix is quadratic and has size $N \times N$.

To describe the path loss on each link and the correlation between links, it is necessary to identify the unique links in the link matrix. Let $\mathbf{l} = [l'_{1,1} \ l'_{1,3} \ \dots \ l'_{1,N} \ l'_{2,3} \ l'_{2,4} \ \dots \ l'_{2,N} \ \dots \ l'_{N-1,N}]^T$ then \mathbf{l} contains all unique links with a nonzero pathloss, which is the elements of the upper triangle of the link matrix excluding the diagonal. It is worth noting that the length of \mathbf{l} :

$$length(\mathbf{l}) = \sum_{i=1}^{k=N-1} i = \frac{N(N+1)}{2} \quad (21)$$

Now, from the single link model description, Equation (8), in Section 3.1.4, the path loss of each link can be described as a sum of a distance dependent part and a fading part.

$$\mathbf{l} = \mathbf{l}_d + \mathbf{l}_{fading} \quad (22)$$

The distance dependent path loss, \mathbf{l}_d , is calculated as (17) using the values obtained by the TLS regression:

$$l_{d,i,j} = -55 \log_{10}(d_{i,j}) + 44.8 \quad (23)$$

where $d_{i,j}$ is the distance between node i and j .

The fading term, \mathbf{l}_{fading} , is where the correlation comes into play. The vector of fading components is simply a zero mean multivariate Gaussian vector, i.e. $\mathbf{l}_{fading} \sim N(\mathbf{0}, \mathbf{\Sigma})$ [38]. First, let \mathbf{C} be the correlation matrix determining the correlation coefficient between links as

$$\mathbf{C} = \begin{cases} r_{i,j}(\theta_{i,j}) & \text{if } i \neq j \text{ and } \text{commonNode}(i,j) \\ 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (24)$$

where $\theta_{i,j}$ is the angle between link i and j , and $\mathit{commonNode}(i,j)$ is a binary function which returns true if link i and j share a common node and false otherwise. The covariance matrix, $\mathbf{\Sigma}$, is related to the correlation matrix by $\mathbf{\Sigma} = \sigma^2 \mathbf{C}$. A standard deviation of $\sigma = 11.4\text{dB}$ is used.

In practice, generation of correlated multivariate Gaussian variables is done in two steps. Firstly, an independent and identically distributed (iid) multivariate Gaussian vector is drawn, $\mathbf{x} = \left[x_1 \ x_2 \ \dots \ x_{\frac{N(N+1)}{2}} \right]^T \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$. Secondly, the independent variables are made dependent by multiplication with the lower triangular Cholesky decomposition of the correlation matrix, $\mathbf{Q} = \mathit{cholesky}(\mathbf{\Sigma})$.

$$\mathbf{l}_{\text{fading}} = \mathbf{Q}\mathbf{x} \quad (25)$$

Now it is possible to generate a single realization of the link matrix, which accounts for the distance dependent path loss and the spatial correlation. Next, the simulation is expanded to include the temporal aspect by implementing the temporal correlation. The temporal index is specified in parenthesis, i.e. $\mathbf{l}_{\text{fading}}(t)$ is the vector describing the fading path loss at time t . The distance dependent path loss is time invariant and needs no modification. The fading incurs temporal correlation as described in (14) with a decorrelation distance of 20m. The path loss at time $t + \Delta t$ can be computed based on the temporal correlation coefficient, $\rho_{\Delta t} = r(d_t, d_r)$, describing the correlation after transmitter and receiver moved distances d_t and d_r , respectively.

$$\mathbf{l}_{\text{fading}}(t + \Delta t) = \overbrace{\mathbf{Q}(t + \Delta t)\mathbf{x}(t + \Delta t)}^{\text{spatial correlation}} \overbrace{\sqrt{1 - \rho_{\Delta t}} + \mathbf{l}_{\text{fading}}(t)\rho_{\Delta t}}^{\text{temporal correlation}} \quad (26)$$

5.1.1 Computational aspects/other observations

Purposely, the node placement and mobility has been left open. The mobility model is of course important for the performance of the network but bears no importance for the implementation of the path loss model. If link distances, link angles and node movements can be extracted, the link model functions.

However, the particular model bears a practical limitation, which has not yet been addressed. In the creation of correlated Gaussian random variables, Cholesky factorization of the covariance matrix is performed. The computational complexity of Cholesky factorization is $O(k^3)$, where k is the size of the input matrix. The size of the covariance matrix is determined by the number of nodes in the network and equal to the length of the unique link vector \mathbf{l} . Therefore, in terms of the number of nodes in the network, N , the computational complexity of the Cholesky factorization is $O((N^2)^3) = O(N^6)$. Obviously, the computational complexity becomes an inhibitor for mid to large scale simulations. Thus, for larger simulation, it might be feasible to consider only parts of network, in order to reduce the number of considered links.

5.2 Model verification and validation

Before applying the model to simulations, it is important to verify that the model adheres to the specifications and validate that the model accurately depicts reality. Naturally, this requires a specification of what is meant by reality and accuracy.

5.2.1 Verification

The purpose of the verification is confirming that the implementation of the model matches the specification of the model. For the path loss model, three specifications should be verified:

1. The distance dependent path loss must follow a power law with path loss exponent as specified in (23)
2. The standard deviation of the shadow fading must be preserved at around 11.4 dB
3. The temporal correlation must obey the correlation as specified by the extended Gudmundson model, Equation (14)
4. The spatial angle correlation must obey the correlation specified by the exponential in Equation (19)

The verification is performed by implementing the model in MatLab and examining the properties of the output at given inputs. The inputs for the model are link distances and angles between links. Instead of generating the distances and angles independently, they are derived from realizations of node positions. Therefore, the input for the verification script is the placements of several nodes in a two-dimensional coordinate system. Specific placement and subsequently node mobility has little relevance in this context, as these would determine the input to the model, not the function. Therefore, node placement is uniform within a 2000x2000m area and the nodes are all moving in the same direction at the same speed. Thereby, the angles between nodes and the temporal correlation are held constant, which makes it possible to interpret the fading variables as multivariate with static correlation.

Because the distance dependent and fading pathloss are separate, they are verified independently. The distance dependent path loss does follow (23). The verification points 2, 3 and 4 relate to the fading path loss. The standard deviation of the fading path loss is verified by calculating the sample standard deviation across all realizations of the fading path loss. As the number of realizations grow, the standard deviation converges to the specified 11.4dB. By calculating the correlation coefficients of the generated fading variables and comparing with the correlation coefficients used to generate angular and temporal correlation respectively, it is easily verified that the rms-error between the generated coefficients and the calculated coefficients approaches zero as the number of realizations increase. Thus, it has been verified that the implemented model follows the specification.

5.2.2 Validation

The purpose of validation is confirming that the model represents the system, which is being modeled. Since the model is designed based on real life measurements from the Philippines and literature which is also based on real life measurements, it is obvious that the specifications from the verification are consistent with reality. However, aspects not covered by the specifications might impact the realism of the model.

The ideal way to validate a model is comparing it to real data. The challenge of this approach is obtaining suitable data. Specifically, the challenges for validating the model against the measurements from the Philippines are (1) the statistical properties of the environment, (2) protocol influence on the measurements and (3) lack of link reciprocity.

(1) The standard deviation of the fading path loss is independent of the link length. This might be an issue. Consider devices separated a few meters apart. Shadowing on short links in an urban environment is likely to be caused by people, trees and other objects of similar size. Reflections of the signal reduce the effect of objects shadowing the line of sight component. Therefore, it is often argued that the standard deviation of the fading is lower at short links. Another consideration is the long links ($>500\text{m}$). In urban environments, where links are established between hand held devices (operation height of 1.5-2m) street layout and open space between buildings are determining factors for the possibility of a line of sight path between transceivers. Because of the twisted streets and closely spaced buildings, the probability of line of sight paths is near zero on long links. Additionally, the signal path on long links in an environment such as the one in the Philippines is likely to go around several corners and past numerous obstacles. Combined, these factors deteriorate the signal strength severely on longer links. This effect can be observed, upon analysis of the data from the Philippines. Although devices are separated more than 500m for a majority of the measurement period, few links are established on this distance.



Figure 14 Street view of a typical street in Manilla, Philippines.

(2) The protocol has a large impact on the measurements performed in the Philippines. As noted previously, a maximum of 12 neighbors are registered. Additionally, the update process of the neighbor list is unknown, but it is reasonable to believe that neighbors are weighted based on the link quality. This protocol behavior complicates the interpretation of the measured data, as less optimal, though still functional, links are absent in the measurements.

(3) In the measured data, links are not guaranteed to be equal in both directions. This doubles the possible datapoints as each link can be associated with two RSSI measurements. In the data set, 1 040 links have RSSI reported twice of which 97% are shorter than 300m

In order to compare the model to reality, the node placements from the Philippines measurements are used in simulation. Figure 15 shows a contour plot of the datapoints, illustrating the positions of the majority of datapoints. Both the measured and the simulated data has had links with an RSSI of less than -90dBm removed. Even though both datasets are based on the same node positions, many more links are established in the simulation $\sim 35\ 000$ against the 16 600 links recorded in the measurements. Unsurprisingly, the majority of the additional established links are longer than 200m and have a clear bias towards low RSSI.

In Figure 15, contour lines of the measured datapoints, and the simulated datapoints are drawn in solid lines. Figure 16 shows a similar plot, where the simulated datapoints have been filtered. The y-axis on both figures indicates the RSSI values and the x-axis the distance. Below the contour plot, a line plot shows the distribution of datapoints relative to the distance. To the right of the contour plot a line plot shows the distribution of datapoints relative to RSSI-value.

The contour lines of the measured and simulated datapoints are not aligned. The difference of the peak positions and number of measurements fall well in line with the expected effects of (1), (2) and (3). However, for accurate validation the simulated data should be “filtered” through the same process, as the measured data. NeoCortec has agreed to provide a piece of software, which can perform the filtering, but it was not available in time for use in this project. The next best thing then, is to implement a filter that tries to imitate the NeoCortec filter.

A simple filtering was performed on the simulated datapoints. As explained in pseudo code in Algorithm 1, the measurements from each node is reduced to a maximum of 12 neighbors. Resulting contour of the data is plotted in green lines in Figure 16. The contours of the filtered datapoints fall between the measured and simulated datapoints. This attest to the fact, that the difference is likely to be caused by the neighbor filtering performed in the NeoCortec protocol. Further comparison of the model and the measurements is left for future work.

Algorithm 1 Neighbor filter

```

1: function NEIGHBORFILTER( $M$ )
2:    $F = \text{sort}(M, 'descending')$ 
3:   return  $F[1 - 12][1 - 2]$  ▷ Return the first 12 rows of F
4: end function

```

$$M = \left. \begin{bmatrix} ID_1 & RSSI_1 \\ ID_2 & RSSI_2 \\ \vdots & \vdots \\ ID_n & RSSI_n \end{bmatrix} \right\} \begin{array}{l} \text{All measurements} \\ \text{in arbitrary order} \end{array}
\qquad
F = \left. \begin{bmatrix} ID_1 & RSSI_1 \\ ID_2 & RSSI_2 \\ \vdots & \vdots \\ ID_{12} & RSSI_{12} \end{bmatrix} \right\} \begin{array}{l} \text{12 largest values} \\ \text{of } M \end{array}$$

Algorithm 1 The neighbor filter tries to imitate the behavior of the NeoCortec neighbor filter, by returning the 12 highest measured RSSI-values in the matrix F.

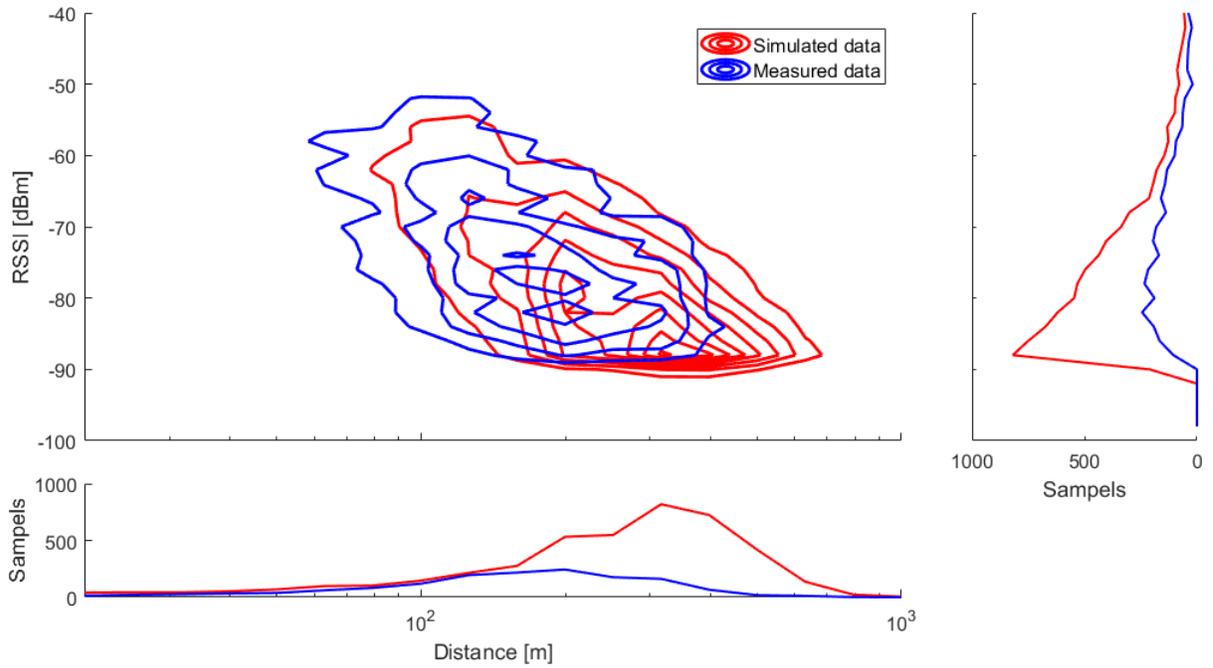


Figure 15 Contour plot of the measured data and the simulated data based on identical node positions.

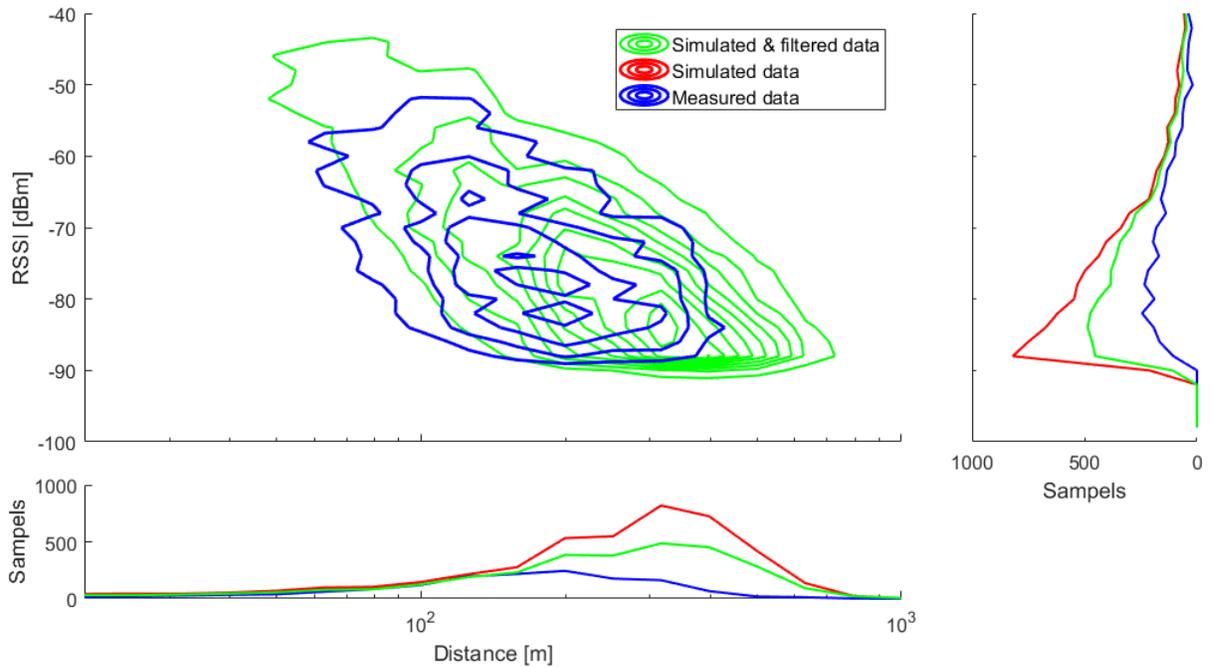


Figure 16 Contour plot of the measured data and the simulated & filtered data based on identical node positions. The raw simulated data is included for comparison.

6 | Performance evaluation

In this chapter, the performance of likely scenarios in a MANET is evaluated based on simulation with the proposed path loss model. The results are compared with a simpler version of the model, where the spatial correlation is disregarded. The performance of the simple scenarios will give insight to the expected performance of a full-size MANET.

The chapter is round off with a presentation of the layered network protocol stack and a brief discussion of data link layer and network protocol design considerations for MANETs given the results of the simulations.

6.1 Evaluation scenarios

It has been established, that angular correlation is one way to describe the spatial correlation in the data set from the Philippines. However, carrying out simulations which account for the correlations bears a computational penalty. The question that arise is, whether the penalty of including correlations can be justified by the advantages of having a more realistic model. Of course, there is no straight answer to this question, as it depends on the situation. Therefore, a comparison of the correlated and the equivalent non-correlated model is performed. Specifically, the difference between the models are in the correlation described by equation (19) in Chapter 3 |. In what is from this point referred to as the correlated model, abbreviated *corr*, the angular correlation of (19) is kept as $r_{i,j}(\theta) = 0.595e^{-0.064\theta} + 0.092$. In the non-correlated or the independent and identically distributed model, abbreviated *iid*, the angular correlation is set to zero, $r_{i,j}(\theta) = 0$.

Comparison of the models are performed as follows. Simple scenarios of how nodes might move in a MANET is described in the following sections. The descriptions are based on situations from (possible) relief work scenarios but can easily apply to other use cases like festivals, industrial IoT or military applications. Monte Carlo simulation with both models are performed on the simple scenarios, and performance metrics are extracted. This enables direct comparison of the models. Additionally, as the simple scenarios are all subsets of a full network, a rough assessment of a potential MANET can be made based on its composition of the simple scenarios. The method is reminiscent of the divide and conquer approach, used in certain algorithms. Here, it is more the case of divide and understand. In contrast, another approach to the comparison could be to specify the behavior of all nodes and run full-scale simulations. While this might show the same differences between the models, understanding the cause of the differences will be more complex in a full-scale simulation.

Performance of a network can be measured by multiple metrics, the most common being related to the network traffic: latency, throughput, jitter and error rate. Additionally, a network may be constrained by several factors, e.g. battery powered devices, interference, legal limits on maximum effective radiated power and other limits caused by the network conditions - environmental conditions, device limitations and mobility of the network devices. However, investigating the network performance of these traffic-related performance metrics require specification of the network protocols and traffic generation.

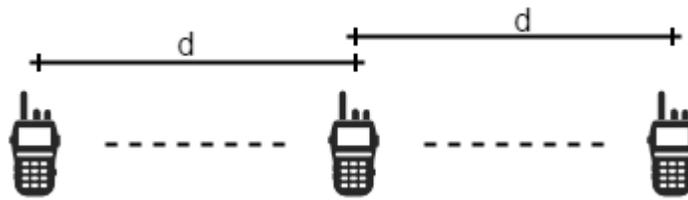
Therefore, the chosen evaluation metrics relate to the connections in the network. Important definitions and metrics used in the evaluation is explained below:

- **Link failure:** A link is said to be connected if the RSSI is above the threshold of -99dBm . Link failure is when a link is not connected.
- **Average max link length (AMLL):** The distance at which the distance dependent path loss on a link becomes -99dBm is the average max link length. This is because the fading path loss is zero mean. The AMLL in both path loss models is 432.6m .
- **Connection failure probability (CFP):** Two (sets of) nodes are connected if any path of connected links can be formed between the nodes. Thus, a connection failure signifies that no path of connected links can be formed between the two nodes(/sets). The probability of path failure is calculated on Monte Carlo simulations with a certain number of realizations. Thus, $CFP = \frac{\text{realizations with connection failure}}{\text{total number of realizations}}$.
- **Distance to first connection failure (DTFCF):** Assume that a connection between two (arbitrary) nodes in a network exists. Mobility in the network might lead to a connection failure between the two nodes. The distance moved by either node between establishment of the connection until connection failure is the distance to first connection failure.
- **Distance to half neighbor connectivity (DTHNC):** Assume that a node has a number of neighbors, n . Mobility of the network might lead to link failure between the node and some neighbors. The distance traveled by each neighboring node, when the number of neighbors has dropped to $n/2$, is the distance to half neighbor connectivity.

6.1.1 Static scenarios

The first set of scenarios exclude mobility and therefore look at the instantaneous conditions in the network. The scenarios are described individually with a figure demonstrating the scenario, a description and an explanation of the metric used to evaluate the performance in the scenario.

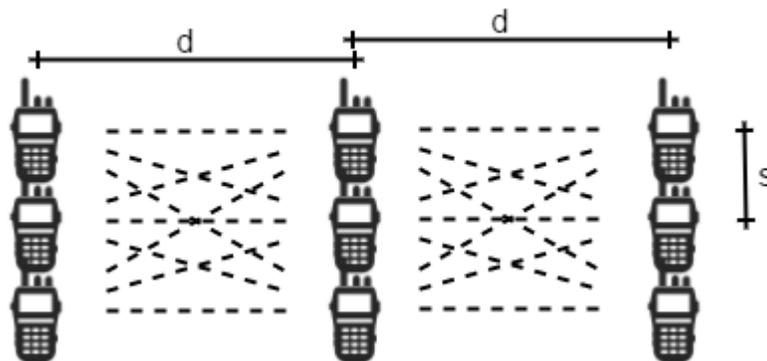
Scenario 1



Description 3 nodes are located on a straight line. The center node can act as a relay between the left- and rightmost nodes, possibly extending the range at which all nodes are able to communicate.

Metric The point of interest in this node constellation is whether the left and right endpoints can communicate (directly or through the relay). The connection failure probability of the connection between the left- and rightmost devices is determined at different link distances, d . d is specified in average max link lengths (AMLL).

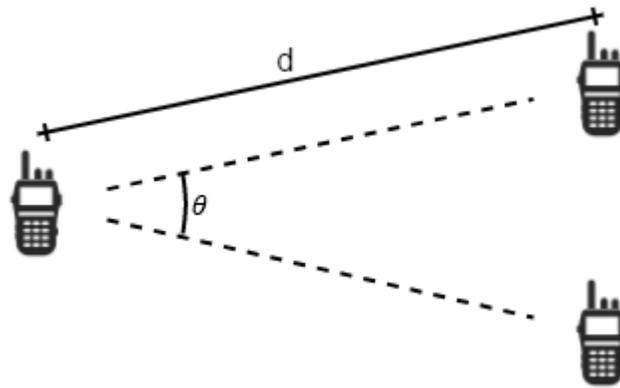
Scenario 2



Description In many scenarios it will be advantageous for relief workers to work together in groups. This increases their own safety and enhance their capabilities in terms of e.g. clearing a road for debris or searching through remains of a collapsed housing area. Therefore, in this scenario nodes are placed in clusters of 3. Within a cluster, nodes are placed on a vertical line with **1m** spacing between adjacent nodes, $s = 1\text{m}$.

Metric The point of interest in this node constellation is whether the left- and rightmost clusters can communicate (directly or through the middle cluster). The connection failure probability between endpoint clusters is determined at different link distances, d . d is specified in average max link lengths (AMLL).

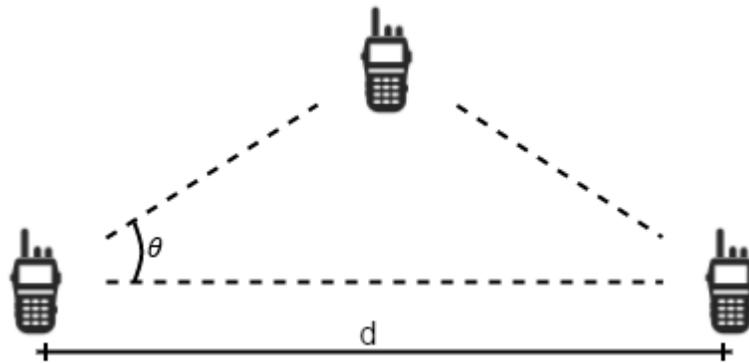
Scenario 3



Description A relief worker has wandered into a remote neighborhood. The nearest peers are located approximately in the same direction. The two peers on the right are the only gateways to the rest of the network. The length of both links to the remote node is $d = 432.6\text{m}$ - the average max link length.

Metric In this scenario, the connection of the remote node on the left to the set including both nodes on the right is investigated. The connection failure probability at various angles, θ , is calculated.

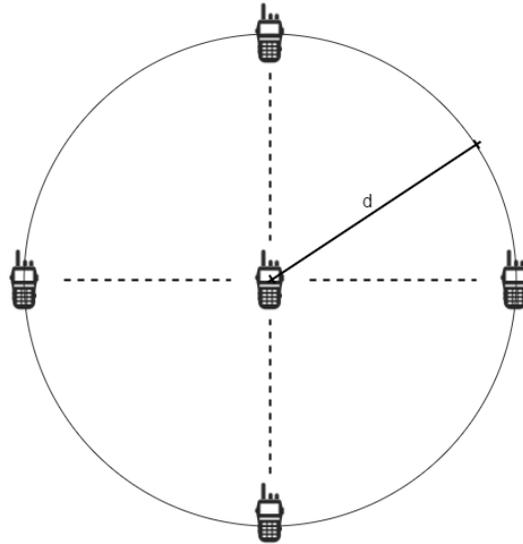
Scenario 4



Description Two nodes are trying to communicate. A third node can be used as a relay, as the individual link distance on the path using the relay is shorter than the direct link distance between endpoints. In contrast to scenario 1, the distance between the endpoints is kept constant. Instead, the relay is placed at various positions. All positions are on the same line perpendicular to the direct link between endpoints. The links from either endpoint to the relay is always equal length, thus the three links in the network form an isosceles triangle. The distance between endpoints is equal to the $4/3$ of the average max link length, $d = 576.8\text{m}$

Metric	The connection failure probability refers to the connection between endpoints (left and rightmost nodes). The CFP is determined for various angles, θ , between the direct link between endpoints and the link between an endpoint and the relay.
---------------	--

Scenario 5



Description	A node is placed in the center of a circle. Additional nodes are placed uniformly on the circumference of the circle. The radius of the circle is the average max link length, $d = 432.6\text{m}$.
--------------------	--

Metric	Here, the interesting parameter is how many neighbors the center node has. Thus, the metrics in this scenario is both the mean number of neighbors and the standard deviation of the number of neighbors. The number of nodes in the circumference is varied in different simulations.
---------------	--

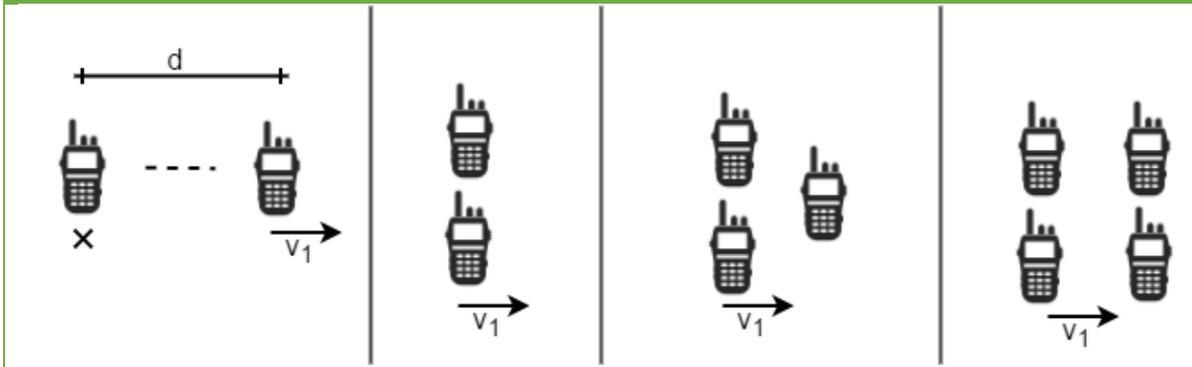
6.1.2 Mobile scenarios

The second set of scenarios include mobility and therefore look at the network conditions over time. The scenarios are described individually with a figure demonstrating the scenario, a description and an explanation of the metric used to evaluate the performance in the scenario. In the figures, movement is indicated by an arrow next to the node or cluster of nodes.

Scenario 6	
Description	Three relief workers are clustered. They split up. One stays in the original position, the others move in opposite directions, v_1 and v_2 .
Metric	As the nodes move away, at some point the connection between the left moving and the right moving nodes will break. At the point of connection failure, the distance, d , is noted. The metrics in this scenario is the mean and standard deviation of the distance to first connection failure (DTFCF).

Scenario 7	
Description	Three clustered relief workers split up. Two nodes move in opposite directions along a line with constant and equal speed. The last node moves perpendicular to the line on which the other move.
Metric	Here, the concerned connection is between the horizontally moving nodes. The distance to first connection failure is measured from the origin to either horizontally moving node, i.e. d is noted at connection failure. The calculated metrics are the mean and standard deviation of the DTFCF. The speed, $ v_3 $, of the last node is varied in different simulations. Thus, the DTFCF is specified for various speed ratios, $ v_3 / v_1 $.

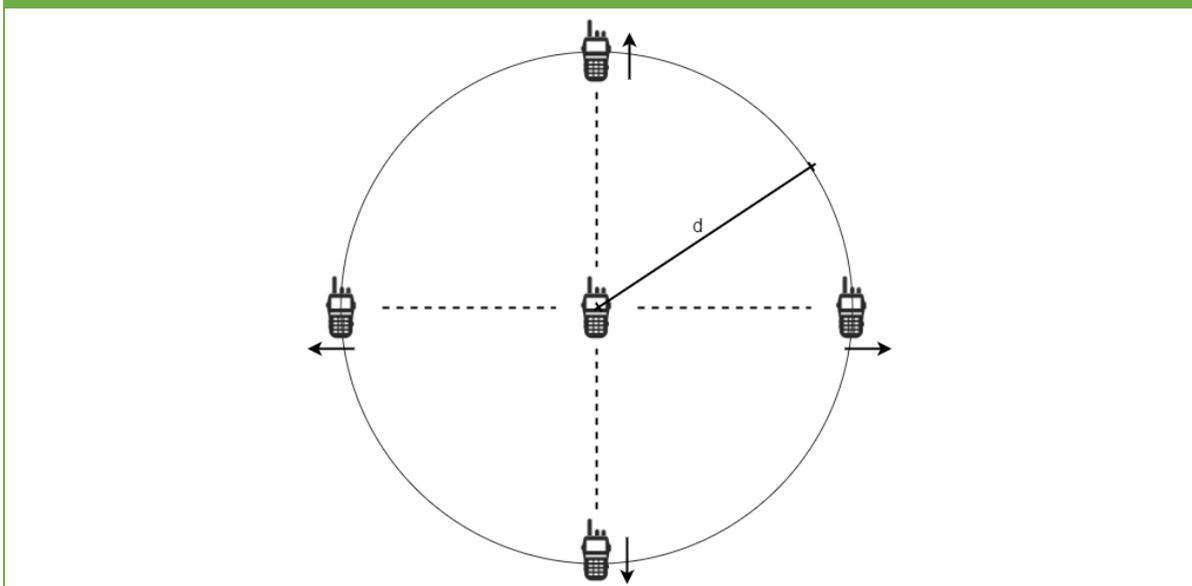
Scenario 8



Description One relief worker remains stationary, while a group moves away. The group size is varied from 1 to 4, and group constellations can be seen in the figure above. Horizontal and vertical distances between adjacent nodes is 1m.

Metric The connection failure refers to the connection between the stationary node and the group. At the point of link failure, the distance is recorded, and mean and standard deviation of the DTFCF is calculated.

Scenario 9



Description A group of nodes start at the same position. One node remains stationary, while the others move away from the origin at equal speed. The moving nodes remain on the circumference of a growing circle. Nodes remain uniformly distributed on the circumference.

Metric The distance at which the center node has lost connection to half of the nodes is recorded. Mean and standard deviation of the distances are calculated.

6.2 Results and discussion

Monte Carlo simulations of the 9 scenarios were performed and metrics extracted. For the static scenarios, 10000 realizations were performed. In mobile scenarios, 2000 realizations were performed. In the following subsections, results are presented in the form of graphs.

6.2.1 Static scenarios

Figure 17 shows the connection failure probability found by simulation of scenarios 1 where nodes are placed on a straight line and scenario 2 where clusters are placed on a straight line. In both graphs, the y-axis denotes the calculated connection failure probability and the x-axis indicates the distance between adjacent nodes used in the simulation. The unit on the x-axis is average max link length (AMLL). The red and blue line represent the results from the correlated model (corr) and the uncorrelated model (iid), respectively

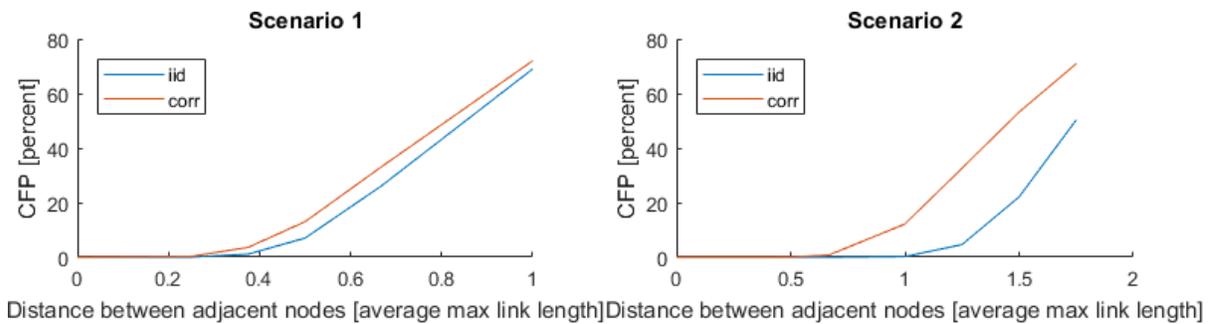


Figure 17 Connection failure probability (CFP) for scenarios 1 and 2.

In scenario 1 the connection failure probability is effectively 0 for both models when adjacent nodes are separated by a quarter average max link length (AMLL). From there, the connection failure probability worsens faster in the correlated model than in the uncorrelated model. Notably, the difference in CFP between the two models is largest around 0.5 to 0.7 average max link lengths. At these lengths, the connection failure probability remains under 50%, i.e. a connection is more likely than a connection failure. Depending on the context, a low CFP might be desired. At the 10% CFP, the difference in distance between the two models is $0.06\text{AMLL} \approx 25\text{m}$ in the simulation of the Reachi system. As the two endpoints in the simulation is separated by a relay node, the total difference in range is 50m.

The performance in scenario 2 shares the same tendency as in scenario 1. Though, the introduction of clusters in scenario 2 extends the distance at which the connection failure becomes severe. This is caused by the fact, that clustering of nodes result in multiple links between clusters. Multiple paths between the endpoint clusters decrease the probability of connection failure. The difference between the two models is more severe in scenario 2, where clusters are communicating, than scenario 1, where single nodes are communicating. The maximum CFP

difference between the models is reached at a cluster separation of 1.5 AMLL. Here, the CFP of the iid model is just above 20%, while the corr model exceeds 50%.

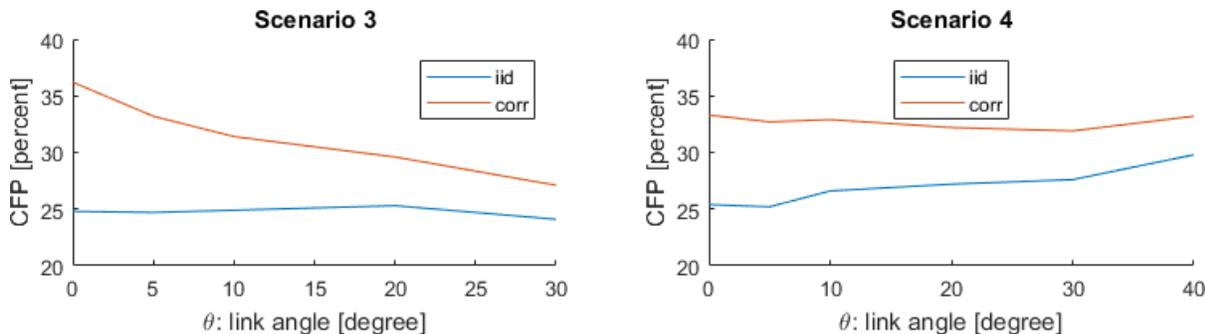


Figure 18 Connection failure probability for scenarios 3 and 4.

In scenario 3 and 4, the variable parameter is the angle, which is indicated on the x-axis of Figure 18. The y-axis is the connection failure probability. Scenario 3 is a remote node trying to communicate with two nodes (one or the other). Scenario 3 is two nodes communicating, with a relay placed perpendicular to the direct communication link. Unsurprisingly, the difference between the models decreases as the angle gets wider. In scenario 3 the uncorrelated model has constant CFP. The fact that it lies at 25% is intuitive, as the probability of a link failure is 50%, a connection failure is defined as a failure on both links, and links fail independently. The connection failure in the correlated model decrease with the growing angle. This is expected as the correlation decrease when the angle grows. The difference between the models is significant, with more than 20 percentage point difference at zero angle.

In scenario 4 the distance between the communication nodes is constant, but the length of the links going to the relay increases as the angle between the links increases. This explains the growth in connection failure probability assumed by the uncorrelated model. However, the correlated model does not follow the same trend as the uncorrelated model in this scenario. Instead, connection failure probability is lowest at angles around 30 degrees. The apparent sweet spot is noteworthy, as it is a result of the interactions of two effects:

- Link correlation cause CFP to increase at shallow angles
- Increase in link lengths cause an increase of CFP at wide angles

As such, the significance of the link angle is much lower in the correlated model, than in the uncorrelated model. Still, the results from the correlated model are consistently worse than in the uncorrelated model.

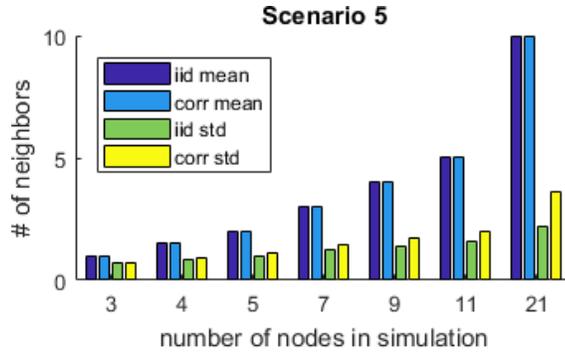


Figure 19 Statistics on the number of neighbors in simulation of scenario 5.

In scenario 5 the node connectivity of the center node is simulated. The mean number of neighbors is indicated on the y-axis of Figure 19. The x-axis indicates the number of nodes in the simulation, including the center node. The bars represent the mean and standard deviation of the correlated and uncorrelated models. The mean number of nodes is equivalent for the two models. The difference between the nodes is exposed by the standard deviation of the two models. As the number of nodes in the simulation grow, so does the standard deviation in both models. However, the correlated model experiences a faster growth than the uncorrelated model. This means, that the use of the correlated model leads to more extreme realizations, where the number of neighbors is more likely to be further from the mean.

6.2.2 Mobile scenarios

In the mobile scenarios, the metrics rely on a distance until either first connection failure or first time the number of neighbors drops to half of the initial number of neighbors. The use of a distance as opposed to time is chosen to abstract movement speed. Whether relief workers are scaling collapsed buildings or running along an open road is immaterial. Distance can be converted to time once the speed is defined. The fact that the metrics depend on the *first* connection failure or neighbor halving, means that the distances will seem comparably lower than the distances observed in the results of the static simulations.

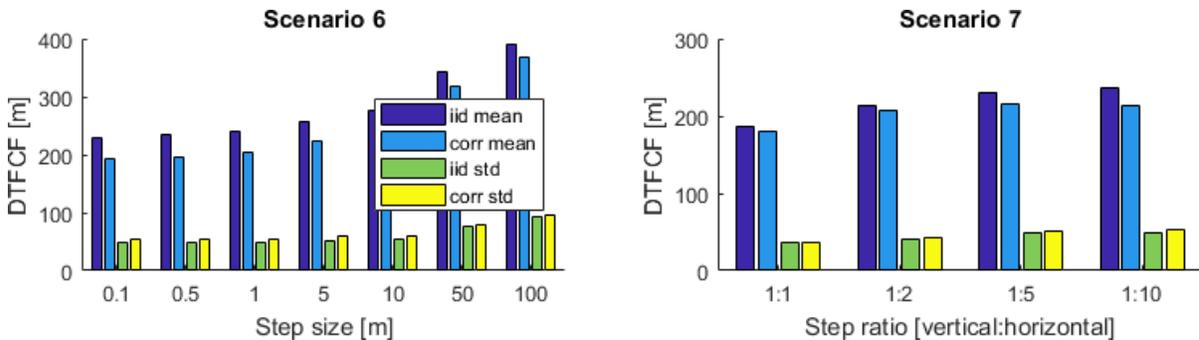


Figure 20 Distance to first connection failure (DTFCF) for scenarios 6 and 7.

In Figure 20 the distance to first connection failure is shown on the y-axis for scenario 6 and 7. The x-axis in scenario 6 is various steps sizes used in the simulation. In scenario 7, the x-axis indicates the ratio between the speed of the vertically moving nodes, and the horizontally moving node.

Interestingly, the step size varied in scenario 6 impacts the DTFCF. As the step size increase, so does the mean DTFCF. However, this can be explained by the following rationale. Consider a realization, where a link will fail at 253m. If the step size is 10m, the link failure is discovered at 260m. In contrast, using 50m step size the link failure is discovered at 300m. Thus, larger step sizes will result in over estimation of the DTFCF. In the rest of the mobile scenarios, step sizes of 1m has been utilized.

Scenario 7 is similar to scenario 4 but includes mobility. Both the communicating nodes and the relay are moving. The counteracting effects, which resulted in a near constant CFP in scenario 4, are difficult to observe in scenario 7. A close comparison of the mean DTFCF for the correlated model and uncorrelated model, reveals that the DTFCF of the correlated model plateaus – i.e. the DTFCF at different step ratios is not varying as much as in the uncorrelated model. Notably, at the 1:10 step ratio, the DTFCF is 2m shorter than at the 1:5 step ratio. The mean DTFCF of the uncorrelated model seems to grow as the step ratio decrease. This is an expected behavior, as lower step ratios will result in shorter links to the relay node. Again, the standard deviation is larger in the simulations performed with the correlated model.

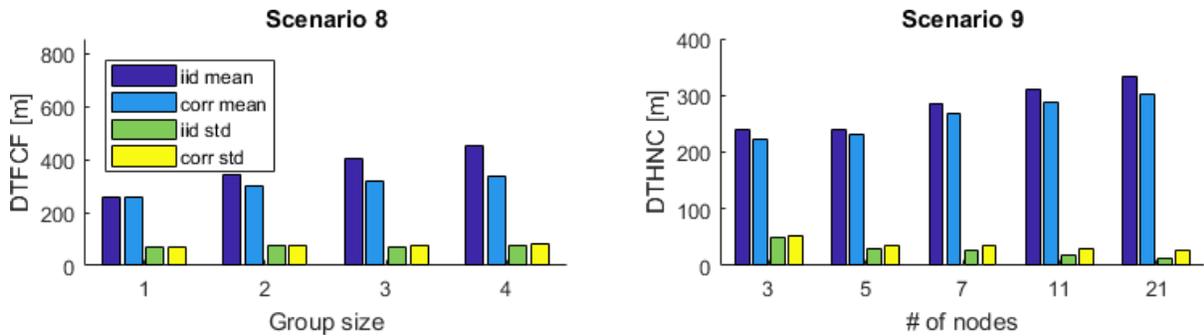


Figure 21 Distance to first connection failure (DTFCF) for scenario 8, and distance to half neighbor connectivity (DTHNC) for scenario 9.

Scenario 8 considers the simple case of one hop connections. Figure 21 shows the distance to first connection failure (y-axis) at various sizes of the group (x-axis) moving away from the stationary node. Naturally, when the group size is 1, only one link is present in the simulation and correlation has no effect. Therefore, both mean and standard deviation is equal for both models when the group size is 1. Increasing the group size immediately introduces difference in the two models. The DTFCF increase with the group size in both models, but the DTFCF in the uncorrelated model increase faster than that of the correlated model. At group size 4,

the uncorrelated model has a DTFCF, which is $113\text{m} \approx \frac{1}{4}\text{AMLL}$ longer than the correlated model. The differences in standard deviation between the models are small, though the correlated model has the highest standard deviation consistently.

Figure 21 also shows the distance to half neighbor connectivity (y-axis) of scenario 9, where one node is static in the center of a circle. The other nodes are moving away from the center, uniformly distributed on the circumference of the circle. Various simulations with different number of nodes in the simulation (x-axis) were performed. The distance to half neighbor connectivity increase with the number of nodes in the simulation. The explanation for this is twofold. Firstly, the metric DTHNC identifies the first time the neighbor connectivity drops to $\frac{1}{2}$. Because of the stochastic nature of the shadow fading, it is possible that a disconnected link can reconnect – even though the distance dependent path loss is strictly increasing with time. This leads to the second part of the explanation, which relates to this volatility of links. With few nodes in the simulation, the DTHNC is more vulnerable to disconnects, as a disconnection of a single node has relatively higher impact on the overall connectivity. Therefore, a higher number of nodes acts as a guard against early disconnects, thus increases the DTHNC.

It is clear across all simulations, that the correlated model produces worse result, whether it be higher connection failure probability, shorter mean distance to first path failure or larger standard deviation in the number of neighbors, DTFCF or distance to half neighbor connectivity.

6.3 MANET protocol design considerations

Communication in any network is managed by a set of protocols, each handling a specific task. Network protocols for MANETs should be suited to meet the needs of the network in terms of performance while operating within the constraints. The results from the previous section provides insight to the conditions, which could be expected in a scenario like disaster relief in the Philippines.

The remainder of this section will present the layered protocol stack and briefly discuss considerations, which should be made in the design of MANET network protocols for disaster relief in the Philippines.

6.3.1 The layered protocol stack

The open system interconnection (OSI) model divides communication into 7 layers. The OSI model is common knowledge among networking professionals, thus explanation of the OSI model is omitted from this report¹. The 7 layers are listed here for convenience:

¹ For an explanation of the OSI model, please refer to the model specification, [42].

1. Physical layer
2. Data link layer
3. Network layer
4. Transport layer
5. Session layer
6. Presentation layer
7. Application layer

Layers 1 thru 3 are also called the media layers, while 4 and up are called the host layers. From the perspective of the host layers, communication happens directly from source to destination (host to host). The media layers allow the abstraction of how information travels through the network. Therefore, the protocols of the media layers are responsible for dealing with the network conditions and these are determining factors of the performance of a MANET.

The physical layer deals with conversion of the digital information to and from the transmission medium. In wireless communication this is the modulation of digital information onto an electromagnetic waveform. This typically includes prevention of bit errors through forward error correction. The function and services provided by the physical layer has been abstracted throughout this report. As such it will not be subject for further discussion.

6.3.2 Data link layer protocol design considerations

The data link layer is responsible of establishing, maintaining and closing links between individual nodes. In disaster relief networks, maximizing the lifespan of the network has high priority. This requires that the power usage of network devices is minimized, because devices are battery powered. This poses a challenge for wireless networks, as both data transmission, reception and idle listening leads to significant power consumption. Most radios have a sleep mode, which is the most energy efficient state of the device. Maximizing the time spent in sleep mode, without missing transmissions from other devices, is the goal. Therefore, sleep periods should be synchronized between devices.

Establishing a link is only possible if devices are within radio range. Even if devices are within range, establishment of a link is challenged by the fact that the wireless medium is shared between devices, and therefore prone to collisions. The medium access protocol (MAC) addresses the issue of having multiple devices trying to utilize the same medium at one time. Three considerations should be made when designing a medium access protocol for the disaster relief network:

- Data transmissions are expected to be infrequent
- Synchronization requires periodic transmission

- The number of nodes within radio range differs, because of node mobility.

Synchronization messages must be sent periodically. However, transmission of synchronization messages cannot be left to a specific node, as mobility might result in some nodes getting out of radio range. Synchronization is a network wide responsibility. Additionally, mobility in the network will vary the node density, which is why static assignment of transmission schedules is not possible. Carrier sensing is a technique, where a node asserts whether other devices are utilizing the medium before transmitting. However, delays associated with e.g. signal propagation and switching between medium sensing and transmission and the possibility of multiple nodes deciding to transmit at the same time, means that carrier sensing does not completely solve the MAC issue. Carrier sense is better suited to minimize collision of aperiodic traffic. Therefore, medium access control in a MANET is better off with a MAC scheme which rely on randomization to minimize the collision probability. The dynamic of the node density also requires that the randomization behavior should adapt to the density of the network. The dynamic property of the network is demonstrated by the simulation of scenario 5, where the number of neighbors of a node was investigated. The correlated link model demonstrated that the number of neighbors in the same scenario varied with a larger standard deviation than the uncorrelated model.

Once a link has been established, the challenge becomes determining when the link is broken. Randomization of synchronization messages means that nodes might remain quiet for an unpredictable amount of time. Thus, it is unknown for other nodes if the absence of messages from a node is due to MAC, or because the node has left radio range. Transmissions from other nodes should be monitored and used to determine the likelihood of being within their radio range. Additionally, information about the location, direction and speed of other nodes and the signal strength of the link could be used in conjunction with the link model, to determine how long a link is expected to last. For example, consider scenario 8, where one node remains stationary, while another (/a group) moves away. The link should be expected to last for at least **250m** of movement, which at an average walking speed of **5km/h** translated into **3min**.

In summary:

- *Design considerations of data link layer protocols for the Reachi system, should consider the traffic patterns, power saving requirements and effects of node mobility in the network.*
- *It is recommended that the network is synchronized to maximize sleep time of the communication radios.*
- *The medium access control protocol should rely on randomization and adjust to the immediate network density.*

6.3.3 Network layer protocol design considerations

The network layer is responsible of facilitating data routes through the network. Routing protocols can adopt one of three routing schemes: proactive, reactive and hybrid. The trade off in a routing scheme lies between maintaining good routes and minimizing routing overhead. Routing overhead is every (part of a) transmission, which is generated by the routing protocol.

A proactive routing scheme will build a routing table in each node, where the next hop destination is specified for each possible endpoint in the network. The disadvantage of this approach is that building and maintaining the routing table generates network traffic periodically. Periodic overhead is a disadvantage as it may be unneeded in periods, where the network has no traffic to serve. In addition, the reactivity in terms of healing broken routes is determined by the interval between period routing traffic. The length of the interval between periodic routing traffic should be determined by the mobility of the network. High frequency network maintenance is necessary in highly mobile networks.

Reactive routing schemes build routes upon request. Routes can be built in local routing tables, or simply be appended to the transmitted data, as it is broadcast through the network. The disadvantage of reactive routing is the latency incurred by having to establish routes before data can be transmitted. The routing overhead can be large relative to the amount of data, which need to be transmitted. This is because, routing requests needs to flood the network to reach the destination. Reactive routing schemes are well suited to deal with mobility in the network, as routes are not assumed, which eliminates the possibility of using a bad route.

Hybrid routing schemes combine proactive and reactive schemes to circumvent the disadvantages of the reactive and proactive approaches. Hybrid schemes often establish routing zones. Inter-zone routing is proactively maintained, while intra-zone routing relies on reactive routing.

The appropriate routing scheme relies heavily on the mobility of a network. Protocols of either routing scheme can be configured to accommodate various degrees of mobility. However, to determine the best protocol or configuration, an evaluation of the protocols with appropriate specifications of the mobility is necessary. In the absence of a mobility model, the guideline could be the mobility in the measurements performed in the Philippines and the simulated scenarios. Other factors than mobility affect the choice of routing scheme, such as the traffic patterns and performance requirements. For the Reachi system, the network traffic is limited to simple text-like messages. The frequency of messages is presumably low – the system allows only one-way communication and messages bear content such as: “food is needed in area X”, “medical assistance is needed in area Y” or “road Z needs to be cleared of debris to allow help to pass”, which is highly dependent on position. As a change in position takes a while on foot,

traffic is assumed to be sparse. It is important that the messages reach their destination, but they should be able to tolerate delays in the order of minutes.

These observations of high mobility, low traffic amounts and high delay tolerance suggest that a reactive routing scheme could be suitable for the Reachi system. Conversely, as the Reachi network is synchronized – which rely on periodic traffic – an appropriately designed proactive routing scheme might suit the system.

In summary:

- *Network layer protocol considerations for the Reachi system should include node mobility, traffic patterns and performance requirements.*
- *Determination of the appropriate routing scheme, protocol and configuration will require further evaluation through realistic simulations with the multi-link model.*

7 | Conclusion

Coordination of disaster relief work is vital to minimize casualties. Existing means of communication are prone to destruction during natural disaster and downtime is prolonged by the fact that reestablishment requires specialized resources. Mobile ad-hoc networks can overcome many of the challenges of traditional cellular networks and is feasible for wide adoption in contrast to satellite networks. The Reachi system utilize mobile ad-hoc networking technology to provide urgent means of communication wherever needed. The mobile ad-hoc network technology requires customization to the deployment environment and scenario.

To aid the optimization of the Reachi system, signal strength measurements were performed. Thorough analysis of the measurements led to the design of a multi-link model, which describes the correlation and conditions of the wireless links comprising the Reachi mobile ad-hoc network. Validation that the model is representative of the actual link conditions in the given environment in the Philippines was ensured through comparison of the model to actual measurements performed on-site. Though the measurements had undergone secret proprietary filtering, agreement between the measurements and the model was shown to some extent, after having filtered the model-generated data through an approximation of the secret filter.

Network simulations is a key method used for design and optimization of mobile ad-hoc networks. The designed multi-link model was used in simulation of simple scenarios which might occur in a disaster relief situation. For comparison, parallel simulations were made of the designed multi-link model and an equivalent model, where the spatial correlation was omitted. Notably, the model with spatial correlation performed consistently worse across all simulation scenarios. Realistic models are of key importance for design and optimization of communication networks, as disagreement between models and reality might lead to poorly operating networks.

Design considerations of data link layer protocols for the Reachi system, should consider the traffic patterns, power saving requirements and effects of node mobility in the network. It is recommended that the network is synchronized to maximize sleep time of the communication radios. Medium access protocol should rely on randomization and adjust to the immediate network density.

Network layer protocol considerations for the Reachi system should include node mobility, traffic patterns and performance requirements. Determination of the appropriate routing scheme, protocol and configuration will require further evaluation through realistic simulations with the multi-link model.

7.1 Future work

Development of the Reachi system is ongoing. Although, prototypes are mature, further optimization and evaluation of the system is required. To this end, multiple tasks are on the horizon.

1. Specification of the node mobility is a vital aspect to allow meaningful evaluation of the system.
2. The medium access control and routing protocols implemented in the prototypes needs optimization to the specific situation in the Philippines.
3. As the Reachi system is made available outside the urban areas of the Philippines, customization for other environments is pending.

Further development in the capabilities of the Reachi system might lead to significant redesign of the communication protocols. For example, the current system enables only one-way communication. This significantly simplifies routing. Bidirectional communication will require redesign of the routing used in the prototypes. Other services, such as voice or video will increase the requirements for latency, jitter and throughput of the network.

Delay tolerance – the ability to handle extended periods of disconnection – has not been discussed in this project. Elucidation of the extent of expected disconnections, as well as methods to cope with disconnection is an obvious topic for investigation.

8 | References

- [1] D. Guha-Sapir, P. Hoyois, P. Wallemacq and R. Below, "Annual Disaster Statistical Review 2016: The Numbers and Trends," Centre for Research on the Epidemiology of Disasters (CRED), Brussels, 2016.
- [2] Disasters Emergency Committee, "Disaster lessons learned," Disasters Emergency Committee, 16 November 2013. [Online]. Available: <https://reliefweb.int/report/philippines/disaster-lessons-learned>. [Accessed 5 June 2018].
- [3] D. Stauffacher, W. Drake, P. Currión and J. Steinberger, "Information and Communication Technology for Peace: The Role of ICT in Preventing, Responding to and Recovering from Conflict," United Nations ICT Task Force, New York, 2005.
- [4] T. J. Rasmussen, "Reachi: ”Vi tror på, at Reachi kan blive forskellen mellem liv og død”, " Vækstcenter for socialøkonomiske virksomheder, 26 08 2015. [Online]. Available: <http://socialvirksomhed.dk/komigang/derfor-starter-ivaersksaettere-socialøkonomisk-virksomhed/reachi>. [Accessed 1 May 2018].
- [5] Markedsmodningsfonden, "Markedsmodningsfonden uddeler 40 mio. kr. til innovative forretningsidéer," 28 October 2016. [Online]. Available: <https://markedsmodningsfonden.dk/markedsmodningsfonden-uddeler-40-mio-kr-til-innovative-forretningsideer>. [Accessed 1 May 2018].
- [6] Philippines Red Cross, "Disaster management services - About," [Online]. Available: <https://www.redcross.org.ph/what-we-do/disaster-management-services/international-humanitarian-law-about>. [Accessed 1 March 2018].
- [7] Philippines Red Cross, "Red cross thanks early supporters of million volunteer run," 2011. [Online]. Available: <https://www.redcross.org.ph/press/news/red-cross-thanks-early-supporters-of-million-volunteer-run>. [Accessed 1 March 2018].
- [8] M. Y. S. Uddin, D. M. Nicol, T. F. Abdelzaher and R. H. Kravets, "A post-disaster mobility model for Delay Tolerant Networking," *Proceedings of the 2009 Winter Simulation Conference (WSC)*, pp. 2785-2796, 2009.
- [9] J. Ebenezer, "A mobility model for MANET in large Scale disaster scenarios," *2014 17th International Conference on Computer and Information Technology (ICCIT)*, pp. 59-64, 2014.
- [10] N. Aschenbruck, E. Gerhards-Padilla and P. Martini, "Modeling mobility in disaster area scenarios," *Performance Evaluation*, vol. 66, no. 11, p. 773–790, 2009.

- [11] S. Arimura, N. Uchida and Y. Shibata, "Self Power Supplied Micro Wireless Ballooned Network for Disaster Recovery," *2013 27th International Conference on Advanced Information Networking and Applications Workshops*, pp. 255-260, 2013.
- [12] Huawei, "Network Stability," [Online]. Available: <http://www.huawei.com/en/about-huawei/sustainability/success-stories-Abandoned%20node%20pls%20dont%20upload%20content%20to%20it/Network%20Stability>. [Accessed 5 March 2018].
- [13] Q. T. Minh, K. Nguyen, C. Borcea and S. Yamada, "On-the-fly establishment of multihop wireless access networks for disaster recovery," *IEEE Communications Magazine*, pp. 60-66, October 2014.
- [14] Z. Lu, G. Cao and T. L. Porta, "TeamPhone: Networking SmartPhones for Disaster Recovery," *IEEE Transactions on Mobile Computing*, pp. 3554-3567, 1 December 2017.
- [15] LinkAiders; Aalborg University; NeoCortec; Attensys.io, *E11589 - Reachi*, Danmarks Innovationsfond, 2017.
- [16] H. T. Friis, "A Note on a Simple Transmission Formula," *Proceedings of the IRE*, vol. 34, no. 5, pp. 254-256, 1946.
- [17] Y. Okumura, E. Ohmori, T. Kawano and K. Fukuda, "Field strength and its variability in VHF and UHF land-mobile radio service," *Review of the Electrical Communication Laboratory*, vol. 16, no. 9-10, pp. 825-873, 1968.
- [18] D. W. Matolak, Q. Zhang and Q. Wu, "Path Loss in an Urban Peer-to-Peer Channel for Six Public-Safety Frequency Bands," *IEEE Wireless Communications Letters*, vol. 2, no. 3, pp. 263-266, 2013.
- [19] J. Turkka and M. Renfors, "Path Loss Measurements for a Non-Line-of-Sight Mobile-to-Mobile Environment," *8th International Conference on ITS Telecommunications*, pp. 274-278, 2008.
- [20] J. Hejselbæk, J. Ø. Nielsen, C. Drewes, W. Fan and G. F. Pedersen, "Propagation Measurements for Device-to-Device Communication in Forest Terrain," *12th European Conference on Antennas and Propagation (EuCAP)*, 2018.
- [21] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, 2002.
- [22] TELETOPIX.ORG, "Slow Fading Margin In LTE With Example Of Standard Deviations In Slow Fading," 28 February 2013. [Online]. Available: <http://www.teletopix.org/4g-lte/slow-fading-margin-in-lte-with-example-of-standard-deviations-in-slow-fading/>. [Accessed May 2018].
- [23] J. D. Parsons, *The Mobile Radio Propagation Channel*, Chichester, West Sussex: John Wiley & Sons Ltd, 2000.

- [24] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electronics Letters*, no. 23, pp. 2145-2146, 7 November 1991.
- [25] A. Dogandzic and B. Zhang, "Dynamic power estimation and prediction in composite fading-shadowing channels," *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. ii-1013-16, 2004.
- [26] Z. Wang, E. K. Tameh and A. Nix, "Simulating Correlated Shadowing in Mobile Multihop Relay/Ad-hoc Networks," *IEEE 802.16 Broadband Wireless Access Working Group*, 30 Juni 2006.
- [27] C. Bettstetter and C. Hartmann, "Connectivity of Wireless Multihop Networks in a Shadow Fading Environment," *Wireless Networks*, pp. 571-579, 01 September 2005.
- [28] P. Agrawal and N. Patwari, "Correlated link shadow fading in multi-hop wireless networks," *IEEE Transactions on Wireless Communications*, pp. 4024-4036, August 2009.
- [29] T. Klingenbrunn and P. Mogensen, "Modelling Cross-Correlated Shadowing in Network Simulations," *Gateway to 21st Century Communications Village. VTC 1999-Fall. IEEE VTS 50th Vehicular Technology Conference*, no. 3, pp. 1407-1411, 1999.
- [30] T. B. Sørensen, "Slow fading cross-correlation against azimuth separation of base stations," *Electronics Letters*, vol. 35, no. 2, pp. 127-129, 21 January 1999.
- [31] D. J. Cichon and T. Kürner, "Chapter 4: Propagation Prediction Models," in *COST 231 final report*, European Cooperation in Science and Technology, 1996, pp. 115-208.
- [32] NIST, "Propagation Channel Models and System Performance," National Institute of Standards and Technology, 6 November 2017. [Online]. Available: <https://www.nist.gov/ctl/pscr/propagation-channel-models-and-system-performance>. [Accessed 1 May 2018].
- [33] LinkAiders, "LinkAiders - optimize disaster communication," [Online]. Available: <http://www.linkaiders.com/>. [Accessed 1 June 2018].
- [34] Neocortec, *NeoCortec-NC1000-9-LR Wireless Mesh Network Module Datasheet version 0.1*, [Unpublished, confidential].
- [35] F. van Diggelen and P. Enge, "The World's first GPS MOOC and Worldwide Laboratory using Smartphones," *Proceedings of the 28th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2015)*, pp. 361-369, September 2015.
- [36] NovAtel, "GPS Position Accuracy Measures," NovAtel, Calgary, Canada, 2003.
- [37] S. Hutsell and C. C. B. McFarland, "One-way GPS Time Transfer: 2002 Performance," *34th Annual Precise Time and Time Interval (PTTI) Meeting*, pp. 137-146, November 2000.

- [38] S. M. Kay, "14.9 Computer Simulations of Gaussian Random Vectors," in *Intuitive Probability and Random Processes using MATLAB*, University of Rhode Island, Springer, 2006, pp. 475-476.
- [39] Neocortec, "User guide," December 2015. [Online]. Available: <http://docplayer.net/53195329-Neo-cortec-user-guide.html>. [Accessed March 2018].
- [40] CNN Philippines, "More islands, more fun in Philippines," CNN Philippines, 20 February 2016. [Online]. Available: <http://cnnphilippines.com/videos/2016/02/20/More-islands-more-fun-in-PH.html>. [Accessed 1 March 2018].
- [41] R. G. L. Narayanan and O. C. Ibe, "A joint network for disaster recovery and search and rescue operations," *Computer Networks*, vol. 56, no. 14, pp. 3347-3373, 2012.
- [42] Andrea Goldsmith, *Wireless Communications*, California: Cambridge University Press, 2005.
- [43] A. S. Tanenbaum, *Computer Networks*, 5th ed., Pearson, 2010.
- [44] H. Zimmermann, "OSI Reference Model - The ISO Model of Architecture for Open Systems Interconnection," *IEEE Transactions on Communications*, vol. 28, no. 4, p. 425 - 432, 1980.
- [45] M. A. Lombardi, L. M. Nelson, A. N. Novick and V. S. Zhang, "Time and Frequency Measurements Using the Global Positioning System," *Cal. Lab. Int. J. Metrology*, pp. 26-33, July-september 2001.

Appendix

A Terminology

Networking terminology is not consistent. Thus, the same phenomena, technology or device will have different names in different literature. In this appendix, some terms and their most common aliases are presented:

Path loss is a term describing signal attenuation caused by the propagation of the signal from transmitter to receiver. The common unit used for path loss is decibel (dB). Other terms used to describe the path loss include propagation loss, path attenuation and path gain. Though, path gain is the negative path loss.

Devices used in wireless communication has various terms. Common terms include: User equipment (UE), terminal, device, mobile phone, endpoint.

Mobile ad-hoc network (MANET) is a terminology used to describe networks of wireless devices. Other terms used to describe similar networks include: wireless mesh network (WMN), peer-to-peer (P2P) network, device to device (D2D) network, multi-hop wireless network, delay tolerant network (DTN), disaster recovery network (DRN) and search and rescue network (SRN)

B Autocorrelation sequence of the first order autoregressive function

The first order autoregressive function, AR(1), is described by

$$y(t) = \phi y(t - 1) + e(t)$$

Or

$$y(t) = \sum_{k=0}^{\infty} \phi^k e(t - k)$$

Where $e(t)$ is a zero mean iid Gaussian process and $\phi \leq 1$ is the correlation coefficient. The autocorrelation function of the AR(1) function is

$$\begin{aligned} r_y(k) &= E[y(t)y(t + k)] \\ &= E[y(t)(\phi y(t - 1 + k) + e(t + k))] \\ &= E[y(t)\phi y(t - 1 + k) + y(t)e(t + k)] \\ &= \phi E[y(t)y(t - 1 + k)] + E[y(t)e(t + k)] \end{aligned}$$

$E[y(t)e(t + k)]$ is zero, because the values of $y(t)$ is only dependent on past values of $e(t)$, i.e. the for $e(t + k)$, $k \geq 1$ we have $E[y(t)e(t + k)] = E[y(t)]E[e(t + k)] = 0$

$$\begin{aligned} r_y(k) &= \phi E[y(t)y(t - 1 + k)] \\ r_y(k) &= \phi r_y(k - 1) \end{aligned}$$

From the sum-version of the AR(1) definition, we can express

$$r_y(0) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \phi^k \phi^l E[e(t - k)e(t - l)]$$

Because of independence, when $k \neq l$, the expected value is 0, leaving only $k = l$

$$\begin{aligned} r_y(0) &= \sum_{k=0}^{\infty} \phi^k \phi^k E[e(t - k)e(t - k)] \\ r_y(0) &= \sum_{k=0}^{\infty} \phi^{2k} E[e(t - k)e(t - k)] \\ r_y(0) &= \sum_{k=0}^{\infty} \phi^{2k} E[e(t)^2] = \sum_{k=0}^{\infty} (\sigma_e^2 + \mu_e^2) \phi^{2k} \end{aligned}$$

$$r_y(0) = \sigma_e^2 \sum_{k=0}^{\infty} \phi^{2k} = \frac{\sigma_e^2}{1 - \phi^2}$$

Because the infinite sum is the geometric series $\sum_{k=0}^{\infty} cr^k = \frac{c}{1-r}$, for $|r| < 1$ where $c = \sigma_e^2$ and $r = \phi^2$. Inserting into the recursive expression for the ACS, for all $k \geq 0$,

$$r_y(k) = \phi^k r_y(0) = \phi^k \frac{\sigma_e^2}{1 - \phi^2}$$

Which can be extended to include negative values of k , because a property of the autocorrelation sequence is that $r_e(k) = r_e(-k)$, yielding

$$r_y(k) = \frac{\sigma_e^2}{1 - \phi^2} \phi^{|k|}$$

C Data format

The format of the fields in the data file supplied by the measurements taken in Philippines are elaborated for future reference.

“Raw” data was provided in 3 separate text files per device. The files from one device were kept in a folder named by the UID of the device. The text files contain only hex values, after a certain log format. The collection of device folders is placed in a folder and opened through the *LogView*-tool version 1.40. The LogView tool can load the datafiles from all devices.

From the LogView-tool a comma separated values (csv) file can be exported. The format of the csv file is given in Table 4 and replicated in Table 5 for convenience.

Field 1	Field 2	Field 3	Field 4	Field 5		
Timestamp	ID	UID	Latitude	Longitude		
Field 6	Field 7	Field 8	...	Field 39	Field 40	Field 41
Neighbor 1 ID	Neighbor 1 RSSI	Neighbor 1 distance	...	Neighbor 12 ID	Neighbor 12 RSSI	Neighbor 12 distance

Table 5 Format of the comma separated data file after pre-processing by the LogView-tool.

The format of each individual field is explained below.

- **Timestamp:** The timestamp is an integer number representing the number of seconds since the Unix epoch. Unix epoch is 00:00.00 01/01/1970 UTC
- **ID:** The (neo) ID is a unique positive base 10 integer in the interval [1,65563]. [39]
- **UID:** The UID is a unique hexadecimal number. The number can take 16^{24} distinct values.
- **Latitude:** The latitude is given as a floating-point number indicating the latitude in degrees with respect to equator. Positive latitude is north of equator, while negative latitude is south of equator. The latitude is defined in the interval $[-90,90]$.
- **Longitude:** The longitude is a floating-point number denoting the longitude in degrees west (negative values) and east (positive values) of the Prime Meridian running through the Royal Observatory, Greenwich in England.
- **Neighbor X ID:** The Neighbor ID is equivalent to the ID of another node. However, the Neighbor X ID is given in hexadecimal rather than base 10.
- **Neighbor X RSSI:** The RSSI is the received signal strength measured on a transmission from neighbor X. The measurement is an average signal strength of the transmitted frame. It is an integer number within the interval $[-1, -128]$.

- Neighbor X distance: The distance is a floating-point number, representing the distance from the current node to node X in kilometers. The distance is equal to or greater than zero.

8.1.2 Peculiarities

The exact method of measuring received signal strength is kept confidential by the producers of the Neocortec radio module, NeoCortec-NC1000-9-LR.

How a node is determined as a neighbor is unknown. One conjecture is that a beacon from a node needs to be heard 3 times in an interval for it to appear on the list of neighbors. Certainly, no more than 12 neighbors are registered at any given time.